

EINS-GEM e il servizio I² plus La ricerca su banche dati professionali e siti web

PAOLA MANNI - GABRIELLA SCIPIONE

L'intervento descrive la fisionomia e le funzioni del servizio EINS-GEM che, nell'ambito di EINS, European Information Network Services, si propone di fornire un servizio avanzato di accesso all'informazione in linea. Particolare attenzione nell'intervento è posta sul ruolo del CINECA che, come centro tecnologico di EINS, a partire dalla tecnologia DIOGENE, ha sviluppato il servizio I² e ha ora realizzato un'applicazione più avanzata, I² plus, che consentirà la ricerca contemporanea sulle banche dati disponibili in EINS-GEM e su una collezione di siti web. Tale applicazione è oggi rivolta al settore aerospaziale, ma sarà estesa, in un prossimo futuro, ad altri settori.

Parole chiave: EINS – CINECA – I² – I² plus – Banche dati – Siti web – Ricerca contemporanea

EINS – European Information Network Services – è l'organizzazione nata all'inizio del 1998 per la fornitura di un servizio di accesso a banche dati fino a quel momento distribuito da ESA-IRS, l'Agenzia Spaziale Europea.

Tale organizzazione, che si è recentemente strutturata in un Gruppo Europeo di Interesse Economico (GEIE), vede come partner la **British Library**, con funzioni di coordinamento, oltre che di centro nazionale per i Paesi di lingua inglese; **COBIDOC**, quale centro amministrativo del consorzio e centro nazionale olandese; **CINECA** [Consorzio Interuniversitario per il Calcolo Automatico dell'Italia Nord Orientale], con la triplice funzione di centro nazionale italiano, centro tecnologico e host di alcune delle banche dati disponibili tramite il servizio.

Scopo di EINS è quello di fornire un servizio avanzato ed allargato di accesso all'informazione *online*, reperibile sia in Europa che negli Stati Uniti, nell'ambito di discipline scientifiche, tecniche, economiche, offrendone una continua, totale disponibilità; tale servizio è denominato EINS-GEM ed è raggiungibile all'indirizzo: <<http://www.einsgem.org>>.

Il consorzio EINS ha come partner servizi *online* europei esistenti, quali **DIMDI** e **FIZ-Karlsruhe**, che mettono a disposizione le loro banche dati, e una rete di centri nazionali, che sulla base di accordi contrattuali formano una rete di supporto all'utenza di tale servizio.

Il contratto fra i partner, assieme ai contratti sottoscritti con i produttori delle banche dati, solleva l'utente dal dover firmare un contratto per ogni host di suo interesse; egli gode inoltre della facilitazione di un unico referente amministrativo, e quindi di ricevere un'unica fattura per tutti i servizi *online* interrogati.

Il CINECA, nella sua funzione di centro tecnologico di EINS, ha sviluppato un sistema completamente nuovo per la gestione dei servizi, denominato Dante, che gestisce appunto l'interconnessione di vari servizi *online* europei: ciò permette all'utente l'accesso, con un'unica interfaccia, un unico contratto, e soprattutto con un unico linguaggio di interrogazione, a fonti informative residenti su host diversi, in modo totalmente trasparente, e gli dà modo di navigare fra queste senza fatica.

Il sistema DANTE, su cui il servizio si basa, permette infatti ricerche complesse ed esaustive grazie all'impiego di un sofisticato e nel contempo semplice linguaggio di interrogazione, il tradizionale linguaggio Quest, che si unisce efficacemente alle potenzialità di Internet, tramite la modalità denominata **Command Search**. È disponibile anche l'accesso semplificato grazie ad una **Easy Interface**, il cui utilizzo non richiede la conoscenza, da parte dell'utente, né del linguaggio d'interrogazione né della struttura dell'informazione. È altresì operativo l'accesso più tradizionale via **Telnet** e **X-25**, mentre è in corso di realizzazione una interfaccia **Advanced** per utenti più esperti del sistema.

I servizi aggiuntivi

Oltre al tradizionale accesso alle banche dati, interrogabili con diverse modalità che rispondono ai bisogni di utenti più o meno esperti, sono inoltre disponibili alcuni servizi aggiuntivi, che forniscono l'accesso, facilitato per gli utenti EINS, a numerosi altri servizi *online*, quali EINSBiosci e Lexis/Nexis, o a strumenti di reperimento dell'informazione su Internet, quali I² (I square), l'osservatorio sul mercato internazionale dell'informazione.

I² (I square)

I² si basa sulla tecnologia DIOGENE, totalmente sviluppata al CINECA, e, nella sua versione disponibile attualmente sul sito di EINS-GEM, è lo strumento che permette di monitorare l'informazione sul mercato dell'informazione disponibile su web.

DIOGENE si compone di due moduli, uno per il recupero dei dati ed uno per la loro analisi. È prevista la possibilità di aggiungere un terzo modulo per l'eventuale trattamento di meta-dati.

