

Digitalia

Anno VI, Numero 2 - **2011**

Rivista del digitale nei beni culturali

ICCU-ROMA

Interoperabilità e biblioteche digitali: Un prontuario tecnico-metodologico¹⁻²

Leonardo Candela - Donatella Castelli

Istituto di Scienza e Tecnologie dell'Informazione "A. Faedo" – Consiglio Nazionale delle Ricerche

La richiesta di biblioteche digitali sempre più ricche, capaci di soddisfare le esigenze che insorgono in vari scenari della nostra società dell'informazione, ha fatto emergere sempre più la necessità di costruire tali sistemi basandoli sul riuso e la condivisione dell'esistente. In questo scenario la problematica dell'interoperabilità svolge un ruolo chiave. Nonostante l'importanza di tale problematica e le molteplici iniziative volte a risolverla, si può affermare che i risultati ottenuti sono, in media, limitati. Tra le principali ragioni di questa lenta evoluzione si annoverano la mancanza di un approccio sistematico e la scarsa conoscenza di quanto prodotto nelle varie iniziative. Al fine di rimuovere queste ragioni di lenta evoluzione, il progetto DL.org ha promosso un'iniziativa internazionale con l'obiettivo di studiare il problema dell'interoperabilità nelle sue sfaccettature e produrre un insieme di buone pratiche e soluzioni esistenti all'annoso problema dell'interoperabilità nel dominio delle biblioteche digitali. Il risultato dell'attività svolta è stato presentato alla Commissione Europea in un rapporto ufficiale. Questo documento rappresenta una sintesi di tale rapporto.

Introduzione

Sempre più spesso viene richiesto che le biblioteche digitali siano capaci di sup-

portare gli approcci moderni al fare scienza.

Questi sono approcci sempre più interdisciplinari che prevedono l'uso di dati ed informazioni di natura eterogenea, spesso su larga scala, nonché di servizi innovativi per la loro analisi. Questa continua e pressante richiesta ha fatto emergere la necessità di sistemi di gestione di biblioteche digitali basati sulla "condivisione" e il "riuso" di risorse esistenti (collezioni di dati, archivi, servizi offerti da terzi). Per soddisfare queste esigenze è di fondamentale importanza che i sistemi siano interoperabili. Solo a tale condizione tali sistemi, o loro componenti, diventano capaci di scambiare informazioni e usare le informazioni scambiate in modo effettivo ed efficace.

Il poter disporre di sistemi interoperabili amplia lo spettro delle scelte possibili nella realizzazione di una biblioteca digitale e offre nuove prospettive a ricercatori, governi e cittadini appartenenti a domini disciplinari e organizzativi diversi. Benché l'importanza dell'interoperabilità sia ben nota e siano stati fatti molti tentativi di dare soluzione ai problemi che essa presenta, manca ancora un approccio sistematico e anche una conoscenza diffusa delle soluzioni che pure già esistono. Molto spesso le soluzioni esistenti restano confinate nell'ambito dei sistemi per i quali sono state pensate e questo produce duplicazione di iniziative e sforzi quando le stesse problematiche si manifestano in contesti diversi.

¹ Questo articolo costituisce la sintesi delle attività svolte dal progetto DL.org, del quale l'Istituto CNR-ISTI è stato coordinatore scientifico, tecnologico e finanziario. <http://www.dlorg.eu/index.php/experts/consortium>.

² Si ringrazia la Dott. Maria Bruna Baldacci (ISTI) per il fondamentale aiuto nella produzione di questo lavoro.

