

Dal documento all'informazione

a cura di Roberto Guarasci

INTRODUZIONE

Forse sarebbe opportuno iniziare la presentazione del volume con la frase: “all’inizio volevo scrivere un manuale”, o, meglio, “era da un po’ di tempo che volevo scrivere un manuale di Documentazione”. Poi, prima il ricordo delle parole del mio maestro, Salvatore Carbone, che ammoniva a ponderare bene una decisione di tal fatta e, poi, l’evoluzione – tutto sommato anomala e disallineata - della disciplina in Italia mi hanno indotto a temporeggiare. All’indecisione non era inoltre estraneo il concepire i manuali – pur se utili – necessariamente riduttivi dello specifico ambito disciplinare e poco adatti ad una disciplina nella quale la trasversalità delle conoscenze e degli apporti disciplinari diventava sempre più il tratto caratterizzante.

I due testi di Paolo Bisogno¹ e il sintetico lavoro di Ferruccio Diozzi² poco si prestavano alle finalità che avevo in mente e che identificavano un target di utenti già in possesso delle cognizioni di base e bisognevoli di specifiche conoscenze o aggiornamento specialistico. Il modello italiano più vicino era la raccolta, ormai datata, *La Documentazione in Italia*³ pur se la struttura del volume era quella di una miscellanea specialistica – per esplicita affermazione del curatore – e non aveva alcuna finalità o strutturazione che lo rendesse specificatamente utilizzabile nei contesti citati.

¹ *Documentazione e il Futuro della Memoria.*

² FERRUCCIO DIOZZI, *Documentazione*, AIB, Roma, 1998.

³ AUGUSTA MARIA PACI (a cura di), *La Documentazione in Italia*, Franco Angeli, Milano, 1996.

Al di là dei confini nazionali il testo di Claire Guinchat e Michel Menou⁴ edito in più lingue dall'Unesco, datato pur se aggiornato nel 1990 da Marie France Blanquet, restava uno dei pochi organici punti di riferimento se non altro per la completezza della strutturazione e dell'impalcatura concettuale. Sul versante non europeo, seppure con una diversa strutturazione, un momento importante nella formalizzazione e nel posizionamento didattico della disciplina era, invece, rappresentato dal volume di Irene S. Farkas-Conn, *From Documentation to Information Science*⁵, edito negli stessi anni, anche se l'approccio alla disciplina era filtrato attraverso l'evoluzione e la storia dell'American Documentation Institute poi diventato American Society for Information Science.

La difficoltà di fondo era cercare di scrivere un testo che spiegasse ad utilizzatori italiani le ragioni dell'esistenza di una disciplina e dei suoi fondamenti epistemologici in un contesto limitatamente umanistico quando il panorama europeo postulava quell'approccio solo come uno dei possibili e non certo come l'esclusivo.

Le riflessioni di Hubert Fondin sulle relazioni tra “les techniques documentaires” e la “science de l'information” che concludono «Il n'y a donc pas une technique documentaire et une science associée, mais des techniques documentaires utilisées pour retrouver les documents utile set des sciences, soit explicatives des phénomènes physique, soit compréhensives des phénomènes humains et sociaux, parmi lesquelles la science de l'information, convoquées autour de celle-ci»⁶ possono forse

⁴ CLAIRE GUINCHAT-MICHAEL MENOU, *Science set Techniques de l'Information et de la Documentation*, Unesco, Parigi, 1990.

⁵ IRENE S. FARKAS-CONN, *From Documentation to Information Science*, Greenwood Publishing group, Westport, 1990.

⁶ HUBERT FONDIN, *La Science de l'Information et la documentation, ou*

rappresentare il canovaccio intorno al quale ho cercato di costruire i contributi a questo volume che cerca di organizzare in forma di manuale specialistico dei contributi di singoli colleghi – ai quali va il mio ringraziamento – sia sulle varie tecniche documentali che sui fondamenti epistemologici delle scienze correlate nel tentativo di presentare non una visione deterministica della disciplina, poco adatta quando non si studiano esclusivamente fenomeni fisici, ma un approccio sistemico basato sull'interrelazione cognitiva che oppone al tradizionale paradigma realtà-teoria-metodo una nuova formulazione epistemologia-teoria concetto – metodo – oggetto⁷.

les relations entre science set technique, in *Documentaliste – Sciences de l'Information*, 2002, vol. 39, n. 3, p. 129.

⁷ Cfr. ALEX MUCCHIELLI, *La nouvelle communication: épistémologie des sciences de l'information-communication*, Paris, Armand Colin, 2000.

Che cos'è la documentazione

ROBERTO GUARASCI

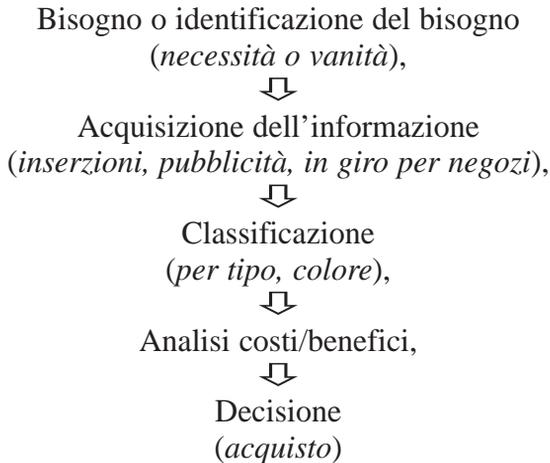
Il punto primo di ogni disciplina, specie se applicativa, dovrebbe essere quello di presentare ai possibili discenti le motivazioni dell'utilità della stessa che poi è un tutt'uno con la sua stessa ragion d'essere. Tanto più queste ragioni sono vicine alla esigenze ed ai bisogni della vita reale tanto più la disciplina sarà diffusa e permeerà i gangli di una società civile¹.

Una delle ragioni dello sviluppo esponenziale delle tecnologie dell'informazione e della comunicazione – ad esempio – è proprio quella di essere pervasive della vita quotidiana; questo genera mercato e la forza di questo, lungi dal banalizzarle – come temono alcuni – attrae investimenti e muove la ricerca. Non sempre le scienze documentarie hanno posto attenzione a questi precetti anzi, più spesso, hanno tratto godimento dall'essere appannaggio di un ristretto pubblico di specialisti.

Se – ad esempio – banalizzando, al posto dell'enunciazione del concetto astratto dell'informazione documentale come sup-

¹ «On peut s'interroger sérieusement à savoir si certains spécialistes de la gestion de l'information administrative n'ont pas perdue de vue, ou n'ont peut être jamais eu en tête les besoins réels de leurs clients et se sont, avec le temps, cantonnés dans une approche nombriliste qui les servait avant tout comme professionnels d'un domaine mal circonscrit». MICHEL ROBERGE, *La Gestion de l'Information Administrative*, Documentor, Québec, 1992, p. 18.

porto decisionale², certo non immediatamente comprensibile ai più, utilizzassimo la classica situazione dello shopping per i negozi e, segmentando l'azione:



estrapolassimo, conseguentemente, dalla realtà esperenziale la valenza teorica, faremmo diventare immediatamente evidente il concetto e, a parte le obiezioni di qualche purista, avremo sicuramente reso l'idea della "fisicità" della disciplina e della sua pratica utilità in una pluralità di contesti non necessariamente specialistici³.

² Ai fini dell'analisi dei processi decisionali e della problematica di un parziale utilizzo delle informazioni da parte dei decisori nelle scelte si veda: ROBERT AXELROD, *The Structure of decision: cognitive maps of political elites*, Princeton University Press, Princeton, 1976.

³ Una sperimentazione didattica in tal senso è stata condotta nell'a.a. 2000/2001 nel corso di Documentazione dell'Università della Calabria facendo registrare significativi miglioramenti in termini di apprendimento. Cfr. ROBERTO GUARASCI *et al.*, *From EDoc to Pekita: The evolution of a platform for e-learning and profiling technology*, in «Journal of e-learning and Knowledge Society», n. 2 (2005), pp. 233-243.

Ogni giorno noi attuiamo numerose sequenze di Bisogno/Informazione/Decisione, indipendentemente dallo specifico settore di attività nel quale lavoriamo. Conoscere le metodologie e le tecniche per far sì che questa azione involontaria diventi un “processo di qualità” è estremamente utile a tutti coloro i quali svolgono una qualche forma di attività relazionale e/o decisionale. Se riusciamo a definire e percepire con chiarezza questa concettualità, la Documentazione assume una sua completa caratterizzazione come base e fondamento teorico delle scienze dell’informazione.

Dovendo, comunque, se non altro per necessità didattiche, formulare una definizione della disciplina, potremmo affermare che «la documentazione si pone come obiettivo il far conoscere ciò che è stato fatto per poter fare e, pertanto, le sue differenziate attività si presentano al mondo esterno come comunicazione di conoscenze offerte sotto forma di dati elaborati con rigorosi criteri e omogenee procedure»⁴. In un’ottica più corretta sarebbe forse il caso di dire che l’obiettivo è quello di far conoscere ciò che è e non solo ciò che è stato fatto, giacché in una società in frenetico divenire, nella quale l’obsolescenza delle informazioni è estremamente rapida è, molto spesso, il presente e non il passato a rappresentare la base per l’astrazione informativa necessaria al supporto decisionale. «L’origine etimologica – afferma il Samain – sottolinea il grande compito della documentazione che è quello di insegnare, informare, trasmettere ad altri le conoscenze, i fatti, le idee, tramite i mezzi più adeguati»⁵. Quasi tut-

⁴ PAOLO BISOGNO, Presentazione a: MARIA PIA CAROSELLA-MARA VALENTI (a cura di), *Documentazione e Biblioteconomia*, Milano, Franco Angeli 1987, p. 16.

⁵ JACQUES SAMAIN, *Qu’est-ce que la Documentation?*, Association Internationale des Documentalistes et techniciens de l’information, Testo dattiloscritto, p. 1, in GIGLIOLA NEGRINI, *Premessa* a TAMARA FARNESI-ANGELO SENSINI (a cura di), *Manuali di Scienza della Documentazione*, CNR, Roma, 1979, p. 7.

te le definizioni correnti sono – in qualche modo – esemplate o ispirate alle affermazioni di Otlet: «Les buts de la documentation organisèe consistent à pouvoir offrir sur tout ordre de fait et de connaissance des informations documentées»⁶. La generica affermazione di una disciplina che ha come scopo e finalità quella di fornire informazioni sullo scibile umano era, sicuramente, già ambiziosa per gli anni in cui venne formulata ma, se non la si colloca nella corretta prospettiva storica, rischia – tanto più – di suscitare il sorriso oggi, in un'epoca nella quale la comunicazione e l'informazione sono la principale merce di scambio e la produzione di oggetti informativi cresce ad un ritmo non immaginabile nel secolo scorso⁷. Sicuramente i concetti di una conoscenza globale – un cervello del mondo – liberamente accessibile e personalizzabile sono strettamente connessi al generale clima di internazionalizzazione e di transnazionalità del primo dopoguerra ed ai movimenti culturali del taylorismo e del positivismo, oltre ad essere frutto delle elaborazioni personali di un ristretto gruppo di intellettuali europei che tentarono un approccio globale alla conoscenza concepita come base ineliminabile per garantire la pace mondiale. La *one best way* di Frederick Winslow Taylor e l'utopia classificatrice di Otlet concordano nell'idea dell'analisi e dell'ottimizzazione dell'agire umano divergendo, ovviamente, sulle finalità d'uso. Anche la collaborazione di Otlet

⁶ PAUL OTLET, *Traité de Documentation, le Livre sur le livre*, Bruxelles, Mundaneum, 1934, p. 6. O anche «Par Documentation on entendra à la fois l'ensemble des documents et la fonction de documenter, c'est-à-dire d'informer à l'aide de documentation». Ivi, p. 373.

⁷ Il 15 settembre 1997 Larry Page e Sergey Brin registrano in California il nome di una nuova azienda: Google.com. La mission strategica dichiarata è "Organizzare la conoscenza del mondo intero e renderla accessibile a tutti". Slogan dell'azienda è: È possibile guadagnare senza far del male a nessuno.

con Le Corbusier⁸ per la redazione di un progetto di Città Mondiale – capitale universale della scienza e del sapere – da realizzarsi a Ginevra prima e ad Anversa poi mirava a generare – mediante la conoscenza e la collaborazione – un clima di fraternità tra i popoli basato sull'interscambio della cultura e del sapere. Presupposto di base per ogni attività cognitiva erano l'ordine e, dunque, la classificazione. Un ordine che – da buon positivista – Otlet immaginava già insito nelle cose. Al saggio competeva solo di svelarlo e renderlo evidente. Anche nell'impianto urbanistico della città ideale di Le Corbusier tutto è suddiviso e organizzato in settori, come nei repertori di Otlet. «Otlet, l'homme qui voulat classer le monde, croyait profondément que le progrès nécessitait l'ordre. Chaque information sur la bonne fiche, chaque être humain à la bonne place»⁹. È il medesimo concetto – con accentuazioni diverse – che compare anche nell'opera di Wilhelm Ostwald¹⁰ che, oltre alla standardizzazione dei formati delle pubblicazioni, ipotizzava la costruzione di una lingua artificiale ed “ausiliaria” come mezzo di comunicazione globale¹¹.

«Sur le plan scientifique et philosophique, l'oeuvre d'Otlet porte la trace d'un ensemble de conception et de caractéristiques

⁸ Le Corbusier – pseudonimo di Charles-Edouard Jeanneret-Gris – (La Chaux-de-Fonds, 6 ottobre 1887-Roquebrune-Cap-Martin, 27 agosto 1965) architetto e urbanista svizzero, pioniere nell'uso del cemento armato per l'architettura e uno dei padri dell'urbanistica.

⁹ JEAN FRANCOIS FUEG, *Ordo ab Chaos, Classer est la plus haute opération de l'esprit*, in «Associations Transnationales», 1.2 (2003), p. 35.

¹⁰ Friedrich Wilhelm Ostwald, comunemente noto come Wilhelm Ostwald (Riga, 2 settembre 1853 - Leipzig, 4 aprile 1932), padre fondatore della chimica fisica, vinse il premio Nobel per la chimica nel 1909.

¹¹ L'idea di una lingua ausiliaria di comunicazione capace di affiancare le lingue naturali è, all'epoca, diffusa. Già nel 1887 Ludovico Lazzaro Zamenhof (1859-1917) aveva pubblicato la versione iniziale dell'Esperanto, lingua ausiliaria di comunicazione.

qui la rattachent à la fois au rationalisme kantien, à la recherche par Leibniz d'un langage universel – plus qu'à celle de l'encyclopédie de Diderot – et au projet romantique de Novalis de l'intégration du savoir humain»¹².

Per Otlet l'organizzazione delle informazioni e l'accesso alla conoscenza non potevano prescindere dalla classificazione¹⁰ – la più dettagliata possibile – e dalla standardizzazione dei formati e dei sistemi di classificazione¹³. Alla fine il processo iterativo di standardizzazione e “meccanizzazione” costruisce una nuova tipologia di sapere. La formalizzazione fisica dell'ordine, derivante da una corretta classificazione¹⁴, è data dal Repertorio documentale e dalla formulazione – innovativa – del concetto di Dossier. Nella definizione di quest'ultimo concetto, avviata nella relazione al Marchese Cappelli¹⁵ e compiutamente definita nel *Traité*, egli sancisce l'autonomia della disciplina dalle altre scienze del libro e del documento e, nello stesso tempo, postulando un'unità documentale dinamica e fruibile liberamente anche da una pluralità di accessi può, a ben ragione, essere indicato come uno dei precursori delle moderne scienze dell'informazione.

«Le Répertoire de documentation est destiné à embrasser dans

¹² PAUL GHILS, *Fonder le monde, fonder le savoir du monde, ou la double utopie de Paul Otlet*, in «Associations Transnationales», 1.2 (2003), p. 39.

¹³ «Classer est la plus haute opération de l'esprit, celle qui implique toutes les autres. L'esprit s'élève à mesure qu'il est susceptible d'abstraction, de systématisation et de synthèse». P. OTLET, *Traité de Documentation*, cit., p. 379.

¹⁴ In tal senso la sua modifica alla classificazione Dewey e la costruzione della Decimale Universale. «Uno speciale equipaggiamento di segni ha la funzione di rendere più precise e articolate le collocazioni dei soggetti: dalle suddivisioni delle lingue alle relazioni generali tra due soggetti, agli aspetti temporali e spaziali, ecc.». ALFREDO SERRAI, *Le Classificazioni*, Olshki, Firenze, rist. 2002, p. 285.

¹⁵ P. OTLET, *Rapport Présenté à monsieur le marquis Cappelli, Président de l'Institut International d'Agriculture*, Annexe A à *Rapports Sur La Bibliographie*, Rome, Institut International d'Agriculture, 1911.

l'unité de ses cadres tous les documents d'études qui n'étant pas des livres, des périodiques ou des brochures, ne trouvent pas leur place naturelle dans la bibliothèque proprement dite [...]

Dans les dossiers du répertoire de documentation sont rassemblés notamment: les notices bibliographique avec notes d'analyse utiles aux services, les découpures de publications, principalement celles des Bulletins de reinsegnements publiés par l'Institut, les copies des lettres recues et expédiées, les feuilles volantes, tableaux statistiques, menus imprimés et en général toutes les notes et notices des rapports établis dans les services intérieurs ou communiqués par le dehors»¹⁶.

Il Dossier è l'unità informativa primaria, frutto della collazione di tutte le fonti disponibili, che non perdono la loro specifica caratterizzazione né vengono snaturate rispetto agli autonomi criteri di ordinamento e gestione. Esse duplicano, in questo contenitore in perenne divenire, la parte di patrimonio informativo necessario al raggiungimento delle finalità per le quali viene creato. In questo senso il dossier non vive autonomamente – come il fascicolo archivistico – ma solo in rapporto a scopi specifici e a determinati domini di conoscenza¹⁷.

¹⁶ Ivi, pp. 9-11.

¹⁷ L'Archivio di Documentazione Politica della Democrazia Cristiana era – ad esempio – ispirato a questa filosofia. Carlo Danè – che ne fu responsabile – ci da un'idea della metodologia di costruzione delle schede: «Un'altra difficoltà da noi incontrata è consistita nella schedatura dei documenti, che per anni abbiamo realizzato attraverso l'elaborazione dei cosiddetti abstract, più o meno ampi. Effettivamente la scheda, se ben fatta molte volte fornisce da sola dati sufficienti da permettere di utilizzare l'informazione anche senza ricorrere al documento originale. Questo metodo però richiedeva l'impiego di diverse persone sia nella fase preparatoria delle schede sia in quella della loro quotidiana sistemazione manuale nello schedario generale, cosicché la mancanza di un numero adeguato di addetti a questi compiti [...] ci costrinse anche qui a rinunciare [...] Si passò allora – dove possibile, naturalmente, distruggendo, almeno parzialmente le fonti, all'o-

È fin troppo semplice e ovvio dire che la definizione della disciplina rimanda immediatamente alla valenza del termine documento che ha conosciuto e conosce una pluralità di accezioni e di significati che stanno, in gran parte, alla base delle diversificazioni esistenti. Anche all'interno delle scienze del libro e del documento non vi è una accettazione univoca della concettualità del termine che oscilla dall'accentuazione giuridico formale di parte di un tutto avente vita propria solo in quanto componente il fascicolo – unità primaria – a quella più estesa e più generale – mutuata da Otlet¹⁸ – di oggetto informativo completamente indipendente dal supporto scrittorio e quindi insieme degli elementi descrittivi di una struttura testuale.

Sulle diverse concettualizzazioni del termine si giocano anche i rapporti – non sempre lineari – tra le varie discipline, rapporti spesso complicati dalla non condivisione di uno stesso vocabolario.

La definizione del documento come mero oggetto informativo permette a Paul Otlet, rispetto agli autori precedenti, come ad esempio, Peignot¹⁹, che concepiva la “bibliologie” come strettamente legata alle scienze librerie, di cambiare la prospettiva, operando una singolare commistione di esperienze e prospettive

perazione del ritaglio. Anche questo sistema fu seguito per diversi anni, abbinato spesso alla fotocopiatura, così da portare alla formazione di voluminosi dossier che raccoglievano, ordinati cronologicamente, sia i ritagli che le indicazioni per il reperimento di altri documenti che non si potevano né ritagliare né fotocopiare». C. DANÈ, *Gli Archivi della Dc*, in *Gli Archivi dei Partiti Politici*, Roma, Ministero per i beni culturali, 1996, p. 120.

¹⁸ «Depuis 1903 en effet, Otlet s'est intéressé au concept de document, affirmant qu'il est erroné de s'en tenir aux seuls imprimés, ce qui doit primer n'est plus le texte lui-même mais bien l'information quel que soit son support. Tout ce qui est porteur d'information devient document». J.F. FUEG, *Ordo ab Chaos*, cit., p. 30.

¹⁹ GABRIEL PEIGNOT, *Dictionnaire raisonné de Bibliologie*, Renouard, Paris, 1802.

diverse, facendo della documentazione l'antesignana delle scienze della comunicazione e dell'informazione documentale²⁰. Egli fonda questi suoi assunti proprio su una concettualità dinamica del termine documento inteso come prisma dalle mille sfaccettature e dai mille punti di accesso e di alimentazione che, oltre ad essere frutto di una condivisione strutturale, utilizza tecnologie e metodi "moderni" per aumentare le sue possibilità di fruizione modellandosi – in qualche modo – sui bisogni dell'utente e intuendo le potenzialità delle nuove scoperte scientifiche.

«En 1937 dans son rapport annuel de l'IID, il cite dans le progrès futurs de la documentation «Les machines intellectuelles et en particulier le cinéma sonore et colorié, le disque et la télévision»²¹.

Nello stesso periodo Otlet commissionerà a Robert Goldschmidt la progettazione di una macchina capace di fotografare i documenti e di renderli – successivamente – fruibili al maggior numero di persone possibili dando così vita reale e teorizzazione concettuale all'uso del microfilm.

«Coup sur coup des inventions merveilleuses – dirà poi nel *Traité* – sont venues étendre immensément les possibilités de la documentation»²².

²⁰ «[...] il rappelle les recherches du suisse Rothlisberger qui le premier élabore la statistique internationale des imprimés dans le droit d'auteur [...] Il rappelle les travaux théorique qu'il a réalisés avec son disciple Iwinski [...] il rappelle les travaux du célèbre chercheur russe Roubakine sur la psychologie bibliologique, sur les rapports mentaux entre la pensée des auteurs et celle des lecteurs qui devaient aboutir à a une théorie fonctionnelle de la création littéraire et aux méthodes de propagande et de publicité [...]»: ROBERT ESTIVALS, *Preface* a P. OTLET, *Traité de Documentation*, cit., rist. 1988, p. IV.

²¹ P. OTLET, *Rapport Annuel Général, IID Communicationes*, Octobres 1938, p. 100. In FRANCOIS LEVIE, *Paul Otlet et le multimedia*, in «Associations Transnationales», 1.2 (2003), p. 53.

²² P. OTLET, *Traité*, cit., p. 216.

Pur se lo strumento di classificazione utilizzato apparenta la nascente disciplina alla biblioteconomia «introducendo il termine documentazione per indicare un nuovo campo di studi e di ricerca, scientificamente fondato ed autonomo rispetto alle discipline bibliografiche e biblioteconomiche tradizionali, Otlet dimostra di concepire la nuova disciplina in termini di progettualità futura che coinvolge globalmente il disegno di un sistema di produzione intellettuale mondiale [...]»²³.

Egli, dunque, concepiva la documentazione come una disciplina autonoma e non come una semplice filiazione delle scienze del libro in quanto la sua concezione del documento non era meramente conservativa ma bensì operativa. Esso è “oggetto”, “magma”, da ridurre a unità informative semplici da integrare per costruire un nuovo sapere basato sulla interrelazione cognitiva.

Questa autonomia della documentazione non sarà però, né pacifica nè unanimemente accettata tanto che la sua concettualità oscillerà tra la definizione di aspetto specialistico della biblioteconomia e l'inserimento nel più ampio ma indistinto panorama delle scienze dell'informazione, aprendo la strada all'utilizzazione massiccia dell' Information Technology degli anni Ottanta del XX secolo²⁴.

Queste sue teorizzazioni e la consapevolezza di aver fatto nascere una nuova scienza Otlet le dettaglia e le esplicita compiute

²³ ANNA BALDAZZI, *Le radici storiche della documentazione in Europa*, in: AUGUSTA MARIA PACI (a cura di), *La Documentazione in Italia*, Milano, Franco Angeli, 1996, p. 44.

²⁴ Emblematica è l'evoluzione della denominazione della Federazione Internazionale di Documentazione. Nata nel 1931 come Istituto Internazionale di Documentazione, su impulso di Otlet e La Fontaine, diventerà, nel 1937, Federazione Internazionale di Documentazione per poi chiamarsi, nel 1988, Federazione Internazionale di Informazione e Documentazione. Cesserà le sue attività nel 2002.

tamente nello studio di fattibilità per la gestione della documentazione dell'Istituto Internazionale di Agricoltura²⁵. La sequenza delle fasi operative – all'epoca delineate - è di estrema modernità:

- definizione del campo di interesse dell'istituzione
- ricognizione di enti affini come potenziali collaboratori e utenti
- costruzione di una biblioteca centralizzata
- costruzione di un repertorio bibliografico
- costruzione di un repertorio della documentazione
- definizione delle regole diversificate per funzioni di accesso/consultazione
- organizzazione degli archivi in sinergia con il repertorio della documentazione
- adozione di un sistema univoco di classificazione e definizione di un manuale dei metodi documentali.

L'intento di Otlet è quello di una «rationalisation du livre et du document», la costituzione di una unità documentaria/cognitiva primaria formata da libri e documenti, ognuno con la sua valenza e la sua carica di informazioni e di specificità. Non a caso nei “Fondamenta” del suo *Traité*, la Documentazione non è solo una evoluzione della Biblioteconomia ma una vera e propria nuova creazione sinergica composta da sette parti principali: *Les documents particuliers*, *La Bibliothèque*, *la Bibliographie*, *Les Archives documentaires*, *Les Archives Administratives*, *Les Archives anciennes*, *Les documents autres que bibliographique et graphique*, *Les collection Muséographiques*, e *l'Encyclopédie*. Di fronte ad una frattura concettuale tra il mondo del libro e quello del documento egli riafferma una concezione unitaria di ar-

²⁵ P. OTLET, *Rapport Présenté à monsieur le marquis Cappelli*, cit., pp. 11-19.

chivio e biblioteca come *Knowledge basket*, sostanzialmente assimilabili, e uniti dalla comune finalità di generatori di conoscenza formalizzata. La diversificazione netta era, comunque, ragionevolmente recente tanto che ancora nel *De Archivis* di Baldassarre Bonifacio si avanza la possibilità che le biblioteche possano essere definite *archivia quaedam librorum* e, nell'introduzione allo stesso testo, curata da L. Sandri²⁶, in più di una occasione si ribadisce non solo la complementarietà ma la loro stessa originaria unicità: «oggi sono due gli istituti che conservano le antiche documentazioni: archivi e biblioteche ma mentre i primi asservantur acta publica le seconde invece doctorum viro- rum lucubrationes; tuttavia una simile distinzione è relativamente recente, in antico alle origini, i due istituti erano una cosa sola»²⁷. E non si poneva certamente l'accento solo sulla fisica coabitazione dei materiali. Tre secoli dopo, però, la posizione del Casanova, pur se non di totale ostilità verso i media diversi dal "documento", inteso nella sua più stretta accezione giuridico formale è, invece, di sostanziale chiusura: «Non neghiamo neppure che diano e daranno sempre maggiore risalto alla fredda dicitura o narrazione dei documenti che gioveranno immensamente alla futura storiografia. Ma, in verità, preferiamo il sistema italiano di farne cosa separata dagli archivi: poiché non hanno alcuno dei requisiti di questi istituti, non ne hanno la multipla finalità, e s'avvicinano assai meglio alle biblioteche e ai musei che non agli

²⁶ LEOPOLDO SANDRI, *Il De Archivis di Baldassarre Bonifacio*, in «Notizie degli Archivi di Stato», X (1950), pp. 95-111. «tuttavia queste notizie frammentarie erano però state sufficienti a mettere in luce come ci siano state epoche nelle quali il vocabolo biblioteca stesse in realtà ad indicare un archivum, in quanto si è cominciato col conservare i documenti fondamentali alla vita dello Stato e come sia avvenuto solo in un secondo tempo la separazione tra il materiale da conservarsi in uno piuttosto che in un altro istituto» p.101.

²⁷ Ivi, p. 102.

archivi»²⁸. È l'opinione che, dieci anni più tardi, sarà ribadita e compiutamente espressa dal Cencetti nel suo *Sull'Archivio come Universitas Rerum*²⁹ che qualifica la biblioteca come destinataria di universalità volontarie e l'archivio come custode di universalità necessarie. È il prevalere della concezione dei giuristi latini, da Servio Mauro ad Ulpiano, per i quali l'archivio è il luogo privilegiato di conservazione degli atti pubblici. È una concezione giuridico formale, consolidatasi nel Medioevo, alla quale è estranea qualunque apertura verso concettualità informative e cognitive ma che, anzi, restringe il campo d'azione alla specificità giuridica e diplomatistica. Senza volersi qui addentrare in una spinosa disquisizione sullo *jus archiviale*, l'origine del problema è anche nella difficoltà di garantire a quanto conservato in biblioteca la pienezza della *fides* che invece era appannaggio del materiale archivistico. Ciò derivava, ovviamente, da una concezione limitante di quest'ultima tipologia di materiali e motiva le ritrosie del Casanova di fronte a tipologie archivistiche per così dire "anomale". Ma quando l'anomalia di queste tipologie è, di fatto, non più esistente e quando anche la limitazione formale della diversità mediale è venuta meno con l'avvento dei documenti post cartacei, quando – infine – assume piena dignità una

²⁸ Ivi, p. 251.

²⁹ GIORGIO CENCETTI, *Sull'Archivio come Universitas Rerum*, in «Archivi», anno IV (1937) pp. 7-13. «I singoli componenti della universalità biblioteca, i libri, oltre ad aver ciascuno la sua origine in un autore, un editore, un libraio che, di regola, sono diversi, hanno altresì fine proprio, raggiungibile con mezzi propri: hanno cioè una propria autonomia originaria [...] Ben diversamente stanno le cose per ciò che riguarda l'archivio. Qui i singoli componenti, le carte [...] non solo provengono dal medesimo individuo, aggregato familiare o ente, ma poiché costituiscono nient'altro che uno fra i mezzi usati dall'ente o individuo per raggiungere i propri scopi, portano in loro stessi, fin dall'origine, il vincolo della destinazione comune».

tipologia *border line*, come la letteratura grigia³⁰, non libro e non documento, forse anche la suddivisione rigida teorizzata in altri contesti assume contorni più sfumati³¹.

Il mutare dei tempi e delle situazioni ha postulato altre possibili sinergie e sollevato, ovviamente, ulteriori dubbi e incertezze. Anche il rapporto tra le scienze dell'Informazione e la Documentazione³² non è riconducibile a parametri univoci di valutazione. Sia perché è la concettualità stessa di scienza dell'informazione a presentare dei problemi epistemologici sia perché all'interno delle scienze documentarie nel mentre è chiaro un possibile utilizzo strumentale dell'ingegneria del *software*, si è ingenerata, in molti casi, una sostanziale commistione e sovrapposizione di competenze³³ per ciò che attiene le scienze dell'informazione.

Del resto la stessa definizione data da Taylor (1967) al momento della creazione dell'American Society of Information

³⁰ Cfr. WILMA ALBERANI, *La letteratura grigia*, Roma, 1992. In tal senso più che il termine di "letteratura grigia" indicante la letteratura non convenzionale – è esplicativo quello di "Ephemeral materials" usato negli Stati Uniti che identifica qualunque tipologia documentale transeunte e non convenzionale.

³¹ Da più parti, ormai, nella letteratura specializzata, si fa strada l'idea della collaborazione tra le varie discipline documentarie specie in presenza di massicce innovazioni tecnologiche. Cfr. TOMMASO GIORDANO, *Futuro No problem. Convergenza tecnologica e patrimonio culturale*, «AIB-Web», n. 4 (1999); MARIA GUERCIO, *Rischi e promesse dell'innovazione tecnologica. I conservatori del patrimonio documentario e la cooperazione fra archivisti e bibliotecari*, «AIB-Web», n. 2 (2001).

³² A tal proposito, pur se datato, si veda JEAN MEYRAT, *La Documentazione: elementi per un riesame*, in A. M. PACI (a cura di), *La Documentazione*, cit., pp. 91-100.

³³ Cfr. HUBERT FONDIN, *La Scienza dell'Informazione: postura epistemologica e specificità disciplinare*, in «AIDA Informazioni», anno 20, n. 1 (gennaio-marzo 2002), pp. 5-24.

Science: «La scienza dell'informazione indaga le caratteristiche e il comportamento dell'informazione, le forze che governano il processo di trasferimento e la tecnologia necessaria al trattamento dell'informazione per ottimizzarne l'accessibilità e l'uso», conferma questa sovrapposizione dichiarando la necessità di una stretta interdipendenza di tali scienze da una pluralità di altri apporti disciplinari, dalla linguistica alla psicologia cognitiva, alla tecnologia, la quale ultima è vissuta con una forte accentuazione strumentale. Partendo proprio dalla necessità dell'interazione multipla ed utilizzando la teoria del Piaget³⁴ «Ingwersen elabora la mappa epistemologica delle interconnessioni delle scienze dell'informazione. Al livello superiore, che chiama dell'interdisciplinarietà, colloca la comunicazione, le scienze cognitive e le scienze dei sistemi; al livello inferiore o livello applicato, situa la documentazione e la biblioteconomia; al livello intermedio, di transizione o disciplinare, trovano posto la psicologia, la linguistica, la scienza dell'informazione, la sociologia, la scienza dei computer. All'interno della rete generale, la scienza dell'informazione ha una sua specificità di link con il livello transdisciplinare e con le discipline della documentazione e della biblioteconomia»³⁵.

Se oggetto delle scienze dell'informazione è, quindi, lo studio di un processo, mentre la documentazione e le scienze documentali studiano un oggetto: il documento, una differenziazione sembrerebbe, quindi, individuabile proprio nella diversificazione dell'ambito di interesse.

«D'altronde lo si vede bene attraverso le formule. Le une difendono l'oggetto documento e/o s'interessano al sistema tecni-

³⁴ JEAN PIAGET, *L'épistémologie des relations interdisciplinaires*, in *L'Interdisciplinarietà: problèmes d'enseignement et de recherche dans les universités*, OCDE, Paris, 1972.

³⁵ A. BALDAZZI, *Le scienze dell'informazione e le teorie della transizione*, in «AIDA Informazioni», anno 20, n. 1 (gennaio-marzo 2002), p. 28.

co (specifiche di un mestiere) mentre le altre vantano il “dispositivo persona” o il “sistema sociale” (quadro di ricerca). Donde le divergenze di punti di vista sulla formazione universitaria. Donde le incomprensioni reciproche sulla certificazione. Donde l’incompatibilità del contenuto tra rivista professionale e rivista scientifica»³⁶.

Ciò che, in punta di penna, sembra chiaramente delineato conosce – però – una ulteriore commistione allorquando Claire Guinchat e Michel Menou parlano, di «sciences et techniques de l’information et de la documentation»³⁷ e si sforzano di trovare una differenziazione concettuale tra il mestiere dell’informatico e le scienze dell’informazione e non già tra queste e la documentazione.

«Certes la ligne de démarcation existant entre le deux disciplines est aujourd’hui mal établie, mais le métiers de l’information de demain ne passent pas seulement par les voies de l’informatique. [...] Des nombreuses facettes du métier échappent totalement à l’ordinateur [...]»³⁸, concludendo – infine – che «Tous ces métiers reflètent dans ces appellations différents le dynamisme d’une unique profession: celle de spécialiste du transfert de l’information».

Sicuramente in queste affermazioni hanno un peso rilevante i condizionamenti e le esperienze personali degli autori, non sempre assimilabili a quelle di altri contesti geografici, ma, il concetto di base è che esistendo un’unica catena documentale che, partendo dai documenti da trattare genera, alla fine, dei prodotti di varia complessità, le scienze dell’informazione e le scienze documentarie si differenziano solo per l’occuparsi di momenti

³⁶ Ivi, p. 16.

³⁷ C. GUINCHAT–M. MENOU, *Introduction générale aux sciences et techniques de l’information et de la documentation*, Unesco, Paris, 1990.

³⁸ Ivi, p. 518.

diversi dello stesso processo non potendosi, in nessun modo, scindere processo ed oggetto dello stesso processo³⁹.

Il valore legale del documento elettronico e la possibilità della gestione totalmente non cartacea del ciclo di vita hanno apportato dei significativi cambiamenti sfumando ulteriormente le polisemie – a volte isteriche – del termine “documento” – obbligando ad una più puntuale definizione concettuale e avviando la concretizzazione di un unico oggetto informativo, risultato virtuale e transeunte della query estemporanea la cui durata corrisponde al ciclo di vita dell'informazione cui si riferisce. Al di là dei problemi legati alla concreta applicazione delle disposizioni sui documenti post cartacei un “oggetto-documento” totalmente immateriale ha – comunque – cominciato a viaggiare nel cyberspazio e la sua valenza informativa ha iniziato a confrontarsi con quella degli altri *e-media* andando a costituire serbatoi di conoscenza capaci di alimentare il web. Il controllo strategico dei contenuti e la loro fruizione e gestione è diventato il vero *core business* degli ultimi anni. La funzione di mediazione culturale tra i contenuti e gli utenti è diventata la professione emergente.

L'irrompere prepotente delle ICT e il mutare del quadro normativo di riferimento ha spostato l'asse di interesse dalla conservazione del supporto materiale, del quale la fruizione era un atto necessario ma conseguente, alla gestione dei contenuti lad-

³⁹ «A la multiplication et à la diversification des utilisateurs répond celle des services et produits fournis. La prolifération des termes qui désignent les diverses unités d'information (bibliothèques, archives, bibliothèques spécialisées, centres ou services de documentation, centre ou services d'analyse, services de liaison, banques de données, bases de données, médiathèques, services d'orientation, services de compilation des données, ecc.) traduit la richesse des possibilités d'information documentaire. En fait, ces diverses unités se différencient réciproquement par l'aspect de la chaîne documentaire su'elles traitent en priorité». Ivi, p. 31.

dove la conservazione dell'oggetto tangibile è completamente scindibile dalla fruizione. La scelta, operata da alcuni istituti di conservazione, dallo Smithsonian Institute agli Archivi Nazionali del Canada⁴⁰, di conservare i documenti su supporto fotografico e di utilizzare il digitale per l'attività corrente non obbedisce solo ad una logica funzionale ma è una precisa scelta culturale che, separando i due momenti, consegna i contenuti al mercato globalizzato dell'informazione ed i "documenti" all'ambito della ricerca storica.

Il vero punto nodale diventa la possibilità e la capacità di padroneggiare e gestire una quantità rilevante di oggetti informativi il cui numero e la cui complessità è in rapidissima crescita. Le tradizionali tecniche di classificazione archivistica, di catalogazione bibliografica e di indicizzazione, pur se utili, mostrano vistosamente i loro limiti di fronte ad una variabilità tipologica e semantica senza precedenti. La conoscenza esplicita prende coscienza dell'esistenza della conoscenza tacita e della necessità della reciproca integrazione. Nasce l'era del Knowledge Management e del Semantic Web.

Alle più tradizionali forme di trasmissione del sapere scritto si aggiungono «les médias nouveaux» che Otlet chiama – genericamente – «substituts du livre». «A la fin de sa vie Otlet développait des thèmes décrits dans son *Traité de documentation* et portant sur des machines qui fonctionneraient comme un cerveau collectif du monde, exploreraient l'espace-temps et pourraient être consultées depuis la maison»⁴¹.

⁴⁰ Cfr. PETER Z. ALDESTEIN, *La Permanence de l'information numérique*, in «Accès à l'information, questions de préservation», «CITRA, XXXIV», (1999), pp. 146-156.

⁴¹ CHARLES VAN DEN HEUVEL-BOYD RAYWARD-PIETER UYT-TENHOVE, *L'Architecture du Savoir; Une recherche sur le Mundaneum et les précurseurs européens de l'Internet*, in «Associations Transnationales», 1.2 (2003), pp. 17-18.

Archivistica e Documentazione tra passato e presente

ANTONIO ROMITI

Il panorama delle discipline attinenti all'archivistica ha fatto registrare in questi ultimi anni significative modificazioni in conseguenza sia dello sviluppo collegato con l'introduzione delle nuove tecnologie, sia delle innovazioni introdotte da recenti disposizioni normative. La necessità di offrire possibili approfondimenti ha condotto ad una situazione nella quale accanto agli aspetti tradizionali, per i quali esistono ormai consolidati parametri teorici e strutturali, si presentano e crescono nuovi settori che trovano la loro ragione di essere nelle sollecitazioni che provengono dalla società.

Le nuove realtà solo eccezionalmente si sviluppano partendo da elementi del tutto innovativi e, anzi, quasi sempre trovano la loro motivazione in altre già esistenti, con le quali anche in passato si sono confrontate al fine di ritagliare propri spazi utili a garantire un'esistenza, per quanto possibile, autonoma.

Con un rapido sguardo al passato, rileviamo come il settore archivistico sia stato contrassegnato da un contesto, nel quale l'Archivistica si trovava a convivere con altri due naturali elementi strutturali rappresentati rispettivamente dalla Paleografia e dalla Diplomatica: a conferma di tale situazione, si può notare come già nella seconda metà dell'Ottocento, nel rispetto di questa triplice configurazione si istituirono presso gli Archivi di Stato le Scuole, gloriosamente ancora esistenti, che avevano lo sco-

po di formare gli archivisti. Nella teoria e nella sostanza queste erano considerate le tre discipline fondamentali, utili e necessarie per la formazione degli archivisti.

È evidente come ognuna di esse fosse contraddistinta da proprie qualificazioni tecniche e scientifiche e come la loro aggregazione fosse strumentale nella prospettiva della formazione dell'archivista. Questa situazione ha assunto un preciso significato in un procedimento che ha mirato da un lato all'autonomia dell'archivistica con riferimento alle altre due discipline e dall'altro ad una distinzione tra la diplomatica e la paleografia. In ambito universitario si verificò inizialmente un particolare sviluppo dei settori *paleografici* e *diplomatistici* mentre, per avere un insegnamento dell'*archivistica* come disciplina autonoma, si dovette attendere la metà degli anni Venti del Novecento, quando Eugenio Casanova ottenne l'incarico di impartirla nell'Università di Roma, presso la Facoltà di Scienze Politiche.

Se è vero che la *paleografia* svolge una funzione in prevalenza strumentale per la formazione della figura professionale dell'archivista, che ha la necessità di conoscere le antiche scritture per poter leggere le carte e quindi per poterle riordinare, è parimenti vero che la *diplomatica* ricopre un ruolo ancora più strutturale, poiché affronta quegli aspetti storici, giuridici e metodologici che, partendo dal documento singolo possono contribuire a risolvere molte delle problematiche che attengono alla gestione ed alla organizzazione della produzione archivistica.

È proprio da alcune riflessioni sul rapporto tra *archivistica* e *diplomatica* che si può giungere ad una maggiore consapevolezza circa le correlazioni che intercorrono tra l'*archivistica* e *documentazione*; tale operazione di raffronto deve tuttavia essere condotta avendo presenti ed evidenti le possibili assimilazioni, ma anche le discrasie che emergono nell'operazione di comparazione tra *diplomatica* e la *documentazione*.

Siamo senza dubbio consapevoli che tali tematiche sono di estesissima portata e che per molti aspetti fuoriescono dalle spe-

cifiche conoscenze che attengono a chi opera nel settore archivistico ma, nonostante tali oggettive difficoltà, crediamo che possano essere affrontate se osservate da un luogo di osservazione, quello archivistico, che consente di introdurre alcune valutazioni in una prospettiva propria che potremmo definire *di settore*.

La dottrina non ha trascurato in passato il quesito riguardante il rapporto tra *archivistica* e *diplomatica*, nel quale l'archivio emerge per la sua caratteristica di corrispondere ad una entità *complessa, organica* e legata da un *vincolo originario e naturale*, mentre con la diplomatica l'indagine scientifica si rivolge all'*elemento singolo* che viene affrontato per conseguire in primo luogo la sua conoscenza strutturale. In sostanza, mentre l'attività *archivistica* si rivolge alla documentazione per riordinarla e renderla fruibile basandosi sulla necessaria conoscenza delle *procedure formative* dei complessi documentari, quella propria del diplomaticista si dirige allo studio delle problematiche generali e particolari che attengono alla formazione, alla organizzazione e alla funzione del *singolo documento*.

È evidente come questa distinzione, proposta dall'osservatorio archivistico, possa avere i limiti di tale prospettiva e non possa essere considerata esaustiva; non può ignorarsi che anche la *diplomatica*, nell'indagine sulla documentazione ad essa pertinente, affronta poi il documento non solo e non tanto nella sua singolarità, bensì lo considera nell'articolato rapporto formativo di una determinata realtà contestuale, nella quale assume senza dubbio una grande rilevanza quella caratteristica della *complessità* che, come si è premesso, rappresenta uno dei principali pilastri della dottrina e della realtà archivistica.

Ne consegue, quindi, che con riguardo agli elementi teorici fondanti, le differenze esistenti tra la *diplomatica* e l'*archivistica* parrebbero tendere ad assumere prospettive sempre più affievolite, sino a ridursi ai minimi termini, producendo un avvicinamento tale da rischiare di annientare quelle oggettive distinzioni che effettivamente esistono. È necessario pertanto considerare

con maggiore profondità tale indicazione di assimilazione per rilevare come, mentre in relazione alla *gestione archivistica* il carattere della *complessità* costituisce una condizione ed una *finalità* essenziale e irrinunciabile, in quanto senza di esso alla documentazione non può attribuirsi la qualificazione archivistica, per il *diplomatista* tale caratteristica, pur essendo di grandissimo rilievo, può anche non essere presente o riconosciuta e non costituisce quindi un preliminare ostacolo assoluto alla attività.

Sappiamo che non è corretto tracciare distinzioni nette e categoriche tra le due figure in quanto molti sono gli elementi di contatto: crediamo d'altra parte che sia sufficiente rilevare come anche l'archivista, nel suo impegno rivolto alla conoscenza ed alla organizzazione delle *complessità organiche*, non possa non affrontare aspetti diretti ad acquisire elementi di approfondimento circa le *singole unità archivistiche*, assumendo quindi un comportamento che per molti aspetti, non esclusi e non secondari quelli più strettamente tecnici, si avvicina a quello del diplomata. La differenziazione tra le due realtà rimane comunque evidente in quanto si riferisce in particolare alle *finalità*.

Le tre discipline, archivistica, paleografia e diplomatica, nelle Università hanno percorso durante il secolo scorso, in linea generale, strade non sempre uniformi, talora procedendo su percorsi vicini, talora distaccandosi in pur paralleli tragitti, determinati in prevalenza da opportunità e orientamenti accademici, ma anche in considerazione delle loro funzioni, senza aprire comunque nuovi spazi di dibattito teorico. Tuttavia, se non grandi interrogativi si sono presentati per la qualificazione del rapporto con la *paleografia latina* che si è mossa in un proprio e ben delineato ambito scientifico e didattico, non altrettanto può dirsi per le interrelazioni della *diplomatica* e dell'*archivistica*, in considerazione della loro visibile contiguità.

D'altra parte, l'archivistica aveva definito molti dei propri aspetti teorici nel corso delle riflessioni che avvennero in particolare tra gli anni Trenta e gli anni Settanta, grazie a Giorgio

Cencetti ed agli archivisti che in seguito hanno operato: l'individuazione del *vincolo* originario, naturale e necessario quale caratteristica distintiva dell'*archivio*, aveva condotto a conclusioni piuttosto definite nelle quali, come si è premesso, assumeva una posizione di assoluta preminenza l'elemento distintivo rappresentato dai menzionati caratteri della *organicità* e dalla *complessità*.

Fu proprio a seguito di tali basilari conclusioni che si aprirono nuovi spazi di analisi e di discussione: se agli archivisti era riconosciuto il compito di occuparsi di materiali complessi, se i diplomaticisti si consideravano impegnati in particolare nello studio di fonti che non necessariamente dovevano presentarsi strutturate in modo organico, per le quali l'ambito cronologico si circoscriveva al medioevo, con estensioni all'età moderna, negli ultimi anni non sono mancate estensioni verso carte recenti, con una disciplina che è stata intitolata *diplomatica contemporanea*.

In proposito, l'orientamento attuale prevalente si mostra contrario a tale ampliamento che rischierebbe di incidere sulla natura stessa di una disciplina che per vocazione e per tradizione è quasi esclusivamente medievistica; nel medesimo tempo deve affermarsi che essa non può essere vista neppure come strettamente archivistica, per la primaria attenzione verso i documenti, osservato più in attinenza alle loro prerogative giuridiche e amministrative, che alle correlazioni che tra loro intercorrono.

È in questo contesto che emerge la necessità di approfondire il rapporto esistente tra *archivistica* e *documentazione*, puntualizzando così aspetti di attualità che si riferiscono all'ambito della gestione pratica, ma che nel contempo sollecitano motivi di un non insignificante preliminare interesse da parte dell'indagine teorica e metodologica. Il tema deve essere affrontato in correlazione con l'analisi della evoluzione della stessa *professione dell'archivista* la quale, almeno fino ai primi anni Settanta del secolo scorso, rimase strettamente ancorata alle figure medievistiche e al massimo modernistiche, per la formazione delle quali la

conoscenza della *paleografia* e della *diplomastica* erano ritenute del tutto essenziali e fondamentali.

In seguito, ne fu un segnale l'abolizione della prova scritta della *lingua latina* tra le discipline richieste nei concorsi per l'accesso alla carriera statale, si ebbe uno sviluppo che, collegato con le innovative tendenze tecnologiche, fece orientare questa figura verso il settore *contemporaneo*; tale procedimento, fu affiancato da molti altri fattori, tra i quali non possiamo non ricordare la drastica riduzione dell'insegnamento del *latino* nelle scuole medie di primo e secondo livello, che ha prodotto una conseguente riduzione degli interessi verso la conoscenza della documentazione medievale e un incremento delle attenzioni verso criteri di organizzazione e di gestione degli archivi contemporanei e di quelli in formazione.

Le nuove tendenze, sostenute dalle recenti disposizioni normative e dalle incidenze prodotte dalla rivoluzione tecnologica, hanno contribuito a sviluppare con maggiore rapidità anche alcuni dei principi che già apparivano presenti nella tradizionale gestione archivistica, ma che mostravano la necessità di adattarsi alle nuove incombenti realtà. In primo luogo si è riproposto l'antico quesito che si è osservato nel rapporto tra *archivistica* e *diplomastica*, per verificare quali potessero essere le assimilazioni e le distinzioni rilevabili, questa volta, tra l'*archivistica* e la *documentazione*.

Da una prima sommaria e superficiale analisi, anche in relazione alla materia lessicale, potrebbe ripetersi quanto già affermato, puntualizzando che mentre l'*archivistica* si occupa degli archivi, considerati nella loro *organicità e complessità*, la *documentazione*, con una assimilazione solamente in linea teorica a quanto si è affermato per la diplomatica, si indirizza verso il *documento*, osservato in prevalenza nella sua realtà autonoma e non nel suo insieme naturale. Tale osservazione tuttavia non soddisfa pienamente in quanto anche con riguardo all'aspetto documentario non può escludersi l'elemento *organico*, mentre per alcune si-

tuazioni divengono basilari quei principi della *totalità* e della *complessità*, magari *volontaria*, che non vivono in contrapposizione, ma in alternativa, con quello della *naturalità*, come ha opportunamente specificato Adolf Brenneke.

La distinzione esistente tra *archivistica* e *documentazione* si può riferire quindi, a nostro avviso, più alle *attività gestionali* che alle *strutture concettuali* riferite alle fondamenta delle due discipline, sebbene non deve ignorarsi che le procedure formative si basano su presupposti spesso decisamente diversificati e diversificanti. Riteniamo tuttavia che il rapporto tra *documento* e *archivio* non possa essere osservato in una visuale contraddittoria e che, anzi, pur nelle naturali prerogative alle quali si è fatto cenno, debba essere preso in considerazione in una ampia visuale di interazione.

L'attività di *documentazione* si può distinguere quindi da quella *archivistica* non solo e non tanto per le procedure formative, ma anche e massimamente per le differenti *finalità* che, nel primo caso, sono mirate ad offrire agli studiosi materiali in prevalenza *informativi* mentre, nel secondo, gli orientamenti sono più dichiaratamente *descrittivi*. A seguito di tale considerazione, si rileva come le *attività archivistiche* risultino in linea generale più strettamente collegate con elementi oggettivamente definiti e determinati e come il *documentalista*, in conseguenza della preminente finalità *informativa* della sua attività, non sia sottoposto a particolari *limiti* nella individuazione delle tipologie delle fonti da affrontare e possa gestire per sua naturale connotazione la *comunicazione* concernente una molteplicità di aspetti culturali e scientifici che attengono a biblioteche, ad archivi, a fototeche, a discoteche, a videoteche, ad audiotecche, a centri di documentazione in generale, depositari e conservatori di una molteplicità di beni culturali e ambientali.

Con queste osservazioni non intendiamo tuttavia porre un limite invalicabile tra le due discipline: tutt'altro; riteniamo opportuno solamente presentare alcuni dei presupposti che possono

caratterizzare gli elementi costitutivi, senza nulla togliere alle possibilità di collaborazione e di completamento che tra loro possono essere individuate. Non possiamo poi dimenticare come l'organizzazione e le tecniche di realizzazione rappresentino i due momenti più importanti, nei quali si sperimentano le differenti metodologie che saranno poi applicate ai singoli oggetti di riferimento.

Quando affermiamo che nella gestione della *documentazione*, il criterio archivistico che attiene al *metodo storico* non trova una necessaria forma di concreta applicazione, in quanto, come si è fatto cenno, le attività non si orientano alla *descrizione oggettiva* del materiale, quanto alla ricerca dei metodi organizzativi dei *dati* che ad esso si riferiscono ai fini della loro diffusione, presentiamo solo un aspetto di un problema ben più complesso.

Le recenti disposizioni normative hanno introdotto in proposito ulteriori elementi di riflessione che possono essere utili per meglio comprendere e per approfondire alcuni degli aspetti in esame. Le definizioni adottate in proposito lasciano ampi spazi interpretativi, dai quali è possibile partire per meglio definire i limiti e le prerogative. Ci riferiamo in primo luogo al recente *Codice dei Beni Culturali*¹, nel quale sono stati individuati i beni culturali oggetto della specifica tutela ministeriale². Tra questi, sono riportati «gli archivi e i singoli documenti dello Stato, delle regioni, degli altri enti pubblici territoriali, nonché di ogni altro ente e istituto pubblico»³, introducendo una prima distinzione tra *archivi* e *documenti* che propone immediatamente qualche interrogativo poiché, considerata la basilare distinzione tra *archivio* e *raccolta* che, come vedremo, caratterizza la descrizione

¹ Codice dei Beni Culturali, Decreto Legislativo n. 42 del 22 gennaio 2004.

² Codice dei Beni Culturali, Parte Seconda. Beni Culturali, Titolo I Tutela. Capo I Oggetto della Tutela. Articolo 10 Beni Culturali e articolo 11 Beni oggetto di specifiche disposizioni di tutela.

³ Codice dei Beni Culturali, cit., articolo 10, comma 2, lettera b.

dei suddetti beni, nel caso specifico lascerebbe intendere una attribuzione della natura di *organicità*, anche a questa tipologia di documenti, pur nella segnalazione della loro individualità.

Altre simili considerazioni si applicano alla definizione che compare là ove si indicano «gli archivi e i singoli documenti, appartenenti a privati, che rivestono interesse storico particolarmente importante»⁴, per i quali i singoli documenti collegati con gli archivi dovrebbero essere ritenuti elementi integrativi e quindi caratterizzati, come quelli appena citati, da una naturale struttura *organica*. Nella sostanza, a livello operativo, la loro gestione dovrebbe competere, così come si prevede per gli archivi ai quali si ritengono strettamente collegati, direttamente agli archivisti. Tuttavia, quando ci chiediamo se tali materiali possano essere gestiti anche dai *documentalisti* siamo orientati per una risposta del tutto positiva, in quanto non si esclude che un archivio, dopo che è stato riordinato e descritto nel rispetto della sua *organicità* e del *vincolo naturale* che necessariamente lo lega, possa essere affrontato e descritto dal documentalista seguendo proprie metodologie che, in linea generale, si fanno riferire in prevalenza alla descrizione delle raccolte di documenti.

La situazione diviene ancora più incerta in occasione delle successive descrizioni, che compaiono nel comma 4 del medesimo articolo e in particolare rispettivamente alla lettera b. ove, nella prima parte del testo, si considerano beni culturali «i manoscritti, gli autografi, i carteggi», alla lettera d. ove ci si riferisce a «le carte geografiche e gli spartiti musicali aventi caratteri di rarità e di pregio», con una commistione piuttosto curiosa tra due categorie di fonti in verità poco collimanti e, infine, alla lettera e. concernente «le fotografie, con relativi negativi e matrici, le pellicole cinematografiche ed i supporti audiovisivi in genere, aventi caratteri di rarità e pregio».

⁴ Codice dei Beni Culturali, cit., articolo 10, comma 3, lettera b.

Analizzando attentamente queste definizioni emergono alcuni aspetti, talora anche non perfettamente condivisibili, che non consentono di operare quella netta distinzione tra *archivio* e *documento* che parrebbe avere guidato le intenzioni del legislatore al momento della stesura di questo testo normativo. Ci riferiamo in primo luogo alla aggregazione in un unico contesto dei *manoscritti*, termine senza dubbio generico che si presta per qualsivoglia interpretazione, con gli *autografi* e con i *carteggi*: sia gli uni, sia gli altri possono infatti fare parte tanto degli archivi quanto delle raccolte, pur con una naturale graduazione in quanto, tralasciando i *manoscritti*, per il loro onnicomprensivo significato e in considerazione di quanto premesso, gli *autografi* per loro natura risultano di solito elementi incardinati e poi estrapolati da archivi, quindi talora riuniti in raccolte, mentre i *carteggi*, costituiti da elementi di regola organicamente strutturati, coincidono comunemente con la realtà archivistica.

In attinenza alle *carte geografiche*, l'indicazione per la quale parrebbe volerle collocare in un contesto non archivistico può apparire piuttosto superficiale, dal momento che la loro origine è spesso legata ad una attività professionale o alla produzione di uffici e solo a seguito di eventi contingenti può aver trovato una diversa collocazione. La medesima riflessione ha valore anche per gli *spartiti musicali* i quali, pur quando non si trovano più presso il soggetto produttore, non di rado sono entrati a far parte di archivi di altri soggetti, che si caratterizzano con tale qualificazione tecnica per essere la memoria dell'attività di tali soggetti. Ci riferiamo, per proporre un esempio, all'Archivio Ricordi, nel quale è vero che si trovano manoscritti prodotti da altri soggetti esterni, ma la loro presenza contribuisce a formare un vero e proprio *archivio*, considerato quale memoria che rappresenta il risultato della attività di un ben connotato soggetto di impresa.

Di altra natura, ma non di univoca collocazione, può considerarsi il materiale *audiovisivo e fotografico*, senza trascurare la

convinzione che tra le due realtà sia necessario introdurre alcune opportune distinzioni. È opportuno premettere che le perplessità non hanno origine dalla valutazione della *natura del supporto*, in quanto siamo pienamente convinti che tale elemento non incida sulla qualificazione del materiale: per stabilire se siamo in presenza o meno di un *archivio*, è necessario infatti analizzare le *procedure formative*, mentre divengono del tutto insignificanti le tipologie dei beni che accolgono le registrazioni, le quali potranno eventualmente incidere sugli *aspetti operativi* del soggetto produttore, quali una maggiore o minore rapidità esecutiva e su *aspetti gestionali* della memoria, quali una maggiore o minore capacità conservativa.

Gli *audiovisivi* per loro natura solitamente non nascono con la struttura di *archivi* e sono di regola inquadrati nella estesa categoria delle *raccolte*. Per tale considerazione questo materiale non viene gestito da archivisti, né si ritiene che a loro debba competere. Ogni elemento che appartiene a questa categoria si caratterizza per la propria individualità e di conseguenza non contribuisce a creare con gli altri una situazione di *organicità*. Si tratta quindi di un settore di possibile pertinenza del *documentalista*, ma che anche in considerazione della sua frequente collocazione, viene gestito dai *bibliotecari*.

Nell'ambito di tale orientamento, riteniamo comunque che in questa documentazione non possano escludersi interessi *archivistici*, con riconoscimenti di interventi non tanto per finalità organizzative, quanto per soddisfare *procedimenti descrittivi*. Non tutte le rappresentazioni audiovisive, ma almeno una parte, riproducono avvenimenti culturali, politici, sociali aventi una valenza storica e costituiscono pertanto veri e propri *documenti* che sono la memoria di azioni di più soggetti. Le *videocassette*, così come le *audiocassette*, così come i *CD*, così come i *DVD* e così a seguire, per la loro funzione di contenere la documentazione relativa alla memoria di altri *soggetti attivi*, possono essere considerati dei veri e propri *archivi di concentrazione*; ricor-

diamo in proposito come gli Archivi di Stato, che appartengono a questa categoria, contengano *archivi prodotti da altri soggetti*.

A seguito di tali considerazioni si ritiene che, specialmente per alcuni livelli descrittivi, gli *archivisti* possano intervenire anche per la trattazione di questi materiali, pur nella consapevolezza che la funzione principale sia di organizzazione, sia di descrizione dovrebbe spettare agli esperti di *documentazione*. Da questa proposta abbiamo escluso tutte quelle rappresentazioni che non hanno valore di *documento*, che si avvicinano alla identificazione con il materiale librario e che spettano per la descrizione ai gestori di *biblioteche*.

Quanto sopra per alcuni aspetti può essere applicato anche al *materiale fotografico* nel quale più complessa è l'individuazione della *organicità* che si ha quando si rappresenta un evento filmato, ma non è del tutto impossibile rilevare tale elemento, specialmente quando si affrontano testimonianze riprodotte in immagini che sono state fissate nel corso di un apposito *servizio*. La presenza dei *rullini dei negativi*, ad esempio, può rappresentare la memoria di una consecutività di azioni di un evento e creare una organica struttura di immagini che testimoniano un avvenimento. In questo contesto non può escludersi quindi la figura dell'*archivista* il quale, a livello integrativo, può collaborare con altre figure.

È opportuno precisare che i materiali *audiovisivi* e *fotografici* compresi nella suddetta elencazione sono quelli che secondo la normativa e sulla base di un discutibile criterio che non può non essere che soggettivo, devono avere caratteristiche di *rarietà* e di *pregio*; gli altri tuttavia non sono del tutto esclusi dall'appartenenza alla qualificazione di beni culturali, purché la loro produzione «risalga ad oltre venticinque anni». Il testo normativo riporta tra tali elementi «le fotografie con relativi negativi e matrici, gli esemplari di opere cinematografiche, audiovisive o di sequenza di immagini in movimento, le documentazioni di

manifestazioni, sonore e verbali, comunque realizzate»⁵.

Questa descrizione, in raffronto con quella precedentemente riportata, amplia le possibilità di partecipazione archivistica in quanto, riferendo situazioni che trovano corrispondenza in particolare nelle *sequenze di immagini in movimento* e nelle documentazioni di *manifestazioni verbali e sonore*, lasciano aperte possibilità di intervenire con metodologie che attengono sia alla cultura della gestione del *documento singolo*, sia a quella della *realtà organica e complessa*. Si ribadisce pertanto che non debba essere la tipologia del materiale a condizionare le possibilità di intervento da parte di differenti, seppur affini, professionalità quali l'*archivista*, il *documentalista* e il *bibliotecario*, ma che debba essere la *finalità* dell'intervento a richiedere la tipologia dello specialista.

In relazione agli elementi sopra proposti, la figura del *documentalista* assume una posizione sempre più delineata, ma non pare poter decollare completamente, considerata la presenza di altre figure tradizionali, quali quella del *bibliotecario* e quella dell'*archivista* che si reggono sopra una secolare tradizione. In effetti nella gestione cartacea questa figura è considerata spesso «uno specialista del recupero e della gestione dell'informazione bibliografica a differenza di quanto accade negli altri paesi europei nei quali è uno specialista del recupero e della gestione dell'informazione tout-court mediante l'utilizzazione preminente di tecnologie informatiche e telematiche»⁶.

È vero nel contempo che la *tradizionale* figura dell'*archivista* sta andando incontro ad un affievolimento del proprio significato nell'ampio e articolato contesto nazionale che si sta svilup-

⁵ Codice dei Beni Culturali, cit., articolo 11, comma 1, lettera f.

⁶ R. GUARASCI, *Documenti, archivi e knowledge management: terminologia e semantica*, in *Documenti e archivi*, Università della Calabria, Centro editoriale e librario, 2002, p. 13.

pando a seguito dell'introduzione delle tecnologie informatiche, ottiche e telematiche, in quanto il mutamento del supporto va inevitabilmente ad incidere sulle attività concrete ed operative, così come è vero che vi sarà la necessità di adeguamenti per raggiungere nuovi equilibri professionali. L'introduzione del *Codice dell'amministrazione digitale*, d'altra parte, non ha arrecato significativi contributi a tale materia.

Le prospettive non sembrerebbero molto rosee se fosse così imminente e reale la prospettiva per la quale «nel corso dei prossimi anni si dovrà passare dalla gestione dei documenti tradizionale non solo e non tanto alla gestione elettronica dei documenti ma a quella che già da adesso viene definita come knowledge management 'un concetto molto avanzato di creazione e scambio di conoscenza, la possibilità di capire i fenomeni e di prendere decisioni in base ad informazioni elettroniche complete, integrate e filtrate secondo le esigenze di ciascuno»⁷.

I tempi non si sono mostrati così rapidi e così pressanti come si prevedeva, ma certamente queste riflessioni devono essere tenute in seria considerazione poiché, sebbene a più lunga scadenza, non potranno mancare innovazioni che incideranno, ci auguriamo positivamente, non solo sul futuro dell'*archivista*, ma anche su quello del *documentalista* attuale, pressato dalla probabile richiesta di figure professionali che attualmente non possono essere limpidamente prevedibili.

L'*archivistica* e la *documentazione* dovranno necessariamente assumere un orientamento rivolto a recepire le istanze che provengono dai diversi settori e che si sviluppano in conseguenza delle modificazioni dei supporti: non dovranno perdersi e disperdersi, come purtroppo si sta rischiando di fare con il manca-

⁷ FEDERICO VALACCHI, *Progettare per tutelare: linee guida per un intervento di valutazione ed ottimizzazione della risorsa archivio*, in *Documenti e archivi*, cit., p. 49.

to rinnovo del personale degli Archivi di Stato, le secolari conoscenze attinenti alle gestioni tradizionali, così come sarà necessario seguire le ricorrenti richieste che provengono dalle continue innovazioni tecnologiche.

Con attenzione a questi ultimi aspetti, ci pare che le distanze tra gli *archivisti* e i *documentalisti* delle nuove generazioni saranno inevitabilmente destinate a ridursi, nella considerazione che il loro lavoro deve comunque valutare gli scambi di interesse tra gli ambiti strettamente descrittivi a quelli più genericamente informativi. Si è scritto infatti che «grazie a tale evoluzione è possibile, da alcuni anni almeno, produrre e diffondere strumenti di ricerca informatizzati che costano poco, semplici nell'utilizzo, ma raffinati nella struttura e nella funzionalità, e soprattutto flessibili nelle forme e nei modi di condivisione del patrimonio informativo» e si è poi aggiunto che con la medesima applicazione si può dare vita «alla pubblicazione elettronica, sulla rete o su supporto ottico, di prodotti testuali di formato tradizionale o si possono predisporre banche dati autonome o integrate, interrogabili a distanza sia a testo libero che utilizzando griglie strutturate»⁸.

Da questa rappresentazione emerge la necessità di individuare non due, ma più nuove figure professionali di *archivista* e di *documentalista* nelle quali devono permanere i caratteri originari e tradizionali e nel medesimo tempo devono evidenziarsi gli elementi di sviluppo di conoscenze e competenze destinate a gestioni decisamente innovative. Come si è accennato, là ove le conoscenze si fondono e si sviluppano, le differenze tra le due basilari figure, quella dell'archivista e quella del documentalista tendono

⁸ MARIA GUERCIO, *Comunicare gli archivi: le opportunità della rete, i rischi della sperimentazione*, in *L'informatizzazione degli archivi storici e l'integrazione con altre banche dati culturali*, Provincia Autonoma di Trento, Servizio beni librari e archivistici, Trento, 2001, p. 14.

ad affievolirsi nella consapevolezza che saranno destinate a perseguire configurazioni che tendono alla nascita di nuove figure.

Recenti normative hanno già provveduto a tali finalità e così, evidentemente nel settore di competenza dell'archivista e del documentalista, compare, ad esempio, nella gestione digitale la figura del *Responsabile della conservazione*, previsto da una deliberazione CNIPA, al quale sono affidati molteplici compiti, alcuni dei quali strettamente legati alla professionalità archivistica, altri del tutto nuovi: deve «a. tenere evidenza delle caratteristiche e dei requisiti del sistema di conservazione in funzione della tipologia dei documenti (analogici o informatici) da conservare; b. organizzare il contenuto dei supporti ottici e gestire le procedure di sicurezza e di tracciabilità che ne garantiscano la corretta conservazione; c. archiviare e rendere disponibili, relativamente ad ogni supporto di memorizzazione utilizzato le seguenti informazioni: descrizione del contenuto, identificazione dei soggetti responsabili e dei compiti ad essi assegnati, indicazione delle copie di sicurezza; d. mantenere e rendere accessibile un archivio del software in gestione, nelle eventuali diverse versioni; e. documentare le procedure di sicurezza da rispettare per l'apposizione del riferimento temporale». Un ultimo sguardo è rivolto alla *formazione* che si ritiene necessaria per poter aspirare a ricoprire tale ruolo, includendo le conoscenze archivistiche, quelle informatiche, quelle concernenti il diritto e la diplomazia contemporanea e quelle relative alla organizzazione ai fini della «rimodulazione degli assetti organizzativi»⁹.

È questa dunque la linea da seguire nella sfida che deve interessare, prima che sia troppo tardi, degli *archivisti* e i *documentalisti* del presente.

⁹ S. PIGLIAPOCO, *La memoria digitale delle amministrazioni pubbliche. Requisiti, metodi e sistemi per la produzione, archiviazione e conservazione dei documenti informatici*, Santarcangelo di Romagna, Maggioli Editore, 2005, pp. 120-121.

Bibliografia, Biblioteconomia, Documentazione

PIERO INNOCENTI

§ 1. 1940. Una *climax* discendente:

Bibliografia → Documentazione → Biblioteconomia →
Archivistica

La scala gerarchica discendente che funge qui sopra da titolo di paragrafo è desunta dal *Vocabolario bibliografico* di Giuseppe Fumagalli (1940), opera che, curata e completata da altri, esce un anno dopo la morte dell'autore, risalente al 1939, e rappresenta la più importante sistemazione, a livello terminologico, delle discipline del libro e del documento nell'Italia d'anteguerra. Il *Vocabolario* di Fumagalli ha avuto una valenza divulgativa non indifferente contribuendo a creare il senso comune disciplinare per un buon numero di anni a venire dalla sua pubblicazione, tant'è così che la sua anastatica, dopo più di un sessantennio, ha avuto anch'essa una certa diffusione. Soprattutto, come è stato detto in letteratura (più volte e da tempo), di quel lavoro sono sostanzialmente tributari i successivi tentativi di raccolta della terminologia disciplinare, generalmente non caratterizzati da grande originalità e neppure coronati da grande successo: le due qualità non sono di necessità correlate (INNOCENTI 1989a, 1989b, 1990, 1994, 2007a).

Questa *climax* discendente porta da un arte-radice (la Bibliografia) a un'arte-applicazione (la Biblioteconomia), passando per la Documentazione, che è vista come espansione e perfezio-

namento dell'arte-radice; è posizionata collateralmente l'Archivistica, che sta in parallelo con la Biblioteconomia per la sottolineatura della sua caratteristica gestionale. In questa serie, insomma, l'Archivistica è un'estranea quanto a contenuto, ma è istanzziata alla Biblioteconomia dal dividerne la caratteristica di ramificazione applicativa (e non di espansione) dell'oggetto-base, cioè la struttura documentaria da cui prendono origine, che per la prima è l'archivio, per l'altra è la biblioteca. In uno schema di classificazione disciplinare ad albero (dendritico) starebbero ambedue al di sotto della voce-base, non al di sopra.

Fumagalli è abbastanza articolato nel trattare la voce *Bibliografia*, per la quale attinge ad una edizione del *Manuale* di Ottino, da lui stesso rivista, di metà anni Trenta (OTTINO 1935): «Della parola *bibliografia*, se è chiara la etimologia [...] – non è ugualmente chiara e sicura la definizione. Il Dupuy la disse, immiserendone il concetto, – “la scienza dei frontespizi”. Forse troppo sottilmente il Tommaseo la definì “l'arte, che può essere scienza, dell'apprezzare e del ben giudicare i libri nel loro valore estrinseco, che non può dall'intrinseco essere superato”. Noi la diremmo invece “lo studio dei libri in sé stessi, non per il loro contenuto intrinseco”, ma non ci sfugge la debolezza di una definizione negativa. L'essenza della bibliografia meglio potrebbe chiarirsi dicendo ch'essa abbraccia due parti: una generale, sintetica, di carattere principalmente erudito, che consiste nella storia del libro, delle forme che ha assunto attraverso le varie epoche della civiltà, delle peculiarità materiali che lo completano; l'altra, speciale ed essenzialmente pratica, consiste nella conoscenza particolareggiata dei libri e poiché la conoscenza di tutti singolarmente non sarebbe possibile per nessun uomo al mondo, essendo il loro numero presso che infinito, s'intende la conoscenza dei libri più rari, più belli, più utili» (FUMAGALLI 1940, p. 49).

La *Documentazione* è definita, sempre da Fumagalli, «l'arte di raccogliere documenti ossia materiale informativo sopra qua-

lunque argomento; così intesa, è un perfezionamento e un'amplificazione della bibliografia. Si veda il *Traité de documentation: le livre sur les livres* (Bruxelles, 1934), pubblicato da Paul Otlet, fondatore e direttore dell'Institut International de Bibliographie che appunto in coerenza di queste più ampie vedute ha cambiato il nome in Institut International de Documentation. A Parigi, dal 16 al 21 agosto 1937, è stato tenuto il Congresso Mondiale della Documentazione universale [*scil.*: Congrès mondial de la documentation universelle], al quale aderirono 460 fra enti e persone, provenienti da 46 paesi diversi. Vi furono rappresentati 31 governi e 48 organizzazioni internazionali. La Dichiarazione generale e le 21 risoluzioni adottate alla unanimità dal Congresso nella seduta di chiusura del 21 agosto sono state stampate in *Archives et Bibliothèques, 1937-1938*, n° 2, pag. 125-133» (FUMAGALLI 1940, p. 150; cfr. BALDAZZI 2003).

La definizione della *Biblioteconomia* è assai più sintetica, quasi limitativa: «l'arte e la dottrina di amministrare una biblioteca. I Tedeschi la chiamano *Bibliothekslehre* (ma lo Zoller, il Petzholdt e altri preferivano dirla *Bibliothekswissenschaft*) e i Francesi *Bibliothéconomie*; gli Inglesi e Americani più propriamente *Library Economy*» (FUMAGALLI 1940, p. 60: ci si esime dall'entrare in questioni relative alla terminologia interlinguistica fra le varie culture biblioteconomiche europee, sulla quale hanno scritto cose bellissime BARBERI 1981, BLUM 2007; per studî di analoga valenza in ambito archivistico-documentario: LODOLINI 1994, 1987).

L'Archivistica, infine, è trattata con questa asciutta definizione: «è la scienza che insegna a bene ordinare gli archivi. Il più recente volume di *Archivistica*, italiano, è quello di Eugenio Casanova (Siena, Lazzeri, 1928)» (FUMAGALLI 1940, p. 29). Si badi che questa è posizione teorica, non sottovalutazione, né tanto meno scarsa informazione: come si vede, viene citato il più aggiornato titolo italiano di archivistica, di assoluto spicco nell'orizzonte disciplinare dell'epoca e destinato a rimanere l'unico fi-

no ai lavori di Lodolini e Navarrini (cfr. LODOLINI 1980, poi 2004-2005; 1984, 1991, 2001; LODOLINI-NAVARRINI 2000-2001).

Potremmo riassumere forse così lo statuto concettuale che presiede alla trattazione definitiva di queste quattro discipline nel pensiero di Fumagalli:

1. in primo luogo si ha un *corpus* teorico, costituito dalla Bibliografia.
2. Se, in secondo luogo, è richiesto un particolare affinamento, esso è fornito dalla Documentazione; ciò si dà, in particolare, quando si applichi il recupero descrittivo e semantico della informazione a materiale più sfuggente e meno normalizzato, per quanto attiene la gerarchizzazione ontologica, di quanto ci abbia abituato il libro, sia manoscritto sia a stampa (di tale definizione si ricorderà Bisogno al momento di formularne una sua propria, più di quaranta anni dopo: «le sue [della Documentazione] differenziate attività si presentano al mondo esterno come comunicazione di conoscenze offerte sotto forma di dati elaborati con rigorosi criteri e omogenee procedure»; così BISOGNO 1987, richiamato in queste pagine anche da GUARASCI; l'antecedente si trova in BISOGNO 1980).
- 3.-4. Vengono, infine, le discipline gestionali, che sono la Biblioteconomia (da *nómos*, ovviamente, suffissato a *bibliothéke*: l'economia non ha a che farci) e l'Archivistica (da Archivio, a sua volta voce-madre, come lo è la biblioteca per la Biblioteconomia).

Potremmo anche concludere che, per Fumagalli, all'ordinamento alfabetico del suo *Vocabolario* è sottostante un preciso pensiero gerarchico, che però non ha il coraggio di spingersi fino alle estreme conseguenze ordinarie: nel che si potrebbe vedere un conato concettuale, che non riesce a compiersi, della logica del *thesaurus* – meglio si dirà «tesoro» (CROCETTI 2001, con ottimi argomenti, sviluppati poi da CALVITTI 2007a-b). Del resto,

come osserva a proposito di lessici, in queste stesse pagine, ADAMO: «A differenza dell'ordinamento alfabetico, privilegiato dai repertori lessicografici per la facilità di consultazione, l'ordinamento concettuale è senz'altro più coerente con l'impianto e le finalità delle raccolte terminologiche, spesso destinate alla consultazione di utenti che conoscono approfonditamente la materia. Ne risulta, infatti, una sorta di quadro sinottico del settore, che permette di valutare l'approfondimento dell'indagine compiuta».

§ 2. **2007.** Una nuova e diversa *climax* discendente:

Biblioteconomia → Bibliografia → Archivistica → Documentazione

Questa è invece la *climax*, sempre discendente, ma coi gradini posizionati diversamente, che ci si trova a percorrere quando, a quasi settanta anni di distanza dal *Vocabolario* di Fumagalli, e mentre ancor ne circolano, come si è visto, sia l'anastatica, sia la memoria, si è cercato di porre mano al lessico generale di un'area disciplinare in qualche modo paragonabile, anche se non con perfetta coestensione, a quella coperta dall'opera fumagalliana. In questo modo, ha preso forma *Biblioteconomia 2007*: vocabolario sistematico, non alfabetico, per il quale si è fatto ricorso ad un gruppo vasto di autori, che hanno dato vita ad un'opera collettiva, ovvero sia, più propriamente, collettanea: a più mani, s'intende, e non emulativa di Wu Ming, Luther Blissett, Bourbaki (o, perché no, *Wikipedia*: inevitabile il rimando alla trattazione di SCENINI 2008). Detto di passaggio, la tendenza a coprire collettivamente tematiche che due-tre generazioni or sono erano gestite da singole personalità della ricerca scientifica è fenomeno dei nostri giorni non limitato al caso in specie, e da ricondurre probabilmente non tanto all'incremento di materia disciplinare quanto al decremento di materia grigia che sembra caratterizzare i nostri tempi.

In *Biblioteconomia 2007* la disciplina eponima, ben lontano dal presentarsi come arte di gestione, occupa uno spazio pari a press' a poco il 70% dell'intera trattazione (700 pagine circa su 1.000 circa), e viene, per così dire, "spacchettata" in una serie di segmenti individuati dai simboli della CDD compresi fra 020 (che etichetta: *Biblioteconomia* → *Biblioteca* → *Tipi di biblioteca* → *Scienza dell'informazione* → *Archivistica* → *Documentazione*), e 028 (che etichetta *Lettura*). Il lettore ne troverà l'indice per notazione classificata alle p. xxiii-xxvi, e l'indice per soggetti considerati alle p. xxix-xxxv. Soprattutto troverà nella parte introduttiva una sintetica quanto chiara definizione: «Se la biblioteca è da intendersi come una raccolta ordinata di risorse documentarie – libri e altri materiali documentali –, la biblioteconomia [...] è la disciplina che si occupa di definire i criteri e le modalità per la loro raccolta e il loro ordinamento [...] La dinamica del 'sistema bibliotecario' comprende dunque tre fattori le cui reciproche interrelazioni costituiscono l'oggetto di studio della biblioteconomia: *materiali* [...], *attività* [...], *funzioni*» (SALARELLI 2007, p. 147).

All'interno di *Biblioteconomia 2007* anche la Bibliografia è stata "spacchettata", in questa successione: *Bibliografia* → *Bibliografia analitica* (*Bibliografia descrittiva*) → *Riferimenti bibliografici* (*Citazione bibliografica*) → *Strumenti e strategie di ricerca bibliografica* → *Bibliografie nazionali* → *Bibliografia Nazionale Italiana (BNI)* ? *Bibliografia di periodici di biblioteconomia*, dopo di che si sfocia in un grande mare che, come si è detto, dalla riva dello 020 (*Biblioteconomia*) porta alla lontana sponda dello 028 (*Lettura*), suddividendosi in un centinaio di soggetti. Al termine di ciascun sub-segmento è data una velocissima informazione bibliografica di approfondimento, consistente di due-tre voci (per il dettaglio, v. BIANCHINI, VADALÀ, FUGALDI-GAMBARI, CRUPI, SARDO, FUGALDI, REVELLI, tutti del 2007).

