

LEAF. Collegare ed esplorare gli *authority file*

Jutta Weber
Staatsbibliothek zu Berlin - Preußischer Kulturbesitz

Il progetto LEAF [1] è finanziato dall'Unione Europea all'interno del quinto programma quadro (programma IST), ha avuto inizio nel marzo 2001 e durerà tre anni.

L'ambito di LEAF riflette il nuovo ruolo che gli *authority data* rivestono nelle biblioteche, negli archivi e nei musei: il loro uso interno al processo di catalogazione, il controllo di autorità che è necessario quando gli *authority data* provenienti da differenti fonti vengono usati in un procedimento di ricupero distribuito e di controllo dei punti d'accesso, che di fatto dipende dalla qualità degli *authority data*. Tutto ciò può essere sintetizzato nella domanda seguente: come possono milioni di *authority data* esistenti e provenienti da fonti molto diverse tra loro essere usati da chiunque (un bibliotecario, un archivista, un museologo, un utente del pubblico) in modo che non ci sia alcuna perdita d'informazione, alcuna accumulazione di dati senza controllo di qualità, ma un collegamento automatico tra informazione di alta qualità sulla spinta dei bisogni effettivi dell'utente?

LEAF tenterà di valorizzare i metodi di ricerca e di ricupero fornendo a chiunque accesso di qualità elevata all'informazione d'autorità. A questo scopo LEAF sta sviluppando un modello di architettura per collezionare, raccogliere, collegare e fornire accesso all'informazione esistente relativa agli *authority record* esistenti per i nomi, indipendentemente dal fatto che la loro creazione sia avvenuta nelle biblioteche, negli archivi o nei musei e senza dipendere dalle differenze nazionali. Lo *scenario* verrà costruito utilizzando *authority file* relativi a nomi di persona, e sta integrando direttamente l'utente nella costituzione di un Central European Name Authority File.

Un modello dimostrativo LEAF sarà costruito e integrato nel motore di ricerca MALVINE [2].

Il consorzio LEAF consiste di:

- Coordinatore: Staatsbibliothek zu Berlin, Berlin,
- Manager: Crossnet Systems Ltd., Newbury ,
- Sviluppatore di sistema: JOANNEUM RESEARCH, Graz,

e Biblioteca Nacional de Portugal, Lisbon; Biblioteca de Universidad Complutense, Madrid; British Library, London; Deutsches Literaturarchiv, Marbach; Forschungsstelle und Dokumentationszentrum für Österreichische Philosophie, Graz; Goethe- und Schiller-Archiv, Weimar; Institut Mémoires de l'édition contemporaine, Paris; Österreichische Nationalbibliothek, Vienna; University of Bergen, Bergen; Swiss National Library, Bern; National and University Library, Ljubljana and Riksarkivet, Stockholm.

Vi è poi un elevato numero di partner esterni che non ottengono finanziamenti ma che desiderano seguire il progetto in quanto istituzioni che operano nel medesimo ambito e sono interessate nel condividere anche i risultati preliminari fornendo, se necessario, consigli: Archives de France pour les technologies de l'information et de la communication, Paris; Det Arnemagnæanske Institut, Copenhagen; Biblioteca Nacional de España, Madrid; Biblioteca nazionale centrale di Roma; Jagiellonian Library, Krakow; Warsaw Public Library, Warsaw; Torun University Library, Torun; Warsaw University Library, Warsaw; Bibliothèque nationale de France, Paris; CIMI Consortium, Chicago; Constantijn Huygens Instituut voor Tekstedities en Intellectuele Geschiedenis, 's-Gravenhage; Dansk Biblioteks Center, Copenhagen; Die Deutsche Bibliothek, Frankfurt am Main; Dokimas Group Holdings Ltd., Nottingham; EKT, (National Documentation Centre), Athens; Franz-Michael-Felder-Archiv der Vorarlberger Landesbibliothek, Bregenz; GTAA, Grupo de Trabajo de Autoridades de Aragón, Zaragoza; Interparty, Boston Spa, Wetherby; Istituto centrale per il catalogo unico delle biblioteche italiane e per le informazioni bibliografiche, (ICCU); Jewish National & University Library, Jerusalem; Det Kongelige Bibliotek, Copenhagen; Koninklijke Bibliotheek, The Hague; Library of Congress, Washington; Magyar Tudományos Akadémia, Budapest; Onderzoekssteunpunt en Databank Intermediaire Structuren in Vlaanderen; Online Computer Library center (OCLC), Dublin/Ohio; Research Biblioteche Group (RLG), Mountain View/California; Tartu University, Tartu; University of Virginia, Charlottesville/Virginia; Wiener Stadt- und Landesbibliothek, Vienna e...