Documenti e discussioni

Il significato di interoperabilità è dipendente dal contesto in cui tale concetto è inserito ed abbraccia diversi livelli, in uno spettro multidimensionale che comprende aspetti *organizzativi, semantici e tecnologici*. Il progetto europeo DL.org³ ha investigato il concetto di interoperabilità nell'ambito delle biblioteche digitali rispetto all'intero spettro che abbraccia questi tre livelli, analizzandolo dal punto di vista di ciascuno dei domini caratterizzanti una biblioteca digitale, cioè dal punto di vista della interoperabilità del *contenuto informativo*, degli *utenti*, della *funzionalità*, delle *politiche di governo*, della *qualità del servizio*, nonché della loro *architettura*. DL.org è stato così il primo progetto a studiare il tema "interoperabilità" da una prospettiva globale. In questo suo sforzo ha coinvolto studiosi a livello internazionale esperti nelle diverse aree coinvolte nella realizzazione di una biblioteca digitale. Il risultato delle ricerche svolte è presentato in un rapporto inviato alla Commissione Europea⁴. Esso presenta un prontuario di "buone pratiche" e soluzioni esistenti per i principali problemi che si incontrano comunemente nello sviluppo di biblioteche digitali "interoperabili". Un aspetto-chiave del rapporto è la definizione di un modello di riferimento in cui inquadrare, in modo sistematico, sia i diversi problemi di interoperabilità, sia le soluzioni e gli approcci attualmente in uso o emergenti. In tale modello, le tematiche legate all'interoperabilità sono discusse dal punto di vista dei diversi domini caratterizzanti le biblioteche digitali. Le varie tematiche del rapporto sono organizzate in modo da facilitare la selezione e la valutazione delle soluzioni che vengono offerte, così da indirizzare i professionisti che affrontano problemi di inter-

operabilità nell'individuare le soluzioni più appropriate e compiere i passi necessari. Questo articolo presenta le tematiche del rapporto in forma abbreviata, organizzate come segue: la Sezione 2 presenta il modello di interoperabilità definito dal progetto DL.org; la Sezione 3 contiene una vasta raccolta di buone pratiche e soluzioni applicabili per risolvere problemi di interoperabilità nei diversi domini delle biblioteche digitali; la Sezione 4 conclude l'articolo con alcune brevi note.

1 Interoperabilità: il modello definito dal progetto DL.org

Una delle maggiori difficoltà che si incontrano nell'affrontare il tema dell'interoperabilità è dovuta alla mancanza di un approccio sistematico basato su un modello comune di biblioteca digitale, da un lato, e, dall'altro, alla mancanza di una definizione di interoperabilità accettata dalla comunità delle biblioteche digitali o da altre comunità costrette ad affrontare questo tipo di problema. Per supplire a questa mancanza, DL.org ha definito un modello capace di servire da guida verso l'analisi di scenari di interoperabilità, l'individuazione sistematica delle problematiche e la realizzazione di soluzioni che risolvano tali problematiche. Il modello ha preso l'avvio dalla definizione di interoperabilità data dall'IEEE⁵: "*capacità di due o più sistemi o componenti di scambiare informazioni e usare l'informazione scambiata*". In essa è messo in risalto che la capacità di interoperabilità tra due sistemi che si scambiano informazioni, l'uno nel ruolo di *fornitore* e l'altro nel ruolo di *consumatore*, dipende da due condizioni: (i) le due entità devono essere capaci di scambiarsi informazioni; (ii) il consumatore

³ <http://www.dlorg.eu/>.

⁴ George Athanasopoulos – Leonardo Candela – Donatella Castelli – Katerina El Raheb – Perla Innocenti – Yannis Ioannidis – Akrivi Katifori – Anna Nika – Seamus Ross – Alice Tani – Costantino Thanos – Eleni Toli – Giuseppina Vullo, *The Digital Library Technology and Methodology Cookbook*, DL.org Project Deliverable, 125 pp. <http://www.dlorg.eu/index.php/outcomes/dl-org-cookbook>.

⁵ A. Geraci, *IEEE Standard Computer Dictionary: Compilation of IEEE Standard Computer Glossaries*, IEEE Press, 1991.

Documenti e discussioni

deve essere capace di usare in modo effettivo l'informazione oggetto di scambio; cioè, il consumatore deve essere capace di svolgere il compito che si era proposto utilizzando l'informazione scambiata.