L'Archivistica occupa in *Biblioteconomia 2007* alcune pagi-

ne, finalizzate ad illustrare questa definizione data in esordio: «L'archivistica ha per fine la conservazione, l'organizzazione, la tutela e la valorizzazione delle attestazioni e delle memorie scritte e si occupa degli archivi nella fase di formazione, di deposito, e in quella comunemente definita storica. I fondamenti della disciplina si caratterizzano per la presenza di elementi teorici, metodologici, tecnici e pratici, che trovano applicazione in una complessa, articolata e diffusa attività professionale. L'archivistica gestisce le testimonianze prodotte da soggetti pubblici e privati, con riguardo alla natura e alla struttura degli archivi, ai criteri e alle procedure di formazione, ai principi di organizzazione, alle modalità di conservazione, ai compiti che attengono alla vigilanza, alla sorveglianza e alla tutela, nonché all'individuazione e alla realizzazione dei mezzi di corredo e degli strumenti per la ricerca» (ROMITI 2007a). Come si vede, una definizione che dice, con più parole e più scolorite, quanto già dice Fumagalli (per trattazioni di maggior peso specifico, anche se cronologicamente precedenti: LODOLINI 1987, nonché il meritamente diffusissimo e fondamentale LODOLINI 1984, arrivato alla 12. ed. nel 2005; PEARCE-MOSES 2005).

Infine, in *Biblioteconomia 2007* la Documentazione è, messa a fuoco con un procedimento di duplice approssimazione successiva, che si arresta, senza entrarvi, sulla soglia della fenomenologia definitoria: «Fatto distintivo che contribuisce a definire l'identità della documentazione è l'uso di tecnologie innovative», «Oggetto della documentazione non sono [*sic*] più un esiguo numero di oggetti "altri", né i nuovi supporti, bensì un nuovo paradigma epistemologico: il digitale, le risorse di rete» (CASTELLUCCI 2007, p. 186 e p. 188, che suggerisce, per approfondimento: DIOZZI 1998, CAROSELLA-VALENTI 1982, BISOGNO 1980). Più forte, come abbiamo visto sopra, la definizione di BISOGNO 1987, mentre ROMITI, in queste stesse pagine, preferisce svuotarla, estraniandola rispetto all'archivistica e delegandola piuttosto ai "gestori di biblioteche": il che conferma, *e contrario*, che la documentazione

può attualmente essere vista come istmo che collega territori attualmente contigui, ancorché in preda a vistose derive e sensibili bradisismi, come rileva – sul versante della Biblioteconomia – Innocenti, che in un rapido cenno ipotizza una schematica valutazione dei movimenti interni alle discipline del libro negli ultimi venti anni: «riesce [...] molto difficile trovare una sola biblioteconomia che vada dalla scienza dell'informazione all'analisi dell'evoluzione della presenza femminile nella professione bibliotecaria, dalla formazione professionale per la documentazione accademica a quella necessaria alla prima informazione e consultazione, dalla crisi della adorazione del totem delle scienze sociali in mezzo al cortile della biblioteca al cambiamento dei supporti da considerare come compito di copertura della biblioteca stessa (e quindi della Biblioteconomia). Nello stesso tempo, è impossibile non riconoscere la pertinenza di area di tali questioni e la leggittimità di una loro etichetta *lato sensu* biblioteconomica, perché il denominatore comune ne è comunque ancora evidente» (INNOCENTI 2007c, p. xxvi-xxviii, in part. p. xxvii). È pacata e probabilmente, come al solito, conclusiva una riflessione di Crocetti che, col pretesto di analizzare archivî e biblioteche, parla di Archivistica e Biblioteconomia, sancendone la forse non reversibile contaminazione: «Da una parte c'è il materiale archivistico di carattere, diciamo così, letterario, quello che illustra l'opera d'un autore o che contribuisce alla conoscenza e alla ricostruzione storica d'un movimento. [...]. Ma non appena il documento tocca più da vicino e specificamente la cultura e la storia d'un territorio, l'ospite d'elezione sarà la biblioteca di quel territorio. [...] Dei molti equivoci che nascono dal vedere una continuità senza soluzioni tra fondo antico e libri moderni nella biblioteca pubblica e dal non riconoscere, invece, la netta frattura, di significato e di gestione, che ormai tra esse esiste, si occuparono Emanuele Casamassima e il sottoscritto nel convegno di Arezzo, del 1977, sul tema "Università e beni culturali". Le indicazioni che in quell'occasione emersero, sul trattamento "archivistico" da riservare al

materiale antico in una biblioteca di pubblica lettura, possono abbastanza bene adattarsi, mi sembra, a quest'altro tipo d'acquisizione per una biblioteca che è della comunità» (CROCETTI 1981; eccellenti riflessioni che sviluppano le considerazioni di Crocetti in ROSSI 2007, CAVALLARO 2007, e cfr. LODOLINI 1995; *contra*, con argomentazione debole, ROMITI 2007b, che tratta il problema in modo nominalistico).

§ 3. **1991-2007.** Nel Web: una situazione piatta (*flat*)

La realtà virtuale è percepita oggi con tale concretezza (se si può usare questo ossimoro), che è ormai oggetto, a sua volta, di enciclopedie, repertori, sussidi introduttivi: il che è di per sé criterio per identificare e dare nome a nuove aree disciplinari, a quanto insegna qualunque guida professionale alla classificazione; si parla comunemente di identità in linea, di alfabetismo al silicio, di retorica della e nella Rete (BURBULES-CALLISTER 1997, BURBULES 1997c, 2002a-b, 2003a-b). Tutto ciò risale, ricordiamo e riassumiamo, all'ottobre del 1990, quando Tim Berners Lee scrive il primo *World Wide Web server (httpd)* e il primo *client (WorldWideWeb)*, *browser* ipertestuale, fondato sul principio "prendi-quel-che-vedi"; in proposito, per introdurre la problematica della affidabilità, potrebbe non essere inutile ricordare, sotto il profilo epistemologico, l'antica contrapposizione fra *dóxa* e *alétheia*. Nel dicembre dello stesso anno si apre una strada che inizia entro il Cern (Conseil Européen pour la Recherche Nucléaire, Ginevra, CH) e porta alla Internet nell'estate del 1991.

La situazione comunicativa (prima ancora che documentaria, almeno inizialmente: ma ognuna di quelle informazioni negoziate costituisce un documento...) che si viene a creare riprende con forza e rilancia, sotto il profilo bibliografico-documentario, tre questioni, di per sé già presenti, sia pure con accentuazioni diverse, nella tradizione della cultura occidentale:

1. Le tecniche di *trasmissione* del sapere, da modificare in conseguenza del modificarsi dei contenuti.
2. La capacità di *descrivere* gli oggetti della conoscenza, che si moltiplicano sia come ontologie, sia come oggetti, sia come istanze di oggetti: e qui sarebbe utile ricordare come pertinente l'antico argomento epistemologico del Terzo Uomo (PLATONE, *Parm.*; ARISTOTELE, *Met.*).
3. La distinzione di *forma* e *contenuto* negli oggetti del sapere.

Vediamo le tre questioni più da vicino.

1. *Adeguamento delle tecniche di trasmissione del sapere.* La prima questione emerge in parallelo col rapido sviluppo della Rete per usi civili e consiste nella necessità di un agguerrito programma di adeguamento delle tecniche educative a contenuti e forme di apprendimento diversi da quelli fino ad allora abituali, come rileva precocemente Nicholas Burbules, destinato a diventare, nell'ultimo quindicennio, uno dei più autorevoli esploratori della problematica, assieme con alcuni suoi co-autori (BURBULES-CALLISTER 1990, 1996; BURBULES 1997d, 1999a-d, 2000a-c; BURBULES-LAMBEIR 2003).

2. *Verifica della capacità di descrivere gli oggetti della conoscenza.* La seconda questione mette in evidenza la necessità di una più complessa analisi formale, cioè *descrittiva* degli oggetti della conoscenza in quanto oggetti fisici (HELLER 1990), e quindi, in quanto tali, veicoli di tradizione delle edizioni delle opere (BURBULES 1997a, FATTAHI 1997, FRBR 1998, SMIRAGLIA 2002b). Soddisfatta la fase descrittiva, occorrerà una più raffinata analisi contenutistica, cioè *semantica*, dei documenti così creati, recuperati, trattati (un approccio esemplare in MORRISSEY 2002; SMIRAGLIA 2002b). Si ricorda che «In termini molto generali, in informatica si suole chiamare “ontologia” lo schema concettuale di un certo dominio. Tale schema concettuale può assumere forme molto diverse, a partire da una semplice tassonomia fino ad arrivare a una vera e propria teoria logica. Normalmente

si tratta di una gerarchia di concetti correlati attraverso relazioni semantiche, ma le ontologie più elaborate forniscono anche regole (come assiomi e teoremi) che aiutano a specificare come è strutturato il dominio. Siamo quindi in presenza di artefatti ingegneristici costituiti da un vocabolario che descrive un certo dominio e un insieme di assunzioni esplicite che vincolano l'interpretazione dei termini del vocabolario in modo da avvicinarsi il più possibile al loro significato inteso», così BENZI-FERRARIO 2006, p. ii (che con qualche ironia lo definiscono anche «ormai un “must” di molti settori dell'informatica»). Il lavoro delle due autrici è rivisitato, in queste stesse pagine, da GAMBARARA-GIVIGLIANO, i quali a loro volta ricordano quattro definizioni di Ontologia che, in quanto largamente circolanti sui prodotti di Rete, sono citate dagli autori come senso comune, senza neppure indicazione della fonte: «1: “An ontology defines the basic terms and relations comprising the vocabulary of a topic area, as well as the rules for combining terms extensions to the vocabulary”. 2: “An ontology is an explicit specification of a conceptualization”. 3: «An ontology is a hierarchically structured set of terms for describing a domain that can be a skeletal foundation for a knowledge base”. 4: “An ontology provides the means for describing explicitly the conceptualization behind the knowledge represented in a knowledge base”». Ai nostri scopi servono la 2. e la 3.

«Istanza»: nel linguaggio politico e giuridico, è una richiesta formale (se ne veda l'ampio uso che ne fa in queste pagine ROLLI); nella logica aristotelica, istanza è una premessa che ne smentisce un'altra; nella programmazione orientata agli oggetti, istanza è per lo più sinonimo di oggetto; in documentazione è una delle *tot* manifestazioni di un oggetto (in queste stesse pagine: DE FRANCESCO, che si sofferma sui rapporti fra istanze documentarie e loro modelli). Verbo derivato è «Istanziare»: termine di largo uso nel linguaggio UML (Unified Modeling Language), comune dal 1996 per finalità di comunicazione aziendale (l'ultima versione, la 2.0, risale al 2004). Il sostantivo che ne indica l'a-

strazione è in inglese «Instantiation» (SMIRAGLIA 2005b, 2007a-b), in italiano «Istanziamento» (SERRA 2004, di ambito accademico napoletano) o «Istanzializzazione» (PONCHIROLI 2004, p. 119, di ambito accademico modenese).

3. *Distinzione di forma e contenuto negli oggetti del sapere.* La terza questione: si prende atto della coincidenza di forma e contenuto nella documentazione che si verrà a creare d'ora in avanti, a partire dal punto zero determinato dalla produzione di supporti digitali e non più analogici, in parallelo e non in sostituzione, per molti decenni e probabilmente per sempre, della documentazione su supporti tradizionali (*Formati* 1997, BURBULES 1997b, 1997e); le due modalità si contaminano fra loro e ciò rende particolarmente necessario attenersi ad un attentissimo protocollo descrittivo (è un intreccio col precedente punto 2.: INNOCENTI 1999, SMIRAGLIA 2008, riassuntivo della bibliografia allo stato).

Nell'arco di questo orizzonte, vengono pertanto a ri-parametrarsi le concezioni disciplinari consolidate. Il nuovo orizzonte afferma orgogliosamente di disporre di tutti i 360° che competono all'orizzonte: non è così, ma si comporta come se lo fosse, determinando intrichi complessi di fermate, percorsi retroflessi ed estroflessi, ripartenze (BURBULES, 1998a-b); ora, se è vero che, per citare un ormai celebre titolo comparso su «La Nación» di Buenos Aires il 17 dicembre 2000, “navigare in Internet è in qualche modo come andare nudi per strada”, è anche vero che all'interno dell'arco di quell'orizzonte si registra comunque la crescita non solo di nuovi contenuti, ma di nuovi linguaggi, il che ha portato, come si è detto sopra, a una nuova grammatica e a una nuova retorica.

Quali riferimenti terminologici fissa il Web, per quanto riguarda questo giardinetto disciplinare? Una rapida perlustrazione delle risorse di prima informazione, propedeutiche alla ricerca e disponibili solo via Web, porta a rilevare queste sintetiche definizioni. Biblioteconomia: «è la disciplina che studia l'organizzazione delle biblioteche. Il suo campo di indagine va dallo

studio dell'acquisizione dei beni librari alla loro catalogazione, indicizzazione e conservazione, dalla localizzazione delle biblioteche alle procedure di accessibilità ai servizi da esse offerte» (*Wikipedia*).

Bibliografia: «Con la parola bibliografia (dal greco antico *bibliograph[?]a*, formata dalle radici di *bíblōs*, libro, e *graphé*, scrittura) si possono intendere concetti alquanto diversi tra loro, e cioè:

- un'opera (o parte di essa) contenente un elenco sistematico di pubblicazioni di vario genere aventi relazione con un argomento specifico, articolato secondo criteri variabili ma sempre definibili chiaramente, basati su regole comuni di ricerca, trascrizione e descrizione delle opere;
- la disciplina che ha il compito di studiare le regole pratiche di compilazione di repertori, cataloghi e inventari;
- la teoria che studia struttura, articolazione e fortuna storica dei repertori.

Per quanto riguarda la prima accezione del termine, va fatta un'ulteriore suddivisione. Nelle pubblicazioni accademiche della tradizione italiana, infatti, sono state di solito ben distinte le sezioni bibliografiche che citano fonti da quelle che menzionano letteratura secondaria»: quest'ultima osservazione, lasciata cadere così, quasi come fatto meramente stilistico, è in realtà concettualmente molto importante, come vedremo poco sotto.

Archivistica: «L'archivistica è una disciplina ausiliaria della Storia, che studia l'archivio, vale a dire l'insieme dei documenti che sono stati prodotti da un ente (pubblico o privato) nel corso della sua attività».

Documentazione: non trattata, e verrebbe di chiosare con un punto esclamativo. Assistiamo, infatti, all'apparente paradosso che la Internet ignora la Documentazione come possibile disciplina-semaforo atta a smistare il traffico variopinto che vi si muove dentro: proprio quella Internet che, interconnettendo, crea la

maggior quantità documentale che si conosca sotto forma di ontologie, che a loro volta si prestano ad essere istanziate, anche solo per mero uso personale, in quantità n , e che a volte divengono ancora, a loro volta, ontologie ed oggetti da cui si producono nuove istanze: e così via, con un processo di *progressione* virtualmente all'infinito, che è reciproco e speculare al processo di *regressione* ad infinito del modello della istanzializzazione: «Dodici pile di libri alte quanto la distanza che separa la terra dal sole o, se preferite, tre milioni di volte la quantità d'informazioni contenute nei libri fin qui scritti dall'uomo: a tanto equivarrebbero i 161 exabyte (milioni di gigabyte) d'informazione digitale prodotta nel corso del 2006; sempre nel 2006 ben 1,1 miliardi di persone hanno frequentato il luogo per eccellenza dove l'informazione digitale viene cercata e recepita: il World Wide Web» (CUNA 2007, p. 915). Il paradosso, solo apparente, deriva dal fatto che quel materiale è documentariamente *flat* (piatto), non ha gerarchie strutturate, ma è accumulato ed esplorato con procedura logica sequenziale o al più trattata linguisticamente con applicazioni estensive della *edit distance* (1965) di Levenshtein.

A testimonianza di come la copertura sia tuttora in costruzione, in *Wikipedia* le pagine ricomprese nella categoria «Biblioteconomia e scienza dell'informazione» sono (15 gennaio 2008): *Bibliotecari*, *Biblioteconomia*, *AIMS*, *AgMES*, *Bibliografia*, *Biblioteca*, *Biblioteca nazionale*, *Biblioteche dell'Abruzzo*, *Cinquecentina*, *Classificazione Colon*, *Classificazione Decimale Universale*, *Classificazione bibliotecaria*, *Classificazione decimale Dewey*, *Collazione*, *Deposito istituzionale*, *Descrittore*, *Emeroteca*, *Festschrift*, *Ibidem*, *Indice KWIC*, *International Standard Bibliographic Description*, *Letteratura grigia*, *MARC*, *Mediateca*, *OPAC*, *OpenUrl*, *Reference*, *Servizio Bibliotecario Nazionale*. Come si vede, si va da concetti generalissimi, vere e proprie etichette di disciplina, a concetti minimi, come la trattazione dell'avverbio *ibidem*, o l'esposizione della definizione di *Festschrift*.

Tralasciamo per un momento di esaminare meglio alcuni termini che abbiamo già introdotto ed usato come ovvî, e che forse tali non sono: ontologia, oggetto, istanza; vi torneremo in conclusione. È di particolare interesse, ai fini del discorso che stiamo argomentando, esaminare, fra le voci di questa pagina enciclopedica, quella che abbiamo messo in evidenza col neretto: «Collazione». La collazione, dice l'anonimo estensore del testo, è termine che può indicare 1. una sezione bibliografica, 2. un'operazione filologica o 3. un istituto giuridico.

1. In Biblioteconomia, si precisa, la collazione è l'area destinata ad ospitare le caratteristiche fisiche e materiali di una pubblicazione.
2. In Filologia, invece, la collazione è il confronto di diverse copie di un manoscritto o di un testo a stampa, atto a rilevare le differenze tra le copie e tra queste e l'originale, con la finalità di ricostruire i rapporti di dipendenza testuale, verticali o contaminati ("stemma").
3. In ambito giuridico, infine, la collazione, disciplinata in Italia dagli articoli 737-751 del Codice civile, è l'obbligo imposto a taluni soggetti (figli legittimi e naturali e loro discendenti legittimi e naturali, nonché il coniuge) che accettino l'eredità, di conferire alla massa che compone il patrimonio del defunto quanto dal medesimo ricevuto in vita per donazione diretta o indiretta, salvo che il testatore non li abbia da ciò dispensati.

In letteratura biblioteconomica, il punto 1. è consolidato nelle abitudini e nella terminologia catalografiche, non occorre soffermarvisi: la collazione è appunto, nel lessico catalografico italiano, l'area della descrizione nella quale si inseriscono, secondo un algoritmo preciso, gli elementi che caratterizzano la composizione fisica del documento: misure e/o formato, paginazione/cartulazione, illustrazioni.

Quanto al punto 2., la riflessione sulla collazione filologica

come metodo per analizzare i rapporti fra il documento e le sue rappresentazioni è stata da tempo introdotta in dottrina bibliografica: ha dell'ovvio. Meno ovvio è che essa serve anche ad analizzare le famiglie di registrazioni all'interno di un catalogo bibliografico (LEAZER-SMIRAGLIA 1999), cioè i rapporti fra la prima descrizione elaborata da una agenzia bibliografica (o documentaria) abilitata, e le descrizioni secondarie che da essa derivino (INNOCENTI 1999, p. 123 fig. 1, 124 fig. 2, 125 fig. 3). La collazione può essere definita dunque come un controllo di autorità in senso globale, che non si limita all'area descrittiva/intestativa destinata all'autore (SMIRAGLIA 1992), normalizzandone l'uso attraverso la costituzione dell'archivio di autorità, né tanto meno si limita alla operazione di descrizione fisica. Si deve, anzi, notare che la riflessione sui vari aspetti del contenuto testuale (*opera*) – e quindi per slittamento concettuale, del documento e la sua rappresentazione grafica, su qualsivoglia supporto, è questione del massimo rilievo per la qualità dell'informazione, sia a prescindere dal supporto, sia a prescindere dai contenuti. Si applica dunque a testi a correre sia di letteratura sia di saggistica (sia di *fiction*, cioè, sia di *non-fiction*), ma anche musica, artefatti (se lo si preferisce: manufatti), documenti d'archivio, ipostasi di riproduzioni documentarie, etc. A questo proposito, si vedano in particolare l'ampia trattazione di SMIRAGLIA 2001, 2002a per il quadro generale, ma anche: VELLUCCI 1997, che per la musica ha individuato due categorie di istanzializzazione, che ha definito rispettivamente «notational transcription» e «musical presentation» (trascrizione mediante notazione ed esecuzione). COLEMAN 2002, che per la modellistica scientifica si spinge a dire che i singoli modelli scientifici potrebbero essere trattati dal punto di vista descrittivo (e successivamente, di conseguenza, semantico) come opere. SMIRAGLIA 2004 e 2005a, che esemplifica l'applicazione a manufatti archeologici, ID. 2006 e 2008 per l'esemplificazione archivistica, tratta in modo particolare dall'archivio della Marina mercantile degli Stati

Uniti. Infine, ICMD 2007 per l'applicazione a oggetti museali.
 Questi gli schemi collazionali a cui si fa qui riferimento:

Sia α l'originale, ossia il testo (documento) citato;
 sia A il primo livello di descrizione, cioè la sintesi bibliografica, catalografica o di astrazione dell'originale, ricavata dall'originale stesso;
 siano $A_1 \rightarrow \dots \rightarrow A_n$ le tante utilizzazioni di A (le tante descrizioni) quante ne abbiamo bisogno, in cui A rimane fonte:

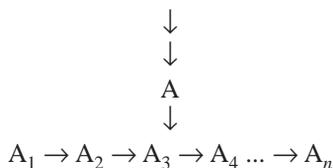


Figura 1.

La verifica dei due passaggi verticali, a valle rispettivamente di α e di A, è affidata al metodo filologico della *collazione*, così come è affidata a questo anche la verifica di eventuali corruzioni nei passaggi orizzontali $A_1 \rightarrow A_n$.

Consegue da questa impostazione che possono verificarsi guasti, anche gravi, quando lo schema si presenti invece così:

Sia α l'originale, ossia il testo (documento) citato;
 sia A la sintesi bibliografica o catalografica dell'originale, ricavata da α ;
 sia B copiato da A, C copiato da B, D copiato da C, e così via a piacimento fino all'ennesima, con moltiplicazione della possibilità d'errore per progressione aritmetica; ogni passaggio potrà infatti contenere almeno un errore in più del precedente:

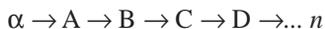


Figura 2.

La conseguenza ultima della mis-applicazione/non-applicazione della procedura di collazione è messa in evidenza in Fig. 3:

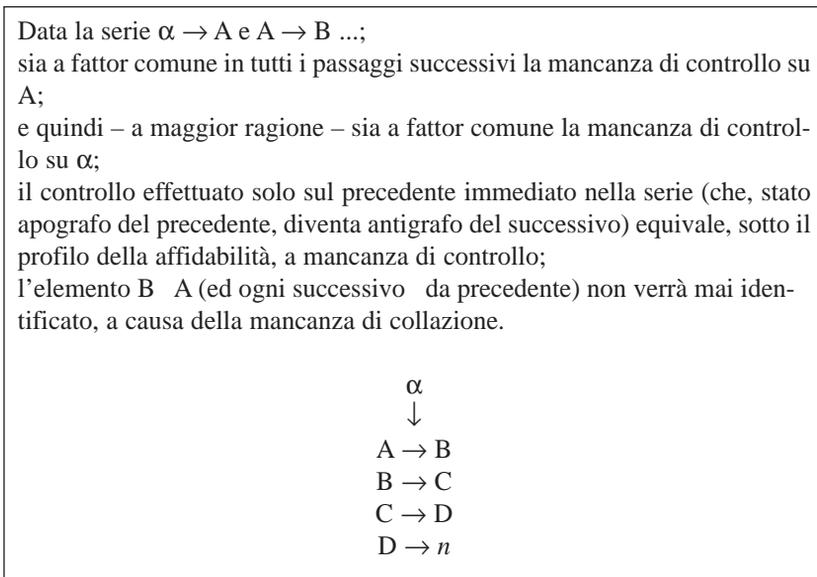


Figura 3.

Il primo livello è quello delle fonti primarie, i successivi rappresentano le fonti secondarie. La mappa che ne consegue è un vero e proprio indice topologico, seguire il quale aiuta a percorrere la trasmissione del dato informativo e la sua variantistica, nel migliore dei casi, o corruzione, nel peggiore (LEAZER-FURNER 1999; sugli aspetti “etic” della affidabilità del dato: BURBULES 2001). Distingueremo dunque:

- Àmbito bibliografico-documentale analogico (cartaceo), in cui si è tradizionalmente sviluppata e consolidata la problematica di cui sopra, in Fig. 1; in esso la procedura di

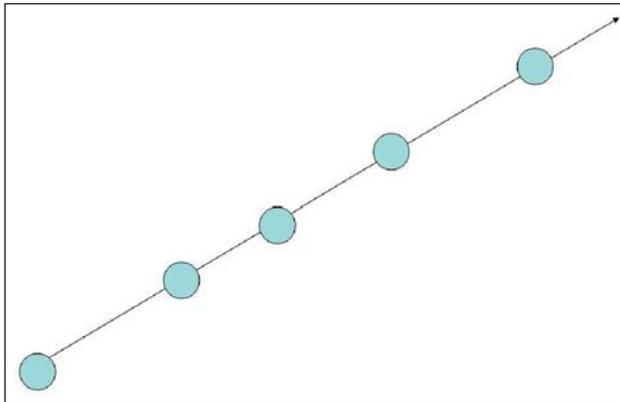
fabbricazione manuale delle descrizioni si è spinta fino alle più complesse applicazioni di meccanizzazione e automazione della scrittura, influenzando l'evoluzione dei codici di catalogazione (FATTAHI 1997); la Documentazione ha, dal canto suo, storicamente incentivato questa sorta di approccio, incrementando la sua sorgiva propensione verso le soluzioni più avanzate e favorendola con ogni mezzo (CASTELLUCCI 2007). È fase che abbiamo irrevocabilmente alle spalle.

- Àmbito bibliografico-documentale ibrido (sia analogico sia digitale). In esso si applicano metodologie sia del precedente àmbito, sia di quello successivo. È il nostro presente.
- Àmbito bibliografico-documentale esclusivamente digitale. Fu descritto come divertente futuribile da un autore di fantascienza la cui vita si è svolta tutta entro il secolo 20. (1904-1988): «Attorno ai Terrestri, semplicemente appoggiate sul pavimento metallico, c'erano altre forme, diverse da quelle incontrate fino a quel momento: forse erano l'equivalente dei mobili, l'arredamento di quell'enorme ambiente [...] – Un ufficio – disse infine Duncan Griffith. – O una stanza di controllo – disse Ted Buckley, l'ingegnere meccanico della spedizione. – Può darsi che fosse la loro zona residenziale – disse Taylor. – Oppure forse era un'officina meccanica – azzardò Jack Scott, il matematico. – I signori hanno considerato – chiese Herbert Anson, il geologo – che potrebbe trattarsi di qualcosa di assolutamente diverso? Potrebbe essere qualcosa che non ha alcuna corrispondenza con le cose a noi note. – L'unica cosa che possiamo fare – disse Spencer King, l'archeologo – è tradurla come meglio possiamo, nei termini a noi conosciuti. La mia ipotesi è che potrebbe essere una biblioteca» (SIMAK). Lo riprende nel 2007 Cuna, con una accurata trasposizione nel presente quotidiano di chiunque si

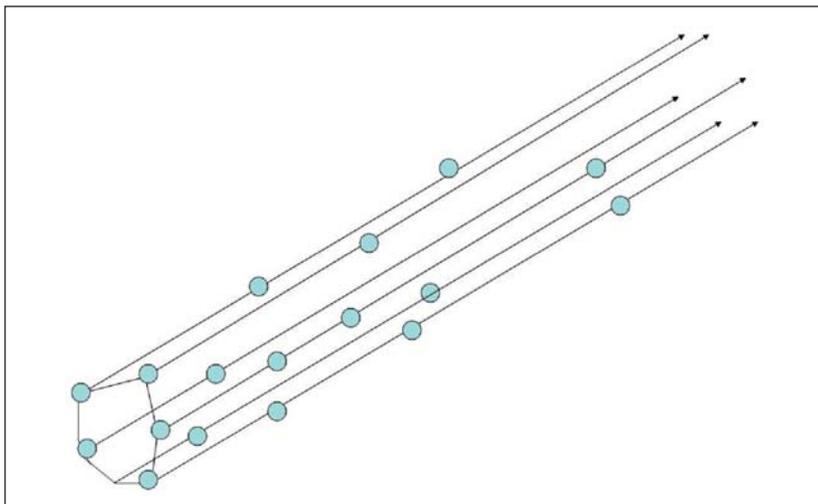
confronta con la propria scrivania elettronica: «l'informazione personale è caratterizzata da attributi per così dire unici, che non hanno riscontro nell'informazione generale. In altri termini e in relazione ai rispettivi trattamenti, gli attributi soggettivi della prima sono cosa ben diversa dagli attributi oggettivi o generali che vengono associati alla seconda in sistemi informativi destinati a servire le molteplici esigenze di un'utenza variegata per professione, educazione e ambiente socio-culturale» (CUNA 2007, p. 916). Nel 2005, è stato messo a punto il concetto di «Personal Space of Information», ovvero “psi”:, come ricorda lo stesso Cuna, secondo il quale esso può forse rappresentare un buon punto di (ri)partenza: «Psi contains a person's books and paper documents, email messages (on various accounts), e-documents and other files (on various computers). A Psi can contain references to Web pages. A Psi also includes applications, tools» (JONES-BRUCE 2005, p. 9). Non è una applicazione di Archivistica, non lo è di Documentazione, non lo è di Biblioteconomia, ma in quanto unifica la gestione (possibile) di istanze di documenti di archivio, di documenti di varia natura, di fonti primarie e secondarie della Bibliografia, implica la necessità di una conoscenza orizzontale dei postulati delle tre discipline, sollecitandole. È in discussione se i confini (ovvero le sfere d'influenza) delle quattro discipline, se vi includiamo l'archivistica, siano per reggere a tale sollecitazione, se non siano destinati ad ulteriormente co-estendersi, se non siano per arroccarsi su posizioni tradizionali (LODOLINI 1990; GUERCIO 2002).

Questa è, infine, il grafo della traiettoria di come una istanzializzazione si sviluppa nel tempo (SMIRAGLIA 2008, p. 21):

Traiettoria nel tempo secondo SMIRAGLIA 2008 (p. 21).



Questa è la rete disegnata da un complesso di istanzializzazioni caratterizzato dalla medesima matrice ideativa, sempre secondo SMIRAGLIA 2008 (p. 22):



È evidente come sia nel primo sia nel secondo caso i nodi del grafo siano occupabili indifferentemente da istanze dell'una, dell'altra o della terza estrazione, che costituiscono insiemi (o sot-

to-insieme), a prescindere dalla sorgente (la prima applicazione della teoria dei grafi alla divulgazione bibliografica si è avuta con *Enciclopedia* 1977-1984). È possibile che il linguaggio a ciò necessario sia considerato ancora come una applicazione del linguaggio bibliografico (su questo insiste la tradizione tedesca secondo la linea Schneider-Blum-Totok), come pure che esso sia considerato una morfologia e una grammatica nuova; certamente non ancora una sintassi, perché sotto il profilo della affidabilità di copertura esso lascia per il momento sensibilmente a desiderare, e invoca procedure miste di trattamento in ogni caso in cui obiettivo sia non tanto l'estensione (ampiezza del raggio di esplorazione) quanto l'intensità (controllabile qualità del dato). Ciò che è essenziale è praticare con sistematicità la logica del trattare i dati documentari sempre e comunque come procedenti da un contesto, che va ricostruito con cautela, perché il processo che ha portato ad estrarne potrebbe aver determinato fattori di squilibrio o di inquinamento del processo a ritroso: diversamente non si hanno dimostrazioni, ma argomentazioni, e non possiamo prescindere dall'applicare loro una logica che individui enunciati veri da enunciati falsi (QUINE 1960a-b; CAROFIGLIO 2007, p. 206).

Ciò di cui si occupano, in qualunque posizione reciproca le si collochino, Archivistica, Bibliografia, Biblioteconomia, Documentazione (menzionate ora in ordine esclusivamente alfabetico) sono da un lato "dati", dall'altro situazioni (riconducibili cioè a persone) in cui i dati vengono usati; il secondo ambito fa parte per definizione dell'argomentazione retorica, cioè della suasiona, che userà i dati per gli scopi che si vorranno raggiungere; ma la raccolta dei dati fa parte della dimostrazione. Dati falsi, nel senso di una rappresentazione inesatta della realtà documentaria che vogliono rappresentare, determineranno non già persuasione, ma inganno.

Come ben sappiamo, è una linea estremamente facile da valicare: *quod demonstrandum erat?*

RISORSE DOCUMENTARIE CITATE:

- ARISTOTELE, *Met.* = ID., *Metafisica*, 1079 a 13; 990 b 17.
- BALDAZZI 2003 = ANNA BALDAZZI, *Donne e documentazione: il respiro di un secolo*, in «AIDAinformazioni, ISSN 1121-0095, trimestrale, 21., n° 1, gennaio-marzo 2003, versione elettronica: <<http://www.aidainformazioni.it/pub/baldazzi12003.html>> (cons. 28 gennaio 2008).
- BARBERI 1981 = FRANCESCO BARBERI, *Biblioteche in Italia: saggi e conversazioni*, Firenze, Giunta regionale Toscana-La nuova Italia, 1981 (Collezione Archivi e biblioteche. 3). Raccolta di scritti risalenti ad anni precedenti.
- BENZI-FERRARIO 2006 = MARGHERITA BENZI, ROBERTA FERRARIO, *Prefazione di «Networks»*, 6, 2006, numero monografico: *Ontologia/Ontology*.
- BIANCHINI 2007 = CARLO BIANCHINI, *Bibliografia*, in *Biblioteconomia 2007*, p. 74-91.
- Biblioteconomia 2007 = Biblioteconomia. Guida classificata*, diretta da Mauro Guerrini, condirettore Gianfranco Crupi, a cura di Stefano Gambari, collaborazione di Vincenzo Fugaldi, presentazione di Luigi Crocetti, Milano, Editrice Bibliografica, 2007.
- BISOGNO 1980 = PAOLO BISOGNO, *Teoria della documentazione*, Milano, Angeli, 1979, rist. 1980.
- BISOGNO 1980 = ID., *Presentazione di*: MARIA PIA CAROSELLA, MARIA VALENTI, *Documentazione e Biblioteconomia*, Milano, Angeli, 1987.
- BLUM 2007 = RUDOLF BLUM, *Bibliografia. Indagine diacronica sul termine e sul concetto*, nota introduttiva di Attilio Mauro Caproni, [2007], 258 p. Tit. or.: *Bibliographia. Eine wort- und begriffsgeschichtliche Untersuchung*, Frankfurt a. M., Buchhändler Vereinigung, 1969; trad. di Maria Letizia Fabbri; recensito da INNOCENTI 2007 e 2008.
- JONES-BRUCE 2005 = WILLIAM JONES, HARRY BRUCE, *A Report on the NSF-Sponsored Workshop on Personal Information Management* (Seattle, WA, 2005): <<http://pim.ischool.washington.edu/final%20PIM%20report.pdf>>, p. 9; cit. in CUNA 2007, p. 916, n. 4.
- BURBULES 1997a = NICHOLAS CONSTANTINE BURBULES, *Misinformation, Malinformation, Messed-up Information, and Mostly Useless Information: How to Avoid Getting Tangled Up in the Net*, in *Digi-*

- tal Rhetorics: Literacies and Technologies in Education-Current Practices and New Directions*, Chris Bigum, Colin Lankshear *et alii*, editors (Canberra, Department of Employment, Education, Training, and Youth Affairs/Brisbane, Queensland University of Technology, 1997), p. 109-120. Rist. col tit.: *Struggling with the Internet*, «Campus Review» (August 13-19, 1997), p. 19-22.
- BURBULES 1997b = ID., *Digital Texts and the Future of Scholarly Writing and Publication*, «Journal of Curriculum Studies», 30., n° 1, 1997, p. 105-124.
- BURBULES 1997c = ID., *Rhetorics of the Web: Hyperreading and Critical Literacy, Page to Screen*, in *Taking Literacy Into the Electronic Era*, Ilana Snyder, ed. (New South Wales, Allen and Unwin, 1997), p. 102-122.
- BURBULES 1997d = ID., *Technology and Changing Educational Communities*, «Educational Foundations», 10., n° 4, 1996, p. 21-32; ripubblicato in forma ridotta come *Technology: What we Haven't Worried About*, in «Education Digest», 62., n° 9, 1997, p. 53-57.
- BURBULES 1997e = ID., *Web Publishing and Educational Scholarship: Where Issues of Form and Content Meet*, «Cambridge Journal of Education», 27., n° 2, 1997, p. 273-282.
- BURBULES, 1998a = ID., *Aporia: Webs, Passages, Getting Lost, and Learning to Go On*, in *Philosophy of Education 1997*, Susan Laird ed., Urbana, Ill., Philosophy of Education Society, 1998, p. 33-43.
- BURBULES, 1998b = ID., *Questions of Content and Questions of Access to the Internet*, «Access», 17., n° 1, 1998, p. 79-89.
- BURBULES 1999a = ID., *Education and Global Communities*, in *Globalisierung: Perspektiven, Paradoxien, Verwerfungen* («Jahrbuch für Bildungs- und Erziehungsphilosophie», Bd. 2., Walter Bauer, Wilfried Lippitz, Winfried Marotzki, Jörg Ruhloff, Alfred Schäfer, Christoph Wulf editors, (Hohengehren, Schneider Verlag, 1999), p. 125-141.
- BURBULES 1999b = ID., *E-education: As Technology Pervades Schools, New Issues Arise*, (intervista), «Northwest Herald» (December 28, 1999), p. 1, 6.
- BURBULES 1999c = ID., *Teaching at an Internet Distance: The Pedagogy of Online Teaching and Learning*, in UNIVERSITY OF ILLINOIS, *Teaching at an Internet Distance*, Seminar: John Regalbutto (presidente), Rachell Anderson, Hassan Aref, N. Burbules, Allan Cook, Cleora D'Arcy, Mark Gelula, David Hansen, Michael Loui, Babet-

- te Neuberger, Linda Smith, Ronald Smith, James Sosnoski, Sandra Theis, Robert Wengert, Donald Wink, Charles Woodbury. Disponibile in linea: <<http://www.vpaa.uillinois.edu/tid/report>> (1999); ultima cons. 20 gennaio 2008.
- BURBULES 1999d = ID., *Technology in Education: Who, Where, When, What and Why?*, «International Journal of Educational Technology» (July, 1999): <<http://www.outreach.uiuc.edu/ijet/v1n1/v1n1feature.html>>; ultima cons. 20 gennaio 2008.
- BURBULES 2000a = ID., *Does the Internet Constitute a Global Educational Community?*, in *Globalization and Education: Critical Perspectives*, N. C. Burbules, Carlos Torres, editors, New York, Routledge, 2000, p. 323-355.
- BURBULES 2000b = ID., *Globalization and Community* (trad.: *Welt-Gemeinschaft: Paradox oder Realitat?*), in «Neue Zürcher Zeitung», 12 Dezember 2000, p. 62.
- BURBULES 2000c = ID., *Navegar por Internet es como andar desnudo por la calle*, Intervista, «La Nación» (Buenos Aires, Argentina, 17 dicembre 2000, p. 3.
- BURBULES 2001 = ID., *Paradoxes of the Web: The Ethical Dimensions of Credibility*, «Library Trends», 49., 3, 2001, p. 441-453.
- BURBULES 2002a = ID., *Like a Version: Playing with Online Identities*, in «Educational Philosophy and Theory», 34. n° 4, 2002, p. 387-393.
- BURBULES 2002b = ID., *The Web as a Rhetorical Place*, in *Silicon Literacies*, Ilana Snyder ed., London, Routledge, 2002, p. 75-84.
- BURBULES 2003a = ID., *Dialogue in Virtual Spaces*, in *Dialog og Naerhet: Ikt og Undervisning*, Yvonne Fritze, Geir Haugsbakk, Yngve Nordkvelle, editors, Kristiansand, Norwegian Academic Press, 2003, p. 19-28.
- BURBULES 2003b = ID., *Virtual Reality*, in *Greenwood Dictionary of Education*, John Collins, Nancy O'Brien editors, Phoenix, Ariz., Oryx Press, 2003, p. 373.
- BURBULES-CALLISTER 1990 = ID., THOMAS A. CALLISTER, *Computer Literacy Programs in Teacher Education: What Teachers Really Need to Learn*, «Computers and Education», 14., 1, 1990, p. 3-7.
- BURBULES-CALLISTER 1996 = ID., *Knowledge at the Crossroads: Alternative Futures of Hypertext Environments for Learning*, «Educational Theory», 46., n° 1, 1996, p. 23-50.

- BURBULES-CALLISTER 1997 = ID., *Who lives here? Access to and Credibility Within Cyberspace*, in *Digital Rhetorics, Literacies and Technologies in Education. Current Practices and New Directions*, Chris Bigum, Colin Lankshear *et alii*, editors, Canberra, Department of Employment, Education, Training, and Youth Affairs, Brisbane, Queensland University of Technology, 1997, p. 95-108.
- BURBULES-LAMBEIR 2003 = ID., BERT LAMBEIR, *The Importance of New Technologies in Promoting Collaborative Educational Research*, in *Beyond Empiricism: On Criteria for Educational Research*, Paul Smeyers, Marc Depaepe, editors, Leuven, University Press of Leuven, 2003, p. 41-52.
- CALVITI 2007a = TIZIANA CALVITI, *Digitalizzazione ed indicizzazione di alcuni periodici di discipline del libro: un progetto di thesaurus*, tesi di dottorato di ricerca in Scienze bibliografiche, archivistiche, documentarie e per la conservazione e restauro dei beni librari ed archivistici (19. ciclo), discussa nell'a.a. 2006-2007 (giugno 2007).
- CALVITI 2007b = ID., *Digitalizzazione ed indicizzazione di alcuni periodici di discipline del libro: un progetto di tesoro di discipline del libro*, in *Una mente colorata Studi in onore di Attilio Mauro Caproni per i suoi 65 anni*, promossi, raccolti, ordinati da P. Innocenti, curati da C. Cavallaro, Roma, Manziana; Il libro e le letterature, Vecchiarelli, 2007, p. 465-484.
- CAROFILIO 2007 = GIANRICO CAROFILIO, *L'arte del dubbio*, Palermo, Sellerio, 2007 (La memoria. 734).
- CAROSELLA-VALENTI 1982 = *Documentazione e biblioteconomia: manuale per i servizi di informazione e le biblioteche speciali italiane*, a cura di Maria Pia Carosella e Maria Valenti, presentazione di P. Bisogno, Milano, Angeli, c1982 (Collezione Manuali professionali. 57). Ininterrottamente ristampato fino al 2000.
- CASTELLUCCI 2007 = PAOLA CASTELLUCCI, *Documentazione*, in *Biblioteconomia 2007*, p. 186-189.
- CAVALLARO 2007 = CRISTINA CAVALLARO, *Fra biblioteca e archivio. Catalogazione, conservazione e valorizzazione di fondi privati*, presentazione di Caterina Del Vivo, saggio introduttivo di Marielisa Rossi, Milano, Bonnard, 2007.
- COLEMAN 2002 = ANITA S. COLEMAN, *Scientific Models as Works*, «Cataloging & Classification Quarterly», 33., 2002, n° 3/4, p. 129-159.
- CROCETTI 1981 = LUIGI CROCETTI, *Problemi dell'intervento regionale*

- nell'acquisizione e sistemazione di archivi privati*, in Davide Lazzeretti e il Monte Amiata: protesta sociale e rinnovamento religioso. Atti del Convegno, Siena e Arcidosso, 11-13 maggio 1979, Firenze, Nuova Guaraldi, 1981, p. 367-372, cit. in SOLIMINE 2007.
- CROCETTI 2001 = ID., *Per un tesoro della biblioteconomia italiana*, «Bollettino Aib», 41., n° 1, marzo 2001, p. 7-19.
- CRUPI 2007 = GIANFRANCO CRUPI, *Strumenti e strategie di ricerca bibliografica*, in *Biblioteconomia 2007*, p. 107-134.
- CUNA 2007 = ANDREA CUNA, *Internet e pc. L'estensione virtuale della «biblioteca» privata*, in *Una mente colorata* cit., p. 915-924.
- DIOZZI 1998 = FERRUCCIO DIOZZI, *Documentazione*, Roma, Associazione italiana biblioteche, 1998 (ET. Enciclopedia tascabile).
- Enciclopedia 1977-1984 = Enciclopedia*, direzione di Ruggiero Romano, Torino, Einaudi, 1977-1984, 16 v.: 1: *Abaco-Astronomia*. 2: *Ateo-Ciclo*. 3: *Città-Cosmologie*. 4: *Costituzione-Divinazione*. 5: *Divino-Fame*. 6: *Famiglia-Ideologia*. 7: *Imitazione-Istituzioni*. 8: *Labirinto-Memoria*. 9: *Mente-Operazioni*. 10: *Opinione-Probabilità*. 11: *Prodotti-Ricchezza*. 12: *Ricerca-Socializzazione*. 13: *Società-Tecnica*. 14: *Tema/motivo-Zero*. 15: *Sistematica*. 16: *Indici* / [a cura di Renato Betti]. Alla tecnica di citazione praticata da questo repertorio è dedicato un capitolo specifico di INNOCENTI 1999.
- FATTAHI 1997 = RAHMATOLLAH FATTAHI, *Relevance of Cataloguing Principles to the Online Environment: An Historical and Analytical Study*, Ph. D. Th., University of New South Wales, Sydney, Australia, 1997. Disponibile all'indirizzo: <<http://profsite.um.ac.ir/~fattahi/fattahi.HTM>>, cons. il 17 settembre 2007).
- Formati 1997 = I formati della memoria. Beni culturali e nuove tecnologie alle soglie del terzo millennio*, a c. di Paolo Galluzzi e Pietro A. Valentino, Firenze, Giunti, 1997.
- FRBR 1998 = INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS, *Functional Requirements for Bibliographic Records*, Munich, K. G. Saur, 1998. Disponibile: <<http://www.ifla.org/VII/s13/frbr/frbr.htm>>, oppure: <<http://www.ifla.org/VII/s13/frbr/frbr.pdf>> (dicembre 2007).
- FUGALDI 2007 = VINCENZO FUGALDI, *Bibliografia nazionale italiana (BNI)*, in *Biblioteconomia 2007*, p. 140-141.
- FUGALDI-GAMBARI 2007 = VINCENZO FUGALDI, STEFANO GAMBARI, *Riferimenti bibliografici (Citazione bibliografica)*, in *Biblioteconomia 2007*, p. 97-107.

FUMAGALLI 1940 = GIUSEPPE FUMAGALLI, *Vocabolario bibliografico*, a cura di Giuseppe Boffito e di Giovanni De Bernard, Firenze, Olshki, 1940 (Biblioteca di bibliografia italiana. 16); rist. anast.: ivi, 1999 (stampa 2000). La 1. ed. fu acquisita da: **Ancona**, Biblioteca comunale Luciano Benincasa. **Biella**, – civica. **Bologna**, Archiginasio, – universitaria. **Cesena**, Malatestiana. Cremona, – statale. **Forlì**, – comunale Aurelio Saffi. **Ferrara**, – comunale Ariosteia. **Firenze**, – di Lettere e filosofia, Bncf, Laurenziana, Riccardiana. **Lucca**, – statale. **Mantova**, – comunale Roberto Ardigò. **Milano**, – centrale della Facoltà di ingegneria del Politecnico, – delle Facoltà di Giurisprudenza e Lettere e filosofia, – di scienze della storia e della documentazione storica dell'Università, Braidense, Centro bibliografico e di documentazione del Servizio biblioteche della Regione. **Napoli**, – nazionale Vittorio Emanuele 3. **Parma**, – Palatina. **Pisa**, – universitaria. **Pordenone**, – comunale. **Potenza**, – nazionale. **Rimini**, – civica Gambalunga. **Roma**, Angelica, – dell'Archivio di Stato, – dell'Istituto centrale per la patologia del libro, – dell'Istituto della enciclopedia italiana, – dell'Istituto italiano per l'Africa e l'Oriente, – di storia moderna e contemporanea, Bnc Vittorio Emanuele 2. **San Giovanni in Persiceto**, – comunale Giulio Cesare Croce. **Sassari**, – universitaria. **Torino**, – dell'Accademia delle scienze, – storica della Provincia. **Treviso**, – comunale. **Trieste**, – generale dell'Università. **Venezia**, – Querini Stampalia, Biblioteche della Fondazione Giorgio Cini. **Vicenza**, Bertoliana, – internazionale La Vigna. La 2. ed. è stata acquisita da: **Macerata**, – comunale Mozzi-Borgetti. **Milano**, – di scienze della storia e della documentazione storica dell'Università. **Padova**, – del Centro interdipartimentale di servizi di Palazzo Maldura. **Perugia**, – comunale Augusta. **Ravenna**, – della Facoltà di conservazione dei beni culturali. **Torino**, – civica centrale. **Viterbo**, – consorziale. **Voghera**, – Civica Ricottiana.

GUERCIO 2002 = MARIA GUERCIO, *Archivistica informatica. I documenti in ambiente digitale*, Roma, Carocci, 2002 (Beni culturali. 25); 2. rist. 2004.

HELLER 1990 = MARK HELLER, *The Ontology of Physical Objects. Four-dimensional Hunks of Matter*, Cambridge, Cambridge University Press, 1990 (Cambridge studies in philosophy).

ICMD 2007 = INTERNATIONAL COUNCIL OF MUSEUMS DOCUMENTATION STANDARDS GROUP, *Definition of the CIDOC Conceptual Referen-*

- ce Model, version 4.2.2*. Disponibile: <http://cidoc.ics.forth.gr/official_release_cidoc.html>, (dicembre 2007).
- INNOCENTI 1989a = PIERO INNOCENTI, *Dimmi quali parole sai e ti dirò che bibliotecario sei*, «Biblioteche oggi», 7., 1989, n° 3, maggio-giugno, p. 317-319.
- INNOCENTI 1989b = ID., *Biblioteca/biblioteche (Italia). La possibile voce di un possibile glossario di discipline del libro*, «Biblioteche oggi», 7., 1989, n° 3, maggio-giugno, p. 325-355.
- INNOCENTI 1990 = ID., *Bibliography: Teaching and Profession*, «Libraries & Culture», 25., n° 3, Summer 1990, p. 461-473 (University of Texas).
- INNOCENTI 1994 = ID., *Dimmi quali parole sai e ti dirò che bibliotecario sei*, «Biblioteche oggi», 12., 1994, n° 5, maggio, p. 68-70.
- INNOCENTI 1999 = ID., *Metodi e tecniche nella ricerca bibliografica (trilogia di Mary Poppins)*, Manziana, Vecchiarelli, 1999 (Bibliografia, Bibliologia e Biblioteconomia. Collana di Testi e Studi), rist. 2002: riprende rielabora ed aggiorna lavori del 1984-1985.
- INNOCENTI 2007a = *Biblioteche e studi per le biblioteche in Italia dopo il 1983. Una rassegna*, in *Scritti in memoria di Raoul Guêze*, a c. di Cristina Cavallaro, Manziana, Vecchiarelli, 2007, p. 209-229.
- INNOCENTI 2007b = ID., *A proposito di Bücherkunde: la Bibliographia di Blum e il Fra biblioteca e archivio di Cavallaro*: <<http://hdl.handle.net/2067/228>>. 12 ottobre 2007. Pubblicazione elettronica sull'Open Archive della Università della Tuscia, Viterbo.
- INNOCENTI 2007c = ID., *Caro Mauro, ovvero: l'imbalsamatore del Quirinale*, in *Una mente colorata* cit., p. xi-xxxii.
- INNOCENTI 2008 = ID., *A proposito di Bücherkunde: 1. La Bibliographia di Blum. 2. Il Fra Biblioteca e archivio di Cavallaro. 3. Due recensioni di Serrai: una risata le seppelli*, «Culture del testo e del documento», 9., 2008, n° 25, gennaio-aprile, p. 27-70.
- INNOCENTI-CAVALLARO 2003 = ID., CRISTINA CAVALLARO, *Passi del leggere. Scritti di lettura, sulla lettura per la lettura: ad uso di chi scrive e di chi cita*, 2 v., Manziana, Vecchiarelli, 2003.
- LEAZER-FURNER 1999 = GREGORY H. LEAZER, JONATHAN FURNER, *Topological Indices of Textual Identity Networks*, in *Proceedings of the 62nd Annual Meeting of the American Society for Information Science*, ed. Larry Woods, Medford (NJ), Information Today, 1999, p. 345-358.

- LEAZER-SMIRAGLIA 1999 = GREGORY H. LEAZER, RICHARD P. SMIRAGLIA. *Bibliographic Families in the Library Catalog: A Qualitative Analysis and Grounded Theory*, «Library Resources & Technical Services», 43., 1999, p. 191-212.
- LODOLINI 1980 = ELIO LODOLINI, *Organizzazione e legislazione archivistica italiana. Dall'unità d'Italia alla costituzione del Ministero per i beni culturali e ambientali*, prefazione di Giovanni Spadolini, Bologna, Pàtron, 1980 (Scienze storico-auxiliarie. 2). Rist. 1983, 1985: 4. ed. 1989.
- LODOLINI 1984 = ID., *Archivistica. Principi e problemi*, Milano, Angeli, c1984 (Manuali professionali. 59); 2. ed. 1985; 3. ed. 1986; 4. ed. 1987; 5. ed. ampliata 1990; 6. ed. ampliata 1992; 7. ed. ampliata 1995; 8. ed. ulteriormente ampliata 1998; 9. ed. 2000; 10. ed. 2002; 11. ed. 2004; 12. ed. 2005.
- LODOLINI 1987 = ID., *Note di terminologia archivistica: "archivio", "archivistica", "archivista", "archivistico"*, «Archivio storico italiano», 145., 1987, n° 532, p. 99-115.
- LODOLINI 1990 = ID., *"Gestione dei documenti" e archivistica: a proposito della convergenza di discipline*, «Rassegna degli Archivi di Stato», 50., 1990, n° 1-2, p. 85-117.
- LODOLINI 1991 = ID., *Lineamenti di storia dell'archivistica italiana. Dalle origini alla metà del secolo 20.*, Roma, NIS, 1991 (Beni culturali. 13); 3. rist. 1998.
- LODOLINI 1994 = ID., *Gli archivisti tedeschi dal 1500 al 1945 e la bibliografia archivistica internazionale in due opere di Wolfgang Leesch*, «Nuovi annali della Scuola speciale per archivisti e bibliotecari», 8., 1994, p. 219-226.
- LODOLINI 1995 = ID., *Manoscritti o documenti? Il caso di Camerino*, «Nuovi annali della Scuola speciale per archivisti e bibliotecari», 9. (1995), p. 167-174.
- LODOLINI 2001 = ID., *Storia dell'archivistica italiana. Dal mondo antico alla metà del secolo 20.*, Milano, Angeli, 2001 (Collezione Temi di storia. 23); 2. ed. 2003; 3. ed. 2004; 4. ed. 2006.
- LODOLINI 2004-2005 = ID., *Legislazione sugli archivi. Storia, normativa, prassi, organizzazione dell'amministrazione archivistica*, Bologna, Pàtron, 2004-2005, 2 v. Tit. dell'ed. precedente: *Organizzazione e legislazione archivistica italiana* (ivi, 1980, v.). Comprende: 1: *Dall'Unità d'Italia al 1997*. 2: *Dal 1. gennaio 1998 al 1. agosto 2004. Con appendice di testi normativi del 2004*.

- LODOLINI-NAVARRINI 2000-2001 = ID., ROBERTO NAVARRINI, *Archivistica speciale moderna e contemporanea*, Bologna, Pitagora, 2000-2001. Documento da proiettare a video, 10 videocassette (VHS): color., son. Comprende: [1]: E. LODOLINI, Lezione 1: *Concetti generali*; Lezione 2: *L'organizzazione archivistica statale*. [2]: ID., Lezione 3: *L'Amministrazione archivistica statale in Italia*; Lezione 4: *Natura degli archivi pubblici in Italia*. [3]: ID., Lezione 5: *La struttura di vertice e gli organi consultivi collegiali dell'Amministrazione archivistica*; Lezione 6: *Gli archivi di Stato (parte 1.)*. [4]: ID., Lezione 7: *Gli Archivi di Stato (parte 2.)*; Lezione 8: *Gli archivi di Stato (parte 3.)*. [5]: ID., Lezione 9: *Gli Archivi di Stato (parte 4.)*; Lezione 10: *Gli archivi di Stato (parte 5.)*, [6]: R. NAVARRINI Lezione 11: *Il concetto di archivio privato nella sua evoluzione storica*; Lezione 12: *Qualità ed estensione del contenuto degli archivi privati*. [7]: ID., Lezione 13: *Il regime giuridico degli archivi privati*; Lezione 14: *La tutela degli archivi privati nella legislazione italiana*. [8]: ID., Lezione 15: *L'ordinamento e l'inventariazione degli archivi privati*; Lezione 16: *Archivi nobiliari e domestici: la conservazione della memoria nell'azienda famiglia*. [9]: ID., Lezione 17: *L'archivio delle persone fisiche*; Lezione 18: *Dall'archivio di famiglia all'archivio d'impresa*. [10]: ID., Lezione 19: *L'archivio d'impresa: caratteri, organizzazione e conservazione*; Lezione 20: *Riflessione sugli archivi d'impresa e sulla loro tutela*.
- MORRISSEY 2002 = FRANCES MORRISSEY, *Introduction to a Semiotic of Scientific Meaning, and its Implications for Access to Scientific Works on the Web*, «Cataloging & Classification Quarterly», 33, 2002, n° 3-4, p. 67-97.
- OTTINO 1935 = *Bibliografia*: rifacimento e ampliamento del *Manuale di bibliografia* di Giuseppe Ottino, 4. ed. riveduta e arricchita da G. Fumagalli e Olga Pinto, con 165 figure, Milano, Hoepli, 1935.
- PEARCE-MOSES 2005 = RICHARD PEARCE-MOSES, *A Glossary of Archival and Records Terminology*, Chicago, Society of American Archivists, 2005, con vasta bibliografia, p. [415]-433.
- PLATONE, *Parm.* = PLATONE, *Parmenide*, 131.
- PONCHIROLI 2004 = PATRIZIA PONCHIROLI, *Progettazione UML nell'ambito del Semantic Web*, tesi di laurea, Università di Modena e Reggio Emilia, Facoltà di Ingegneria-Sede di Modena, Corso di laurea in Ingegneria informatica (rel.: Sonia Bergamaschi), a. a. 2003-2004; disponibile all'indirizzo di Rete: <<http://www.db->

- group.unimo.it/tesi/ponchiroli.pdf>, cons. 16 gennaio 2008.
- QUINE 1960a = ID., *Ontological Relativity*, in *Ontological Relativity and Other Essays*, New York, Columbia University Press, 1969, p. 26-68.
- QUINE 1960b = WILLARD VAN ORMAN QUINE, *Word and Object*, Cambridge, Mass., MIT Press, 1960.
- REVELLI 2007 = CARLO REVELLI, *Bibliografia di periodici di biblioteconomia*, in *Biblioteconomia 2007*, p. 143-146. Un rapido sguardo sulle circa 2.000 testate che controllano l'area disciplinare.
- ROMITI 2007a = ANTONIO ROMITI, *Archivistica*, in *Biblioteconomia 2007*, p. 179-186.
- ROMITI 2007b = IDEM, *Tra archivi e biblioteche: i "fondi archivistici" e i "fondi misti"*, in *Una mente colorata* cit., p. 735-748.
- ROSSI 2007 = MARELISA ROSSI, *Conservare libri e raccolte, oggi*, in CAVALLARO 2007, p. xiii-lii (in versione più semplice e in sede manualistica col titolo *La valorizzazione delle raccolte*, in *Biblioteconomia: principi e questioni*, a c. di Giovanni Solimine e Paul G. Weston, Roma, Carocci, 2007, p. 167-184).
- SALARELLI 2007 = ALBERTO SALARELLI, *Biblioteconomia*, in *Biblioteconomia 2007*, p. 147-162.
- SARDO 2007 = LUCIA SARDO, *Bibliografie nazionali*, in *Biblioteconomia 2007*, p. 135-140.
- SCENINI 2008 = FRANCESCA SCENINI, *Luther Blissett/Wu Ming autori collettivi per un modello di editoria copyleft*, «Culture del testo e del documento», 9., 2008, n° 25, gennaio-aprile, p. 71-86.
- SERRA 2004 = ROBERTO SERRA, *Tecniche basate su paradigma ad oggetti per l'analisi di sistemi complessi*, Laurea in Ingegneria Informatica, Università degli studi di Napoli Federico 2., a. a. 2003-2004.
- SIMAK = CLIFFORD DONALD SIMAK, *Fattore limitativo*, cit. in INNOCENTI-CAVALLARO 2003, p. 750-751.
- SMIRAGLIA 1992 = RICHARD P. SMIRAGLIA, *Authority Control and the Extent of Derivative Bibliographic Relationships*. Ph. D. Dissertation, University of Chicago, 1992.
- SMIRAGLIA 2001 = ID., *The Nature of a Work: Implications for the Organization of Knowledge*, Lanham (Md.), Scarecrow, 2001.
- SMIRAGLIA 2002a = ID., *Further Progress in Theory in Knowledge Organization*, in «Canadian Journal of Information and Library Science», 26., 2002, n° 2/3, p. 30-49.

- SMIRAGLIA 2002b = ID., *Works as Signs, Symbols, and Canons: the Epistemology of the Work*, in «Knowledge Organization», 28., 2002, p. 192-202.
- SMIRAGLIA 2004 = ID., *Knowledge Sharing and Content Genealogy: Extending the "Works" Model as a Metaphor for Non-Documentary Artifacts with Case Studies of Etruscan Artifacts*, in *Knowledge Organization and the Global Information Society*; Proceedings of the Eighth International ISKO Conference 13-16 July London UK, ed. Ian C. McIlwaine, «Advances in Knowledge Organization», v. 9., Würzburg, Ergon Verlag, 2004, p. 309-314.
- SMIRAGLIA 2005a = ID., *Content Metadata – an Analysis of Etruscan Artifacts in a Museum of Archeology*, in «Cataloging & Classification Quarterly», 40., 2005, n° 3/4, p. 135-51.
- SMIRAGLIA 2005b = ID., *Instantiation: Toward a Theory*, in *Data, Information, and Knowledge in a Networked World*: Proceedings of the Canadian Association for Information Science Annual Conference June 2-4 2005, ed. Liwen Vaughan: <<http://www.cais-acsi.ca/search.asp?year=2005>> (cons. dicembre 2007).
- SMIRAGLIA 2006 = ID., *Empiricism as the Basis for Metadata Categorization: Expanding the Case for Instantiation with Archival Documents*, in *Knowledge Organization and the Global Learning Society*; Proceedings of the 9th ISKO International Conference, Vienna, July 4-7 2006, G. Budin, C. Swertz, K. Mitgutsch editors, p. 383-388.
- SMIRAGLIA 2007a = ID., *Performance Works: Continuing to Comprehend Instantiation*, in *Proceedings of the North American Symposium on Knowledge Organization*, June 14-15, 2007, Toronto, Canada, ed. Joseph Tennis, disponibile all'indirizzo: <http://dlist.sir.arizona.edu/view/conference/North_American_Symposium_on_Knowledge_Organization_2007.html>.
- SMIRAGLIA 2007b = ID., *The 'Works' Phenomenon and Best Selling Books*, «Cataloging & classification quarterly», 44 n° 3/4, p. 179-95.
- SMIRAGLIA 2008 = ID., *A Meta-Analysis of Instantiation as a Phenomenon of Information Objects*, «Culture del testo e del documento», 9., 2008, n° 25, gennaio-aprile, p. 5-25.
- SOLIMINE 2007 = ID., *Materiale minore e documentazione in rete questioni non effimere di Biblioteconomia*, in *Una mente colorata cit.*, p. 539-545.

VADALÀ 2007 = MARIA ENRICA VADALÀ, *Bibliografia analitica (Bibliografia descrittiva)*, in *Biblioteconomia 2007*, p. 91-97.

VELLUCCI 1997 = SHERRY L. VELLUCCI, *Bibliographic Relationships in Music Catalogs*, Lanham (Md.), Scarecrow Press, 1997.

Wikipedia registra Archivistica, Bibliografia, Biblioteconomia, non Documentazione (ricerche del 31 dicembre 2007):

<<http://it.wikipedia.org/wiki/Archivistica>>.

<http://it.wikipedia.org/wiki/Bibliografia#_note-0>.

<<http://it.wikipedia.org/wiki/Biblioteconomia>>.

Il Documento

GIUSEPPE SCANDURRA¹

1. Definizione

1.1 Sul significato e sul valore del termine *documento*, occorre riferirsi, in ambito civilistico, alla dottrina, dato che il Codice civile non offre alcuna definizione idonea a fissare il complesso degli elementi che ne caratterizzino sul piano concettuale una nozione estensibile all'intero ordine dei documenti.

Nell'ambito del diritto amministrativo, merita un richiamo alla memoria l'art. 1 del Testo unico delle disposizioni legislative e regolamentari in materia di documentazione amministrativa, approvato con d.P.R. 28 dicembre 2000, n. 445, in cui viene precisata (sia pure ai soli fini dello stesso Testo unico) una definizione normativa del documento amministrativo, del documento informatico e di altri documenti a questi affini².

Fra tutte le definizioni del documento, sembra preferibile quella espressa in modo molto ampio e che riassume in sé i risultati di un procedimento logico discorsivo fondato sull'esame delle varie norme civilistiche³.

¹ Procuratore Generale Militare Emerito presso la Corte Suprema di Cassazione

² V. *infra* n. 1.3.

³ Il termine *documento* è presente in diverse disposizioni del Codice civile, come, ad esempio, negli artt. 736, 1262, 1477, 2235, 2961. Esso è adope-

Secondo tale opinione prevalente, il documento è sempre una cosa formata dall'uomo e rappresentativa di un fatto. Come tale, esso è una cosa corporale, semplice o composta, idonea a fissare, conservare o trasmettere la rappresentazione descrittiva o emblematica o fonetica di un dato fatto idoneo ad avere un'influenza nel mondo giuridico⁴.

Non è riconducibile a questa definizione il riferimento ad una funzione probatoria, che pur costituisce un connotato ricorrente del documento.

Non si può, infatti, ignorare che il documento può essere stato originato da sollecitazioni diverse dalle funzioni di prova, come risulta evidente dai documenti, cosiddetti. costitutivi, nei quali la funzione probatoria è certamente presente, pur non avendone costituito la ragione principale.

Occorre, comunque, chiarire che la documentazione va distinta dal pensiero documentato e che qualunque sia la struttura dell'atto documentato è sempre possibile giungere ad una sua documentazione, che, appunto per questa sua applicabilità ad un intero ordine di fatti, ha una dimensione universale. Poiché ogni atto deve necessariamente manifestarsi in una qualsiasi forma perché abbia esistenza materiale e giuridica rilevanza, si può avere un'esternazione orale, un'esternazione ge-

rato anche in alcuni articoli del Codice di procedura civile, come, ad esempio, negli artt. 165, 166, 184, 372. Nel capo del Codice civile riferito alla prova documentale, sono pur contenute alcune norme, concernenti altri strumenti di fissazione o di riproduzioni di fatti. In altre disposizioni, per indicare un documento, si usano anche altri termini, come atto, scrittura, o scritto, carta, incorrendo, così, in difetto di unitarietà e di esattezza terminologica. Nei casi in cui viene indicato esplicitamente il termine documento, esso è riferito ai documenti scritti, cioè a quelli redatti con l'impiego di un foglio di carta, su cui vengono tracciati i segni convenzionali di una lingua, relativi ai suoni che formano le parole.

⁴ Cfr. AURELIO CANDIAN, in *Documentazione e documento*, *Enciclopedia del diritto*, Milano Giuffrè, vol. XIII, p. 579.

stuale ed un'esternazione scritta. Nel caso dello scritto, l'esternazione assume la forma del documento, in cui possono concludersi anche le altre due forme di esternazione, quando la legge o la volontà dei privati vogliono ulteriormente realizzare, mediante scrittura o mediante registrazione fonografica o audiovisiva, una forma documentale, sia essa *ad probationem* ovvero *ad substantiam*.

1.2 Nella nozione di documenti rientra in modo comprensivo, oltre agli atti in senso stretto contenenti atti giuridici, ogni sorta di materiale documentale di qualsiasi provenienza, anche se parziale, frammentario e non identificato, sia che contenga scritti, sia che riproduca altre raffigurazioni, come disegni, fotografie o altri segni, destinati ai fatti in esso rappresentati.

Vanno, quindi, compresi sia i documenti che tali sono sin dalla loro origine, sia gli altri oggetti destinati ad altri fini, ma che, purtuttavia, comprendono elementi idonei a suffragare indicazioni utili a formulare giudizi sull'esistenza di un fatto giuridicamente rilevante.

Nella pratica comune si sono sempre più imposte, accanto alle rappresentazioni lessicali del pensiero, anche le rappresentazioni ideografiche, aventi il particolare vantaggio rispetto alle scritturazioni alfabetiche di poter trasmettere il pensiero senza doversi servire della parola, senza, cioè, dover tradurre le proprie idee in vocaboli e far ritradurre questi dal lettore.

Un'evoluzione altrettanto notevole è quella relativa al mezzo di rappresentazione che, in origine, era limitato alla scrittura e che si è, poi, sempre più trasformato, mediante i nuovi mezzi continuamente creati dalla moderna tecnologia. Dapprima, si è avuta, infatti, la scrittura a tatto, con la creazione dell'alfabeto per i non vedenti, e successivamente si è aggiunta la registrazione e la riproduzione dei suoni e dell'immagine. Anche queste "riproduzioni meccaniche" devono ritenersi comprese nella catego-

ria dei documenti, purché esse esprimano un pensiero e ne sia riconoscibile l'autore⁵.

1.3 Fra le riproduzioni indicate, con elencazione meramente esemplificativa, dall'art. 2712 C.c.⁶ rientra anche la riproduzione di un atto mediante il servizio telefax, nonché tutti quegli altri atti indicati nel citato art. 1 del Testo unico 445/2000.

Fra tali atti, è opportuno ricordare, soprattutto per sottolinearne gli elementi definitori, il *documento amministrativo* («ogni rappresentazione, comunque formata, del contenuto di atti, anche interni, delle pubbliche amministrazioni o, comunque, utilizzati ai fini dell'attività amministrativa»); il *documento informatico* («la rappresentazione informatica di atti, fatti o dati giuridicamente rilevanti»); il *documento di riconoscimento* («ogni documento munito di fotografia del titolare e rilasciato su supporto cartaceo, magnetico o informatico, da una pubblica amministrazione italiana o di altri Stati, che consenta l'identificazione personale del titolare»); il *documento d'identità elettronico* («il documento analogo alla carta d'identità elettronica rilasciato dal comune fino al compimento del quindicesimo anno di età»); la *dichiarazione sostitutiva di atto di notorietà* («il documento sottoscritto dall'interessato, concernente stati, qualità personali e fatti, che siano a diretta conoscenza di questi»); i *certificati elettronici* («gli attestati elettronici che collegano i dati utilizzati per verificare le firme elettroniche ai titolari e confermano l'identità dei titolari stessi»).

Il documento informatico ha l'efficacia probatoria prevista

⁵ All'obiezione che la registrazione potrebbe essere esclusa dal novero dei documenti, perché falsificabile da un abile imitatore, si può obiettare che anche la sottoscrizione di un documento può essere contraffatto da un falsificatore.

⁶ L'art. 2712 si riferisce a riproduzioni di fatti e cose e non a copie di scritture, per le quali, invece, dispone l'art. 2719 C.c.

dall'art. 2712 del Codice civile, riguardo ai fatti ed alle cose rappresentate. Sottoscritto con firma elettronica, soddisfa il requisito legale della forma scritta.

1.4 In ogni modo, va precisato che il documento ha sempre una sua rilevanza giuridica, che può sussistere sin dal momento stesso in cui esso si forma originariamente o può anche sopravvenire successivamente nel corso del tempo.

In quest'ultimo caso, la scrittura non è forma dell'atto e tanto meno lo è la registrazione; si ha, però, una documentazione che si aggiunge all'esternazione e che risponde soprattutto ad esigenze di tecnica organizzativa, in quanto destinata alla conservazione dell'atto.

Il documento rappresenta, cioè, la veste esteriore mediante la quale l'atto viene materialmente ad esistere, mentre ha solo una funzione di conservazione e di rappresentazione, quando riguardi un atto che ha avuto già una propria esternazione⁷.

2. Il documento come *res* e i diritti sul documento.

2.1 Il documento costituisce sempre una cosa mobile, anche se non si può escludere che in alcune particolarità esso possa essere considerato alla stregua di un bene immobile, come, ad esempio, nel caso di un documento lapideo inserito nella struttura di un edificio.

La rappresentazione del pensiero raccolta nel documento deve, comunque, avere caratteristiche tali da poter essere trasmessa nel tempo e da poter avere una certa continuità per una validità duratura⁸.

⁷ In questi sensi, ANIELLO NAPPI, *Documentazione degli atti processuali*, in *Digesto delle discipline penali*, Torino, Utet, vol. IV, par. 2, p. 164.

⁸ Non si richiede che la trasformazione del pensiero venga immortalata, a

Per la sua concreta esistenza, il documento richiede, di norma, una “base” per la redazione del pensiero e di un “mezzo di redazione”⁹. La base può essere costituita da qualsiasi materiale fisico in grado di recepire la stesura del pensiero e di poterlo tramandare ad altri soggetti contemporanei o anche alle generazioni successive. La carta, la pergamena, la pelle, le tavolette cerate, il tessuto, il vetro, il legno o i metalli sono ritenuti basi funzionalmente utili, come pure è tale qualunque altro oggetto idoneo a raccogliere la registrazione di un pensiero. Il mezzo di redazione può essere formato da segni tracciati sulla cosa dall'uomo o da apparati meccanici da questo predisposti e può essere costituito dal fuoco, dalla grafite, dall'inchiostro e da ogni altra sostanza colorante capace di fissarsi su un supporto.

La stessa sostanza, naturale o artificiale, colorante può essere ritenuta elemento strumentale labile, se viene impiegata per delineare alcune scritte nell'aria mediante aerei o sulla superficie dell'acqua per mezzo di rapide imbarcazioni, ma diviene mezzo idoneo, se viene, invece, impiegata su carta o su legno. Anche un mezzo a rapida scomparsa, come ad esempio il fuoco, può ben essere utilizzato per redigere un documento, se viene adoperato per fissare un'idea o un pensiero su pelli, legno o metallo.

Importante è sempre l'esito conclusivo e definitivo che si raggiunge attraverso l'adattamento del mezzo al supporto¹⁰. Esula, quindi, il carattere documentale in quelle rappresentazioni del pensiero, raccolte in modo labile e destinate a essere dissolte nel

perenne memoria, nel senso che essa debba avere una costante sopravvivenza, malgrado l'opera distruttrice o trasformatrice del tempo.

⁹ Sulla necessità di una “base” per l'incorporazione del pensiero e di un “mezzo di redazione”, v. ALESSANDRO MALINVERNI, *Documento (dir. pen.)*, in *Enciclopedia del diritto*, Milano, Giuffrè, vol. XIII, par. 8, p. 630.

¹⁰ Nota MALINVERNI, *op. cit.*, p. 631, che anche un testamento scritto con il sangue sulla parete di un carcere è sempre un documento.

nulla, come quelle, ad esempio, tracciate sulla neve o sulla sabbia. Oltre al tipo di carta speciale richiesto, la cambiale e il vaglia cambiario devono essere regolarmente bollati sin dall'origine.

2.2 Come tutte le cose, anche il documento può formare oggetto di diritti e, in quanto tale, esso è un bene, secondo la definizione data dal Codice civile (art. 810). Può essere oggetto del diritto di proprietà, nonché dei diritti reali frazionari di godimento o di pegno.

Il proprietario può rivendicarlo nei confronti di chiunque di fatto lo possenga o lo detenga ovvero abbia cessato di possederlo o di detenerlo per fatto proprio, dopo la domanda (art. 948 C.c.). Può anche formare oggetto di azione negatoria per far dichiarare l'inesistenza di diritti affermati da altri sul documento, quando si ha motivo di temerne pregiudizio (art. 949 C.c.). Può essere difeso con l'azione di reintegrazione o con quella di manutenzione, quando abbia formato oggetto di uno spoglio violento o occulto (art. 1168 C.c.) o di una molestia possessoria (art. 1170 C.c.). Può formare oggetto di provvedimenti cautelari, come, ad esempio, del sequestro giudiziario (art. 670 C.p.c.).

Qualora il documento appartenga ad una pluralità di soggetti, si applicano ad esso anche le regole relative alla comunione (art. 1100 e segg. C.c.). Se esso presenta interesse storico, artistico, archeologico o etnografico, vale la normativa risultante dal Codice dei beni culturali e del paesaggio di cui al D.Lgs. 22 gennaio 2004, n. 42 e s.m.i.

2.3 Il documento è un bene strumentale, in quanto esso serve ad esercitare o a fornire la prova delle posizioni giuridiche in esso rappresentate. In relazione a tale sua natura, si è sostenuto che il diritto avente ad oggetto il contenuto del documento prevale sul diritto avente ad oggetto la *res*, su cui il documento è redatto. Il titolare di quest'ultimo è tenuto, perciò, a non apportare

modifiche al documento né ad alterarlo o distruggerlo.

Il titolare dei diritti rappresentati nel documento può legittimamente far uso del contenuto del documento, ancorché non sia titolare del diritto di proprietà o di altro diritto reale sull'elemento materiale. Al riguardo, si deve, però, precisare che il diritto sul contenuto del documento non può essere qualificato come diritto di proprietà o altro diritto reale, potendo tale diritto configurarsi solo con riferimento all'elemento materiale. Esso deve essere ritenuto come diritto di fruttuosa utilizzazione dell'atto in esso contenuto e rappresentato¹¹.

Nell'ipotesi che ricorra un conflitto tra i diversi titolari, come sul caso dell'utilizzazione della materia altrui per la formazione del documento, il conflitto è risolto alla stregua dell'art. 940 C.c., in base al quale, se taluno ha adoperato una materia che non gli apparteneva per formare una nuova cosa, ne acquista per specificazione la proprietà, pagando il prezzo della materia.

3. Distruzione, soppressione, occultamento, danneggiamento o smarrimento del documento.

3.1 In quanto *res*, il documento può andare distrutto, soppresso o danneggiato, con effetti fra loro diversi, in rapporto alla natura del documento ed a seconda se questo sia un documento pubblico o privato, un originale ovvero una copia, un documento costitutivo o semplicemente probatorio.

Si ha la *distruzione* del documento, quando esso viene materialmente eliminato dal mondo fisico, in modo da renderlo definitivamente inutilizzabile o da cancellarne perfino la traccia. È indifferente il mezzo adoperato: combustione, sommersione, lacerazione in frammenti incomponibili, ecc.

¹¹ Su tale tematica, v. LUIGI CARRARO, *Il diritto sul documento*, Padova, Cedam, 1941, p. 39.

Ai fini della *soppressione* non si richiede, invece, la materiale distruzione dell'atto, ma è sufficiente la semplice eliminazione del contenuto del documento dalla sfera giuridica in cui avrebbe potuto operare, cancellandolo o sovrapponendo ai segni grafici altre scritture contorte o illeggibili, sgorbi, ecc., tali da rendere il contenuto incomprensibile. Od anche una immutazione del vero, nel senso che il documento risulti falsamente inesistente.

Rientra nel concetto di *occultamento* ogni altra condotta diretta ad impedire che l'atto vero adempia alla funzione per cui è predisposto. È sufficiente a tale effetto che il documento venga nascosto o portato in luogo ignoto od anche che l'avente diritto venga privato anche temporaneamente della disponibilità del documento, così da mettere in pericolo la sua potenziale utilizzabilità¹².

Mentre la soppressione e l'occultamento sono generalmente volontari, la distruzione o il danneggiamento possono essere volontari o assolutamente fortuiti ed estranei a qualsiasi determinata intenzione dell'autore. Possono essere riferibili al soggetto o ai soggetti interessati o solo ad alcuni tra questi. In quest'ultimo caso, la distruzione del documento non comporta necessariamente l'annullamento della dichiarazione o l'estinzione del rapporto in esso racchiusa. Nessuna incidenza e nessun effetto diretto o concomitante può, del resto, avere la distruzione del documento provocata dalle stesse parti, quando il rapporto giuridico fra loro instaurato ha ricevuto piena esecuzione.

3.2 Comunque, va osservato che l'atto materiale della distruzione non fa venir meno la dichiarazione rappresentata nel docu-

¹² Secondo alcuni, la distruzione e l'occultamento sono due *species* della soppressione che è un *genus*. Non si può, però ignorare che la distruzione consente di scorgere qualche differenza fra soppressione e distruzione e tanto più tra distruzione ed occultamento.

mento, ma solo l'essenza materiale del documento. Ne consegue che il titolare del diritto sul contenuto del documento stesso ha il diritto di ricostruirlo. In tal caso, si produce sia una ricostruzione del contenuto del documento sia una ricostruzione materiale, consistente nella formazione di un nuovo documento avente contenuto identico a quello distrutto¹³.

