

Progetto “Rassegna Stampa Multimediale e Sistema di Classificazione Automatica” – CL.AU.d.io V.2

Tipologia finanziamento: Convenzione operativa tra Cosmic Blue Team SpA (CBT) e Consiglio Nazionale delle Ricerche (C.N.R.) per la realizzazione di un tool per la classificazione semiautomatica di risorse digitali prot.n.56486 del 29/07/2011.

Durata: settembre 2011 – giugno 2013

Obiettivo del progetto è stato lo sviluppo di un sistema di classificazione automatica per documenti digitali e unità informative provenienti dal sistema documentale della rassegna Stampa Multimediale di Barilla. Il sistema realizzato consente all’utente di prendere visione delle possibili categorie da associare ad ogni unità informativa, rendendo possibile un confronto tra la categoria proposta dal provider e quelle suggerite dal sistema di classificazione.

Il funzionamento del classificatore dipende da una base di conoscenza incrementale, la quale inizialmente rappresenta il training set necessario per addestrare il sistema, diventando poi il punto di riferimento per processare ogni nuovo documento entrante.

Risultati conseguiti: L’applicativo fornito dal CNR è una black box che integra la funzione di classificazione semiautomatica del tool di CBT. Il sistema comunica direttamente con il software di CBT dal quale riceverà in input i dati necessari alla classificazione e propone, ovvero assegna, per ogni documento processato, una o più classi di appartenenza, determinate sulla base di uno schema pre-coordinato ed una tassonomia, entrambi definiti sulle reali esigenze e richieste di Barilla Group.