Usando tale definizione come base di partenza, il modello proposto da DL.org introduce i seguenti tre concetti:

1. *Scenario di interoperabilità*, cioè una caratterizzazione del contesto in cui si vuol ottenere interoperabilità che includa l'obiettivo dell'interoperabilità stessa;
2. *Problema di interoperabilità*, cioè l'aspetto che ostacola l'interoperabilità attesa in un 'scenario di interoperabilità';
3. *Soluzione per l'interoperabilità*, cioè l'approccio che rimuove un problema di interoperabilità al fine di realizzare l'interoperabilità attesa.

presenza di un canale di comunicazione e di un protocollo che regoli il suo funzionamento, cioè senza la presenza di un mezzo di comunicazione che renda possibile lo scambio di informazioni e senza la presenza di certe regole che governino il suo uso effettivo durante il passaggio di informazioni da una entità all'altra. Non ci può essere informazione senza una forma di rappresentazione, cioè, l'informazione è "trasportata" dalla rappresentazione o dipende da essa. La significatività delle informazioni dipende dalla *risorsa* e dal *compito* che caratterizza lo scenario, cioè, la risorsa dovrebbe soddisfare le esigenze del consumatore e il consumatore dovrebbe acquisire, su tale risorsa, le informazioni che gli sono necessarie per svolgere il suo compito (*precondizioni per il compito*).



Fig. 1 – Scenario di Interoperabilità

Uno scenario di interoperabilità è presente quando si manifestino le seguenti condizioni:

- Ci sono almeno due entità che devono cooperare all'interno di tale scenario. Una delle entità ricopre il ruolo di *fornitore* mentre l'altra ricopre il ruolo di *consumatore*.
- La cooperazione consiste nel fatto che un consumatore vuole sfruttare una certa *risorsa* – posseduta dal fornitore – per svolgere un certo *compito* – il lavoro che il consumatore intende svolgere facendo affidamento sulla risorsa posseduta da quella terza parte.
- Affinché lo scenario sia realizzabile è necessario che le due entità si scambino informazioni "significative". Nessuno scambio di informazioni può avvenire senza la

- Il funzionamento di ciascuna entità, sia essa il fornitore o il consumatore, dipende da aspetti *organizzativi*, *semantici* e *tecnici*. Gli aspetti organizzativi catturano le caratteristiche degli obiettivi e dei processi dell'istituzione in cui l'entità opera. Esempi di aspetti organizzativi sono le tipologie delle regole che governano l'uso delle risorse, il tipo di funzionalità messe a disposizione dei consumatori, la qualità del servizio che deve essere assicurata a una specifica funzionalità. Gli aspetti semantici catturano il significato delle risorse che vengono scambiate, oltre alle caratteristiche del resto delle informazioni scambiate attraverso il canale di comunicazione. Esempi di aspetti rilevanti sono il significato assegnato a certe regole, il significato assegnato a certi

Documenti e discussioni

parametri di qualità del servizio, il significato assegnato a un certo valore in un record di metadati. Gli aspetti tecnici catturano le caratteristiche della tecnologia che sostiene le operazioni di una entità e quelle del canale di comunicazione, insieme alle informazioni che vi vengono scambiate. Esempi di aspetti tecnici sono il sistema di gestione usato per implementare la biblioteca digitale, il protocollo usato per una certa funzione, il formato di codifica di un certo oggetto informativo. È importante notare che questi tre aspetti si influenzano reciprocamente, con una relazione dall'alto verso il basso, cioè, gli aspetti organizzativi fissano l'intero scenario, caratterizzando il suo scopo e il suo funzionamento generale; gli aspetti semantici definiscono il significato delle entità coinvolte nello scenario in accordo con gli aspetti organizzativi; gli aspetti tecnici riguardano l'implementazione degli aspetti organizzativi e semantici.

Un **problema di interoperabilità** si presenta quando non sono verificate le condizioni che sono necessarie affinché il consumatore possa eseguire il compito che si prefigge utilizzando la risorsa fornita dal produttore. In altre parole, il problema si presenta quando non sono compatibili gli aspetti tecnici, semantici e/o organizzativi che caratterizzano il fornitore e il consumatore in relazione alla risorsa e al compito in oggetto. Ad esempio, problemi di interoperabilità si presentano quando: il formato usato dal fornitore per rappresentare un oggetto informativo differisce da quello che il consumatore si aspetta per eseguire una elaborazione; l'interfaccia che il fornitore mette a disposizione per la funzione di accesso all'oggetto informativo è diversa da quella che il fornitore si aspetta di usare per ottenere contenuto informativo; la semantica della funzione di ricerca implementata dal fornitore è diversa da quella che il consumatore ha intenzione di usare per eseguire una ricerca su più

sistemi; le regole che governano l'uso degli oggetti informativi definite dal fornitore sono incompatibili con quelle che il consumatore si aspetta.