Aspetti diversi si hanno nel caso in cui la distruzione del documento, provocata volontariamente dall'autore della dichiarazione, possa valere come revoca tacita della dichiarazione. Così è, ad esempio, per l'ipotesi in cui un testamento olografo venga distrutto, lacerato, cancellato in tutto o in parte. In tal caso, si ha una presunzione di revoca volontaria del testamento, da equiparare, *quoad effectum*, alla revoca espressa, come si desume dal tenore dell'art. 684 C.c. che, in tema di scheda testamentaria, privilegia l'ipotesi di una provata carenza della volontà di revoca, superabile solo con la prova che il testamento «fu distrutto, lacerato o cancellato da persona diversa dal testatore, ovvero si provi che il testatore non ebbe l'intenzione di revocarlo».

3.3 Nell'ipotesi dello smarrimento involontario o dell'unilaterale distruzione pur essa involontaria, ma comunque incolpevole, la parte interessata potrà sempre fornire la prova per testimoni di quanto risulta nel documento. Per gli atti rispetto ai quali è richiesta la forma scritta (sia *ad substantiam* che *ad probationem*) la legge, infatti, ammette in ogni caso la prova per testimoni, «quando il contraente ha senza sua colpa perduto il documento che gli forniva la prova» (art. 2724, n. 3, C.c.).

Si ha perdita incolpevole del documento, anche nel caso in cui il documento sia stato smarrito da un terzo pienamente affidabile, cui sia stato consegnato per apprezzabili ragioni dal soggetto

¹³ SALVATORE PATTI, voce *Documento*, in *Digesto delle discipline privatistiche*, Torino, Utet, par. 9, p. 7, che richiama, a sua volta, LUIGI CARRARO, *Il diritto sul documento*, Padova, Cedam, 1941, p. 130.

interessato; in tale ipotesi, la mancanza di colpa va valutata in relazione al soggetto contraente che invoca a suo favore il contenuto del documento e non in riferimento al comportamento del terzo che lo ha smarrito.

4. I requisiti del documento

4.1 Per le varie tipologie di documento, la legge prevede diversi e distinti requisiti. In linea generale, l'autografia del documento privato non è richiesta dal diritto positivo, tranne alcune ipotesi specifiche, come, ad esempio, per il testamento olografo, per il quale si richiede che esso «deve essere scritto, datato e sottoscritto di mano del testatore» (art. 602 C.c.). Di norma, quindi, il documento è eterografo: necessaria è solo la sottoscrizione delle parti.

Con la sottoscrizione dell'autore, consistente nell'apposizione autografa dei segni grafici idonei ad individuare l'autore stesso, si realizza il collegamento giuridico tra soggetto e scrittura. Essa, di solito, viene espressa con l'indicazione del nome e del cognome¹⁴. Per la validità del documento è, comunque, richiesta qualunque altra sottoscrizione, purché idonea ad indicare con ragionevole certezza l'autore del documento.

Pienamente valide devono considerarsi, a tale effetto, ai sensi dell'art. 1 del D.P.R. n. 445/2000: sia *la firma elettronica*, e cioè «l'insieme dei dati in forma elettronica, allegati oppure connessi tramite associazione logica ad altri dati elettronici, utilizzati come metodo di autenticazione informatica»; sia *la firma elettronica avanzata*, costituita dalla «firma elettronica ottenuta attraverso

¹⁴ La sottoscrizione viene definita da ENRICO MARMOCCHI, *Scrittura privata*, in *Enciclopedia Giuridica*, Istituto della Enciclopedia italiana, vol. XXVIII, par. 5.1, p. 3 come «la scrittura autografa del proprio nome che una persona appone in fine del testo della scrittura».

so una procedura informatica che garantisce la connessione univoca al firmatario e la sua univoca identificazione»; e sia *la firma elettronica qualificata*, e cioè «la firma elettronica avanzata che sia basata su un certificato qualificato e creata mediante un dispositivo sicuro per la creazione della firma».

Alla sottoscrizione collocata in calce ad un documento, la legge attribuisce l'effetto giuridico di collegare il documento, e quindi la relativa dichiarazione, al soggetto indicato nella sottoscrizione. Un atto che contenga la semplice indicazione dell'autore, cui non si accompagni l'indicazione di un ulteriore contenuto di pensiero, non costituisce un documento. Come pure la sola rappresentazione di un pensiero, pur permeato di elaborati contenuti mentali, ma privo di ogni indicazione riferibile all'autore, non integra alcun documento. Non rappresenta parimenti un documento, il contenuto di uno scritto che non consenta l'identificazione del suo autore¹⁵. Al contrario, l'indicazione di colui che ha redatto un documento mediante un suo pseudonimo o un soprannome o altri segni identificativi, ricavati dal documento stesso o da elementi esplicitamente richiamati da esso non escludono la natura del documento.

Il semplice crocesegno apposto in fine del documento da parte della persona che non sa o non può comporre il segno del nome, non può essere ritenuto valida sottoscrizione, a meno che non sia eccezionalmente ritenuto dalla legge, in concomitanza con altri elementi formali, come segno avente natura equivalente alla sottoscrizione. Ovviamente, la parte è sempre libera di riconoscere come propria una sottoscrizione che giuridicamente non sarebbe tale.

La scrittura privata realizzata da due parti e sottoscritta solo da una di esse si deve considerare valida, quando è prodotta in

¹⁵ A tale conclusione deve pervenirsi anche per l'atto anonimo, da intendersi non già come quell'atto di cui si ignora l'autore, ma quello di cui non risulta l'autore, vale a dire l'atto non sottoscritto.

giudizio dalla controparte. Per l'atto pubblico, assume particolare rilievo la sottoscrizione dall'ufficiale rogante, la cui mancanza rende l'atto inesistente.

In tema di testamento olografo, l'art. 602 C.c. dispone che, anche quando la sottoscrizione difetti dell'indicazione del nome e del cognome dell'autore, la sottoscrizione «è tuttavia valida quando designa con certezza la persona del testatore». Deve, pertanto, ritenersi valida la manifestazione delle ultime volontà espresse in uno scritto avente forma di lettera, sottoscritta con l'indicazione del rapporto di parentela con i beneficiari delle disposizioni, quando tale indicazione comporti la certezza sull'identità della persona del testatore.

Non sempre la data viene ad assumere per il documento un rilievo decisivo. Di norma, essa costituisce un elemento esterno all'atto ed alla sua mancanza si può normalmente ovviare, a meno che non si richieda una data certa. Non va, comunque, trascurato che, in alcune ipotesi, la data del documento rappresenta un elemento essenziale, come, ad esempio, avviene nei casi di sopravvenuta incapacità di uno dei soggetti, o di sopravvenienza di una nuova legge o della anteposizione di un soggetto rispetto ad un altro, traente causa dallo stesso autore e per la medesima *res*.

Nel concetto di data è compreso, in senso ampio, anche l'indicazione del luogo in cui il documento è stato sottoscritto¹⁶.

5. Documento autogeno e documento eterogeneo; autografo ed eterografo

5.1 Considerato nella sua entità strutturale, il documento consta sia dell'elemento materiale su cui, o con il quale, la rappresentazione dell'entità giuridica viene realizzata, sia del contenuto di quest'ultima. Ovviamente, per dar luogo ad un docu-

¹⁶ CANDIAN, *Documentazione*, cit., p. 581.

mento, entrambi tali elementi devono risultare espressi in termini sufficientemente chiari e comprensibili.

La distinzione fra documento autogeno e documento eterogeneo si fonda sull'essere, oppure no, l'autore del documento lo stesso soggetto cui è riconducibile la paternità del pensiero documentato. Ai fini di questa distinzione, non occorre che l'artefice del documento sia anche l'autore materiale, ma è sufficiente che ne sia l'autore ideale. In altre parole, l'autore del documento può averlo redatto materialmente oppure può averne anche affidata la redazione, sia pure per interposta persona, a persona diversa. Comunque, esso deve essere consapevole di essere effettivamente l'autore del pensiero documentato¹⁷.

Nell'ipotesi che una persona si assuma la paternità di un documento ed un'altra la paternità della redazione, appare argomento risolutivo per la determinazione dell'autore non tanto la redazione del documento, quanto la volontà di dare origine al contenuto del pensiero documentato.

Bisogna, inoltre, tener presente che il documento, quando rappresenta l'unica forma di esternazione dell'atto, non è altro che la veste esteriore attraverso cui l'atto determina la sua stessa esistenza. Assume, invece, una semplice funzione di rappresentazione e di conservazione, quando esso costituisce un elemento aggiuntivo rispetto ad un'esternazione già avvenuta, fondata solitamente sulla voce o sui gesti.

La trasposizione dell'atto in un documento, pur costituendo sempre un'operazione materialmente distinta dalla esternazione dell'atto, può ricevere dalla legge un trattamento diverso, a seconda se venga privilegiata come elemento essenziale dell'esternazione o come operazione distinta dalla formazione dell'atto¹⁸.

¹⁷ Secondo una teoria formale della redazione, la qualità di autore va, invece, riconosciuta solo al redattore del documento.

¹⁸ In dottrina A. NAPPI, Voce *Documentazione degli atti processuali*, in *Digesto delle discipline penalistiche*, Torino, Utet, Vol. IV, par. 2, p. 1649),

Occorre, però, concludere che, nella sua realtà essenziale, il documento viene creato solo per una testuale conservazione di tutte o di alcune dichiarazioni intervenute in ordine ad un determinato fatto, e, comunque, in vista dell'esigenza di giungere alla realizzazione di una certezza giuridica. Natura ben diversa ha, invece, l'interpretazione del documento, con cui si cerca di accertare la volontà degli interessati, sulla base delle varie clausole del documento e secondo gli scopi riconoscibili di esso.

6. Documento costitutivo e documento probatorio

6.1 In conformità alle esigenze sistematiche che il tema esige e in relazione alle funzioni che il documento viene a svolgere rispetto al rapporto giuridico in esso attestato, si impone, anzitutto, una distinzione fondata sulla relazione intercorrente tra la forma materiale del documento e il sotteso negozio giuridico: fra il documento probatorio ed il documento costitutivo. Nel primo, e cioè nel documento probatorio, la formulazione del documento assume essenzialmente alla semplice funzione probatoria del negozio giuridico in esso eretto. Non rivestendo alcuna funzione privilegiata, esso viene valutato alla stregua di qualsiasi altro mezzo di prova e non svolge un'influenza maggiore di quella data, ad esempio, alla prova testimoniale. La forma scritta è prevista solo per la prova dell'atto ed è richiesta solo *ad probationem*. Si deve, pertanto, considerare come documento probatorio quello creato esclusivamente per esigenze di certezza giuridica, a semplice scopo di prova. In tal caso, il documento si presenta in termini logi-

è stato notato che la distinzione materiale fra esternazione e documentazione è rivelata dal fatto che normalmente la documentazione è affidata ad un soggetto diverso da quello che compie l'atto documentato, rispetto al quale il primo svolge una funzione ausiliaria. Tuttavia, in alcuni casi la documentazione è affidata allo stesso autore dell'atto documentato.

ci e cronologici come formalità creata successivamente alla nascita dell'atto e serve solo per fotografare e cristallizzare in modo talora incontestabile un atto già formato¹⁹. Fra i documenti che si riconducono a questa tipologia, andrebbero annoverati soprattutto i processi verbali, che sono rappresentativi di tutte le attività svolte nella sede dell'organo in cui sono stati predisposti.

6.2 Nel documento costitutivo, invece, la fattispecie negoziale non può prescindere dalla documentazione, che si presenta come un elemento integrativo del negozio e viene ad assumere una funzione costitutiva nei riguardi del negozio in esso documentato. A differenza del documento probatorio, non vi è alcuna altra alternativa né per provare né per costituire il negozio giuridico, se non a mezzo del documento. L'atto materiale formativo, cioè, è richiesto *ad substantiam* e la stessa validità dell'atto giuridico è subordinata alla sua esistenza²⁰. In tale formalità costitutiva, l'atto giuridico, in tanto esiste ed in tanto è produttivo dei previsti effetti, in quanto viene esternato nel documento previsto dalla legge: esso nasce, in sostanza, in coincidenza con la formazione del documento. Seppure la funzione essenziale del documento costitutivo sia quella di dar luogo al negozio in esso documentato, essa, comunque, non esclude anche una funzionalità probatoria che non costituisce, però, la finalità essenziale per cui l'atto è stato creato. Giurisprudenza costante della Corte di Cassazione stabilisce che «quando la legge richiede l'atto scritto *ad substantiam*, non è sufficiente che la volontà risulti da un qualunque documento, ma è necessario che questo consacri la volontà di concludere il contratto e cioè che il documento medesi-

¹⁹ In questi sensi, GIULIANO AMATO, *Documentazione costituzionale*, in *Enciclopedia del diritto*, Giuffrè, vol. XIII, par. 2, lett. b., pag. 599.

²⁰ Per un'indicazione dei vari atti interessanti l'ordinamento costituzionale e rientranti fra i documenti aventi siffatte tipicità, v. AMATO, *op. cit.*, par. 3, pag. 600.

mo sia stato posto in essere dalle parti al fine specifico di manifestare per iscritto tale volontà»²¹.

Il documento costitutivo o essenziale è tale o per volontà della legge (ad esempio, art. 1350 C.c., in cui sono indicati gli atti che devono farsi per atto pubblico o per scrittura privata, sotto pena di nullità) o per volontà delle parti (art. 1352 C.c.), in cui è lasciata all'autonomia privata la possibilità di convenire per iscritto, per un futuro contratto, l'adozione di una determinata forma.

6.3 Accanto al documento costitutivo dobbiamo considerare anche il documento non costitutivo, ritenendo per tale quel documento che presenti imperfezioni o difetti rispetto al tipo voluto dalla legge (v. ad esempio, art. 2297 C.c.).

Non si può trascurare di considerare, infine, anche un terzo tipo di atto in cui l'esternazione documentale è indispensabile non già per l'esistenza dell'atto, ma per la progressiva produzione dei suoi effetti, a fini procedimentali ovvero a fini pubblicitari. Così è, ad esempio per la legge, che nasce e si manifesta in virtù di un atto orale, quale è la proclamazione della votazione da parte del presidente della seconda assemblea che l'ha approvata, mentre i successivi documenti di esternazione (rappresentati dal testo inserito nella raccolta delle leggi e da quello pubblicato nella G.U.) costituiscono formalità di manifestazione esterna o di comunicazione a fini di notorietà, per la concreta operatività delle norme in esso contenute.

7. Documenti pubblici e documenti privati.

7.1 Fra i documenti eterogenei, e cioè fra i documenti in cui l'autore della rappresentazione è persona diversa da quella che è

²¹ C. Cass., sentenze n. 2766 del 1960, n. 782 e 3216 del 1962.

interessata alla sua formazione, va compreso il documento pubblico, cui si adatta la definizione di atto pubblico offerta dall'art. 2699 del C.c.²². Secondo tale definizione, «l'atto pubblico è il documento redatto con le richieste formalità da un notaio o da altro pubblico ufficiale autorizzato ad attribuirvi pubblica fede nel luogo ove l'atto è formato». Sulla base di questa definizione normativa, si possono precisare le due caratteristiche essenziali dell'atto pubblico costituite dalla natura di documento e dalla sua provenienza da un pubblico ufficiale, esercente una pubblica funzione di certificazione.

Conformemente al dato normativo, l'atto pubblico si configura, quindi, come un documento, nel quale è racchiuso, mediante segni grafici e in forma descrittiva, l'attestazione di avvenimenti svoltisi in presenza e ad opera del pubblico ufficiale documentatore²³. Dalla definizione normativa dell'atto pubblico si deduce negativamente quella della scrittura privata, da qualificare come il documento scritto e sottoscritto da colui o da coloro che lo hanno redatto.

²² Con osservazione puramente formale è stata rilevata l'inesattezza terminologica del dettato normativo rappresentata dall'uso del termine "atto", anziché dell'altro "documento", posto che il termine "atto" potrebbe essere identificato nell'atto giuridico della parte, inteso nel suo significato rigorosamente tecnico, mentre meglio si prestava, nel momento definitorio dell'art. 2699 C.c., il riferimento alla scrittura pubblica o anche al documento pubblico. Quest'ultima definizione risulta, peraltro, confermata dal testo letterale dell'art. 2699 C.c. che definisce, appunto, l'atto pubblico come «il documento redatto». Sul punto, v. CANDIAN, *Documentazione*, cit., par. 14, p. 583, nonché FRANCESCO CARNELUTTI, *Documento (Teoria moderna)*, in *Novissimo Digesto italiano*, Utet, vol. VI, p. 85. Quest'ultimo autore riconduce l'inesattezza terminologica del legislatore alla circostanza che, nel caso degli atti pubblici, la legge cede all'uso di scambiare il contenuto con il contenente e, nel caso della scrittura privata, non suppone l'esistenza di documenti diversi dalle scritture.

²³ In questi sensi, GIORGIO CRISCI, *Atto pubblico* (dir. civ.) in *Enciclopedia del diritto*, Milano, Giuffrè, vol. IV, p. 265.

7.2 Caratteristica essenziale dell'atto pubblico è l'estrinseco connotato di attendibilità attribuito sia all'identità dei soggetti in esso indicati sia al contenuto risultante dal testo, la cui validità può essere eccepita solo con querela di falso (art. 2700 C.c.). L'atto pubblico rende, infatti, piena prova sia della provenienza di esso, sia degli atti compiuti dal pubblico ufficiale nonché di quelli avvenuti alla sua presenza. Non prova, però, la veridicità delle dichiarazioni delle parti, le quali, pertanto, possono essere contrastate ed accertate con tutti i mezzi di prova consentiti dalla legge, senza che, all'uopo, occorra, o possa proporsi, querela di falso. Questa non preclude, peraltro, l'indagine su circostanze o fatti che nello stesso atto non risultino né positivamente né negativamente acquisiti.

Qualora il contenuto di un documento pubblico sia contraddetto o smentito da altre risultanze processuali, non è sufficiente la semplice contrapposizione di queste ultime, anche se pur esse dotate dell'efficacia probatoria di una prova legale, per togliere all'atto pubblico l'efficacia che gli è propria, ma è necessario il previo esperimento della querela di falso. Quest'ultima costituisce, infatti, lo strumento imprescindibile per neutralizzare il valore probatorio del documento. Deve essere proposta sia in via principale sia in corso di causa in qualunque stato e grado di giudizio. Deve contenere a norma dell'art. 221, secondo comma, C.p.c., a pena di nullità, l'indicazione degli elementi e delle prove della falsità, deve essere proposta personalmente dalla parte oppure a mezzo di procuratore speciale, con atto di citazione e con dichiarazione da unirsi al verbale di udienza.

7.3 L'intervento del pubblico ufficiale nella redazione del documento pubblico non può in nessun caso essere posto a confronto con la partecipazione alla redazione dell'atto, esplicita dallo scrivano o dal dattilografo. Infatti, costituiscono atti pubblici soltanto gli atti che il notaio o l'altro pubblico ufficiale formano nell'esercizio di pubbliche funzioni certificative delle qua-

li siano stati investiti dalla legge. Esulano dalla previsione della norma dell'art. 2699 C.c. i documenti dei pubblici ufficiali che non siano espressioni di tali funzioni certificative. Il notaio o l'altro pubblico ufficiale interviene nella formazione dell'atto non già come mandatario delle parti²⁴, ma nell'esercizio di un ufficio pubblico; di un istituto, cioè, cui è estranea ogni dipendenza ovvero ogni soggezione alla volontà del titolare dell'interesse. La parte o le parti non hanno alcun potere di ingerenza sulle modalità di esercizio di quell'ufficio.

Gli altri pubblici ufficiali autorizzati ad attribuire pubblica fede al contenuto del documento possono identificarsi negli ufficiali dello stato civile, nei cancellieri, nei segretari comunali, negli ufficiali giudiziari, nei conservatori delle ipoteche, negli agenti delle imposte, nei magistrati e in tutti quegli altri delegati dall'ordinamento giuridico alla pubblica attestazione dei fatti da loro compiuti o delle dichiarazioni rese in loro presenza.

Perché l'atto pubblico possa considerarsi valido, occorre che l'ufficiale rogante sia competente per territorio e per materia e che siano state, inoltre, osservate le formalità prescritte dalla legge. Qualora nel documento pubblico manchi per qualsiasi motivo, volontario o meno, la sottoscrizione del documento da parte del pubblico ufficiale, viene a mancare un requisito essenziale dell'atto pubblico. Ai sensi dell'art. 2701 C.c., in tema di conversione dell'atto pubblico, non è, però, escluso che lo stesso atto abbia l'efficacia probatoria della scrittura privata, ove risulti che il consenso degli stipulanti, riferito al contenuto del documento, sia stato espresso compiutamente mediante la sottoscrizione²⁵.

²⁴ Nota CANDIAN, *Documentazione*, cit., paragrafo 14, p. 583, che il pubblico ufficiale è stato talora qualificato, quale mandatario della parte. L'improprietà di tale qualificazione risulta evidente dalla constatazione che mancherebbe in tal caso un mandante, a meno che questo non venga individuato nella legge.

²⁵ È stato notato in dottrina (PATTI, *Documento*, cit., parag. 11, p. 8) che l'in-

8. Scrittura privata e figure affini.

8.1 In mancanza di una definizione normativa, si è giunti attraverso la dottrina e la giurisprudenza a determinare in modo abbastanza preciso il concetto di scrittura privata²⁶, per fissarlo in «qualunque documento sottoscritto dalle parti, senza la partecipazione di un pubblico ufficiale nell'esercizio di una pubblica funzione certificativa»²⁷. La sua disciplina è compresa, fra gli artt. 2702 – 2708 del Codice civile, nel capo dedicato alla prova documentale, ed ha corrispondenza processuale negli artt. 214 e segg. del Codice di procedura civile.

Non sempre i documenti trattati nella sezione riservata alla scrittura privata presentano note distintive riconducibili alla definizione ora indicata. Bisogna, tuttavia, pur sempre prendere atto che tali documenti presentano caratteristiche sufficientemente comuni e che la legge ne equipara l'efficacia probatoria a quella propria della scrittura privata. Si deve, però, tener presente che la normativa sulla scrittura privata presenta, comunque, un'accentuata possibilità di adattamento, utile ad attrarre nella sfera della propria capacità regolatrice anche tutti quei nuovi documenti

dicata disposizione dell'art. 2701 C.c. presenta un aspetto normativo di grande rilievo, soprattutto per la sua rilevanza pratica, in quanto una delle parti non potrebbe mai far valere la nullità dell'atto per sottrarsi alle obbligazioni assunte, giacché l'atto l'avrebbe sempre come scrittura privata.

²⁶ Nota, infatti, MARMOCCI, *Scrittura privata*, cit., vol. XXVIII, par. 1.1, p. 1, che il codice civile, diversamente che per l'atto pubblico, si limita a disciplinare la scrittura privata, attribuendo ad essa il nome tecnico senza enunciarne la definizione, ricavabile dall'interprete mediante l'analisi del fenomeno e il confronto delle fattispecie.

²⁷ Per una definizione analoga, v. PATTI, *Documento*, cit., par. 12, p. 9. Secondo AURELIO SCARDACCIONE, *Scrittura privata*, in *Novissimo Digesto Italiano*, vol. XVI, par. 1, p. 810, scrittura privata è qualunque documento consistente in uno scritto redatto da un soggetto che non sia pubblico ufficiale e che sia sottoscritto.

che, di volta in volta, sorgono nella realtà quotidiana per effetto delle continue evoluzioni della tecnica e dei più progrediti mezzi di trasmissioni delle dichiarazioni delle parti.

8.2 Requisiti essenziali della scrittura privata sono la *res* su cui vengono tracciati i segni grafici della scrittura in forma di linguaggio, l'espressione scritta del pensiero costituente il vero e proprio contenuto giuridicamente rilevante del documento e la sottoscrizione²⁸. L'ultimo di questi elementi costituisce il punto nodale nel quale si accentra il complesso di tutti i problemi sorti intorno alla scrittura privata. Da un lato, infatti, si è ravvisato nella sottoscrizione il dato di collegamento tra il documento ed i soggetti cui attribuire il documento²⁹, dall'altro, la sottoscrizione appare sempre più superata dall'evoluzione della tecnica, tanto da far sorgere la necessità di dover giungere all'individuazione dei mezzi che siano equivalenti alla sottoscrizione e siano al pari di questa, idonei ad assicurare lo stesso grado di certezza sull'attribuzione della paternità del documento³⁰.

8.3 Per quanto attiene all'efficacia probatoria della scrittura privata, risulta fissato nell'art. 2702 C.c. il principio in base al quale la scrittura privata, al pari del documento pubblico, fa pie-

²⁸ Come è stato ben specificato dalla giurisprudenza (C. cass., sent. n. 6133/1992), la sottoscrizione ha la funzione di individuare l'autore del documento, nonché quella di far assumere ad un soggetto la paternità dello scritto. Per raggiungere tale duplice finalità, essa deve sostanziarsi nell'apposizione di segni grafici idonei, tra i quali non può essere compreso il crocesegno, ancorché apposto in presenza di testimoni.

²⁹ La giurisprudenza della Cassazione ha, infatti, sempre affermato (tra le altre, sent. n. 9289/2001) che «elemento essenziale per la validità di una scrittura privata è la sottoscrizione della stessa da parte del suo autore, che non trova equipollente nella autografia del testo, né nel segno di croce apposto dal soggetto da cui il documento proviene».

³⁰ In questi sensi, PATTI, *Documento*, cit., par. 12, p. 10.

na prova, fino a querela di falso, della provenienza della dichiarazione da parte di chi l'ha sottoscritta, purché colui contro il quale essa è prodotta in giudizio riconosca la sottoscrizione ovvero questa sia legalmente considerata come riconosciuta, in virtù di un'autenticazione operata dal notaio o da altro pubblico ufficiale a ciò autorizzato. L'efficacia probatoria così attribuita alla scrittura privata riguarda, però, le sole scritture a carattere negoziale ed essa si riferisce al caso in cui il documento sia prodotto nei confronti del sottoscrittore ovvero di un suo erede o avente causa. La sottoscrizione vale, infatti, ai sensi dell'art. 2702 C.c., a ingenerare una presunzione *iuris tantum* di consenso del sottoscrittore al contenuto dell'atto e di assunzione della paternità dello scritto, indipendentemente dal fatto che la dichiarazione non sia stata vergata o redatta dal sottoscrittore. Ben può, infatti, la scrittura privata essere redatta da soggetto distinto dal sottoscrittore, od anche da più soggetti, essendo l'autografia richiesta solo per scritture aventi natura particolare.

La scrittura non sottoscritta, priva cioè di ogni indicazione nominativa del suo autore, non è soggetta alla disciplina fissata negli artt. 2702 C.c. e 214 e segg. C.p.c. Ne consegue che la parte contro cui è prodotta non ha l'onere del disconoscimento, né il suo comportamento può assumere il valore di un riconoscimento tacito³¹.

8.4 Per quanto riguarda la data in cui la scrittura è formata, essa non costituisce un elemento essenziale di validità³² del do-

³¹ Ciò varrebbe anche nell'ipotesi di scrittura autografa. L'assenza di ogni sottoscrizione in calce al documento dovrebbe essere intesa nel senso che la parte non ha inteso assumersi la paternità del documento. In questi sensi, BRUNETTO CARPINO, *Scrittura privata*, in *Enciclopedia del diritto*, Giuffrè, XLI, parag. 14, p. 805.

³² In dottrina (CARNELUTTI, *La prova civile*, Roma, 1974, p. 208) si è sostenuto che l'indicazione della data costituisce un requisito essenziale anche per la scrittura privata.

cumento, salvo alcuni casi espressamente previsti dalla legge (art. 602 C.c.). La sua funzione è «quella di indicare le circostanze di luogo, e soprattutto di tempo, in cui il documento è stato formulato e di fissare, in tal modo, gli effetti giuridici della dichiarazione documentata nella duplice dimensione dello spazio e del tempo»³³.

Sulla base di tale nozione, si deve preliminarmente distinguere tra data del documento e data dell'atto, in relazione alla circostanza che la dichiarazione della volontà espressa nel documento rappresenta un'entità autonoma rispetto al documento stesso che ne costituisce la cristallizzazione formale. Tale distinzione rivela tutta la sua portata nell'ipotesi in cui le parti, una volta concluso un negozio, si risolvano di operarne la redazione mediante un atto pubblico, in cui trasfondere una precedente convenzione verbale o una scrittura privata. Essa, peraltro, rileva tutta la sua particolare rilevanza riguardo agli effetti che possono scaturire dall'uno e dall'altro, anche in relazione agli eventuali vizi che ciascuno di essi possa presentare e che, di regola, non si comunicano dall'uno all'altro³⁴.

Nell'atto pubblico, la data assume un ulteriore elemento di rilievo, nella considerazione che la particolare fede privilegiata attribuita a codesto tipo di documento, comprende anche la data, nel senso che questa risulta accertata fino a querela di falso. Ovviamente, ciò non comporta alcun limite nella valutazione e distinzione della data dell'atto documentato, cui va riconosciuto, anche sotto questo aspetto, piena autonomia³⁵.

Qualora, però, il procedimento di documentazione assuma particolare rilievo, perché da esso derivano effetti di natura so-

³³ UGO NATOLI e ROMEO FERRUCCI, *Della tutela dei diritti*, in *Commentario del codice civile*, VI, t. I. Torino, sub art. 2704.

³⁴ Su questo punto, v. GIOVANNI GIACOBBE, *Data (in generale)*, in *Enciclopedia del diritto*, Milano, Giuffrè, Vol. XI, paragraf. 3, p. 693.

³⁵ Ivi, p. 694.

stanziale relativi alla opponibilità dell'atto ai terzi, il legislatore stesso detta le regole attinenti alla data, indicando i criteri ed i limiti perché questa possa essere rilevante rispetto ad essi. In particolare, la data della scrittura privata di cui non sia stata autenticata la sottoscrizione «non è certa e computabile riguardo ai terzi, se non dal giorno in cui la scrittura è stata registrata o dal giorno della morte o della sopravvenuta impossibilità fisica di colui o di uno di coloro che l'hanno sottoscritta o dal giorno in cui il contenuto della scrittura è riprodotto in atti pubblici o, infine, dal giorno in cui si verifica un altro fatto che stabilisca in modo egualmente certo l'antiorità della formazione del documento» (art. 2704, 1° comma, C.c.). Quanto alla qualifica di *terzo*, deve ritenersi tale qualunque soggetto che sia portatore di un interesse giuridicamente perfetto e differenziato dagli interessi delle parti del rapporto medesimo, in guisa da rendere possibile la configurazione di un conflitto anche solo potenziale con questi ultimi.

8.5.1 Figure affini alla scrittura privata sono alcuni documenti indicati espressamente nella stessa sezione del Codice civile dedicata alla scrittura privata. Trattasi, in particolare, del telegramma, delle carte e dei registri domestici, e delle annotazioni in calce, in margine o a tergo di un documento.

8.5.2 Come è noto, nel telegramma non vi è l'autografia del mittente né è da questi sottoscritto. Eppure, ad esso viene riconosciuta dal legislatore l'efficacia probatoria della scrittura privata, con parziale deroga alle statuizioni generali. Dispone, infatti, l'art. 2705 C.c. che il telegramma ha la stessa efficacia probatoria della scrittura privata, non solo quando l'originale consegnato all'ufficio postale di partenza è sottoscritto dal mittente, ma anche quando si provi che esso è stato consegnato o fatto consegnare dal mittente medesimo, anche senza sottoscriverlo. È stata in tal modo, concretizzata, nell'intento di favorire la rapi-

dità dell'incontro di volontà negoziali tra persone tra loro distanti un'equivalenza tra la sottoscrizione del documento e la sua materiale consegna all'ufficio di partenza. Con la conseguenza che solo in caso in cui ciò sia contestato, il mittente medesimo è tenuto, ove intenda valersene quale scrittura privata, a fornire la prova delle condizioni poste dall'art. 2705 C.c., mentre, ove nessuna contestazione vi sia stata circa la provenienza del telegramma, questo ha, a tutti gli effetti, il valore di scrittura privata, senza che il mittente sia tenuto a dare alcuna ulteriore prova³⁶. Con l'introduzione del sistema della dettatura (o dell'incarico di dettatura) del telegramma per telefono da parte del mittente all'ufficio telegrafico, si può dire realizzato, sia pure con la collaborazione di terzi, il recapito del proprio messaggio all'ufficio telegrafico, con la conseguenza di una prova della spedizione più agevole, in quanto l'ufficio spedisce per posta al mittente una copia del modulo redatto sotto dettatura³⁷. La fedeltà tra l'originale e il documento riproduttivo viene, in base all'art. 2706 C.c., presunta, sino a prova contraria.

8.5.3 Con l'art. 2707 C.c. si ammette, invece, per le carte ed i registri domestici, un'efficacia probatoria limitata sia sotto l'aspetto soggettivo sia sotto quello del contenuto.

Dispone, infatti, l'art. 2707 C.c. che «Le carte ed i registri domestici fanno prova contro chi li ha scritti: 1) quando enunciano espressamente un pagamento ricevuto; 2) quando contengono la menzione espressa che l'annotazione è stata fatta per supplire alla mancanza di titolo in favore di chi è indicato come creditore». Sotto l'aspetto soggettivo, si richiede, infatti, che il documento sia olografo. Per la limitazione di carattere oggettivo, l'art. 2707 C.c. richiede espressamente che quanto enunciato nel documen-

³⁶ In tali termini, C. Cass., sent. n. 12128 del 1992.

³⁷ In questi sensi, CANDIAN, *Documentazione*, cit., par. 13, p. 583.

to deve riferirsi ad un pagamento ricevuto o alla finalità di supplire alla mancanza di titolo.

8.5.4 Analoga considerazione va fatta per le annotazioni fatte dal creditore in calce, in margine o a tergo di un documento rimasto in suo possesso: fanno prova, anche se non sottoscritte dal creditore, quando tendono ad accertare la liberazione del debitore (art. 2708 C.c.).

Nel richiedere un'annotazione in calce, in margine o a tergo di un documento rimasto in possesso del creditore, la norma in esame ha, così, inteso attribuire valore di prova liberatoria ad una nota o ad un appunto che non faccia parte integrante del testo del documento e che non sia stato redatto nello stesso contesto di tempo, ma sia stato apposto dal creditore, od anche da un suo incaricato, su una parte periferica del documento in un secondo momento, dopo il suo completamento. Lo stesso valore ha l'annotazione fatta dal creditore in calce, in margine o a tergo di una quietanza o di un esemplare del documento del debito posseduto dal debitore.

9. Scrittura privata autentica e scrittura privata riconosciuta

Oltre alle particolarità che differenziano pur sempre la scrittura privata autenticata dall'atto pubblico³⁸ vanno, comunque, considerate alcune specificità che differenziano la scrittura privata autenticata dalla scrittura privata riconosciuta.

³⁸ Sulla individuazione di tali peculiarità, v., in particolare anche quanto risulta dalla legge sull'ordinamento del notariato e degli archivi notarili, nonché dal Codice penale (artt. 28, comma 1, n. 1; 46, lett. b); 47, comma 2, e 72 L. 16.2.1913, n. 89, come modificati dall'art. 12, comma 1, lett. b) ed e) L. 28.11.2005, n. 46; artt. 476 e 485 C.p.).

La nozione di autenticazione della sottoscrizione ci viene data dall'art. 2703, comma 2, C.c., mentre il concetto di scrittura privata riconosciuta si può desumere dall'art. 2702 C.c. Il concetto di quest'ultima, che si ha per riconosciuta ai fini processuali, è più ampio di quello della scrittura privata autenticata, in cui è rappresentata una delle ipotesi in cui la scrittura è considerata legalmente riconosciuta dalla legge. Ne consegue che non può costituire titolo esecutivo ex art. 474, comma 2, n. 2 C.p.c., la scrittura privata riconosciuta in giudizio da una delle parti o da questa non disconosciuta nei termini di cui all'art. 215, comma 1, n. 2, C.p.c. ovvero prodotta nei confronti di parte rimasta contumace. Né può considerarsi titolo esecutivo la scrittura privata sottoposta con esito positivo al giudizio di verifica di cui agli artt. 216 e segg. C.p.c.. L'efficacia probatoria prevista dall'art. 2702, C.c. non va, cioè, confusa con l'efficacia esecutiva di cui all'art. 474 C.p.c., la cui elencazione deve essere considerata tassativa³⁹.

In conclusione, le scritture private autenticate fruiscono della stessa disciplina prevista per le cambiali e gli altri titoli di credito, per i quali non è richiesta l'apposizione di alcuna formula esecutiva e per la cui azione è sufficiente la presentazione diretta dell'originale del titolo esecutivo all'ufficiale giudiziario cui spetta solamente esaminare se il precetto ne riporti fedelmente il contenuto (art. 480, comma secondo, secondo periodo, C.p.c.).

10. Documenti dichiarativi e documenti narrativi; documenti contestuali e documenti mutuamente complementari; documenti con data non vera e documenti ricognitivi

10.1 In relazione al contenuto del documento, si distingue fra

³⁹ In questi sensi, ALBERTO CARDINO, *La riforma della legge n. 80/2005 in tema di esecuzioni civili*, in *Incontro di studi indetto dal C.S.M.*, in Roma 10-11 aprile 2005.

documenti dichiarativi e documenti narrativi: i primi contengono dichiarazioni espresse dall'autore del documento, mentre i secondi contengono soltanto l'esposizione di un fatto. Nell'ambito, poi, dei documenti dichiarativi, si opera un'ulteriore distinzione fra documenti cosiddetti testimoniali, in cui la dichiarazione rappresentata costituisce una dichiarazione di scienza, e documenti dispositivi, in cui la dichiarazione esprime, invece, un atto di volontà. Qualora la dichiarazione di volontà riproduca una dichiarazione già effettuata, ricorre la figura del documento riproduttivo.

10.2 Un'ulteriore distinzione è quella operata fra documenti contestuali e documenti mutuamente o reciprocamente complementari, a seconda se le dichiarazioni delle parti interessate siano riportate in un unico documento o siano, invece, desumibili da due o più scritture, fra loro integrate. Pur quando le dichiarazioni delle parti debbano *ad substantiam* risultare da scrittura, non è necessaria la contestualità delle dichiarazioni, ma è sufficiente che esse risultino da scritture, ancorché plurime.

10.3 Per i documenti qualificati dalla legge come solenni, si richiede la sottoscrizione di tutte le parti che ne siano soggetti. Diffusa e accettata opinione ammette la validità di un documento sottoscritto da una sola parte e prodotto dalla parte non sottoscritta allo scopo di avvalersi delle attestazioni in esso contenute.

10.4 In merito alla data, v'è da rilevare che essa viene inserita nel documento, perché ne vengano desunti il luogo e il momento della sua formazione. Ovviamente, la data può essere vera oppure no. Nel caso difetti l'elemento della verità, si distingue se la difformità è fortuita oppure intenzionale. In questo secondo caso, si ha l'ipotesi della data simulata.

Tuttavia, parrebbe, secondo quanto riportato dal Candian⁴⁰, che la simulazione sussiste soltanto quando vi sia stata l'intenzione, e cioè un fermo proposito e la consapevolezza di ingannare i terzi e non anche nell'ipotesi in cui la volontà dei soggetti artefici sia stata quella di imporre semplicemente all'atto una indicazione temporale, anteriore o posteriore all'atto, senza alcuna volontà di indurre altri in errore. A questa distinzione si oppone altra dottrina, secondo cui si tratterebbe sempre di simulazione, ancorché distinguibile in simulazione legittima (quando non consegua alcuna lesione di diritti altrui) e in simulazione fraudolenta⁴¹.

Altre tipologie di documenti sono quelle che fanno capo ai c.d. documenti riproduttivi o ricognitivi. Trattasi di documenti redatti dalle parti e suggeriti dall'intento di riprodurre, per i più svariati motivi, un precedente documento o anche per esprimere in modo tecnicamente compiuto e corretto un ulteriore accordo o per trasformare o annullare una clausola, oppure anche per tradurre in termini razionali accessibili una dichiarazione di volontà già formulata o per conferire una maggiore efficienza ad un diritto di credito o per prolungare oltre il termine stabilito gli effetti di un contratto.

10.5 Ma oltre che con dichiarazioni di volontà, le parti possono riaffermare precedenti dichiarazioni iniziali, con altre dichiarazioni che ben potrebbero essere qualificate come dichiarazioni di verità. Senza voler trasformare o mutare lo stato giuridico precedente, esse hanno per scopo soltanto quello di constatare l'avvenuta attuazione di un determinato fatto, come, ad esempio, un avvenuto pagamento o la consegna di una determinata

⁴⁰ Aurelio CANDIAN, *op. cit.*, p. 582.

⁴¹ In questi sensi, Aurelio CANDIAN, *op. cit.*, p. 582.

merce ovvero di ribadire, a fini di convalida, la precedente conclusione di un contratto⁴².

Sulla natura giuridica di siffatte dichiarazioni, sono sorte in dottrina notevoli e complesse divergenze. Sembra accettabile la tesi che le riconduce nell'ambito delle dichiarazioni confessorie⁴³, alla stregua delle quali esse non possono più essere revocate, se non si prova che sono state determinate da errore di fatto o da violenza (art. 2732 C.c.). A tale riguardo, la Corte di Cassazione ha precisato (sent. n. 3010 del 2002) che, al fine della revoca della confessione, è necessario non solo l'elemento oggettivo, costituito dalla non rispondenza al vero del fatto confessato, ma anche l'elemento soggettivo, cioè la prova dello stato di errore in cui il confidente si trovava nel momento in cui il fatto veniva confessato. È indubbio, comunque, che il documento ricognitivo è inidoneo a svolgere le funzioni di un documento costitutivo nei casi in cui questo sia richiesto dalla legge.

11. Titoli esecutivi

11.1 Come è noto, il titolo esecutivo è il documento che nel processo civile consente di promuovere l'esecuzione forzata. Oltre ai titoli esecutivi che normalmente si formano nell'ambito di

⁴² CANDIAN, *Documentazione*, cit., par. 27, p. 590, cita, come casi frequenti di documenti ricognitivi, quelli in cui due persone, una volta perfezionato il contratto, tornino a trovarsi insieme e traducano il negozio (che dopo un sufficiente scambio epistolare si era già perfezionato) in una scrittura contestuale, privata o notarile; ovvero quei fatti specifici in cui venga a concretarsi, ad esempio, un contratto già stipulato per scrittura privata e riprodotto per atto autentico; o anche la dichiarazione di benessere al saldo di conto corrente, o la dichiarazione di benessere al risultato dei conti periodicamente trasmessi dal banchiere al beneficiario di apertura di credito.

⁴³ In tal senso, CANDIAN, *ivi*, par. 27, p. 591.

un processo o comunque all'esito di una attività giurisdizionale (c.d. titoli giudiziali), il nostro ordinamento riconosce altri titoli esecutivi che vengono a formarsi fuori del processo (c.d. titoli stragiudiziali) ed a cui è riconosciuta dalla legge, in vista di una facilitazione alla circolazione creditizia, un'efficacia esecutiva corrispondente a quella dei provvedimenti giurisdizionali. Trattasi, in particolare, delle cambiali, nonché degli altri titoli di credito e degli atti ai quali la legge attribuisce espressamente tale efficacia (artt. 1684, 1790, 1791 C.c.). La cambiale, il vaglia cambiario e l'assegno bancario non hanno, però, la qualità di titoli esecutivi se non sono stati regolarmente bollati sin dall'origine e, qualora si tratti di titoli provenienti dall'estero, prima che se ne faccia uso. Qualora non sia stata corrisposta l'imposta di bollo dovuta, e pagato le relative sanzioni amministrative, il portatore o possessore non può esercitare i diritti cambiari inerenti al titolo (art. 20, D.P.R. 26.10.1972, n. 642, parzialmente modificato dall'art. 5, comma 1, lett. *a*) del D. Lgs. 18 dicembre 1997, n. 473)⁴⁴.

11.2 Fra i titoli di formazione privata, occorre menzionare espressamente le scritture private autenticate relativamente alle obbligazioni pecuniarie in esse assunte dalla parte, ed introdotte (art. 2, comma 3, lett. *e*) del D. L. 14.3.2005, n. 35, convertito dalla L. 14.5.2005, n. 80) in sede di riforma del procedimento di esecuzione forzata, mediante l'aumento del novero dei titoli esecutivi di formazione stragiudiziale. A differenza dei titoli giudiziali per i quali l'azione esecutiva può essere promossa solo quando siano muniti della formula esecutiva, essi sono esenti da tale formalità (art. 475 C.p.c.).

⁴⁴ Con ordinanza n. 133 del 26 aprile 2004, la Corte Costituzionale ha dichiarato la manifesta inammissibilità della questione di legittimità costituzionale (sollevata in riferimento all'art. 3 della Costituzione) dell'art. 20 del D.P.R. n. 642/1972.

Pur conservando i caratteri intrinseci della scrittura privata, essi ricevono, inoltre, un'ulteriore tutela dalle norme del Codice penale (art. 491) per una loro equiparazione, *quoad poenam*, agli atti pubblici. A tale effetto, l'espressione cambiaria contenuta nell'art. 491 C.p., va intesa, secondo la giurisprudenza della Corte di Cassazione, nel significato più ampio, comprensivo sia del pagherò cambiario che della cambiale tratta, sia del titolo completo che di quello incompleto, poiché l'obiettività giuridica del reato è la tutela della pubblica fede inerente ai titoli di credito transmissibili per girata. Agli effetti del falso documentale, è cambiaria anche il documento privo di alcuno dei requisiti prescritti dalla legge, purché esso contenga la firma dell'emittente e sia trasferibile per girata.

12. Il documento nel diritto penale

12.1 Poiché il documento rappresenta, ai sensi dell'esposizione che precede, un valore, il diritto penale vi accorda protezione attraverso la previsione di una numerosa serie di ipotesi criminose, inserite, soprattutto, nella tipologia dei reati contro il patrimonio o in quella dei reati di falso, sotto il duplice aspetto della falsità materiale o della falsità ideologica. La nozione di atto pubblico o di scrittura privata sotto il profilo penalistico è autonoma e più ampia di quella prevista dall'ordinamento civile, in quanto la legge penale tutela il documento nella sua genuinità e veridicità, sia quale strumento probatorio, sia in sé, come principale espressione del bene giuridico della fede pubblica.

La definizione dettata dall'art. 2699 C.c. per l'atto pubblico non esaurisce l'intera categoria di tali documenti, sia perché il nostro ordinamento prevede accanto alla figura del pubblico ufficiale (art. 357 C.p.) quelle della persona incaricata di un pubblico servizio (art. 358 C.p.) e delle persone esercenti un servizio di pubblica necessità (art. 359 C.p.), sia perché i pubblici uf-

ficiali pongono in essere, nell'esercizio delle loro funzioni, anche atti per scopi diversi da quello di attribuire ad essi pubblica fede. Infatti, nella tutela penale rientrano due tipi di pubblico, rispettivamente previsti nell'art. 476 commi 1 e 2: quello formato dal pubblico ufficiale per uno scopo inerente alla funzione esercitata e quello facente fede fino a querela di falso. I successivi artt. 477 e 478, prevedono, inoltre, una specifica tutela dei certificati, delle autorizzazioni amministrative e degli attestati del contenuto di atti.

Nella previsione criminosa ipotizzata nell'art. 476, primo comma, C.p., l'espressione normativa dell'esercizio delle proprie funzioni deve essere intesa non in senso specifico, bensì generico di ambito delle funzioni, nel senso cioè che l'atto contraffatto o alterato deve rientrare, per sua natura, nella competenza funzionale del pubblico ufficiale o della persona incaricata di un pubblico servizio; onde il requisito sussiste se il soggetto ha il potere di collaborare in qualsiasi modo ed anche occasionalmente alla formazione dell'atto. Alla locuzione così specificata non può, inoltre, essere attribuito un significato strettamente limitato al tempo ed al luogo in cui il pubblico ufficiale svolge la sua attività, ma in essa va compreso ogni atto rientrante nell'ambito delle funzioni del pubblico ufficiale, in qualsiasi momento posto in essere, anche se abusivamente o con violazione dei propri doveri.

Anche per quanto riguarda la scrittura privata, la sua nozione, ai fini della giustizia penale ha un contenuto più ampio di quello previsto dall'ordinamento civile: non può restringersi a quegli atti che contengono dichiarazioni o manifestazioni di volontà idonee a produrre la nascita, la modificazione o l'estinzione di un diritto soggettivo, ma comprende anche tutte le scritture formate da un privato che si riferiscono a situazioni dalle quali possono derivare effetti giuridicamente rilevanti, vantaggiosi o dannosi per un determinato soggetto.

Ai fini dell'art. 485 C.p., il concetto di scrittura privata, non definito neanche dalla legge penale, va desunto dalla sua funzio-

ne specifica, che è quella di fissare in un documento, redatto senza l'assistenza di un pubblico ufficiale, qualsiasi dichiarazione di volontà o di scienza avente rilevanza nell'ambito di un rapporto giuridico.

Inteso nel suo significato più ampio, il documento può essere, quindi, precisato in una pluralità di significati e di funzioni. La scrittura privata può, infatti, riguardare non solo la nascita, l'esercizio, l'estinzione di un diritto soggettivo, ma anche qualsiasi circostanza idonea a spiegare effetti giuridici nell'ambito di un rapporto giuridico. Può, dunque, avere un valore culturale o scientifico e può possedere un valore patrimoniale, economico od anche solo affettivo riguardante tanto il contenuto ideale del documento quanto la materia di cui è costituito.

Sotto quest'ultimo aspetto, il concetto di documento può essere esteso sino a comprendere svariati oggetti, come, per esempio, una fotografia, un quadro, uno spartito musicale ed altri oggetti simili. In questi casi, nel documento si concretizza non solo il valore protetto dalle ipotesi criminose contro il patrimonio, quali il furto (art. 624 C.p.), l'appropriazione indebita (art. 646 C.p.) o il danneggiamento (art. 635 C.p.), ma anche quelle altre ipotesi criminose volte a prevenire e punire azioni contro altri beni, quali la malversazione (art. 315 C.p.) o il peculato (art. 314 C.p.) ovvero l'esercizio arbitrario delle proprie ragioni con violenza sulle cose (art. 392 C.p.).

12.3 Per quanto riguarda il gruppo dei delitti di falso documentale, è essenziale, al fine di poter giungere ad una delimitazione dei reati in esame, pervenire ad una determinazione del documento nel suo duplice significato di oggetto materiale della condotta e di bene giuridico protetto. Sotto la prima prospettiva, il documento viene considerato come una cosa sulla quale cade la condotta del falsario, mentre sotto il secondo profilo, il documento viene preso in esame quale bene penalmente pro-

tetto attraverso l'incriminazione del falso documentale⁴⁵: in quanto *res* su cui viene a svolgersi l'attività del falsario, il documento viene preso in esame quale oggetto materiale; in quanto bene penalmente protetto, il documento viene esaminato per la qualità che esso presenta.

In altre parole, in tema di falsità, la legge penale ha per oggetto la tutela del documento non solo come strumento probatorio, ma anche in sé come espressione del bene giuridico della fede pubblica, che viene ad avere un diverso contenuto secondo che si tratti di scrittura privata o di atto pubblico: rispetto alla prima, consiste nel credito che il documento trova nei confronti del pubblico per le necessità della vita sociale; rispetto al secondo, consiste, invece, in quel particolare credito che attiene agli atti formati dal pubblico ufficiale nell'esercizio della sue funzioni⁴⁶.

Comunque e in conclusione, va precisato che la falsità che non incide in alcun modo sull'esistenza, sull'efficacia e sul contenuto di un determinato documento è penalmente irrilevante, mentre è ipotizzabile nelle sue varie ipotesi normative, quando il documento conserva tutte le sue originarie caratteristiche di struttura e di contenuto. Del pari, è penalmente irrilevante la falsità grossolana e la falsità innocua. Ricorre la prima, quando l'inidoneità del documento a ledere la pubblica fede è assoluta, nel senso che la falsità sia rilevabile *ietu oculi* da chiunque, e sia tale che resti esclusa non la semplice probabilità, ma la stessa possibilità dell'inganno e quindi del documento alla pubblica fede. Il falso innocuo è quello che si rivela in concreto inidoneo a ledere l'interesse tutelato dalla genuinità del documento e che non ha la capacità di conseguire uno scopo antiggiuridico.

⁴⁵ Sul punto, v. le perspicue osservazioni di ALESSANDRO MALINVERNI, *Documento (dir. pen.)* in *Enciclopedia del diritto*, Milano, Giuffrè, vol. XIII, par. 2, p. 623.

⁴⁶ Cfr. la giurisprudenza formatasi in materia, e, in particolare, C. Cass. sent. n. 164744/84.

Il diritto di accesso ai documenti amministrativi

RENATO ROLLI

1. Considerazioni preliminari

Il diritto di accesso ai documenti amministrativi sarà analizzato alla luce delle innovazioni apportate con la legge di riforma del procedimento amministrativo (l. n. 15 del 2005) e con il contributo della giurisprudenza amministrativa formatasi sul punto, modulato secondo le linee di sviluppo di massima garanzia disegnate dal DPR n. 184 del 2006. Le nuove disposizioni della legge n. 15 del 2005 hanno destato alcune considerazioni critiche di carattere generale, posto che il progetto di riforma non sembra il prodotto di un piano coordinato, bensì della composizione di indirizzi giurisprudenziali, per lo più stabilizzati. D'altro canto, tenendo in conto che la realizzazione di regole per via giurisprudenziale è ritenuta atta a perfezionare, per via d'interpretazione, le formulazioni di principio della legge sul procedimento (l. n. 241 del 1990), è sorto – sin da subito – l'interrogativo sulla necessità della riforma medesima che, da un lato, abbonda di modifiche sulla forma, dall'altro circoscrive i contenuti di tipo sostanziale. In secondo luogo appare che l'intento sia stato quello di inserire, in un complesso normativo teso – in via generale – ad accogliere consolidati indirizzi giurisprudenziali, norme orientate alla concretizzazione di una sorta di controriforma della legge n. 241 del 1990. Risulta conseguirne, per un verso, la riduzione degli ambiti di garanzia, trasparenza e tutela nei confronti dei

privati in relazione al procedimento amministrativo e, per l'altro, l'incrinatura dei doveri di responsabilità delle amministrazioni pubbliche, determinando forme di aggravamento nell'agire dell'amministrazione, con difetto di bilanciamento dei benefici nei confronti del cittadino. In terzo luogo, l'interesse verso l'ordinamento, con riferimento ad alcune parti della normativa, sembrerebbe porsi soltanto in forma attuale e non in forma proiettiva. L'attestazione più palese è offerta dalla disciplina dell'accesso ai documenti¹.

¹ Il diritto di accesso e la relativa disciplina è oggetto di un intenso dibattito dottrinale. Si vedano, almeno, MARIA ALESSANDRA SANDULLI, *Accesso alle notizie e ai documenti amministrativi*, in «Enciclopedia del diritto», Agg., IV, 2000; ALDO SANDULLI, *Il procedimento*, in «Trattato di diritto amministrativo» (a cura di SABINO CASSESE), Parte generale, Milano, 2003, spec. p. 1157; RICCARDO VILLATA, *La trasparenza dell'azione amministrativa*, in «Diritto processuale amministrativo», 1987, p. 528; GREGORIO ARENA, *Trasparenza amministrativa*, in «Enciclopedia giuridica», XXXI, Roma; MARCO MAZZAMUTO, *Sul diritto d'accesso nella l. n. 241/90*, in «Foro amministrativo Consiglio di Stato», 1992, p. 1571; CARLO EMANUELE GALLO, SERGIO FOA', *Accesso agli atti amministrativi*, in «Digesto discipline pubblicistiche» 2000, p. 10; ROBERTO CAVALLO PERIN, *Validità del provvedimento e dell'atto amministrativo*, in «Digesto Discipline pubblicistiche», 2000, XV, p. 613; FRANCESCO CARINGELLA- ROBERTO GAROFOLI- MARIA TERESA SEMPREVIVA, *L'accesso ai documenti amministrativi. Profili sostanziali e processuali*, 2° ed., Milano, 2003; ANTONIO ROMANO TASSONE, *A chi serve il diritto di accesso (Riflessioni su legittimazione e modalità d'esercizio del diritto d'accesso nella l. n. 241 del 1990)*, in «Diritto amministrativo», 1995; BARBARA GAGLIARDI, *Il diritto di accesso alle informazioni ambientali e la nozione di "ambiente" ai sensi dell'art. 2 d.lg. n. 39 del 1997*, in «Foro amministrativo Consiglio di Stato», 2003, p. 1032; in particolare sul diritto di accesso da parte delle associazioni dei consumatori come mezzo per la protezione degli interessi che rappresentano, LUIGI CARBONE, *Diritto di accesso delle associazioni di tutela dei consumatori*, in «Corriere giuridico», 2002, p. 24; SALVATORE MINE-TOLA, *Il diritto di accesso riconosciuto alle associazioni dei consumato-*

La legge n. 241 del 1990, nel disciplinare i rapporti tra cittadini e pubblica amministrazione, ha riconosciuto il diritto di accesso ai documenti amministrativi al fine di assicurare la trasparenza dell'attività amministrativa e di favorirne lo sviluppo imparziale; la giurisprudenza ha chiarito che il riconoscimento nel nostro ordinamento del principio di pubblicità dei documenti amministrativi segna un totale cambiamento di prospettiva, perché comporta che se un tempo il segreto era la regola e la pubblicità l'eccezione, ora è vero il contrario. Il "diritto di informazione" nei riguardi degli enti pubblici ha luogo, in generale, attraverso il diritto di accesso ai documenti amministrativi. Su tale istituto, dopo il riconoscimento legislativo della legge 8 giugno 1990, n. 142 e poi dalla legge 7 agosto 1990, n. 241, sono stati apportati incisivi cambiamenti, determinati intanto dall'approvazione della disciplina sulla tutela della riservatezza (l. 31 dicembre 1996, n. 675 e il d.lgs del giugno 2003, n. 196, c.d. "Codice della privacy"); in secondo luogo di quella sull'accesso alle informazioni riguardanti l'ambiente²; infine per effetto della legge 11 febbraio 2005, n. 15, con cui il Capo V della l. n. 241 del 1990 è stato in sostanza riformulato³.

Va subito rilevato che l'istituto in questione non ha mantenuto il suo primitivo aspetto unitario; si presenta attualmente suddiviso in varie tipologie, perché le disposizioni contenute nei citati testi legislativi hanno inserito sul diritto di accesso una molteplicità di discipline. In verità, già nel 1990 erano presenti alcu-

ri non può trasformarsi in potere esplorativo nei confronti dell'azione della p.a., in «Consiglio di Stato», 2001, p. 2179; ROSARIO FERRARA, *Consumatore (protezione del) nel diritto amministrativo*, in «Digesto discipline pubblicistiche», III, 1998, p. 515; VINCENZO ZENO ZENCOVICH, *Consumatore (tutela del)*, in «Enciclopedia giuridica», VIII, Roma.

² Direttiva 2003/4/CE del Parlamento Europeo e del Consiglio del 28 gennaio 2003 e d.lgs 24 febbraio 1997, n. 39.

³ Riviste, inoltre, dal D.P.R. n. 184 del 2006.

ne diversità tra l'accesso agli atti degli enti locali (art. 7 l. 142/90 trasfuso nell'art. 10 d.lgs. 18 agosto 2000 n. 267) e quello disciplinato attraverso la l. n. 241/90. Le diversità concernono in special modo la legittimazione attiva, che la legge relativa all'ordinamento degli enti locali riporta allo *status* di cittadino piuttosto che alla titolarità di un "interesse per la tutela di situazioni giuridicamente rilevanti" (trasformato, in modo più circoscritto in "interesse personale e concreto per la tutela di situazioni giuridicamente rilevanti" durante la vigenza del D. P. R. 27 giugno 1992, n. 352) risolte dalla giurisprudenza che ha privilegiato la normativa offerta dal capo V della l. n. 241/90. Questo processo di *reductio ad unitatem* è divenuto comunque più difficoltoso per il crescente spezzettamento del diritto di accesso, frutto della sovrapposizione alle informazioni ambientali delle regole riguardanti la riservatezza e l'accesso. Attualmente, a seguito dei cambiamenti prodotti dalla l. n. 15/2005, sembra verosimile preferire la differente qualificazione di "diritti di accesso" in luogo di "diritto di accesso", nei riguardi degli enti pubblici. I "diritti di accesso" vengono disciplinati attraverso la l. 241/90, il d. lgs. 196/2003, la direttiva 2003/4/CE e il d.lgs. n. 39/1997. Va evidenziato che la particolare normativa di alcuni tipi di accesso è regolamentata non solo attraverso norme relative ad uno dei tre testi legislativi, ma è anche orientata attraverso le norme della l. 241/90 e del Codice della privacy che si coniugano insieme. In effetti, non è sempre determinabile la preponderanza di "gruppi di articoli" derivati soltanto da una delle due leggi, al punto da non considerare le altre disposizioni in materia, in modo da creare un (quasi) inestricabile nodo gordiano.

Secondo la normativa vigente, la dottrina distingue i diritti di accesso in: 1) diritto di accesso extraprocedimentale (o informativo), azionato dal detentore dei dati personali al fine di riceverne la comunicazione in forma chiara; 2) diritto di accesso extraprocedimentale al fine di conoscere documenti su dati personali di titolarità di un soggetto distinto dall'istante; 3) diritto di ac-

cesso extraprocedimentale al fine di conoscere documenti su dati sensibili e giudiziari di titolarità di un soggetto distinto dall'istante; 4) diritto di accesso extraprocedimentale al fine di conoscere documenti su "dati supersensibili" di titolarità di un soggetto distinto dall'istante; 5) diritto di accesso extraprocedimentale al fine di conoscere documenti su "dati superipersensibili" di titolarità di un soggetto distinto dall'istante; 6) diritto di accesso endoprocedimentale; 7) diritto di accesso alle informazioni ambientali⁴.

2. Il diritto di accesso dopo la riforma sul procedimento amministrativo: spunti problematici

Diverse sono le novità apportate dalla legge di riforma sul procedimento amministrativo; così come diversi i dubbi e le perplessità che essa ha destato⁵. Un importante elemento di in-

⁴ Cfr. MASSIMO OCCHIENA, *I diritti di accesso dopo la riforma della l. n. 241 del 1990*, in «Foro amministrativo Tar», 2005, p. 905, ss.

⁵ Nella vasta letteratura formatasi a seguito della riforma del procedimento amministrativo, in senso critico, FILIPPO SATTA, *La riforma della legge n. 241/1990: dubbi e perplessità*, in <www.giustamm.it>; VINCENZO CERULLI IRELLI, *Osservazioni generali sulla legge di modifica della l. n. 241 del 1990*, in <www.giustamm.it>; SILVIO GAMBINO, *Il diritto di accesso. La nuova disciplina legislativa (l.n. 15/2005) alla luce della revisione costituzionale del Titolo V, con particolare riferimento all'ambito di applicazione regionale e a quello degli enti locali*, in «Le Corti calabresi» n. 2 del 2005, p. 419, ss.; MASSIMO OCCHIENA, *Il diritto di accesso dopo la riforma della l. n. 241 del 1990*, cit.; GREGORIO ARENA-MARCO BOMBARDELLI, *Il diritto di accesso ai documenti amministrativi*, in «La disciplina generale dell'azione amministrativa» (a cura di VINCENZO CERULLI IRELLI), Napoli, 2006, p. 412; FRANCO GAETANO SCOCA, *Il punto sulla disciplina legislativa del procedimento amministrativo*, in Giustizia amministrativa, <www.giustamm.it>; SALVATORE CIMINI,

novazione e garanzia è rappresentato dalla comunicazione di avvio al procedimento di accesso ai soggetti controinteressati. (art. 3, D.P.R. n. 184/06). Ciò comporta, per l'amministrazione la necessaria valutazione dei documenti, per verificare se vi siano soggetti contro interessati, le comunicazioni dei quali possono concretarsi anche in opposizioni motivate alla concessione dell'accesso (con conseguente valutazione del responsabile del procedimento ai fini della decisione finale). Inoltre il regolamento n. 184 del 2006 (art. 6, comma 5), assicura una ulteriore garanzia procedurale per cui ove la richiesta di accesso sia irregolare o incompleta, l'amministrazione, entro dieci giorni, ne dà comunicazione al richiedente con raccomandata con avviso di ricevimento ovvero con altro mezzo idoneo a comprovarne la ricezione. In tal caso, il termine del procedimento ricomincia a decorrere dalla presentazione della richiesta corretta. Da una analisi globale del regolamento di attuazione del 2006, emerge

Diritto di accesso e riservatezza: il legislatore alla ricerca di nuovi equilibri, in <www.giustamm.it>; MARIA ALESSANDRA SANDULLI, *La riforma della legge sul procedimento amministrativo tra novità vere ed apparenti*, in <www.federalismi.it>; ALDO TRAVI, *La l. n. 15 del 2005: verso un nuovo diritto amministrativo*, in «Corriere giuridico», n. 4 del 2005; GREGORIO ARENA, *Trasparenza amministrativa*, in «Dizionario di diritto pubblico» (a cura di Sabino Cassese), Milano 2006, Vol. VI; BERNARDO GIORGIO MATTARELLA, *Le dieci ambiguità della l. n. 15 del 2005*, in «Giornale di diritto amministrativo», 2005, p. 821; MASSIMO CIAMMOLA, *Il diritto di accesso all'informazione ambientale: dalla legge istitutiva del Ministero dell'ambiente al decreto legislativo n. 195 del 2005*, in «Foro amministrativo Consiglio di Stato», 2007, p. 657; Id., *La legittimazione ad accedere ai documenti amministrativi (prima e dopo la l. 11 febbraio 2005, n. 15)*, in «Foro amministrativo Tar», n. 3 del 2007, p. 1181, ss.; MARIA GRAZIA DELLA SCALA, *Il diritto di accesso all'attività privatistica dell'amministrazione e dei soggetti privati alla luce della l. n. 15 del 2005. Condanna all'esibizione degli atti e ottemperanza al giudicato*, in «Diritto amministrativo», 2006, p. 173.

una duplice possibilità di accesso: formale e informale. La distinzione è determinata dalla presenza o meno di soggetti controinteressati. Ebbene, nel caso in cui, in base alla natura del documento richiesto non risultino controinteressati, il diritto di accesso sarà esercitato in via informale mediante richiesta, anche verbale, all'ufficio dell'amministrazione competente a formare l'atto conclusivo del procedimento (o a detenerlo stabilmente, secondo l'art. 5, comma 1). Il richiedente dovrà indicare gli estremi del documento che richiede; ovvero gli elementi che ne consentano l'individuazione; specificare e, se del caso, comprovare l'interesse connesso all'oggetto della richiesta, dimostrare la propria identità e, se del caso, i propri poteri di rappresentanza del soggetto interessato (art. 5, comma 2). Allorquando l'amministrazione, sulla base del contenuto del documento richiesto, verifichi l'esistenza di controinteressati, invita l'interessato (art. 6) a presentare richiesta formale di accesso. Da qui si apre una procedura formale (art. 6, comma 1), e dunque: 1) qualora non sia possibile l'accoglimento immediato della richiesta in via informale; 2) qualora sorgano dubbi sulla legittimazione del richiedente o sulla sua identità o sui suoi poteri rappresentativi, sulla sussistenza dell'interesse alla stregua delle informazioni e delle documentazioni fornite; 3) qualora sorgano dubbi sull'accessibilità del documento o sull'esistenza di controinteressati. Al verificarsi di tali situazioni, l'amministrazione invita l'interessato a presentare richiesta d'accesso formale, di cui l'ufficio rilascia ricevuta (secondo il procedimento descritto dall'art. 6, comma 4, D.P.R. n. 184 del 2006). Intanto, anche a seguito della novella, non risulta rilevabile alcuna evoluzione circa la legittimazione ad accedere, perché in luogo dell'espressione meno restrittiva «chiunque sia titolare di una situazione giuridicamente rilevante» è stata inserita una formulazione più rigorosa «tutti i soggetti privati [...] che abbiano un interesse diretto, concreto ed attuale, corrispondente ad una situazione giuridicamente tutelata e collegata al documento al quale è chiesto

l'accesso»⁶. Tale locuzione che, da un lato, riflette l'art. 2, c.1, D.P.R. n. 352/1992⁷ (ora abrogato dal D.P.R. 12 aprile 2006, n. 184) e, dall'altro, accetta alcuni indirizzi giurisprudenziali, sembra meno estensiva dell'altra. Ciò si evince sia in rapporto al giudizio concernente gli aspetti concreti ed attuali dell'interesse, sia in rapporto alla sostituzione, rispetto alla situazione giuridica, della espressione "rilevante" con la voce "tutelata". Da qui si ripresenta la questione sulla possibilità di richiedere l'accesso soltanto ove sussista una posizione sottesa associata ad una situazione giuridica soggettiva che l'ordinamento tutela. Sul punto, è necessario sottolineare che, secondo l'impostazione della Commissione Nigro⁸ e del legislatore del 1990, si tendeva alla "casa di vetro" turatiana mediante un percorso grada-

⁶ Il diritto di accesso ai documenti amministrativi «non si atteggia come una sorta di azione popolare diretta a consentire una forma di controllo generalizzato dell'amministrazione, giacché, da un lato, l'interesse che legittima ciascun soggetto all'istanza, da accertare caso per caso, deve essere personale e concreto e ricollegabile al soggetto stesso da uno specifico nesso e, dall'altro, la documentazione richiesta deve essere direttamente riferibile a tale interesse, oltre che individuata o ben individuabile; pertanto, l'accesso de quo non è uno strumento di controllo ispettivo, utilizzabile al solo scopo di sottoporre a verifica generalizzata l'operato dell'amministrazione». Cons. Stato, Sez. VI, 17 marzo 2000, n. 1414; Cons. Stato, Sez. V?, 16 dicembre 1998, n. 1683, in <www.giustizia-amministrativa.it>.

⁷ Regolamento per la disciplina delle modalità di esercizio e dei casi di esclusione del diritto ai documenti amministrativi, in attuazione dell'art. 24 comma 2, l. 7 agosto 1990, n. 241.

⁸ Ed in base agli studi dell'autorevole studioso: MARIO NIGRO, *Il nodo della partecipazione*, in «Politica e regioni», 1981, ora in «Scritti giuridici», Milano, 1996, II, p. 707; ID., *Procedimento amministrativo e tutela giurisdizionale contro la pubblica amministrazione (il problema della legge generale del procedimento amministrativo)*, in «Rivista di diritto processuale», 1980, ora in «Scritti giuridici», cit., p. 1429; ID., *Il procedimento amministrativo tra inerzia legislativa e trasformazioni dell'amministrazione. A proposito di un recente disegno di legge*, ivi, p. 2039.

to, teso a realizzare una completa espansione della legittimazione del diritto ad accedere⁹. Si era ritenuto allora, che non sussistessero le condizioni, per l'ordinamento giuridico italiano, ai fini di un ampliamento del principio di trasparenza ed inoltre che fosse necessario differire un'applicazione più densa delle disposizioni in materia di accesso, in considerazione delle opposizioni delle pubbliche amministrazioni. Invero non sembrano emergere tra il 1990 ed oggi rilevanti modificazioni, se il legislatore del 2005, non richiamando alcune discipline (ad esempio quelle riguardanti l'informazione ambientale e le autonomie locali) e non considerando altri ordinamenti europei, ha stabilito una disciplina parimenti limitata per quanto riguarda il diritto di accesso. In virtù degli artt. 22, c.2, e 29 c.2 l. *proc.amm.*, le regioni e gli enti locali decidono in modo indipendente sull'ampliamento delle garanzie nei confronti dei privati. Infine sembra problematica l'interpretazione di alcune norme di cui si darà conto in seguito: si pensi all'art. 22, c.3 ove si specifica «tutti i documenti amministrativi sono accessibili, ad eccezione di quelli indicati all'art. 24, commi 1, 2, 3, 5 e 6». Vi è da chie-

⁹ Così ALDO SANDULLI, *L'accesso ai documenti*, in «Giornale di diritto amministrativo», n. 5/2005, p. 494, ss. In giurisprudenza si veda Tar Lombardia, sent. 11 aprile 2005, in «Foro amministrativo Tar», 2005, p. 1845, secondo cui l'impostazione della legge n. 241 del 1990, in materia di accesso «ha indotto la giurisprudenza stessa ad affermare che il diritto di accesso ai documenti amministrativi riconosciuto dalla l. n. 241 del 1990 prevale sull'esigenza di riservatezza del terzo ogniqualvolta l'accesso venga in rilievo per la cura o la difesa di interessi giuridici del richiedente (v. Cons. Stato, Ad. Plen. n. 5 del 1997) anche se, beninteso, l'interesse vantato deve essere concretamente collegato alle specifiche esigenze del richiedente e, come tale, deve essere serio, non emulativo, e non riconducibile a mera curiosità, nonché riferibile al medesimo in quanto titolare di una posizione soggettiva giuridicamente rilevante e qualificata dall'ordinamento come meritevole di tutela».

dersi allora se un documento tutelato da Segreto di Stato o comprendente «informazioni di carattere psico-attitudinale relative a terzi» possa escludersi dall'accesso solo ove le singole pubbliche amministrazioni lo includessero nella serie dei documenti sottratti previsti dall'art. 24 c.2. Ed ancora l'art. 22, c.4, l. *proc.amm.* detta che «non sono accessibili le informazioni in possesso di una pubblica amministrazione che non abbiano forma di documento amministrativo, salvo quanto previsto dal d.lgs. n. 196/2003 in materia di accesso a dati personali da parte della persona cui di dati si riferiscono». Da qui, essendo previsto dalla legge l'accesso ai documenti e sussistendo una normativa specifica che disciplina l'accesso ai propri dati personali, sorge l'interrogativo se debbano ritenersi segrete tutte le informazioni non incluse nei documenti. Inoltre, l'art. 22, c.6 l. *proc.amm.*, stabilisce che «il diritto di accesso è esercitabile fino a quando la pubblica amministrazione ha l'obbligo di detenere i documenti amministrativi ai quali si chiede di accedere». Orbene, come è noto, alcune disposizioni stabiliscono che, dopo un certo periodo di tempo, alle pubbliche amministrazioni non compete più l'obbligo di possedere la documentazione, ma possono eseguire la relativa archiviazione¹⁰. Per cui emerge l'ulteriore interrogativo se l'amministrazione, ove detenesse ancora il documento richiesto, abbia il diritto, essendo terminato il periodo di obbligatoria detenzione, di negare l'accesso al documento.

3. Accesso ai documenti e riservatezza dei terzi

La relazione tra diritto di accesso ai documenti e tutela della riservatezza dei terzi è regolata da una disciplina che presenta

¹⁰ Artt. 67-68 d. P.R. n. 445/2000; artt. 122, ss., d.lgs n. 42/2004.

una sequenza di rinvii incrociati tra la legge n. 241/90 e il d. lgs. n. 196/2003 (Codice sulla privacy)¹¹. Le disposizioni che interessano – e delle quali si darà conto in seguito – sono da un lato l'art. 24, c.6, lett. d) e l'art. 24, c. 7), appartenenti alla legge n. 241/90; dall'altro gli art. 59 e 60 inclusi nel Codice sulla privacy. Questo articolato gruppo di norme consente tuttavia di individuare quattro forme di relazioni tra accesso e riservatezza. La prima riguarda l'eventualità secondo cui l'istanza di accesso non

¹¹ Sul tema cfr., almeno, CESARE MASSIMO BIANCA -FRANCESCO DONATO BUSNELLI (a cura di), *Tutela della privacy (l. 31 dicembre 1996, n. 675), Commentario*, in «Le nuove leggi civili commentate», 1999; FRANCESCO CARINGELLA, *Riservatezza ed accesso. I rapporti tra le l. n. 241 del 1990 e l. n. 675 del 1996 in tema di tutela dei dati personali*, in FRANCESCO CARINGELLA- ROBERTO GAROFOLI – MARIA TERESA SEMPREVIVA, *L'accesso ai documenti amministrativi*, Milano, 2003. ID., *Ancora fitta la nebbia sull'accesso ai dati personali (sensibili e non)*, in «Il Foro Italiano», 1998, III, p. 610; MARCO BOMBARDELLI, nota a TAR Abruzzo, Pescara, n. 681/97, in «Giornale di diritto amministrativo», n. 6/98, p. 532; MASSIMO OCCHIENA, *Diritto di accesso, atti di diritto privato e tutela della riservatezza dopo le legge sulla privacy*, in «Diritto processuale amministrativo», 1998, p. 377; ROSARIO FERRARA, *Premesse ad uno studio sulle banche dati della pubblica amministrazione fra regole di concorrenza e tutela della persona*, cit., 1997, p. 556; MARCO LIPARI, *L'amministrazione pubblica tra trasparenza e riservatezza*, *Diritto amministrativo*, in <www.giustizia-amministrativa.it> ; MARIA CRISTINA CAVALLARO, *Diritto d'accesso ai documenti amministrativi e tutela della privacy: un difficile equilibrio?*, in «Foro amministrativo Tar», 2002, p. 2929; SALVATORE GIACCHETTI, *Accesso e riservatezza: separati in casa*, in «Consiglio di Stato», 2004, II, p. 463; ANDREINA SCOGNAMIGLIO, *Diritto di accesso e tutela della riservatezza*, in «Foro amministrativo Tar», 1998, p. 983; STEFANO TARULLO, *Diritto di accesso ai documenti amministrativi e diritto alla riservatezza: un difficile rapporto*, in «Jus», 1996, p. 250; SALVATORE BELLOMIA, *Un passo avanti e uno indietro della giurisprudenza amministrativa in materia di accesso*, in «Giustizia civile», 2002, I, p. 249.

contempli documenti «la cui conoscenza sia necessaria per curare o difendere i propri interessi giuridici»: ne consegue che la riservatezza prevale sull'accesso. La seconda fa riferimento al caso per cui la richiesta di accesso concerne documenti indispensabili per la tutela di interessi giuridici: ne deriva che l'accesso ha prevalenza sulla riservatezza. La terza riguarda l'ipotesi per cui dati sensibili e giudiziari siano presenti nei documenti; in questo caso sussiste la prevalenza dell'accesso dove la conoscenza dei documenti si qualifichi come “strettamente indispensabile”. Da ultimo, la quarta considera l'istanza di accesso che attenga a dati cosiddetti iper-sensibili; segnatamente a «dati idonei a rivelare lo stato di salute e la vita sessuale». Sul punto – come si vedrà – la l. *proc.amm.*, rimanda all'art. 60, d.lgs. 196/2003, che prevede la prevalenza dell'accesso ove il caso giuridicamente rilevante da garantire risulti di livello se non altro uguale al diritto sulla riservatezza di cui il controinteressato è titolare («ovvero consiste in un diritto della personalità o in un altro diritto o libertà fondamentale o inviolabile»). A tal punto va evidenziato in primo luogo che gli elementi di separazione tra “necessità” e “stretta indispensabilità” non sono facilmente determinabili; in secondo luogo il legislatore, con riferimento all'ultimo caso, ha rinviato la decisione alla pubblica amministrazione, offrendo semplicemente alcuni indicatori. Pertanto, dal momento che è meno frequente il primo caso (dove si presenta ben delineato il risultato della scelta che il legislatore ha eseguito), sembra plausibile evidenziare che un alto livello di potere discrezionale amministrativo caratterizzi la disciplina dei rapporti intercorrenti tra accesso e riservatezza. Sul punto, sorge l'interrogativo se tale situazione non confligga con la qualificazione di diritto soggettivo perfetto dell'accesso ai documenti.

3.1 Diritto di accesso e dati sensibili e giudiziari

Ai sensi degli artt. 20 e 21 del Codice della privacy, è ammesso il “trattamento” dei dati sensibili e giudiziari ad opera di

soggetti pubblici soltanto in presenza dell'autorizzazione di esplicita disposizione di legge o di un provvedimento del Garante (che nell'ipotesi dei dati sensibili ha il compito di riconoscere le «attività, tra quelle demandate ai medesimi soggetti dalla legge, che perseguono finalità di rilevante interesse pubblico e per le quali è conseguentemente autorizzato [...] il trattamento dei dati sensibili» art. 20, comma 3, del Codice) in cui si precisano le tipologie di dati oggetto di trattamento e di procedure attuabili e gli obiettivi di notevole interesse pubblico prefissati¹². Secondo la definizione di “trattamento” ex art. 4, comma 1, lett. a, le pubbliche amministrazioni hanno la facoltà di rendere noti a terzi – quindi consentire l'accesso – documenti su dati sensibili soltanto nelle situazioni che la legge stabilisce in modo esplicito.

Ove esista la legge o l'autorizzazione del Garante per cui è consentito ad un ente pubblico il “trattamento” (da ritenere in proposito nei significati di “consultazione”, “estrazione” e “comunicazione”) di dati sensibili e giudiziari, nell'ipotesi di istanza di accesso a documenti che comprendono tali dati, l'amministrazione detentrica ha l'obbligo di applicare l'art. 24, comma 7, l. 241/90. Viene invero previsto che «nel caso di documenti con-

¹² Utilizzare i dati di cui si tratta impone – come si è visto – soluzioni a delicati problemi. Il Codice sulla privacy, all'art. 22, ne consente l'impiego, solo «per svolgere attività istituzionali che non possono essere adempiute, caso per caso, mediante il trattamento di dati anonimi o di dati personali di natura diversa». La norma prevede, altresì, forme di conservazione elettronica che apprestano maggiori garanzie, «richiedendo, in particolare, tecniche di cifratura, utilizzazione di codici identificativi o altre soluzioni che, considerato il numero e la natura dei dati trattati, siano atti a renderli temporaneamente indecifrabili anche a chi è autorizzato ad accedervi, permettendo l'identificazione degli interessati solamente in caso di necessità». Così FRANCESCO GAVERINI, *Accesso (e accessi) ai documenti amministrativi*, cit., p. 1851.

tenenti dati sensibili e giudiziari, l'accesso è consentito nei limiti in cui sia strettamente indispensabile e nei termini previsti dall'art. 60 del d.lgs 30 giugno 2003, n. 196, in caso di dati idonei a rivelare lo stato di salute e la vita sessuale». Tale disposizione sembra distinguere l'accesso ai dati sensibili e giudiziari da quelli semplicemente personali e da quelli sensibili concernenti la vita sessuale e lo stato di salute. Da qui l'ente pubblico, che è in possesso dei documenti contenenti quei dati, giudica se essi siano realmente necessari ai fini della cura e tutela di interessi giuridici del richiedente. Va anche rilevato per l'istante una rigorosa segnalazione delle ragioni a sostegno della "stretta indispensabilità" di conoscere i dati sensibili e giudiziari di soggetti terzi; nonché per l'amministrazione il potere e l'obbligo¹³ di stabilire che tale conoscenza sia inevitabile ai fini della cura e tutela degli interessi del richiedente. In più, atteso che i controinteressati sono stati riconosciuti in modo chiaro da parte della legge, l'ente pubblico avrà anche l'obbligo di rendere noto al soggetto detentore dei dati sensibili e giudiziari l'esistenza di una istanza di accesso a quei dati riservati, perché egli intervenga – nel ruolo di "controinteressato sostanziale" – al conseguente procedimento secondo l'art. 10 della l. n. 241/90.