Un robot ha il compito di cercare l'informazione all'interno di una collezione di siti selezionati e di scaricare le pagine pertinenti. Tale informazione viene raccolta

in una banca dati, che è quindi continuamente ampliata dai nuovi dati aggiunti, all'interno della quale le *query* vengono effettuate da un potente motore di ricerca.

Sulle pagine reperite dal robot vengono effettuati una serie di controlli al fine di scaricare le pagine più aggiornate dei siti, sostituendole a quelle già esistenti nella collezione.

Il contenuto dei dati reperiti sui siti viene indicizzato ed è su questi indici che la ricerca viene effettuata; ciò consente perciò di formulare domande precise, pur interrogando dei documenti *full-text*, e di ottenere risposte rapide.

I², all the information on information

I² offers free searching through the full-text of a number of leading information resources, providing an observatory on the information market. See the I² Collections for a list of the sources accessed.

Search

Search keywords:

Options

Order by:

All the words Any of the words The exact phrase

Proximity: (No. of words)

Help

A service offered by EINS

Based on DIOGENE Technology by CINECA

fig.1 – Il motore di ricerca I²

I² è un'applicazione della tecnologia DIOGENE mirata al mercato dell'informazione, già disponibile con accesso libero sulla *home page* del servizio EINS-GEM (<<http://www.einsgem.net>>).

I² plus

- Perché I² plus

Sulla base della logica e dell'implementazione del servizio I² è ora in fase di sviluppo un'applicazione più avanzata, che permette la ricerca contemporanea sulle banche dati disponibili in EINS-GEM e su una collezione di siti web.

Riteniamo infatti che il passo da compiere in questo momento, per rispondere ai bisogni dei nostri utenti, sia quello di integrare – nel modo più trasparente possibile

– fonti informative eterogenee per provenienza e per struttura dei dati.

Negli ultimi anni Internet ha catturato l'attenzione di chi ha necessità di reperire informazione; tuttavia, mentre è noto che le banche dati rimangono le fonti di informazione per eccellenza grazie alla loro esaustività, al loro regolare aggiornamento e alla struttura dei dati in esse contenuti, è anche vero che alcuni siti web contengono informazioni di grande interesse per chi opera in un certo settore disciplinare.

I² plus nasce con l'intento di eliminare il *gap* fra le diverse fonti d'informazione; l'integrazione dei due strumenti o, meglio, delle due fonti, che diventano ricercabili con un unico strumento, è sicuramente la risposta per le varie esigenze informative.

Poiché è noto che cercare informazione su Internet può essere poco efficace a causa della sua vastità e della genericità dei dati in essa presenti, ci preme sottolineare che i siti web integrati al servizio banche dati, e quindi ricercabili contemporaneamente grazie ad **I² plus**, fanno parte di una collezione, cioè vengono selezionati ed aggiunti su segnalazione dei nostri partner o su richiesta degli utenti.

Questa restrizione consente di ottenere anche dal web risultati di alta qualità, evitando risposte sovrabbondanti.

- *Come agisce I² plus*

I² plus, come s'è accennato in precedenza, è un servizio completamente integrato nel sistema EINS e di conseguenza ne sfrutta tutte le potenzialità.

Tutti gli utenti registrati ai servizi EINS-GEM a breve avranno a disposizione questo strumento, incluso nel servizio.

La *query* formulata dall'utente sulle banche dati EINS-GEM viene tradotta, automaticamente ed in parallelo, in un linguaggio comprensibile al database di **I² plus**; egli riceve quindi i risultati della banca dati interrogata ed ottiene simultaneamente i risultati dai siti Web facenti parte della collezione, e può decidere di visualizzare gli uni e/o gli altri a sua discrezione.



fig.2 – Risultati dalle banche dati e da I² plus

Occorre tuttavia rimarcare adeguatamente la natura tanto diversa dell'informazione presente rispettivamente sulle banche dati e su web: nel primo caso si parla infatti di informazione strutturata, in cui la ricerca può essere effettuata su campi precisi e secondo regole definite, nel secondo di informazione non strutturata.

Attualmente il motore di ricerca I² plus traduce la *query* fatta su un campo in una *query* sull'intero testo del documento disponibile su web. In futuro, con l'utilizzo sempre più massiccio del linguaggio XML, è prevedibile che le fonti informative raggiungeranno una sempre maggiore somiglianza nella loro struttura.

Il prototipo di I² plus attualmente realizzato riguarda il settore dell'aerospazio, ma in un prossimo futuro lo stesso servizio verrà esteso ad altri ambiti.

I siti raccolti nella collezione sono i più rilevanti del settore disciplinare cui fa riferimento, ed altri siti possono essere aggiunti anche su richiesta dell'utente, il quale può inoltre definire una propria collezione, personalizzata sulle proprie esigenze.