Una soluzione per l'interoperabilità è un approccio atto a rimuovere le differenze evidenziate dal problema di interoperabilità e realizzare lo scenario. Concettualmente, l'approccio si basa su una funzione di trasformazione che agisce su qualsiasi livello dell'interazione fra fornitore e consumatore – organizzativo, semantico o tecnico – al fine di rendere compatibili le caratteristiche del fornitore e le esigenze del consumatore agendo sulle une o sulle altre o su ambedue. Esempi di soluzioni per l'interoperabilità sono: la trasformazione e presentazione dei metadati degli oggetti informativi attraverso il protocollo ed il formato che il consumatore si aspetta; l'implementazione di un "client" che sia conforme con le specifiche dell'interfaccia di ricerca implementata dal fornitore; l'implementazione di regole che garantiscano che le operazioni di ricerca distribuita avvengano con la qualità del servizio pattuita.

1.1 Modelli di Interoperabilità

Le soluzioni ai problemi di interoperabilità nel dominio delle biblioteche digitali, qualsiasi sia lo scenario, si possono sempre ricondurre a due classi di base: "soluzioni basate su accordi tra le parti" e "soluzioni basate su servizi di mediazione". Nella pratica, scenari complessi possono richiedere la combinazione di soluzioni appartenenti ad entrambe le classi. In alcuni casi, approcci basati su accordi e approcci basati su servizi di mediazione si combinano gli uni con gli altri dando luogo ad approcci misti. Ad esempio questo avviene quando un servizio di mediazione implementa una parte delle sue funzionalità rispettando situazioni e regole stabilite da accordi.

Approcci basati su accordi: costituiscono la strada tradizionale e uno dei metodi più effi-

Documenti e discussioni

caci per raggiungere l'interoperabilità, che è ottenuta con l'accordo su un insieme di principi che rendono possibile un grado seppur limitato di omogeneità fra entità eterogenee. Appartengono a questa categoria gli *standard*. Il maggior inconveniente di tali approcci risiede nel fatto che sugli *standard* e, più in generale, sugli accordi molto difficilmente si raggiunge una intesa fra le diverse organizzazioni e le diverse biblioteche digitali, specie quando queste operano in contesti diversi. Le soluzioni perciò finiscono spesso per essere complicate combinazioni di politiche riflettenti gli interessi delle varie parti. Inoltre, per loro natura, queste soluzioni infrangono l'autonomia delle entità che le adottano.

Approcci basati su servizi di mediazione: sono stati proposti per dare soluzioni a scenari nei quali è necessario garantire un alto grado di autonomia alle entità partecipanti. Questi approcci consistono nell'isolare il "meccanismo" di interoperabilità e implementarlo attraverso componenti specificamente concepiti per collegare le varie entità dello scenario. Queste soluzioni sono state inizialmente proposte nel dominio dei sistemi informativi⁶, ma oggi sono usate in molteplici casi e realizzate in diversi modi.

La parte più importante di questa tipologia di approcci è rappresentata dalla "funzione di mediazione", cioè il meccanismo che essi implementano. L'uso più frequente di questi approcci è quello di fornire funzioni primarie per la trasformazione dei formati di dati e delle modalità di interazione. Nella maggior parte dei casi, lo sviluppo di una funzione di mediazione richiede molta perizia e molto tempo. Ad esempio, in uno scenario nel quale le entità non collaborino, è lo sviluppatore della funzione di mediazione che deve acquisire le conoscenze necessarie

e sviluppare il collegamento fra fornitore e consumatore. In altri casi lo sviluppo può essere semi-automatico. Ad esempio, in caso di scenari collaborativi, le entità interessate presentano i loro dati secondo certe regole e lo sviluppatore può collegarle in base alle caratteristiche dei loro dati. Diversamente dagli approcci basati su accordi, gli approcci basati su mediazione sono molto efficaci ai fini della garanzia dell'autonomia nelle entità coinvolte, tuttavia la loro efficacia dipende dalla dinamicità delle parti interessate: ogni cambiamento nelle entità che devono interagire implica conseguentemente un cambiamento nel meccanismo di mediazione che assicura l'interoperabilità.