3.2 Segue: il diritto di accesso ai "dati supersensibili"

Secondo l'art. 24, comma 7, della legge sul procedimento amministrativo, l'accesso «ai dati idonei a rivelare lo stato di salute e la vita sessuale» si attiene al disposto dell'art. 60 del Codice privacy. Tale norma ha contemplato una particolare disciplina relativa a questi dati, classificati come "dati ipersensibili"¹⁴. Un

¹³ Cfr. MARCELLO CLARICH, *Trasparenza e protezione dei dati personali nell'azione amministrativa*, in «Foro amministrativo Tar», 2004, p. 3890.

¹⁴ Si vedano sul punto le considerazioni di GIUSEPPE MORBIDELLI, *Il procedimento amministrativo*, in «Diritto amministrativo», a cura di LEO-

soggetto, che non ne sia possessore, può esaminarli «se la situazione giuridicamente rilevante che si intende tutelare con la richiesta di accesso ai documenti amministrativi è di rango almeno pari ai diritti dell'interessato, ovvero consiste in un diritto della personalità o in un altro diritto o libertà fondamentale e inviolabile». Si pone così di nuovo in evidenza il “potere arbitrare”¹⁵ dell'ente pubblico di analizzare e stabilire i dovuti bilanciamenti e le proporzioni tra situazioni giuridiche differenti. Segnatamente, da un lato esiste la situazione giuridica di colui che ha esibito l'istanza di accesso¹⁶; dall'altro quella dell'interessato dei “dati supersensibili” implicati nella istanza medesima. È certo che l'amministrazione, in virtù della attribuzione da parte delle legge, può stabilire in ordine ad ogni singolo caso i criteri di organizzazione degli interessi sottostanti all'istanza di accesso e alla difesa della riservatezza¹⁷. In altri termini, l'ente che possiede il documento – di cui è richiesta la presentazione – ha il compito di decidere applicando un potere – dovere, esercitando una valutazione discrezionale dando prevalenza all'uno o all'altro diritto, alla luce delle situazioni implicate nel caso specifico¹⁸.

POLDO MAZZAROLLI, GIUSEPPE PERICU, ALBERTO ROMANO, FABIO ANTONIO ROVERSI MONACO, FRANCO GAETANO SCOCA, Bologna, 2005, p. 681.

¹⁵ Ancora MASSIMO OCCHIENA, *I diritti di accesso*, cit., p. 921.

¹⁶ Posto che, in generale, la domanda di accesso deve essere finalizzata alla tutela di uno specifico interesse giuridico di cui il richiedente è portatore, secondo Cons. Stato, Sez. VI, 30 settembre 1998, n. 1346.

¹⁷ La domanda di accesso non può essere un mezzo per compiere una indagine o un controllo ispettivo, cui sono ordinariamente preposti organi pubblici, perché in tal caso nella domanda di accesso è assente un diretto collegamento o specifiche situazioni giuridicamente rilevanti (Cons. Stato, Sez. IV, 29 aprile 2002, n. 2283; Tar Lazio sez. II, 22 luglio 1998, n.120).

¹⁸ Ampie forme di tutela sono estese in relazione ai dati idonei a rivelare lo stato di salute e la vita sessuale, che sono sottoposti alle stesse regole che riguardano i dati giudiziari, non solo per il caso di trattamenti informatici,

3.3 Segue: il diritto di accesso ai “dati superipersensibili”

Si discosta, dai precedenti, il diritto di accesso extraprocedimentale relativo alla conoscenza di documenti amministrativi su «procedimenti selettivi [...] concernenti informazioni di carattere psico-attitudinale relativi a terzi», per i quali l'art. 24 comma 1, lett. *d*, l. n. 241/90 stabilisce l'esclusione dall'esercizio del diritto di accesso. Non contemplando la legge nessun caso particolare sulla limitazione dell'accessibilità, sembra opportuna la qualificazione di “dati superipersensibili”¹⁹ che evidenzia la diversità da quelli ex art. 60 del Codice della privacy. In più, il divieto di accesso a tali dati, secondo la dottrina parrebbe derivare da una disattenzione del legislatore. Effettivamente, benché la legislazione in materia presenti anche atti segreti, non risulta ragionevole la sottrazione di tali dati al diritto di accesso, se si fa riferimento in special modo, a quanto stabilito in ordine ai “dati superipersensibili”. Ed invero sembra incomprensibile che il principio della trasparenza si debba limitare davanti a documenti che rendono edotti sulle facoltà psico-attitudinali di candidati a concorsi selettivi e, considerando segnatamente l'eccezione in ordine alla riservatezza dei dati personali e sensibili stabiliti attraverso l'art. 24, comma 7, della legge n. 241/90. Non risulta chiaro, ad esempio, che davanti ad una non soddisfacente posizione in graduatoria, non venga consentito a chi si sia candidato in una procedura selettiva, di esaminare elementi indispensabili allo scopo di tutelare l'interesse legittimo che vanta. Sul punto, si può evidenziare anche che in frequenti concorsi di pubblico impiego è

ma anche nel caso in cui gli elenchi, i registi o le banche dati siano tenuti senza l'ausilio di strumenti elettronici (art. 34 Codice privacy). È poi prevista una conservazione separata delle informazioni di cui si tratta rispetto a quelle trattate per finalità che non richiedono il loro utilizzo, con la finalità di evitare conoscenza occasionale, casuale dei dati medesimi.

¹⁹ Così definiti da MASSIMO OCCHIENA, *I diritti di accesso*, cit., p. 922.

contemplata la prova di test o colloqui di natura psico-attitudinale per valutare la capacità dei partecipanti ad occupare i posti messi a concorso. A tal punto, la norma in questione indurrebbe a ritenere che questa fase del concorso sia inaccessibile. Orbene, sarebbe conveniente che l'entità di tale regola venisse ridotta ad opera della giurisprudenza. La dottrina avanza l'ipotesi dell'assegnare alla espressione "nei procedimenti" (selettivi) il senso di "durante le procedure" (selettive): in tale guisa l'applicazione del divieto si limiterebbe soltanto ai casi di accesso ai documenti circa le capacità psico-attitudinali di soggetti terzi nel corso del procedimento selettivo, tranne ammetterne la conoscenza a conclusione del concorso²⁰.

3.4 Accesso endoprocedimentale

Resta da esaminare, a questo punto, un ulteriore caso di accesso, regolato dalla legge sul procedimento. Si tratta dell'accesso eseguito all'interno della partecipazione al procedimento amministrativo (art. 10 l. 241/90). Al di là della tipologia dei dati (personali, sensibili, giudiziari, supersensibili, superipersensibili) dell'istante oppure di un terzo, chi interviene al procedimento si rapporta sempre con un potere amministrativo; pertanto il diritto di accesso acquisterebbe la natura di interesse legittimo e non di diritto soggettivo. D'altro canto, durante l'esplicazione del potere, l'amministrazione che adotta la decisione può vietare la conoscenza di documenti in base a motivi istruttori. Sul punto, la legge attribuisce all'amministrazione il potere di differimento (artt. 24, comma 4, e 25, comma 3 l. n. 241/90). Questo si applica soprattutto all'interno della procedura amministrativa, allo scopo di garantire gli effetti della decisione conclusiva; inoltre viene generalmente utilizzato allorquando la conoscenza di alcuni dati ad opera dei cittadini implicati potrebbe ostacolare la

²⁰ Ivi, p. 923.

decisione finale nella manifestazione dei suoi effetti. Salvo che non venga in modo esplicito statuito²¹, il potere di vietare la conoscenza degli atti endoprocedimentali è vincolato a delle limitazioni affinché non siano vanificate le norme sulla partecipazione e sulla trasparenza. Anche la Corte costituzionale con la sentenza del 3 novembre 2000, n. 460²² ne ha colto la rilevanza analizzando la normativa speciale relativa alla intermediazione finanziaria. In ordine alla procedura esplicativa del diritto di accesso endoprocedimentale, l'art. 41, comma 2, d. lgs. 7 marzo 2005, n. 82, allo scopo di agevolare la partecipazione ai processi decisionali pubblici, ne ha stabilito lo svolgimento in forma telematica. Sarebbe conveniente applicare tale innovazione anche nell'accesso extraprocedimentale, conseguendone indubbi benefici nei confronti dei privati e delle pubbliche amministrazioni.

Le ulteriori innovazioni in materia di documento informatico e di gestione elettronica dei dati – volendo parificare le tipologie di documenti amministrativi – pongono certamente difficili problemi che, tuttavia, andranno risolti caso per caso, agendo a tutela di diritti di riservatezza, in modo da non ledere basilari principi di rango costituzionale.

4. Diritto di accesso, enti pubblici e tutela della riservatezza

Il diritto di accesso azionato da chi ne ha interesse (cioè «la persona fisica, la persona giuridica, l'ente o l'associazione cui si riferiscono i dati personali» ai sensi dell'art. 4, comma 1, lett i) d.lgs n. 196/2003) su dati personali che lo interessano, viene di-

²¹ Si pensi all'art. 22 della legge 11 febbraio 1994, n. 109, "Legge Merloni".

²² In «Giornale di diritto amministrativo», 2001, p. 261, con nota di BERNARDO GIORGIO MATTARELLA, *Diritto di accesso e Consob: l'interpretazione della Corte costituzionale*.

sciplinato attraverso il titolo II del Codice della privacy. Segnatamente ai sensi dell'art. 7 del Codice spetta al soggetto interessato il «diritto di ottenere la conferma dell'esistenza o meno di dati personali che lo riguardano, anche se non ancora registrati» ed anche «la loro comunicazione in forma intelligibile». Alla luce dell'art. 4, comma 1, lett. *d* ed *e* del Codice, si è rilevato che nei dati personali sono inclusi i dati sensibili e giudiziari, da ritenersi questi, al pari dei dati personali. Sul tema risulta interessante la problematica secondo cui il diritto di accesso nei riguardi degli enti pubblici non debba sempre osservare le procedure stabilite dalla legge n. 241 del 1990. La questione emerge perché l'art 59 del Codice detta che, tranne quanto stabilito dall'art. 60 relativamente ai «dati supersensibili», «i presupposti, le modalità, i limiti per l'esercizio del diritto di accesso a documenti amministrativi contenenti dati personali, e la relativa tutela giurisdizionale, restano disciplinati dalle legge 7 agosto 1990, n. 241». Ed ancora, secondo l'art. 22, comma 4, l. n. 241/1990 non può sussistere l'esercizio del diritto di accesso nei riguardi delle «informazioni in possesso di una pubblica amministrazione» ove esse non «abbiano forma di documento amministrativo, salvo quanto previsto dal decreto legislativo 30 giugno 2003, n. 196». Tenendo in conto che in base all'art. 22, comma 1, lett. *d*, l. n. 241/1990 si interpreta come documento amministrativo «ogni rappresentazione grafica, fotocinematografica, elettromagnetica o di qualunque altra specie del contenuto di atti, anche interni o non relativi ad uno specifico procedimento, detenuti da una pubblica amministrazione e concernenti attività di pubblico interesse, indipendentemente dalla natura pubblicistica o privatistica della loro disciplina sostanziale»²³, si potrebbe dedurre che il diritto di accesso nei riguardi del-

²³ Indubbiamente, l'informatizzazione delle procedure crea alcuni interrogativi quali ad esempio l'archiviazione di documenti informatici su supporti magnetici o digitali accessibili mediante connessioni di reti, «che permet-

le pubbliche amministrazioni possa esplicarsi unicamente secondo i criteri ex capo V della legge n. 241/90, escludendo i criteri regolati dal Codice della privacy²⁴.

A ben vedere, in virtù della ampiezza della nozione offerta dal legislatore di documento amministrativo (stabilita – come detto – dall'art. 22 della legge sul procedimento che inserisce altresì quelli prodotti da privati ove usati nei processi decisionali pubblici: ad esempio gli elaborati dei candidati nei concorsi pubblici, ecc.), non risulta individuabile a quali tipologie di informazioni appartenenti alla pubblica amministrazione, che non rappresentino “documenti amministrativi”, il cittadino possa avere accesso²⁵. In base alla definizione, le informazioni rin-

tono di ipotizzare problemi riferiti non solo all'accesso c.d. “esterno” di cui alla l. n. 241/90, ma anche all'indiscriminato accesso “interno”, ossia effettuato da parte del personale della p.a. al di là degli scopi per i quali la connessione ad un determinato sistema – per esempio di archiviazione – è stata apprestata e garantita». Così FRANCESCO GAVERINI, *Accesso (e accessi) ai documenti amministrativi: linee-guida per le p.a. a garanzia del diritto alla privacy*, in «Foro amministrativo Tar», 2005, p. 1846, ss.

²⁴ La nozione di documento informatico è contenuta nell'art. 1, comma 1, lett. b) D.P.R. 445/2000 e poi ripresa dall'art. 1 comma 1 lett. p) d.l. 7 marzo 2005, n. 82, – Codice dell'amministrazione digitale) – nei termini seguenti «la rappresentazione informatica di atti, fatti, o dati giuridicamente rilevanti».

²⁵ In tale contesto si colloca la questione in parte disciplinata dal D.P.R. 445/2000. Infatti, l'art. 58 dispone che le operazioni di accesso ai sistemi informatici siano disciplinate da criteri di abilitazione in grado di limitare l'utilizzo di risorse informative ai soli soggetti autorizzati. L'art. 61 dispone che ogni amministrazione istituisca un servizio per la tenuta del protocollo informatico, della gestione dei flussi documentali e degli archivi in ciascuna delle grandi aree organizzative omogenee individuate ai sensi dell'art. 50; a tale servizio sono attribuite una serie di funzioni: tra queste l'attribuzione di livelli di autorizzazione «per l'accesso alle funzioni della procedura, distinguendo tra abilitazioni alla consultazione e abilitazioni all'inserimento e alla modifica delle informazioni».

tracciabili presso gli enti pubblici, che non presentino la caratteristica di documenti amministrativi, sono esclusivamente i pareri personali dei dipendenti e degli amministratori, e gli elementi conoscitivi di cui non sia stata eseguita la riproduzione (“registrazione”, terminologia prevista dall’art. 7 comma 1 del Codice della privacy), in atti non identificabili, pertanto, come componenti dell’azione conoscitiva dell’ente pubblico²⁶. Essi, tuttavia, vanno oltre l’area di applicazione del diritto di accesso, perché il privato esercita questo diritto per ottenere notizie sui dati informativi usati nel processo decisionale pubblico, definiti attraverso atti e documenti. E proprio l’art. 18 come 2, del Codice sulla privacy sembra escludere l’accesso a dati personali non inclusi in documenti amministrativi. Infatti, secondo tale norma, si specifica che «qualunque trattamento di dati personali da parte di soggetti pubblici è consentito soltanto per lo svolgimento delle funzioni istituzionali». Ciò comporta che quei dati vengano collocati su supporti (cartacei, fotografici, elettroma-

²⁶ Il diritto di accesso, secondo la giurisprudenza del Consiglio di Stato, Sezione VI, 10 febbraio 2006, in «Foro amministrativo Consiglio di Stato», 2006, p. 1270, “non si configura mai come un’azione popolare (fatta eccezione per il peculiare settore dell’accesso ambientale) ma postula sempre un accertamento concreto dell’esistenza di un interesse differenziato della parte che richiede i documenti. La titolarità (o la rappresentatività) degli interessi diffusi non giustifica un generalizzato e pluricomprendivo diritto alla conoscenza di tutti i documenti riferiti all’attività di un gestore del servizio e non collegati alla prestazione dei servizi all’utenza, ma solo al più limitato diritto alla conoscenza di atti, rivolti ai consumatori, che incidono in via diretta ed immediata, e non in via meramente ipotetica e riflessa, sugli interessi dei consumatori. Dagli art. 1 e 7 del d.lgs. n. 196 del 2003 (Codice della privacy), e dall’art. 2 d.lg n. 206 del 2005 (Codice del consumo) si desume che le associazioni di tutela dei consumatori possono agire a tutela della privacy indifferenziata dei consumatori, ma solo a tutela di un proprio specifico interesse alla privacy o a tutela della privacy di soggetti determinati da cui abbiano specifica procura scritta”.

gnatici, e così via) atti ad assicurarne la conservazione e la circolazione, esclusivamente nella struttura organizzativa dell'ente interessato. L'impossibilità di discernere, in riferimento agli enti pubblici, tra accesso ai dati non registrati e accesso ai dati registrati, e cioè ai documenti amministrativi, non consente di concludere che l'accesso debba esplicarsi soltanto in base alla disciplina della l. n. 241/90 mantenendosi in sostanza non applicabile in campo pubblico il Codice privacy ad eccezione delle ipotesi di dati personali che non presentino l'aspetto di documenti amministrativi. Oltre a particolari tipologie di diritto di accesso agli atti amministrativi riconosciute a soggetti identificati in base al ruolo istituzionale che ricoprono – il caso dei consiglieri comunali e provinciali²⁷ che possono accedere alle notizie oltre gli schemi rigidamente dettati dalla disciplina generale in tema²⁸, è noto che l'analisi deve tenere in considerazione l'evoluzione delle procedure informatiche, l'aumento della creazione di banche dati, l'interconnessione delle reti, che hanno aumentato la verificabilità dell'azione amministrativa interna della p.a. (in termini di trasparenza). E, tuttavia, il Codice della privacy ha sul punto individuato le figure «soggettive aventi rilevanza nei procedimenti di trattamenti dei dati personali, distinguendo il titolare, il responsabile, e gli incaricati»; figure queste

²⁷ Essa si denoterebbe come tipologia di accesso a metà strada tra interna ed esterna, con alcune differenze piuttosto evidenti. Cfr. in giurisprudenza Tar Abruzzo, Pescara, 16 dicembre 2004, n. 1100; RENATO CAMELI, *Considerazioni in tema di accesso agli atti delle amministrazioni locali*, in *Foro amministrativo*, «Consiglio di Stato», n. 10 del 2004, p. 3578.

²⁸ Si veda l'art. 43 d.lgs 18 agosto 2000, n. 267 che prevede «i consiglieri comunali e provinciali hanno diritto di ottenere dagli uffici, rispettivamente del Comune e della provincia, nonché delle loro aziende ed enti dipendenti, tutte le notizie e le informazioni in loro possesso, utili all'espletamento del proprio mandato. Essi sono tenuti al segreto nei casi specificamente determinati dalla legge».

che devono farsi carico di trattare i dati, con possibili lesioni del diritto alla riservatezza²⁹.

L'esame della normativa in vigore, infatti, sembrerebbe implicare una conclusione singolare: in altri termini, in virtù dell'impossibilità di accedere a dati personali che non siano inseriti in documenti amministrativi, la disciplina di riferimento si identifica con quella del Codice della privacy. Tranne che non si voglia ritenere l'inapplicabilità nei riguardi dei soggetti pubblici, assunto in evidente contrasto con le norme del Codice, secondo le quali è – in modo chiaro – riconosciuto all'interessato il diritto di acquisire dalle amministrazioni la comunicazione dei dati personali che lo interessano in forma intellegibile³⁰. Tra l'altro, sebbene il coordinamento tra le due discipline sia inadeguato, emergono altri elementi su quanto si sta argomentando, al di là del fatto che il d.lgs n. 196/2003 si rivolge in modo chiaro all'accesso nei riguardi delle pubbliche amministrazioni. Intanto si desume che il diritto di accesso nei riguardi degli enti pubblici è da esercitarsi con riferimento a documenti amministrativi, perché il capo II del d.lgs n. 196/2003 sancendo «regole ulteriori per i soggetti pubblici» contempla norme rivolte a regolare in modo specifico i criteri di “trattamento” dei dati in possesso degli enti pubblici. Il “trattamento” secondo l'art. 4, comma 1, lett. a, del Codice privacy, è qualificato come «qualunque operazione o complesso di operazioni, effettuati anche senza l'ausilio di strumenti elettronici, concernenti la raccolta, la registrazione, l'organizzazione, la conservazione, la consultazione, l'elaborazione, la modificazione, la selezione, l'estrazione, il raffronto, l'utilizzo, l'interconnessione, il blocco, la co-

²⁹ Sul tema, *amplius*, FRANCESCO GAVERINI, *Accesso (e accessi) ai documenti amministrativi*, cit., p. 1846-1847.

³⁰ Diffusamente, cfr. MASSIMO OCCHIENA, *I diritti di accesso dopo la riforma della l. n. 241 del 1990*, cit., p. 909.

municazione, la diffusione, la cancellazione e la distruzione di dati anche se non registrati in una banca dati»: criteri di impiego delle informazioni che esigono la loro inclusione in quegli atti che l'art. 22, l. n. 241/90, denota come "documenti amministrativi". Inoltre, la normativa in vigore non delimita l'esercizio dei diritti ex art. 7, ss del Codice, ove si utilizzzi nei riguardi di soggetti pubblici. In proposito, secondo l'art. 8 del d.lgs. n. 196/2003 tali diritti «sono esercitati con richiesta rivolta senza formalità al titolare o al responsabile», soggetti di cui fanno parte anche le pubbliche amministrazioni, per come le definizioni contenute nell'art. 4 comma 1 d.lgs cit., di "titolare" e di "responsabile" indicano in modo chiaro. Ciò ben si coniuga con la struttura generale del Codice della privacy: esso, al fine di tutelare la riservatezza nelle sue forme più svariate, in linea generale non differenzia i soggetti pubblici dai privati, ove "l'interessato" voglia conoscere i dati personali che gli concernono. Infine si può osservare che il legislatore ha operato in modo chiaro nel circoscrivere l'ambito di applicazione del Titolo II del Codice della privacy nei riguardi degli enti pubblici. Sul punto, la restrizione contemplata nell'art. 18, comma 3, d.lgs cit., esige una chiara "norma di legge o di regolamento" allorquando un "soggetto pubblico" debba avviare la "comunicazione"³¹ e la "diffusione" verso "privati o a enti pubblici economici" su "dati diversi da quelli sensibili e giudiziari". Quindi l'art. 18 comma 3 cit., comporta che l'accesso ai dati "personali", "sensibili" e "giudiziari" ad opera dell'interessato si uniformi alla normativa del Codice sulla privacy, potendo prescindere dal richiamo a disposizioni che si differenziano da quelle espresse nel d.lgs

³¹ In base all'art. 4 comma 1, lett. l del Codice, s'intende «dare conoscenza dei dati personali a uno o più soggetti determinati diversi dall'interessato» e la "diffusione" definita dallo stesso art. 4 comma 1 lett. m, come «il dare conoscenza dei dati personali a soggetti indeterminati».

196/2003. Non sembra dunque eseguibile una rigorosa interpretazione letterale sul combinato disposto dell'art. 59 del Codice e dell'art. 22, comma 4 della legge sul procedimento amministrativo. Se si ritenesse che il diritto dell'interessato alla comunicazione in forma comprensibile dei propri dati personali ex art. 7 del Codice privacy possa esercitarsi soltanto nei casi di accesso ai dati personali che non abbiano la natura di "documenti amministrativi", ne conseguirebbe l'assunto che il diritto di accesso ai dati personali ad opera di colui che ha titolo sugli stessi abbia luogo soltanto secondo le norme della legge sul procedimento, perché l'ampia nozione di "documento amministrativo" contiene qualsiasi informazione che sotto forma di atto sia posseduta da una pubblica amministrazione. Tra l'altro, ciò si discosta sia da altre disposizioni del Codice, sia dalla ratio su cui si fonda. Pertanto il diritto di accesso volto ad acquisire la comunicazione in forma comprensibile di dati personali (inclusi i dati sensibili e giudiziari) per il soggetto interessato si esplica secondo le norme del Codice stesso quando sono presenti in atti di un procedimento condotto a termine. Altrimenti tale tipologia di diritto di accesso non può esercitarsi, perché incompatibile con la sussistenza di un potere amministrativo³². Segnatamente, in virtù dell'art. 7 del Codice, l'interessato ha il diritto di acquisire non solo se esistano o meno dati personali che lo interessano (anche i dati non ancora registrati), ma anche – in special modo – la loro comunicazione in termini comprensibili. Ed ancora l'art. 7 prevede per l'interessato il diritto ad acquisire l'indicazione dell'origine dei dati personali, degli scopi e metodi del trattamento, della procedura utilizzata ove il trattamento sia eseguito mediante mezzi elettronici, degli elementi che individuano il titolare e i responsabili, nonché dei soggetti cui è possibile comunicare i dati personali. In riferimento alla procedura

³² Cfr. MASSIMO OCCHIENA, *I diritti di accesso*, cit., p. 910.

di accesso, gli articoli 8 e 9 del Codice stabiliscono che il soggetto interessato possa indirizzare all'ente pubblico una istanza cui deve corrispondere tempestivamente un adeguato riscontro. In virtù dell'art. 9 comma 5, l'istanza può essere rinnovata «salva l'esistenza di giustificati motivi, con intervallo non minore di novanta giorni»; inoltre essa può essere eseguita in forma orale («e in tal caso è annotata sinteticamente a cura dell'incaricato o del responsabile», per come detta l'art. 9 comma 1), o in forma scritta; in quest'ultimo caso si può utilizzare lettera raccomandata, telefax o posta elettronica. L'interessato, nell'esercitare l'accesso, può agire personalmente oppure delegare altro soggetto. I commi 2 e 3 dell'art. 10 del Codice prevedono una serie di modalità ai fini della conoscenza dei dati personali nei confronti del soggetto interessato. Infatti la richiesta dell'interessato può essere soddisfatta intanto attraverso comunicazione orale; oppure permettendogli la «visione mediante strumenti elettronici, sempre che in tali casi la comprensione dei dati sia agevole, considerata anche la qualità e la quantità delle informazioni»; o trasferendo i dati «su supporto cartaceo o informatico»; ed ancora comunicandogli «per via telematica»; infine attraverso «l'esibizione o la consegna in copia di atti e documenti contenenti i dati personali richiesti», ove il prelievo dei dati fosse «particolarmente difficoltoso». Inoltre il comma 3 dell'art. 10 stabilisce la regola secondo cui «salvo che la richiesta sia riferita ad un particolare trattamento o a specifici dati personali o categorie di dati personali, il riscontro all'interessato comprende tutti i dati personali che riguardano l'interessato comunque trattati dal titolare». Sembra evidente, tenendo in conto la quantità di dati personali in possesso degli enti pubblici, al fine di impedire la stasi per l'attività e l'organizzazione amministrative, l'inapplicabilità di tale norma relativamente all'accesso eseguito nei riguardi di una pubblica amministrazione. Peraltro ciò trova riscontro nell'art. 24, comma 3, l. n. 241 del 1990 che non ammette «istanza di accesso preordinate ad un controllo generaliz-

zato dell'operato delle pubbliche amministrazioni»³³; ed anche nel combinato disposto dell'art. 22 comma 1, lett. *b* – che richiede la titolarità di un «interesse diretto, concreto e attuale, corrispondente ad una situazione giuridicamente tutelata e collegata al documento al quale è chiesto l'accesso» –, e dell'art. 25, comma 2, 1. cit. È previsto invero che deve sussistere la motivazione della richiesta, allo scopo di mettere in risalto la posizione giuridica che legittima l'accesso su qualche documento, e non su tutti quelli che detiene l'ente pubblico verso cui sia indirizzata la richiesta di esibizione. Con riferimento alla tutela giurisdizionale³⁴, l'art. 145, comma 1 del Codice privacy prevede che il diritto in questione possa essere azionato «dinanzi all'autorità giudiziaria ordinaria o con ricorso al Garante». Inoltre, secondo l'art. 150, competono al Garante poteri di cognizione e di condanna, che arrivano anche sino all'ottemperanza. Sotto l'aspetto giurisdizionale, definita l'area di competenza del giudice ordinario, in virtù dell'art. 152 del Codice gli viene conferito intanto il potere di accettare o respingere l'istanza integralmente o parzialmente; in secondo luogo il potere di disporre i criteri necessari, infine il potere di emettere la condanna nella ipotesi di risarcimento dei danni. Tale disciplina è palese per quanto concerne la natura di diritto soggettivo dell'accesso teso alla acquisizione della comunicazione dei propri dati personali in mo-

³³ Con riferimento, altresì, ai soggetti di diritto privato, su cui in dottrina VINCENZO CERULLI IRELLI, *Osservazioni generali sulla legge di modifica della l. n. 241/90*, in *Giustizia amministrativa*, <www.giustamm.it>; MARIA GRAZIA DELLA SCALA, *Il diritto di accesso all'attività privatistica dell'amministrazione*, cit.; FRANCESCO DE LEONARDIS, *Il concetto di organo indiretto: verso nuove ipotesi di applicazione dell'esercizio privato di funzioni pubbliche*, in «Diritto amministrativo», 1995, p. 347.

³⁴ Su cui, sia consentito – anche per i profili processuali – il rinvio a RENATO ROLLI, ALBINO SACCOMANNO, *Difesa civica e protezione dei diritti fondamentali*, in «Le Corti calabresi», n. 2 del 2005, p. 521, ss.

do comprensibile. Ed invero, stante il suo esercizio nei riguardi sia dei soggetti privati sia di enti pubblici, ed al riparo da limitazioni derivanti dall'esercizio di un potere amministrativo, tale tipologia di accesso si caratterizza libera in questa ultima ipotesi, per cui spetta alle amministrazioni l'osservanza in materia dei comportamenti tipici di tutti gli altri soggetti dell'ordinamento³⁵.

Diversa questione si pone in relazione alla richiesta di accesso extraprocedimentale di dati il cui titolare è soggetto diverso dal richiedente.

In virtù dell'art. 10, comma 5, del Codice sulla privacy si specifica che «il diritto di ottenere la comunicazione in forma intellegibile dei dati non riguarda dati personali relativi a terzi salvo che la scomposizione dei dati trattati o la privazione di alcuni elementi renda incomprensibili i dati personali relativi all'interessato». Questa disposizione non consente in modo chiaro la possibilità di usare, in tali situazioni, nei riguardi degli enti pubblici il diritto di accesso che il titolo II del d.lgs. n. 196/2003 disciplina. Sul punto, tenendo in conto che quei dati sono inclusi in documenti amministrativi e che «tutti i documenti amministrativi sono accessibili» (ex art. 22, comma 3, l. n. 241/90), tranne i casi evidenziati nell'art. 24, si può dedurre che l'accesso extraprocedimentale al fine della conoscenza di documenti comprendenti dati personali di soggetti distinti da chi ne faccia istanza venga azionato in base alla normativa stabilita attraverso il capo V della legge sul procedimento. Il diritto di accesso viene regolamentato dalla l. n. 241/90 con modalità diverse da quelle contemplate nel Codice della privacy. Al riguardo, circa la legittimazione attiva, ridimensiona la serie dei soggetti che godono del diritto di accesso in confronto alle disposizioni del Codice della privacy sull'accesso da parte del privato ai dati personali che lo

³⁵ Cfr. MASSIMO OCCHIENA, *I diritti di accesso*, cit., p. 911-912.

interessano. L'art. 22 comma 1, lett. b, sulla linea delle statuizioni previste dall'art. 2, comma 1, del d. P.R. 27 giugno 1992, n. 352, individua come «interessati [...] tutti i soggetti privati, compresi quelli portatori di interessi pubblici o diffusi, che abbiano un interesse diretto, concreto e attuale, corrispondente ad una situazione giuridicamente tutelata e collegata al documento al quale è chiesto l'accesso». In ordine all'oggetto, va evidenziato che dal concetto di "dati personali" sono esclusi quelli sensibili e giudiziari, che sono disciplinati da una precisa normativa. Va anche sottolineato che l'art. 24, comma 1, l. 241/90 non contempla nell'accesso alcune tipologie di documenti, indicati alle lettere *a*, *b*, e *c* (e non *d*), che ogni pubblica amministrazione ha l'obbligo di identificare in modo reale, in considerazione degli atti che usualmente vengono usati nei processi decisionali o custoditi nei loro archivi³⁶.

Per quanto i documenti appartenenti a tale elenco non riguardino in modo diretto la riservatezza di terzi – in quanto non contemplano obbligatoriamente dati personali su soggetti distinti da chi ne faccia istanza – sono tuttavia esclusi dall'accesso; o, al massimo, la loro accessibilità è realizzabile soltanto nel caso in cui venga ammessa in modo chiaro dalle leggi di settore ed in base ai criteri in esse stabiliti. In entrambe le situazioni risulta che non può esercitarsi il diritto di accesso così come regolamentano gli artt. 7 ss. contenuti nel Codice privacy. Con riferimento alla tutela della riservatezza, in virtù del nuovo testo dell'art. 24, comma 6, lett. *d*, della l. n. 241/90, è affidata al Governo la determinazione attraverso specifico regolamento di delegificazione (ex art. 17, comma 2, della legge 23 agosto 1988, n. 400) dell'esclusione dell'accesso per documenti amministrativi ove «ri-

³⁶ Sebbene la legge non si esprima sul punto, è presumibile che questa identificazione si attui per lo più attraverso regolamento, affidando i casi particolari ad atti amministrativi puntuali.

guardino la vita privata o la riservatezza di persone fisiche, persone giuridiche, gruppi, imprese e associazioni, con particolare riferimento agli interessi epistolare, sanitario, professionale, finanziario, industriale e commerciale di cui siano in concreto titolari, ancorché i relativi dati siano forniti all'amministrazione dagli stessi soggetti cui si riferiscono». Non può classificarsi, d'altra parte, come una sottrazione totale, dal momento che il comma 7 specifica che «deve comunque essere garantito ai richiedenti l'accesso ai documenti amministrativi la cui conoscenza sia necessaria per curare o per difendere i propri interessi giuridici». Naturalmente è competenza dell'ente pubblico – ricevente l'istanza di esibizione dei documenti – giudicare sulla funzionalità dell'accesso per la tutela degli interessi giuridici del richiedente. Da qui, sembra chiaro che l'esistenza di queste disposizioni comportino che l'istante sia soggetto al potere; nell'ipotesi di accesso relativo a dati personali di terzi, appare verosimile che l'attuazione dell'interesse (ad accedere) del richiedente sia sottoposta all'esplicazione di un potere, che la legge non gli assicura in modo definitivo. In effetti il potere si manifesta secondo svariate modalità e su piani diversi. Intanto è evidenziabile il potere regolamentare, che spetta ad ogni amministrazione per determinare le tipologie di documenti che essa custodisce o di cui può disporre: documenti esclusi dall'accesso perché inclusi nei casi previsti dall'ex art. 24, comma 1, lett. da *a* fino a *c*. In secondo luogo, è da sottolineare anche il potere di distinguere l'eventualità per cui un atto – che non sia di norma oggetto di “trattamento istituzionale” – sia incluso o meno in quella tipologia, qualora l'amministrazione si trovi “per caso” nella condizione di possederlo. In effetti, sebbene i regolamenti siano particolareggiati ed espressi con attenzione, non saranno in grado di presumere – teoricamente – la tipologia della totalità dei documenti che l'ente pubblico potrà detenere e che potrebbero contenere dati sulla riservatezza. In queste ipotesi l'ente pubblico, davanti ad una istanza di accesso, dovrà determinare tangibilmente

se quel tipo di documento faccia parte di quelli esclusi dall'accesso; aziona così il potere di decidere la norma di prevalenza o meno intercorrente tra accesso e *privacy* attribuibile a quella unica tipologia. Vanno evidenziati, ancora, i poteri stabiliti nell'art. 25, comma 3, l. n. 241/90, con i quali si rinvia, si circoscrive o si nega con motivazione l'accesso nei casi che l'art. 24, l. n. 241/90 contempla.

Un ulteriore potere è teso a determinare da un lato il limite tra la tutela della riservatezza su dati personali inclusi nei documenti e la certezza di accesso a tali dati per l'istante allo scopo di tutelare gli interessi giuridici di cui è detentore, per converso. Non essendo possibile la previsione di tutti i casi di risoluzione delle questioni tra accesso e riservatezza in termini generali e tecnici, la legge affida alle amministrazioni stabilire – nel caso concreto – se il dato personale riservato sia da rendere noto o meno all'istante sulla base dell'esigenza di curare gli interessi di cui è detentore e di mettere in atto forme di tutela. Dal momento che la soluzione assunta dall'ente pubblico, detentore del documento sul dato personale riservato, determina la soddisfazione di uno dei due interessi (la *privacy* oppure l'accesso), sembra palese che l'esercizio del potere amministrativo dia luogo ad una serie di vicende giuridiche. E, sebbene questo potere sembri mostrare una connotazione “poco usuale”, con funzione arbitrale³⁷, gli effetti giuridici che caratterizzano la difesa della riservatezza o la garanzia della trasparenza amministrativa sono generati proprio dall'esercizio di questo potere. Al contrario, l'assunto secondo cui nelle ipotesi di contrasto tra accesso e *privacy* siano totalmente assenti aspetti volti all'interesse pubblico, sembra contrapporsi con la normativa in vigore. In realtà, il nuovo testo dell'art. 22, comma 2 l. n. 241/90 specifica che «l'accesso ai documenti ammini-

³⁷ Sul punto, FABRIZIO FRACCHIA, *Riti speciali a rilevanza endoprocedimentale*, Torino, 2003, p. 30.

strativi, attese le sue rilevanti finalità di pubblico interesse, costituisce principio generale dell'attività amministrativa al fine di favorire la partecipazione e di assicurarne l'imparzialità e la trasparenza»³⁸. Da ultimo, tenendo in conto che l'art. 24, comma 7, 1. cit., dispone che "deve" essere garantito l'accesso ove tenda a curare o difendere un interesse giuridico dell'istante, non sembra neanche plausibile asserire l'insussistenza del potere amministrativo. In altri termini, non può ritenersi che la pubblica amministrazione sia carente di un potere di valutazione dal momento che, secondo la decisione del Consiglio di Stato, Sez. IV, del 4 febbraio 1997, n. 82³⁹ «la strumentalità dei documenti richiesti alle esigenze di difesa dell'interessato non può essere che un controllo estrinseco, in quanto il diritto di difesa nel suo concreto esercizio rientra nella sfera di autonomia della persona».

Invero, in tale situazione, nel confronto tra riservatezza ed accesso, il legislatore ha preferito considerare utilizzabile il diritto di accesso, in cui l'ente pubblico può azionare il giudizio di necessità al fine della tutela e cura di interessi giuridici; attività questa che, di *contra*, non può svolgere nei riguardi del diritto alla riservatezza. Tra l'altro, i commi 6 e 7 dell'art. 24, cit., statuiscono di norma l'esclusione dell'accesso, salvo che non sia con-

³⁸ Sul punto, MARIA ALESSANDRA SANDULLI, *Accesso alle notizie e ai documenti amministrativi*, cit.; SILVIO GAMBINO, *Il diritto di accesso. La nuova disciplina legislativa (l. n. 15/2005) alla luce della revisione costituzionale del Titolo V*, cit., osserva che con la nuova «scelta operata dal legislatore di riforma la trasparenza e, più in generale il principio di pubblicità dell'azione amministrativa cessano di essere finalità strettamente connesse alla garanzia dello svolgimento imparziale dell'attività amministrativa [...], per essere elevati, con l'art. 1.1. della legge n. 15/2005, a principio generale dell'attività amministrativa», vincolante – oltre che per le pubbliche amministrazioni e gli enti pubblici nazionali – per le regioni e gli enti autonomi della repubblica, unitamente al principio di legalità e “ai principi dell'ordinamento comunitario”.

³⁹ In «Diritto processuale amministrativo», 1998, p. 377.

cretamente – dopo il giudizio dell’ente pubblico – rispondente alla cura e difesa di interessi giuridici del richiedente. Tra l’altro, l’amministrazione può sempre nascondere i dati personali oppure utilizzare espedienti simili allo scopo di escludere dei documenti su dati riservati che non influiscono in sé sul diritto di difesa. Senza dubbio l’ente pubblico potrebbe porre dei limiti illegittimamente all’accesso, al fine di indebolire le risorse di difesa del privato che vuole garantirsi nei confronti dello stesso ente. Tuttavia, davanti ad un tale comportamento dell’amministrazione, il privato può in ogni caso richiedere la tutela prevista dall’art. 25 della legge n. 241/90. Infatti, in previsione di un accertamento regolare, l’ente pubblico ha l’obbligo di assicurare il diritto di accesso, quale istituto applicativo del principio di trasparenza amministrativa che l’art. 2, comma 2, della legge sul procedimento amministrativo riconosce in modo chiaro.

5. La natura del diritto di accesso ai documenti

Il D.P.R. n. 184 del 2006, giova rilevare, ha centrato l’attenzione sul responsabile del procedimento di accesso, soggetto che non coincide necessariamente con il responsabile dei procedimenti amministrativi, ma deve essere individuato nel dirigente, nel funzionario preposto all’unità organizzativa o in altro dipendente addetto all’unità competente a formare il documento, o a detenerlo stabilmente. Questa distinzione non è priva di carica innovativa. Invero l’attenzione del legislatore si è rivolta verso una linea di semplificazione tra soggetti interessati e amministrazione, attraverso la predisposizione, per i primi, di una serie di modalità per l’esercizio del diritto, strutturate secondo al quadro di articolazione organizzativa dell’ente. Del resto, l’inosservanza di alcune disposizioni del D.P.R. n. 184/06 determinano l’illegittimità degli atti adottati a seguito del procedimento per l’esercizio del diritto di accesso. In via generale la soluzione le-

gislativa sembra andare nell'ottica della facilitazione dell'esercizio del diritto di accesso, come ad esempio quelle finalizzate a garantirlo anche per via telematica (art. 13). Il rinvio ad altro atto normativo sulla definizione dei casi di esclusione dell'accesso (art. 10 D.P.R. n. 184/06), apre la strada, inoltre, ad un intervento mirato a dettagliare le fattispecie di riferimento in coerenza con quanto stabilito dall'art. 24, commi 1 e 6. Sembra verosimile che il legislatore abbia preferito configurare l'accesso come diritto soggettivo perfetto (non risultando di conseguenza necessario l'obbligo della notifica del ricorso ad almeno uno dei controinteressati). E ciò potrebbe essere dedotto da due passaggi della normativa⁴⁰. Invero, secondo l'art. 22, c.1 lett. a) l. *proc.amm.* il diritto di accesso deve ritenersi «il diritto degli interessati di prendere visione e di estrarre copia di documenti amministrativi». E, in virtù dell'art. 22, c. 2 l. cit., l'accesso è definito come un «principio generale dell'attività amministrativa al fine di favorire la partecipazione e di assicurarne l'imparzialità e la trasparenza, ed attiene ai livelli essenziali delle prestazioni concernenti i diritti civili e sociali che devono essere garantiti su tutto il territorio nazionale»⁴¹. Tra l'altro, se il legislatore sembra aver

⁴⁰ La qualificazione del diritto di accesso quale «diritto soggettivo ad un'informazione qualificata», in dottrina, MARCELLO CLARICH, *Diritto d'accesso e tutela della riservatezza: regole sostanziali e tutela processuale*, in «Diritto processuale amministrativo», 1996, p. 430; LORENZO IEVA, *Diritto di accesso ai documenti amministrativi e tutela della riservatezza della persona umana*, in «Foro amministrativo», 2001, p. 2643; MASSIMO OCCHIENA, *Diritto di accesso, atti di diritto privato e tutela della riservatezza dopo la legge sulla privacy (il diritto di accesso e riforma dell'organizzazione della p.a.)*, cit.; PELINO SANTORO, *Diritto di accesso partecipativo e diritto di accesso conoscitivo: posizioni soggettive ed effettività*, in «Foro amministrativo», 1992, p. 1779.

⁴¹ Sulla qualificazione in termini di diritto soggettivo, in giurisprudenza, cfr. Cons. Stato, Sez. VI, ord. 7 giugno 2005, n. 2954 e ord. 9 settembre 2005, n. 4686.

privilegiato la natura di diritto soggettivo, gli aspetti normativi maggiormente ostativi a tale qualificazione non sono stati modificati in modo opportuno. Intanto, nell'art. 25, l. n. 241/90 non è stato specificato se il giudizio riguardante il ricorso contro il diniego di accesso faccia parte della giurisdizione generale di legittimità oppure si svolga nella sede della giurisdizione esclusiva del giudice amministrativo. Inoltre, dal momento che soltanto precise necessità di riservatezza possono offrire un limite all'accesso, va evidenziato, di *contra*, che il legislatore sembra aver rafforzato gli ambiti di remissione della decisione concernente il bilanciamento tra i due interessi alla potestà discrezionale della pubblica amministrazione⁴². Da ultimo, le pubbliche amministrazioni hanno inserito nuovi fattori di valutazione relativi al diritto di accesso dell'istante. L'amministrazione controlla sia che l'interesse sia diretto, concreto ad attuale, sia che la situazione giuridica tutelata si connetta al documento che è stato richiesto. L'art. 22, c.2 della l. *proc. amm.*, tuttavia suscita altre osservazioni. Intanto è necessario sottolineare che il diritto di accesso rappresenta il principale mezzo di attuazione di due principi generali dell'attività amministrativa, che la nuova enunciazione dell'art. 1, comma 1 l. *proc. amm.* cita in modo chiaro: segnatamente i principi di pubblicità e trasparenza⁴³. Inoltre non

⁴² In questi termini, ALDO SANDULLI, *L'accesso ai documenti amministrativi*, cit., p. 495.

⁴³ Difatti, il diritto di accedere agli atti dell'amministrazione comunale «è legato non tanto alla residenza del soggetto richiedente, quanto al collegamento fra l'interesse sotteso alla richiesta d'accesso e l'attività del Comune, con la conseguenza che un imprenditore avente sede fuori dal territorio comunale ha diritto all'esibizione degli atti riguardanti un'attività posta in essere nel Comune da un'impresa concorrente, in quanto la detta ostensione consente di valutare l'eventuale esistenza di irregolarità che possano tradursi in pregiudizi alla concorrenza»: Tar Lombardia, sent. 11 aprile 2005, in «Foro amministrativo Tar», 2005, p. 1846. Cfr. altresì Tar Sicilia, Sez. Catania, 15 gennaio 2004, n. 28.

sembra calzante richiamare il principio di partecipazione, perché l'autonomia del diritto di accesso, sempre sostenuta dalla giurisprudenza, potrebbe essere compromessa. Infatti, l'art. 10, c. 1 lett. a) disciplina l'accesso partecipativo che deve essere separato dall'accesso conoscitivo. E ciò al fine di eludere, per effetto della commistione tra le due tipologie, il diniego della richiesta di accesso, ove la conoscenza del documento non possa determinare per il richiedente benefici partecipativi. Infine, la collocazione dell'accesso ai documenti nei livelli essenziali (sul piano costituzionale) delle prestazioni, potenzia in modo conveniente la concezione dell'azione amministrativa quale contributo offerto al cittadino, nell'osservanza di criteri e metodologie da assicurare su tutto il territorio nazionale.

Con riferimento alle situazioni giuridiche soggettive, quindi, poiché l'accesso extraprocedimentale teso alla conoscenza di documenti su dati personali appartenenti ad un soggetto diverso esige l'esplicazione di un potere amministrativo, si può affermare che il cittadino sia detentore di un interesse legittimo. In proposito, sembra sostenibile che il novellato comma 2 dell'art. 22, l. *proc. amm.*, vieti di delineare il diritto di accesso come interesse legittimo, perché contempla che l'accesso «attiene ai livelli essenziali delle prestazioni concernenti i diritti civili e sociali che devono essere garantiti su tutto il territorio nazionale ai sensi dell'art. 117, secondo comma, lett. m), della Costituzione»⁴⁴.

⁴⁴ SILVIO GAMBINO, *Il diritto di accesso*, cit., p. 439, osserva che «anche gli enti locali, nell'ambito della relativa competenza, vedono condizionata la potestà statutaria e regolamentare (loro riconosciuta dall'art. 114, II co.e dall'art. 117 VI co., Cost.) dall'obbligo del "rispetto del sistema costituzionale e delle garanzie del cittadino nei riguardi dell'azione amministrativa, così come definite dai principi stabiliti dalla presente legge", laddove per tali principi, in quanto ora considerati principi di rango costituzionale, – sia pure la formulazione appaia del tutto involuta – non pare potersi fare riferimento ai soli principi desumibili dalla medesima legge di riforma (co-

Orbene, sebbene si discuta vivacemente in dottrina e in giurisprudenza circa la natura del diritto di accesso (stante il fatto che la giurisprudenza amministrativa non ha sposato una interpretazione univoca, quasi rinunciando alla propria funzione di fornire al giurista salde proposte per l'applicazione), non sembra che la norma apporti cambiamenti alla soluzione fornita dalla decisione dell'Adunanza Plenaria che ha stabilito il carattere di interesse legittimo. Va anche evidenziato che il riferimento all'art. 117, 2° comma lettera *m*) della Costituzione sembra finalizzato non a sostenere il carattere di diritto soggettivo dell'accesso, ma a garantire l'applicazione delle disposizioni del capo V della legge n. 241/90 da parte di ogni amministrazione (e dei soggetti «di diritto privato limitatamente alla loro attività di pubblico interesse disciplinata dal diritto nazionale o comunitario», ai sensi dell'art. 22, comma 1, lett. *e*) agente sul territorio nazionale, evitando in tal modo il rischio di cambiamenti da parte di leggi regionali e di regolamenti di enti locali, che hanno la possibilità di «garantire livelli ulteriori di tutela». La norma, piuttosto che definire la natura del diritto di accesso, sembra evitare per quest'ultimo l'applicazione dell'art. 29 della *l. proc. amm.*, e, quindi, della normativa regionale e degli enti locali, al fine di regolarlo per converso con la normativa legislativa e regolamentare che pone lo Stato. Non si può neanche eccepire che il carattere di diritto soggettivo consegua dal motivo che l'*actio ad exhibendum* è stata affidata alla giurisdizione esclusiva amministrativa dopo l'introduzione nell'art. 25, comma 5, l. n. 241/90 della precisazione per cui «le controversie relative all'accesso ai documenti amministrativi sono attribuite alla giurisdizione esclusiva del giudice amministrativo»; e ciò da parte dell'art. 3, comma 6-*decies*, l. 14

me poteva affermarsi per il prevedente testo della legge n. 241/90), ma alla stessa disciplina di dettaglio accolta nella legge di riforma (l. n. 15/2005)».

maggio 2005, n. 80. Sembra in astratto erroneo che questa modificazione possa determinare la tacita ammissione del carattere di diritto soggettivo dell'accesso, dal momento che la qualificazione della situazione giuridico-soggettiva non proviene dalle procedure che l'ordinamento stabilisce allo scopo di garantirla processualmente. Ed invero, l'attività giurisdizionale è deputata a realizzare il diritto sostanziale o in modo subordinato, nel caso che sia oggetto di violazione, o in modo sostitutivo ove esso non sia concretizzato. In più, ritenendo che l'accesso sia un diritto, la giurisdizione sarebbe esclusiva e l'assenza di chiamata in giudizio del terzo implicherebbe ai sensi dell'ex art. 102 c.p.c. il dovere dell'integrazione del contraddittorio *iussu iudicis*. E ciò sembra opposto all'ammissione, azionata dall'art. 22, comma 1, lett. c, l. 241/90, della sussistenza di controinteressati, cioè quelli «individuati o facilmente individuabili in base alla natura del documento richiesto, che dall'esercizio dell'accesso vedrebbero compromesso il loro diritto alla riservatezza».

L'art. 22, comma 1, lett. c, cit., al di là dell'evidenziare l'esigenza che il terzo detentore di dati personali, di cui l'istante richiede l'accesso, partecipi (attraverso la comunicazione di avvio ex art. 7 l. n. 241/90) al conseguente procedimento amministrativo, riproduce l'espressione caratterizzante il processo di impugnazione, che stabilisce l'onere di notificazione del ricorso ai controinteressati, il cui inadempimento produce l'inammissibilità (art. 21 l. 6 dicembre 1971, n. 1034). Per contro, sebbene esistano fattori atti a sostenerne la qualificazione come processo di impugnazione caratteristico della giurisdizione di legittimità a garanzia degli interessi legittimi, sembra più opportuno rilevare il carattere particolare del giudizio sul diritto di accesso, per come gli aspetti tipici del rito incidentale inserito dall'art. 1 comma 1, l. 21 luglio 2000, n. 205, e attualmente riproposto all'interno dell'art. 25, comma 5, l. n. 241/90 indicano in modo chiaro. Da ultimo, poiché non sembra sostenibile sotto l'aspetto sostanziale che sia inadeguato il potere amministrativo, qualora un

privato eserciti il diritto di accesso, l'ammissione della giurisdizione esclusiva in tale ambito sembra giustificata dal fatto che al giudice amministrativo, nel relativo processo, sovente compete la conoscenza sia della situazione giuridica sostanziale dell'istante, sia di quelle dei terzi detentori di dati personali, implicati nell'istanza di esibizione del documento.

La difficoltosa correlazione tra accesso e riservatezza (ed i relativi intrecci e rimandi legislativi) assume centralità nelle questioni di qualificazione circa il carattere della situazione giuridica del diritto di accesso e dei conseguenti profili processuali. Orbene, fino a quando la l. legge *proc. amm.* ha regolamentato il contrasto tra accesso e *privacy* (si pensi all'iniziale art. 24, comma 2, lett. *d*) – con cui veniva assicurato l'esame degli atti su dati riservati, che fosse indispensabile al fine di vigilare o tutelare interessi giuridici dei richiedenti – non sussisteva alcun veto a ritenere l'accesso quale diritto soggettivo, dal momento che non competeva all'amministrazione nessun potere davanti all'istanza di esaminare documenti da essa detenuti. Eppure, allorquando la l. n. 675/1996 ha vietato l'accessibilità, in via generale, dei cosiddetti “dati sensibili” e la susseguente regolamentazione (tra cui emerge il d.lgs 11 maggio 1999, n. 135 sul trattamento dei dati sensibili ad opera dei soggetti pubblici) ha attribuito una funzione agli enti pubblici circa l'individuazione della concreta dimensione del diritto di accesso, la qualificazione di esso come diritto soggettivo ha messo in luce aspetti limitativi. In considerazione dell'esigenza di tutelare la *privacy* sui dati sensibili e della difficoltà del legislatore nell'individuare tutte le forme risolutive dei contrasti tra riservatezza e accesso ai documenti, l'unica soluzione era affidare all'amministrazione il potere di scelta sull'interesse che dovesse prevalere. Donde la finalità che ha suggerito la sentenza n. 16/1999 dell'Adunanza Plenaria del Consiglio di Stato: di fronte alla riservatezza, qualificata abitualmente come diritto soggettivo e quindi regolata attraverso la l. 675/1996, il Giudice amministrativo ha potuto soltanto operare

sul concetto del diritto di accesso, idoneo a rapportarsi con l'azione di un potere amministrativo.

I profili problematici esaminati dall'Adunanza Plenaria non hanno ricevuto vantaggio dalla nuova normativa sul diritto di accesso e sulla tutela della riservatezza contenuta nel nuovo testo del Capo V della l. n. 241/90 e nel Codice della privacy. Sembra sufficiente il riferimento, al di là dell'accesso ai dati personali di individui distinti da colui che chiede l'accesso, all'art. 60 del Codice che, davanti ai dati riservati concernenti l'ambito più privato della persona, demanda agli enti pubblici il giudizio su quando l'accesso debba tuttavia essere accordato, in quanto rapportato a situazioni giuridiche che si innalzano al livello di diritti della personalità o ad altre libertà costituzionali fondamentali e inviolabili. La caratterizzazione della giurisdizione del giudice amministrativo come esclusiva, che il legislatore ha tracciato, da ultimo, attraverso la l. n. 80/2005 sembra in teoria ammissibile, perché nella controversia in materia di accesso il giudice amministrativo ha il compito di conoscere casi reali in cui, nell'occasione della legittimazione, si intersecano diritti soggettivi e interessi legittimi. Tale circostanza, secondo autorevole dottrina, rappresenta l'unico motivo che si armonizza con i precetti costituzionali, per cui è attribuibile la giurisdizione esclusiva al giudice amministrativo⁴⁵. Infine, sorge l'interrogativo se, a fronte di altri motivi, venga assegnata la giurisdizione esclusiva al giudice amministrativo. Orbene, il conferimento di tale giurisdizione potrebbe trovare spiegazione alla luce del progetto generale con cui sembra che il legislatore negli ultimi tempi (situazione comunque che la Corte costituzionale non condivide) si allinei nel distinguere le giurisdizioni secondo i cosiddetti blocchi di materie, piuttosto che in funzione del-

⁴⁵ FRANCO GAETANO SCOCA, *Sopravvivrà la giurisdizione esclusiva?*, in «Giurisprudenza costituzionale», 2004, p. 2209, ss.

la natura delle situazioni giuridiche soggettive, su cui viene richiesta garanzia processuale. E ciò comprova, anche sotto tale aspetto, l'errore nell'intenzione di riconoscere il carattere della situazione giuridica soggettiva dalla tipologia di garanzia giurisdizionale apprestata dalla legge.

6. L'ambito oggettivo e soggettivo di applicazione del diritto di accesso

In ordine all'area soggettiva di applicazione dell'accesso ai documenti amministrativi, appare pacifica l'estensione della serie di soggetti, pubblici e privati tenuti ad osservare le norme sull'accesso⁴⁶. Segnatamente, il legislatore ha utilizzato un concetto ampio di pubblica amministrazione, che include l'organismo di diritto pubblico, secondo l'orientamento della disciplina europea: anche a tal proposito è stato condiviso l'orientamento prevalente della giurisprudenza amministrativa⁴⁷. E, in riferimento all'area oggettiva dell'accesso, l'art. 22, c.1 lett. d), ha innovato il concetto di documento amministrativo («atti detenuti da una pubblica amministrazione e concernenti attività di pubblico interesse, indipendentemente dalla natura pubblicistica o privatistica della loro disciplina sostanziale»). Inoltre sembra palese che il legislatore si sia orientato verso una elaborazione atta ad accogliere consolidati indirizzi giurisprudenziali sanciti dall'Adunan-

⁴⁶ Va rilevato che già l'art. 2, comma 1 D.P.R. n. 352 del 1992, disponeva: «Il diritto di accesso ai documenti amministrativi è esercitato nei confronti di tutte le pubbliche amministrazioni e dei concessionari di pubblici servizi da chiunque vi abbia un interesse personale e concreto per la tutela di situazioni giuridicamente rilevanti». Il nuovo regolamento sull'accesso (D.P.R. 12 aprile 2006, n. 184), ha abrogato il D.P. R. n. 352, salvo che per l'art. 8.

⁴⁷ Ad es. TAR Lombardia, Sez. III, 3 marzo 1998, n. 440, in «Urbanistica e appalti», 1998, p. 746.

za Plenaria del Consiglio di Stato, del 22 aprile 1999, n. 4 e 5⁴⁸ (in relazione sull'accesso all'azione di diritto privato della pubblica amministrazione)⁴⁹.

7. Diniego di accesso e forme di tutela

I modelli di tutela che presentano cambiamenti rilevanti concernono la tutela non giurisdizionale. Sul punto, è possibile ricorrere contro il veto dell'accesso delle amministrazioni centrali e periferiche dello Stato, alla Commissione per l'accesso ai documenti amministrativi (C.A.D.A.). Il procedimento presso la Commissione è simile a quello che si attua con riferimento al veto espresso da amministrazioni comunali, provinciali e regionali, presso il difensore civico competente per territorio. È previsto che la CADA richieda il giudizio del Garante per la tutela dei dati personali, ove il veto riguardi dati personali. Sembra opinabile l'aver posto, nell'art. 25 c.4. della l. n. 241/90, una norma di completamento del Codice della privacy. Tale norma prevede che, ove un procedimento sul trattamento di dati personali riguardi l'accesso ai documenti, il Garante abbia l'obbligo di richiedere il giudizio necessario e non vincolante della CADA. In ordine alla tutela giurisdizionale, si possono soltanto evidenziare due integrazioni nel rito speciale delineato dall'art. 25 l. n. 241/90, infatti sono state incluse nei commi 5 e 5-bis, due disposizioni precedentemente prodotte con la l. n. 205/2000. La prima riguarda l'impugnazione in materia di accesso all'interno di un

⁴⁸ In «Giornale di diritto amministrativo», 1999, p. 946.

⁴⁹ In tema MARIA GRAZIA DELLA SCALA, *Il diritto di accesso all'attività privatistica dell'amministrazione e dei soggetti privati alla luce della l. n. 15 del 2005. Condanna all'esibizione degli atti e ottemperanza al giudizio amministrativo*, in «Diritto amministrativo», 1996, p. 173.

ricorso principale già instaurato; la seconda contempla che il ricorrente possa stare in giudizio personalmente, senza l'assistenza del difensore (e che un dirigente, munito di autorizzazione del rappresentante legale dell'Ente, rappresenti e difenda l'amministrazione). I cambiamenti apportati alla struttura ed al funzionamento della Commissione per l'accesso suscitano alcune riflessioni. Intanto va osservato che, con riferimento alla nuova struttura, è aumentata la presenza politica all'interno della Commissione (cinque membri su tredici, in luogo di cinque su diciassette), ed è invece divenuto più esiguo il numero di membri (passando da diciassette a tredici). Tale fenomeno desta qualche perplessità perché, in virtù dell'art. 25, c.4, l. n. 241/90 alla CADA, tra l'altro, sono state assegnate funzioni semi-contenziose. In secondo luogo, sembrerebbe più conveniente da un lato, affidare ad un difensore civico statale l'attività paragiurisdizionale conferita alla Commissione; dall'altro, potenziare, di *contra*, le funzioni di vigilanza per la realizzazione della legge sul procedimento all'interno della CADA, con aumento dell'area di applicazione della trasparenza fino ai principi appartenenti alla legge n. 241/90 e con attribuzione alla medesima di poteri impositivi nei riguardi delle amministrazioni inosservanti. Infine, circa l'inquadramento strutturale della Commissione all'interno della Presidenza del Consiglio dei Ministri, sarebbe stato più vantaggioso costituire un'Autorità amministrativa indipendente, con caratteri di neutralità nei confronti degli interessi rappresentati.

Produzione ordinamento e gestione dei documenti

ANNA ROVELLA

Alcune questioni preliminari

L'elaborazione di un testo a più mani sulle problematiche del documento e dell'informazione è certamente, per chi scrive, una magnifica occasione di esperienza conoscitiva in cui la contestualizzazione della pratica didattica e professionale si valorizza attraverso la rielaborazione dell'apprendimento dei testi scritti da coloro che su questi argomenti hanno già formalizzato i loro studi e le loro riflessioni.

Il tema della produzione, dell'ordinamento e della gestione dei documenti rappresenta non solo l'elemento centrale di una copiosa quanto autorevole letteratura archivistica ma anche il *core business* di una notevole quantità di applicativi per la gestione documentale. Certamente, il considerevole sviluppo di teorie, l'elaborazione di standard, la realizzazione di sistemi di certificazione e la produzione di software dedicato degli ultimi venti anni postula e rinvia ad una reale esigenza di elaborazione di regole sicure, di certezze metodologiche, di strategie tecnologiche e di strumenti necessari a tutte le attività che a questo specifico ambito appartengono. Numerosi autori hanno ribadito, a più riprese, la validità – anche in un contesto variato rispetto a quello di elaborazione – degli assunti concettuali che da oltre un secolo presiedono all'organizzazione dei documenti e degli archivi, ribadendo la necessità della funzione documentale come mez-

zo attraverso il quale l'amministrazione esplica e concretizza la propria attività nello scenario socio-economico in cui si trova ad operare. Appare evidente, quindi, che ogni evoluzione e modifica della forma istituzionale e della sua interazione con la realtà ha una rilevante ricaduta anche sulle attività di formazione e di organizzazione dei documenti nelle amministrazioni pubbliche e private¹ proprio in relazione alla obbligatorietà di documentazione del soggetto produttore nel suo rapporto con l'ambiente di interesse. In particolare, la naturale evoluzione che ha accompagnato le democrazie occidentali nel corso dell'ultimo quarto di secolo ha determinato una crescente complessità dei processi amministrativi e produttivi aumentando, conseguentemente, il fabbisogno informativo delle organizzazioni. Tuttavia, l'incremento esponenziale e accelerato dei volumi di documenti² e il trend di crescita della cosiddetta "rivoluzione digitale" hanno, talvolta, favorito approcci e soluzioni caratterizzati dall'incompletezza teorica e dalla parzialità, conseguenze di una logica che ha inteso privilegiare il mercato e fronteggiare le emergenze piuttosto che sistematizzare soluzioni in grado di governare la molteplicità e la complessità dei problemi che andavano via via

¹ A proposito della specularità del rapporto tra Archivio e soggetto produttore si veda il noto contributo di CLAUDIO PAVONE, *Ma è poi tanto pacifico che l'Archivio rispecchi l'Istituto*, in «Rassegna degli Archivi di Stato», XXX, 3, 1970, pp. 145-149, nonché l'originale approccio di CARLO LAROCHE, *Que signifie le respect des fonds? Esquisse d'une archivistique structurale*, «La Gazette des Archives», Supplement au n. 73, 1971.

² Il rapporto dell'osservatorio sul documento elaborato nel 2007 da *Unioncamere* e *Net Consulting* indica che, in Italia nelle organizzazioni pubbliche e private, ogni anno si stampano circa 115 miliardi di pagine, di cui 19,5 miliardi inutilizzate, per un costo annuo di 287 milioni di euro; ogni documento viene duplicato in media 9/11 volte, per un costo di 18 euro/documento; il 3% dei documenti sono archiviati erroneamente la percentuale media del tempo di ogni impiegato speso in attività di gestione documentale e dell'ordine del 40%.

emergendo nella trasformazione di una società che è sempre più attenta al valore competitivo della conoscenza e al controllo dell'informazione³. Ne è conseguita un'insoddisfazione generale, causata da un'eccessiva, fiducia nella tecnologia e nelle sue possibilità di rendere efficienti processi di per se non efficaci a discapito della globale gestione del progetto archivio e dei problemi ad esso annessi, nella complessa eterogeneità di piani e di elementi che organizzazione e gestione di documenti e informazioni normalmente comportano all'interno delle amministrazioni pubbliche e private. A questo sentire comune della Pubblica Amministrazione non è estranea – comunque – la consapevolezza della farraginosità di alcune norme e la mancata sedimentazione di soluzioni pur tecnologicamente adeguate, problemi ai quali si è inteso rispondere accelerando il trend di innovazione/sostituzione senza aver definito i motivi dell'inefficacia della soluzione dismessa. Una simile prospettiva non ha favorito la consapevolezza che non è la singola tecnica o la specifica metodologia ad essere inefficace, quanto piuttosto la mancanza di una visione globale, capace di collocare correttamente i metodi e le tecniche nel più generale contesto funzionale nel quale devono esplicare la propria azione. La fase di transizione dal cartaceo al digitale è stata approcciata cercando di trasferire integralmente le regole e le metodologie esistenti per il cartaceo al digitale al quale, peraltro, sono state chieste verifiche e certezze mai pensate per il primo supporto. E' verosimile supporre che, un procedere siffatto, contenga nella sua stessa strutturazione gli elementi di freno e di limitatezza rispetto al pieno raggiungimento degli obiettivi imposti dalla teoria e dal mutato quadro legislativo. La carta e il di-

³ Sul valore strategico dell'informazione documentale ed il rapporto di questa con uno sviluppo possibile si veda il recente contributo di HAMAL HABIB – CLAUDE BALTZ, *Quelle Information pour piloter le développement durable?*, in «Documentaliste Sciences de l'Information», n. 1, Février 2008, pp. 4-13.

gitale dimostrano la loro diversità e le loro distanze non solo e non tanto nell'impiego di materiali e supporti quanto piuttosto nelle logiche cui i sistemi obbediscono e che devono essere attentamente ponderate anche prima della scelta dell'uno o dell'altro sistema, tenendo conto che la normativa in molti casi consiglia fortemente ma non impone un particolare tipo di media. Bisogna, tuttavia, riconoscere che lo sforzo profuso fin ora nella verifica di quanto la teoria archivistica e le metodologie elaborate per i documenti cartacei fossero utili ed efficaci anche per i documenti digitali certo non è stato vano. Esso è servito a riflettere sui nodi fondamentali della disciplina in termini di una sua applicabilità anche al mutare delle tecnologie. Ciò malgrado, di frequente, lo sforzo di adeguare, comunque, ogni assunto disciplinare al mutamento tecnologico ha prodotto risultati non sempre positivi sul piano teorico rendendo azzardate o pleonastiche alcune tesi e problematica la difendibilità di altre. L'esperienza maturata sul cartaceo rappresenta, al contrario, un'importante risorsa, una conoscenza fondamentale per l'elaborazione delle nuove regole. Le prassi consolidate, generandosi da competenze pratiche e applicative maturate nel corso del tempo, hanno un solido fondamento di legittimazione ma, proprio perché fortemente contestualizzate, devono servire da base certa per una rielaborazione funzionale, che tenga conto delle innovazioni intervenute nel contesto di applicazione. Quelle riflessioni, divenute consapevolezza di elementi noti e sperimentati, integrano il nuovo e concorrono a plasmarlo⁴.

⁴ «la conoscenza viene impiegata dai soggetti economici come risorsa produttiva, ma conserva la sua autonomia rispetto a questa formalizzazione. Dunque, perché l'economia della conoscenza possa affermarsi come disciplina autonoma, c'è bisogno di riconoscere le sue caratteristiche specifiche, adattando ad esse la teoria, e non viceversa. La pratica ha acquisito familiarità con la conoscenza impiegata, grazie anche alla lunga esperienza fatta nel corso della modernità. La teoria è ancora agli inizi, ci sta provando,

«Nel 1991, Albert Hirschmann, sociologo di Princeton, scrisse un libro di successo intitolato “La Retorica di Reazione” (The Rhetoric of Reaction) in cui egli classificava in diverse categorie di retorica i principali argomenti esposti per respingere le principali riforme del periodo post – illuminismo quali il riconoscimento dei diritti umani, l’universalizzazione del diritto di voto [...]»⁵. Egli classificava l’opposizione in cinque tesi: La Tesi di Rifiuto, la Tesi Perversa, la Tesi del Pericolo, la Tesi della Futilità, la Tesi delle Alternative e, di ognuna specificava il dettaglio di applicazione identificandole – alla fine – come categorie non necessariamente legate allo specifico contesto ma generalizzabili in maniera quasi paradigmatica. Senza voler indebitamente paragonare l’argomento qui trattato ai diritti fondamentali delle persone, forse non è necessaria molta immaginazione per calare quelle tesi nel mondo dei documenti e della controversa transizione dalla carta al digitale alla quale non è estranea la vera o presunta dequalificazione di alcuni tipi di lavoro in conseguenza dell’introduzione delle tecnologie informatiche, a mo’, quasi, di un luddismo digitale. «L’introduzione dei supporti informatici nei luoghi di lavoro incide profondamente sull’organizzazione del lavoro: nuovi processi produttivi, infatti, implicano spesso nuovi tipi di relazioni e ciò, a sua volta, presenta diversi aspetti. Se l’adozione di nuove tecnologie eleva la qualità di alcuni impieghi, essa comporta anche nuove separazioni del lavoratore dal prodotto del suo lavoro e dai suoi compagni. In alcuni casi i nuo-

e promette di imparare presto a vivere pericolosamente nel nuovo mondo della conoscenza, oltre il confine del noto e del rassicurante. *Hic sunt leones*: e non se ne andranno presto». ENZO RULLANI, «Economia della conoscenza. Creatività e valore nel capitalismo delle reti», Roma, Carocci, 2004, pp. 400-401.

⁵ BERNARD J. HIBBITTS, *Oggi come ieri, Scettici, scribi e la fine delle riviste giuridiche*, trad.it., in «Informatica e Diritto», Anno 2, n. 97 (1997), p. 83.

vi processi produttivi portano a nuovi livelli di cooperazione tra lavoratori, in altri ad un maggior controllo direzionale»⁶.

Le pagine che seguono delineano sinteticamente alcuni tra i principi, le teorie e i metodi generalmente applicati agli archivi in formazione con l'obiettivo di definire gli elementi fondamentali del disegno di un sistema di gestione documentale teso alla massimizzazione del valore amministrativo ed economico del patrimonio cognitivo e attento però, sin dalla fase di progettazione, alla corretta gestione del suo intero ciclo di vita.

Archivio, gestione documentale e documenti

Di recente lo standard UNI ISO 15489 - 2:2007 ha affermato che «le organizzazioni dovrebbero definire e documentare politiche di gestione documentale e dovrebbero assicurarsi che tali politiche siano recepite e applicate a tutti i livelli»⁷. Lo stesso standard chiarisce, poco più avanti, che le azioni necessarie all'attuazione di tale politica sono: regole condivise, coinvolgimento e sostegno del management e, una calibrata assegnazione di risorse economiche e umane. Lo scenario delineato è quello dell'archivio come elemento strategico delle organizzazioni e come strumento fondamentale per la governance. In realtà archiviare è senza dubbio una esigenza remota dell'uomo in quanto la conoscenza postula, necessariamente, un ordinamento per poter meglio comprendere le cose. Così, mentre sul piano formale l'archivio è una risposta a requisiti giuridico amministrativi e storico culturali⁸,

⁶ DAVID LYON, *La Società dell'Informazione*, Il Mulino, Bologna, 1991, p. 124.

⁷ UNI ISO 15489 - 2:2007, p. 1.

⁸ «La documentazione archivistica, detto in altri termini, è quasi sempre prodotta per finalità pratiche-operative, giuridiche, amministrative, connesse all'esercizio di specifiche e concrete attività [...] Troppo semplicistico è

sul piano operativo esso si pone come soluzione alle istanze di adeguatezza a principi normativi e funzionali⁹. Il concetto di archivio come risorsa¹⁰ consente il superamento di una prospettiva che per anni ha privilegiato la conservazione rispetto alle esigenze funzionali e la documentazione pubblica rispetto a quella privata essendo, quest'ultima, per molto tempo regolata – de facto – solo da norme fiscali e tributarie con un ciclo di vita estremamente breve. «La documentazione non statale perciò è stata a lungo pressochè esclusa dalla circolazione culturale e di fatto sottratta, in quanto memoria-fonte, all'uso storiografico»¹¹. Solo il

comunque affermare che la documentazione archivistica nasce “amministrativa” e diventa “storica”. Troppo semplicistico ritenere che il passaggio dall'una all'altra area di appartenenza sia da attribuire al mero trascorrere del tempo: col passare del tempo la documentazione perderebbe il suo primitivo significato pratico-amministrativo per acquistare quello storico-culturale. Entrambi i significati sono invece insiti nella documentazione archivistica nel momento stesso in cui è posta in essere». ISABELLA ZANNI ROSIELLO, *Gli archivi tra passato e presente*, Bologna, Il Mulino, 2005, p. 13.

⁹ «Riuscire a comprendere simultaneamente i due approcci significa capire l'archivistica nei suoi molteplici aspetti, valutare le sue possibilità di intervento, scegliere e calibrare i suoi strumenti metodologici e strategici, coniugare assieme tradizione e innovazione». GIORGETTA BONFIGLIO DOSIO, *Primi passi nel mondo degli archivi. Temi e testi per la formazione di primo livello*, Padova, Cleup, 2005, p. 11.

¹⁰ «Inoltre le pubbliche amministrazioni e i privati trovano nell'archivio uno strumento di lavoro eccezionale: da esso, infatti, possono ricavare i “precedenti”, conoscere cioè l'antefatto specifico o generale; con le informazioni ricavate dall'archivio possono valutare la situazione di contesto, programmare l'attività amministrativa, organizzare il lavoro verificare i risultati e controllare il rendimento della propria organizzazione”. BONFIGLIO DOSIO, cit. p. 24.

¹¹ ZANNI ROSIELLO, cit., p. 9. Si veda anche ELIO LODOLINI, *Archivistica. Principi e problemi*, Milano, Franco Angeli, 1995. e ANTONIO ROMITI, *Archivistica generale, Primi elementi: modulo di base*, Torre del Lago, Civita Editore, 2002.

dibattito culturale più recente ha rivalutato, con pari dignità, la documentazione privata modificando il concetto stesso di “privato” nel mondo della gestione documentale. Ciò è diventato particolarmente evidente quando le modificazioni legislative sulla gestione dei documenti amministrativi, hanno reso palese il necessario intreccio tra pubblico e privato e l’attenuarsi della linea di demarcazione tra i due mondi in un universo di casi e di situazioni per i quali è necessaria una comprensione e gestione delle interazioni piuttosto che solo delle specificità giuridiche dei singoli soggetti produttori. La migliore garanzia del rispetto della totalità dei caratteri fondamentali dell’archivio risiede, comunque, nella sua corretta impostazione e gestione fin dal processo di formazione: ciò garantisce il perseguimento delle finalità amministrative e giuridiche in vista del potenziale e successivo valore culturale del documento.

Tra la fine degli anni ottanta e gli inizi degli anni novanta, il dibattito presente negli Stati Uniti e in Europa sulla trasparenza amministrativa si è palesato, in Italia, in numerosi rapporti e lavori di commissioni parlamentari che hanno portato all’emanazione, nel 1990 della *Legge n. 241*. I principi di efficacia, efficienza e trasparenza amministrativa promossi dalla norma, unitamente ad una diversa definizione della concettualità di documento, hanno accentuato l’interesse verso *l’archivio corrente e di deposito* in quanto luoghi di conservazione di scritture ancora “vive”, ovvero di quei documenti che attivano e legittimano la comunicazione tra amministrazione e cittadino, posta alla base della riforma della pubblica amministrazione avviata in Italia proprio con la citata legge. Con lentezza ma inesorabilmente ci si è resi conto che nessuna trasparenza amministrativa era possibile in assenza della capacità di rendere fruibili, e quindi trasparenti, gli elementi primi di formalizzazione delle azioni amministrative: i documenti e la loro forma organizzata, gli archivi. È appena il caso di ricordare la polisemia del termine archivio che, mentre nella realtà amministrativa definisce univoca-

mente un elemento noto, nell'uso comune e nell'era digitale, esso, pur ricorrendo con estrema frequenza, assume una gamma di significati piuttosto ampia e non sempre coincidenti quello che gli è riconosciuto in ambito archivistico¹². In questo specifico dominio – infatti – l'archivio è «il complesso dei documenti prodotti o ricevuti da una persona fisica o giuridica nell'esercizio delle sue funzioni. Tali documenti sono legati da un vincolo originario e spontaneo di contenuto e di competenza»¹³. Tale definizione non rimanda quindi alla sua precipua natura fisica ma fa piuttosto riferimento ad una astrazione concettuale: un complesso di documenti in cui la caratteristica fondamentale è conferita dalla relazione causa – effetto, il *vincolo archivistico*, che collega in modo necessario le parti costituenti fin dalla loro produzione creando una sedimentazione in cui si palesa una corrispondenza – non priva di criticità – tra ordine dei documenti e funzionamento del soggetto produttore. La relazione tra archivio e soggetto produttore è quindi un elemento determinante che definisce e governa anche il sistema di classificazione dell'archivio stesso. In conclusione, ogni organizzazione produce i documenti perché necessariamente deve farlo per poter agire e, altrettanto necessariamente, li tratta, li gestisce e li ordina secondo un autonomo schema che è quello delle proprie specifiche funzionali. Il recupero dell'informazione e la gestione della co-

¹² Cfr. ELIO LODOLINI, *Questioni di base dell'Archivistica* «Rassegna degli Archivi di Stato» XXX, n. 2, maggio – agosto 1970, pp. 325-361, e le osservazioni sul testo di CLAUDIO PAVONE, *Questioni di Base o Questioni verbali*, «Rassegna degli Archivi di Stato», XXX, n. 3, settembre-dicembre 1970, pp. 660-662. In relazione alle definizioni del concetto di archivio rimanda, inoltre, alla bibliografia contenuta in *Scritti di Teoria Archivistica Italiana – Rassegna Bibliografica*, Ministero per i Beni e le Attività Culturali, Roma, 2000.