Vi sono anche partner finanziatori che si sono accordati per procurare dati per il test quando sarà richiesto: K.G. Saur Verlag, Munich; J.A. Stargardt, manuscript dealer, Berlin; Library of Congress, Washington.

Il motore di ricerca MALVINE realizza la raccolta di database che forniscono informazioni relative a lettere scritte da persone famose e conservate in istituzioni europee molto differenti. A causa della mancanza di una migliore informazione fornita dalle istituzioni partecipanti si può effettuare una ricerca per nome, non per individuo. Gli studiosi che operano nel settore delle risorse manoscritte del nostro patrimonio culturale e che svolgono ricerche in MALVINE sono contenti dei risultati forniti loro nella misura in cui i nomi che stanno utilizzando nella ricerca sono usati per una singola persona. Quando due o più persone usano lo stesso nome gli studiosi desiderano distinguerli. Questo termine di differenziazione può consistere di differenti date di nascita, di una diversa professione, ecc., in altre parole tutta quell'informazione che gli *authority file* nazionali tentano di fornire. Il problema è che le piccole istituzioni non hanno normalmente accesso a questi *authority file* nazionali. Così non sarà sorprendente il fatto che solo pochi partecipanti di MALVINE sono in grado di fornire questa specie di informazione d'autorità.

Il modello LEAF sarà basato su un'idea molto semplice: i dati provenienti da diversi fornitori sono immagazzinati in un *server* centrale, che verrà caricato di dati regolarmente oppure un meccanismo di raccolta aggiungerà nuovi dati ogni volta un utente eseguirà una ricerca in LEAF. Tutti i dati saranno collegati automaticamente, la qualità di tali collegamenti verificata nel momento in cui una ricerca è sottoposta al sistema. Questa verifica può essere svolta dall'utente stesso oppure dal fornitore dell'informazione che sarà automaticamente informato con una nota. Sono previste le seguenti caratteristiche:¹

1. Caricare le *authorities* distribuite in un sistema centrale.

Authority record locali (LAR) saranno caricati dai *server* locali delle organizzazioni partecipanti nel sistema centrale LEAF dove sono immagazzinati nel formato EAC (Encoded Archival Context), attualmente emergente, come *authority file* locali (LAF). Un aggiornamento a intervalli regolari dei dati che sono stati caricati assicurerà che il dato nel sistema centrale LEAF è per quanto possibile aggiornato.

2. Collegare le *authorities* che si riferiscono a una stessa entità. Con l'aiuto di regole di collegamento automatico definite all'interno del progetto gli *authority record* che si riferiscono a una medesima entità sono collegati tra loro. Naturalmente sarà possibile controllare questi legami generati automaticamente e se necessario annullarli manualmente. Ogniquale volta un utente interroga il sistema LEAF utilizzando una sequenza di caratteri che costituisce un nome come argomento della ricerca, questa sequenza rappresenterà così un'entità o, nel linguaggio di LEAF, vari *authority record* locali che rappresentano la stessa entità saranno aggregati per formare uno *shared name authority record* (SLAR), essi stessi venendo a costituire il *shared name authority file* (SLAF). È decisivo notare che gli *authority record* locali per i nomi non saranno fusi in un definitivo "*corporate*" *record*, ma aggregati o collegati in riconoscimento del fatto che, malgrado ogni differenza locale che possa sussistere, si riferiscono alla stessa entità. In tal modo, il mantenere le tradizioni d'autorità locali (che ha molti vantaggi pratici) può essere considerato compatibile con un desiderio di maggiore accuratezza e coerenza per l'utente finale.