Approcci misti: le due classi di approcci appena descritte non sono mutuamente esclusive, anzi, in scenari concreti possono combinarsi per dar luogo ad approcci misti. La necessità di tale combinazione nasce in relazione alle peculiarità di ciascun scenario o entità interessata. Così può succedere che l'interoperabilità non sia raggiungibile attraverso accordi o standard, e ad essi si debba aggiungere uno specifico meccanismo di interoperabilità implementato da una funzione di mediazione, o che una funzione di mediazione si basi su uno o più standard per regolare l'interazione con il fornitore o con il consumatore.

2 Interoperabilità organizzativa, semantica e tecnica: soluzioni e buone pratiche

Esiste una vasta gamma di buone pratiche e di soluzioni che si sono dimostrate efficaci per risolvere i problemi di interoperabilità che si presentano nei diversi domini delle biblioteche digitali. Di queste viene presentata, nel seguito, una scelta significativa.

⁶ G. Wiederhold e M. Genesereth, *The Conceptual Basis for Mediation Services*, "IEEE Expert: Intelligent Systems and Their Applications", vol. 12 n.5, 1997, p. 38-47.

2.1 Interoperabilità relativa al dominio “contenuto informativo”: buone pratiche e soluzioni

I problemi di interoperabilità legati al contenuto informativo sorgono quando due o più sistemi di biblioteche digitali (o loro componenti) intendono interagire sfruttando l'una le risorse informative dell'altra. Tra i problemi più significativi che si presentano in questo tipo di interoperabilità si possono individuare: (i) come il fornitore può esporre la caratterizzazione degli oggetti informativi (pubblicazione di descrizioni di oggetti informativi) per rendere possibile ai consumatori l'implementazione dei servizi richiesti; (ii) come stabilire un'intesa sulla caratterizzazione degli oggetti informativi basandosi su accordi (standard per la rappresentazione di oggetti informativi e metadati); (iii) come raggiungere un'intesa sugli schemi per la caratterizzazione degli oggetti informativi basandosi su accordi (profili applicativi); (iv) come mediare tra diverse caratterizzazioni degli oggetti informativi (*mapping/cross-walk*), e (v) come associare identificatori univoci e persistenti a risorse di sistemi diversi (identificatori di risorse).

Gli approcci relativi all'esposizione di una caratterizzazione degli oggetti informativi che ne permetta il consumo da sistemi terzi sono basati su meccanismi per la pubblicazione di “metadati” degli oggetti informativi stessi. Sistemi terzi possono realizzare servizi sugli oggetti informativi basandosi sulle informazioni disponibili in questi metadati. Esempi concreti di tale tipologia di soluzione sono: (i) il protocollo OAI-PMH⁷, un approccio che permette ai *fornitori* di esporre, e ai *consumatori*

di raccogliere record di metadati (le *risorse*) in vari formati, tra cui il Dublin Core, in virtù della condivisione delle nozioni di “elemento appartenente a un archivio (o deposito)”, “record di metadati” e “formato di metadati”; (ii) OAI-ORE⁸, un approccio secondo cui il *fornitore* presenta sul Web oggetti informativi composti (*risorse*) in modo tale che essi possano essere acquisiti da qualsiasi *consumatore*. A tal fine *fornitore* e *consumatore* dovrebbero avere lo stesso intendimento sia delle entità che formano il modello ORE, cioè “Aggregazione”, “Mappa delle Risorse”, e “Proxy, sia delle loro relazioni; (iii) Linked Data⁹, un insieme di buone pratiche – basate su meccanismi e protocolli standard del Web – per pubblicare sul Web dati di vario genere mantenendo le relazioni tra essi esistenti. L'approccio Linked Data permette a *fornitori* e *consumatori* di condividere qualsiasi tipo di informazione riguardante una qualsiasi risorsa, ad esempio la sua struttura, i suoi metadati e le sue relazioni con altri risorse. *Fornitore* e *consumatore* dovrebbero condividere la conoscenza della semantica dei termini caratteristici del *framework* RDF¹⁰ come pure il significato dei termini di altri vocabolari; (iv) Open Data Protocol¹¹, un protocollo Web per la condivisione dei dati e la loro modifica che permette al *fornitore* di presentare ai *consumatori* qualsiasi tipo di dato, sia esso strutturato o no (la *risorsa*), ivi compresi i metadati e le possibili operazioni ad esso associate. *Fornitore* e *consumatore* dovrebbero condividere la conoscenza della semantica degli elementi che costituiscono il modello astratto dei dati (Entity Data Model).