¹³ La definizione si richiama a quanto affermato da GIORGIO CENCETTI, *Sull'archivio come universitas rerum*, in «Archivi», IV, 1937, p.7-13.

noscenza del soggetto produttore sono evidentemente presupposto e garanzia della sua efficienza ed efficacia. In breve tempo la conoscenza è diventata una risorsa strategica che le organizzazioni e le persone utilizzano per creare valore economico e generare vantaggio competitivo. Al concetto di archiviazione si è andato quindi via via affiancando quello di organizzazione e gestione dei documenti. Il citato standard UNI ISO 15489 - 1:2006 definisce la gestione documentale come «settore delle attività gestionali responsabile dell'efficiente e sistematico controllo della produzione, ricezione, mantenimento, uso e destinazione finale dei documenti, incluse le operazioni per rappresentare e mantenere il valore probatorio e le informazioni circa le attività e le azioni in forma di documenti»¹⁴. L'utilizzo dei documenti digitali, la loro organizzazione in sistemi di gestione documentale comporta, però, necessariamente alcune valutazioni sul concetto di documento. Innanzi tutto, emerge una considerazione di ordine culturale che riguarda l'associazione, scontata quanto frequente, del termine documento al supporto cartaceo. Tale concezione, così fortemente radicata nell'immaginario collettivo, influenza, significativamente, anche l'elaborazione di regole, metodi, tecniche e tecnologie per la gestione dei documenti. In realtà, com'è noto, la storia ha conosciuto una pluralità di materiali scrittori la cui utilizzazione ha sicuramente influenzato le metodologie di conservazione ma difficilmente ha mutato le concettualità giuridiche e formali, pur se cesure e dubbi sono emersi ad ogni passaggio specie se non garantito da lunghi periodi di compresenza. È pur vero, però, che nel passaggio dal cartaceo al digitale non si è modificato solo un materiale scrittorio bensì, si è palesato un modo diverso di concepire e "vedere" lo scritto. La necessità di oggetti di decodifica spesso

¹⁴ UNI ISO 15489 - 1:2006, p. 3.

estranei alla comprensione dei redattori e la completa scissione del dato/informazione dal suo supporto e dalla fisicità del media ha prodotto una modifica unica – ad oggi – nella storia della scrittura. La definizione di documento data dalla citata Legge 241/90 ha rotto esplicitamente il rapporto tra il documento ed un supporto tangibile rimandando l'utilizzabilità del termine alla funzione di documentazione delle amministrazioni e sancendo, così, la piena legittimità di una pluralità di media ad entrare in archivio senza dover necessariamente operare distinguo ed interpretazioni estensive legate, almeno in Italia, alla particolare tipologia utilizzata per registrare la memoria. Oggi il concetto di documento non fa, quindi, più esclusivo riferimento alla sua collocazione spazio-temporale quanto piuttosto ai contenuti in esso memorizzati, alla rappresentazione dei fatti o degli atti di cui esso è testimonianza. Ciò tanto più nei documenti digitali immateriali e non immediatamente intellegibili. L'indeterminatezza concettuale ha determinato – conseguentemente - una oggettiva difficoltà nel definire con chiarezza il “materiale d'archivio”: ovvero la definizione dei documenti che per loro natura devono considerarsi parte, fin dall'origine, di uno specifico complesso documentario. Lunga ed emblematica – ad esempio – la discussione sulla natura giuridica e sulla validità probatoria del fax ieri e, oggi sui messaggi di posta elettronica. La diachronia tra definizione concettuale ed evoluzione reale comporta il rischio concreto di periodi più o meno lunghi nei quali la mancata inclusione di tipologie documentali nel novero dei materiali da preservare e conservare produce considerevoli depauperamenti della memoria delle organizzazioni. È necessaria, pertanto, una ridefinizione organica ed estensiva della concettualità di documento al fine di non disperdere – escludendolo - parte di quanto quotidianamente prodotto. Questa necessaria ridefinizione impatta, ovviamente, sulle problematiche dell'ordinamento, della conservazione e della consultazione. Sempre lo standard UNI ISO 15489 - 1:2006 definisce il documento quale «infor-

mazione registrata o oggetto che può essere trattato come un'unità»¹⁵. La stessa norma specifica poi, che i documenti di archivio o *record* sono «informazioni prodotte, ricevute e conservate a fini probatori e informativi da una persona fisica o giuridica per soddisfare obblighi legali o per lo svolgimento delle proprie attività»¹⁶ precisando, poi, anche gli elementi che determinano la scelta dei documenti da acquisire nel sistema documentale: analisi del quadro normativo, esigenze relative alle attività e alle responsabilità, valutazione del rischio rappresentato dalla mancata acquisizione. Il sistema di gestione documentale, di fatto, riconosce quindi una centralità dinamica al documento quale unità informativa, spostando significativamente quell'ottica puramente conservativa che nel concetto di unità attribuiva enfasi al nesso organico esistente in un insieme di documenti piuttosto che al singolo pezzo o, tanto meno, alle informazioni e ai contenuti in esso rappresentati. Il documento e il suo contenuto vengono, dunque, posti al centro del sistema archivio. La sua registrazione attiva tutte le «proprietà» e gli «attributi» che ne consentono la gestione: le relazioni con gli altri documenti e l'appartenenza al fascicolo, la classificazione in base a specifiche competenze del soggetto produttore, il controllo del ciclo di vita ovvero la posizione che esso occupa nel piano di conservazione; i procedimenti, l'iter, la tempistica e le risorse necessarie alla lavorazione; i diritti relativi all'accesso e alla consultazione. La natura dei documenti digitali e le procedure informatiche necessarie alla loro lavorazione ed organizzazione impongono una diversa articolazione del processo in cui i principi e le metodologie tradizionali (vincolo archivistico, classificazione d'archivio, ciclo vitale del documento) mantengono la loro centralità anche se approcciati da un diverso punto di partenza. L'aver

¹⁵ UNI ISO 15489 - 1:2006, p. 2.

¹⁶ Ivi, p. 3.

per lungo tempo considerato centrali e preponderanti, se non esclusive, le problematiche inerenti la conservazione e l'ordinamento della documentazione storica¹⁷, ha limitato fortemente la partecipazione degli archivisti all'evoluzione dei sistemi di gestione documentale lasciando ampi spazi di interventi ad approcci di altro genere e tipo¹⁸. L'affermazione, nel tempo, della cultura incentrata sul riordino e scarsamente attenta alla gestione dell'archivio in formazione ha fortemente condizionato il concetto di classificazione e ha minimizzato la portata di strumenti tipici dell'archivio corrente quali, ad esempio, la registrazione. Tuttavia, la recente normativa italiana sul protocollo informatico e i fabbisogni informativi delle organizzazioni private hanno individuato proprio nella registrazione un processo di importanza strategica nella operatività delle organizzazioni poiché esso avvia le procedure di certificazione della provenienza e di acquisizione del documento – identificandolo in maniera univoca nell'ambito di una sequenza numerica collegata con l'indicazione cronologica – e attiva e gestisce tutti i flussi di lavoro endogeni ed esogeni delle amministrazioni coinvolte.

¹⁷ Quasi a voler ribadire il principio enunciato dal Brenneke della netta distinzione tra Registratur ed Archiv, ovvero tra archivio corrente ed archivio Storico a tutto discapito del primo, cui viene assegnato un interesse meramente pratico mentre solo il secondo assurge a dignità scientifica. Cfr LEOPOLDO SANDRI, *L'Archivistica*, in «Rassegna degli Archivi di Stato», XXVII (1967), 2-3, pp. 410-429; VITTORIO STELLA, *La Storiografia e L'Archivistica, il lavoro d'archivio e l'archivistica*, in «Rassegna degli Archivi di Stato», XXXII (1972), 2, pp. 269-284. MARIA GUERCIO, *Gli Archivistici Italiani e la Sfida dell'automazione: archivi correnti e nuovi documenti*, in «Archivi per la storia», V (1992), 2, pp. 39-58. Per una valutazione della prospettiva di altri paesi si veda la sezione Electronic Records in «Atlanti», vol. 17 (2007) n. 1-2.

¹⁸ Cfr. *La formazione professionale dell'Archivista*, Atti del convegno di Eri-ce 2-4 novembre 2006, in «Archivi», anno II, n. 1 (gennaio-giugno 2007).

Registrazione e gestione del ciclo di vita dei documenti

La ratifica legislativa dell'obbligo dell'uso del registro di protocollo cartaceo nelle pubbliche amministrazioni in Italia avviene con l'emanazione del Regio Decreto n. 35 del 25 gennaio 1900. In realtà il protocollo si diffonde nella prassi archivistica nazionale tra la fine del XVIII secolo e gli inizi del XIX per effetto della riorganizzazione amministrativa francese che ne impone l'uso prima nella Repubblica Cisalpina e poi nel Regno d'Italia anche se le origini del protocollo come strumento amministrativo sono da ricercarsi in area tedesca/austriaca. Il citato Regio Decreto che stabiliva il regolamento per gli uffici di registratura e di archivio delle amministrazioni centrali resterà, tra l'altro l'unico riferimento legislativo di regolamentazione del protocollo per circa cento anni, fino all'emanazione del DPR n. 428/1998.

Malgrado la presenza di norme specifiche, nella prassi, i soggetti produttori hanno affrontato da sempre il problema della protocollazione con scarsa attenzione¹⁹, non considerandola come una delle fasi cruciali nella gestione dei sistemi documentali quanto piuttosto un fastidioso obbligo²⁰. A partire dal 2000 una rinnovata motivazione nel mondo pubblico e privato verso l'at-

¹⁹ Secondo una indagine Gartner Group del marzo 2003 le pubbliche amministrazioni ed, in misura minore, i privati hanno sottostimato le problematiche del document management evidenziando, contestualmente, che la gestione del protocollo ai sensi del R.D. 35/1900, con la quasi totale assenza di supporti informatici, assorbiva 250.000 addetti, pari al circa 7% delle risorse umane della p.a. con un costo di diverse migliaia di miliardi di vecchie lire.

²⁰ Già la volontà napoleonica teorizzata, nel 1806, nelle specifiche istruzioni ai Vice Prefetti del Regno d'Italia, era quella di legare la registrazione di protocollo alla classificazione di archivio e quindi in un qualche modo alla gestione della conoscenza dell'ente.

tuazione delle politiche dell'Unione Europea in materia di trasparenza amministrativa e gestionale²¹, una cospicua produzione di norme legislative in materia di gestione documentale, l'emanazione di standard e le accresciute necessità di circolazione e fruizione di informazioni hanno determinato una maggiore attenzione ed una rafforzata sensibilità verso la funzione che le attività di protocollazione/registrazione dei documenti rivestono all'interno dei processi di produzione di beni o di servizi.

In particolare, in ambito pubblico, il D.P.R. n. 445 del 28 dicembre 2000 e il Decreto Legislativo n. 82 del 7 marzo 2005, attribuiscono al termine protocollo informatico un significato estensivo, rispetto alla originaria concettualità, identificandolo come lo strumento di registrazione, di gestione dei documenti e dell'archivio delle amministrazioni. Alla funzione giuridica da sempre riconosciuta al protocollo, si viene ad affiancare quindi quella di sistema di gestione documentale, determinata dal fatto che i documenti digitali, ancor più di quelli cartacei necessitano per la loro gestione di processi fortemente controllati. Ciò malgrado, ad oggi le pubbliche amministrazioni continuano a privilegiare una versione fortemente limitativa dello strumento²². Non dissimile il panorama delle imprese italiane che rivelano

²¹ Basti pensare agli stimoli e agli incentivi che la pubblica amministrazione italiana ha ricevuto mediante le iniziative per l'attuazione dell'obiettivo «e-Europe, amministrazioni e cittadini nella digital age» e per la realizzazione dei piani di e-government; rilevante anche, per la sua visione prospettica la Risoluzione del Consiglio d'Europa 162/02 «Sulla conservazione della memoria del domani - conservazione dei contenuti digitali per le generazioni future».

²² Si veda, a questo proposito, Centro di Competenza per il Protocollo Informatico e Trasparenza Amministrativa (CNIPA), a cura di, «Valutazione delle azioni delle pubbliche amministrazioni sul tema del protocollo informatico e della gestione dei flussi documentali», *I Quaderni*, n. 34 ottobre 2007, consultabile alla data del 10 aprile 2008, all'URL <http://www.cnipa.it>.

una diffusa difficoltà nella realizzazione di sistemi di gestione e comunicazione delle informazioni e della conoscenza, causata, in gran parte, da una politica nella quale il vantaggio competitivo è assegnato alla raccolta e alla codifica di informazioni esterne piuttosto che alla loro integrazione con capitale cognitivo interno ed alla loro necessaria condivisione nei processi produttivi. I documenti e i dati acquisiti possono essere valorizzati e ottimizzati solo in un sistema organico in grado di controllare e di massimizzare la produzione, la fruizione e la condivisione delle informazioni presenti all'interno dell'organizzazione pena il mancato raggiungimento, sul mercato, delle performance sperate pur in presenza dell'impiego di ingenti risorse umane e finanziarie. Tuttavia, l'attuazione di processi di gestione documentale non è semplice né tantomeno scontata. Essa, al contrario, lungi dall'essere un mero insieme di accorgimenti tecnici, postula profondi cambiamenti organizzativi e procedurali, nuove metodologie operative, e fa nascere la pressante esigenza di nuove professionalità. Questi elementi, non disgiunti dalle necessarie risorse economiche, hanno limitato lo sviluppo di strumenti e di sistemi adeguati alla registrazione/protocollazione informatica dei documenti nell'archivio corrente digitale, determinando, di fatto, la scarsa attuazione di quanto disposto dal vigente quadro normativo. Peraltro, l'assenza di sanzioni che accompagna le suddette norme induce alla supposizione che, il legislatore, nella consapevolezza della difficoltà di applicazione e, a fronte di un esplicito gap presente tra principi affermati dalla norma e contesto applicativo - sul piano delle innovazioni, della tecnologia, delle infrastrutture e degli strumenti organizzativi - abbia legiferato per imprimere un'accelerazione al processo piuttosto che sancire, regolamentandola, una prassi emergente della società civile. In effetti, da più parti si è sottolineato che sia il DPR 445/2000 che il Decreto Legislativo 82/2005 appaiono estremamente generali, centrati più sull'aspetto innovativo che non su quello riorganizzativo, privi di elementi necessari a completare

quel quadro di insieme, quella corretta architettura concettuale, quella necessaria integrazione tra il piano giuridico amministrativo e la prospettiva storico culturale che è, di solito, garanzia di successo. In conclusione, alle norme vigenti sembra mancare un approccio adeguato che permetta di collocare correttamente la tecnologia, all'interno di un sistema di gestione documentale, quale strumento in grado di garantire efficienza, efficacia ed integrato in una più generale finalità di qualità globale che riservi la dovuta attenzione ad una coerente preservazione, conservazione e fruibilità futura delle fonti. Il concetto di qualità ricorre ampiamente anche nel più volte citato standard UNI ISO 15489 che, oltre a definire i requisiti di un sistema di gestione documentale di qualità²³, ne determina con chiarezza anche gli stru-

²³ «La norma Iso 15489 specifica per un sistema di gestione documentale i seguenti requisiti; esso deve:

- permettere l'acquisizione dei documenti insieme ai metadati che ne esplicitano la struttura (formato e relazioni tra le parti che lo compongono) e il contesto nel quale sono stati prodotti;
- permettere di organizzare i documenti secondo un ordine che rispecchia l'attività del soggetto produttore;
- contenere le informazioni su tutte le operazioni eseguite su un determinato documento;
- favorire l'accesso al patrimonio documentario registrato garantendone al contempo la riservatezza e l'integrità;
- registrare gli audit trail che provano il regolare funzionamento delle sue componenti e il mantenimento nel tempo delle caratteristiche di autenticità, affidabilità, integrità e riservatezza dei documenti;
- assicurare l'attendibilità dei documenti gestiti anche a seguito di eventi disastrosi, modifiche al sistema o migrazioni verso nuove piattaforme tecnologiche;
- disporre di supporti di memorizzazione adeguati al tempo di conservazione previsto per i documenti registrati;
- presentare funzionalità idonee per l'eliminazione dei documenti ritenuti inutili e la selezione di quelli da destinare alla conservazione a lungo termine;

menti necessari e tra questi attribuisce un ruolo primario alla classificazione²⁴. Indubbiamente un sistema di gestione documentale di qualità rappresenta un valido supporto dei processi di creazione e condivisione della conoscenza, anche se il fenomeno dell'information overflow rischia spesso di minimizzarne la portata. In sintesi in assenza di idonee soluzioni tecnologiche ed organizzative maggiore è la quantità delle informazioni minore è la capacità di trattarle e di attribuire ad esse significato e valore per l'azione. Ad Archivi scarsamente aggiornati e poco integrati si aggiungono spesso strumenti di classificazione non efficaci ed esaustivi. Lo standard UNI ISO 15489 - 2:2007 precisa che «i sistemi di classificazione forniscono all'organizzazione uno strumento per:

- a) organizzare, descrivere e collegare tra loro i documenti;
- b) mettere in relazione e condividere documenti di interesse comune sia all'interno sia all'esterno dell'organizzazione;
- c) migliorare l'accesso, il recupero, l'uso e la distribuzione dei propri documenti in maniera adeguata. Sostenuti da strumenti quali i vocabolari controllati e tecniche di indicizzazione, i sistemi di classificazione favoriscono la coerenza della intito-

- abilitare soluzioni architettoniche che prevedano la conservazione dei documenti ad opera di istituzioni diverse da quelle che li ha prodotti». STEFANO PIGLIAPOCO, *La qualità nella gestione dei documenti: la norma ISO 15489*, cit., p. 714.

²⁴ «Gli strumenti principali utilizzati nelle operazioni di gestione documentale sono:

- a) un piano di classificazione basato sulle attività;
- b) un piano di conservazione;
- c) uno schema di classificazione dei livelli di accesso e di sicurezza.

Le organizzazioni possono utilizzare ulteriori strumenti specifici di gestione documentale, quali:

- a) un thesaurus dei termini preferiti;
- b) un glossario dei termini o altri vocabolari controllati». UNI ISO 15489 - 2:2007, p. 9.

lazione e della descrizione facilitando così il recupero e l'uso dei documenti. I sistemi di classificazione possono essere usati per sostenere una varietà di processi di gestione documentale oltre a facilitare l'accesso e l'uso, come nel caso della memorizzazione e protezione e della conservazione ed eliminazione dei documenti»²⁵. Se la corretta classificazione è centrale per la conoscenza e la corretta utilizzazione dell'informazione e se sempre maggiori sono i volumi da gestire una delle sfide più importanti è proprio l'implementazione di strumenti capaci di automatizzare parte del processo diminuendo parallelamente il necessario utilizzo di personale specificatamente qualificato. I progetti di automazione della classificazione mostrano, tuttavia, un elevato grado di complessità, anche perchè una loro coerente realizzazione richiede una sinergia di competenze – informatiche, linguistiche, terminologiche, archivistiche – sicuramente non facile da realizzare specie con le difficoltà interdisciplinari tipiche del nostro paese. A ciò si aggiunga che la stringa di testo sulla quale operare è spesso limitata alla sola voce oggetto del registro di protocollo quale unico testo digitale in un universo ancora fortemente dominato dalla carta e quindi l'ambito sul quale poter operare la necessaria estrazione terminologica è fortemente penalizzato nell'utilizzo di algoritmi statistici appositamente dedicati. Nel mondo aziendale dove più forte e significativa è la

²⁵ UNI ISO 15489-2:2007, p. 9. La definizione riportata completa e chiarisce i termini di applicazione della classificazione d'archivio. Nel 2005, la Scuola Superiore della Pubblica Amministrazione definiva la classificazione d'archivio «attività di organizzazione sia logica che fisica di tutti i documenti correnti, protocollati e non di un soggetto produttore secondo uno schema articolato di voci che identificano funzioni, attività e materie specifiche del soggetto con la finalità precipua di individuare, per ciascun documento, la voce finale di appartenenza». S.S.P.A., *La metodologia per la definizione dei piani di classificazione in ambiente digitale*, Roma, 2005.

percentuale di documenti digitali si assiste ad un crescente numero di scelte tecnologiche tendenti alla gestione integrata del patrimonio informativo, pur se delocalizzato, generando però, a volte, soluzioni, sviluppate con logiche tra loro disomogenee e quindi non interoperabili non contribuendo certo alla capitalizzazione della conoscenza. Il problema è naturalmente destinato ad ingigantirsi nella realizzazione di progetti collaborativi di filiera, elementi oramai appartenenti non solo alla realtà di aziende produttive, ma fortemente operativi anche all'interno di organizzazioni preposte all'erogazione di servizi. La cosiddetta interoperabilità dei documenti, delle informazioni e dei dati, è ancora percepita come un'azione lontana, appartenente al futuro, mentre nel presente la sua realizzazione è ostacolata dalla considerazione che i sistemi di gestione documentale sono tradizionalmente e ideologicamente considerati esclusivo appannaggio delle singole organizzazioni complice anche la crescente attenzione verso le più generali problematiche della sicurezza. La norma UNI ISO 15489 - 2:2007 stabilisce che i livelli di accesso e di sicurezza richiedono una precisa regolamentazione che diventa ancor più stringente in presenza di organizzazioni complesse quanto ad attività e a quadro normativo di riferimento e prevede uno specifico piano di disaster recovery, ampiamente diffuso per i dati ma non altrettanto per i documenti²⁶.

Conclusioni

Le argomentazioni esposte invitano ad alcune riflessioni circa le prospettive e le opportunità future di ricerca in questo specifi-

²⁶ Cfr ALESSANDRA RUGGIERO (a cura di), *Conservazione delle memorie digitali rischi ed emergenze, sei casi di studio*, Dossier ICCU, 2004.

co ambito. Molti sono ancora gli elementi che aspettano di essere implementati o sottoposti a standardizzazione. La più volte citata UNI ISO 15489, che pur ha il merito di aver focalizzato alcuni aspetti importanti necessita di essere completata con una regolamentazione meglio definita sotto il profilo operativo in una logica più ampia di integrazione tra gestione documentale ed economia della conoscenza specie in ordine alla produzione di norme standardizzate di riferimento. L'operatività del sistema dipende dalla capacità di tutti gli attori di condividere finalità, obiettivi, tecnologie e metodologie. Le politiche di gestione documentale lungi dall'essere esclusivamente tecniche impattano fortemente con l'organizzazione della quale sono il prodotto postulandone necessari cambiamenti ed adeguamenti non sempre possibili e accettati. Maggiore è la condivisione delle scelte e il coinvolgimento dei decisori e maggiori sono, dunque, le possibilità di successo. Le scelte tecnologiche necessariamente orientate verso sistemi aperti e flessibili sotto il profilo metodologico sono un passo necessario ma successivo. La flessibilità e l'apertura al nuovo sono due accentuazioni necessarie in uno scenario fortemente dinamico nel quale si sperimentano continuamente nuove interazioni comunicative e, quindi, nuove concretizzazioni documentali. La transizione sempre più marcata, in alcuni ambiti, dal testo al visuale e la conseguente necessità di gestione di queste tipologie di documenti, la loro classificazione e l'indicizzazione dei contenuti in essi rappresentati è senza dubbio una ulteriore frontiera della gestione della conoscenza. Ma le sfide sono veramente tante: la comunicazione ha trovato altre forme importanti di rappresentazione negli strumenti di collaborazione che progressivamente si integrano sulle piattaforme web delle organizzazioni. VOIP, forum, instant messaging e, strumenti collaborativi come Blog e Wiki esprimono un nuovo panorama di comportamenti, di logiche organizzative e di comunicazione che in ragione della loro immediatezza, e dell'apparente veste informale sono in grado di superare quelle barriere alla condivisione

delle informazioni esistenti anche tra organizzazioni diverse ma cooperanti, rivelando una significativa capacità di incidere e una piena giustificazione alla loro integrazione in un sistema di gestione della conoscenza.

I linguaggi di descrizione documentale

EDUARDO DE FRANCESCO

Premessa

Esistono diverse formalismi per rappresentare informaticamente i documenti, i principali sono:

1. I sistemi di codifica dei caratteri;
2. I sistemi di marcatura dei testi.

I caratteri sono il livello elementare della rappresentazione d'un testo su supporto digitale, infatti ogni documento è rappresentato da un insieme di caratteri (stringa) che può rappresentare sia una sequenza di elementi lessicali, sia una sequenza di elementi di controllo.

Nel corso degli anni sono stati codificati molti sistemi per gestire le sequenze di controllo, alcuni di essi si limitarono a dialogare con il computer per fornire delle informazioni di controllo, altri si sono specializzati per potere interpretare il significato strutturale dei testi. Questi ultimi hanno preso il nome di linguaggi di marcatura (*tagging*) e si sono specializzati in due categorie:

1. Linguaggi procedurali indicati anche come “specific markup language”;
2. linguaggi dichiarativi detti anche come “generic markup language”.

Alla prima categoria fanno capo linguaggi come Script, TROFF, TEX o RTF. Essi permettono di indicare la struttura visuale (talvolta detta topografica) del testo e riportano informazioni relative al corpo del carattere, alla sua spaziatura, alla sua posizione rispetto alla riga, alla sua dimensione, al tipo di font utilizzato, alla strutturazione delle pagine, ecc. Alla seconda categoria fanno capo linguaggi come SGML, XML e parzialmente l'HTML; essi permettono di descrivere la struttura astratta del documento, dichiarando gli elementi che lo costituiscono (ad esempio: sezione capitolo, paragrafo, nota, enfasi, ecc.).

Nel seguito, verranno approfonditi questi concetti, sia attraverso la loro storia, sia attraverso un approfondimento tecnico specifico. Verrà – in particolare – analizzata la loro evoluzione nel tempo fino ad arrivare alla situazione attuale ed all'individuazione delle future linee di sviluppo.

Cenni storici

L'evoluzione dei sistemi di codifica dei caratteri

I caratteri vengono rappresentati mediante una codifica numerica binaria. Essa stabilisce una associazione biunivoca tra gli elementi di una collezione di simboli (*character repertoire*) e un insieme di codici numerici (*code set*): l'insieme risultante viene denominato tecnicamente **coded character set** (*set*). Per ciascun *set*, poi, si definisce una codifica dei caratteri (*character encoding*) basata su un algoritmo che mappa una o più sequenze di bit al numero intero che rappresenta un dato carattere in un *set*.

Alcuni *set* sono stati definiti da enti di standardizzazione nazionali e internazionali (ISO e ANSI in primo luogo) che si differenziano per il numero di cifre binarie che utilizzano, e dunque per il numero di caratteri che possono codificare. Tra questi il più diffuso è il cosiddetto codice ASCII (American Standard Code for Information Interchange), la cui versione in-

ternazionale si chiama ISO 646 IRV. Nato in USA, utilizza solo 7 *bit*; il *set* è composto da solo 128 caratteri, tra cui i simboli alfabetici dell'alfabeto anglosassone e alcuni simboli di punteggiatura.

Può sembrare strano l'uso di soli 7 *bit* (ormai tutti ragionano in *byte*, ossia su base 8) ma l'inglese americano non aveva (e non ha) la necessità di gestire caratteri particolari o accentati come negli alfabeti europei e a quel tempo ogni *bit* era prezioso¹.

La diffusione dei computer ha naturalmente determinato l'esigenza di rappresentare i caratteri di altri alfabeti. Sono stati così sviluppati molteplici code set che utilizzano 8 *bit* (un intero otetto e quindi 256 posizioni) e che hanno di volta in volta accolto i simboli dei vari alfabeti latini. Tra di essi ricordiamo la famiglia ISO 8859, nel cui ambito particolarmente diffuso è il *code set* ISO 8859-1, meglio conosciuto come ISO Latin 1. Esso contiene i caratteri principali delle lingue occidentali con alfabeti latini, ed è usato da molte applicazioni su Internet (ad esempio World Wide Web) e da molti sistemi operativi.

Tra gli altri set è da notare l'EBCDIC (dall'inglese Extended Binary Coded Decimal Interchange Code): esso indica un sistema di codifica dell'informazione a 8 bit usato in numerosi sistemi operativi di produzione IBM, sia per elaboratori di classe *mainframe*. Nel seguito di questo capitolo è stato inserito un piccolo approfondimento su questo *code set* perché utilizzato in modo significativo da IBM, in associazione al codice di marcatura GML, negli anni '70 ed '80.

Il più completo ed evoluto standard per la codifica di caratteri attualmente disponibile è l'ISO 10646-1, rilasciato nel 1993. Es-

¹ Pochi sanno che a tutto oggi il WEB usa ancora la codifica a 7 *bit*, con significativi problemi di conversione nei confronti di *set* di caratteri particolari.

so definisce lo *Universal Character Set*, un *coded character set* basato su una codifica a 31 *bit*. Esso coincide praticamente con l'omologo *set* a 16 *bit* (65536 combinazioni) Unicode, sviluppato autonomamente da una organizzazione privata, lo Unicode Consortium.

Occorre spendere qualche ulteriore parola su Unicode che attribuisce un numero univoco a ogni carattere, indipendentemente dalla piattaforma, dall'applicativo, dalla lingua. Lo standard Unicode è stato adottato da leader di mercato del calibro di Apple, HP, IBM, JustSystem, Microsoft, Oracle, SAP, Sun, Sybase, Unisys e molti altri. Esso è alla base di molti moderni standard, come XML, Java, ECMAScript (JavaScript), LDAP, CORBA 3.0, WML eccetera, e costituisce l'implementazione ufficiale dello standard internazionale ISO/IEC 10646. Il suo emergere, unito alla recente disponibilità di strumenti che lo supportano, è da considerarsi fra i più significativi sviluppi della tecnologia della globalizzazione del *software*.

L'adozione di Unicode sui siti web e nelle applicazioni *client/server* o *multi-tiered*, rispetto all'utilizzo dei set di caratteri tradizionali, permette un significativo abbattimento dei costi di gestione. Esso consente che un'unica versione di un *software* o di un sito web siano fruibili con piattaforme, lingue e paesi diversi, evitando la necessità di reingegnerizzare il prodotto per ogni situazione specifica. Permette, inoltre, il trasporto del testo fra sistemi diversi senza che abbia luogo alcuna corruzione dei dati.

Cronologia dei coded character set

Ufficialmente, il primo fu il codice Morse, nato nel **1840**, seguito dal linguaggio delle bandiere usato in marina.

L'introduzione delle telescriventi porta al codice Baudot del **1930**, a 5 bit.

Nel **1963** nasce lo standard ASCII, a 7 bit, che è utilizzato per la nascita di Internet e dei protocolli connessi. Nel 1965 viene approvato lo US-ASCII, che nel 1972 diventa lo standard ISO 646.

Tra il 1963 e il **1964**, è stata presentata da IBM. la codifica EBCDIC. Si trattava di una codifica a 8 bit in grado di estendere in modo rilevante il sistema BCD a 6 bit allora in uso; il suo sviluppo è stato indipendente rispetto alla ASCII, che è una codifica a 7 bit.

Nel **1981** le estensioni al codice ASCII per i caratteri dal 128 al 255 vengono identificate dai *codepage* PC-DOS e *traslate* poi per retrocompatibilità nello MS-DOS. Nel 1985 la ISO approva gli standard *codepage* come ISO 8859-n, dove n è un numero che identifica il particolare *codepage*.

L'affermarsi di Windows, anche in Asia, porta alle estensioni alle lingue orientali, nel **1990**, dei *codepage* di Windows.

La comprensibile babele risultante dallo scambio di email e documenti tra paesi a *codepage* diverso fu oggetto di attenzione dell'ISO prima con lo standard del 1986 ISO 2022, in vigore ma scarsamente utilizzato, e poi con la proposta del **1991** dell'Unicode 1.0, la cui versione 1.1 del 1993 diviene lo standard ISO 10646, lo Universal Character Set o UCS.

Lo UCS-2, che usa due *byte* per ogni carattere, fu utilizzato dalla Microsoft in Windows NT sin dal **1995** e poi esteso a tutte le altre versioni.

Le definizioni dei formati UTF-8 e UTF-16 datano al **1996**, con la versione 2.0 di Unicode. Lo UTF (Unicode Transformation Format) divenne lo standard POSIX *de facto*, ed essendo ratificato dalla RFC 3629, è anche riconosciuto dal W3C. Esistono anche lo UTF-7 e l'UCS 4. Lo UTF-16 è un'estensione dello UCS 2.

La successiva versione Unicode 3.0 del **1999** introduce la bidirezionalità e la composizione di tabelle, mentre la 4.0 del 2001 include anche le lingue antiche. La 4.1 del **2005** è la versione attualmente in uso.

L'evoluzione dei sistemi di marcatura

La funzionalità della rappresentazione digitale del testo dipende dalla qualità con cui l'informazione può essere modellizzata e dalle operazioni ad essa applicabili. L'informazione testuale viene rappresentata mediante sequenze lineari, o stringhe, di caratteri codificati in forma binaria. Il *markup*, ossia l'inserimento di marcatori (*tag*), permette di assegnare una struttura alla rappresentazione del testo distinguendo, nella sequenza dei caratteri codificati, parti diverse con funzioni diverse. Il modello assegnato all'informazione testuale, dipende dal sistema di *markup*.

L'espressione *markup* deriva dalla analogia tra questi linguaggi e le annotazioni inserite da autori, curatori editoriali e correttori nei manoscritti e nella bozze di stampa di un testo, al fine di indicare correzioni e trattamenti editoriali, chiamate appunto in inglese *mark up*. In modo simile, i linguaggi di marcatura sono costituiti da un insieme di istruzioni, dette *tag* (marcatori), che servono a descrivere la struttura e la formalizzazione mediale del documento. I marcatori sono sequenze di normali caratteri ASCII e vengono introdotti, secondo una determinata sintassi, all'interno del documento, accanto alla porzione di testo

cui si riferiscono. Secondo i vari linguaggi di marcatura si possono rappresentare o il modello strutturale dei contenuti del documento o il modello di rappresentazione mediale. Il testimone più illustre dei linguaggi di markup è lo *Standard Generalized Markup Language* (SGML). Ideato da Charles Goldfarb, SGML è divenuto lo standard ufficiale adottato dalla *International Standardization Organization* (ISO) per la creazione e l'interscambio di documenti elettronici. Da esso si sono poi originati altri linguaggi di marcatura (HTML, XML ed altri), approfonditi nel seguito del testo. Quanto segue è la storia dell' SGML (e quindi di fatto dei linguaggi di marcatura) così come raccontata da suo fondatore e nome tutelare Charles Goldfarb, sulle pagine del SGML User Group (www.isgmlug.org) nel giugno del 1990.

Il concetto generico di codifica

Storicamente, i manoscritti elettronici contenevano dei codici di controllo o delle macro che permettevano a un documento di essere formattato in un modo particolare (*Specific coding*). Al contrario, le metodologie di codifica generalizzata, che iniziano nel tardo 1960, usano marcature descrittive (ad esempio la marcatura "titolo" al posto di un "formato-17").

Molta della spinta verso la codifica generica è dovuta ad una presentazione effettuata da William Tunnicliffe, *chairman* della Graphic Communications Association (GCA), durante un *meeting* presso il Canadian Government Printing Office nella September nel 1967. Durante questo *meeting* viene infatti affermato per la prima volta uno dei concetti fondamentali dei linguaggi di marcatura moderni ossia la separazione completa fra il concetto di contenuto informativo del documento e la sua formalizzazione mediale. Sempre durante l'ultimo periodo del 1960 Stanley Rice (progettista di libri a New York) propose l'idea di un catalogo universale di elementi di marcatura parametrici (*tag*). L'idea fu

raccolta da Norman Scharpf, Director della GCA che propose un "committee" sull'argomento. Il "committee" concluse che era necessario utilizzare diversi sistemi di marcatura per differenti tipi di documenti e che documenti più piccoli potevano essere incorporati come elementi di documenti più grandi. Il progetto dette origine al GenCode Committee che fu, di fatto, la base operativa per la creazione dello standard SGML.

Nel 1969 Charles Goldfarb dirigeva un progetto di ricerca IBM su un sistema informativo per la gestione delle leggi. In quel contesto inventò il Generalized Markup Language (GML) come metodo per permettere la scrittura, la formattazione il recupero e la condivisione di documenti. Questo linguaggio era basato sull'idea originale di Rice e Tunnicliffe, ma introduceva anche il concetto di strutture ricorsive esplicite all'interno del documento. Questo linguaggio, utilizzato dai mainframe IBM, raggiunse una larga diffusione, tanto che la stessa IBM diventò, in quegli anni, il secondo publisher a livello mondiale.

La nascita dell'SGML come standard mondiale

Nel 1978, l'American National Standards Institute (ANSI) decise la nascita di un comitato per lo studio del "Computer Languages for the Processing of Text" capeggiato da Charles Card. Goldfarb fu invitato a partecipare al progetto per favorire la nascita di uno standard basato su GML. La prima bozza dello standard vide la luce nel 1980 e si concluse con la pubblicazione dello stesso nell'ottobre del 1985 come ISO 8879. È da rimarcare come durante questa fase si instaurò una cooperazione con l'"European Particle Physics Laboratori" (CERN) di Ginevra dove, di fatto, verrà poi generato lo standard HTML.

Nei primi anni successivi alla nascita dello standard, ci furono delle importanti evoluzioni dovute alla adozione dell'SGML da parte della Association of American Publishers (AAP), tramite il

progetto “Electronic Manuscript Project”, e del Ministero della Difesa Americano, tramite l’iniziativa Computer-aided Acquisition and Logistic Support (CALs). Questi due eventi permisero la diffusione dello standard nel mondo degli editori e nel mondo della difesa/spazio.

L’hyperlink

il termine “hyperlink” fu coniato nel 1965 da Theodore Nelson all’inizio del progetto “Xanadu”. Nelson fu ispirato da una storia in cui si descriveva una macchina per microfilm, nella quale si potevano collegare due qualsiasi pagine in una “catenella” di informazioni e quindi muoversi liberamente in questa sequenza. In un corpus di libri pubblicati tra il 1964 e il 1980, Nelson trasferì questi concetti in un contesto informatizzato rendendo applicabile questo concetto di collegamento non più alle pagine intere ma a singoli elementi (stringhe) del testo e, poco dopo, ai differenti paragrafi di diversi documenti. Di fatto era nato il concetto di ipertesto.

Un caso particolare ed importante di “specific markup language”

Uno dei più comuni linguaggi di codifica procedurali è l’RTF (Rich Text Format, “formato di testo ricco”), realizzato dalla Microsoft che ne è proprietaria, al fine di agevolare lo scambio di documenti fra diverse applicazioni. L’RTF è un tipo di codifica del testo “tagged”, in cui il documento viene codificato come puro testo e degli “identificatori” (*tag*) indicano quale tipo di strutture di formattazione (allineamento, impaginazione, tipo di carattere) applicare al semplice testo. Per mezzo della marcatura, il testo del documento RTF (definito sfruttando soltanto set di

caratteri molto semplici e diffusi, tipo l'ASCII) viene arricchito di nuove informazioni riguardanti la formattazione, di cui i formati di puro testo sono privi: da ciò la spiegazione del nome di "formato di testo ricco".

I tag RTF sono in genere definiti dal carattere \ (*back slash*) seguito da alcune lettere che specificano il *tag*. Un esempio di testo marcato con *tag* RTF è:

```
{\b Questo testo è in grassetto,  
{\i e questo è anche in corsivo.}}
```

Microsoft ha a lungo sostenuto (fino circa al 1994) questo standard in alternativa all'SGML, l'RTF infatti ha perso la sua battaglia contro l' SGML quando la diffusione del WEB ha imposto di fatto i linguaggi di marcatura generica. È da notare comunque che oggi tutti i *word processor* più diffusi hanno la possibilità di importare e di esportare dei file in RTF e che i manuali ipertestuali realizzati con la tecnologia dell'"help" di Windows (formato *.HLP) sono tutti basati sul suddetto formato.

La standardizzazione del metodo per descrivere gli elementi costituenti il documento

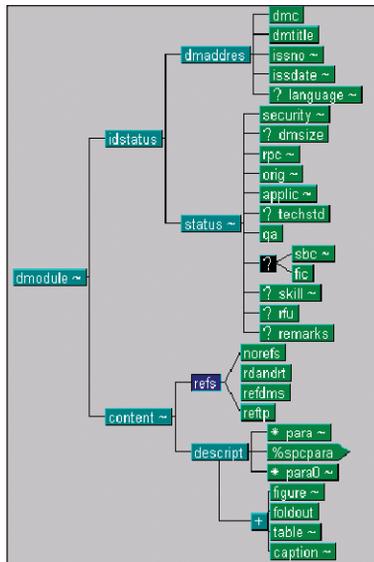
Negli anni tra il 1960 e il 1985 si sviluppa quindi un movimento che porta alla definizione del primo standard di un linguaggio di marcatura, l' SGML.

Vediamo ora in concreto che cosa significa scrivere tramite un linguaggio di marcatura; nell'esempio che segue, utilizzando un documento a tutti noto quale la *Divina Commedia* di Dante, si è provato a rappresentarlo in termini strutturali e tramite elementi di marcatura, da qui in seguito chiamati, secondo l'uso internazionale, "tag".



La struttura logica riportata nella parte sinistra della figura, ri-trae una delle possibili modellizzazioni con cui si può rappre-sentare il documento in esame.

Nell'esempio in figura, tratto da una normativa per la model-lizzazione di documenti tecnici in ambito aeronautico (S1000D), si vede la strutturazione di un documento complesso attraverso la sua rappresentazione grafica.



Ogni elemento rappresentato in figura, tipicamente composto da *tag* o altri elementi quali attributi della *tag*, si descrive attraverso un linguaggio specifico di marcatura.

Il progettista documentale analizza il documento nelle sue componenti e ne definisce una struttura o modello, descrivendolo quindi attraverso un linguaggio formale specifico; nel nostro caso appunto l'SGML. In particolare:

- Ogni documento SGML deve contenere un solo elemento radice, cioè un elemento che racchiude tutti gli altri elementi secondo una struttura di relazione padre-figlio, cioè strettamente gerarchica;
- L'elemento radice è il primo elemento di un documento SGML;
- Le relazioni gerarchiche esistenti fra i vari elementi danno vita al cosiddetto albero SGML.

Più che un linguaggio, lo SGML è un *metalinguaggio*. Esso prescrive precise regole sintattiche per definire un insieme di marcatori e di relazioni tra marcatori, ma non dice nulla per quanto riguarda la loro tipologia, quantità e nomenclatura. Questa astrazione costituisce il nucleo e la potenza dello SGML. In sostanza, SGML serve non già a marcare direttamente documenti, ma a costruire, rispettando standard comuni e rigorosi, specifici linguaggi di marcatura adatti per le varie esigenze particolari. Un linguaggio di marcatura che rispetti le specifiche SGML viene definito 'applicazione SGML' (*SGML application*).

Una applicazione SGML, a sua volta, descrive la struttura logica di una classe di documenti e non la loro forma fisica. Tale struttura astratta viene specificata dichiarando gli elementi che la costituiscono (ad esempio: *capitolo*, *titolo*, *paragrafo*, *nota*, *citazione*, ecc.) e le relazioni che tra questi intercorrono, relazioni che possono essere *gerarchiche* o *ordinali*.

Le dichiarazioni sono contenute in un *file*, denominato *Document Type Definition* (DTD), che costituisce una sorta di gram-

matica dei documenti che ad essa si riferiscono e rispetto alla quale debbono essere convalidati. Più precisamente nella DTD sono elencati e definiti tutti gli oggetti necessari all'elaborazione di un determinato tipo di documento:

- **Elementi:** i nomi dei marcatori adottati per identificare i blocchi strutturali costituenti l'albero dei dati.
- **Content model:** il modello di contenuto per ciascun elemento indica quali altri elementi, quando ed in che numero, possono comparire dentro un elemento. Indica inoltre quali attributi possono essere associati ad un elemento. Gli attributi permettono l'aggiunta all'elemento di informazioni addizionali, che possono essere utili per descrivere più accuratamente la struttura dei dati rappresentata. Gli attributi possono essere paragonati a degli aggettivi.
- **Entità:** le entità permettono di creare dei riferimenti a dati esterni non necessariamente testuali, ad altri documenti, a *file* grafici e ad altri media. Molto usate sono le entità per rappresentare caratteri, che non sono presenti nella tabella codici del documento o che si teme potrebbero essere rappresentati non correttamente nel passaggio da un sistema all'altro.

Riprendendo le parole di Yury Rubinsky, creatore del primo editor commerciale SGML, si può dire che un documento realizzato in SGML non presenta sorprese poiché:

- Esso deve cominciare con una "SGML declaration". Questa dichiarazione formale standard permette al ricevente del documento, di sapere esattamente quale "character set" si aspetta di ricevere e quali sono le regole specifiche di marcatura applicabili.
- Deve indicare (contenuto o riferito) qual è il modello documentale (DTD) applicabile all'istanza documentale che segue.
- L'istanza documentale deve seguire le dichiarazioni

suindicate, sapendo quale *set* di caratteri e quale *set* di regole sono applicabili.

Un'ultima annotazione può essere interessante: si è parlato di GML come precursore dell'SGML. Esiste una precisa correlazione tra il linguaggio GML e il *character set* EBCDIC, in quanto entrambi espressione del grande sviluppo commerciale dei *mainframe*. L'SGML invece, pur potendo accettare qualsiasi *character set*, è più legato alla codifica ASCII ed alle sue evoluzioni.

Le codifiche EBCDIC e ASCII non sono compatibili tra loro. Poiché i computer sono in grado di elaborare solo dati numerici, entrambe le convenzioni assegnano specifici caratteri a tali numeri, ma identici valori numerici sono interpretati come caratteri diversi in funzione della "codepage" utilizzata. Per utilizzare dei dati memorizzati in EBCDIC è necessaria una conversione tra "codepage" per visualizzare correttamente le informazioni su elaboratori basati sulla codifica ASCII.

Nell'esempio di seguito riportato si è analizzato il documento "ricette gastronomiche"; questa formalizzazione ed i relativi impatti a livello informatico saranno spesso usati come esempio nelle successive pagine.

```

<ricettario>
  <ricetta>
    <ricetta idref="Alici Marinate">
      <ingredienti>
        <ingrediente idref="Alici Fresche" quantita="600 g"/>
        <ingrediente idref="Aglio" quantita="2 spicchi"/>
        <ingrediente idref="Prezzemolo" quantita="qb"/>
        <ingrediente idref="Sale marino" quantita="qb"/>
        <ingrediente idref="Origano" quantita="qb"/>
        <ingrediente idref="Olio extra vergine" quantita="qb"/>
        <ingrediente idref="Peperoncino rosso" quantita="qb"/>
        <ingrediente idref="Limon" quantita="2"/>
        <ingrediente idref="Aceto Bianco" quantita="2 bicchieri"/>
      </ingredienti>
      <preparazione>
        <fase idref="1" testo="Pulire le alici, togliendo..."/>
        <fase idref="2" testo="Sopra ogni strato mettere sale..."/>
        <fase idref="3" testo="Aggiungere a ogni strato il succo..."/>
        <fase idref="4" testo="All'ultimo strato a tutti gli in..."/>
        <fase idref="5" testo="Far riposare per mezz'ora e poi..."/>
        <fase idref="6" testo="Servire aggiungendo un filo d'olio..."/>
      </preparazione>
    </ricetta>
  </ricette>
</ricettario>

<DOCTYPE Ricettario>
<ELEMENT Ricettario (Glossario,Ricette)>
  <ELEMENT Glossario .....>
  <ELEMENT Ricette (Ricetta)>
    <ELEMENT Ricetta (Titolo, Ingredienti+, Preparazione)>
      <ELEMENT Titolo (#PCDATA)>
        <ELEMENT Ingredienti EMPTY>
          <ATTRIBUTE Ingredienti IDREF #REQUIRED>
        <ELEMENT Preparazione (Fasi+)>
          <ELEMENT Fasi (Testo, Ingredienti)>
            <ELEMENT Testo (#PCDATA)>

```

In figura, nella parte superiore viene presentata l'istanza del documento, nel caso specifico la ricetta, nella parte inferiore la sua modellizzazione formale (DTD).

Una volta definito un determinato linguaggio, esso può essere utilizzato per rappresentare infiniti documenti in base ad una sintassi rigorosa. A ciascun elemento corrisponde una coppia di marcatori. La sintassi standard prevede che i marcatori siano racchiusi tra i simboli di maggiore e minore. Ogni elemento viene identificato da un marcatore iniziale e uno finale (costruito premettendo una barra al nome del marcatore iniziale), a meno che non sia un elemento vuoto (ne qual caso è identificato solo dal marcatore iniziale).

Le *tag* (SGML e suoi successori), oltre alla loro potenza espressiva, offrono una serie di vantaggi dal punto di vista del trattamento informatico dei documenti. Un documento marcato è composto esclusivamente da una sequenza di caratteri in un dato *code set* ed è quindi facilmente portabile su ogni tipo di computer e di sistema operativo. Lo stesso testo marcato può essere utilizzato per scopi differenti: stampa su carta o video, rappresentazione su media diversi (audio, tatto, etc.), analisi tramite *software* specifici, elaborazione con *database*, correlazione automatica di *corpus* linguistici. Ciò, anche in tempi diversi, senza dovere pagare i costi di complesse, costose e spesso inaffidabili conversioni tra formati spesso incompatibili.

Particolare rilevanza assumono i linguaggi di marcatura per la produzione e la manutenzione di pubblicazioni articolate, come quelle relative alla documentazione tecnica, nelle quali la sicurezza del dato è elemento spesso determinante per la sicurezza dell'uomo, o al settore legale, nel quale ogni singola parola ha spesso valenza specifica.

Per queste sue caratteristiche, SGML ha trovato impiego soprattutto in contesti industriali e militari, nei quali la gestione efficiente e sicura dei documenti tecnici ha una funzione critica. Ma non mancano applicazioni SGML in ambito scientifico e

umanistico, nel quale ultimo una applicazione SGML denominata *Text Encoding Initiative* è divenuta lo standard per la codifica e l'archiviazione dei testi su supporto digitale e per la creazione di biblioteche digitali.

Senza dubbio – comunque – l'applicazione SGML che gode della diffusione maggiore, anche se molti non la percepiscono come SGML, è il linguaggio attuale del WEB ossia l' *HyperText Markup Language* (HTML).

L'evoluzione delle tag: dalla identificazione degli elementi costituenti la singola istanza al collegamento tra più istanze (*iperlink*)

Il modello documentale legato all'SGML è relativo a una singola istanza del documento; ciò significa, in pratica, che è possibile modellizzare un libro, un documento tecnico, la descrizione di un farmaco, un ricettario, ma di per sé non è possibile collegare dei libri, dei documenti tecnici, delle descrizioni di farmaci, dei ricettari. Ossia non esiste, originariamente, all'interno della codifica di marcatura SGML, un metodo per il collegamento tra più istanze. Questa enorme limitazione che, di fatto, non permetteva la creazione di ipertesti fra più istanze, viene rilevata intorno alla fine degli anni '80 e trova una sua esplicitazione in una serie di studi che culminano con la creazione del linguaggio HTML, il cui padre fondatore viene considerato Tim Berners Lee.

Tim Berners Lee era un ricercatore del CERN² di Ginevra e,

² Il CERN, Organisation européenne pour la recherche nucléaire, (storicamente il nome è l'acronimo di Conseil Européen pour la Recherche Nucléaire), è il più grande laboratorio al mondo di fisica delle particelle. Si trova alla periferia ovest della città di Ginevra. La convenzione che istituiva il CERN fu firmata il 29 settembre 1954 da 12 stati membri. Oggi fanno parte del CERN 20 stati membri più alcuni osservatori anche extraeuropei.

come si è accennato precedentemente, esisteva una forte correlazione culturale tra i comitati ISO che avevano portato alla nascita dello standard SGML e l'istituto di ricerca svizzero. In tale contesto – da Berners Lee – venne definito un progetto “interno” per di correlare più documenti tra loro.

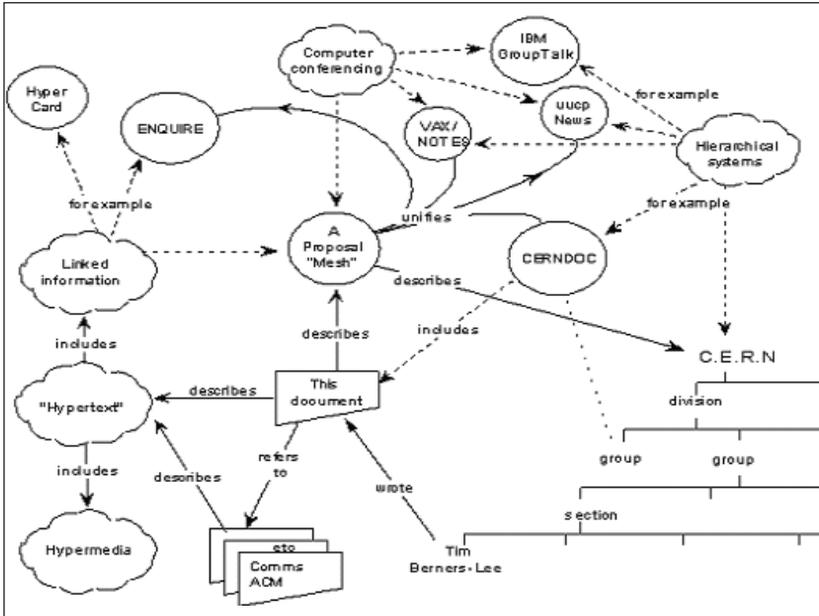
Il testo seguente è ispirato e tratto proprio dal documento originale che ha dato origine al progetto “Enquire” del CERN (1989) e quindi di fatto a HTML, diventato lo standard mondiale del Web.

Il problema del CERN

Il CERN era una magnifica organizzazione, coinvolgeva svariate migliaia di persone, molte delle quali veramente creative, tutte al lavoro verso obiettivi comuni. Per quanto esse fossero nominalmente organizzate in una struttura gerarchica, non c'erano limiti tra le varie persone per comunicare, scambiare informazioni, apparati e *software* trasversalmente ai gruppi. La struttura era come una ragnatela multipla interconnessa, le cui interconnessioni variavano in continuazione. Il personale infatti variava ruolo ogni circa due anni. Ciò creava un ambiente estremamente stimolante, ma inevitabilmente le varie riunioni finivano con la seguente considerazione: OK è tutto posto, ma come possiamo tenere traccia di un progetto così grande?

Se il CERN fosse stato un esperimento statico, si sarebbe potuto provare a mettere tutte le informazioni in un “grande libro”, ma il CERN variava continuamente e sempre nuove idee venivano generate mano mano che nuove tecnologie si rendevano disponibili. Ciò rendeva impossibile depositare le informazioni in un solo posto centrale.

La figura che segue può rendere l'idea del problema.



Le informazioni che servivano durante le riunioni erano del tipo:

- dove è utilizzato questo modulo?
- chi ha scritto questo codice e come funziona?
- che documenti ci sono su questo argomento?
- quali laboratori sono coinvolti con questo progetto?
- a chi si riferisce questo documento?

Come si può notare il problema della perdita di informazioni era particolarmente grave, ma in effetti il CERN di allora rappresentava solo in miniatura il problema attuale dell'informazione in WEB. Il problema era la correlazione tra più elementi informativi tra loro fisicamente separati, oggi lo definiremmo un tipico problema di ipertesto o addirittura di ipermedia, Tim Berner Lee usa e, di fatto, battezza la parola

“WEB” «The organization is a multiple connected web». In questo contesto Tim Berners Lee propone un “linked information system” che possa evolvere con l’organizzazione ed il progetto che descrive. Per rendere possibile ciò, il metodo di memorizzazione delle informazioni non può porre restrizioni all’informazione stessa. Ciò perché una ragnatela “web” di note, ognuna con le proprie connessioni “links”, può essere molto più utile di un sistema gerarchico centrale. Era nato il concetto di informazione diffusa in Web, che sarebbe diventato il modello attuale del mondo moderno.

Una seconda considerazione di Lee era quella che non si poteva parlare di modelli predefiniti di documento, ma di volta in volta occorreva riferirsi a documenti molto differenziati tra di loro. Arriva quindi alla conclusione che occorre sviluppare un modello generico e non un modello specializzato di documento, mentre doveva essere esaltato al massimo il concetto della correlazione fra più tipi di informazioni e quindi di documenti.

Non era stato ancora però definito il modello informatico, che poteva supportare questa architettura, ma occorre ricordare che il CERN era stato fortemente coinvolto durante il progetto SGML e quindi aveva piena competenza dei linguaggi di marcatura.

La nascita dell’HTML

Da tutte queste considerazioni emerse quindi il modello informatico, basato ovviamente sull’SGML, tramite il quale si realizzò un modello di documento assolutamente generale, dotato di una “testatina”, contenente le parole chiave, e di un corpo generico. Ma la grande novità è la formalizzazione e quindi la normalizzazione dei *link* fra più documenti. Un grande passo avanti è stato effettuato: non si concepisce più soltanto il grande

documento monolitico (il grande manuale tecnico, la grande enciclopedia) ma anche una serie di piccoli documenti, diversi tra loro ma collegati nel modo più flessibile possibile. È nato l'HTML in cui l'accento è posto sulla parola HyperText (HT) ed ovviamente in Mark-up Language (ML).

L'evoluzione dei linguaggi di marcatura (SGML, HTML, XML) *SGML ed HTML*

Anche in questo caso occorre fare un po' di storia.

Abbiamo visto come nasce l' SGML, ne abbiamo capito la forza ma anche le limitazioni ed abbiamo seguito la successiva nascita del HTML, quasi 10 anni dopo la definizione dello standard ISO 8879.

Riassumiamo un po' la situazione. L'SGML si pone come linguaggio di marcatura generica standard pensato per la realizzazione e gestione di grandi quantità di documenti, composti da singole istanze di significativa entità. La sua forza sta nella standardizzazione, nella capacità di interscambio tra sistemi diversi, nella possibilità di ipotizzare processi molto più complessi, in quanto il documento diventa dato. Il suo limite sta nella complessità del processo che porta alla realizzazione di una simile architettura documentale. Al di là di enti che hanno una grande capacità di spesa (mondo della difesa, mondo aeronautico, mondo automobilistico, grandi editori e pochi altri), tutti gli altri trovano tale modello troppo complesso da affrontare, troppo rigido, troppo costoso.

Anche un certo approccio, forse troppo accademico o elitario, della comunità SGML contribuisce a rendere questo mondo distante dall'utenza normale e quindi a limitarlo solo ad ambienti molto specifici. A ciò si somma un importante elemento commerciale: Microsoft sta tentando di affermare nel frattempo il suo standard RTF e quindi fa una guerra feroce alla comunità

SGML. In questo scenario l'arrivo del HTML è dirompente, in quanto l'utente della strada scopre improvvisamente che questi linguaggi complessi permettono di generare delle semplici pagine e di correlarle rapidamente tra di loro (è nato lo sviluppatore di pagine Web). La tecnologia viene in soccorso e si sviluppano rapidamente dei visualizzatori di codice HTML capaci, non solo di visualizzare il semplice documento, ma anche di spostarsi su un altro documento semplicemente cliccando sulla parola ipertestuale. È nato il browser Web.

Nel giro di poco tempo si affermano sul mercato due *browsers*: Mosaic³ e Netscape. Fino a questo punto Microsoft è rimasta alla finestra continuando a proporre il suo formato RTF ed il suo *browser* per la visualizzazione della sua versione ipertestuale, appunto l'Help di Windows.

Ma il mondo parla ormai Web e Microsoft capisce che la partita è persa e che rischia di perdere l'appuntamento con Internet, ormai diventata una realtà. Netscape si è intanto affermata e sta diventando una realtà per certi aspetti concorrenziale con la stessa Microsoft. A questo punto compie due mosse: compra il rivale di Netscape (NCSA Mosaic), gli cambia il nome in Explorer e, contemporaneamente (spalleggiata in questo dalla SUN), entra con forza nei comitati di sviluppo degli standard ed impone la nascita di un nuovo standard, denominato XML (eXtensible Mark-up Language).

³ Mosaic è un *browser* per il World Wide Web scritto al National Center for Supercomputing Applications (NCSA). Il suo sviluppo è incominciato nel 1992 ed è ufficialmente finito il 7 gennaio 1997. Il leader del gruppo che lo sviluppò, Marc Andreessen lasciò NCSA e fondò la Mosaic Communications Corporation che diventò Netscape Communications Corporation produttrice di Netscape Navigator. Spyglass, Inc. acquistò la tecnologia e i marchi da NCSA per produrre il suo *web browser* ma non usò alcun codice sorgente di NCSA Mosaic. Spyglass Mosaic fu poi venduto alla Microsoft che lo modificò e lo rinominò come Internet Explorer.

L'XML

Cerchiamo quindi di capire che cosa è XML: è solo un'operazione commerciale? o c'è qualcosa di più tecnico alle spalle? Ci sono entrambi: da un lato, si cancella la dizione SGML (oggi addirittura cancellato dalla lista del W3C-World Wide Web Consortium), dall'altro si integrano all'interno della struttura SGML due nuovi componenti fondamentali.

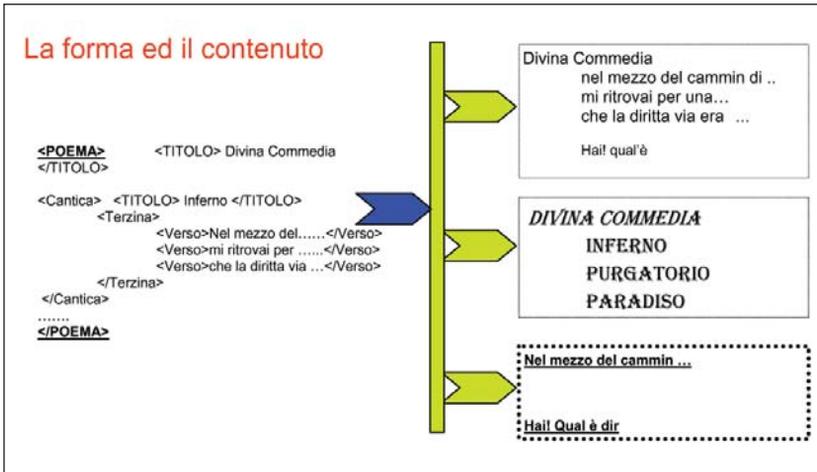
Il primo componente è il recupero dell' *hyperlink*, così come modellizzato, sviluppato e fatto crescere dall'ambiente HTML. Il secondo è l'eliminazione dell'obbligo della DTD: essa infatti diventa opzionale mentre si inserisce un nuovo concetto, quello di un documento "well formed" ossia "ben costruito". Di fatto questo concetto è molto semplice: esso consiste nel fatto che ogni *tag* aperta dev'essere anche correttamente chiusa.

A latere a queste varianti fondamentali, vengono anche eliminate molte "incrostazioni" che negli anni si erano aggiunte (o non erano state eliminate) allo standard SGML, quali i concetti di "omittag", di ricorsività illimitata di entità SGML e varie altre. Molte di queste erano già state di fatto eliminate dalla prassi applicativa dello standard.

Di fatto quindi SGML ed XML sono estremamente vicini, ma la mancanza dell'obbligo del modello formale (DTD) apre la strada all'uso integrato dell'SGML con i DBMS (come verrà meglio chiarito di seguito) e quindi una migliore interfacciabilità dei documenti con i *data base* tradizionali a tabelle. Questa interfacciabilità era di specifico interesse dei grandi gestori dei *data base* e quindi dei fornitori *hardware* a loro collegati (nel caso specifico SUN) che non a caso erano entrati nei comitati degli standard. Ovviamente, nel giro di poco tempo, i fornitori di *browser* dichiararono la loro compatibilità al nuovo standard ma qui si pose il grosso problema della visualizzazione dell'informazione, a tutt'oggi solo parzialmente risolto.

I browser

Occorre fare alcune considerazioni di base: il linguaggio SGML dichiarava esplicitamente una totale separazione tra contenuto e forma. In figura viene indicato un esempio del concetto di contenuto e forma.



Nell'esempio all'istanza "divina commedia" del modello "poema" vengono applicati tre stili: ognuno di essi evidenzia caratteristiche particolari dell'istanza. Non soltanto possono essere variati gli stili dei caratteri e il loro corpo, ma si può decidere se mostrare o non mostrare parti coerenti del documento.

I browser SGML puri, del periodo tra il 1986 ed il 1996, erano dotati di uno "style editor" necessario per creare la forma da applicare al contenuto. Gli stili, che di fatto erano dei *file* in SGML che descrivevano la formalizzazione della visualizzazione/medializzazione, non erano standard ma venivano realizzati tramite tecnologia proprietaria degli sviluppatori dei *browser*. Era nata una iniziativa di standardizzazione dello stile chiamata DSSSL (in gergo Diesel) ma non aveva avuto un grande successo. Questa iniziativa era poi di fatto morta a causa dell'HTML.

L'HTML infatti, per non complicare la vita al “popolo delle pagine Web”, non aveva mantenuto questa separazione rigida tra forma e contenuto, ma le *tag* HTML trattavano indifferentemente forma e/o contenuto. Ciò aveva semplificato notevolmente anche lo sviluppo dei *browser* HTML, che non doveva occuparsi di questa complessa gestione della esplicitazione di un qualsiasi modello di contenuto tramite una qualsiasi forma di stile. In questo senso i *browser* SGML erano estremamente più complessi e di fatto ad oggi mai raggiunti dalla tecnologia attuale. L'HTML può essere infatti considerato quasi un linguaggio procedurale (*specific markup language*) e come tale è infatti molto più semplice da trattare.

Tornando ora all'XML, i *browser* HTML si trovarono a dover trattare di nuovo un problema di separazione tra contenuto e forma, a cui non erano pronti a reagire, proprio perché progettati a fronte di un linguaggio procedurale. Si è quindi riesumato il concetto di stile, che, attraverso varie evoluzioni (CSS, XSL, XSLT), ha riproposto il problema della standardizzazione della formalizzazione. Come il suo predecessore DSSSL, ha avuto vita dura, in quanto i produttori di *browser* hanno tentato di risolvere questo problema riconvertendo in HTML il sorgente XML, interpretato alla luce del *file* di stile. Di fatto, ad oggi 2007, non è ancora stato riproposto un *browser* che risolva in modo nativo (ossia non passando attraverso HTML o suoi derivati) lo stile. Le varie iniziative che descrivono gli stili, gestiti dall'ente che ha standardizzato il Web, ossia il World Wide Web Consortium (W3C.org) verranno descritti nell'ultima parte di questo capitolo.

L'XML come elemento di congiunzione tra linguaggi di descrizione documentale e metodi di descrizione dei contenuti dei DBMS

Approfondiamo ora alcune caratteristiche di XML. Cerchi-

amo di comprendere il motivo per cui viene considerato un vantaggio la non obbligatorietà della DTD. Per capirne il significato occorre risalire all'ultima fase di sviluppo di SGML all'inizio degli anni '90, in quel periodo si tentava l'integrazione tra basi dati documentali e basi dati tabellari.

Il problema che è subito emerso è che i *data base* avevano ovviamente un loro modello di riferimento, formulato dall'analista durante la relativa fase di progettazione. Quando si tentava di integrarlo con il documento, occorreva che entrambi ragionassero secondo un modello comune. Per ottenere ciò, si ricorreva a complicate tabelle di correlazione il cui risultato non era sempre efficace.

L'XML, non avendo la necessità di esporre il suo modello, può tranquillamente accettare e ospitare il modello del data base integrandolo all'interno della sua struttura. Formuleremo in termini teorici questo concetto nel successivo paragrafo; ma ora prendiamo confidenza con le relazioni tra documento, modello del documento e base dati giocando un po' con le parole.

Per far ciò ci riferiremo al modello di ricetta già utilizzato all'inizio di questo percorso.

LA RICETTA

I sapori del sud

ALICI MARINATE (6 persone)

Ingredienti	Preparazione
Alici: 500 gr Limoni: 2 Clio di oliva: qb Aglio: 1 spicchio Prezzemolo: qb Menta: qualche fogliolina Sale: qb	* Frigate le alici di testa, dicscati, lavate più volte sotto acqua fredda, corrente. * Il pesce va sunitato su un piatto e coperto con succo di limone per 2 ore fino a che cambia superficialmente colore. * Trascorse 2 ore, svuotando il piatto dove sono disposte le alici, si faccia rotolare via il succo e lo si continui con battuto finissimo di prezzemolo, aglio, 2-3 foglie di menta, olio d'oliva e sale.



documento, database, dato?

o l'insieme di tutti e tre?

Poniamoci alcune domande: in questo contesto la ricetta è un documento o un *database*?

Forse nessuno dei due, questi dati possono essere espressi sotto forma di documento o possono essere classificati tramite un *data base*; nella buona sostanza è un insieme di dati.

Al di là del gioco di parole, la realtà è che i tre concetti (documento, *data base*, dato) sono confondibili poiché in assenza di fisicità cartacea hanno realtà fortemente assimilabili.

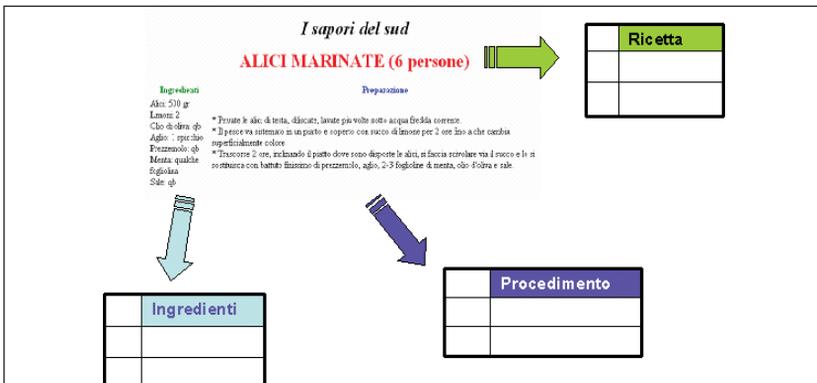
Introduciamo un terzo concetto: la ricetta è un processo? La risposta non può che essere positiva.

Proviamo a fare un po' di ordine. La moderna civiltà industriale si basa sui processi, i processi hanno bisogno dei dati, in molti casi i dati sono rappresentati tramite documenti. Tutto questo mette in ordine le parole con cui abbiamo liberamente giocato, ma pone il problema concreto di realizzare e di modellizzare questa interazione tra processo dati e documenti.

Quindi riassumendo graficamente:

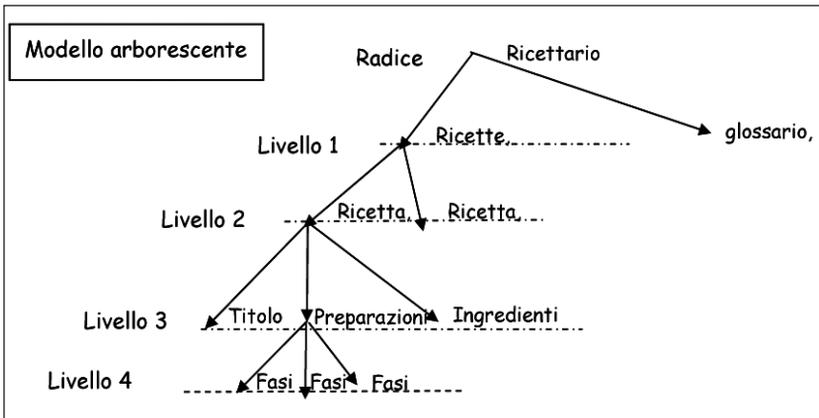
- ⇒ La ricetta è storicamente e fisicamente un documento
- ⇒ La ricetta contiene i dati degli ingredienti
- ⇒ La ricetta contiene il processo di cottura

Riesaminando quindi la ricetta, e guardandola come una base dati otteniamo le informazioni indicate nella figura sottostante.



presentato il modello dei rami principali della ricetta, in cui peraltro non tutti i rami sono stati esplosi né in profondità né nella loro molteplicità. Ad esempio sono state mostrate soltanto due ricette e di una sola di queste sono state mostrati i livelli inferiori (livello tre e livello quattro). Se avessimo dovuto presentare il modello reale, l'albero sarebbe stato molto più complesso ed articolato. È da notare che volutamente abbiamo lasciato la voce glossario separata e non sviluppata per preparare i successivi argomenti.

Questo modello è sviluppato a fronte di una DTD (ossia quelle riportata al precedente paragrafo) ed è possibile effettuare un controllo tra l'istanza del documento e il suo modello. La possibilità di un controllo è un concetto assolutamente lecito (anzi obbligatorio) in SGML mentre, come abbiamo spiegato, è solo opzionale in XML. Cerchiamo di capire graficamente perché questo è un vantaggio.

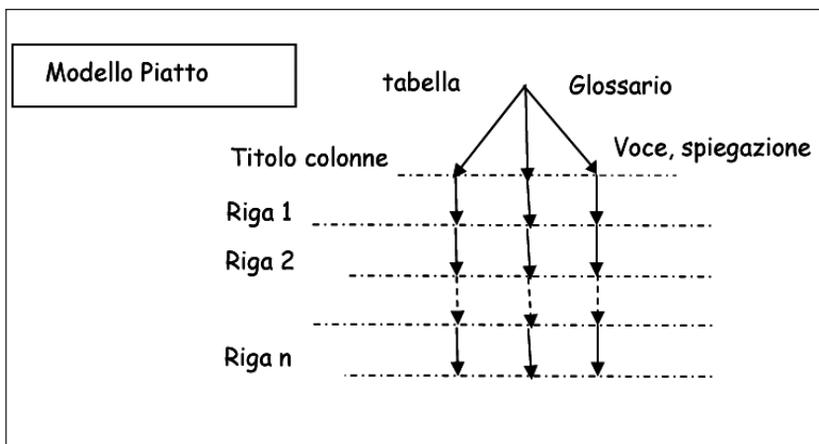


Osserviamo ora una struttura che apparentemente ha poche varianti, ma vedremo che queste sono sostanziali.

La struttura, definita a modello piatto, è tipica di una tabella nella quale il nodo padre (radice) è il nome della tabella, e il pri-

mo elemento della struttura piatta e tipicamente il titolo della colonna. Tutte le successive ripetizioni della stessa struttura rappresentano le righe della tabella. Possiamo ripetere lo stesso concetto parlando di campi e di *record*, ossia parlando di *data base*.

Ipotizziamo ora l'utilizzazione di questo tipo di struttura per realizzare la base dati di un glossario nel quale ogni riga sia composta da una voce e dalla sua spiegazione.

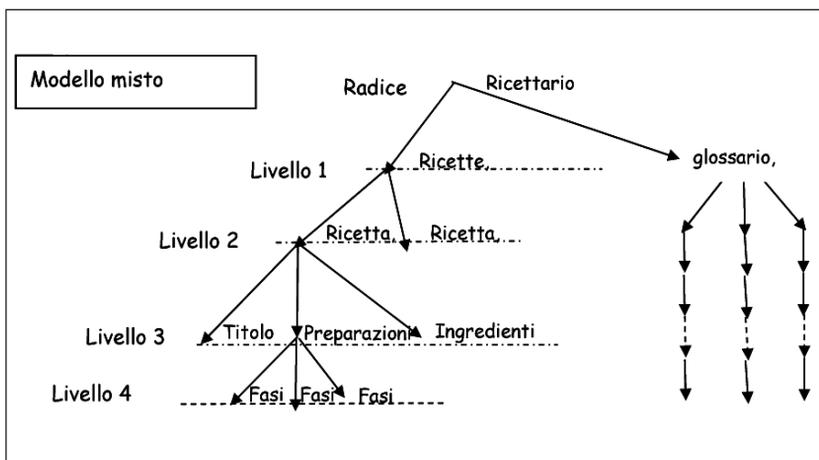


Come si può notare, e peraltro come era facile intuire, c'è una correlazione precisa tra una tabella ed un potenziale *data base*. Ovviamente il *data base* ha il suo modello.

Inseriamo ora detta tabella all'interno del mondo documentale arborecente precedentemente definito. Come possiamo vedere in figura, abbiamo applicato il modello tabella all'elemento glossario, lasciato precedentemente appositamente libero. Siamo quindi riusciti a realizzare un modello misto composto da una struttura documentale e da una struttura tabellare. Per realizzare ciò in SGML, il modello del *data base* avrebbe dovuto essere assolutamente identico al modello della tabella descritta documentalmente.

In XML invece, possiamo usare il modello documentale per la parte documentale ed il modello proveniente dal *data base* per la parte gabbellare. Quando il gioco si fa più complicato ed i modelli documentali devono integrare modelli di tabelle provenienti da *data base* estremamente complessi e sofisticati, questo approccio diventa vincente.

Non si sottovaluti questa problematica. Per rendersi conto di quanto questo sia importante, basta guardare con occhio critico qualsiasi pagina di un sito Web che fornisca elenchi di voli, alberghi, orari, parti, conti bancari, informazioni assicurative eccetera. Senza questo approccio il WEB avrebbe dovuto fare a meno (o almeno avrebbe avuto difficoltà a gestire) di una delle informazioni più importanti disponibili oggi: l'enorme quantità di dati accumulati negli ultimi cinquant'anni da *data base* di ogni tipo e dimensione.



La proliferazione dei linguaggi di marcatura

Da quanto annunciato nei precedenti paragrafi, dovrebbe risultare evidente l'importanza che i riguardi di marcatura hanno

oggi. La dimensione assunta dal Web ha portato al fatto che la standardizzazione dei linguaggi di marcatura sia, di fatto, spostata sotto l'ente che oggi ha il compito di promuovere, gestire e custodirne gli standard, il World Wide Web Consortium. È qui che oggi dobbiamo andar a cercare la continuazione delle iniziative prese nell'ambito della ISO. In questo senso il sito istituzionale del consorzio (W3C.org) è il punto di riferimento assoluto per chi intenda operare in questo settore.

I linguaggi di marcatura oggi non sono più esclusivamente connessi ai problemi documentali; l'evoluzione di XML, che ha portato alla integrazione con le basi di dati esistenti, ha anche dimostrato che il linguaggio di marcatura si prestava alla codifica di sorgenti informative diversificate. Sono quindi nati una serie di nuovi linguaggi per la descrizione di grafiche, musiche, voce, sincronizzazione multimediali, ecc. il cui comune punto di riferimento è l'uso di linguaggi di marcatura per la loro descrizione formale.

Per l'eventuale esame di questi standard consolidati o emergenti, si rinvia al sito istituzionale del World Wide Web Consortium (www.w3c.org).

Indicizzazione e classificazione: concetti generali

ROBERTO GUARASCI

Il *Dictionary of Archival Terminology* definisce una classe come: «Ensemble d'articles appartenant à un fonds regroupés en fonction d'une ou de plusieurs caractéristiques communes», laddove il termine «articles» sta per «Unité de base pour la cotation, le rangement et l'inventaire des documents d'archives [...]»¹. Più in generale, quindi, una classe è l'insieme degli oggetti raggruppati sulla base di un predeterminato livello di similarità degli attributi scelti come significativi in funzione delle finalità dell'attività di classificazione. Quando gli oggetti sono testi scritti gli attributi altro non sono che termini (nomi o sintagmi nominali) che rappresentano in maniera sintetica i concetti e le azioni espresse nei testi dai quali provengono². L'estrazione terminolo-

¹ FRANK B. EVANS, FRANCOIS J. HIMLY, PETER WALNE, *Dictionary of Archival Terminology*, ICA Handbooks Series, vol. 3, Munchen, 1984, *sub voce*. Dello stesso tenore «Group of things sharing common characteristics; members of a set», il *Glossary of Archival e Records Terminology* curato da RICHARD PEARCE MOSES, The Society of American Archivists, Chicago, 2005.

² Paolo Bisogno ritiene più appropriato non usare lo specifico termine di classe – pur invalso nell'uso – ma bensì parlare di «attributi intesi come elementi definitori di ciò che, per non usare il termine classe, si può chiamare gruppo, raggruppamenti. Sembra più fondato considerare la categoria come insieme degli oggetti che hanno in comune l'essenza costitutiva e la capacità di svolgere eguali funzioni. La suddivisione in classi, infatti,

gica finalizzata alla definizione dei termini di indicizzazione o descrittori informativi assume – quindi – un ruolo centrale nel processo di aggregazione delle classi. Nel composito mondo delle scienze del libro e del documento quest'ultima definizione e quella, correlata, di classificazione ha, comunque, conosciuto delle accentuazioni diverse.

Per Raffaele De Felice la **classificazione** «è una operazione quanto mai delicata, che deve consentire di ricondurre la molteplicità degli affari trattati ad un determinato numero di categorie gerarchicamente ordinate, in modo che gli atti prodotti quotidianamente vengano a sedimentarsi secondo un ordine logico che rispecchi, storicamente, lo sviluppo e l'evoluzione dell'attività svolta, nonché le finalità raggiunte»³.

Nel *Traitè* di Otlet, per contro, «On entend par **classement** bibliographique, l'art de disposer les ouvrages d'après leur matière et par **classification**, le tableau ou les tables qui disposent les connaissances dans l'ordre où doivent l'être les ouvrages eux-mêmes»⁴.

Anche Eugenio Casanova dedica ampia parte del suo testo all'argomento. La buona classificazione, afferma, «[...] serve al rinvenimento degli atti quanto gli indici e gli inventari. Se per ogni genere di affari si avesse un titolo preciso, e per ogni affare, un fascicolo con le carte disposte in ordine di tempo, le ricerche sarebbero possibili senza altri soccorsi»⁵. Negli stessi termi-

presenta una indubbia utilità quando usata per indagini e rilevamenti sociologici, statistici, economici, ecc. pur essendo soggetta ad eccessiva arbitrarietà, mentre la categoria così intesa può costituire un fondamento chiaro». PAOLO BISOGNO, *Il Futuro della Memoria, elementi per una teoria della documentazione*, Franco Angeli, Milano, 1994, p. 107.

³ R. DE FELICE, *La classificazione degli atti negli archivi moderni*, in *Rassegna degli Archivi di Stato*, n. 2 (1964), pp. 10-11.

⁴ PAUL OTLET, *Traitè de Documentation, le Livre sur le livre*, Bruxelles, Mundaneum, 1934, p. 379.

⁵ E. CASANOVA, *Archivistica*, Lazzeri, Siena, 1928, p. 141.

ni si esprimono anche – nel 1806 – *le Istruzioni ai vice Prefetti del Regno d'Italia* che, nel prescrivere l'obbligo del registro di protocollo per tutte le Pubbliche Amministrazioni, legano questa prescrizione, per la prima volta, alla necessità di correlare l'indice numerico ad un sistema di classificazione capace di consentire la collocazione degli atti in un reticolo fisico/logico che, oltre a categorizzare il sapere, permettesse, praticamente, il recupero delle unità documentarie⁶. La classificazione della quale parla il Casanova posiziona l'unità archivisticamente rilevante – serie o fascicolo – in una strutturazione ordinatoria giuridico formale discendente dalle specifiche funzioni del soggetto produttore delle quali costituisce una astrazione. L'ordine è costruito sulla base della concretizzazione pratica delle stesse funzioni e non in base ai contenuti o alla segmentazione dell'azione svolta dagli attori. Non vi è una strutturazione precostituita del sapere all'interno della quale trovano posto le unità informative né vi è la possibilità di costruire, estemporaneamente, *corpus* di contenuti sulla base di bisogni emergenti. Lo schema di classificazione è dato dalle funzioni svolte dal soggetto produttore nella sua interazione con il contesto nel quale si è trovato ad operare. Il singolo elemento informativo – il documento o parte di esso – ha realtà solo in quanto componente di un tutto archivisticamente rilevante: il fascicolo. Questa tipologia di classificazione, messa in essere

⁶ «La decima (colonna del protocollo) contiene le annotazioni per trovare le carte riposte in archivio: essa è suddivisa in due. Nella prima si marca il titolo, o sia la materia a cui appartiene la carta; nella seconda il numero di fascicolo» [...] «Perché l'archivio riunisca il doppio scopo di conservare con sicurezza le carte e di offrire il modo di rinvenirle prontamente, vi saranno esse distinte per materia. Ogni materia sarà collocata in archivio per ordine alfabetico, cominciando da un punto della camera e proseguendo gradatamente. Sul dorso di ciascuna cartella sarà scritto il titolo della materia, per esempio censo, acqua, strade, amministrazione dei comuni, imposta distrettuale, ecc[...]»: *Istruzioni del Ministero degli Interni ai vice-prefetti del Regno d'Italia*, n. 16046 del 1 gennaio 1806.

al momento della registrazione di protocollo, prevede – come subordinati – metodi di ordinamento⁷ quali quello cronologico o alfabetico, basati su una voce indice che – di solito – non discende da una analisi concettuale del testo ma bensì dalla maggiore riconducibilità del termine prescelto alla funzione esercitata (l'intestatario di un fascicolo personale, la località di esazione di una imposta, ecc.). È comunque da dire che questo *modus operandi* che la dottrina italiana ha, a volte, accentuato non sempre trova riscontro nei modelli europei e nella più recente strutturazione degli standard Iso che prevedono la possibilità di derivare l'intitolazione dei livelli divisionali da procedure standardizzate di estrazione terminologica⁸ e, quindi, dall'analisi dei contenuti.