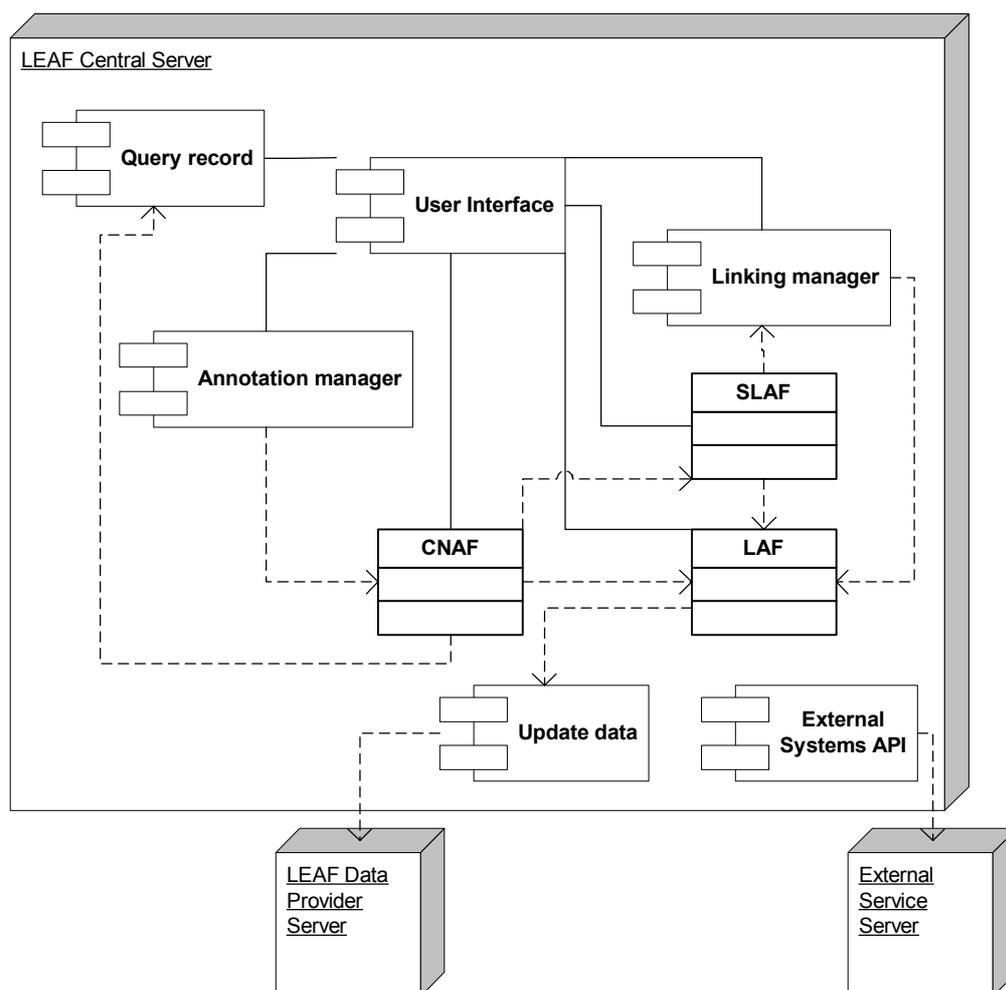
3. Corredare di note le *authorities* nell'intento di migliorare il contenuto e fornire informazioni aggiuntive. Ogni utente registrato del sistema LEAF sarà in grado di aggiungere annotazioni a particolari record di dati nel sistema stesso. Questa funzionalità è orientata principalmente al miglioramento degli *authority record* locali e ci si aspetta che richieda una qualche negoziazione tra l'utente che produce la nota e il proprietario del record in questione. LEAF fornirà un contesto in cui svolgere facilmente questi processi di negoziazione. Oltre a ciò sarà possibile allegare informazioni addizionali a un *data record* specifico, per es. istituzioni di piccole dimensioni senza un'offerta di dati in formato elettronico loro propria possono così informare gli utenti di LEAF che manoscritti collegati a un'entità specifica sono conservati in quella particolare istituzione. Inoltre chi commercia nel settore dei manoscritti può informare che i manoscritti di un certo autore sono in vendita, ecc.

4. Sostenere servizi esterni. Le applicazioni Internet esistenti potrebbero, in molti casi, chiaramente trarre beneficio dall'integrazione dell'informazione di autorità. Poiché i nomi rappresentano il più comune punto di accesso a database e reti bibliografiche, il recupero in linea sarà migliorato in considerevole misura dal collegamento degli *authority name record* ai record bibliografici. Per dimostrarlo, LEAF sarà integrato nel servizio esistente MALVINE.

5. Salvare i risultati della ricerca in un Central Name Authority File (CNAF) pan-europeo. L'informazione recuperata come risultato di una ricerca sottoposta a LEAF sarà conservata in un Central Name Authority File pan-europeo. Poiché ogni nuova ricerca genererà un nuovo record da conservare, questo Central Name Authority File crescerà a ogni ricerca e allo stesso tempo rifletterà esattamente a quale tipo di record gli utenti di LEAF sono interessati. Le biblioteche e gli archivi che intendono migliorare l'informazione d'autorità saranno così in grado di realizzare un elenco di priorità per il loro lavoro di *editing*.

Il seguente diagramma dell'architettura del sistema LEAF, semplificato, illustra i componenti principali del sistema stesso:

¹ Il testo seguente è basato, con l'aggiornamento di alcuni passi, sull'articolo di Hans-Jörg Lieder [1].