⁷ Open Archives Initiative, *Open Archives Initiative - Protocol for Metadata Harvesting*, 2002. <http://www.openarchives.org/pmh/>.

⁸ Open Archives Initiative, *Open Archives Initiative Object Reuse and Exchange*, 2008. <http://www.openarchives.org/ore/>.

⁹ C. Bizer e T. Heath e T. Berners-Lee, *Linked Data - The Story So Far*, “International Journal on Semantic Web and Information Systems - Special Issue on Linked Data”, vol. 5, n.3, 2009. p. 1-22.

¹⁰ G. Klyne e J. Carroll, *Resource Description Framework (RDF): Concepts and Abstract Syntax*. <http://www.w3.org/TR/rdf-concepts/>.

¹¹ <http://www.odata.org>.

Documenti e discussioni

Gli approcci per stabilire un'intesa sulla caratterizzazione degli oggetti informativi sono basati su accordi (standard) per la loro rappresentazione. Esempi concreti di tale tipo di soluzione sono: (i) *Dublin Core*¹², uno standard per metadati basato su un insieme ristretto di elementi usabile per la descrizione di qualsiasi tipo di risorsa; con esso, *fornitore* e *consumatore* possono condividere i metadati di qualsiasi risorsa descritta in accordo con lo schema; (ii) il modello *EDM (Europeana Data Model)*¹³, un modello di dati proposto per la strutturazione degli oggetti digitali e dei metadati gestiti da "Europeana", la "Biblioteca Digitale Europea". Usando EDM, il sistema Europeana riesce a gestire in modo integrato i diversi modelli usati per il trattamento dei dati "culturali", mantenendo la loro forma originale ma assicurandone contemporaneamente l'interoperabilità semantica. In questo contesto, il *fornitore* espone oggetti digitali e relativi metadati (*le risorse*) in modo conforme a EDM, e questi oggetti possono essere utilizzati da qualsiasi *consumatore* capace di conformarsi a tale modello; (iii) *CERIF (Common European Research Information Format)*¹⁴, un modello formale pensato per rendere interoperabili i sistemi informativi progettati per il trattamento dei dati relativi alle attività di ricerca. In questo contesto, il *fornitore* espone

oggetti digitali e relativi metadati (*le risorse*) in modo conforme a CERIF, e questi oggetti possono essere utilizzati da qualsiasi *consumatore* capace di conformarsi a tale modello.

Gli approcci relativi ai profili applicativi¹⁵ sono basati su uno schema di metadati che lo sviluppatore crea combinando elementi descrittivi tratti da uno o più schemi esistenti mantenendo esplicitamente il riferimento allo schema sorgente. I profili applicativi adattano o combinano schemi pre-esistenti per confezionare schemi adatti a soddisfare i requisiti funzionali di una particolare applicazione, mantenendo l'interoperabilità con gli schemi originari. In questo contesto, un *fornitore* espone metadati di generiche risorse in conformità con un profilo applicativo concordato. Tali metadati possono essere utilizzati da qualsiasi *consumatore* che rispetti tale profilo applicativo. Profili applicativi ampiamente noti sono: *Scholarly Works Application Profile (SWAP)*¹⁶, *Dublin Core Collections Application Profile*¹⁷, *DC-Library Application Profile*¹⁸, *AGRIS Application Profile*¹⁹, *Biological Data Profile*²⁰, *Darwin Core*²¹, *DCMI Government Application Profile (DC-Gov)*²².

Gli approcci relativi a *mapping/cross-walk*^{23,24} consistono nel definire una corrispondenza fra gli elementi, la semantica e la sintassi di uno schema di metadati ed elementi, semantica e

¹² <http://dublincore.org>.