«Ma il titolare dell'ufficio di registrazione era scarso ai bisogni delle indagini amministrative e culturali insieme, e difficile da applicare agli archivi antichi. E poi sempre nuovi problemi, sempre nuove questioni si presentavano all'esame degli studiosi, dopo che gli atti non venivano più considerati soltanto sotto l'aspetto giuridico, ma altresì sotto quello scientifico. Nacque, pertanto, in mente all'ordinatore di disporre gli atti affidati alla sua

⁷ «Presso la maggior parte degli archivi correnti il termine classificazione viene interpretato quale indicativo delle operazioni di riunione di un atto ai suoi precedenti o di formazione di un nuovo fascicolo con l'attribuzione di un numero o di una sigla necessaria ai fini della ricerca. Manca assolutamente la valutazione della posizione che il fascicolo stesso deve assumere nell'ambito dell'archivio, per assolvere una funzione giuridico-amministrativa connessa con le competenze specifiche dell'ufficio nella cui storia si inquadra. La classificazione viene, pertanto, confusa con l'operazione di ordinamento per la quale si adottano metodi vari, dall'alfabetico al cronologico, dal geografico al numerico, a seconda delle preminenti esigenze di ricerca degli atti e per corrispondere alle pressanti richieste provenienti dal funzionario che tratta l'affare ed alle necessità di conservare più o meno ordinatamente i fascicoli». R. DE FELICE, *La Classificazione*, cit., pp. 9-10.

⁸ *Model Requirements for the Management of Electronic Records*, <<http://www.dlmforum.eu.org>>. Cfr. anche, in tal senso, ISO 15489/01.

custodia in tanti gruppi quanti secondo lui potessero essere gli scopi, i soggetti degli studi e delle ricerche: sicché, in un attimo, ogni domanda potesse essere soddisfatta»⁹. Tutta la letteratura archivistica ha sempre ritenuto esecrabile questo metodo¹⁰ – comunemente definito per materia – per il semplice motivo che, distruggendo l'originaria composizione delle serie archivistiche, scompaginava la strutturazione giuridico-formale della documentazione consegnando quest'ultima all'arbitrarietà dell'operatore. Parimenti non condivisibile appariva l'utilizzazione della Classificazione Decimale Universale (CDU) e del sistema Dewey, pur riconoscendo la validità dei presupposti degli esperimenti, anche su larga scala, tentati in molti paesi europei. Queste opposizioni avevano certamente ragion d'essere nella misura in cui un tipo siffatto di aggregazione comportava – di fatto – la completa scompaginazione dell'archivio ed una sua ricomposizione fisica utilizzando criteri sostanzialmente artificiali e non codificati¹¹. È opportuno precisare che metodi concettualmente simili – ad esempio il *block numeric system* – sono largamente usati nei paesi nordamericani anche se alcune criticità e mutamenti di rotta stanno emergendo negli ultimi decenni¹². Queste posizioni divergenti, in parte legate al momento contingente nel quale vengono formulate, fanno registrare, nel corso degli anni, dei tentativi più o meno timidi di trovare delle soluzioni capaci di coniugare la teoria archivistica con una logica funzionalmente più rispondente alle richieste di accesso ai contenuti.

⁹ E. CASANOVA, *Archivistica*, Lazzari, Siena, 1928, p. 141, p. 210.

¹⁰ «Ogni tentativo in proposito non può che essere artificioso, instabile, insufficiente e, del tutto, soggettivo». *Ivi*, p. 210.

¹¹ Parlando dei metodi di ordinamento da adottare Casanova dimostra di conoscere il lavoro di Paul Otlet e dell'Istituto Internazionale di bibliografia di Bruxelles che espressamente cita e al quale – pur non condividendolo – dedica alcune pagine del suo lavoro.

¹² Cfr. PAUL SABOURIN, *Functional File Classification*, march 2001, <www.Rmicanada.com>.

Nel 1953 la Commissione sulla Riforma della Pubblica Amministrazione propone una “classificazione ideologica”. Essa «prende in considerazione per unità di misura la divisione ministeriale quale unità mobile da sistemare come si vuole e si deve, senza toccare gli atti che la compongono, come le unità di un esercito che si muovono, si raggruppano, si spostano, senza far loro perdere né forze né aspetto»¹³.

Questo sistema, previsto per l'Archivio Centrale dello Stato, si ispirava a criteri molto vicini alla classificazione decimale universale ma lasciava aperto ed irrisolto il problema della classificazione interna dei singoli fondi e della gestione del livello di aggregazione degli archivi correnti in quanto postulava il concetto di “archivio divisionale o ministeriale” come unità archivistica primaria o livello divisionale primario.

Sullo stesso argomento Muller, Feith e Fruin, calati in un altro contesto, pur ribadendo il principio dell'ordinamento originario non ignorano altre prospettive: «E qui cade opportuna un'altra osservazione: l'antica organizzazione dell'archivio si deve conservare per i suddetti motivi; ma entro questi confini non è necessario né perfino desiderabile, di accettare l'antico ordinamento degli atti in tutti i particolari. Le antiche direzioni archivistiche che nei loro inventari miravano ad uno scopo diverso dal nostro fecero, ancora nel secolo XVIII, un lavoro che per i nostri bisogni non importa affatto; in questi casi abbiamo, non solo il diritto, ma anzi il dovere di cambiare cotale loro superficiale ordinamento»¹⁴. Riecheggiano – in queste affermazioni – anticipazioni della concettualità di “corpo archivistico” del Brenneke che cerca una difficile conciliazione tra le posizioni con-

¹³ PRESIDENZA DEL CONSIGLIO DEI MINISTRI, *Stato dei Lavori sulla Riforma della Pubblica Amministrazione*, vol. III, Roma, 1953, in R. DE FELICE, *La Classificazione*, cit., p. 14.

¹⁴ S. MULLER-J.A. FEITH, R. FRUIN, *Ordinamento e Inventario degli archivi*, trad. it., Unione Tipografico Editrice Torinese, Torino, 1908, p. 29.

trapposte mirando alla costruzione di un ordinamento “ideale” che salvaguardi sia il principio della rispondenza alla strutturazione funzionale sia quello della modifica di un ordinamento originario scorretto ed artificioso¹⁵. Del resto l’ordinamento per materie è la norma negli archivi francesi post rivoluzionari e il *Tableau Systématique* redatto dal Daunou nel 1811 si limita ad aggiungere alle «prime quattro lettere dell’alfabeto utilizzate dal Camus per gli archivi delle Assemblee [...] le altre nelle quali si dovevano inquadrare gli archivi correnti delle amministrazioni e quelle degli uffici cessati»¹⁶. Tutto ciò, nonostante i tentativi del capo-sezione degli archivi nazionali Natalis de Wailly che, dopo aver tracciato un *project de classement* chiaramente ispirato alle metodologie del Tableau, continua il tradizionale ordinamento per materie ed il suo celebre richiamo al “*respect des fonds*” appare più un monito per il futuro che un’idea operativa per il suo tempo.

Oltre alla vincolante connotazione giuridico-formale dei complessi archivistici la remora maggiore ad adottare sistemi di classificazione che miravano a costruire partizioni non presenti nell’originaria strutturazione delle carte era ed è data, sicuramente dall’impossibilità – se non a prezzo di grandi difficoltà – di riprodurre il processo inverso. Non a caso Otlet quando parla del dossier non postula la necessaria scompaginazione dei complessi di provenienza ma bensì la realizzazione di sintesi concettuali su supporti standardizzati, e per la documentazione archivistica riafferma la necessità del rispetto del principio di provenienza¹⁷.

¹⁵ ADOLF BRENNEKE, *Archivistica*, trad. it. a cura di R. Perrella, Milano, 1968, p. 43.

¹⁶ SALVATORE CARBONE, *Gli Archivi Francesi*, Quaderni della Rassegna degli Archivi di Stato, n. 3, Roma, 1960, p. 38.

¹⁷ Anni dopo l’amministrazione pubblica belga tenterà una trasposizione della CDU nell’ambito della documentazione amministrativa mediante la realizzazione del sistema Dècasépel. «C’est surtout en vue des administrations

Le singole unità documentarie – nella prospettiva di Paul Otlet – vengono disaggregate in unità informative per poi essere ricomposte in una nuova realtà «[...] che assume carattere unitario attraverso il reticolo delle relazioni. In altre parole, nel concetto si ha una identità tra l'esistere del dato e le sue relazioni potenziali e pertanto le unità informative sono da considerare come gli elementi interni di quella sola realtà che è l'universo documentaristico, proiezione della realtà totale esterna»¹⁸. La classificazione avviene sulle astrazioni informative concretizzate nelle fiches lasciando inalterati i complessi di provenienza. È – per usare un termine moderno – virtuale. Centrale e innovativo è il concetto della “fiche” che segmenta logicamente i contenuti informativi, supera la partizione puramente fisica dei formati e ne permette, così, una utilizzazione in più contesti, anticipando il documento virtuale e delocalizzato della società dell'informazione¹⁹. Questo approccio – poco considerato ai suoi tempi – apre la strada ad un possibile superamento della opposizione tra i si-

communales dont les relations avec l'Etat et la Province évoluent constamment que l'organisation d'une classification doit être conçue d'une façon suffisamment large pour pouvoir tenir compte de cette évolution. En effet, il faut en premier lieu que ce soit une classification pratique, don't les fonctionnaires doivent pouvoir se servir dans leurs occupations journalières comme d'un instrument efficace et indispensable». *Le Systeme Decasepel: manuel pour l'organisation et l'emploi d'une documentation courante et retrospective, conçu d'après la Classification decimale universelle*, publié sous la direction de A. BUCQUE, Sepeli, Gand, 1949, p. III.

¹⁸ PAOLO BISOGNO, *Il Futuro della memoria. Elementi per una teoria della Documentazione*, Franco Angeli, Milano, 1994, p. 71.

¹⁹ «Nous possédons dans le Répertoire un instrument nouveau approprié à l'enregistrement analytique et monographique des faits, des idées, des informations. Le système s'est perfectionné par des fiches divisionnaires variées de formes et de couleurs disposées de manière à exprimer extérieurement le cadre de la classification adoptée et à réduire ainsi au minimum le temps de la recherche»: OTLET, *Traité*, cit., p. 386.

stemi di classificazione facendo venir meno la necessarietà della scompaginazione dell'ordine fisico in favore di una aggregazione temporanea e virtuale centrata sulle richieste informative. Trasferendo concettualità e problemi nell'universo del documento informatico ne deriva che, quand'anche venga ricostituita o predefinita una strutturazione formale dei *files* sulla base di una organizzazione logica discendente dalle funzioni amministrative questa altro non è che un "Avatar"²⁰, una finzione mediatica giacché, in deroga e indipendentemente da quella, è possibile ogni aggregazione e relazione sulla base di una semplice *query*. «Il punto centrale dell'ordine originario, naturalmente, è che la collocazione fisica di un documento contestualmente ad un altro documento veicola un contenuto intellettuale. Poiché i documenti elettronici non esistono come entità fisiche distinte e poiché l'effettivo immagazzinamento dei segnali elettronici che costituiscono un documento raramente conserva qualche connessione con un documento che viene mostrato su di un monitor o stampato, la collocazione fisica dei bit e delle porzioni di documenti elettronici non richiede né veicola contenuto intellettuale. Sebbene mantenere l'ordine originario dei documenti elettronici non abbia alcuna conseguenza, documentare le relazioni logiche è assolutamente cruciale»²¹.

²⁰ L'*avatar* è un'immagine scelta per rappresentare la propria utenza in community, luoghi di aggregazione, discussione, o di gioco on-line. La parola, che è in lingua sanscrita, è originaria della tradizione induista, nella quale ha il significato di incarnazione, di assunzione di un corpo fisico da parte di un dio: per traslazione metaforica, nel gergo di internet si intende che una persona reale che scelga di mostrarsi agli altri, lo faccia attraverso una propria rappresentazione, un'incarnazione: un *avatar* appunto.

²¹ CHARLES M. DOLLAR, *La teoria e la prassi archivistica di fronte all'informatica. Alcune considerazioni*, in: ODDO BUCCI (a cura di), *L'Archivistica alle soglie del 2000*, Atti della conferenza internazionale di Macerata, 3-8 settembre, 1990, Pubblicazioni dell'Università di Macerata, Ancona 1992, p. 306.

L'attenzione viene quindi spostata sui contenuti e su ciò che, nei contenuti, permette di costruire o mantenere relazioni logiche siano esse contenute nel testo o nei metadati²²: i termini di indicizzazione. Va, comunque, fatta una premessa. Nella costruzione dei reticoli logici di significati esistono, necessariamente, due opzioni. La prima più squisitamente documentale che li determina sulla base di una partizione a priori del sapere che è dinamica in quanto funzione della domanda informativa ed una seconda, statica, che estrae le voci di indicizzazione avendo come ottica il mantenimento di relazioni funzionali tra i documenti. La seconda può contenere la prima come subset subordinato, più difficilmente il contrario.

La norma ISO 5963/85²³ definisce in maniera univoca e standardizzata le procedure per l'estrazione dei termini di indicizzazione dai documenti, specificando, in apertura, che essa «[...] descrive le procedure raccomandate per l'analisi dei documenti, la determinazione del loro contenuto e la scelta dei termini d'indicizzazione appropriati, [...] Descrive anche le tecniche generali per l'analisi dei documenti applicabili a tutti i casi di indicizzazione. Le tecniche descritte dalla norma sono adottabili nell'ambito dell'indicizzazione 'umana' e non automatica» La suc-

²² I metadati sono informazioni strutturate che descrivono, esplicano, localizzano qualsiasi elemento utile per recuperare, usare o gestire una risorsa informativa. Un record di metadati consiste in un insieme di attributi o elementi, necessari a descrivere tale risorsa. La strutturazione del *record* è definita nello standard ISO/IEC 11179 *Metadata registry*, che si compone di varie parti: 11179-1: *Framework*, 11179-2: *Classification*, 11179-3: *Registry metamodel and basic attributes*, 11179-4: *Formulation of data definition*, 11179-5: *Naming and identification principles*, 11179-6: *Registration*. Per la definizione del *set* di elementi descrittivi necessari, in ambito testuale è largamente usato lo standard Dublin Core, ANSI/NISO Z39.85 - 2007 *The Dublin Core Metadata Element Set*.

²³ UNI ISO 5963/1985, *Metodi per l'analisi dei documenti, la determinazione del loro soggetto e la selezione dei termini di indicizzazione*.

cessiva norma ISO 15489/01 “Records Management” però, allorché parla dell’attività di indicizzazione nei sistemi di gestione documentale, nell’affermare che l’attività può essere – ovviamente – condotta sia manualmente che automaticamente, esplicitamente rinvia – *erga omnes* – alla ISO 5963 ricomprendendola nel più generale quadro della gestione documentale²⁴ e affrancandola – sostanzialmente – dalla limitazione “umana”.

Più nel dettaglio, l’attività di indicizzazione, è definita nella citata norma come l’insieme delle procedure necessarie per l’analisi dei documenti, la determinazione del loro contenuto e l’individuazione dei termini di indicizzazione; essa prefigura – quindi – una serie di azioni che mirano a riassumere la significatività semantica dell’unità documentale in una stringa di testo o in uno o più termini singoli rappresentativi dell’intero contesto. Primo *step* metodologico di tale attività è l’analisi concettuale, che consiste, proprio «nell’esame del documento, nella determinazione del suo contenuto concettuale e nell’identificazione dei concetti che si ritiene possano utilmente rappresentare quel contenuto in uno strumento di [...] recupero»²⁵. Lo specifico richiamo, nello standard del Record management, alla procedura standardizzata di estrazione dei termini di indicizzazione assume però un’altra e ben più rilevante valenza concettuale in quanto salda – unificandole – una metodologia di definizione dell’indice di classificazione, tradizionalmente basata sull’astrazione delle funzioni, con una specificatamente orientata all’analisi concettuale del contenuto informativo. Il rapporto tra funzioni e contenuto informativo viene quindi reimpostato su basi diverse e convergenti.

²⁴ 9.5.4 Indexation - L’indexation peut être pratiquée manuellement ou générée automatiquement. Elle peut intervenir à différents niveaux au sein d’un système d’archivage. Un guide de l’indexation est donné dans l’ISO 5963, *Documentation - Méthodes pour l’analyse des documents, la détermination de leur contenu et la sélection des termes d’indexation*.

²⁵ Cfr. anche *Guida all’indicizzazione per soggetto/ GRIS*, AIB, Roma 1996.

«In questa prospettiva – afferma Dollar – una questione essenziale per il servizio di consultazione archivistica riguarda, pertanto, le modalità per organizzare un servizio di consultazione orientato alla domanda che soddisfi le aspettative dei ricercatori e abbia bisogno di informazioni registrate, non di semplici documenti. Questo approccio presume che gli archivisti guardino in modo nuovo al problema e si applichino per sviluppare una strategia relativa al servizio di consultazione per l'informazione elettronica che comporti la modifica delle tradizionali modalità di rapporto con i ricercatori. Un aspetto di questa strategia dovrebbe riflettere l'impegno a trasformare gli archivi da istituti di deposito a strutture di intermediazione»²⁶.

²⁶ CHARLES M. DOLLAR, *Archivistica e Informatica*, Università di Macerata, Ancona 1992, pp. 68-69.

La terminologia

GIOVANNI ADAMO

*L'origine e il progresso delle nostre conoscenze
dipendono interamente dalla maniera
con la quale ci serviamo dei segni.*

Condillac

Terminologia è voce registrata – fino agli ultimi anni del XX secolo e in tutta la produzione lessicografica europea – per fare riferimento all'«insieme dei termini propri di una scienza, di un'arte, di un autore». Questa è senza dubbio l'accezione più diffusa, nonostante si sia in presenza di un termine composto con l'elemento formante *-(o)logia*, che esprime il significato di «studio sistematico, trattazione, teoria, dottrina». Solo da alcuni anni i dizionari hanno accolto anche il riferimento più specialistico al «sette degli studi linguistici che si occupa degli aspetti teorici della formazione e dell'uso sistematico dei termini nell'ambito di una scienza o di una disciplina». Peraltro, già nei primi decenni del XIX secolo, un repertorio dedicato al lessico tecnico e alle sue implicazioni etimo-filologiche definiva la *terminologia* come la «dottrina dei termini o dei vocaboli propri di una data arte o scienza»¹. Altrettanto significativa è una segnalazione di Alain Rey:

¹ MARCO AURELIO MARCHI, *Dizionario tecnico-etimologico-filologico*, Milano, Giacomo Pirola, 1828-1829, 2 vol., xxiv-619-745 p.

«L'uso moderno del concetto oggettivo [di "terminologia"] sembra determinarsi in Inghilterra. La definizione di William Whewell [epistemologo e moralista inglese, 1794-1866], nel 1837, attribuisce alla parola il suo valore scientifico, ancora limitato a un gruppo di scienze: "sistema di termini usati nella descrizione degli oggetti della storia naturale". La connessione dei concetti di "sistema", di "oggetto" e di "scienza" a quello di "termine" dà a questa definizione troppo dimenticata un'attualità sorprendente»². Il periodo più ricco di fermenti determinanti per il concepimento della moderna teoria terminologica è, però, da collocare nei primi decenni del XX secolo, connotati da una forte espansione dell'attività industriale e da una vivace attività culturale. Sono gli anni in cui nascono e si affermano il Circolo di Vienna³ e il Circolo lin-

² «L'emploi moderne de la notion objective s'élabore, semble-t-il, en Angleterre. La définition de William Whewell, en 1837, donne au mot sa valeur scientifique, encore restreinte à un groupe de sciences: "système des termes employés dans la description des objets de l'histoire naturelle". L'articulation des notions de "système", d'"objet" et de "science" à celle de "terme" donne à cette définition trop oubliée une actualité surprenante», ALAIN REY, *La terminologie. Noms et notions*, Paris, Presses Universitaires de France, 1992², pp. 6-7.

³ Il Circolo, attivo tra il 1922 e la fine degli anni Trenta, raccolse attorno a Moritz Schlick (filosofo e fisico) molti tra i più prestigiosi intellettuali e studiosi dell'epoca, di formazione culturale diversa. Tra essi: Rudolf Carnap (filosofo), Philipp Frank (fisico teorico), Kurt Gödel (logico), Hans Hahn (matematico), Otto Neurath (economista e sociologo), Friedrich Waismann (filosofo). Anche altri insigni studiosi, tra i quali Karl R. Popper, John Von Neumann, Carl G. Hempel, Federigo Enriques e Hans Reichenbach, ebbero contatti o parteciparono occasionalmente alle riunioni del Circolo, che si fece conoscere dall'opinione pubblica internazionale con un manifesto intitolato *Wissenschaftliche Weltauffassung. Der Wiener Kreis*, pubblicato da Hahn, Neurath e Carnap nel 1929. La produzione più significativa del Circolo fu affidata alle pagine del periodico «Erkenntnis» (nelle ultime annate «The Journal of Unified Science»), pubblicato tra il 1930 e il 1940, in collaborazione con un gruppo di studiosi berlinesi organizzata da Reichenbach.

guistico di Praga⁴. Uno degli elementi di maggiore interesse consiste nell'ambizioso progetto del Circolo viennese di rifondare la conoscenza umana su basi esclusivamente logiche e empiriche, attraverso un linguaggio unificato della scienza⁵. È difficile pensare che la formazione e il pensiero dell'ingegnere austriaco Eugen Wüster (1898-1977) non abbiano risentito di influssi tanto profondi e innovativi. Ma occorre attendere il 1979, anno in cui viene pubblicata postuma la sua *Einführung in die allgemeine Terminologielehre und terminologische Lexikographie*⁶, per vedere defini-

⁴ La sua prima riunione si tenne il 6 ottobre 1926, sotto la presidenza di Vilem Mathesius. Tra i membri fondatori: Roman Jakobson, Nikolaj S. Trubeckoj e René Wellek. Da ricordare il grande impulso dato dal Circolo praghese agli studi di linguistica sincronica e descrittiva.

⁵ Cfr. *Foundations of the Unity of Science. Toward an International Encyclopedia of Unified Science*, edited by Otto Neurath, Rudolf Carnap, Charles Morris, Chicago-London, The University of Chicago Press, 1971 (Third impression), 2 voll. Nei *Travaux du IX^e Congrès International de Philosophie. Congrès Descartes. L'Unité de la Science: la Méthode et les méthodes*, publiés par les soins de Raymond Bayer, Paris, Hermann, 1937, sono raccolte, tra le altre, due comunicazioni di grande interesse per il nostro argomento: RUDOLF CARNAP, *Einheit der Wissenschaft durch Einheit der Sprache*, pp. 51-57; OTTO NEURATH, *Prognosen und Terminologie in Physik, Biologie, Soziologie*, pp. 77-85.

⁶ L'opera, pubblicata in due volumi a Vienna-New York per i tipi di Springer e subito tradotta in francese (*Introduction à la théorie générale de la terminologie et à la lexicographie terminologique*, Québec, Université de Laval, 1979), era stata anticipata da EUGEN WÜSTER, *Die allgemeine Terminologielehre, ein Grenzgebiet zwischen Sprachwissenschaft, Logik, Ontologie, Informatik und den Sachwissenschaften*, «Linguistics. An International Review», 1974, n. 119, pp. 61-106. Ma già EUGEN WÜSTER, *Die vier Dimensionen der Terminologiearbeit*, «Mitteilungsblatt für Dolmetscher und Übersetzer», 15, 1969, n. 2, pp. 1-6; n. 5, pp. 1-5, riproponeva la relazione presentata a un Colloquio tenutosi presso l'Istituto per traduttori e interpreti di Gernersheim (Università di Mainz) nell'ottobre dell'anno precedente. Vi si definivano le quattro dimensioni della prassi terminologica: il settore specialistico, le lingue, l'obiettivo (trattare documenti, utilizzare una terminologia, condurre ricerche su un campo concettuale), e il grado di astrazione.

tivamente elaborati il nucleo teorico dello studio delle terminologie specialistiche, di cui lo stesso Wüster è il codificatore. Ma è opportuno sottolineare come egli parlasse ancora di *terminologische Lexikographie*, quasi a preservare il legame che, per alcuni secoli, si era instaurato in via pressoché esclusiva tra la prassi lessicografica, più o meno attenta – occorre riconoscerlo –, e le nomenclature dei settori specialistici.

Prospettive teoriche della terminologia

A partire dagli anni Settanta del XX secolo, in concomitanza con la riflessione teorica elaborata da Wüster, si è andata consolidando una *tripartizione di orientamenti e tendenze*, che riflette prospettive e finalità diverse, e che ha avuto conseguenze significative anche dal punto di vista del riconoscimento dello statuto scientifico-disciplinare della terminologia.

Helmut Felber individua i *tre orientamenti*, considerandoli come indirizzati rispettivamente ai settori specialistici, alla filosofia, e alla linguistica⁷. Il *primo*, caratterizzato dall'interdisciplinarietà, è quello espresso nella *Teoria generale della terminologia* elaborata da Wüster. Vi si pongono al centro dell'attenzione: il concetto e le sue relazioni con le altre unità del medesimo sistema concettuale, la corrispondenza tra concetto e termine, l'assegnazione dei termini ai concetti. Il *secondo* orientamento, piuttosto vicino al primo, preferisce porre l'accento sulle classificazioni dei concetti in base a categorie filosofiche. Esso si propone, quindi, di studiare gli aspetti connessi con le teorie della classificazione, che lo accomunano alle problematiche della documentazione⁸. Il *terzo*, orientato agli studi linguistici, si fonda

⁷ HELMUT FELBER, *Terminology Manual*, Paris, Unesco-Infoterm, 1984, p. 47.

⁸ Cfr. INGETRAUT DAHLBERG, *Grundlagen universaler Wissensord-*

sul principio che le terminologie, essendo sottocomponenti del lessico di un linguaggio speciale, rientrano nei sottocodici delle varie lingue.

Qualche anno più tardi, Teresa Cabré riprende uno schema già proposto da Pierre Auger e delinea un quadro più complesso e ricco di sfumature, articolato anch'esso in *tre grandi tendenze*, che considerano la terminologia dal punto di vista del sistema linguistico, da quello della traduzione, e, infine, da quello della pianificazione linguistica⁹. La *prima* tendenza, linguistico-terminologica, si caratterizza per la propensione alla standardizzazione concettuale e denominazionale ed è rappresentata da tre scuole, ritenute 'classiche': la scuola di Vienna, che ha come figura di riferimento Eugen Wüster, considera il lavoro terminologico in vista della normalizzazione di concetti e termini, e vede negli specialisti dei differenti settori i responsabili delle diverse terminologie; la scuola di Praga, rappresentata da Lubomír Drozd e ispirata agli studi di linguistica funzionale del noto Circolo linguistico, si occupa della descrizione strutturale e funzionale dei linguaggi settoriali, considerati come uno «stile professionale»; la scuola di Mosca (D. Lotte), molto attenta alle posizioni di Wüster, si interessa alla normalizzazione di concetti e termini nella cornice dei problemi connessi con la situazione di plurilinguismo dell'ex Unione Sovietica. La *seconda* tendenza, orientata alla traduzione, è molto sviluppata nei Paesi francofoni ed è alla base dei lavori terminologici prodotti dai grandi organismi internazionali multilingui, con l'intento di stabilire equivalenze terminologiche tra le diverse lingue, per orientare i traduttori nel loro lavoro. La *terza* tendenza, quella *aménagiste* (secondo la denominazione coniata nel Québec), considera la terminologia all'in-

nung. Probleme und Möglichkeiten eines universalen Klassifikationssystem des Wissens, München, Verlag Dokumentation Saur KG, 1974.

⁹ M. TERESA CABRÉ, *La terminología. Teoría, metodología, aplicaciones*, Barcelona, Editorial Antártida/Empúries, 1993, pp. 39-41.

terno di un più ampio processo di pianificazione linguistica¹⁰. La terminologia, ritenuta uno degli elementi fondamentali di ogni codice comunicativo, diviene strumento per la rivitalizzazione delle lingue minoritarie, ma anche per il rilancio della funzione comunicativa internazionale di lingue stabili e consolidate, che si trovano però a dover fronteggiare la concorrenza pervasiva delle lingue di Paesi in situazione di forte predominio culturale, economico o tecnologico. Per raggiungere questo obiettivo, si ritiene indispensabile un progetto d'azione sistematico e strategico, sostenuto da adeguati interventi legislativi.

Su posizioni diverse si trova Juan C. Sager, che non riconosce alla terminologia lo statuto di disciplina autonoma, pur sottolineando l'importanza e il valore dell'attività terminologica e la necessità della sua presenza nei percorsi formativi¹¹. D'altro canto, Teresa Cabré affermava di recente la necessità di una rivisitazione dei fondamenti teorici della terminologia, sottolineando l'inadeguatezza alla realtà attuale dell'istanza riduzionista e uniformatrice, già presente nella teoria di Wüster e amplificata da alcune successive prese di posizione di studiosi a lui vicini. Inoltre, ribadendo l'esigenza di considerare più attentamente i riflessi sociali della comunicazione specialistica, Teresa Cabré ha fornito nuovi contributi per la formulazione di una proposta teorica e ha evidenziato il ruolo e l'influsso dell'organizzazione del-

¹⁰ Per una comprensione più agevole, il concetto di «pianificazione» può essere ritenuto equivalente a quello di «progettazione». Un approfondimento di tale concetto si trova in M. TERESA CABRÉ, *La terminología. Teoría, metodología, aplicaciones*, Barcelona, Editorial Antártida/Empúries, 1993, p. 108: «La normalizzazione di una lingua deve consistere in un processo pianificato che parta da una situazione esplicita, che si proponga alcuni obiettivi concreti da perseguire in un periodo di tempo determinato, che si basi su canali di diffusione e su risorse di "disseminazione" adeguate e che disponga di una legislazione che favorisca questo processo di scambio».

¹¹ JUAN C. SAGER, *A practical course in terminology processing*, Amsterdam-Philadelphia, John Benjamins, 1990, p. 1.

la società in rapporto alle due dimensioni fondamentali della terminologia: la rappresentazione e la trasmissione delle conoscenze trattate¹².

Il quadro complessivo appare, dunque, piuttosto frammentato, fino a far pensare che il solo elemento di unitarietà possa consistere in una prassi di lavoro comune, più che in un'improbabile conciliazione delle differenti impostazioni teoriche. A questo si aggiungano alcuni tentativi di portare alle estreme conseguenze il ruolo di scienza e disciplina autonoma attribuito alla terminologia¹³, ruolo che Wüster sembra avere comunque molto temperato, elaborando le *Teorie speciali della terminologia*, con l'intento di andare incontro alle peculiarità di ciascuno dei diversi settori specialistici e di ogni singola espressione linguistica¹⁴, e dichiarando esplicitamente l'appartenenza della terminologia alla linguistica applicata¹⁵.

¹² MARIA TERESA CABRÉ CASTELLVÍ, *Elementos para una teoría de la terminología: hacia un paradigma alternativo*, «El lenguaraz» (Buenos Aires), 1, 1998, 1, pp. 59-78.

¹³ Cfr. HERIBERT PICTH, *En record d'E. Wüster. La multidisciplinarietat de la terminologia*, in: M. T. Cabré (dir.), *Terminologia. Selecció de textos d'E. Wüster*, Barcelona, Universitat de Barcelona - Servei de la llengua catalana, 1996, pp. 253-287, in particolare il paragrafo 6. *Resum i perspectives*, pp. 277-278.

¹⁴ Per un esame più completo delle caratteristiche dei linguaggi scientifici, si veda M. DARDANO, *I linguaggi scientifici*, in: Luca Serianni e Pietro Trifone (a cura di), *Storia della lingua italiana*, Torino, Einaudi, vol. II, 1994, pp. 497-551, in particolare alle pp. 497-498. Sui «contrassegni di qualità», che rendono conto del carattere specifico dei linguaggi scientifici, si veda MAURIZIO DARDANO, *I linguaggi scientifici nell'italiano di oggi*, in: Ministero dell'Università e della Ricerca Scientifica e Tecnologica, *La terminologia tecnica e scientifica. Attualità e prospettive*, (Supplemento al Notiziario «Università e Ricerca»), Roma, Istituto Poligrafico e Zecca dello Stato, p.12.

¹⁵ Cfr. EUGEN WÜSTER, *Einführung in die allgemeine Terminologielehre und terminologische Lexikographie*, Wien, Springer, 1979, pp. 64-65.

Obiettivi della terminologia

La terminologia ha come finalità quelle di individuare e determinare le unità concettuali che costituiscono il sistema strutturato di conoscenze proprio di un settore specialistico e, successivamente, di occuparsi della denominazione di quelle unità concettuali mediante i termini. In altre parole, la terminologia si propone di esaminare e descrivere la struttura e l'organizzazione delle unità di conoscenza elaborate all'interno dei diversi settori del sapere specialistico: da quelli che, per il rigore metodologico e l'ambito di interesse più nettamente circoscritto, hanno raggiunto un livello di codificazione tradizionalmente consolidato e molto evoluto, come avviene nel caso delle scienze, a quelli delle tecnologie, delle arti e delle tecniche, ai settori di attività professionale, artigianale e pratica. Si tratta, per usare un'espressione di carattere onnicomprensivo, dei più diversi campi dell'esperienza umana, che tendono a distinguersi per un interesse omogeneo e condiviso e che sviluppano la necessità di esprimere le conoscenze e le esperienze maturate al loro interno attraverso un *linguaggio speciale*¹⁶, che si diversifica via via

¹⁶ Si registrano anche altre espressioni affini (tecnoleto, linguaggio settoriale, linguaggio della scienza e della tecnica, linguaggio scientifico, linguaggio specialistico), che si è soliti ricondurre all'ambito dei sottocodici linguistici. Rispetto al codice di base, questi ultimi presentano tratti distintivi peculiari. Occorre tuttavia notare il favore crescente che incontra la posizione teorica di quanti considerano lo studio della lingua da un punto di vista unitario: «La lengua es única, los llamados “lenguajes de especialidad”, o tecnoletos, no son más que especializaciones de determinados elementos de la única lengua general, y, por lo tanto, forman parte de ella. Por tal razón considero que es una inconsecuencia con los postulados más generales de la lingüística hablar de “lengua especializada”, “lengua de especialidad” y expresiones similares. [...] El término es un signo lingüístico como cualquier unidad léxica de las áreas no especializadas, y los hablantes científicos y técnicos son tan hablantes como el resto. Las diferencias entre “término” y

dalla lingua dell'uso comune, producendo quello che Tullio De Mauro ha definito «un uso speciale della lingua»¹⁷. Quest'uso tende a essere codificato sulla base di tratti particolari che lo caratterizzano nei tre piani: formale, funzionale e del significato. In particolare, e soprattutto per i linguaggi scientifici, da un punto di vista *formale*, sono assai numerosi i termini formati mediante il ricorso a affissi e confissi, molti dei quali di origine greca e latina: risulta in tal modo favorita la comprensione e la circolazione interlinguistica dei termini. Su un altro piano, si riscontra un'inclinazione all'essenzialità dei moduli *sintattici*, riducendo l'uso dei tempi e dei modi verbali e privilegiando una struttura lineare e nominale, che raramente ricorre a costruzioni ipotattiche. Per quanto riguarda l'aspetto *semantico*, si cerca di stabilire un rapporto diretto e univoco tra termine e designato, evitando che i termini si carichino di uno spettro ampio di significati (*polisemia*), che vengano usati, cioè, per designare più oggetti o concetti diversi tra loro. Si osserva, inoltre, che il significato è trasmesso senza ricorrere alle situazioni proprie dell'espressività comunicativa. Tali caratteristiche nascono fondamentalmente dal bisogno di rappresentare e di esprimere – nel modo più chiaro, sintetico e preciso possibile – la struttura concettuale del settore specialistico: si pensi alle grandi classifica-

“no término” son de tipo pragmático, no formal, funcional ni semántico. Esto hace que muchos de los postulados tradicionales de la terminología sean desmentidos, o al menos puestos en entredicho, por la práctica cotidiana», RODOLFO ALPÍZAR CASTILLO, *¿Cómo hacer un diccionario científico-técnico?*, Buenos Aires, Editorial Memphis, 1997, p. 8.

¹⁷ «Quando un cospicuo numero di accezioni di parole diverse e, eventualmente, di parole tecniche e neologismi siano in nesso tra loro e usati da gruppi specifici di parlanti per trattare di argomenti determinati, nasce ciò che la linguistica storica e sociologica chiama una 'lingua speciale' o 'linguaggio speciale' (o 'settoriale') e che meglio si dirà un 'uso speciale della lingua'», TULLIO DE MAURO, *Minisemantica. Dei linguaggi non verbali e delle lingue*, Roma-Bari, Laterza, 1995³, p. 131.

zioni tassonomiche degli esseri viventi elaborate da Carl von Linné (1707-1778) e alla riforma della nomenclatura chimica, condotta sul finire del XVIII secolo da Guyton de Morveau, Lavoisier, Berthollet e Fourcroy. Nello stesso periodo, Diderot pensava a una «grammatica delle arti», come fattore di regolazione sistematica delle terminologie e dei linguaggi delle scienze e delle tecniche, e nella sua voce *Encyclopédie* notava: «la conoscenza della lingua è il fondamento di tutte queste grandi speranze; esse rimarranno incerte, se la lingua non è stabilmente definita e trasmessa alla posterità in tutta la sua perfezione; e questo è lo scopo primario tra quelli di cui gli Enciclopedisti dovrebbero occuparsi profondamente. Noi ce ne siamo accorti troppo tardi; e tale inavvertenza ha proiettato un'imperfezione su tutta la nostra opera. Il versante della lingua è rimasto debole (dico *della lingua*, e non *della Grammatica*); e perciò questo dev'essere l'argomento principale in una voce in cui si esamina imparzialmente il proprio lavoro, e in cui si cercano i mezzi per correggerne i difetti»¹⁸.

Quanto più elevato è il rigore metodologico e specificamente circoscritto l'ambito di interesse di un settore specialistico, come avviene nel caso delle scienze, tanto più le caratteristiche

¹⁸ «La connoissance de la langue est le fondement de toutes ces grandes espérances; elles resteront incertaines, si la langue n'est fixée & transmise à la postérité dans toute sa perfection; & cet objet est le premier de ceux dont il convenoit à des Encyclopédistes de s'occuper profondément. Nous nous en sommes aperçus trop tard; & cette inadvertance a jetté de l'imperfection sur tout notre ouvrage. Le côté de la langue est resté foible (je dis de *la langue*, & non de *la Grammaire*); & par cette raison ce doit être le sujet principal, dans un article où l'on examine impartialement son travail, & où l'on cherche les moyens d'en corriger les défauts», *Encyclopédie ou Dictionnaire raisonné des sciences, des arts et des métiers. Nouvelle impression en facsimilé de la première édition de 1751-1780*, Vol. 5, Stuttgart-Bad Cannstatt, Friedrich Frommann Verlag (Günther Holzboog), 1966, p. 637^r.

di quest'uso speciale della lingua si manifestano in modo evidente e peculiare. Al punto che alcuni studiosi arrivano a distinguere nettamente i linguaggi scientifici da quelli settoriali, osservando che «le argomentazioni della scienza, i metodi di indagine, le nuove prospettive del pensiero, modificando il campo della percezione, creano nuove forme enunciative e linguistiche»¹⁹. Anche altre ragioni contribuiscono a determinare il lento processo di diversificazione dei linguaggi specialistici dalla lingua d'uso: in primo luogo, la necessità di una comunicazione rapida e efficace, in grado di superare agevolmente le barriere delle espressioni linguistiche nazionali, ma anche il bisogno di trasmettere a nuovi cultori le nozioni che formano il patrimonio di conoscenze sviluppate nel settore al quale intendono dedicarsi.

Occorre, a questo proposito, almeno ricordare l'impegnativo tema della varietà dei discorsi scientifici, che possono essere utilmente classificati in lingua della teoria (*Theoriesprache*), lingua del laboratorio (*Werkstattsprache*) e lingua della divulgazione (*Verteilersprache*)²⁰. Considerata da un punto di vista terminologico, questa diversificazione di registri del discorso scientifico, che si basa sul pubblico dei destinatari e quindi sulla rilevazione del diverso grado di specializzazione dei testi, si riflette profondamente anche nell'uso dei termini. Per fare un solo esempio, si pensi alla denominazione dei prodotti chimici: il composto individuato e determinato dalla formula H_2O può essere anche denominato *monossido di diidrogeno* o, più semplicemente, *acqua*. Si

¹⁹ M. DARDANO, *I linguaggi scientifici*, in: Luca Serianni e Pietro Trifone (a cura di), *Storia della lingua italiana*, Torino, Einaudi, vol. II, 1994, p. 499.

²⁰ Cfr. WOLFGANG PÖCKL, *Französisch: Fachsprachen. Langues de spécialité*, in: G. Holtus – M. Metzeltin – Ch. Schmitt (hrsg.), *Lexikon der Romanistischen Linguistik (LRL)*, Tübingen, Max Niemeyer Verlag, 1990, pp. 268-269.

è inoltre riscontrato che, quanto più è elevato il grado di specializzazione di un testo scientifico, tanto maggiore risulta la densità dei termini specialistici in esso presenti, proprio a motivo di bandire ogni possibile equivoco o incomprendione nella comunicazione.

Scienziati e tecnologi hanno come obiettivo primario l'avanzamento delle conoscenze e – nei limiti posti dagli interessi particolari e dalla concorrenza imprenditoriale – tendono a mettere in comune il frutto del proprio lavoro. Hanno bisogno di comunicare in modo efficace, rapido e preciso con i loro colleghi, prescindendo dalle limitazioni imposte dallo spazio e dal tempo. Proprio per questo sono portati a privilegiare una lingua di comunicazione internazionale che, a costo di comprimere le diversità culturali dei singoli, sia in grado di favorire le relazioni e gli scambi. Per essere conosciuti in ambito internazionale, scienziati e tecnologi devono far circolare le loro ricerche e i loro studi su riviste pubblicate in inglese, lingua che utilizzano abitualmente anche per presentare le loro relazioni ai congressi internazionali. Occorre, però, tenere ben presente che il fatto che la maggior parte della produzione scientifica si esprima in inglese può risultare fortemente penalizzante per lo sviluppo armonico delle singole lingue nazionali. Se uno scienziato non curasse l'adeguata circolazione dei termini specialistici del suo settore d'interesse anche nella propria lingua madre, ne determinerebbe un progressivo impoverimento che potrebbe arrivare fino alla completa atrofia lessicale dei settori maggiormente specializzati, nei quali opera un più ristretto numero di addetti. Ma anche i grandi settori specialistici più direttamente coinvolti nel fenomeno della globalizzazione – si pensi all'informatica e alla telematica, alle nuove tecnologie in genere, all'economia e alla nuova economia – sono quelli che presentano nell'italiano contemporaneo il più alto tasso di prestiti integrali, di parole cioè originariamente estranee al sistema linguistico.

Metodo di lavoro della terminologia

Il metodo di lavoro proprio della terminologia è detto *onomasiologico* perché si propone di denominare le singole unità di conoscenza individuate in un settore specialistico. Accade di rado, però, che tale attività assuma anche un'effettiva funzione *onomaturgica*, ovvero che il terminologo sia chiamato a imporre un nome alle unità rilevate. Più abitualmente, il suo lavoro consiste in un'opera paziente di inventariazione e sistematizzazione delle conoscenze accumulate in un settore specialistico per classificarle mediante i termini che, denominandole, le rappresentano. Nella maggior parte dei casi, si tratta di espressioni linguistiche, semplici o complesse, ma possono essere anche segni, simboli, formule o elementi grafici. I termini sono convenzionalmente scelti per rappresentare – nella forma più precisa, concisa e adeguata – la definizione di ciascun nucleo della conoscenza elaborata in un determinato settore²¹ e, al tempo stesso, essi costituiscono il veicolo della comunicazione specialistica, sia tra gli specialisti dello stesso settore, sia nel trasferimento delle conoscenze acquisite, a livello divulgativo o di scambio interlinguistico²².

²¹ «Y es así porque en todas las representaciones los símbolos utilizados – igual que los términos, que son en definitiva representaciones – no son más que substituciones de la definición subyacente del concepto de que se trate; y las definiciones del concepto están establecidas con palabras», BERTHA M. GUTIÉRREZ RODILLA, *La ciencia empieza en la palabra. Análisis e historia del lenguaje científico*, Barcelona, Ediciones Península, 1998, p. 30. L'argomento è ripreso poco dopo: «Como ya hemos dicho, los términos son marcas substitutivas de sus correspondientes definiciones, porque es la definición – completa y precisa – la única representación lingüística adecuada de un concepto. Reemplazan las definiciones – normalmente largas – por una expresión más económica pero de igual precisión; esto significa que no son intrínsecamente precisos, sino que sólo lo son si su definición lo es en sí misma», *ivi*, p. 32.

²² M. TERESA CABRÉ, *La terminología. Teoría, metodología, aplicaciones*, Barcelona, Editorial Antártida/Empúries, 1993, p. 71.

L'attività terminologica consiste, dunque, nell'ordinamento e nella classificazione del patrimonio cognitivo e comunicativo di un dominio. Pertanto, la rappresentazione concettuale e terminologica di un settore specialistico realizzata da un terminologo può costituire una via d'accesso privilegiata attraverso la quale anche il documentalista e il traduttore possono avvicinarsi al settore specialistico di cui debbono occuparsi per il proprio lavoro, per farsene un'idea complessiva o per approfondire aspetti di più specifico interesse.

Il prodotto dell'attività terminologica consiste nella rappresentazione linguistico-documentaria delle unità concettuali individuate mediante l'analisi di un *corpus* di fonti scritte, le più rappresentative e autorevoli del dominio specialistico esaminato: trattati e manuali, documentazione tecnica, riviste di settore e, se esistono, anche fonti legislative e normative (leggi, trattati, protocolli, convenzioni, norme tecniche, nazionali e internazionali). Secondo la natura specifica di ciascun dominio, possono essere presi in considerazione anche glossari o repertori terminologici precedentemente realizzati e, talvolta, anche periodici divulgativi. L'intento è quello di prelevare materiale da un *corpus* che offra il panorama più fedele possibile delle conoscenze sviluppate nel settore, attraverso i termini effettivamente utilizzati dagli specialisti.

La rilevazione di forme diverse di lessicalizzazione dei termini che designano le medesime unità concettuali (come nei casi di varianti formali, sinonimi, varianti di scuola, varianti diatopiche o geografiche, impiego di termini abbreviati e acronimi) permette di fornire un quadro che consente di evidenziare il ventaglio di sfumature che si delineano all'interno di quello che si tenderebbe a considerare un linguaggio unitario, monolitico e prescrittivo. In esso emergono, in misura sempre più significativa, le variazioni che si producono all'interno di scuole di pensiero diverse, delle differenti aree regionali nelle quali si parlano lingue a diffusione intercontinentale, di consuetudini editoriali che im-

pongono o consentono agli autori dei contributi anche l'uso di sigle estemporanee, spesso del tutto arbitrarie e prive di una reale efficacia che non sia quella dettata dal mero tentativo di risparmiare qualche riga di testo.

La registrazione delle varianti non comporta di per sé l'abbandono della vocazione originaria della terminologia: la tendenza all'unificazione e alla normalizzazione dei termini, che è esigenza avvertita principalmente nella comunicazione internazionale, e che costituisce uno dei tratti che maggiormente connotano l'attività terminologica. A questa istanza, ancora oggi fortemente condivisa, si sono affiancati anche altri punti di vista, che tendono a considerare con maggiore attenzione le terminologie nel loro rapporto con le singole culture, e quindi con le lingue che le esprimono. Si va affermando, insomma, un orientamento maggiormente improntato al pragmatismo e alla funzionalità.

Occorre precisare che, anche a causa della sempre più frequente divulgazione attraverso la rete Internet, nella forma di banche dati, le risorse terminologiche disponibili per la consultazione non sempre sono presentate seguendo l'ordinamento in precedenza più consueto, ovvero quello detto *concettuale*, *tematico*, *metodico* o *sistematico*. A differenza dell'ordinamento alfabetico, privilegiato dai repertori lessicografici per la facilità di consultazione, l'ordinamento concettuale è senz'altro più coerente con l'impianto e le finalità delle raccolte terminologiche, spesso destinate alla consultazione di utenti che conoscono approfonditamente la materia. Ne risulta, infatti, una sorta di quadro sinottico del settore, che permette di valutare l'approfondimento dell'indagine compiuta dal terminologo e facilita anche le operazioni di modifica e aggiornamento del repertorio stesso²³.

²³ Konferenz der Übersetzungsdienste westeuropäischer Staaten (Arbeitsgruppe *Terminologie und Dokumentation*), *Empfehlungen für die Terminologearbeit*, Bern, Schweizerische Bundeskanzlei, 1990, Kap. 6, Klassifikation.

Questo tipo di ordinamento, ispirandosi a un criterio vagamente enciclopedico, consente anche a chi non è esperto di farsi un'idea d'insieme e di esplorare i nessi della rete concettuale del settore specialistico preso in esame.

Repertori terminografici

Come appena ricordato, la produzione terminografica risponde, nella gran parte dei casi, alle esigenze di un pubblico esperto, che ricorre ad essa per controlli e verifiche, ma anche per cercare gli equivalenti di un termine in altre lingue. I repertori terminografici sono, infatti, prevalentemente multilingui e si propongono di rispondere alle esigenze di una rapida comunicazione interlinguistica. Recentemente hanno avuto diffusione anche piccoli prontuari terminologici destinati alla divulgazione di terminologie specialistiche, perlopiù relative alle nuove tecnologie di grande impatto sociale. Si tratta spesso di repertori monolingui o corredati del solo equivalente nella lingua d'origine delle tecnologie stesse, ma anche di piccole raccolte terminologiche multilingui: l'essenzialità del contenuto è dettata dalle esigenze di una diffusione capillare e al tempo stesso economica, che tiene conto della rapida obsolescenza cui vanno soggette le tecnologie e delle conseguenti ripercussioni nel cambiamento delle terminologie. L'orientamento alla pianificazione linguistica e l'attenzione ai criteri della socioterminologia hanno introdotto elementi innovativi rispetto alla tradizionale presentazione terminografica: la segnalazione delle varianti e la presenza dei contesti d'uso sono quelli di maggior rilievo.

La terminografia raccoglie e presenta i termini dei linguaggi specialistici e tende all'uniformazione dei concetti e all'unificazione delle denominazioni, per consentire una comunicazione più diretta e efficace. La norma che rappresenta ha quindi la funzione di modello prescrittivo, benché l'orientamento di tipo va-

riazionista vada raccogliendo un favore sempre più diffuso.

Il criterio privilegiato è quello sincronico, che testimonia l'impegno a contribuire a una comunicazione sempre aggiornata; per questa ragione, le norme internazionali sono sottoposte a revisione periodica²⁴.

Il *corpus* terminografico è selezionato con criteri rigorosi di rappresentatività e generalmente si limita alla documentazione scritta, privilegiando le fonti più autorevoli.

I repertori tendono a escludere la polisemia: un termine si connota per la relazione di univocità stabilita con il suo designato e solo in casi piuttosto rari può avere sinonimi. È possibile però riscontrare casi di omonimia, cioè di termini identici che designano oggetti distinti, ma questo può accadere in settori specialistici diversi tra loro.

La frequenza d'uso di un termine non ha alcuna rilevanza ai fini dell'inclusione in un repertorio terminografico: è stato osservato che il grado di specializzazione di un termine è inversamente proporzionale alla sua frequenza d'uso²⁵.

Le entrate coincidono con le unità di conoscenza elaborate in un settore specialistico e, quindi, sono rappresentate da termini, che possono essere semplici o composti, e talvolta accompagnati dall'indicazione della categoria grammaticale.

La definizione descrive le caratteristiche e le funzioni dell'oggetto o del concetto denominato dal termine.

La contestualizzazione dei termini non è obbligatoria. Quan-

²⁴ «Der Vorrang der Begriffe hat zwangsläufig dazu geführt, daß die terminologische Sprachbetrachtung synchronisch ist. Für die Terminologie ist das Wichtigste an einer Sprache das Begriffssystem, das ihr zugrunde liegt», E. WÜSTER, *Einführung in die allgemeine Terminologielehre und terminologische Lexikographie*, cit., p. 2.

²⁵ Cfr. BODO MÜLLER, *Le français d'aujourd'hui*, traduit de l'allemand par Annie Elsass, édition révisée et augmentée par l'auteur, Paris, Éditions Klincksieck, 1985, p. 188.

do è presente, appare improntata all'esigenza della verifica documentaria, testimoniando l'uso effettivo dei termini in fonti di diversa natura (ufficiale, divulgativa) e attestando le variazioni geografiche e culturali.

L'ordinamento terminografico è prevalentemente concettuale, perché destinato alla consultazione di utenti che conoscono approfonditamente la materia²⁶. Questo tipo di ordinamento può rivelarsi utile nell'estendere controlli e verifiche a elementi contigui a quello da cui ha avuto origine la ricerca e fornisce, inoltre, una sorta di quadro sinottico di un dominio del quale può giovare anche chi, non esperto, volesse farsene un'idea d'insieme. Tuttavia, le finalità divulgative e l'uso delle tecnologie informatiche sono alla base dell'impiego dell'ordinamento alfabetico, che sembra incontrare una diffusione in costante crescita.

Documentazione della terminologia

L'attività terminologica si ispira a un criterio di natura sincronica, che presuppone una costante attualizzazione del patrimonio terminologico-concettuale raccolto per ciascun settore. Si pensi, per esempio, alle norme tecniche, nazionali e internazionali, istituzionalmente sottoposte a revisione e aggiornamento periodico, anche nella parte relativa alla terminologia e al vocabolario, allo scopo di assicurare precisione e efficacia alla comunicazione in ciascun settore. Gli enti nazionali e internazionali preposti all'unificazione, alla normalizzazione e alla standardizzazione sono ormai dotati di siti web presso i quali è possibile consultare il catalogo delle norme prodotte per ogni settore, con le relative date di aggiornamento: si tratta di una fase indispensabile e prelimi-

²⁶ Cfr. EUGEN WÜSTER, *Einführung in die allgemeine Terminologielehre und terminologische Lexikographie*, Bonn, Romanistischer Verlag, 1991³, Par. 9.42, pp. 126-127.

nare a ogni lavoro di documentazione o di traduzione tecnica.

Allo stesso modo, è opportuno conoscere l'attività svolta da molte associazioni scientifiche, professionali e di categoria, sul piano nazionale e internazionale. Spesso esse dispongono di centri di documentazione per il settore di cui si interessano. Talvolta, mettono a disposizione, anche attraverso la rete Internet, repertori bibliografici, indicazioni relative ai dizionari specialistici, alla regolamentazione legislativa e normativa del settore. E, in molti casi, i siti di queste istituzioni sono dotati di motori di ricerca, che permettono di compiere interrogazioni puntuali mediante l'impiego di termini specifici, oppure presentano glossari o liste selezionate di parole chiave, che possono rivelarsi di grande utilità per il documentalista e il traduttore.

Negli ultimi anni sono state costituite molte associazioni nazionali e internazionali di terminologia, con l'obiettivo di censire, incrementare e coordinare la produzione di risorse terminologiche nei vari settori specialistici. Esse forniscono un servizio informativo e di documentazione sui repertori terminologici esistenti, sia a stampa, sia disponibili per la consultazione in linea. Raccolgono anche informazioni bibliografiche sulla produzione teorica e applicativa nel settore della terminologia. Una segnalazione particolare merita l'iniziativa intrapresa dall'Associazione europea per la terminologia (Aet-Eaft: <http://www.eaft-aet.net/>) per la costituzione di un Server europeo di informazione terminologica (Etis: <http://www.computing.surrey.ac.uk/ai/etis/>), che si propone come centro unitario di raccolta e smistamento di informazioni in ambito europeo. Questa iniziativa intende soddisfare una delle esigenze oggi più fortemente avvertite, quella di ridurre l'alto tasso di frammentazione delle risorse terminologiche. Infatti, nonostante l'opera preziosa e efficace svolta dalle associazioni nazionali e dai centri di consulenza terminologica, non esiste ancora oggi un nucleo unitario di riferimento e di orientamento per chi voglia entrare in contatto con il mondo complesso e articolato della terminologia.

Nella stessa direzione sono stati rivolti anche gli sforzi di coloro che hanno contribuito alla formazione di vere e proprie *reti di lavoro* tra quanti operano nei più disparati settori dell'attività terminologica: *Nordterm*, un circuito che collega i Paesi del Nord Europa (<http://www.nordterm.net/>); il *Rint*, Rete internazionale di neologia e terminologia tra i Paesi dell'area francofona (<http://catalogue.cdeacfc.ca/Record.htm?idlist=1&record=19115309124919335819>); *Riterm*, Rete iberoamericana di terminologia tra i Paesi di espressione linguistica spagnola e portoghese (<http://www.riterm.net/>); e *Realiter*, la Rete panlatina di terminologia, che si pone come luogo di raccordo e di scambio tra i Paesi di lingua neolatina (http://dtil.unilat.org/realiter_spip/).

Ma la risorsa più preziosa, che spesso si rivela anche la più ardua da raggiungere e da consultare, è costituita dalle numerose banche di dati terminologici. Nel 1996, il Rint ha realizzato un *Inventaire des banques de terminologie*, disponibile su Internet e attualmente in corso di aggiornamento, che raccoglie schede informative su un gran numero di banche terminologiche, grandi e piccole, private e pubbliche. Ciascuna scheda mostra le finalità, la consistenza, il tipo e la quantità di dati raccolti, e offre notizie sulle formalità e le condizioni di accesso alla banca dati (<http://www.inventerm.com/>).

La letteratura tecnica che descrive le basi di dati terminologici seleziona all'interno di esse un gruppo ristretto, che viene abitualmente designato con l'appellativo di «grandi banche dati». Si tratta di banche dati multilingui, le più importanti e consistenti, le prime a essere implementate, tra la fine degli anni Sessanta e l'inizio dei Settanta: *Eurodicautom*, oggi in via di sostituzione con

IATE (<http://europa.eu/languages/it/document/100>), *Termium*, (<http://www.bureaudelatraduction.gc.ca/index.php?lang=francais&cont=001>), *le Grand dictionnaire terminologique* (<http://www.olf.gouv.qc.ca/ressources/gdt.html>), *Lexis* e *Team* (Gagnon 1994, pp. 498-499). Queste prime banche dati erano sta-

te ideate con lo scopo di superare la rapida obsolescenza dei dizionari e repertori a stampa in rapporto alla continua evoluzione delle terminologie, e per approntare uno strumento che facilitasse il lavoro di documentalisti, traduttori e interpreti. L'impianto di queste banche dati richiedeva però l'investimento di risorse considerevoli, dal punto di vista finanziario, informatico e del lavoro umano, senza tenere conto della complessità e lentezza delle operazioni di aggiornamento, o della necessità di una rete capillare per la diffusione e la consultazione dei dati: soltanto da alcuni anni, grazie a Internet, quest'ultimo aspetto ha trovato una positiva soluzione.

Il clima di sfiducia e scetticismo che si era prodotto attorno a queste grandi banche dati, nella seconda metà degli anni Settanta e nei primi anni Ottanta, ne aveva determinato un lento e progressivo abbandono, culminato con la diffusione dei personal computer, che hanno contribuito allo sviluppo di banche dati di dimensioni più ridotte, e finalizzate alla raccolta di terminologie di settori particolari. Presto, però, si è dovuto constatare che la nuova soluzione costringeva a reduplicare il lavoro necessario per costruire piccole banche dati settoriali ad uso personale da parte dei vari professionisti impegnati nel medesimo settore di attività. In questo modo finiva per essere incentivata anche la mancanza di omogeneità delle terminologie, troppo spesso affidate alla sensibilità e all'intuizione dei singoli, e prive di ogni forma di controllo e di validazione. Di conseguenza, nell'ultimo decennio si è registrato un forte impulso a varie forme di collaborazione: associazioni, centri di consulenza e di servizio. La contemporanea diffusione di nuove tecnologie per la distribuzione e comunicazione dei dati, il cd-rom e le reti informatiche, ha riportato alla ribalta l'utilità delle grandi banche dati terminologiche centralizzate.

Thesauri, mappe semantiche, ontologie. Problemi semantici e costruzioni concettuali

DANIELE GAMBARARA - ALFREDO GIVIGLIANO¹

1. I documenti e il linguaggio come conoscenza

Tratteremo in questo paragrafo le caratteristiche di lingua storico-naturale, specie semantiche, di documenti, testi linguistici, rete/i di saperi/conoscenze². Essi non possono essere compresi e interpretati al di fuori di qualunque contesto (non contengono cioè tutti i dati, e soprattutto, tutti i metadati necessari alla loro comprensione): sono, in effetti, oggetti e dimensioni di una dinamica in costante tensione con le determinazioni del parziale rispetto al totale (depositate nell'individuo rispetto alla collettività) e del locale rispetto al globale (determinato geograficamente e temporalmente). Filologia, semiotica del testo, ermeneutica cercano di recuperare da diverse posizioni per quanto possibile singole configurazioni in contesto. Ma è ben diverso dallo stabi-

¹ I due autori hanno progettato e discusso insieme questo contributo. A Daniele Gambarara vanno attribuiti i parr. 1, 2, ad Alfredo Givigliano vanno attribuiti i parr. 3, 4.

² Chiameremo qui quello che si esprime nei significati delle parole, sapere generico o conoscenza linguistica (inestricabilmente connessa con le conoscenze enciclopediche comuni) per distinguerlo dai concetti delle terminologie che organizzano un sapere più o meno formalizzato.

lire queste condizioni a priori, a cui ambiscono solitamente i trattamenti automatici dei testi, proseguendo l'oggettivazione di questi saperi, che già operano grammatiche, dizionari, ed enciclopedie. Un possibile ostacolo, una possibile serie di ostacoli, al loro trattamento oggettivato può essere rintracciata nei concetti di variabilità ed apertura, ineliminabili dalla semantica delle lingue naturali.

Un documento è un testo, e la strutturazione che ce ne fa riconoscere la testualità è essenzialmente quella del linguaggio verbale. Certo, esistono documenti composti principalmente o esclusivamente da figure o diagrammi, ma di fatto essi sono trattati come se queste componenti fossero la realizzazione non verbale di un testo verbale: nome e date dell'autore, titolo dell'opera e collocazione per un disegno o una fotografia, legenda o didascalia che dichiara l'oggetto per una illustrazione o una figura tecnica.

Non approfondiremo qui la questione del rapporto fra immagini (o in genere elementi iconici) e parole, su cui la discussione è accesa da anni. Ci limitiamo a segnalare che, a nostro parere, solo alcune immagini sono propriamente indipendenti dal linguaggio verbale, e dotate di autonoma e diversa strutturazione, e quindi vanno trattate 'come se' lo fossero, con gli espedienti di cui sopra; altre immagini (e forse tutti i diagrammi) hanno effettivo rapporto con la strutturazione del linguaggio verbale, e possono esserne espresse senza residui.

Entriamo, quindi, all'interno del mondo, dell'universo del discorso, che si declina nelle regole della testualità. I documenti/testi in lingue verbali, storico naturali, costituiscono sempre in qualche forma espressione di un sapere specifico, determinato, costruito con gli elementi di sapere generico costitutivi degli stessi sistemi linguistici (Dizionario, ma anche Enciclopedia), che sono comunque determinati spazialmente e temporalmente. Il nostro punto di partenza non può, quindi, che essere quello della considerazione del linguaggio verbale umano nella sua *na-*

tura di prassi³. In questo modo possiamo anche rendere conto del fatto che «Se potessimo abbracciare la somma delle immagini verbali immagazzinate in tutti gli individui, toccheremmo il legame sociale che costituisce la lingua. Questa è un tesoro depositato nella pratica della *parole* nei soggetti appartenenti a una stessa comunità»⁴.

Questo fatto, essenziale al funzionamento delle lingue, e alla loro capacità di produrre testi, incontra due ordini di problemi:

- a. variabilità nel tempo e nello spazio (diversità delle lingue, variabilità diacronica e diatonica);
- b. usi informali e usi formali del linguaggio.

In altri termini, si può anche concludere, contestualizzando che: «Questa comunicazione richiede un primo grado di meta-rappresentazioni, e su questa base possono presentarsi stati intenzionali che costituiscono atteggiamenti ‘quasi-proposizionali’, ricordando sempre i limiti della sua semantica e della sua sintassi»⁵.

1.1 Sincronicità

Un primo caso rilevante di determinazione nel continuo mutare delle lingue e delle conoscenze che esse supportano è l'estrazione da questo delle sincronie. L'italiano contemporaneo somiglia all'italiano di Dante, ma ne è un falso amico: “Tanto gentile e tanto onesta pare la donna mia...” non dice in italiano oggi, ciò che Dante voleva dire; si dovrebbero in realtà distinguere

³ Cfr. DANIELE GAMBARARA., *Comunicazione, cognizione, socialità*, in FRANCESCO FERRETTI – GAMBARARA (a c. di), 2005, *Comunicazione e scienza cognitiva*, Laterza, Roma-Bari, p. 193.

⁴ FERDINAND DE SAUSSURE, *Cours de linguistique générale*, Payot, Paris, trad. it. a c. di TULLIO DE MAURO, *Corso di linguistica generale*, Laterza, Roma-Bari, p. 23.

⁵ Cfr. GAMBARARA, *Comunicazione*, cit., p. 23.

gli italiani di Dante: quello della Vita Nova non è quello del Convivio, e ancora di più non è quello della Commedia. «È sincronico tutto ciò che si riferisce all'aspetto statico della nostra scienza, è diacronico tutto ciò che ha rapporti con le evoluzioni. Similmente, *sincronia* e *diacronia* designeranno rispettivamente uno stato di lingua ed una fase di evoluzione»⁶. L'individuazione delle singole sincronie, cioè delle sistematicità in cui i termini assumono valore correlativamente, è operazione preliminare e delicata.

Vedremo in seguito come la modalità della sincronia, in correlazione con quella diacronia, sia uno degli aspetti caratteristici della costruzione stessa e della gestione della conoscenza, delle conoscenze. Costruzione e gestione che viaggiano sul filo della riduzione, ma che nello stesso tempo intrecciano (anche in virtù di questa eventualità) relazioni molto strette con le tecnologie a disposizione. Conseguenza rilevante: sapere nel testo e sul testo, sia in maniera contestualizzata (*ad hoc* per quel testo) che in modo tale da poter unire più conoscenze; quindi, in forma e contenuto.

2. Parole e termini, significati e concetti

In questo paragrafo tratteremo questioni semantiche, perché gli stessi sistemi semantici delle lingue storico-naturali, pur essendo sistemi, non vanno intesi nel senso in cui trattiamo sistemi di saperi organizzati. Si co-determinano, ma lo fanno in un quadro di apertura e in qualche modo il loro aspetto dinamico è intrinseco. Essi devono potersi adattare a bisogni diversi in condizioni mutate. Mentre ciò che facciamo, in parziale astrazione, in un sistema formale ne determina anche, sotto alcuni rispetti, la

⁶ DE SAUSSURE, *Cours*, cit., p. 100.

possibilità di crollo, i sistemi semantici non devono crollare mai.

Ancora, apertura perché la determinazione relativa è locale su campi semantici ristretti e la adattabilità ad una forma di autoorganizzazione potrebbe far considerare che vi sia in essi una sorta di autopoiesi⁷: uno dei supporti organici alla nostra intelligenza, non una protesi propriamente esterna.

Problema di fondo è, quindi, il rapporto tra significato e concetto. Le coppie in qualche modo oppostive delle quali bisogna trovare una possibilità di risoluzione risultano essere:

- a. discipline scientifiche e terminologie scientifiche;
- b. parola e termine;
- c. usi informali e usi formali (distinto, ma in rapporto, formalizzazione).

In questo modo, si può arrivare alla possibile costruzione di una organizzazione gerarchica di una terminologia, alla possibilità di strumenti metalinguistici per la sua trattazione; partendo da (e ritornando a) lingue storico naturali che presentano variabilità e apertura degli stessi sistemi semantici. Esempio: le trattazioni per campi semantici (forse le più vicine, in termini linguistico descrittivi, agli oggetti che tratteremo nel par. 4).

Problema di fondo che si complica ulteriormente nel momento in cui la relazione è a tre tra *pensiero*, *concetto* e *significato*⁸. Tre oggetti che sono stati variamente intrecciati tra loro nel corso della riflessione filosofica, linguistica come anche nell'universo di discorso identificato come scienze dell'informazione. E che servono a classificare e comprendere gli oggetti della prassi documentaria.

⁷ Non nei termini propri di Maturana e Varela, ma in un percorso che viaggiando non in modo propriamente parallelo presenta anche possibili intersezioni.

⁸ Cfr. GAMBARARA., *Comunicazione*, cit.

2.1 La dinamica tra parola e termine

Non si può prescindere, quindi, dall'identificare nella tensione tra *parola* (nell'uso informale quotidiano di una lingua storico naturale) e *termine* (in un uso più formale, in un sottoinsieme chiuso di lingua) uno dei nodi più significativi da analizzare. Mentre da un lato la parola, che rinvia necessariamente ad una dimensione di linguaggio naturale, è di per se stessa una possibilità di significazione sempre ancora aperta, dall'altro, il *termine* è il fissare, all'interno di un discorso (linguaggio) formale, il significato stesso cercando di farlo coincidere al massimo con il concetto che si intende usare.

Come vedremo in seguito, anche in questi termini si può porre l'ulteriore tensione tra livello e modalità di *sincronia* e livello e modalità di *diacronia* degli e negli stessi modelli di rappresentazione.

2.2 Dalla vaghezza alla precisazione

Anche considerato sincronicamente il linguaggio naturale è *naturalmente* vago. Questa affermazione che potrebbe sembrare del tutto ovvia, fino anche ridondante, racchiude in sé gran parte della complessità del problema. Problema che, partendo dalla tensione tra *parola* e *termine* arriva ad includere, in una sorta di gioco dei tre corpi, anche il *concetto*. Il tutto nell'ottica della *usabilità* dei modelli operativi⁹.

Vedremo nel paragrafo 3.2.1. come tutto questo si iscriva anche all'interno di una concretizzazione del simbolico. Concretizzazione che pone come problematica la posizione del produttore, anche e soprattutto in relazione a quella del cosiddetto *utente*.

Ma qui ciò che ci interessa è che, in modi diversi, da quelli che tengono più dell'espedito empirico, a quelli che adottano metodologie esplicite e sofisticate, anche la possibilità di preci-

⁹ Cfr. GAMBARARA., *Comunicazione*, cit, p. 198, nota 3.

sazione ed esplicitazione degli usi è interna alle possibilità della lingua. Sapendo ciò che si fa, e informandone gli utenti, è possibile procedere, e di fatto si procede utilmente, alla costruzione di modalità che consentono un trattamento (quasi-) automatico dei testi linguistici.

3. Modellizzazioni di significati e concetti

I concetti circolano,
si spostano in campi problematici e oggettuali
assai lontani dal loro punto di partenza,
si arricchiscono e mutano natura,
ritornano così arricchiti vicino al punto di partenza
rivelando aspetti e problemi del loro oggetto
prima completamente inopinati e non indagabili,
allargano il quadro di riferimento, generalizzano,
si travestono, scompaiono, rinascono...¹⁰

All'interno del discorso riguardante il nostro percorso, gli oggetti da tenere insieme e, quindi, da analizzare sia singolarmente che nelle loro relazioni reciproche, ai fini di una strutturazione-costruzione di strumenti per l'analisi dei testi e dei documenti (Thesaurus¹¹, Reti/Mappe semantiche¹², Mappe topi-

¹⁰ MAURO CERUTI, *Il vincolo e la possibilità*, Feltrinelli, Milano, 1996, p. 29.

¹¹ La descrizione di un Thesaurus sarà presentata in seguito.

¹² «Le relazioni tassonomiche sono rappresentabili in modo naturale tramite reti semantiche, in cui le varie categorie sono usate per etichettare i nodi della rete, che a loro volta sono connessi per mezzo di archi rappresentanti la relazione di inclusione fra categorie. Quando tale relazione di inclusione non è stretta, ma solo plausibile, si hanno reti semantiche non monotone in cui si possono rappresentare le eccezioni», ALDO ANTONELLI, *La logica del ragionamento plausibile*, in LUCIANO FLORIDI (a c. di), *Linee di ricerca*, SWIF, Sito Web Italiano per la Filosofia – ISSN 1126-4780 – www.swif.it/biblioteca/1r, 2003, p. 11.

che¹³, Mappe cognitive¹⁴, Ontologie¹⁵) possono essere identificati, come abbiamo visto, in *in pensiero, concetto e significato*¹⁶. Per quanto ci riguarda, il problema risulta essere la concretizza-

¹³ Esempi di Mappe topiche sono illustrati in:

<<http://www.lmtm.de/>;

http://www.ifi.uio.no/info-design/info_design_uio.php;

<<http://jncicancerspectrum.oxfordjournals.org/>;

<<http://www.nzetc.org>; <http://www.austlit.edu.au/>>.

¹⁴ Una discussione estremamente stimolante sulla *rappresentazione* e la *co-gnizione*, con riferimento ad una visione filosofica delle mappe cognitive può essere ritrovata in ROBERTO CORDESCHI, *Vecchi problemi filosofici per la nuova intelligenza artificiale*, in «Networks», 1, pp. 1-23, 2003, <www.swif.it/lei/ai/networks>.

¹⁵ «In termini molto generali, in informatica si suole chiamare “ontologia” lo schema concettuale di un certo dominio. Tale schema concettuale può assumere forme molto diverse, a partire da una semplice tassonomia fino ad arrivare a una vera e propria teoria logica. Normalmente si tratta di una gerarchia di concetti correlati attraverso relazioni semantiche, ma le ontologie più elaborate forniscono anche regole (come assiomi e teoremi) che aiutano a specificare come è strutturato il dominio. Siamo quindi in presenza di artefatti ingegneristici costituiti da un vocabolario che descrive un certo dominio e un insieme di assunzioni esplicite che vincolano l'interpretazione dei termini del vocabolario in modo da avvicinarsi il più possibile al loro significato inteso», MARGHERITA BENZI – ROBERTA FERRARIO, *Prefazione*, in «Networks», 6, 2006, <www.swif.it/lei/ai/networks>, pp. i-iv, p. ii. Per ulteriori contestualizzazioni sia di utilizzo che di costruzione, teoretica e tecnica si rimanda a (ogni URL è preceduto dall'acronimo del singolo oggetto cui si riferisce):

Protégé: <<http://protege.stanford.edu/>>;

OWL: <<http://www.w3.org/TR/owl-features/>>;

RDF: <<http://www.w3.org/RDF/>>;

SWRL: <<http://www.w3.org/Submission/SWRL/>>;

SPARQL: <<http://www.w3.org/TR/rdf-sparql-query/>>;

Jena Api: <<http://jena.sourceforge.net/>>;

Sesame Api: <<http://www.openrdf.org/>>;

OWL-S: <<http://www.w3.org/Submission/OWL-S/>>;

Semantic Web: <<http://www.semwebcentral.org/>>.

¹⁶ Cfr. GAMBARARA, *Comunicazione*, cit.

zione, in termini operativi e strumentali, del modo in cui questi tre oggetti si rendono vivi all'interno di un determinato momento storico culturale, in modo tale da essere *formalizzati*, meglio *modellizzati* all'interno e per mezzo delle tecnologie *ora* a disposizione. Lo sfondo filosofico, linguistico, teoretico in generale, diviene concretizzazione negli strumenti di volta in volta ritenuti i più idonei (anche in relazione al livello di conoscenze tecniche specifiche, da un punto di vista operativo) a svolgere il lavoro per il quale si chiede loro la stessa ragione di esistenza.

In quale modo il *pensiero* viene ad essere catturato in strumenti adatti alla sua gestione? In quale modo il *concetto* viene ad essere oggetto attraverso, per mezzo e tramite il quale costruire tali strumenti, sia in termini di pura progettazione che di singola applicazione (di singola situazione applicativa)? In quale modo i/il termine/i può essere descritto/i come una possibile concretizzazione, formale, all'interno di un linguaggio formale, della relazione tra pensiero e concetto?

Queste domande hanno le loro varie risposte a seconda, in ragione e funzione, della singola fattispecie operativa che si vuole utilizzare, che si ritiene essere la più idonea a formalizzare, quindi, precisare ed organizzare.

3.1 Dalla *vaghezza* alla *precisazione*: la *modellizzazione*

La posta in palio riguarda ciò che si può arrivare a descrivere come modifiche nell'informazione, del contenuto informativo. Modifiche che riguardano necessariamente il processo che dalla *vaghezza tout court* porta alla propria gestione in *forma* di modellizzazione, passando, ma non soffermandosi (o limitandosi) attraverso una modalità¹⁷ di precisazione, che cerchi di circoscrivere al massimo ogni forma di riduzione.

¹⁷ Quella propria di ogni strumento: Thesaurus, Rete/Mappa semantica, Mappa topica, Mappa cognitiva, Ontologia. Alcuni spunti, per quanto ri-

Che cosa guadagno, ad esempio, attraverso la modellizzazione in reti semantiche? Che cosa perdo nella modellizzazione in reti semantiche? Rispondere a queste domande¹⁸ vuol dire incamminarsi lungo un sentiero che possa essere percorso con la luce fornita, di volta in volta, dalla opzione di modello scelta.

Ancora, in che modo, questi modelli, identificano *forme* di concetti e *forme* di significati? Una possibile risposta a queste domande va nella direzione di ciò che può essere identificato come una *new social ontology*: diversi studiosi che pensando tratti e legami in base ai propri risultati contribuiscono alla progettazione, costruzione, di un sistema in grado di estrarre i pesi medi di e per ogni singola occorrenza, in modo da avere un sistema, quindi, con adattamento dinamico¹⁹.

Come abbiamo visto: «[Il linguaggio verbale umano] è essenzialmente prassi, e solo in modo derivato le lingue e i testi linguistici possono essere considerati come oggetti autonomi»²⁰.

Con l'operazione della *modellizzazione*²¹, tramite e per mez-

guarda le Reti semantiche, possono essere rintracciati, per approfondimenti in chiave logico-formale, in MARCELLO FRIXIONE, *Logica, significato e intelligenza artificiale*, Franco Angeli, Milano, 1984.

¹⁸ Ad esempio cfr. RONALD BRACHMAN – HECTOR LEVESQUE, (eds.), *Readings in Knowledge Representation*, Morgan Kaufman, Los Altos (CA), 1985.

¹⁹ «Se potessimo abbracciare la somma delle immagini verbali immagazzinate in tutti gli individui, toccheremmo il legame sociale che costituisce la lingua. Questa è un tesoro depositato nella pratica della *parole* nei soggetti appartenenti a una stessa comunità», DE SAUSSURE, *Cours*, cit., p. 23.

²⁰ GAMBARARA, *Comunicazione*, cit., p. 193.

²¹ Cfr. GIORGIO ISRAEL, *La mathématisation du réel. Essai sur la modélisation mathématique*, Éditions du Seuil, Paris, 1996 trad. it. dell'autore, *La visione matematica della realtà. Introduzione ai temi e alla storia della modellistica matematica*, Laterza, Roma-Bari 2003, nonché dello stesso autore *Modelli matematici. Introduzione alla matematica applicata*, Muzio, Roma, 2002.

zo di essa, entriamo nell'ambito di una processualità che si concretizza in una (forse) possibile distinzione tra *significati* e *concetti* all'interno del contesto specifico nel quale ci troviamo e secondo le linee guida della riflessione che stiamo conducendo²².

3.2 Modellizzazione

Operazione di *modellizzazione*, come operazione propria della interazione tra ciò che è presente nel mondo esterno, nello specifico nel mondo concretizzato del documento e ciò che sono le necessità, le direttive, le possibilità di chi, con gli oggetti del primo tipo, deve costruire un qualcosa che serva da guida ed orientamento funzionale, specifico, determinato per possibili attori²³; con la specificazione della ovvia previsione di ulteriori attori che volessero entrare all'intero di questa stessa dinamica.

Mondo *esterno* e mondo *interno* sono una pluralità di linguaggi: quello naturale e quelli specialistici; con l'aggiunta funzionale ed operativa del linguaggio stesso della modellizzazione²⁴. Linguaggi che sono tutti declinazioni di quello naturale,

²² Specificando il passaggio del par. 2.2: «Il significato non va confuso con il concetto: ogni concetto ha una radice percettiva (il concetto di mela si forma mediante il contatto con le mele), mentre ogni significato si forma nell'intreccio con altri significati (il concetto libertà non nasce dall'esperienza con un fantomatico oggetto 'libertà', perché un oggetto simile semplicemente non esiste come oggetto, e quindi nemmeno se ne può avere una qualche esperienza). Nella CNV i concetti percettivo-operativi prevalgono sui significati: non vi è arbitrarietà in senso tecnico saussuriano, e questa è la migliore prova della piena continuità tra mondo non comunicativo e mondo comunicativo negli animali non umani.», GAMBARARA, *Comunicazione*, cit., p. 198, nota 3.

²³ «Un 'processo interno' abbisogna di criteri esterni», LUDWIG WITTGENSTEIN, *Philosophische Untersuchungen. Philosophical Investigations* (Herausgegeben von G. E. M. Anscombe und R. Rhees), Blackwell, Oxford, 1953, trad. it. a c. di MARIO TRINCHERO – ROBERTO PIOVESAN, *Ricerche filosofiche*, Einaudi, Torino 1995, § 580.