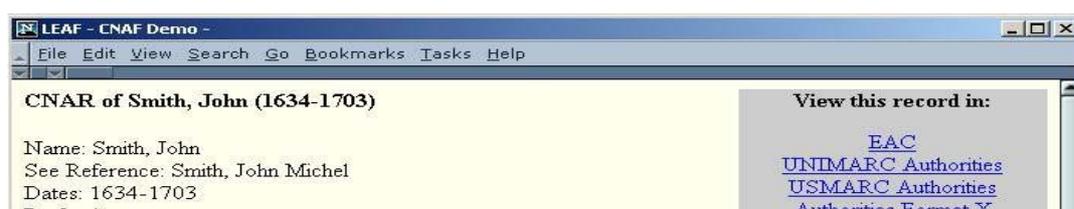


Il *manager* del caricamento dati trasferirà i dati locali dai *server dei data provider* di LEAF al database centrale LEAF EAC dove i record saranno conservati nel formato EAC. Il *manager* dei collegamenti svilupperà il collegamento automatico dei record che si riferiscono alla stessa entità, il *manager* delle note si occuperà dell'elaborazione e dell'amministrazione delle note ai record che sono state segnalate dagli utenti di LEAF. Tramite un'interfaccia utente ogni tipo di utente, compresi i fornitori di servizi esterni, sarà in grado di interagire con il database centrale LEAF. Un esempio di un'operazione fittizia di ricerca può illustrare le principali funzionalità:

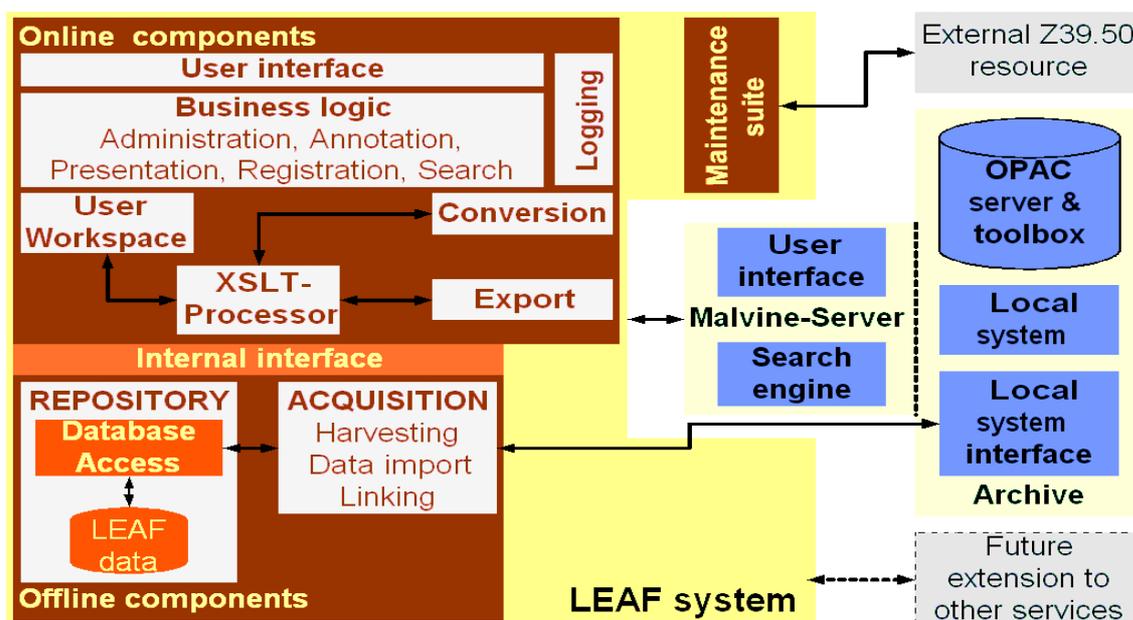
un utente svolge una ricerca usando "Smith, John". I server locali dei fornitori di dati LEAF contengono un numero di *authority record* che si riferiscono a "Smith, John". Attraverso il *linking manager* questi record sono raggruppati in un modo simile:

Smith, John (1542-1598)
 Smith, John (1634-1703)
 Smith, John (1712-1788)
 Smith, John (fl. 1912)

Ogni voce di questo breve elenco sarà espandibile per visualizzare un Central Name Authority Record aggregato, che sarà conservato a livello centrale e che assomiglierà al seguente esempio:



Le conseguenze tecniche saranno molteplici; il modello dimostrativo LEAF verrà sviluppato nei prossimi mesi sulla base di quanto detto. Il seguente diagramma ne mostra le caratteristiche principali:



Si visiti il sito Web di LEAF [1], dove tra breve saranno disponibili informazioni più dettagliate sul modello dimostrativo e sulle verifiche sperimentali di LEAF.

Note bibliografiche

[1] LEAF, <<http://www.leaf-eu.org>>.

[2] MALVINE, <<http://www.malvine.org>>.

[1] Hans-Jörg Lieder. *LEAF in a nutshell*. "LEAF newsletter", [????].