¹³ A. Isaac (Ed.), *Europeana Data Model Primer*, 2010.
<http://www.europeanconnect.eu/news.php?area=News&pag=48>.

¹⁴ Keith G. Jeffery, *CERIF – the Common European Research Information Format*, "ERICIM News", n. 68, 2007.

¹⁵ Rachel Heery, Manjula Patel, *Application profiles: mixing and matching metadata schemas*, "Ariadne", n. 25, 2000.

¹⁶ Julie Allinson, Pete Johnston, Andy Powell, *A Dublin Core Application Profile for scholarly works*, "Ariadne", n. 50, January 2007.

¹⁷ <http://dublincore.org/groups/collections/collection-application-profile>.

¹⁸ <http://dublincore.org/documents/library-application-profile>.

¹⁹ <http://www.fao.org/docrep/008/ae909e/ae909e00.htm>.

²⁰ <http://www.fgdc.gov/standards/projects/FGDC-standards-projects/metadata/biometadata>.

²¹ <http://rs.tdwg.org/dwc/>.

²² <http://www.dublincore.org/dcgapwiki>.

²³ NISO (National Information Standards Organization), *Understanding Metadata*, NISO, 2004.

²⁴ M. Zeng e L. Xiao, *Mapping metadata elements of different format*, in: *E-Libraries 2001, Proceedings*, May 15-17, 2001, p. 91-99. New York: Information Today, Inc.

sintassi di un altro schema. I *cross-walk* sono comunemente usati per rendere interoperabili schemi di metadati diversi. Esiste una grande quantità di *mapping* fra schemi diversi (ad esempio da MARC21 a Dublin Core) e molte altri ne arriveranno. Attualmente, i *cross-walk* sono l'approccio di gran lunga più usato, attraverso diverse modalità, per rendere possibile l'interoperabilità fra schemi di metadati. Ad esempio, con l'uso di uno schema di riferimento o di una "lingua franca" è sufficiente che, dato uno schema di metadati, si definisca una sola trasformazione tra lo schema sorgente e lo schema di riferimento o lingua franca. Basandosi su queste trasformazioni è possibile realizzare catene di trasformazioni senza la necessità di dover definire una trasformazione esplicita tra due schemi di metadati. Oltre a ciò, *mapping* e *cross-walk* possono essere usati per trasformare oggetti informativi cambiando il loro formato nativo in altri formati, ad es. OAI-ORE.

Approcci relativi agli identificatori di oggetti informativi (*risorse*) sono normalmente basati su accordi riguardanti il modo di riferirsi alle risorse. Esempi di questa tipologia di soluzioni per l'interoperabilità includono: (i) *Uniform Resource Identifier (URI)*, che offre a *fornitore* e *consumatore* un comune strumento per identificare qualsiasi tipo di *risorsa*, sia astratta che fisica; (ii) *Handle System*²⁵, secondo il quale è il *fornitore* stesso a fungere da *Handle System* implementandone il relativo protocollo, il *consumatore* è qualsiasi *client* che desidera identificare la risorsa Web associata a un identificatore "Handle", mentre la *risorsa* condivisa è un identificatore persistente – *handle* – associato a qualsiasi tipo di risorsa Web. La maggiore e migliore implementazione dell'*Handle System* è quella dell'International DOI Foundation (IDF)²⁶.

²⁵ <http://www.handle.net/>.

²⁶ <http://www.doi.org/index.html>.

²⁷ D. Heckmann, T. Schwartz, B. Brandherm, M. Schmitz, M. Wilamowitz-Moellendorff, *GUMO - the general user model ontology*, in: *Proceedings of the 10th International Conference on User Modeling*, Edinburgh, Scotland, 2005.

2.2 Interoperabilità relativa al dominio "utenti": buone pratiche e soluzioni

Per interoperabilità delle biblioteche digitali nel contesto del dominio degli utenti si intende la capacità di due o più biblioteche digitali (o parti di essi) di scambiare informazioni sullo stesso utente e di usare tali informazioni in modo significativo e accurato al fine di produrre risultati definiti "utili" dagli utenti di tali sistemi. Nel dominio "utenti" l'interoperabilità riguarda: (i) la rappresentazione condivisa di modelli di utente (formato condiviso di modelli di utente); (ii) la trasformazione di modelli e profili di utenti (conversione di modelli e profili di utente); e (iii) l'autenticazione e autorizzazione fra sistemi diversi (autenticazione e autorizzazione).