²⁴ Modello come linguaggio specialistico di secondo livello: linguaggio/i

ma che assumono funzioni, si sviluppano lungo dimensioni e determinano relazioni di volta in volta specifiche e differenti. Il modello stesso può essere identificato, quindi, tramite, per mezzo e come una modalità del linguaggio naturale; come linguaggio esso stesso.

Funzioni: le funzioni che assumono i linguaggi specialistici nella e della modellizzazione, sono del tipo non solo *rappresentazionale*, ma, sotto determinati rispetti, *fondativi* essi stessi della complessità della tensione tra significato e concetto. *Rappresentazionali* nel momento in cui necessariamente devono tenere in considerazione ciò che è già presente, ciò che è già stato, in differenti modi e secondo differenti funzionalità, processato ed operazionalizzato. *Fondativi*, perché sulla base della modalità *rappresentazionale*, viene a nascere, come prodotto, ma anche come possibile fonte di ulteriore produzione, una nuova, differente, modalità conoscitiva (che ricomprende le prime determinazioni, ma che nello stesso tempo risulta essere altro); modalità conoscitiva esplicativa e comprendente i contenuti e le forme della conoscenza stessa.

Dimensioni: le singole dimensioni possono essere individuate come le singole discipline che formano il sostrato, la base, delle determinazioni di conoscenza. I linguaggi (specialistici), sotto questo rispetto, sono i processi, i risultati e le possibilità ulteriori della complessità del rapporto tra *concetto* e *significato*, rapporto che è nello stesso tempo di separazione/differenziazione e compresenza/identificazione. Due oggetti, meglio un solo oggetto

specialistico/o a sua/loro volta risulta/no da operazioni di secondo livello sul linguaggio naturale e su quelli specialistici di primo livello; ma che nello stesso tempo si pone/pongono esso/loro stesso/i, ad un differente livello: è/sono linguaggio/i specialistico/i di primo livello. (Il gioco tra singolare e plurale è dato dalla necessità di interrogarsi sulla dinamica multi-uno che può essere espressa dalla analisi degli strumenti quali Thesauri, Reti/Mappe semantiche, Ontologie, etc.).

to su due livelli in relazione e rapporto alla produzione, gestione, uso della conoscenza, delle conoscenze e dei modi in cui tutto questo avviene e viene rappresentato.

Relazioni: le relazioni sono, quindi, una collezione di procedure, di regole, di modalità in cui concretizzare le possibilità fornite dalle interazioni da un lato tra linguaggio ordinario e linguaggi scientifici; tra linguaggi scientifici e linguaggi scientifici dall'altro. Procedure, regole, modalità che sono lo stesso linguaggio/processo di modellizzazione. La tensione tra concetto e significato, in questo modo, non può essere risolta nella scissione in due *entità* nettamente distinte e separate.

Funzioni, dimensioni e relazioni che sono altrettante determinazioni della dimensione di diagramma del linguaggio scientifico, del modello stesso, nei confronti del linguaggio naturale. Questo punto sarà analizzato ulteriormente nel prossimo paragrafo.

Primo: il mondo *esterno* come mondo concretizzato²⁵, il mondo *interno*, come esplicitazione e contestualizzazione, sotto determinati rispetti ed in prima istanza, rilettura del mondo *esterno*²⁶. Secondo: questa processualità, che non deve (dovrebbe) essere mai riduzione, anche se ne può assumere alcuni caratteri e determinare alcune modalità, deve rendere conto del senso stesso della tensione tra i due *mondi*. Tensione che in ultima analisi, risulta essere tra linguaggi. Quindi, due livelli della stessa realtà.

Detto in altri termini, nel momento in cui si deve porre come storicamente fondata a livello teoretico, o quanto meno discutere di una distinzione, di un rapporto, di una tensione tra *significati* e *concetti*, il primo nodo da affrontare riguarda la possibilità che si stia parlando non di differenti tipi di oggetto, mutuamente

²⁵ Il documento, il testo, il dato che si vuole processare, utilizzare, descrivere, come oggetto in un linguaggio in *interazione* con altri linguaggi che lo *concretizzano* congiunturalmente.

²⁶ Le operazioni di processazione, utilizzo, descrizione nel e del modello.

esclusivi, separati nettamente, ma di un qualcosa di diverso. Un qualcosa che tenga conto di uno sfumare del concetto e del significato, dei concetti e dei significati. Meglio, un qualcosa che rappresenta, fonda, si dispiega e costruisce relazioni su differenti livelli della stessa *realtà*.

In questo modo è possibile dare ragione della determinazione delle dimensioni sintattiche e semantiche, delle modalità attraverso le quali, in funzione e ragione delle quali, le modalità sintattiche di questi linguaggi, e le rispettive modalità semantiche siano esse stesse in un rapporto che discende direttamente (e nello stesso tempo ricomprende) da tutto quanto analizzato fino ad ora.

Rapporto tra sintassi: il rapporto tra le sintassi dei modelli e quelle dei linguaggi specialistici (sempre identificati come dimensioni di quello naturale) è un rapporto, in questo caso di isomorfismo *verso il basso*. Con questa descrizione intendiamo una base di partenza che deve essere necessariamente isomorfica, quindi, da un punto di vista processuale ed operativo deve poter delineare la stessa potenzialità di calcolo, la stessa potenzialità di manipolazione formale, la stessa applicabilità di regole di quello di partenza; ma che nel risultato (in questo si precisa la specificazione *verso il basso*) evidenzia un minore numero di oggetti e di modalità di combinazione/gestione di questi stessi oggetti. Esempi espliciti di questo i Thesauri, le Reti semantiche, le Mappe semantiche, le Ontologie e le altre forme assunte dalla modellizzazione.

Se da un lato la riduzione è presente, non è tale nella forma classica di una riduzione di significato partendo dalla descrizione e tramite la precisazione di un concetto; ma piuttosto, nella forma di una riscrittura delle possibilità, meglio dei criteri di possibilità di produzione significativa della *forma* nella dinamica processuale che da una conoscenza acquisita porta verso una nuova modalità di presentazione di tale conoscenza.

Rapporto tra semantiche: si potrebbe assistere ad una intera-

zione, forse anche una relazione tra una semantica del concetto ed una semantica dei significati. Questo vorrebbe dire identificare in maniera netta e separata i due oggetti. Questo è forse ciò che avviene, quanto meno in termini di diagramma a livello della costruzione dei modelli oggi presenti. Tuttavia, la possibilità di una architettura che tenga presente come concetto/i e significato/i siano due livelli della stessa realtà potrebbe essere una sfida verso la quale indirizzare gli sforzi sia teoretici che operazionali.

Rapporto tra sintassi e semantica: in prima istanza sembrerebbe quasi una predominanza della semantica nei confronti della sintassi. Il rapporto, in un certo senso canonico, di formalizzazione della sintassi, quindi, la sovraimposizione di una semantica, sembrerebbe se non ribaltato, quanto meno posto sotto un forte vaglio critico. Una possibile risposta potrebbe essere quella fornita attraverso la modellizzazione presentata da Jackendoff²⁷. Modellizzazione, che, tuttavia, risente, è frutto, di una prospettiva generativista che nel momento in cui accentua determinati caratteri del rapporto sintassi-semantica-cognizione, dall'altro ne tace di altrettanto fondamentali.

La proposta che viene fuori dalla nostra discussione, va invece nella direzione di una co-determinazione di sintassi e semantica. Una co-determinazione che non deve essere necessariamente speculare a quella del linguaggio naturale, ma declinandosi da questa deve essere in grado di tenere in debito conto e *formalizzare* in modo appropriato le dinamiche che abbiamo appena illustrato. Il vincolo tecnologico non è un qualcosa da assumere in maniera dogmatica, ma un qualcosa che ci dice fino a dove siamo arrivati e cosa ancora ci resta da fare.

²⁷ RAY JACKENDOFF, *Semantics and Cognition*, The MIT Press, Cambridge, 1983, trad. it. di MARIA GIOVANNA TASSINARI, *Semantica e cognizione*, il Mulino, Bologna, 1989.

3.2.1 La natura iconico-diagrammatica-metaforica dei modelli di rappresentazione della tensione tra concetto e significato

«Le ipoicone possono essere sommariamente suddivise a seconda del modo di Primità di cui partecipano. Quelle che partecipano di qualità semplici ovvero della Prima Primità, sono *immagini*; quelle che rappresentano le relazioni (principalmente diadiche, o considerate tali) delle parti di una cosa per mezzo di relazioni analoghe fra le loro proprie parti sono *diagrammi*; quelle che rappresentano il carattere rappresentativo di un rappresentamen mediante la rappresentazione di un parallelismo con qualcos'altro sono *metafore*»²⁸.

In quale misura i Thesauri, le Reti semantiche, le Mappe semantiche, le Mappe cognitive, le Mappe topiche, le Ontologie sono in rapporto iconico, diagrammatico, metaforico con le *forme* e le modalità di conoscenza (che identificano una tensione tra i livelli del concetto ed i livelli del significato) che da molte diventano una? Rispondere a questa domanda potrebbe sembrare molto più complicato di quanto in effetti non sia descrivere le differenti interrelazioni tra le componenti di immagine, di diagramma e di metafora che corrispondono alla natura iconica²⁹ di tali modelli.

²⁸ CHARLES S. PEIRCE parti non pubblicate di, *Syllabus of Certain Topics in Logic*, Alfred Mudge & Son, Boston, 1903, pubblicate in CHARLES S. PEIRCE, *Collected Papers* (C. Hartshorne, P. Weiss eds. Voll. I–VI; A. W. Burks ed. Voll. VII–VIII), vol. II, The Belknap Press, Cambridge, 1931-1959, trad. it. a. c. di MASSIMO BONFANTINI, in CHARLES S. PEIRCE, *Opere*, Bompiani, Milano, 2003, § 2.277.

²⁹ Cfr. C.S. PEIRCE, *On the Algebra of Logic. A Contribution to the Philosophy of Notation*, in “*The American Journal of Mathematics*”, 2, vol. 7, 1885, pp. 180-202, ristampato in CHARLES S. PEIRCE, *Collected Papers*, cit., vol. III, trad. it. a. c. di MASSIMO BONFANTINI, *Sull'algebra della logica: un contributo alla filosofia della notazione*, in CHARLES S. PEIRCE, *Opere*, cit., § 3.362; C.S. PEIRCE, *That Categorical and Hy-*

Nel momento in cui, sotto determinati rispetti, sono tutte forme di rappresentazioni originali (con l'aggiunta della dimensione fondativa, quindi, simbolica della funzione), tutte partecipano del particolare tipo di natura iconica. Ad un livello base si potrebbe anche affermare che sono tutte icone con differenti modalità e gradazioni della natura di indice e simbolo delle specifiche *forme* (forma, in senso classico e *contenuto*, come co-determinati e co-determinanti) di conoscenza. Oggetti che presentano, quindi, nel loro essere determinati in termini processuali sia le caratteristiche dell'immagine, che quelle del diagramma, che quelle della metafora.

Immagini: è la modalità rappresentativa più immediata con la quale vengono rappresentate. Si ricorre a vere e proprie immagini per fissare la relazione, la tensione tra significato/i e concetto/i sia per quanto riguarda ogni singola *forma* di conoscenza, sia per quanto riguarda possibili interazioni-sovrapposizioni tra differenti *forme*.

Diagrammi: le declinazioni in *funzioni*, *dimensioni*, *relazioni*, che abbiamo analizzato nel paragrafo precedente sono tutte specificazioni dell'effettiva costruzione diagrammatica della conoscenza attraverso, per mezzo e secondo le modalità specifiche del modello.

Metafore: il livello metaforico è quello più vicino alla determinazione simbolica, quindi, alla convenzionalità ed alla costruzione della tensione con il momento iconico del modello esplicitato nella forma operativa di Thesaurus, Rete semantica, Mappa semantica, Mappa topica, Mappa cognitiva, Ontologia.

Tutto questo non solo nel momento in cui si prende in considerazione una sola possibile modalità di conoscenza, ma l'intero

pothetical Propositions are one in essence, with some connected matters, in ID. *Collected Papers* cit., vol. II, trad. it. a c. di MASSIMO BONFANTINI, in C.S. PEIRCE, *Opere*, cit., § 2.278; C.S. PEIRCE, parti non pubblicate di, *Syllabus of...*, cit., § 2.276.

insieme degli universi del discorso che da molteplici diventano uno, tramite, per mezzo e secondo le modalità dell'immagine, del diagramma e della metafora. Ad un singolo concetto si vuole porre in relazione di rappresentazione (quindi anche fondativa) una pluralità di significati specifici, come si potrebbe anche voler risalire da ogni singolo significato ad una pluralità di concetti che potrebbero *generarlo*. In questo modo ciò che ne risulta contiene il proprio punto di partenza, ma se ne differenzia passando da una somma di singole dimensioni ad una multidimensionalità non necessariamente linearmente determinata.

Entrambe queste specificazioni (concetto-significati; significato-concetti) non sono che due parti di un tutto complesso, analizzarne solo una per arrivare ad una determinazione univoca in senso forte, vorrebbe dire disincarnare non solo i significati: farli diventare all'interno del singolo universo di discorso un qualcosa di completamente distaccato da chi e perché lo produce; ma, nello stesso tempo, non tenere in considerazione i differenti portati esperienziali, di significato, concettuali che ogni singola acquisizione di conoscenza porta con sé, ne è determinata e li ridetermina per diventare un nuovo punto di partenza³⁰.

Possiamo, quindi, concludere questa parte della nostra analisi affermando che «Questa comunicazione richiede un primo grado di metarappresentazioni, e su questa base possono presentarsi stati intenzionali che costituiscono atteggiamenti 'quasi-proposizionali', ricordando sempre i limiti della sua semantica e della sua sintassi»³¹.

3.2.2 Sincronicità dei modelli di rappresentazione

Riprendiamo la citazione di de Saussure del par. 1.1 «È sincronico tutto ciò che si riferisce all'aspetto statico della nostra

³⁰ GAMBARARA, *Comunicazione*, cit., p. 204; 206.

³¹ Ivi, p. 203.

scienza, è diacronico tutto ciò che ha rapporti con le evoluzioni. Similmente, *sincronia* e *diacronia* designeranno rispettivamente uno stato di lingua ed una fase di evoluzione»³².

La sincronia e la diacronia possono, all'interno di questo ambito, essere analizzate e discusse in relazione ai differenti modi ed alle diverse modalità di conoscenza in gioco. Possono, anzi devono, in quanto risultano esse stesse dinamiche e modalità fondanti (da un punto di vista strutturale) le modellizzazioni che rappresentano, quindi, non una riduzione *tout court*, bensì una ridefinizione con un universo di possibilità che ha vincoli più forti nella propria concretizzazione.

Se da un lato si ha un processo sincronico di costruzione di modelli tramite, per mezzo, ed in funzione della tecnologia e delle modalità di rappresentazione della conoscenza presenti in un dato momento storico dello sviluppo della riflessione teorica; dall'altro si ha un'evoluzione di tali costruzioni (tecnologia e modalità di rappresentazione), evoluzione che ripropone di volta in volta nuove articolazioni delle regole di costruzione, che, tuttavia, forniscono risultati commensurabili in *forma* e *contenuto*.

Per quanto riguarda la sincronia, l'analisi che dobbiamo compiere riguarda ciò che succede in un determinato momento storico culturale. Non necessariamente si hanno processi di concretizzazione e determinazione che viaggiano alla stessa velocità³³, identificare un preciso istante dello sviluppo della conoscenza vuol dire fissare una ben precisa *forma* di modellizzazione, quindi, analizzare i comportamenti e le determinazioni delle varie forme di conoscenza che vogliamo far entrare in gioco.

Ogni singola *forma* che si declina in modello rappresenta l'insieme delle conoscenze e delle determinazioni tecnico-teoriche di quel preciso istante.

³² DE SAUSSURE, *Cours*, cit., p. 100.

³³ Sia per quanto riguarda ciò che avviene all'interno di un singolo universo di discorso, sia per quanto avviene confrontando differenti universi.

Gli altri universi di discorso devono, sotto determinati rispetti, diventare la controparte di un rapporto iconico, per come lo si è descritto precedentemente, che presenta come vertice opposto lo stato dell'arte *attuale* della modellizzazione. Ovviamente si può anche scegliere di determinare che in un preciso momento, in presenza di una pluralità di modelli possibili, ad ogni singolo insieme di universi di discorso sia più opportuno far corrispondere uno ed uno solo dei modelli a disposizione.

Strettamente intrecciata con questa modalità della sincronia, vi è quella dell'andamento diacronico del processo di modellizzazione. Prendendo in considerazione questa dimensione si può affermare che sono i singoli universi di discorso, le singole modalità conoscitive, le singole discipline scientifiche, nel loro procedere verso forme nuove di elaborazione teoretica che determinano lo schiudersi di nuove possibilità nella progettazione e costruzione dei modelli.

La dinamica tra sincronia e diacronia, non deve essere vista in termini puri e semplici di cumulatività della/e scienza/e o di rivoluzioni traumatiche nel loro sviluppo/procedere. È possibile arrivare ad una determinata forma di modellizzazione perché vincolati al momento presente di uno sviluppo che non necessariamente procede simultaneamente per ogni singola disciplina.

Il risultato è sempre e comunque che lo stato attuale è determinato da quelli precedenti³⁴ e corrisponde solo ad una differenza nell'articolazione delle funzioni, dimensioni, relazioni che costituiscono il processo di modellizzazione ed il modello stesso.

³⁴ Secondo i modelli del Thesaurus, delle Reti semantiche, delle Mappe semantiche, delle Mappe topiche, delle Mappe cognitive, delle Ontologie.(ad oggi).

4. Costruzioni concettuali

Utilizzando le dinamiche, le processualità tra i differenti linguaggi (ordinario, specialistico/i, modelli), si possono identificare, a questo punto, una serie di modalità di determinazioni e processualità ulteriori, nel momento in cui si prendono in considerazione le varie concretizzazioni che rispondono ai nomi di Thesaurus, Reti/Mappe semantiche, Ontologie³⁵. Oggetti differenti, modalità e determinazioni differenti, ma che, come abbiamo visto, derivano, sono determinate, a loro volta determinano, un rapporto con le conoscenze tecnologiche e teoretiche di un determinato momento nello sviluppo storico-culturale delle differenti discipline in gioco.

Operando tra e con le determinazioni dei concetti e dei significati, nell'oscillazione della determinazione in *forma* di *termine*, si hanno le differenti fattispecie³⁶. Fattispecie che assolvono, determinano e modellizzano secondo la stessa processualità e/o processualità non solo commensurabili, ma del tutto, in forma e contenuto, sovrapponibili.

Come punto di partenza possiamo, quindi, indicare le seguenti declinazioni e determinazioni:

- a. Il Thesaurus come *inclusione* del *termine*;
- b. Le Reti/Mappe semantiche come *rappresentazione* della tensione tra *termine* e *concetto*;
- c. Le Ontologie come *inclusione* del *concetto*.

Queste declinazioni sono date partendo da una *esemplificazione riduzionista*, non del tutto omogenea con il portato delle

³⁵ Per una prospettiva più specificatamente linguistico-computazionale si vedano i lavori dell'I.L.C. (Istituto di Linguistica Computazionale del C.N.R. sede di Pisa), <<http://www.ilc.cnr.it>>.

³⁶ Strumenti di indicizzazione e reperimento a base semantico lessicale in documentazione.

considerazioni sviluppate fino a questo punto dei *desiderata* riguardanti la processualità di costruzione, descrizione, identificazione e presentazione della conoscenza e delle modalità e dinamiche di tali operazioni. Infatti, secondo questa descrizione il significato (come oggetto) viene ad essere, sotto alcuni rispetti, sostituito dal termine (come oggetto). Da sottolineare è il carattere sempre e comunque vincolato al livello teoretico e tecnologico del *presente* nel quale si opera.

Le differenti strade identificate per passare dal livello di pura riflessione ad un livello di pura operatività, necessariamente, fino a questo punto, hanno dovuto sottostare al dazio di una qualche forma di compromesso con la vaghezza, compromesso che ha preso le sembianze di differenti gestioni della ambiguità, essa stessa riduzione della vaghezza del linguaggio naturale e della tensione tra questo ed i linguaggi specialistici. Per dare ulteriore consistenza a queste affermazioni si può notare come i *termini* utilizzati nella precedente identificazione e descrizione delle declinazioni siano termini appartenenti ad un linguaggio logico-formale con possibili connotazioni matematiche.

Esempi di questo primo passaggio verso la costruzione degli effettivi strumenti operativi: Thesauri, Mappe semantiche, Ontologie, possono essere le analisi, gli usi, i trattamenti analoghi a quelli presentati nell'antologia curata da Kittredge e Lehrberger³⁷, sullo studio e l'analisi delle relazioni e dei rapporti tra differenti domini semantici ristretti ed il linguaggio.

Questo risulta essere contemporaneamente un punto di forza ed uno di debolezza delle strumentazioni che di volta in volta prendono il posto ed il ruolo del modello all'interno del processo. Punto di forza in quanto presentano un *range* ed un *target* di utilizzo, quindi, di effettiva applicabilità fortemente ancorato al-

³⁷ RICHARD. KITTREDGE – JOHN LEHRBERGER, *Sublanguage. Studies of Language in Restricted Semantic Domains*, de Gruyter, Berlin-New York, 1982.

la posizione teoretica di riferimento. Una sorta di autovalidazione che permette la possibilità di comprensione e di passaggio di informazione e conoscenza fra livelli differenti. Punto debole, nel momento in cui lo si sostituisce alla complessità della tensione tra significato e concetto all'interno di una riduzione che di volta in volta risulta essere funzionale, ma anche congiunturale e sicuramente non definitiva.

Nel momento in cui si progettano, si costruiscono, si utilizzano questi strumenti, bisogna sempre tenere presente l'aspetto rappresentativo e fondativo analizzati in precedenza. Così come anche che le dinamiche non sono e non si presentano necessariamente in termini lineari, ma si dispiegano lungo dimensioni in interazione anche ricorsiva tra loro; ancora, le relazioni che in questo modo prendono corpo, si concretizzano e formano un insieme di vincoli e possibilità, non statiche, ma dotate di una propria dinamica strutturale.

L'esempio sviluppato in termini insiemistici con fondamento logico-matematico è un buon punto di osservazione e, nello stesso tempo, una buona possibilità di comparazione per valutare non solo la portata degli strumenti, ma anche come alle loro spalle operi una collezione di posizioni scientifiche non ristrette ad un singolo ambito disciplinare. Descrivere i singoli linguaggi, come anche le componenti sintattiche e semantiche di questi, in termini puramente insiemistici è sicuramente una modalità interessante, vincolata, tuttavia, nelle proprie possibilità ad una costruzione ben precisa della matematica e della logica di base.

Passiamo, quindi, ad alcuni esempi che renderanno ancora più espliciti i passaggi argomentati e descritti.

4.1 Thesaurus

Come esempio di Thesaurus, possiamo prenderne in considerazione uno progettato come supporto per una iniziativa della Camera dei Deputati della Repubblica Italiana. Questo esempio illustra come la tecnologia, declinata nel linguaggio di program-

mazione XML, in stretta relazione con la costruzione di un Thesaurus descriva l'intero percorso che abbiamo tracciato.

La costruzione, non solo del corpus del Thesaurus, ma anche della stessa architettura che ne doveva essere, allo stesso tempo, contenitore e possibilità di utilizzo, è stata portata avanti da programmatori ed *esperti* delle varie aree tematiche oggetto della iniziativa. Nel momento in cui si accedeva ad ogni singola relazione, si aveva la possibilità di utilizzo di una serie di *termini* chiave predisposti per la *comprensione* e la *omogeneizzazione* dei contenuti in discussione. Termini chiave a cui si poteva accedere tramite e per mezzo di una serie di *link* incrociati e rimandi contestuali.

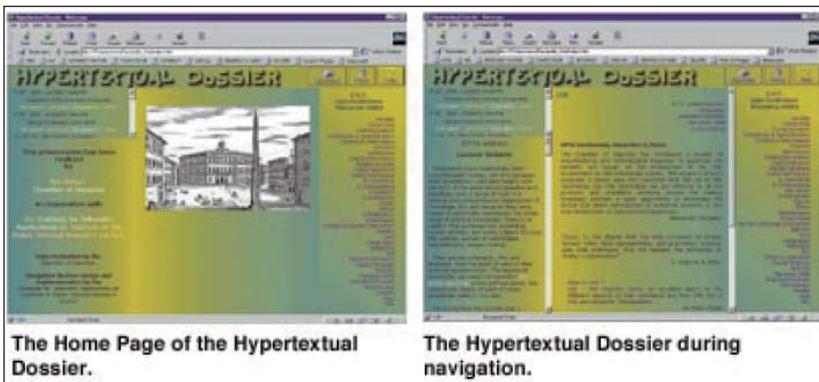


Fig. 1: Esempio di Gestione di un Thesaurus - <http://www.ercim.org>

4.2 Rete Semantica

Come esempio di Rete semantica, possiamo prenderne in considerazione una progettata per la categorizzazione dei concetti rappresentati nel Unified Medical Language System® (UMLS, *Sistema unificato per il linguaggio medico*). Categorizzazione di un'ampia terminologia, i cui referenti, appartengono ad una pluralità di universi di discorso.

les for combining terms extensions to the vocabulary»³⁸.

Definizione 2: «An ontology is an explicit specification of a conceptualization»³⁹.

Definizione 3: «An ontology is a hierarchically structured set of terms for describing a domain that can be a skeletal foundation for a knowledge base»⁴⁰.

Definizione 4: «An ontology provides the means for describing explicitly the conceptualization behind the knowledge represented in a knowledge base»⁴¹.

Tutte e quattro queste definizioni rientrano all'interno della tensione tra sincronia e diacronia. Infatti prendono il via da specifiche posizioni teoretiche che costituiscono la controparte di riferimento per la costruzione dello strumento, modello, identificato come *Ontologia*. Possiamo notare che a seconda del modo di saturare le possibili occorrenze libere identificate ad esempio dai termini *rule*; *concept*; *term*; *knowledge*, si potrebbe parimente costruire, in forma diagrammatica, reti semantiche, mappe semantiche, e le altre *forme* che, in un dato momento, possono risultare essere le più *funzionali*, lungo le *dimensioni* della cono-

³⁸ ROBERT NECHES – RICHARD FIKES – TIM FININ – THOMAS GRUBER – RAMESH PATIL – TED SENATOR – WILLIAM R. SWARTOUT, *Enabling Technology For Knowledge Sharing*, <<http://www.isi.edu/isd/KR-Sharing/vision/AIMag-small.html>>.

³⁹ THOMAS R. GRUBER, *A Translation Approach to Portable Ontology Specifications*, <<http://kls-web.stanford.edu/Ksl-92-71.html>>.

⁴⁰ WILLIAM R. SWARTOUT – RAMESH PATIL – KEVIN KNIGHT – TOM RUSS, *Toward Distributed Use of Large-Scale Ontologies*, <http://ksi.cpsc.ucalgary.ca/KAW/KAW96/swartout/Banff_96_final_2.htm>.

⁴¹ AMAIA BERNARAS – IÑAKI LARASGOITI – JOSE MANUEL CORERA, *Building and Reusing Ontologies for Electrical Network Applications*, in *ECAI96. 12th European Conference on Artificial Intelligence*, John Wiley & Sons, London, 1996, pp. 298-302; citazione in EVA BLOMQUIST, *State of the Art: Patterns in Ontology Engineering*, <http://www.diva-portal.org/diva/getdocument?urn_nbn_se_hj_diva-56-1_fulltext.pdf>.

scienza, per instaurare *relazioni*. Costruire, quindi, strumenti di indicizzazione semantico lessicale per il trattamento dei documenti.

La stessa terminologia utilizzata, in modo specifico nella definizione 3 rappresenta un *uso* della modellistica (anche) logico-matematica della quale si è discusso nel paragrafo precedente.

In ultima analisi, vediamo come la tensione tra significato e concetto si risolve nei termini della costruzione di un modello che è linguaggio, a sua volta, in relazione con quello naturale, con quelli delle singole scienze (universi di discorso) con quello della dimensione tecnica.

Ulteriore esempio può essere quello del MRDA *corpus* sviluppato alla Stanford University in relazione alla architettura NOMOS per l'uso di ontologie. Nello specifico MRDA include concetti quali: *utterances, dialogues acts, etc.*

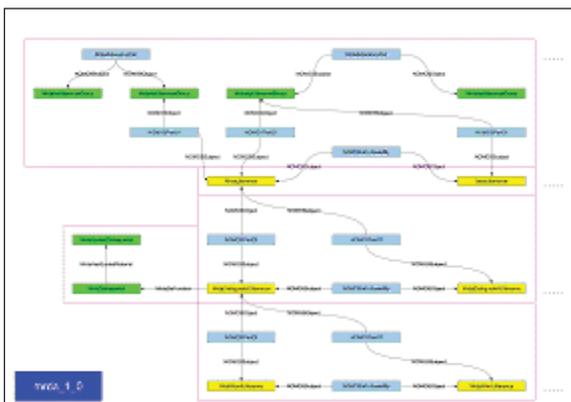


Fig. 3: Esempio di Ontologia - <http://go-del.stanford.edu>

5. Conclusioni

Il rapido esame, gli spunti di riflessione e ulteriore approfondimento, che abbiamo fornito nella presentazione di alcuni tra gli strumenti concettuali attualmente utilizzati, ci permette di iden-

tificare, comunque, una serie di linee guida che li accomuna, ma nello stesso tempo, li differenzia a seconda della loro gestione. I nodi riguardano sempre e comunque la dinamica-tensione tra significato, concetto e termine.

In questo modo, avremo, in aggiunta alle considerazioni precedenti, ed in qualche modo come loro conseguenza, una differente gestione e costruzione degli stessi concetti, come elenchi di concetti più o meno chiusi. I sistemi come sistemi chiusi in determinati casi, con possibilità di maggiore apertura in altri. La necessità di organizzare e costruire percorsi più o meno gerarchizzati, quindi, con una differente costruzione e gestione delle sovra e sotto determinazioni (delle relazioni interne e verso l'esterno dei concetti declinati e/o sostituiti ai termini).

Nel modo in cui si risponde a queste domande, partendo da una posizione teorica iniziale che racchiude, come abbiamo sottolineato più volte, sia l'aspetto di conoscenza teoretica che quello di conoscenza tecnico-empirica, si avranno le differenti fattispecie di strumenti concettuali.

Data Mining e Text Mining

MAURIZIO LANCIA - ANDREA LAPICCIRELLA

Introduzione

Lo sviluppo tecnologico ha consentito di produrre ed immagazzinare enormi quantità di dati con una ricchezza di informazioni potenzialmente accessibili che eccede quanto era nelle intenzioni di chi li aveva raccolti, ovvero aveva progettato la banca dati. Informazioni che possono essere particolarmente preziose per i processi decisionali.

Un esempio eclatante di cosa possa significare conoscenza nascosta viene dal mondo dell'astrofisica. In quell'ambito esistono vere e proprie biblioteche digitali contenenti gli insiemi di dati raccolti con molto dispendio in occasione di campagne attuate tramite l'utilizzo di satelliti. Le "strisciate sperimentali" di questo tipo di dati sono molto ridondanti (*id est* raccolgono moltissimi parametri) ed il loro contenuto informativo eccede di molto quanto strettamente necessario al raggiungimento dell'obiettivo primario per il quale la campagna di raccolta era stata disegnata. Di conseguenza oggi le banche in cui sono raccolti i dati delle campagne suddette sono oggetto di studi ed analisi nell'ambito dei quali si vogliono verificare ipotesi ed idee completamente diverse da quelle che si volevano studiare o provare all'atto del disegno della campagne stesse. In questa maniera, lavorando unicamente su dati già raccolti, si effettuano vere e proprie campagne virtuali (si parla di osservatori virtuali) che spa-

ziano e fondono assieme tutto ciò che già c'è, arrivando a conclusioni e scoperte di conoscenze completamente nuove e assolutamente non considerate all'atto del disegno dell'esperimento originale. È inutile dire che in questa maniera si risparmiano un sacco di soldi ed il rapporto benefici su costo di tali nuove indagini è assolutamente favorevole.

Tuttavia, se non si dispone di strumenti che consentano di individuare ed accedere alle informazioni nascoste, queste enormi quantità di dati rischiano di essere inefficaci per prendere decisioni, arrivando così a costituire più una perdita che un *asset*. Per far fronte a questo problema si è sviluppata negli ultimi anni un'area disciplinare che, prendendo spunto da altre discipline, si è posta l'obiettivo di fornire gli strumenti e le metodologie di estrazione della conoscenza all'interno di banche dati di grandi dimensioni.

Il processo strutturato di estrazione di conoscenza dai dati, ovvero il processo di selezione, esplorazione e modellazione di basi dati molto consistenti al fine di scoprire correlazioni, similitudini (*patterns*), sequenze e tendenze (*trends*) è conosciuto come *knowledge discovery in databases* (KDD).

Il Data Mining è una particolare fase di questo processo e consiste nell'applicazione di analisi di dati e di algoritmi di scoperta che portano all'individuazione di *pattern* tra i dati. Per far questo usa tecniche statistiche ed impiega una o più tecniche di apprendimento computerizzato con lo scopo di produrre un modello che formalizzi le non ovvie conoscenze individuabili nei dati.

C'è molta confusione circa l'esatto significato dei termini 'Data Mining' e 'KDD': vengono spesso considerati sinonimi. Diversi autori tuttavia, nell'ultimo periodo, hanno preferito usare il termine KDD per descrivere l'intero processo strutturato di estrazione della conoscenza ed il Data Mining come la fase di applicazione all'interno del processo di uno specifico algoritmo per l'individuazione dei "patterns".

Il processo di scoperta della conoscenza è strettamente legato allo sviluppi di sistemi di *data warehousing*¹ e può essere considerato una sua estensione, una sua evoluzione. Lo sviluppo di un sistema di Data Mining dovrebbe avvenire partendo da un Data Warehouse già implementato, o comunque da ambienti certificati dove i dati siano stati regolarizzati, in modo che l'analisi possa essere fatta su dati accurati, integri e omogenei, il più possibile "ripuliti" da incongruenze che potrebbero influenzare la precisione dei risultati.

Il Text Mining consiste nell'estrazione di *pattern* utili da testi liberi. Non è altro che l'applicazione di tecniche di Data Mining a testi non strutturati (agenzie stampa, pagine web, *e-mail*, libri e articoli in versione digitale, ... più in generale a qualsiasi *corpus* di documenti). Questo tipo di applicazione è particolarmente utile per individuare gruppi tematici, classificare i documenti, scoprire associazioni nascoste (legami tra argomenti, o tra autori, *trend* temporali, ...), addestrare motori di ricerca, estrarre concetti per la creazione di ontologie (*ontology learning*).

Il processo di estrazione della conoscenza (KDD)

Il processo di estrazione della conoscenza è un processo interattivo ed iterativo e, indipendentemente dal tipo di applicazione specifica, può essere schematizzato attraverso le seguenti fasi:

1. Definizione dell'obiettivo
2. Selezione dei dati
3. *Pre-processing* (Pulizia dei dati)
4. Trasformazione e Riduzione

¹ Il Data Warehouse è una collezione di dati finalizzata, integrata, soggetta alla variazione del tempo e non volatile, per il supporto del processo decisionale (W.H. Immon, Building the Data Warehouse, Wiley, New York, 1996).

5. Data Mining
6. Interpretazione / valutazione dei risultati
7. Rappresentazione dei risultati

La natura iterativa del processo viene evidenziata dallo schema seguente che mette in evidenza come la fase di valutazione può portare da una semplice ridefinizione dei parametri di analisi utilizzati, ad una ridefinizione dell'intero processo (a partire dai dati estratti).

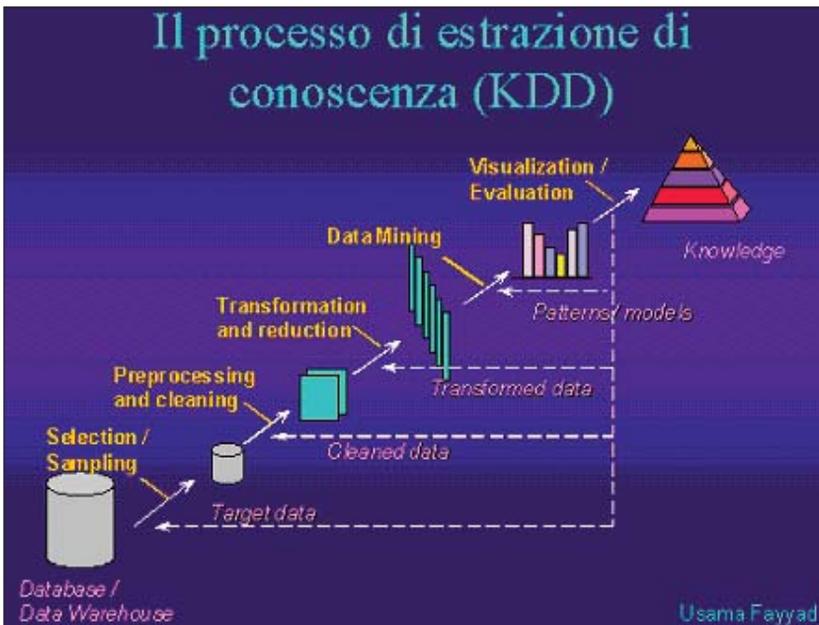


Fig. 1. Il processo KDD dal sito <http://www.cineca.it>

All'interno del processo di estrazione della conoscenza alcune funzioni, come l'individuazione dei *pattern*, possono essere svolte automaticamente, altre richiedono il coinvolgimento di

varie professionalità (esperti del dominio applicativo, specialisti in analisi dati, informatici).

Definizione degli obiettivi

Nel processo di estrazione della conoscenza, per ottenere i migliori risultati è importante comprendere al meglio il dominio applicativo di riferimento ed individuare chiaramente gli obiettivi che si intendono raggiungere attraverso tale attività.

La perfetta definizione degli obiettivi da raggiungere è forse la fase più delicata del processo e fondamentale per la riuscita dell'intero progetto. Condizione essenziale è parlare con i cosiddetti esperti del settore (esperti di dominio) per capire con esattezza cosa va fatto e quali sono le esigenze da soddisfare. Si può ad esempio voler incrementare il numero di possibili clienti che risponderanno positivamente ad una campagna promozionale. Ciò si traduce in due differenti obiettivi: “incrementare il tasso di risposta” e “incrementare il valore di una risposta”. Per essi sono necessari due modelli completamente diversi.

Selezione dei dati

Il secondo passo del processo consiste nella selezione dei dati su cui sarà attivato il processo di scoperta. In questa fase risulta particolarmente importante individuare le fonti da cui reperire i dati necessari per il raggiungimento degli obiettivi precedentemente definiti e valutare l'eventuale ambiente informatico che ne consenta la gestione. L'integrazione dei dati provenienti da fonti di dati diverse – che possono essere interne o esterne, in funzione della necessità di utilizzare nuovi elementi non presenti nel sistema informativo – deve portare alla creazione di una matrice di dati (*data set*) disegnata coerentemente con il modello dati definito sulla base degli obiettivi individuati.

Un ambiente che favorisce l'integrazione dei dati è il Data Warehouse. Fornisce un unico punto di accesso a dati di tipo operativo che possono avere un immenso valore strategico nel

processo decisionale se opportunamente organizzati e certificati. Se non si dispone di un Data Warehouse, si dovrà ricorrere all'utilizzo di alcuni *tool* che offrono la possibilità di pre-elaborare i dati e prepararli per la fase successiva. Da sottolineare comunque che, non è possibile utilizzare un Data Warehouse direttamente per il Data Mining, poiché è necessario un gran lavoro di preparazione dei dati che devono essere utilizzati per le analisi.

Pre-processing (Pulizia dei dati)

Una volta selezionati i dati dalle varie fonti interne ed esterne ed organizzati in un unico contenitore (la matrice di dati) si rende necessario valutarne la qualità e procedere ad una attenta preparazione dei dati.

Poiché i modelli di Data Mining si pongono l'obiettivo di cogliere caratteristiche presenti nei dati e non note a priori, è molto importante individuare eventuali contaminazioni e cercare di porvi rimedio per non correre il rischio di trarre conclusioni che siano basate su caratteristiche che derivano da vizi nella raccolta o nella registrazione.

È particolarmente applicabile in questo ambito il detto "Garbage In, Garbage Out" che sta a significare che se si ha della spazzatura in ingresso al sistema non si può che ottenere della spazzatura in uscita.

È quindi necessario effettuare uno studio preliminare che consenta di identificare le caratteristiche dei dati ed eliminare le eventuali contaminazioni presenti al fine di eseguire la fase di Data Mining in modo effettivo ed efficiente. Solitamente si individuano tre categorie di problemi:

- *valori mancanti*: la presenza di dati mancanti può essere affrontata in modi diversi. Nella maggior parte dei casi essi indicano informazioni perse o non inserite all'origine. Alcune tecniche consentono di trattare i valori mancanti scartando i *record* contenenti valori mancanti oppure me-

dianete sostituzione del valore mancante con valori calcolati (es. la media dei valori delle altre osservazioni corrispondenti) o rilevati per altre osservazioni molto simili. Queste soluzioni, tuttavia, potrebbero non essere coerenti con gli obiettivi definiti.

- *dati anomali*: rappresentano errori causali nei valori delle variabili. Quando si analizzano grandi *data set* tali valori possono assumere svariate forme e tipologie. Possono essere presenti dati replicati (*record* doppi), valori errati (non ammissibili) e/o valori che si discostano moltissimo dagli altri valori, collocandosi al di fuori dell'intervallo atteso, e che, infilandosi malignamente nelle analisi, possono modificare significativamente i risultati. Purtroppo non si hanno delle tecniche precise per la loro gestione.
- *dati incerti*: la precisione dei dati è di particolare rilevanza nei sistemi di scoperta della conoscenza. Se i dati si trovano in una situazione molto grave di incertezza, la migliore soluzione è riciclare sull'individuazione delle fonti per ottenere dei dati migliori.

Come già detto la “pulizia” dei dati, in molti casi, può essere fatta automaticamente. Non è tuttavia realistico pensare di essere capaci di rimuovere tutte le contaminazioni in un'unica fase: alcune anomalie nei dati potranno essere scoperte solo durante il processo di Data Mining stesso. Ciò dimostra la natura iterativa del processo.

Trasformazione e Riduzione

Un trattamento dei dati particolare è richiesto per quegli strumenti di Data Mining che possono accettare solamente specifici valori in *input*. Ad esempio classificatori come le reti neurali e le regole di associazione – che verranno descritti in seguito – riescono a lavorare meglio se i dati hanno un valore numerico e sono compresi tra 0 e 1. In questo caso vanno effettuate operazio-

ni di normalizzazione dei dati cambiando i valori numerici, per fare in modo che essi cadano all'interno di un *range* specificato, e di conversione dei dati trasformando i dati categorici in dati numerici.

Quando il numero di variabili o caratteristiche che descrivono i dati da analizzare è particolarmente elevato può essere necessario applicare alcune tecniche che consentono di selezionare solo quelle caratteristiche che sono utili e rilevanti, nel processo di Data Mining, scartando quelle che sono ridondanti e/o insignificanti. Alcuni algoritmi di Data Mining, infatti, non possono analizzare dati con molte variabili, ed in alcuni casi l'eccessivo numero di variabili può rallentare le prestazioni *run-time* della costruzione dei modelli.

La riduzione del numero delle variabili oltre a portare ad un miglioramento delle prestazioni può facilitare la visualizzazione dei dati e favorire la selezione del modello da utilizzare. L'attività di identificazione delle variabili dipendenti, indipendenti e correlate viene indicata col termine di "data pruning".

Data Mining

È la fase del processo KDD in cui vengono applicati iterativamente particolari metodi per estrarre conoscenza ovvero cercare interessanti *pattern*, regole, o sequenze ripetute all'interno di grandi quantità di dati.

La conoscenza estratta da una sessione di Data Mining è data da un modello o da una generalizzazione dei dati. Un modello non è altro che la rappresentazione delle relazioni chiave che sono scoperte durante il processo di Data Mining. In pratica, è una topologia delle relazioni che mappa quali condizioni di *input* hanno influenza su specifiche condizioni di *output* e come alcune relazioni possono influenzare significativamente altre relazioni. È la rappresentazione formale, attraverso un'espressione matematica, di un certo aspetto o fenomeno e ne riproduce le caratteristiche essenziali: nel nostro caso spiega l'effetto degli *input*

sugli *output*. Come tale, può essere raffinato e messo a punto per una maggiore precisione attraverso un processo iterativo di comprensione dei dati che sono alla sua base.

Scelta dei metodi

Il primo passo di questa fase consiste nell'individuazione di quei metodi che meglio possono portare agli obiettivi che si vogliono raggiungere e che meglio si adattano al tipo di dati da analizzare all'interno di una strategia di estrazione. Esistono numerose tecniche (metodi) di Data Mining. Tuttavia tutti i metodi di Data Mining applicano l'apprendimento basato sull'induzione. Questo è un processo che genera definizioni di concetti generali sull'osservazione di esempi specifici che necessitano di essere appresi.

Le strategie di Data Mining possono essere classificate come strategie supervisionate e strategie non supervisionate. Il Data Mining supervisionato è un approccio *top down* applicabile quando sappiamo cosa stiamo cercando e assume spesso la forma di modelli previsionali. In pratica vengono costruiti modelli tramite l'utilizzo di attributi di *input* per predire i valori degli attributi di *output*. Il Data Mining non supervisionato è un approccio *bottom up*, vale a dire in cui si lascia che i dati stessi indichino un risultato. Non esiste un attributo di *output*: tutti gli attributi utilizzati per costruire i modelli sono variabili indipendenti (attributi di *input*). Quando l'apprendimento non è supervisionato, una volta individuati i modelli, dipende dall'utente stabilirne l'importanza.

Le strategie di apprendimento supervisionato possono essere ulteriormente classificate sia in funzione del fatto che gli attributi di *output* siano discreti o categorici, sia in base al fatto che i modelli siano progettati per descrivere una condizione corrente o per predire una condizione futura. In ogni caso occorre dire che

le due strategie non si escludono a vicenda, anzi in molti casi è necessario ricorrere ad entrambi.

Una strategia di Data Mining si applica a un insieme di dati utilizzando una o più tecniche (metodi). Una tecnica specifica di Data Mining è definita da un algoritmo e da una struttura di conoscenze associata (es. alberi, insieme di regole). I metodi sono applicabili a qualsiasi ambito di indagine, tuttavia la scelta di quale metodo utilizzare nella fase di analisi dipende essenzialmente dal tipo di problema oggetto di studio e dal tipo di dati disponibili per l'analisi.

Sulla base di tali considerazioni, per orientarsi meglio nella scelta dei metodi alcuni dei quali saranno descritti in seguito, si preferisce² suddividere gli stessi in quattro grandi classi:

- *Metodi esplorativi* che presentano forti analogie con le tecniche di tipo OLAP (On Line Analytical Processing) e con i metodi propri dell'analisi dei dati (*query tools*). Tali metodi si basano su metodologie interattive e, solitamente, visuali, che hanno lo scopo di trarre le prime conclusioni ipotetiche dalla massa di dati disponibili, oltre che fornire indicazioni su eventuali trasformazioni dei dati disponibili.
- *Metodi descrittivi* (detti anche non supervisionati), basati su modelli simmetrici, privi di ipotesi di causalità, che hanno lo scopo di descrivere l'insieme dei dati, in un modo parsimonioso. Attraverso tali metodi è possibile sintetizzare delle osservazioni che vengono pertanto classificate in gruppi non noti a priori (analisi *cluster*, mappe di Kohonen) o delle variabili, che vengono tra loro relazionate, secondo legami non noti a priori (modelli log-lineari, modelli grafici).
- *Metodi previsivi* (denominati anche supervisionati): grup-

² PAOLO GIUDICI, *Data mining e statistica*, in «Statistica & Società», (2004), vol. III, n. 1.

po di metodologie con l'obiettivo di spiegare una o più variabili in funzione di tutte le altre, ricercando, sulla base dei dati, delle regole di segmentazione e di valorizzazione (*scoring*) delle osservazioni.

- *Metodi locali* dove l'obiettivo dell'analisi non è, come nei casi precedenti, la descrizione delle caratteristiche del *database* nel suo complesso (analisi globale), ma l'individuazione di caratteristiche peculiari, relative a sottoinsiemi di interesse del *database* (analisi locali).

Interpretazione, valutazione dei risultati

Lo scopo dell'interpretazione e della valutazione è determinare la validità di un modello e la relativa applicabilità a problemi esterni all'ambito del test. Se si ottengono risultati accettabili si trasforma la conoscenza acquisita in termini comprensibili dagli utenti. La fase di interpretazione, attraverso un processo iterativo, può suggerire di ritornare ai passi precedenti per ulteriori attività di raffinamento. In questa fase si può ricorrere a tecniche di visualizzazione per analizzare i modelli estratti.

Una volta che il modello è stato creato e raffinato ad un accettabile grado di accuratezza, esso può essere usato in due modi:

1. in modo **descrittivo**, permettendo agli utenti finali di studiare le relazioni scoperte tra i dati per migliorare la comprensione dei fattori chiave che influenzano il *business*;
2. in modo **predittivo**, determinando la più probabile condizione di output associata agli *input* forniti.

Rappresentazione dei risultati

L'ultimo obiettivo del processo di estrazione della conoscenza consiste nell'utilizzazione di ciò che è stato appreso attraverso la sua applicazione ad altri sistemi per le azioni del caso.

L'attività di *reporting* dei risultati è particolarmente importante. Essa può assumere diverse forme. In genere è possibile

usare un qualsiasi *report writer* o *tool* grafico per rendere accessibili i risultati del processo.

L'attività di *reporting* si sviluppa attraverso due differenti funzioni:

- analisi dei risultati degli algoritmi di *pattern recognition* (riconoscimento di forme);
- applicazione dei risultati dell'algoritmo di *pattern recognition* a dati nuovi.

Non si vuole solamente esaminare ciò che si è appreso, ma si vorrebbe applicare le informazioni classificate e segmentate che si sono ottenute. In molti casi il *reporting* può essere effettuato usando i tradizionali *query tools* per *database*, tuttavia, si stanno affermando nuove tecniche.

Data Mining (Modellistica e Tecnologie)

La fase di “scoperta” ed estrazione della conoscenza da un coacervo amplissimo e multiforme di informazioni variamente raccolte e variamente immagazzinate è certamente il momento più affascinante ed importante dell'intero processo di KDD. In questa fase emerge come fondamentale il Data Mining che, a propria volta, oggi è sinonimo di un insieme multiforme di algoritmi, anche indicati con il termine di metodologie poiché si fa riferimento al metodo generale cui l'algoritmo si richiama, ognuno dei quali garantisce l'ottenimento di specifici obiettivi significativi anche in ambiti più generali quali:

- *il machine learning*;
- *la pattern recognition*;
- *il data cleansing*;
- etc.

Come sopra accennato il Data Mining è un mondo composito

nel quale diverse famiglie di algoritmi di analisi lavorano in modo differente sull'intera massa di dati e colgono di volta in volta obiettivi specifici e parzialmente differenti l'uno dall'altro. Le singole estrazioni di dati ottenute tramite l'applicazione di una data metodologia possono essere integrate con estrazioni derivate dall'applicazione di una metodologia differente e così modificate possono far da base all'applicazione di una terza metodologia. In questa maniera si viene a creare quella che può essere definita come una vera e propria catena di scoperta della conoscenza.

All'interno delle metodologie riteniamo interessante analizzare in modo particolare i seguenti metodi:

- *i query tools*;
- la visualizzazione;
- gli strumenti OLAP (On-Line Analytical Processing);
- I K primi vicini (*the k-nearest neighbors*);
- le regole di associazione;
- gli alberi di decisione;
- le reti neurali;
- gli algoritmi genetici.

I query tools

Questi sono gli strumenti più semplici ed essi dovrebbero essere usati come primi anelli di quella catena di scoperta della conoscenza cui prima si accennava. Nei fatti, come quando esplorando per la prima volta una landa sconosciuta si cercano i riferimenti fondamentali, così nella ricerca della conoscenza nascosta in un insieme multiforme di dati si inizia, laddove la struttura del DWH lo permetta, con delle *query SQL* dirette con cui si estraggono grosse masse di dati che rispondono a livello macro agli obiettivi finali che ci si proponeva. In questo modo si arriva a scoprire l'80% circa dell'informazione che si cerca. Il 20% dell'informazione ricercata, peraltro quella che potrebbe essere di natura vitale per il raggiungimento dell'obiettivo finale, rima-

ne nascosta e deve essere raggiunta tramite tecniche ed algoritmi più avanzati.

Nonostante queste limitazioni, come primo strumento di ricerca i query tools permettono di esplorare il “terreno” in maniera sicura, evidenziando dettagli primari e strutture generali di riferimento, elementi questi che serviranno di base per l'applicazione di strumenti di ricerca via via più sofisticati.

La Visualizzazione

Un altro strumento, peraltro più moderno dei Query Tools, di prima esplorazione dell'insieme di dati per trovare *pattern* ovvero regolarità nella struttura dei dati stessa è la visualizzazione.

Per capire la natura dei processi di visualizzazione poniamoci di fronte ad un classico problema di *marketing*: capire chi è propenso ad acquisire un dato bene. Per far ciò possiamo farci aiutare creando un semplice grafico cartesiano di *data scatter*, funzione di grafica ormai standard anche in Excel. I passi che portano dal nostro insieme di dati ad una loro rappresentazione in un *data scatter diagram* significativo per il nostro proposito finale sono i seguenti:

- Si individua nel *data base* quell'insieme di persone per cui sono noti due differenti attributi di natura numerica, per esempio le persone di cui si conoscono al contempo l'età e l'entità delle loro entrate personali annue;
- Per le persone individuate in questo insieme si individua, sempreché detta informazione sia presente nel *data base*, il valore della caratteristica che si vuole investigare i.e. la propensione o meno ad acquistare un dato bene;
- Si rappresentano le singole persone come punti in un piano cartesiano in cui l'ascissa rappresenta l'età e l'ordinata rappresenta l'entità delle entrate personali annue. Le singole persone trasformate in punti si posizionano nello spazio cartesiano a due dimensioni in funzione della loro specifica età ed entità delle entrate;

- Si può dare ai punti/persona un colore differente a seconda che la singola persona rappresentata sia propensa o non sia propensa ad acquisire il bene su cui si sta svolgendo l'indagine di mercato.

Se il quadro grafico finale vedrà punti di colore differente distribuirsi sul piano in maniera assolutamente casuale e sovrapposta, ciò vorrà dire che età ed entità annue delle entrate personali non sono due variabili correlate all'acquisizione del bene voluto.

Laddove, invece, emergano insieme definiti di punti di ugual colore confinati a diverse zone dello spazio cartesiano rappresentato, ciò sta a significare che età e entità delle entrate sono correlabili alla propensione delle persone ad acquisire il prodotto in questione e laddove il nostro *data base* contenga un campione sufficientemente ampio di persone possiamo ragionevolmente pensare di applicare la relazione individuata (la propensione o meno ad acquisire un dato prodotto in funzione di età e reddito) a persone non contenute nel campione ma che possiedano le stesse caratteristiche.

Il processo può essere esteso a coppie di caratteristiche numeriche delle persone differenti da età ed entità delle entrate (ad esempio distanza dell'abitazione dal centro commerciale e numero di figli idi età inferiore ai diciotto anni presenti in famiglia), sempreché detti attributi abbiano la natura di un indice numerico continuo, per poi fare considerazioni analoghe a quelle sopra descritte.

Si può passare alla grafica tridimensionale esplorando triplette di caratteristiche e si può utilizzare una tavolozza di colori o sfumature per rappresentare i singoli punti in funzione della caratteristica che si sta investigando (ad esempio rappresentare in una scala a quattro colori la propensione ad acquisire un dato bene, sempre laddove questa informazione numerica sia presente nel nostro *data base*) etc.

Da quanto sopra descritto si capisce come la visualizzazione rappresenti uno strumento di primo intervento di esplorazione del nostro spazio dei dati che sostituisce l'analogo ma più antico strumento dei *query tools*.

L'accoppiamento tecnologico di grafica avanzata e *query* dati sta negli ultimi anni ricevendo un grande impulso ed è solo recentemente che sono arrivati sul mercato strumenti quali l'*object oriented three-dimensional tool kit*, ovvero la grafica tridimensionale interattiva con la quale si rappresentano in tre dimensioni i risultati di *query* che individuano oggetti con tre attributi numerici che servono per l'appunto a definire le tre coordinate spaziali dell'oggetto da rappresentare.

Una volta che siamo passati dal dato alla sua posizione in uno spazio (bi o tridimensionale) è possibile assumere che punti vicini l'uno all'altro siano simili ed a volte è possibile individuare dei *pattern* interessanti ovvero insieme significativi di punti anche con un semplice esame visivo.

A volte, invece, l'individuazione di "nuvole" di punti che emergano dal tutto in modo statisticamente significativo deve essere ricercato mediante algoritmi specifici. Ricordiamoci a questo punto di cosa succede se vogliamo esplorare in via grafica oggetti in funzione di un numero di caratteristiche superiori a tre. È lo stesso problema che si ha se si vuole rappresentare graficamente una funzione di quattro variabili:

$$f(x, y, z, t).$$

Per farlo in maniera completa ed assumendo che si possa rappresentare in via grafica solo una funzione di al massimo tre coordinate (x,y,z), si devono fissare valori successivi di t. Per esempio porre t=1 e rappresentare interamente la f(x,y,z,1). Quindi porre t=2 e rappresentare interamente la f(x,y,z,2).

Si può ripetere la stessa procedura, esplorando passo passo l'intero intervallo nell'ambito del quale la variabile t è compresa

e così si arriva a rappresentare interamente la funzione a quattro variabili.

La stessa struttura di azione si può adottare nel nostro spazio dei dati solo che gli attributi (i.e. il numero di variabili indipendenti della nostra funzione) in funzione dei quali dobbiamo esplorare lo spazio a volte sono molte di più di quattro, ciò pone il problema di scelta del “taglio” rispetto al quale rappresentiamo la funzione: i.e. la scelta di quali attributi svolgono il ruolo di x,y,z e di quali invece svolgono il ruolo della t fissata, di volta in volta, ad un dato valore.

L'OLAP

Abbiamo visto come la visualizzazione ci abbia naturalmente portato a considerare il concetto di dimensionalità del dato; nei fatti una tabella costituita da n colonne che contengono ognuna un attributo indipendente dagli altri e da una $n+1$ -ma colonna in cui è contenuto un dato che possiamo considerare una variabile dipendente dagli n precedenti attributi può essere vista alla stregua di una funzione in uno spazio a n dimensioni.

Se ci si pensa, i *managers* in genere fanno sempre domande per la cui risposta è necessario analizzare un problema a n dimensioni. In genere essi non si limitano a chiedere solamente quanto è stato venduto nel complesso il mese scorso (la domanda a dimensione zero), ma anche che tipo di rivista è stata venduta ogni mese, per regione e che età aveva il compratore (domanda a quattro dimensioni), vendita in funzione di prodotto (tipo di rivista), tempo (ogni mese), zona geografica (per regione) ed età dell'acquirente.

Questo tipo di informazione è per propria natura intrinseca multidimensionale e le relazioni in questi contesti sono difficili ad individuarsi laddove l'unico strumento di rappresentazione ed indagine siano tabelle a due dimensioni. Peraltro, quando dobbiamo trovare le relazioni fra più dimensioni nemmeno i *data base* relazionali standard ci aiutano (i.e. i tradizionali *query*

tools). Nei fatti, l'SQL identifica i *record* tramite chiavi e il concetto di multidimensionalità si traduce nel concetto di chiavi multiple per singolo *record*; purtroppo c'è un limite al numero di chiavi che in *data base* relazionali possono essere associate ad una singola tabella mentre non vi è un limite alla fantasia del *manager* nel chiedere dati aggregati per insiemi di dimensioni ogni volta differenti.

Il Manager ad un dato tempo t chiede i dati delle vendite in funzione di area geografica, età dell'acquirente e relative entrate annue, al tempo t più un minuto lo stesso Manager chiede sempre i dati di vendita ma questa volta in funzione di carta di credito ed età, possibilmente interrogando *on line* il *data base* di gestione delle casse dei supermercati della compagnia.

Si può ricorrere ai cosiddetti *tool* OLAP. Si tratta di strumenti che per prima cosa prelevano i dati contenuti nel *data base* e poi li portano in memoria in un formato intrinsecamente multi dimensionale; ciò fatto possono rispondere a qualsivoglia stravagante domanda venga eventualmente fatta dal Manager di cui sopra. È ovvio che, proprio perché è avvenuto un processo di copia dal *data base* originale a quello immagine in memoria, si possono perdere i cambiamenti che in quello stesso istante avvengono nel *data base* parente. Si deve, perciò, con una certa frequenza aggiornare la copia in memoria al fine di avere dati il più possibile freschi. Sarebbe peraltro impossibile lavorare direttamente sul *data base* origine perché i *tool* OLAP lancerebbero *query* così complesse da "far sedere" DBMS anche potenti ed attrezzati, con grave degrado delle prestazioni gestionali e conseguente allungamento delle file di clienti alle casse del supermercato sopra citato.

I *tool* OLAP rivestono un ruolo importante nel Data Mining anche se sono caratterizzati da un limite intrinseco importante: i Tool OLAP non imparano, essi non creano nuove conoscenze e non possono ricercare nuove conoscenze. Forniscono dati in risposta a domande complesse ma non offrono o individuano interpretazioni.

I K primi vicini

Con questa notazione un poco criptica si indica una famiglia di algoritmi che ancora una volta trasportano in uno spazio cartesiano i nostri dati e ci fa arrivare a delle conclusioni sulla base della vicinanza dei punti.

Per fare un esempio pratico si pensi a due differenti oggetti che possiedono i valori di tre attributi numerici molto simili fra di loro. Una volta che detti attributi siano tradotti in coordinate e che tramite questa trasformazione essi possano venire rappresentati in uno spazio cartesiano a tre dimensioni i due punti di cui sopra verranno rappresentati uno molto vicino all'altro.

Rifacciamoci, estendendolo, all'esempio già fatto con due coordinate: prendiamo due persone che sono ugualmente propense ad acquisire un bene e ci accorgiamo che tre attributi che caratterizzano queste due persone (età, entità delle loro entrate annue e quantità del loro indebitamento con terzi) sono molto simili. Ecco che se adesso andiamo a rappresentare le due persone in uno spazio cartesiano a tre dimensioni (età, entità delle loro entrate annue e quantità del loro indebitamento con terzi) i punti che per l'appunto rappresentano le due persone appariranno l'uno vicino dell'altro. Abbiamo così definito il concetto di "vicinato" (*neighborhood*): *record* che, una volta rappresentati nello spazio cartesiano adottando la metodologia suddetta, appaiano l'uno vicino all'altro si dice che appartengono allo stesso vicinato.

Poniamoci ora il problema di voler predire il comportamento di un insieme di persone relativamente, per esempio, alla possibile acquisizione di un bene, e per far ciò abbiamo a disposizione un *data base* in cui ciascun potenziale cliente è caratterizzato da una serie di attributi (ad esempio età, entità delle loro entrate annue e quantità del loro indebitamento con terzi etc.).

L'ipotesi base che guida la nostra predizione è che persone dello stesso tipo è probabile si comportino in maniera simile ri-

spetto ad una eventuale acquisizione del bene. Nella nostra traduzione attributi \Rightarrow coordinate, l'appartenere ad uno stesso tipo può essere tradotto nell'appartenere ad una stessa regione dello spazio, *id est* appartenere allo stesso vicinato. Basandoci su questa semplice constatazione è possibile costruire un potente algoritmo di auto apprendimento denominato: i K primi vicini (*the k-nearest neighbors*).

La filosofia base dell'algoritmo è la seguente: in genere una persona fa ciò che vede fare al suo vicino. In altre parole laddove si voglia attribuire una misura ad un dato comportamento di un punto/persona (i.e. la propensione all'acquisizione di un dato bene), questa viene calcolata facendo la media della misura effettiva di quello stesso comportamento nei 10 punti/persona che sono più vicini nel *data space* (i.e. nello spazio cartesiano nel quale rappresentiamo i punti persona) al punto in questione.

La lettera *k* nel nome dell'algoritmo "i k primi vicini" sta ad indicare quanti punti (*k* per l'appunto) si prendono per calcolare la proprietà voluta sulla base della media delle proprietà effettive (i.e. sperimentali, ovvero i valori naturali di quella proprietà contenute nel *data base* originale). Nell'esempio sopra fatto si lavora con un algoritmo a 10 primi vicini.

Il principale vantaggio della presente metodologia è il seguente: è una tecnica di ricerca che ha come proprio riferimento interno lo stesso insieme di dati su cui si effettua la ricerca. Peraltro gli svantaggi, soprattutto dal punto di vista della complessità numerica, sono non pochi e fra di essi maggiormente rilevano:

- In un insieme di oggetti n , per poter decidere quali sono i k primi vicini di una dato oggetto bisogna misurare la distanza fra l'oggetto preso in considerazione e gli altri $n-1$ oggetti che fanno parte dell'insieme. Ciò porta a dover calcolare $n-1$ distanze per ogni oggetto dell'insieme e di conseguenza circa n^2 distanze se si deve prendere in con-

siderazione l'intero insieme. La complessità quadratica è molto pesante; basti pensare che per un insieme di un milione di oggetti si devono fare mille miliardi di operazioni di definizione della distanza. In genere la complessità degli algoritmi migliori nel Data Mining hanno complessità dell'ordine $n \cdot \log(n)$ dove n è il numero di *record* o oggetti nel Data Base. Proprio per questo motivo si tende ad usare questa metodologia su piccoli sottoinsiemi degli oggetti da analizzare (i.e. per indagini di natura cosiddetta locale);

- Il numero di attributi di un oggetto e, quindi, il numero di “dimensioni” è spesso molto più alto di tre. Lavorare in uno spazio altamente multidimensionale porta un elevato grado di complessità poiché esso, fra l'altro, elude le nostre percezioni abituate a lavorare solo in tre dimensioni. Per esempio in tre dimensioni un milione di punti affollano lo spazio occupabile in maniera tale da avere veramente un punto non molto distante dall'altro. Mentre quello stesso milione di punti in uno spazio a 20 dimensioni lo occupa con una densità molto bassa ed è difficile parlare a questo punto di vicinanza di punti e quindi di similarità di comportamento di punti/individui.

Ancora una volta si riconferma il seguente principio: non esiste algoritmo di Data Mining in assoluto migliore di tutti gli altri e per ciò stesso tale da far con sicurezza trovare ciò che si cerca. La realtà è la seguente: si devono di volta in volta usare algoritmi specifici per ottenere dei risultati parziali nell'ambito di quella catena di scoperta della conoscenza sopra citata.

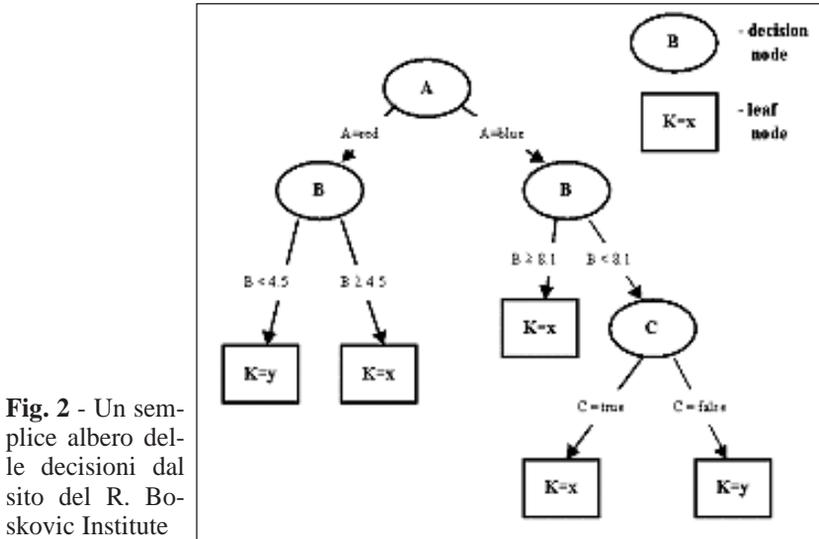
Gli alberi di decisione

Per descrivere questo algoritmo partiamo, come al solito, pensando di dover trattare una tabella nella quale è data la propensione di n individui ad acquisire un bene, l'automobile ad esem-

pio, ed al contempo l'individuo è caratterizzato da tre ulteriori attributi, i.e. dati numerici continui, in questo caso l'età, le entrate annue e la capacità di ricevere credito. Il tentativo di predire un certo tipo di comportamento è sempre basato sul concetto che se un individuo appartiene ad un gruppo esso si comporterà in genere come il gruppo cui appartiene.

Per fare la predizione voluta ci si chiede, in prima istanza, se dei tre attributi a disposizione l'età sia il fattore determinante nel caratterizzare il comportamento degli individui dell'insieme rispetto alla loro disponibilità ad acquisire una macchina. Questo fatto di per sé implica che solamente conoscendo l'età degli individui del campione si è in grado di predire se un individuo sia propenso o meno ad acquisire una macchina. Laddove ciò fosse, è possibile, di conseguenza, spaccare il nostro insieme di individui in due distinti sottoinsiemi facendo uso della sola età. Si deve investigare se esista o meno un'età limite che possa dividere i possibili compratori dai probabili non compratori di macchine. Analizzato l'insieme rispetto al primo degli attributi a disposizione ed eventualmente rilevata la presenza di questa età limite, si passa al secondo degli attributi a disposizione e si ripete rispetto ad esso la stessa analisi sopra descritta. Analizzato l'insieme rispetto al secondo attributo a disposizione ed, anche in questo caso, rilevata l'eventuale presenza di entrate annue limite con cui si può spaccare l'insieme, si passa al terzo attributo e si fanno le stesse domande.

Il suddetto tipo di analisi può essere fatto a cascata, cioè si può analizzare rispetto al secondo attributo, le entrate annue, il sottoinsieme di persone propense ad acquisire una macchina derivante dall'analisi rispetto al primo attributo, l'età, e così via. Alla fine di questo processo iterativo, attributo per attributo si arriva a creare il cosiddetto albero delle decisioni per l'insieme sotto esame in relazione ad una data proprietà, in questo caso la propensione ad acquistare una macchina.



Vi sono molti algoritmi e, quindi, molti programmi che eseguono i *task* sopra descritti con grande efficienza anche perché la complessità degli algoritmi stessi è proporzionale ad $n \cdot \log(n)$ in cui n è il numero di oggetti all'interno dell'insieme sotto esame. Stante questa situazione gli algoritmi di strutturazione di un albero di decisione lavorano molto bene in insiemi molto popolati (al contrario di quanto accadeva per gli algoritmi dei k primi vicini a causa della loro dipendenza da n^2). Un vantaggio ulteriore dei risultati ottenuti con questa metodologia è che essa è di diretta e naturale comprensione.

Le regole associative

In genere ai *manager* piacciono delle regole semplici che presumono siano dettate dal mercato tipo: al 90% delle signore che guidano macchine *spider* rosse e possiedono un cagnolino di piccola taglia piace lo Chanel n° 5. Questo tipo di regole piacciono poiché danno un chiaro profilo del cliente sul quale finalizzare

una precisa strategia di promozione. Alcune delle metodologie di Data Mining, quelle soprannominate regole associative, promettono di estrarre da grandi basi dati regole come quelle sopra descritte.

Si lavori per esempio su un *data base* nel quale per ogni individuo è memorizzato il genere, il tipo di macchina, il colore della macchina, la razza del cane posseduto, ed il possesso o meno di una serie di beni o di preferenze (l'uso o non uso del profumo, tipo di profumo etc.). La regola che piace tanto ai manager letta dal *data base* suddetto si traduce in: il 90% dei *record* in cui il genere è femminile, la macchina è una *spider*, il colore della macchina è rosso, il cane è di taglia piccolo ha alla colonna profumo la voce Chanel N°5.

Per prima cosa bisogna dire che le regole associative si trovano in *data base* in cui gli attributi hanno valore binario (sì, no, compra non compra, alto basso etc) e qualora gli attributi fossero di natura numerica questi ultimi dovrebbero essere "appiattiti" fino ad assumere una forma binaria; se ad esempio uno degli attributi fosse l'età ed essa avesse una distribuzione nell'insieme sotto esame fra i 30 ed i settanta anni dovremmo, prima di applicare un algoritmo associativo, tramutare l'età nel parametro binario "sopra i cinquanta: sì/no". Una volta fatta questa trasformazione vi sono molti algoritmi che possono ricavare associazioni come quelle sopra descritte e dette associazioni possono comprendere un attributo, due attributi, tre attributi etc.

È, però, vero che il numero di regole associative cresce esponenzialmente con il numero di differenti attributi che compongono il *data base* oggetto di investigazione. Peraltro non è difficile trovare regole associative quanto definire se dette regole abbiano o meno un qualche valore. Per esempio nel nostro precedente esempio è vero che al 90% delle signore che guidano macchine *spider* rosse e possiedono un cagnolino di piccola taglia piace lo Chanel n°5, ma può risultare anche vero che il numero delle signore che guidano macchine *spider* rosse e possiedono un

cagnolino di piccola taglia sia solamente 10 sui 2 miliardi di individui registrati nel *data base* sotto esame. Quindi la regola associativa estratta non è poi così significativa.

Nei fatti gli algoritmi di ricerca di regole associative trovano con facilità moltissime associazioni, il problema è quello di trovare alcuni indici che misurino il peso delle associazioni trovate aiutandoci a individuare solo quelle significative.

Gli indici in questione sono in genere due:

- gli indici di supporto;
- gli indici di confidenza.

Facciamo un esempio, in una *data base* abbiamo tre attributi binari per individuo registrato: lettore o non lettore di giornali sportivi (*sport, no sport*); lettore o non lettore di giornali generalisti (*gen no gen*), lettore o non lettore di riviste di moto e motori (*moto no moto*)

Abbiamo ricavato la seguente regola associativa: chi legge sia giornali sportivi sia giornali generalisti legge anche giornali di moto e motori. In termini formali *sport and gen => moto* ovvero *sport gen moto*.

Quanto è solida questa nostra associazione, quanto i dati contenuti nel *data base* la supportano? L'indice di supporto ci dà informazioni in questo senso fornendo il numero di lettori che contemporaneamente leggono *sport gen e moto* diviso per l'intero numero di *record* contenuti nel db.

Purtroppo l'indice di supporto non basta di per sé a darci indicazioni sufficienti. Per esempio lavorando sulla associazione ci possiamo accorgere che, sì, l'indice di supporto della nostra associazione a tre (*sport gen e moto*) è significativo, ma che è anche di molto inferiore all'indice di supporto della sola associazione a due attributi *sport e gen*. In parole povere il fatto che il numero di lettori di tre riviste contemporaneamente è molto piccolo in confronto di quello binario di partenza, i.e. i lettori contemporanei delle sole due riviste di sport e generaliste, ci dice

che la nostra associazione a tre benché rilevata è anche debole. Ecco che entra in gioco un ulteriore indice quello di confidenza che non è altro che il rapporto fra il numero di lettori di *sport gen e moto* diviso per il numero di lettori di *sport e gen*.

Quindi lavorando per ogni regola e ricavando al contempo per ogni associazione i relativi indici di supporto e di confidenza si può arrivare a discriminare le associazioni importanti da quelle assolutamente inutili. Ancora una volta la ricerca di regole associative deve partire da un obiettivo prefissato come ad esempio la conferma di una ipotesi di esistenza di una specifica associazione, altrimenti si è sommersi da una mole di informazioni che confonde più che chiarire. L'applicazione di un algoritmo che trova tante regole non garantisce che tutte quelle trovate siano significative. D'altra parte se partiamo con un algoritmo più rigido troviamo sì meno associazioni, ma possiamo aver mancato quelle veramente importanti. Si devono quindi provare con criterio diversi algoritmi, valutare i risultati parziali con attenzione per poi salire di una maglia sulla catena di scoperta della conoscenza.

Le reti neurali

Quando si arriva a voler creare un *tool* di analisi descrittivo ovvero che implichi processi di tipo *machine learning* si derivano modelli e strutture provenienti da aree della scienza completamente differenti dalla matematica e dall'informatica tradizionale. Tali sono gli algoritmi basati sulle reti neurali che trovano la loro origine nella ricerca e nella scienza del funzionamento del cervello in uomini ed animali.

Nei fatti la struttura del cervello vede la presenza di un gran numero di cellule dette neuroni, nell'uomo circa 10^{11} , ognuno dei quali è connesso a circa 2000 altri neuroni tramite le cosiddette sinapsi. Basandosi su questa struttura base che è al contempo semplice, flessibile ed altamente ridondante, il cervello può gestire compiti estremamente complessi.

Tentando di ricreare nelle macchine gli elementi che funzionano come neuroni e sinapsi, dapprima con *hardware* specifici oggi tramite programmi s/w, si è aperto il campo di ricerca sulle reti neurali che sono impiegate per eseguire in modo automatico compiti complessi simili a quelli eseguiti dal cervello.

Tipicamente una rete neurale consiste in una serie di nodi interconnessi fra loro:

- i nodi di *input* ricevono segnali di ingresso,
- i nodi di *output* passano segnali in uscita,
- i nodi intermedi o nascosti, raggruppati in un numero variabile di strati intermedi, svolgono funzioni elaborative.

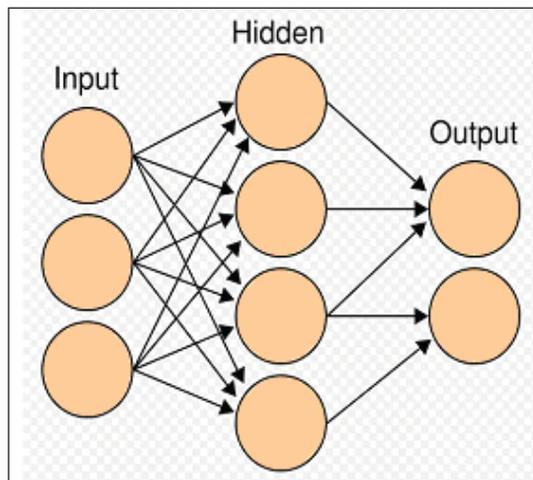


Fig. 3 - Rete neurale (da Wikipedia)

L'uso delle reti neurali prevede due fasi distinte:

- La fase di cosiddetto *encoding* nella quale si “insegna” alla rete ad eseguire un determinato compito;
- La fase di cosiddetto *decoding* nella quale si utilizza la rete per eseguire in via ripetitiva il compito che le è stato insegnato (classificare esempi, fare predizioni, etc).

Vi sono tante differenti architetture di reti neurali ognuna delle quali è caratterizzata da specifiche strutture interconnettive e da differenti strategie di addestramento ed acquisizione dell'esperienza. In questo contesto parleremo di tre strutture neurali che hanno un ruolo oggi nella tecnologia informatica:

- I Perceptrons;
- I Back Propagation Networks (BPN, letteralmente le reti a retro propagazione);
- Le mappe auto organizzanti di Kohonen.

Una macchina a Perceptrons consiste di una semplice struttura a tre strati con i nodi di *input* chiamati foto ricettori, i nodi del livello intermedio chiamati associatori ed i nodi di *output* chiamati risponditori. Una macchina a Perceptrons può imparare e, quindi, eseguire compiti semplici quale può, ad esempio, essere una semplice attività di catalogazione.

I BPN possiedono architetture più complicate ed essi non possiedono solamente nodi di *input* e di *output* ma ha anche un numero variabile di strati nascosti di nodi intermedi.

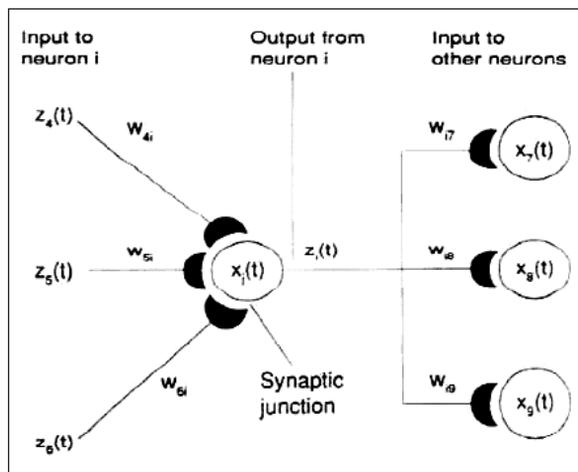


Fig. 4 - Nodo di un BPN, figura da Christos Stergiou e Dimitrios Siganos in www.doc.ic.ac.uk

Allo stato iniziale il BPN hanno sinapsi con pesi (w_{ij}) scelti a caso. Durante la fase di *encoding* il network è esposto ad un insieme di dati di *training* in *input* (insieme di dati per cui si conosce la risposta ovvero l'*output*) in istanze ripetute di cosiddetto *training*. Per ogni istanza di *training*, l'*output* dal BPN è paragonato all'*output* corretto che ci si aspetta, se vi è una differenza fra *output* effettivo ed *output* voluto i pesi delle sinapsi di tutti i nodi intermedi coinvolti sono di volta in volta mutati sino a che la differenza fra *output* effettivo ed *output* desiderato non scende al di sotto di un valore di soglia predeterminato. Una volta che la struttura del BPN si è stabilizzata (i.e. i pesi delle sinapsi non variano più e si hanno i dati di output voluti per i dati di training) si possono far elaborare al BPN nuovi insiemi di dati per ottenere da essi un *output* che li categorizzi, operando su di essi in maniera analoga a quando si era operato su i dati di *training*.

L'architettura BPN ha rappresentato un grande passo in avanti rispetto alla semplice struttura a Perceptrons. Comunque si rilevano ancora diversi svantaggi, uno di essi è che la fase di *training* è lunga e necessita di insiemi di *training* molto popolati. Peraltro, i BPN danno i risultati senza dare alcuna motivazione per razionalizzare il risultato finale fornito (si ricordi ad esempio quanto accadeva con il metodo degli alberi di decisione).

Le Reti Auto Organizzanti di Kohonen (RAOK) possono essere pensate come la controparte artificiale della mappa di neuroni attivati che esistono in zone diverse del cervello. Una RAOK è un insieme di neuroni, le unità, ognuno dei quali è connesso ad un piccolo numero di altre unità chiamati vicini. Per la maggior parte del tempo la RAOK si mantiene in due dimensioni; nella RAOK ogni nodo è caratterizzato da un fattore che è correlato allo spazio oggetto dell'investigazione. Nella fase iniziale la RAOK è caratterizzata da una scelta di vettori, i fattori di cui sopra, casuale. Via via che il *training* avanza i vettori sono aggiustati al fine di garantire la migliore mappatura dello spazio da esplorare.

Le reti neurali funzionano bene in compiti di classificazione ed in questo ambito sono molto usate nel Data Mining, anche se sono delle vere scatole nere da cui scaturiscono i risultati senza che l'utente possa ricostruire la via logica che ha portato il programma ad un dato risultato.

Gli Algoritmi Genetici

Gli algoritmi genetici hanno tre illustri padri:

- Darwin che ha individuato nel processo di selezione/evoluzione un meccanismo di ottimizzazione, scelta ed adattamento di una serie di attributi da lui riferito agli esseri viventi ma in via di principio applicabile a qualsivoglia sistema di cui si voglia trovare un *optimum*;
- Francis Crick che ha scoperto il meccanismo molecolare con il quale si attua negli esseri viventi il processo di selezione;
- Nils Aall Barricelli, virologo genetista italo norvegese, che ha passato praticamente molte delle sue notti nelle sale calcolo dei *super computer* di tutto il mondo, ivi compreso lo storico calcolatore dell'Istitute for Advanced Study di Princeton (USA NJ) diretto da Johnny von Neumann, e che ha inventato e provato algoritmi che tentano di riprodurre col calcolatore il processo di selezione naturale.

La formula per la costruzione di un algoritmo genetico per la soluzione di un qualsivoglia problema di ottimizzazione è la seguente:

- Strutturare una buona ed elegante codifica del problema scegliendo un alfabeto (i.e. le triplette di basi del DNA ognuna delle quali corrisponde ad un amminoacido naturale) limitato ma efficace;
- Inventare un ambiente artificiale nel calcolatore dove le possibili diverse soluzioni al problema posto possano competere fra di loro. Per far ciò bisogna anche strutturare un

buon algoritmo di *rating* che possa giudicare il successo od il fallimento di una possibile soluzione, è cioè quella che viene in gergo chiamata una funzione di *fitness* della possibile soluzione;

- Sviluppare modi in cui diverse soluzioni possano accoppiarsi e quindi figliare una nuova soluzione. In questo l'algoritmo di riproduzione più usato è quello detto del *cross over* con il quale le stringhe "padre" e "madre" sono tagliate, cambiate e riattaccate l'una all'altra. Però in questo contesto sono anche stati applicati tutta una serie di operatori di mutazione;
- È importante che la simulazione evolutiva parta con una popolazione di possibili soluzioni ben bilanciata e varia. Ad ogni passo evolutivo si devono scartare le soluzioni cattive eventualmente rimpiazzandole con la loro "progenie" ovvero con mutazioni che possiedono una buona *fitness*.
- Il processo termina quando tutte le soluzioni rimaste in gioco sono buone soluzioni.

Per quanto, in via di principio, estranei al mondo del calcolo, i passi suddetti sono facilmente programmabili e ciò è dimostrato dal successo che questa algoritmica ha avuto a partire dagli anni '60-'70. Il buon programmatore si vede nel momento della codifica (un problema di ingegneria della rappresentazione) e nel trovare operatori di mutazione efficaci.

I meriti ed i demeriti della famiglia degli algoritmi genetici riproducono quelli del processo di selezione che si opera in natura. I due principali svantaggi del metodo sono la larga produzione di soluzioni/individui e la casualità del processo di mutazione. In generale servono grandi potenze di calcolo, e di questo Barricelli se ne era ben accorto, al fine di raggiungere una soglia di accettabilità delle soluzioni generate. D'altra parte il processo è solido; se una soluzione esiste e si ha a disposizione un buon

calcolatore è ragionevole pensare che l'algoritmo genetico la scopra. Negli ultimi tempi si è visto il nascere di una serie di metodi ibridi, nel cui ambito le reti neurali fornivano un risultato che serviva da *input* per gli algoritmi genetici. Ad oggi le applicazioni più importanti degli algoritmi genetici si trovano nello studio degli andamenti dei mercati finanziari e in applicazioni specifiche per il mondo assicurativo.

Text Mining

Il Text Mining è una variazione del Data Mining e consiste nell'applicazione delle sue tecniche a testi non strutturati. La differenza tra il Data Mining ed il Text Mining è che nel Text Mining i *pattern* (sequenze di parole) sono estratti da testi (documenti) in linguaggio naturale piuttosto che da *database* strutturati di fatti. I testi esprimono, in realtà, una vasta e ricca gamma di informazioni, ma codificano queste informazioni in una forma che è difficile decifrare automaticamente.

Oggi, con la diffusione delle tecnologie proprie della società dell'informazione e con la crescita smisurata del numero di documenti che possono contenere informazioni interessanti e quindi da analizzare (agenzie stampa, pagine web, *e-mail*, libri e articoli in versione digitale, ...) risulta quanto meno importante disporre di strumenti automatici per la loro catalogazione ed analisi. Dal momento che i dati testuali non sono strutturati, l'analisi automatica appare abbastanza difficile. Tuttavia, l'estrazione di informazioni utili è possibile laddove, ad esempio, si cerca unicamente di classificare dati testuali senza cercare di capire completamente il contenuto.

Un problema comune che può essere risolto con il Text Mining consiste nella determinazione dell'appartenenza o meno di un determinato documento ad uno specifico argomento (*classificazione, raggruppamento tematico*). L'individuazione di gruppi

tematici consente di dare un'organizzazione all'informazione disponibile e di individuare argomenti minori, che anche ad una lettura attenta potrebbero sfuggire. Le relazioni, inoltre, possono mettere in evidenza *associazioni* nascoste considerando legami tra argomenti apparentemente separati ma che hanno una terminologia comune (legami tra argomenti, o tra autori, *trend* temporali,...). Altri problemi che possono essere affrontati con il Text Mining riguardano la possibilità di *identificare entità* (es. nomi di geni, nomi di aziende, ...) contenute nei documenti (*information extration*) ed *estrarre concetti* per la creazione di ontologie (i.e la struttura fondamentale e necessaria sottesa ad un insieme di dati che però prescinde dalla specifica ed individuale manifestazione dell'insieme di dati stesso, *ontology learning*).

Il Text Mining può essere usato per analizzare documenti in linguaggio naturale riguardanti qualsiasi soggetto, benché l'area applicativa più attiva e più promettente sia nel campo bio-medico e specificatamente nella ricerca di interazioni (attivazione, inibizione, ...) tra le proteine attraverso l'analisi della letteratura biomedica. Il Text Mining non solo può estrarre informazioni dai documenti sulle interazioni fra le proteine, ma può anche andare un passo oltre per scoprire *pattern* nelle interazioni estratte. Queste informazioni potrebbero aiutare a rispondere a importanti domande o suggerire nuovi percorsi da esplorare.

Il Text Mining coinvolge l'applicazione di tecniche da aree come l'*information retrieval*, elaborazione del linguaggio naturale, estrazione delle informazioni e Data Mining. Tecniche che possono costituire fasi di un unico processo di Text Mining e combinate insieme in un singolo *workflow*. Si può affermare che, analogamente a quanto visto per il KDD, il Text Mining percorre le fasi classiche di un processo di estrazione di conoscenza.

Vediamo ora come si struttura un processo di Text Mining descrivendo con maggior dettaglio ognuna delle fasi in cui si suddivide:

- *Information retrieval (IR)*: identifica all'interno di una col-

lezione una serie di documenti che soddisfano *query* definite dall'utente. I sistemi IR più conosciuti sono i *search engine* come Google™, che identifica quei documenti nel mondo World Wide Web che contengono un certo insieme di parole chiave; i sistemi IR sono spesso usati in ambito bibliotecario, dove i documenti non sono i libri stessi ma *record* digitali contenenti informazioni relative ai libri. Ciò tuttavia sta cambiando con la diffusione delle biblioteche digitali (*digital libraries*) dove i documenti da recuperare sono la versione digitale dei libri e delle riviste (*journal*). I sistemi IR consentono di accedere più velocemente all'argomento di interesse e individuare i legami con altri argomenti riducendo la quantità di documenti da analizzare. Per esempio, se siamo interessati solo a informazioni riguardanti le interazioni tra proteine, possiamo restringere l'analisi a quei documenti che contengono il nome di una proteina, o alcune forme verbali del verbo "interagire" o uno dei suoi sinonimi.

- *Elaborazione del linguaggio naturale (NLP in inglese)*: è uno dei problemi più indagati e difficili nel campo dell'intelligenza artificiale. Si occupa dell'analisi del linguaggio umano al fine di consentire la comprensione automatica del linguaggio naturale da parte del *computer* così come farebbe un essere umano. Benché l'obiettivo finale sia ancora lontano da raggiungere, esistono sistemi che possono effettuare alcuni tipi di analisi con un certo grado di successo. Il ruolo di questi sistemi nel Text Mining è quello di curare la fase di estrazione delle informazioni.
- *Estrazione delle informazioni*: è il processo che consente di ottenere dati strutturati da un documento non strutturato in linguaggio naturale. L'obiettivo è l'estrazione di termini specifici (e predefiniti) dal testo. Ciò può essere effettuato definendo uno o più *template* con i quali guidare il processo di estrazione. Al termine di questa fase, per ogni docu-

mento si ottengono informazioni (meta-informazioni aggiuntive), che possono essere memorizzate in un dizionario o ontologia, consistenti nella lista delle proteine presenti nel testo e nel tipo di relazione esistente.

- *Data Mining*: la fase di Data Mining vera e propria si concretizza nell'applicazione di algoritmi a fatti generati dalla fase di estrazione delle informazioni. Generalmente si tratta di algoritmi di *clustering* (per il raggruppamento tematico) oppure di algoritmi di *machine learning* (per la classificazione automatica). Facendo riferimento all'esempio relativo alle interazioni tra proteine, potremmo considerare l'ampio numero di interazioni di proteine estratte dalla collezione di documenti e memorizzate come fatti in un *database*. Applicando il Data Mining a questo *database* potremmo identificare alcuni *pattern* nei fatti. Ciò può portare alla scoperta di nuove interazioni che possono o no verificarsi, oppure la relazione tra tipi di interazioni e particolari malattie.

Per meglio descrivere il processo e le tecniche che si possono applicare si ritiene opportuno proporre l'esempio di costruzione di un prototipo³ per un sistema di Text Mining che consente di analizzare la letteratura biomedica, nel campo della genetica, allo scopo di individuare le eventuali interazioni tra geni.

L'uso del Text Mining in questa specifica area applicativa risulta particolarmente importante poiché la conoscenza che si estrae analizzando le pubblicazioni specialistiche può essere considerata una fondamentale sorgente di informazioni che il ricercatore usa per interpretare e comprendere meglio i risultati sperimentali ottenuti usando la tecnologia DNA *microarray*⁴.

³ Progetto MedMole, disponibile su <<http://medmole.cineca.it/>>.

⁴ Per informazioni vedi <<http://it.wikipedia.org/wiki/Microarray>>.

Questa tecnologia infatti si propone di esaminare il profilo d'espressione di un gene o per identificare la presenza di un gene o di una breve sequenza in miscela di migliaia

L'approccio che viene utilizzato si pone l'obiettivo di produrre una rappresentazione strutturata dell'informazione testuale predefinendo entità e relazioni da ricercare nei testi (ad esempio le entità possono essere proteine e farmaci e le relazioni attivazione, inibizione, ...). Si adottano tecniche che utilizzano sia l'analisi sintattica che l'analisi semantica. Alla fine del processo viene generato un *database* che può essere analizzato utilizzando le tecniche di *data mining*.

Come primo passo, attraverso un sistema IR, può essere effettuata una selezione tra i documenti disponibili (Pubmed) per ottenere l'insieme completo delle pubblicazioni che trattano di ciclo cellulare utilizzando uno specifico insieme di parole chiave (cell cycle OR cell proliferation OR cell death OR oncogenes OR tumor suppressor , ...).

Una volta delimitato l'insieme di documenti voluto si passa alla loro cosiddetta preparazione; che può essere effettuata attraverso tre fasi:

- innanzitutto l'individuazione delle diverse parti del documento consente di separare l'informazione testuale dalla metainformazione (organismo di appartenenza, data e rivista di pubblicazione, tipo di pubblicazione, paese, ecc...);
- sulla parte testuale dei documenti così individuata può essere effettuata un'*analisi grammaticale*;
- successivamente possono essere applicate tecniche di *estrazione dell'informazione*.

In particolare con l'*analisi grammaticale* di un testo si selezionano tutti i sostantivi che all'interno di un discorso sono generalmente i termini con maggiore contenuto semantico. Questi vengono a costituire l'insieme di parole chiave che caratterizza e descrive ciascun documento. È sulla base di queste parole chia-

ve che verrà giudicato il grado di somiglianza tra i documenti. Tutte le altre parti del discorso (aggettivi, verbi e nomi propri) vengono mantenute come informazioni aggiuntive, così come la meta-informazione che era stata estratta durante la prima fase. Poiché nell'ambito della ricerca biologica è molto importante poter identificare i nomi dei geni, e poiché l'analisi grammaticale li classifica, genericamente, come nomi propri, è necessaria analizzare ulteriormente il testo per estrarre questo tipo d'informazione. Con la fase di *estrazione dell'informazione*, ci si pone l'obiettivo di estrarre termini specifici (e predefiniti) dal testo, utilizzando un dizionario che nel nostro caso è quello contenente i nomi ufficiali di geni ed i relativi *alias*.

Il dizionario non è altro che una tabella (*file*) di due colonne dove nella prima colonna, "GENE", è contenuto il nome ufficiale, nella seconda, "ALIAS", è contenuto un *alias* dello stesso gene: il nome del gene è ripetuto tante volte per quanti sono gli *alias* esistenti. Lo strumento che si utilizza cerca nel documento la presenza di qualsiasi occorrenza presente nella colonna ALIAS e memorizza il corrispondente nome presente nella colonna GENE associandogli l'identificatore del documento analizzato. Si ottiene così una lista di nomi di gene che sono referenziati in ogni documento. Questa lista è integrata con la lista di parole chiave precedentemente individuate.

Alla fine di questa fase ogni documento è descritto in formato sintetico da una lista di parole chiave e da una lista di nomi di geni. Questo formato consente di rappresentare l'informazione in una matrice binaria che contiene, sulle righe, ciascun documento, sulle colonne, ciascun sostantivo, e all'interno i valori 1 o 0 ad indicare la presenza di una determinata parola chiave in un determinato documento oppure la sua assenza (la metainformazione è rappresentata in maniera analoga).

Siamo così giunti alla fase di *data mining* vera e propria che si concretizza nell'applicazione di un algoritmo di *clustering*. Poiché l'informazione disponibile è tutta di tipo qualitativo, si

può scegliere un algoritmo partitivo basato sull'analisi relazionale. Questo metodo consente di confrontare tutte le coppie di documenti e di calcolare, per ogni coppia, un indice di somiglianza basato sul numero di co-occorrenze delle parole chiave. I *cluster* vengono formati in modo da massimizzare la somiglianza complessiva dei documenti raggruppati e minimizzare la somiglianza dei documenti che vengono separati. Su questo processo si può intervenire agendo attraverso alcuni parametri (la soglia di somiglianza ed il sistema di ponderazione).

Supponiamo di essere interessati alle eventuali interazioni tra due geni, BRCA1 e BRCA2, e selezioniamo tutti i documenti che contengono almeno uno di questi geni. Può essere lanciata un'analisi di *data mining*, che, in tempo reale, confronta le parole chiave dei documenti selezionati (in quanto questi sono gli attributi descrittivi che abbiamo deciso di usare per il *clustering*) e li raggruppa secondo i *pattern* individuati. La visualizzazione dei risultati rende evidente che i documenti selezionati tendono a trattare i due geni separatamente, come era prevedibile, dato che un gene è implicato nel tumore al polmone e l'altro nel tumore al seno e che le interrelazioni dovrebbero essere minime. Viceversa in alcuni casi viene evidenziato che, quando compare BRCA1, spesso compare anche BRCA2, quindi tra questi due geni può esserci una interrelazione.

Gli stessi passi possono essere generalizzati e, quindi, essere fatti su un qualsivoglia altro tipo di documenti posto che si sia in possesso di un dizionario significativo per il settore nel cui ambito si svolge la ricerca.

I sistemi informativi

STEFANO PIGLIAPOCO

1. Introduzione

La convergenza delle politiche socio-economiche degli Stati verso la globalizzazione dei mercati e l'enorme sviluppo delle tecnologie dell'informazione e della comunicazione (ICT) stanno cambiando radicalmente i sistemi industriali. «Le macchine a vapore sono state le protagoniste, tra il 1780 e il 1840, della prima rivoluzione industriale. L'acciaio e le ferrovie hanno caratterizzato, fra il 1840 e il 1890, la seconda. L'energia elettrica e la chimica sono alla base della terza rivoluzione industriale degli anni 1890-1930. Il petrolio e l'automobile sono i protagonisti, fra il 1930 e il 1980, della quarta onda lunga. Il microprocessore e la Rete sono i protagonisti della quinta rivoluzione industriale che stiamo vivendo»¹.

Nell'economia globale, il modello d'impresa passa da una concezione monolitica, dove la catena del valore è interamente controllata da un'organizzazione gerarchica e centralizzata, a una concezione a rete, dove "piccole" aziende specializzate cooperano sulla base di accordi specifici per affrontare il mercato

¹ MAURO MARCONI, *Economia e finanza delle imprese marchigiane, in Sviluppo e internazionalizzazione dell'industria marchigiana*, Ancona, Il lavoro editoriale, 2001, p. 43.

con maggiore flessibilità e competitività. Illuminante è la descrizione del fenomeno dell'*outsourcing* proposta da Rifkin: «in un numero crescente di aziende, la gestione e la manutenzione dei computer, la formazione, lo sviluppo applicativo, la consulenza e la progettazione vengono affidati a fornitori esterni (così come) la gestione della posta, la stampa e la riproduzione dei documenti, la gestione degli archivi e delle scorte, gli acquisti e la contabilità generale [...], le vendite e il marketing»².

Il trasferimento della produzione dei beni materiali nei Paesi in via sviluppo, attuato per conseguire forti riduzioni di costi, e l'affermarsi di un modello d'impresa meno gerarchico, più flessibile e decentrato, attribuiscono ai sistemi informativi un ruolo strategico. L'esigenza di avere, in qualsiasi momento della giornata e in ogni luogo, in tempo reale, le informazioni necessarie per svolgere i processi di natura operativa o per prendere decisioni, è sentita a tutti i livelli aziendali. Oggi, se un'impresa vuole essere protagonista sul mercato non può fare a meno di avere un sistema informativo che, oltre a supportare efficacemente i tradizionali processi produttivi e di amministrazione, renda possibile la piena condivisione tra i soggetti interessati dei dati e dei documenti rilevanti per l'azienda.

L'esigenza di un sistema informativo efficiente e tecnologicamente avanzato è ancora più sentita nelle cosiddette *imprese digitali*, i cui prodotti hanno un contenuto prevalentemente, se non totalmente, informativo e tecnologico, e dove tutte le relazioni con i clienti, i fornitori e gli stessi dipendenti si sviluppano su base informatica. Anche nella pubblica amministrazione, dove l'informazione rappresenta in molti casi il prodotto finale e i processi si sviluppano attraverso l'acquisizione e la produzione di documenti, il sistema informativo rappresenta una risorsa fonda-

² JEREMY RIFKIN, *L'era dell'accesso. La rivoluzione della new economy*, Milano, Mondadori, 2000.

mentale. Non a caso il legislatore italiano ha emanato, nell'ultimo decennio, un complesso di norme volto a promuovere l'innovazione tecnologica nelle pubbliche amministrazioni e a creare le condizioni per la erogazione di servizi *on line* e la dematerializzazione dell'azione amministrativa³.

2. Il sistema informativo

Prima di definire il sistema informativo è opportuno richiamare i concetti di risorsa, processo e struttura organizzativa.

Una risorsa è una qualsiasi entità, fisica o immateriale, che un'organizzazione utilizza per raggiungere i suoi scopi. Sono risorse: il personale, gli immobili, le attrezzature, il capitale sociale, il *know how*, i clienti, il mercato, etc. Anche le informazioni sono risorse, e sicuramente tra le più preziose.

Un processo può essere visto come «una sequenza di attività tra loro interdipendenti e finalizzate al conseguimento di un obiettivo comune; esso riceve un certo input (materiali, istruzioni e specifiche del cliente), vi apporta le trasformazioni che aggiungono valore utilizzando le risorse aziendali, ossia le persone, i materiali e le strutture, ed infine trasferisce all'esterno l'output richiesto»⁴. Un'altra definizione interessante identifica il processo come «un insieme di attività tra loro interrelate, finalizzate alla realizzazione di un risultato definito e misurabile (il prodotto/servizio interno o esterno) che contribuisce al raggiungimento

³ Per dematerializzazione dell'azione amministrativa si intende la riduzione del volume della documentazione cartacea prodotta da un ente pubblico nell'esercizio delle sue funzioni, conseguita attraverso processi di semplificazione amministrativa e l'impiego delle tecnologie dell'informazione e della comunicazione.

⁴ Cfr. PIERO DE RISI, *La gestione per processi e i suoi riflessi organizzativi in azienda*, Roma, Nuovo Studio Tecna, 2002.

della missione dell'organizzazione e che trasferisce valore al fruitore del servizio (il cliente, anche in questo caso, interno o esterno)»⁵. Ogni processo è suddivisibile in fasi, che a loro volta possono essere scomposte fino ad individuare le attività nel cui ambito si svolgono i compiti. Le attività di un processo sono legate tra loro dalle informazioni e dai prodotti/servizi che si scambiano, e possono essere svolte da più persone anche appartenenti ad unità organizzative diverse.

La struttura organizzativa è una «suddivisione dell'organizzazione in differenti unità organizzative alle quali sono attribuiti compiti e obiettivi specifici e che cooperano attraverso legami di tipo gerarchico e funzionale»⁶. A questa visione tradizionale si tende oggi a sostituire una visione trasversale, concependo l'organizzazione come un insieme di processi più che come unione di funzioni da coordinare. Le recenti norme sulla qualità considerano l'organizzazione per processi un elemento determinante per la soddisfazione del cliente. In particolare, la norma ISO 9001:2000, che definisce i requisiti dei sistemi di qualità, richiede esplicitamente un approccio basato sui processi, sostenendo l'importanza centrale della loro identificazione, monitoraggio e miglioramento continuo. D'altra parte, nella maggioranza dei casi, un processo non inizia e finisce all'interno di una stessa unità organizzativa, ma si sviluppa trasversalmente, coinvolgendo più funzioni come mostrato in figura 1. Pertanto, in un'organizzazione strutturata in modo tradizionale ci si trova nella situazione in cui nessuno controlla l'intero iter del processo e ciascun manager si preoccupa di gestire solo la fase di sua competenza, perdendo di vista ciò che accade prima e dopo.

⁵ Cfr. GAETANO SANTUCCI, *Introduzione ai sistemi informativi*, in *Sistemi informativi*, Milano, Franco Angeli, 2001, p. 17.

⁶ Cfr. *Ibidem*.

software utilizzati per la gestione della risorsa informazione, ma è un'entità composita che incorpora il sistema informatico. Inoltre, da non confondere il sistema informativo con il complesso delle informazioni necessarie all'organizzazione per raggiungere i suoi scopi, in quanto il secondo è il prodotto del primo.

3. Il sistema informatico

Se è vero che il sistema informatico non coincide con il sistema informativo di un'organizzazione, è altrettanto vero che esso ne costituisce la parte più rilevante. Si può affermare che l'evoluzione dei sistemi informativi è resa possibile dallo sviluppo delle tecnologie e allo stesso tempo sostenere il contrario, cioè che l'esigenza di soddisfare i bisogni informativi crescenti delle organizzazioni determina la realizzazione di nuove soluzioni informatiche.

Un sistema informatico, inteso come il componente base di un sistema informativo, è costituito da quattro elementi:

- un'infrastruttura di rete;
- un insieme di *computer* interconnessi secondo una determinata architettura;
- un complesso tecnologico per la gestione dei dati;
- un insieme di moduli *software*.

3.1 L'infrastruttura di rete

Una delle più grandi innovazioni dell'informatica è stata la possibilità di collegare tra di loro più *computer* tramite supporti fisici per la trasmissione di segnali, e quindi di creare un'infrastruttura di rete capace di supportare forme di comunicazione di dati, scambio di informazioni e condivisione di risorse.

Tradizionalmente, le reti di calcolatori vengono classificate in base alla loro estensione e complessità. Si definiscono *LAN (Local Area Network)*, se sono gestite all'interno delle organizzazio-

ni e coprono un'area limitata di qualche centinaio di metri; *MAN* (*Metropolitan Area Network*), se sono gestite da fornitori di servizi di comunicazione e coprono un'area territoriale di qualche decina di Km; *WAN* (*Wide Area Network*), se si estendono su aree geografiche molto vaste e presentano una struttura articolata e complessa.

Internet è una rete planetaria ottenuta dall'interconnessione di reti eterogenee sulla base di regole comuni, i cosiddetti *protocolli Internet*, che rendono possibile la comunicazione di dati per via informatica. La sua nascita risale al 1969 quando il Ministero della Difesa Statunitense creò un'agenzia per i progetti di ricerca avanzata, chiamata ARPA (*Advanced Research Projects Agency*), con lo scopo di sviluppare una rete di *computer* che potesse reggere al bombardamento nucleare, garantendo la continuità di comunicazione tra località diverse. Il progetto coinvolse centri di ricerca, Università e qualche azienda privata: tutti in qualche modo legati all'attività militare e dotati di moderna tecnologia informatica. Si giunse così alla nascita di ARPANET, il precursore di Internet. A distanza di un trentennio, Internet conta diverse centinaia di milioni di utenti quotidianamente collegati e veicola giornalmente un volume di dati digitali dell'ordine del Terabyte. Il suo più grande merito è quello di aver imposto, con la legge dei grandi numeri, le regole tecniche per l'interoperabilità tra i sistemi informatici di produttori diversi, rendendo disponibile il canale di comunicazione più esteso e a basso costo che sia mai esistito.

L'enorme successo di Internet ha portato alla realizzazione delle cosiddette *Intranet*, cioè di reti private interamente basate sui protocolli e sulle applicazioni Internet, che svolgono una funzione comunicativa e di scambio di informazioni tra i comparti di un'organizzazione. La realizzazione di reti aziendali secondo il modello Intranet migliora la capacità delle organizzazioni di comunicare con il mondo esterno e di collegare più sottoreti, dislocate in aree geografiche diverse, a costi estremamente contenuti.

Il mondo delle reti è in continua trasformazione. L'avvento della cosiddetta *banda larga*, cioè di un ambiente tecnologico che consente un'elevata capacità e velocità di trasmissione, con un alto livello di interattività, rende possibile la comunicazione di contenuti multimediali composti da testo, grafica, audio e video variamente combinati. Già da tempo, esistono strumenti per digitalizzare e trasmettere immagini in movimento attraverso Internet, rendendo disponibili su ogni *computer* eventi dal vivo (*web casting*) o registrati (*on demand*), videoconferenze, presentazioni multimediali interattive di prodotti, etc.

Oggi si assiste anche alla convergenza di tutti i canali di comunicazione sulla tecnologia digitale. Risalgono alla metà degli anni '90 i primi *VoIP gateway*, che sono dispositivi di rete dotati di interfacce verso il mondo tradizionale della telefonia e in grado di trasportare attraverso Internet anche il traffico vocale. Attualmente, questa tecnologia sta ulteriormente evolvendo verso la *telefonia IP*, che realizza la totale integrazione tra le reti di *computer* e gli apparecchi telefonici. Il centralino si trasforma in uno dei tanti *server* di rete, o in un *software* da installare su un *server* esistente, mentre i telefoni diventano apparati digitali sempre più sofisticati e intelligenti, o addirittura *software* installati su PC.

Un altro fenomeno di grande portata è rappresentato dalla diffusione della telefonia mobile, realizzata attraverso cellulari, *smartphone*, palmari e altri dispositivi in grado di attivare una comunicazione digitale in ogni luogo e in qualsiasi momento della giornata. Con la tecnologia di comunicazione GPRS (*General Packet Radio Service*), oggi implementata su quasi tutti i dispositivi di telefonia mobile, e ancora di più con l'UMTS (*Universal Mobile Telecommunication System*), che costituisce lo standard delle comunicazioni mobili di terza generazione, è possibile avere un accesso veloce a Internet, ricevere e trasmettere contenuti informativi multimediali di ogni genere e fruire di servizi appositamente predisposti da operatori economici.

3.2 L'architettura di sistema

La diffusione dei *personal computer* (PC) e l'aumento della potenza di calcolo in sistemi a basso costo⁷ hanno spinto le organizzazioni verso l'informatica distribuita. I *mainframe* hanno lasciato il posto a sistemi *server* di più piccole dimensioni, i *personal computer* hanno sostituito i terminali, le applicazioni monolitiche residenti su *host* sono evolute verso modelli di tipo *client-server* per sfruttare la capacità di elaborazione propria dei PC, la rapida diffusione degli strumenti di *office automation* ha moltiplicato la capacità dei singoli di produrre documenti, memorizzare dati e scambiare informazioni. In sintesi, dalle architetture di sistema centralizzate, nate negli anni '60 e '70, si è passati nell'arco di pochi decenni alle architetture di sistema distribuite.

Un'architettura di sistema di tipo centralizzato vede un insieme di terminali, più o meno "intelligenti", collegati ad un *host computer* (*mainframe*) sul quale risiedono le applicazioni, i dati e ogni altro elemento necessario per la piena operatività dell'impianto. In un'architettura di sistema di tipo distribuito, invece, le applicazioni, i dati e la capacità di elaborazione risiedono su più sistemi che comunicano tra di loro e cooperano attraverso un'infrastruttura di rete. I vantaggi del secondo modello rispetto al primo sono evidenti: minor costo di impianto e di manutenzione, perché non c'è la necessità di avere un grande sistema centrale "factotum"; maggiore scalabilità dell'impianto, che può essere ampliato senza dover potenziare i sistemi esistenti, ma semplicemente aggiungendone altri in rete; aumento dell'adattabilità del sistema alle esigenze degli utenti, che possono sempre contare sull'attivazione in rete di apparecchiature informatiche specializzate per il loro settore di attività; possibilità di combinare insie-

⁷ Recenti rilevazioni sul livello di diffusione e sviluppo delle tecnologie dell'informazione e della comunicazione (ICT) ci mostrano che i processori raddoppiano la loro potenza ogni 18 mesi, mentre la capacità di memorizzazione aumenta di circa il 20% ogni anno a parità di costo.

me sistemi di produttori diversi, scegliendo le soluzioni che di volta in volta presentano il miglior rapporto costi/benefici.

Un particolare tipo di architettura distribuita, oggi ampiamente utilizzata, è l'architettura *client-server*. Questo modello vede la presenza di due entità: i *client*, rappresentati dai sistemi utilizzati dagli utenti per formulare le loro richieste di servizio, e i *server*, che sono i sistemi erogatori dei servizi richiesti dai *client*. Un sistema *server* deve essere in grado di soddisfare le richieste di servizio provenienti da più *client* contemporaneamente e quindi deve essere progettato con estrema cura per non causare rallentamenti inaccettabili. Le architetture *client-server* possono essere implementate in diverse configurazioni che vanno dal *fat client*, in cui la logica applicativa è implementata in gran parte sul sistema client e il *server* svolge solo le funzioni di gestione dei dati, al *thin client*, dove a livello *client* si eseguono solo le operazioni di presentazione dei risultati delle elaborazioni sviluppate sui sistemi *server*.

In un'architettura distribuita la complessità si sposta dall'elaboratore centrale al sistema visto come un insieme di componenti eterogenei chiamati a cooperare in funzione delle esigenze applicative. Questa complessità di progettazione e conduzione dei sistemi distribuiti aumenta i costi di realizzazione e riduce i vantaggi del *downsizing*. Come soluzione a queste problematiche è stato sviluppato un modello di elaborazione denominato *Distributed Object Computing*, nel quale i sistemi sono progettati come oggetti autonomi e cooperanti, distribuiti fisicamente su più *computer* connessi in rete. Ciascun oggetto rappresenta una risorsa che svolge un ruolo nei processi di un'organizzazione ed è in grado di comunicare con gli altri oggetti del sistema attraverso lo scambio di messaggi⁸.

⁸ Per approfondimenti si veda ANTONIO MASSARI, *Sistemi distribuiti*, in *Sistemi informativi*, Milano, Franco Angeli, 2001.

La transizione verso l'informatica distribuita, e in particolare verso le architetture di sistema di tipo *client-server*, ha influenzato le tecniche di progettazione e sviluppo del *software* applicativo. In una moderna applicazione informatica si individuano tre strati: l'interfaccia grafica verso l'utente (GUI), che risiede sui sistemi *client* e gestisce il colloquio uomo-macchina nel modo più semplice e intuitivo possibile; le regole gestionali, che sono implementate sui cosiddetti *application server* e costituiscono la logica applicativa; la gestione dei dati, realizzata con sistemi DBMS⁹ attivati su appositi *database server*.

3.3 Le basi di dati

Come sottolineato nel paragrafo 2, la finalità di un sistema informativo è quella di rendere disponibili le informazioni necessarie per lo svolgimento dei processi di un'organizzazione; pertanto, quello che più ogni altro determina la qualità e la fisionomia di un sistema informativo è il complesso tecnologico utilizzato per la gestione dei dati¹⁰.

Fino alla fine degli anni '70, le attività di registrazione, modifica e accesso ai dati erano gestite con applicazioni informatiche che producevano *file* organizzati in modo sequenziale, o ad indici per garantire un accesso più veloce. I *file* si componevano di *record*, che a loro volta erano suddivisi in *campi*, ciascuno caratterizzato da un certo *valore*. Il *record* rappresentava l'insieme delle informazioni relative ad una singola entità¹¹, mentre il *file* era l'insieme delle informazioni su tutte le entità dello stesso ti-

⁹ Si veda il successivo paragrafo.

¹⁰ Il dato è una rappresentazione originaria e non interpretata di un fenomeno, mentre l'informazione è un insieme di dati, opportunamente aggregati, che assume un significato ben preciso in un determinato contesto. Le informazioni, insomma, sono il risultato della elaborazione dei dati.

¹¹ Un'entità può essere una persona, un luogo, una cosa materiale per la quale si devono gestire informazioni.

po. I programmi finalizzati all'informatizzazione di una determinata area applicativa, quale ad esempio la contabilità, gli stipendi, la fatturazione e il magazzino, producevano e aggiornavano un insieme di *file* dove venivano memorizzati i dati. Questo comportava la modifica di un certo numero di programmi applicativi ogni volta che veniva richiesta la gestione di dati non previsti nel disegno originario. Inoltre, si osservava una forte ridondanza del patrimonio informativo per il fatto che non veniva perseguita l'integrazione tra i *file* gestiti da applicazioni diverse, anche perché gli ambienti tecnologici utilizzati non erano interoperabili.

Negli anni '80, come soluzione alle problematiche evidenziate, si è affermato il concetto di *basi di dati*, cioè di collezioni di dati correlati, gestite con un *software* specifico, denominato *Data Base Management System* (DBMS). Il primo sistema di gestione di basi di dati, progettato da Charles Bachman alla General Electric negli anni '60, fu chiamato *Integrated Data Store* e costituì la base per lo sviluppo del modello reticolare. Sul finire degli anni '60, la IBM sviluppò l'*Information Management System* (IMS), introducendo il modello gerarchico per la rappresentazione dei dati in alternativa a quello reticolare. Nel 1970 Edgar Codd, nei Laboratori di Ricerca di San Jose, propose una nuova struttura chiamata modello relazionale, che ebbe un enorme successo e contribuì in misura determinante all'affermazione dei sistemi di gestione delle basi di dati. Oggi il settore dei DBMS rappresenta uno dei più grandi e vigorosi segmenti del mercato dell'Information Technology e lascia intravedere prospettive interessanti come la gestione delle basi di dati multimediali, le *digital libraries*, i *data warehouse*, gli *XML database*, etc.

Un Data Base Management System è un *software* specializzato nella gestione di basi di dati di grandi dimensioni, persistenti, nel senso che hanno un tempo di vita non limitato all'esecuzione dei programmi che le utilizzano, e normalizzate, dove

cioè la ridondanza dei dati è ridotta al minimo indispensabile¹². Esso assicura l'affidabilità e la riservatezza dei dati, aumentando l'efficienza e l'efficacia dell'intero sistema informativo. Ogni DBMS implementa un determinato *modello di dati*, inteso come l'insieme delle regole che determinano l'organizzazione dei dati e la loro descrizione con un linguaggio comprensibile dall'elaboratore. I modelli più diffusi sono: il modello *gerarchico*, il modello *reticolare*, il modello *relazionale* e il modello *ad oggetti*.

Il modello gerarchico è basato sull'uso di strutture ad albero, dove le entità sono descritte secondo la logica padre-figlio: un figlio può avere un solo padre, mentre un padre può avere più figli. La *radice* è l'elemento da cui prende origine la catena delle relazioni. Il modello reticolare è così chiamato perché ogni suo schema può essere rappresentato con un grafo, dove gli oggetti informativi sono i nodi e le relazioni gli archi. Esso differisce dal modello gerarchico per la possibilità di applicare relazioni del tipo molti a molti: un padre può avere più figli e un figlio può avere più padri. Il modello relazionale rappresenta la base di dati come una collezione di tabelle denominate relazioni. Le righe di una tabella prendono il nome di *tuple* e contengono l'insieme dei valori che identificano le entità del mondo reale. Il modello ad oggetti è il modello di dati più recente, che è stato ideato per rispondere all'esigenza di gestire non solo i dati alfanumerici, ma anche quelli multimediali, le immagini, le mappe geografiche, le registrazioni audio e video, etc. Esso rappresenta di fatto l'evoluzione del modello relazionale secondo la logica della programmazione ad oggetti.

I DBMS relazionali devono molto del loro successo alla disponibilità del linguaggio di interrogazione SQL (*Structured*

¹² La teoria della normalizzazione si occupa di definire criteri formali per valutare la qualità degli schemi di rappresentazione dei dati e di sviluppare algoritmi per trasformare uno schema in un altro equivalente, ma privo di anomalie.

Query Language), che nel corso degli anni è stato oggetto di un'intensa opera di standardizzazione da parte degli organismi ANSI (American National Standards Institute) e ISO (International Organization for Standardization). Attraverso l'SQL è possibile interrogare basi di dati gestite con DBMS di fornitori diversi in quanto tutti sono in grado di interpretare ed eseguire le stesse istruzioni. In realtà, i vari produttori di DBMS, a partire da una base standard minima, hanno arricchito l'SQL con funzionalità aggiuntive che aumentano le potenzialità del loro prodotto a scapito, però, dell'interoperabilità con i sistemi concorrenti.

Indipendentemente dal modello di dati implementato, un generico DBMS presenta un'architettura a tre livelli:

- *livello fisico o interno*, è la rappresentazione dei dati in rapporto alle strutture fisiche di memorizzazione utilizzate, ovvero al sottosistema di *storage* (*file* sequenziali, ad indici, *hashing*, etc.);
- *livello logico o concettuale*, è la rappresentazione dell'intera base di dati secondo il modello logico adottato: gerarchico, reticolare, relazionale o ad oggetti;
- *livello utente o esterno*, è la rappresentazione dei dati come sono percepiti dall'utente finale. Uno schema esterno può prevedere organizzazioni di dati diverse da quella proposta a livello logico.

Quest'architettura a tre livelli, schematizzata in figura 2, garantisce l'indipendenza dei dati dal sistema di gestione, sia a livello fisico che logico. L'indipendenza fisica riguarda la possibilità di interagire con il DBMS senza conoscere l'organizzazione fisica dei dati, ovvero di riorganizzare l'allocazione dei file nei dispositivi di memorizzazione senza incidere minimamente sulla descrizione dei dati ad alto livello e quindi sul *software* applicativo che interagisce con il DBMS. L'indipendenza logica, invece, riguarda la possibilità di avere più viste esterne dello stesso schema concettuale dei dati. Si sottolinea che gli utenti pos-

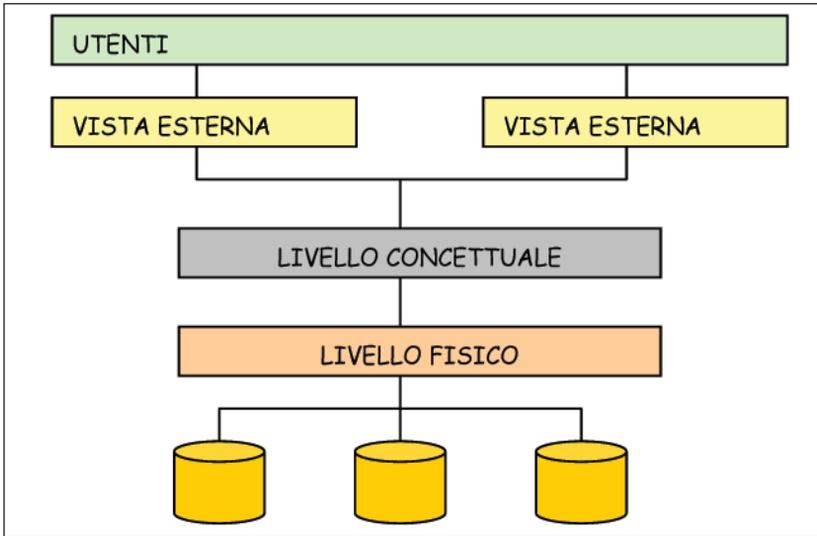


Fig. 2 - Architettura a tre livelli di un DBMS

sono accedere ai dati memorizzati nel *database* sempre e solo attraverso il livello esterno: è il DBMS che traduce le richieste in un linguaggio comprensibile ai livelli sottostanti.

Un DBMS presenta anche una serie di funzionalità volte a garantire la sicurezza e la riservatezza della base di dati. Le minacce più significative sono riconducibili a tre macro-aree: perdita d'integrità, perdita di disponibilità e perdita di riservatezza. La perdita d'integrità si ha quando ai dati vengono apportate modifiche accidentali o intenzionali che, se non rilevate con tempestività, possono compromettere l'affidabilità del *database*; la perdita di disponibilità si riferisce al fatto che il patrimonio informativo deve essere mantenuto accessibile con la massima continuità possibile; l'esigenza di riservatezza nasce dalla necessità di proteggere i dati da accessi non autorizzati. Tra le funzionalità che un buon DBMS mette a disposizione per la verifica dell'integrità della base di dati, assume particolare rilevanza la produ-

zione dei cosiddetti *audit trail*, che sono registrazioni cronologiche sufficienti per consentire la ricostruzione, la revisione e l'esame della sequenza di situazioni e di attività che hanno riguardato il patrimonio informativo gestito.

Gli attuali DBMS garantiscono prestazioni elevate e permettono di gestire basi di dati *distribuite* su più *server* interconnessi in rete e dislocati in aree geografiche diverse. Al limite si può arrivare ai cosiddetti DBMS *federati* (o sistemi *multidatabase*) in cui la base di dati è distribuita su più *server*, in ciascuno dei quali è presente un DBMS indipendente che può implementare un modello di dati diverso da quello degli altri DBMS della stessa federazione. Inoltre, in applicazioni complesse, dove i tempi di ricerca delle informazioni devono essere ridotti al minimo e dove si ha l'esigenza di soddisfare contemporaneamente un numero molto elevato di richieste, si possono utilizzare i DBMS dotati di *parallelismo*, che può essere di tipo *inter-query*, se vengono eseguite più interrogazioni diverse in parallelo, oppure *intra-query*, se vengono eseguite in parallelo più parti di una stessa interrogazione.

La tecnologia dei Data Base Management System è finalizzata prevalentemente alla gestione efficiente e affidabile di grandi quantità di dati *on line*, cioè sono sistemi indirizzati verso l'*On Line Transaction Processing* (OLTP). I dati gestiti, però, possono rivelarsi utili non solo per l'esecuzione delle transazioni, ma anche per la pianificazione e il supporto delle decisioni. All'inizio degli anni '90, a fianco dei sistemi OLTP si sono sviluppati i sistemi dedicati alla elaborazione e all'analisi dei dati, denominati *On Line Analytical Processing* (OLAP). L'elemento principale dell'architettura OLAP è una particolare base di dati denominata *data warehouse* (DW) nella quale vengono raccolte tutte le informazioni che, opportunamente analizzate, possono fornire un supporto alle decisioni. Un *data warehouse*, quindi, è una base di dati costituita attingendo da sorgenti informative preesistenti, è alimentata con meccanismi di tipo asincrono, contiene

informazioni di carattere storico-temporale e dati in forma aggregata. Con la disponibilità dei *data warehouse* si sono sviluppate nuove tecniche di analisi dei dati. Tra queste, assume particolare rilevanza il *data mining*, una tecnica che mira a cercare le informazioni “nascoste” in un DW ed è utilizzata nell’analisi di mercato, nell’analisi comportamentale e nell’analisi di previsione. Un processo di *data mining* viene svolto in modo iterativo e adattativo, innescando un meccanismo di costruzione progressiva della conoscenza basato sull’analisi statistica e sulle tecniche di intelligenza artificiale.

La crescente esigenza delle organizzazioni di gestire in modo integrato dati e documenti, cioè sia informazioni strutturate che destrutturate, unitamente all’affermarsi del formato standard XML per la produzione documentaria, ha determinato lo sviluppo dei cosiddetti *XML databases*. L’XML è un linguaggio di *markup*¹³ il cui scopo è quello di consentire la definizione di linguaggi personalizzati; si tratta di uno standard indipendente mantenuto dal W3C (*World Wide Web Consortium*) e comunemente utilizzato per l’interscambio di dati tra sistemi eterogenei. Con il termine *XML database* (XDBMS) si fa riferimento a due tipologie di *database*:

- *database XML-enabled*: sono *database* in cui i dati vengono memorizzati all’interno di strutture di tipo tradizionale (*database* relazionali, ad oggetti, gerarchici, etc.), dotate però di un apposito *middleware* che si occupa di estrarli e presentarli sotto forma di documenti XML e viceversa;
- *database XML-native*: sono *database* che memorizzano i documenti XML nella loro forma nativa, ovvero come documenti XML.

¹³ I linguaggi di *markup* prevedono l’inserimento di informazioni, frammiste al testo, che forniscono istruzioni relative alle modalità con cui questo deve essere trattato.

Gli *XML databases* si propongono come la soluzione ideale per l'archiviazione di oggetti informativi di tipo *document-centric* (o *document-oriented*), che contengono informazioni normalmente pensate per la fruizione diretta da parte dell'uomo, quali gli articoli di riviste, gli *e-book*, i manuali, i testi normativi, i messaggi di posta elettronica, i documenti scritti a mano, etc. Essi sono caratterizzati da una struttura irregolare e l'informazione è rappresentata dal documento *in sé*. Sebbene la tecnologia dei *database XML-native* sia ancora giovane mentre quella dei *database* relazionali è ben consolidata, si prevede che questo svantaggio sarà colmato in tempi piuttosto brevi visto l'interesse che l'XML e le tecnologie ad esso correlate stanno suscitando in tutti i settori applicativi.

4. Classificazione dei sistemi informativi

L'evoluzione delle tecnologie informatiche e il dinamismo delle organizzazioni chiamate a soddisfare le esigenze di una società in continua trasformazione, hanno cambiato la fisionomia dei sistemi informativi. Dai primi prototipi prevalentemente orientati al supporto dei processi di natura operativa si è arrivati ai sistemi integrati, dove le informazioni gestite a un qualsiasi livello dell'organizzazione sono elaborate, archiviate e rese disponibili a tutti gli altri livelli, compresi quello manageriale e direzionale. Oggi l'offerta dei sistemi informativi è così ricca che non è semplice assegnare a ciascuno di essi una ben precisa area di applicazione.

Per classificare i sistemi informativi si può fare riferimento alla famosa catena del valore di Porter, secondo la quale un'azienda produce valore per il cliente svolgendo una serie di attività che riguardano: l'acquisizione delle materie prime, la loro trasformazione in prodotti/servizi, le vendite, la consegna dei prodotti ai clienti, l'assistenza post vendita. A queste attività speci-

ficamente mirate alla produzione del valore (attività verticali) si aggiungono quelle di supporto (attività trasversali) che riguardano: gli approvvigionamenti, le tecnologie, le risorse umane e le infrastrutture. Lo schema risultante è riportato in figura 3.

La catena del valore di Porter permette di costruire una prima mappa dei sistemi informativi prendendo come criterio di classificazione l'area funzionale a cui sono principalmente destinati: la produzione piuttosto che le vendite, la gestione delle risorse umane anziché gli approvvigionamenti, e così via. Questa mappa può essere ulteriormente rifinita introducendo come ulteriori parametri di classificazione il livello dell'organizzazione e il tipo di utenza che il sistema informativo mira a supportare. Sulla base di questi due parametri, i sistemi si distinguono in: sistemi informativi *operativi*, sistemi per la gestione della *conoscenza*, sistemi di supporto alle *attività manageriali*, sistemi di supporto alle *attività direzionali*.

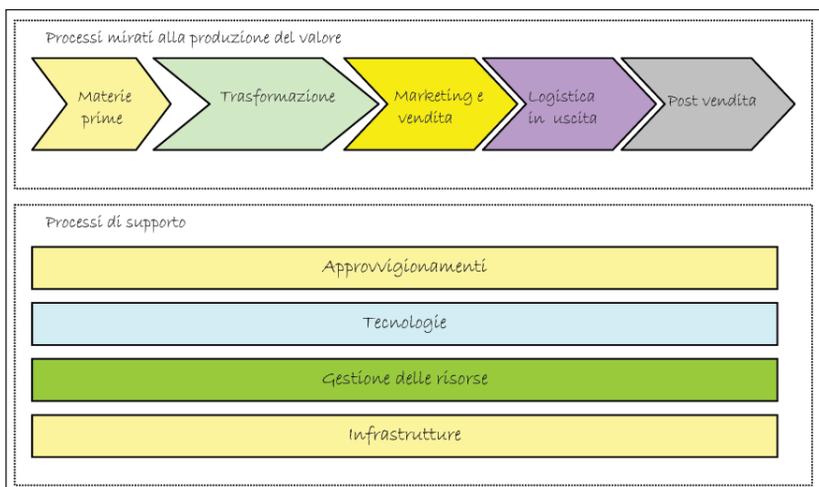


Fig. 3 - La catena del valore di Porter

I sistemi informativi operativi sono quelli disegnati per supportare le transazioni di natura operativa, quali ad esempio la gestione degli ordini, il calcolo degli stipendi, la gestione dei pagamenti e degli incassi. I processi supportati da questi tipo di sistemi sono caratterizzati da una definizione piuttosto accurata e sono eseguiti in maniera ripetitiva. I sistemi informativi per la gestione della conoscenza, invece, comprendono le applicazioni per l'*office automation*, il *knowledge management* e la gestione elettronica dei documenti. Essi mirano a supportare le attività di coloro che producono, gestiscono e diffondono conoscenza nell'organizzazione. I sistemi informativi a livello manageriale sono nati per aiutare i *manager* nello svolgimento delle attività di monitoraggio e controllo, oltre che nello sviluppo di processi decisionali di varia natura. Infine, i sistemi informativi a supporto delle attività direzionali rispondono alle necessità dell'alta direzione, che è chiamata a prendere decisioni non di *routine* e ad elaborare piani strategici di medio o lungo periodo.

Combinando i tre criteri di aggregazione sopra esposti si ottiene la classificazione dei sistemi informativi mostrata in figura 4. Dallo schema proposto si nota che i sistemi informativi di livello più alto, cioè quelli che mirano a supportare il livello manageriale e strategico, hanno bisogno di attingere alle informazioni create e gestite dai sistemi di livello più basso. E questo ci fa capire che la gestione ottimale della risorsa informazione si ottiene con un processo di automazione che parta dal basso, cioè dall'informatizzazione delle transazioni operative, e proceda verso l'alto, cioè verso la soddisfazione delle esigenze dell'alta direzione, integrando perfettamente i diversi livelli di sistema informativo. D'altra parte, le prime applicazioni informatiche hanno riguardato proprio le transazioni operative con l'intento di automatizzare quelle attività che erano ben formalizzate, ripetitive e onerose in termini di risorse impiegate.

L'insieme dei moduli applicativi utilizzati da un'azienda per supportare i suoi processi interni, dall'acquisizione delle materie

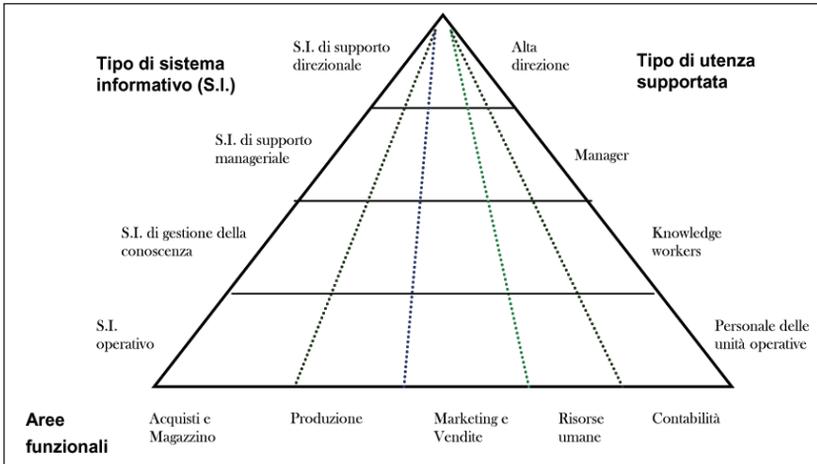


Fig. 4 - Classificazione dei sistemi informativi¹⁴

prime alla produzione, vendita e post vendita, prende il nome di sistema ERP (*Enterprise Resource Planning*). Tali sistemi sono caratterizzati dall'unitarietà dell'informazione, dalla modularità e dalla modellazione dei processi. Per unitarietà dell'informazione si intende la condivisione da parte di tutti i moduli applicativi della *suite* ERP di una stessa base di dati, ottenuta attraverso l'uso di un DBMS di ultima generazione e di meccanismi per la sincronizzazione dei dati. La modularità attiene alla tecnica di progettazione dei sistemi ERP che privilegia la scomposizione della *suite* in applicazioni informatiche autosufficienti, ciascuna mirata alla soluzione di problematiche specifiche, ma tutte pienamente integrate tra di loro. Questa modularità permette alle organizzazioni di realizzare gradualmente il loro sistema informativo, aggiungendo moduli quando la base applicativa sottostante

¹⁴ Schema tratto da KENNETH LAUDON, JANE LAUDON, *Management dei sistemi informativi*, Milano, Pearson Education Italia Srl, 2004, p. 46.

è consolidata, coerentemente con i piani di sviluppo di medio e lungo periodo. La modellazione dei processi è implicita nei sistemi ERP in quanto sono specializzati per area applicativa e quindi realizzati sulla base di schemi operativi comuni alla maggior parte delle aziende che operano nello stesso settore di attività. Ciò comporta che molto spesso sono le organizzazioni a dover adeguare i loro processi agli schemi implementati nei sistemi ERP, e non viceversa. Da un lato questo riduce la flessibilità di questa tipologia di sistemi informativi, ma dall'altro garantisce la piena integrazione tra le applicazioni che li compongono in quanto queste non vengono "stravolte" per soddisfare le esigenze particolari di questo o quel cliente.

Oggi quasi tutte le aziende puntano alla soddisfazione e fidelizzazione del cliente per il semplice motivo che un cliente soddisfatto resta fedele nel tempo e percepisce la qualità del servizio. Il "marketing transazionale", che vedeva ogni rapporto tra produttore del bene e acquirente come una specifica e occasionale negoziazione, ha lasciato il posto al "marketing moderno", che punta sulla continuità del rapporto anche in assenza di una transazione di vendita-acquisto. I sistemi CRM (*Customer Relationship Management*) sono nati recentemente per supportare i processi finalizzati al miglioramento delle relazioni con i clienti; essi sono complementari ai sistemi ERP e interessano in modo particolare il *marketing*, le vendite e il supporto post-vendita. Un sistema CRM è caratterizzato dalla multicanalità, cioè dalla disponibilità di un complesso tecnologico che permette all'azienda di interagire con il cliente attraverso il canale di comunicazione che egli preferisce: telefono (*call center*, *smartphone*, SMS, MMS), Internet (Web, *e-mail*), canale televisivo (DTT, IP-TV, TV-SAT). Nelle forme più avanzate di CRM, i moduli *software* interagiscono direttamente con gli applicativi dei sistemi ERP per favorire l'integrazione dei processi di *front office* con quelli di *back office*, garantendo l'unicità dell'informazione sui clienti e sui prodotti.

Il grande successo di Internet ha favorito anche lo sviluppo dei sistemi informativi per l'*e-commerce*, ovvero per supportare le transazioni commerciali, di vendita o acquisto, che avvengono per via telematica. Sulla base dei soggetti coinvolti nei rapporti e nelle transazioni, si possono distinguere due diverse tipologie di *e-commerce*: il B2B e il B2C. Il commercio elettronico di tipo *B2B* riguarda lo scambio di beni e servizi tra le imprese e può essere di tipo *diretto*, quando Internet è il mezzo attraverso il quale transitano tutte le informazioni tipiche di un rapporto di compravendita (dall'ordine di acquisto fino al pagamento), oppure *indiretto*, quando Internet si propone come semplice punto di incontro tra domanda ed offerta, ma poi la transazione viene completata con mezzi tradizionali. Il *B2C* è quella parte del commercio elettronico che si riferisce ai rapporti tra aziende e consumatori finali.

Nella categoria dei sistemi informativi per l'*e-commerce*, assumono una fisionomia a sé stante quelli per l'*e-procurement*, ovvero per la gestione degli approvvigionamenti *on line*. L'*e-procurement* si è sviluppato in diversi modelli, ognuno dei quali è caratterizzato da pregi e difetti. Nel modello *sell-side one to many*, il venditore crea un proprio sito Internet dove gli acquirenti possono entrare ed acquistare i prodotti in modalità *on line* ed in tempo reale, con contratti di acquisto specifici per ciascun prodotto. Nel modello *buy-side one to many*, l'azienda acquirente mantiene al proprio interno i cataloghi di diversi venditori ed è responsabile dell'invio degli ordini e dei relativi pagamenti. I *marketplace* sono luoghi di intermediazione che favoriscono l'incontro tra la domanda e l'offerta di beni e servizi. Sono realizzati con una piattaforma tecnologica che permette agli operatori economici di scambiarsi informazioni, presentare o cercare prodotti/servizi, perfezionare e concludere transazioni economiche. I *marketplace* si definiscono *verticali*, se sono progettati per settori merceologici specifici e mirano a soddisfare i bisogni di una comunità ben definita, oppure *orizzontali*, se accolgono pro-

dotti e servizi di aree merceologiche diverse. Le *aste on line* forniscono la possibilità di scambiare prodotti/servizi con la prospettiva di ottenere le migliori condizioni economiche. La peculiarità delle *aste on line* è rappresentata dal fatto che gli acquirenti e i fornitori si incontrano per contrattare la compravendita di un prodotto o servizio in un luogo “virtuale” accessibile solo attraverso Internet. Come nelle aste tradizionali, anche in quelle *on line* i beni e i servizi trattati sono inizialmente definiti in termini di quantità, prezzo e specifiche, ma nel corso dell'asta una o più di tali caratteristiche (di solito il prezzo) possono cambiare. I modelli d'asta più usati *on line* sono: l'*asta inglese*, che vede il banditore fissare un prezzo dal quale gli offerenti partono con le proprie proposte, al rialzo o al ribasso, fino ad arrivare a quella più vantaggiosa del vincitore; l'*asta olandese*, dove il banditore parte da un prezzo che riduce progressivamente fino ad ottenere un'offerta di acquisto dai partecipanti; l'*asta con offerta segreta*, nella quale i partecipanti propongono segretamente un prezzo unico e l'aggiudicazione è fatta al miglior prezzo pervenuto al banditore dell'asta entro una scadenza prefissata. Un sistema di *e-procurement* cambia radicalmente il modo di gestire gli approvvigionamenti nelle aziende come nelle pubbliche amministrazioni, velocizzando il processo, che può concludersi in pochi minuti anziché richiedere giorni o settimane, semplificandolo, in quanto tutto avviene con l'ausilio di un *software* che arriva a gestire anche i pagamenti *on line*, e assicurando un buon livello di sicurezza nelle transazioni.

Nella pubblica amministrazione l'*e-procurement* rientra nell'insieme più vasto dei sistemi informativi per l'*e-government*, che sono finalizzati al miglioramento dei servizi pubblici e al rafforzamento del processo democratico, in sintonia con quanto previsto dal Consiglio d'Europa riunito nella conferenza di Lisbona (marzo 2000). In questo contesto trovano applicazione le tecnologie più innovative, quali la firma elettronica, alla quale il legislatore italiano ha attribuito una forza probatoria equivalente

alla sottoscrizione autografa, la posta elettronica certificata, che fornisce le stesse certezze giuridiche delle tradizionali raccomandate con ricevuta di ritorno, i sistemi di autenticazione *on line* basati sull'uso della carta d'identità elettronica o della carta nazionale dei servizi, i sistemi di pagamento elettronici, che permettono di eseguire *on line* l'intera procedura che porta alla erogazione del servizio richiesto dall'utente¹⁵.

Ikujiro Nonaka, uno dei maggiori teorici del *knowledge management*¹⁶, afferma che «in un'epoca in cui l'unica certezza è l'incertezza, l'unica fonte sicura per il vantaggio competitivo è la conoscenza». La conoscenza è da sempre considerata una delle principali fonti di ricchezza per ogni individuo, qualcosa che permette di distinguersi e di competere nel contesto sociale, ma oggi viene vista anche come una ricchezza aziendale e le organizzazioni sono chiamate ad adottare metodi e sistemi per identifi-

¹⁵ L'Unione Europea ha definito quattro possibili livelli d'interazione *on line* tra l'utenza e la pubblica amministrazione:

- *livello 1: informativo*. Disponibilità *on line* solo delle informazioni necessarie per avviare la procedura che porta alla erogazione del servizio;
- *livello 2: download modulistica*. Possibilità di scaricare *on line* i moduli necessari ad avviare la procedura che porta all'erogazione del servizio (interazione *one way*).
- *livello 3: inoltrare istanze*. Possibilità di scaricare la modulistica e di avviare *on line* la procedura che porta alla erogazione del servizio (interazione *two way*);
- *livello 4: esecuzione e transazione, compreso l'eventuale pagamento e la consegna*. Possibilità di eseguire *on line* l'intera procedura che porta alla erogazione del servizio, compreso l'eventuale pagamento dei relativi oneri, la notifica e la consegna dei provvedimenti.

¹⁶ Il *knowledge management* è un processo sistematico usato per creare, ricercare, sintetizzare, acquisire, condividere informazioni, intuito ed esperienze per il conseguimento degli obiettivi dell'organizzazione. In altre parole, è l'insieme delle attività e dei processi che attengono alla generazione, mappatura, selezione, organizzazione e diffusione delle conoscenze.

care e convertire le esperienze, le specialità e le abilità individuali in risorse per l'organizzazione. Da qui lo sviluppo dei sistemi informativi per il *knowledge management* (KM), che mirano a rendere la conoscenza, intesa come "capitale intellettuale" costituito dalle esperienze, professionalità, patrimonio informativo e documentario di un'organizzazione, ampiamente accessibile e utilizzabile ad ogni livello. Sotto il profilo tecnologico, un sistema di knowledge management assomma le funzionalità di un sistema di gestione delle risorse umane (HRM), un sistema di gestione elettronica dei documenti (EDMS), un sistema di gestione dei flussi di lavoro (WFMS) e un sistema di *content management* (CMS).

In un'organizzazione basata sulla conoscenza è importante che ogni singola persona impiegata sia valutata per le conoscenze che possiede, per la sua capacità di analisi, di diagnosi, di azione, e pertanto il sistema informativo HRM deve supportare efficacemente le politiche dell'organizzazione volte ad incentivare la trasformazione della conoscenza tacita in conoscenza esplicita e a favorire i processi di apprendimento e condivisione del sapere. Un classico sistema EDMS permette di produrre documenti elettronici e digitalizzare quelli cartacei, registrandoli su supporto informatico insieme ai metadati che li identificano e ne permettono la ricerca e l'accesso. Esso, inoltre, consente di formare i fascicoli elettronici e di renderli disponibili *on line* a tutti i soggetti che sono autorizzati ad accedervi. I sistemi WFMS supportano il lavoro collaborativo e l'automazione dei processi dell'organizzazione; essi sono definiti come «sistemi in grado di attivare e gestire l'esecuzione di flussi di lavoro con strumenti software capaci di interpretare il formato di definizione dei processi, di interagire con i partecipanti al flusso e, se necessario, di invocare l'esecuzione di applicazioni informatiche». Un Workflow Management System si compone di tre elementi:

- *un descrittore dei processi*, che permette di disegnare graficamente la sequenza delle attività che li compongono, in-

dicando per ciascuna attività le regole di esecuzione;

- *un gestore degli stati di avanzamento*, che consente di dare avvio ad un processo e di monitorare il suo stato di avanzamento, registrando le azioni dei soggetti coinvolti;
- *un generatore di report e statistiche*, che permette di analizzare i tempi di svolgimento dei processi e valutare l'efficienza dell'organizzazione.

La creazione, l'archiviazione "intelligente", l'aggiornamento continuo e la diffusione di contenuti informativi digitali attraverso siti e portali Web hanno trovato nei sistemi di content management (CMS) una soluzione molto efficace. Per contenuti informativi si intendono documenti, comunicati stampa, descrizioni di prodotti, manualistica, informazioni finanziarie, ordini, fatture, modulistica elettronica e in genere tutti gli oggetti informativi che circolano tra gli attori che compongono la catena delle relazioni dell'organizzazione. Tali sistemi realizzano la gestione di contenuti in quattro fasi:

- *creazione del content*: è la fase che attiene alla raccolta, selezione e classificazione delle informazioni che si vogliono rendere disponibili attraverso il sito o il portale Web;
- *personalizzazione del content*: è la fase nella quale si riorganizza il *content* per soddisfare al meglio e più velocemente le richieste dei futuri fruitori;
- *gestione del content*: riguarda la capacità del CMS di attivare flussi di lavoro automatizzati che guidano il *content* attraverso le persone che hanno la responsabilità di crearlo, verificarlo, certificarlo, pubblicarlo ed eliminarlo quando obsoleto;
- *distribuzione del content*: è la fase che comprende l'attività di *deployment*, cioè la selezione e l'aggregazione dei contenuti in base a regole predefinite, l'attività di *publishing*, che determina il layout grafico e la logica di presentazione dei contenuti, e l'attività di *delivery*, che rappresenta la di-

stribuzione vera e propria dei contenuti verso i dispositivi utilizzati dagli utenti.

Un buon CMS consente ai creatori dei contenuti informativi di concentrarsi sul loro lavoro senza essere “distratti” dalle problematiche tecnologiche connesse all’impianto dove tali contenuti troveranno collocazione. Soprattutto, facilita l’aggiornamento delle basi informative pubblicate su Web, in quanto questo può essere eseguito direttamente da coloro che hanno la responsabilità della creazione del contenuto, senza l’intermediazione del personale informatico.

Al livello più alto della piramide riportata nella figura 4, che mostra lo schema di classificazione dei sistemi informativi, troviamo quelli orientati al supporto manageriale e direzionale. I primi sistemi dedicati al supporto delle decisioni furono i *Management Information Systems* (MIS), che fornivano informazioni, solitamente sotto forma di *report* generati a scadenze programmate, sull’andamento dell’azienda vista nel suo complesso. Successivamente, hanno trovato spazio i DSS (*Decision Support Systems*), che possono supportare le attività decisionali dei *manager*¹⁷, basandosi su modelli previsionali e comportamentali, oppure sfruttando la tecnologia dei *data warehouse* e la tecnica del *data mining* per analizzare enormi quantità di dati. Oggi esistono sistemi DSS evoluti che si propongono efficacemente per la gestione delle decisioni relative ai processi di fornitura, ai rapporti con i clienti, alla gestione dei gruppi di lavoro¹⁸. I sistemi ideati per supportare le decisioni dell’alta direzione, che normal-

¹⁷ Le decisioni supportate dai sistemi DSS possono essere di tipo strutturato, se sono basate su regole chiare e definite, oppure semi-strutturato e non-strutturato, se sono parzialmente o totalmente affidate alla creatività e all’esperienza del *manager*.

¹⁸ Per approfondimenti si veda K. LAUDON, J. LAUDON, *Management dei sistemi informativi*, cit., pp. 421-455.

mente sono di tipo non-strutturato, prendono il nome di *Executive Support Systems* (ESS). Essi combinano i dati sull'andamento dei diversi comparti dell'organizzazione con le informazioni che provengono dall'esterno, quali quelle relative al *trend* dei mercati, alle strategie delle aziende concorrenti, ai fenomeni che incidono positivamente o negativamente sulla capacità di produzione, e così via. Il patrimonio informativo così ottenuto viene messo a disposizione dei *senior manager* insieme agli strumenti che li aiutano a formulare i problemi, a valutare i dati, ad analizzare le tendenze e, quindi, a prendere le decisioni strategiche.

5. Ciclo di vita dei sistemi informativi

Nei paragrafi precedenti si è evidenziata da un lato la valenza strategica dei sistemi informativi per le moderne organizzazioni pubbliche o private, dall'altro la loro complessità derivante dal fatto che la risorsa informazione interessa tutti i settori applicativi. Inoltre, la continua evoluzione delle tecnologie e la trasformazione dei modelli organizzativi, indotta dai repentini mutamenti dei mercati e dei bisogni della società, rendono necessaria la verifica continua del livello di efficacia e di efficienza di un sistema informativo. Da queste considerazioni consegue la necessità di adottare una metodologia ben definita per la progettazione e la realizzazione dei sistemi informativi. Quella tradizionalmente utilizzata è basata sul "ciclo di Deming", che prevede lo svolgimento ciclico di quattro fasi: pianificazione, realizzazione, verifica ed intervento.

La fase di pianificazione ha lo scopo di raccordare il piano di sviluppo del sistema informativo con gli obiettivi strategici dell'organizzazione, tenendo conto delle priorità stabilite dall'alta direzione e delle risorse realmente disponibili. Il prodotto della pianificazione è il *piano per l'informatica*, che individua le aree applicative d'intervento, delinea un quadro tecnologico di riferi-

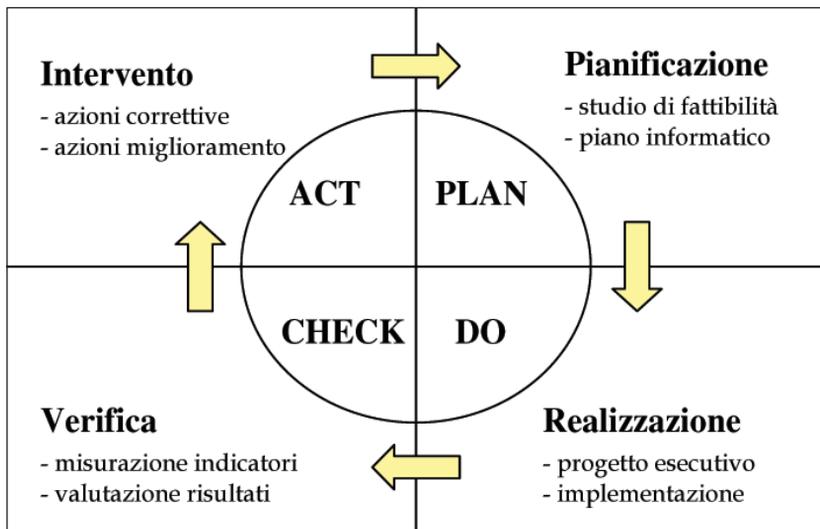


Fig. 5 - Il ciclo di Deming

mento, assegna le risorse, distribuendole su un arco temporale che normalmente va da tre a cinque anni.

L'individuazione delle aree applicative d'intervento, ovvero dei progetti da realizzare nell'arco temporale coperto dal piano per l'informatica, richiede una valutazione attenta dei costi, dei benefici e delle criticità connesse a ciascuna realizzazione. Lo *studio di fattibilità* ha proprio lo scopo di fornire a chi ha la responsabilità del sistema informativo quelle informazioni che sono necessarie per valutare l'opportunità e la fattibilità di un progetto. Esso «rappresenta in generale un documento che raccoglie l'insieme delle informazioni giudicate, nei vari contesti, sufficienti per decidere consapevolmente sull'investimento e per attivare i passi realizzativi veri e propri»¹⁹. Un generico studio di fattibilità comprende:

¹⁹ Cfr. GABRIELE LAZZI, *Lo studio di fattibilità*, in *Sistemi informativi*, Milano, Franco Angeli, 2001.

- l'analisi dello stato attuale, che ha come finalità la descrizione della problematica e la definizione delle esigenze, dei vincoli e delle criticità;
- l'esplicitazione degli obiettivi che si intendono conseguire;
- il progetto di massima, che delinea la soluzione da adottare in termini tecnologici, applicativi e organizzativi;
- l'analisi del rischio, dove si mostrano e si quantificano gli elementi che possono mettere a rischio il successo dell'iniziativa;
- l'analisi dei costi/benefici, dove i benefici devono trovare una quantificazione per essere confrontati con i costi stimati;
- il piano di attuazione, che proietta le fasi di realizzazione del progetto sul piano temporale di riferimento, fissando anche i punti di decisione (*milestones*), cioè i «momenti in cui si dovranno prendere le decisioni sulla prosecuzione del progetto, analizzando il lavoro effettuato»²⁰.

La fase di realizzazione prende avvio dal piano per l'informatica, dove sono individuate le linee di sviluppo del sistema informativo nel medio o lungo periodo, e riguarda l'attuazione concreta di ogni singolo intervento pianificato, prendendo l'anno come base di riferimento temporale. Il primo *step* è rappresentato dalla elaborazione di un *progetto esecutivo* a partire da quello di massima predisposto nell'ambito dello studio di fattibilità. In questo nuovo documento si devono descrivere dettagliatamente le soluzioni tecnologiche, procedurali ed organizzative da implementare, fissando un piano di realizzazione che individui i tempi e le modalità di svolgimento di ogni attività. Alla progettazione esecutiva seguono lo sviluppo del *software*, la sua implementazione, l'avviamento operativo e la gestione.

²⁰ *Ibidem.*

Più volte nei paragrafi precedenti si è affermato che un sistema informativo deve garantire la gestione della risorsa informazione per lo svolgimento ottimale dei processi dell'organizzazione. È buona norma, quindi, anteporre alla progettazione del sistema informativo il *reengineering* dei processi interessati, in modo da combinare l'introduzione dei nuovi strumenti tecnologici con la rimodulazione degli assetti organizzativi e la revisione delle procedure. Il *Business Process Reengineering* (BPR) è un approccio al cambiamento che si è affermato intorno agli anni '90 ed ha come caratteri distintivi l'organizzazione per processi, la riprogettazione radicale, la discontinuità delle prestazioni e l'uso delle tecnologie dell'informazione e della comunicazione (ICT). La visione per processi, come rilevato nel paragrafo 2, si presta bene a soddisfare le esigenze di miglioramento della competitività di un'organizzazione, in quanto l'ottimizzazione di intere catene di attività si ripercuote direttamente e positivamente sui tempi di produzione e sui costi dei prodotti/servizi. Nell'ambito del *process management* i possibili approcci sono due: il miglioramento radicale o il miglioramento continuo (incrementale). Il primo punta alla riprogettazione radicale dei processi, senza lasciarsi condizionare dalla situazione esistente, con l'obiettivo di conseguire un miglioramento discontinuo delle prestazioni rispetto ai valori attuali; il secondo, invece, propone interventi di miglioramento continuo, che portino gradualmente alla eliminazione delle inefficienze rilevate in fase di analisi. Molto spesso gli interventi di *reengineering* combinano entrambi i tipi di approccio al fine di minimizzare i rischi dell'intervento e ottimizzare il rapporto costi/benefici. Il BPR trova il suo naturale completamento nell'introduzione delle tecnologie informatiche e in particolare dei sistemi di *workflow management* descritti nel precedente paragrafo 3, che automatizzano le attività di un processo e gestiscono le comunicazioni tra gli attori coinvolti, prescindendo dalla tradizionale organizzazione per aree funzionali.

Un esempio di metodologia di BPR è quella adottata dall'Autorità per l'Informatica nella Pubblica Amministrazione (AI-PA)²¹, che prevede lo svolgimento di tre fasi:

- *definizione e campo di applicazione*: questa fase ha come obiettivo la definizione dei confini dell'intervento di *reengineering*, specificando il campo di applicazione, le motivazioni, le opportunità e i vincoli;
- *diagnosi delle criticità e delle priorità*: l'obiettivo è quello di individuare, attraverso l'analisi dello stato attuale e la modellazione dei processi, il *gap* esistente tra le esigenze delle parti interessate e i servizi erogati, definendo di conseguenza le criticità e le priorità d'intervento;
- *riprogettazione dei processi*: questa è la fase che riguarda il ridisegno dei processi con l'approccio della riprogettazione radicale. È importante che i nuovi flussi di lavoro siano definiti unitamente ai piani di monitoraggio dei livelli prestazionali e alle attività di *change management*, in modo da assicurare/verificare l'effettiva applicazione delle nuove logiche operative.

La schematizzazione dei processi e della base di dati dovrebbe essere realizzata con un linguaggio standard, sia per documentare con diagrammi universalmente comprensibili la logica degli applicativi da realizzare e sia per mantenere distinti i risultati della fase di analisi dalle specifiche tecniche del *software*. Tra questi linguaggi, assume particolare rilevanza l'*Unified Modeling Language* (UML), che è il risultato delle migliori esperienze di ingegneria del *software* per specificare, visualizzare, costruire e documentare sistemi. L'UML può essere utilizzato a tre livelli: schema, progetto e linguaggio di programmazione.

²¹ Oggi Centro Nazionale per l'Informatica nella Pubblica Amministrazione (CNIPA).

Nella maggior parte dei casi viene utilizzato a livello schema per documentare alcuni aspetti del sistema, favorendo la comprensione e la comunicazione tra coloro che partecipano alla fase di analisi. L'uso dell'UML a livello progetto è finalizzato a descrivere in modo dettagliato il sistema da implementare, arrivando a produrre la documentazione per i programmatori che dovranno sviluppare il *software*. Il terzo livello dell'UML permette di trasformare i diagrammi descrittivi del sistema in codice eseguibile. I vantaggi connessi all'utilizzo dell'UML nella realizzazione dei sistemi informativi possono essere riassunti nei seguenti punti:

- permette di disegnare e documentare il *software* prima ancora di svilupparlo, e quindi di valutare in anticipo il risultato finale del progetto che si sta realizzando;
- permette di documentare efficacemente i processi, la base informativa e le relazioni tra tutte le entità del sistema, riducendo i tempi di sviluppo e i costi del *software*;
- facilita la comunicazione e l'interazione tra il personale coinvolto nello sviluppo del sistema informativo, attenuando i rischi di incomprensione.

Alla fase di realizzazione segue quella di verifica dei risultati (*assessment*), definita come «un processo strutturato che tende ad identificare lo stato e le eventuali criticità di un sistema informativo automatizzato, nonché le cause e gli interventi da mettere in atto per correggere le disfunzioni e conseguire miglioramenti in termini di efficienza ed efficacia dei processi di servizio supportati»²². La fase di verifica dovrebbe essere eseguita sia per i sistemi informativi di nuova attivazione, per i quali ovviamente è necessario stabilire se sono stati ben concepiti e realizzati, e

²² Cfr. ALESSANDRO ALESSANDRONI, *L'assessment dei sistemi informativi*, in *Sistemi informativi*, Milano, Franco Angeli, 2001.

sia per i sistemi informativi già in uso, che potrebbero non essere più rispondenti alle esigenze dell'organizzazione in quanto queste si modificano nel tempo.

I risultati dell'*assessment* costituiscono l'*input* della fase di intervento che riguarda la progettazione delle azioni correttive e di miglioramento necessarie per eliminare le criticità rilevate. Queste azioni si ripercuotono ovviamente sul piano per l'informatica, che dovrà essere rimodulato per tenere conto degli interventi aggiuntivi rispetto alla pianificazione iniziale. Il ciclo di vita dei sistemi informativi, quindi, è un ciclo ripetitivo che vede la formulazione di un piano di sviluppo e il suo continuo aggiornamento in base ai risultati delle realizzazioni effettuate e alle mutevoli esigenze informative dell'organizzazione.

Indice

<i>Introduzione</i>	Pag. 5
<i>Roberto Guarasci</i> Che cos'è la documentazione	“ 9
<i>Antonio Romiti</i> Archivistica e Documentazione tra passato e presente	“ 27
<i>Piero Innocenti</i> Bibliografia, Biblioteconomia, Documentazione	“ 43
<i>Giuseppe Scandurra</i> Il Documento	“ 77
<i>Renato Rolli</i> Il diritto di accesso ai documenti amministrativi	“ 113
<i>Anna Rovella</i> Produzione ordinamento e gestione dei documenti	“ 157
<i>Eduardo De Francesco</i> I linguaggi di descrizione documentale	“ 179
<i>Roberto Guarasci</i> Indicizzazione e classificazione: concetti generali	“ 211
<i>Giovanni Adamo</i> La terminologia	“ 223

<i>Daniele Gambarara - Alfredo Givigliano</i> Thesauri, mappe semantiche, ontologie. Problemi semantici e costruzioni concettuali	Pag. 245
<i>Maurizio Lancia - Andrea Lapicciarella</i> Data Mining e Text Mining	“ 273
<i>Stefano Pigliapoco</i> I sistemi informativi	“ 311