L'approccio relativo alla condivisione del formato dei modelli di utente impone la condivisione della sintassi e della semantica usate per rappresentare i modelli di utente. La comunità interessata ai modelli di utente ha recentemente posto l'accento su approcci basati su ontologie. Ad esempio, GUMO (General User Model Ontology)²⁷ è una ontologia costruita mediante il linguaggio OWL. La costruzione di GUMO è basata sul concetto di dividere le descrizioni delle caratteristiche dei modelli di utente in tre elementi ("auxiliary", "predicate" e "range"). La descrizione risultante è chiamata "dichiarazione situazionale" ("situational statement"). In questo contesto sono coinvolti due tipi di attori: il servizio di gestione del modello di utente (il *fornitore*) e le applicazioni (il *consumatore*). La *risorsa* che le due entità intendono scambiarsi è la dichiarazione situazionale. Il servizio di gestione del modello di utente è erogato da un server, indipendente dall'applicazione, attraverso cui le informazioni sull'utente vengono accedute, me-

Documenti e discussioni

morizzate e scambiate tra le diverse applicazioni. Un'applicazione può aggiungere o richiedere le informazioni memorizzate in tale server.

Gli approcci relativi alla conversione dei modelli e profili di utente impiegano metodi adatti a trasformare la sintassi e la semantica del modello di utente usato in un certo sistema in quelle usate da un sistema diverso. L'approccio GUC (Generic User model Component)²⁸ usa tecnologie del Web semantico per fornire funzionalità di memorizzazione e condivisione di modelli di utente. Nell'architettura che esso propone sono presenti due tipi di attori: GUC (il *fornitore*) e le applicazioni che necessitano di usare il modello di utente (il *consumatore*). GUC è un componente generico che offre funzionalità per memorizzare modelli di schemi per le applicazioni e per scambiare informazioni sull'utente (profilo di utente) tra questi modelli. Ogni applicazione che usa un modello di utente è un'entità esterna a GUC che usa GUC per memorizzare informazioni sull'utente stesso. Se viene memorizzato in GUC uno schema di applicazione, l'applicazione può trasmettere profili utente conformi a tale schema, chiamati "user application-view"(UAV). Per ciascun utente, GUC mantiene un archivio dove memorizza una UAV per ciascuna applicazione usata dall'utente.

Approcci relativi ad Autorizzazione e Autenticazione sono basati preferibilmente sulla nozione di "identità federata". Due approcci molto importanti ai fini dell'interoperabilità sono: (i) *OpenID*²⁹ – nel suo contesto sono coinvolti due tipologie di attori: *OpenID Providers* (il *fornitore*) e *Relying Parties* (il *consumatore*). Un *OpenID Provider* è un server di autenticazione a cui le *Relying Parties* si affidano per accertare che un utente abbia

il controllo di un identificatore (la *risorsa*). Una *Relying Party* è un'applicazione Web che richiede l'accertamento che un certo utente abbia il controllo di un identificatore; e (ii) *Security Assertion Markup Language* (SAML)³⁰ – questo standard definisce un ambiente basato su XML per descrivere e scambiare informazioni relative alla sicurezza tra entità coinvolte in una cooperazione. Tali informazioni sono specificate in forma di asserzioni conformi allo standard SAML su cui possono fare affidamento le diverse applicazioni che operano in domini diversi. In questo contesto sono coinvolte due specie di attori: il *SAML Asserting Party* (il *fornitore*) e il *SAML Relying Party* (il *consumatore*). Un *Asserting Party* è una entità del sistema che fa asserzioni. Un *Relying Party* è una entità del sistema che riceve asserzioni (la *risorsa*). Al cuore della maggior parte delle asserzioni SAML si trova un soggetto su cui si asserisce qualcosa (essendo tale soggetto un'entità che può essere autenticata nel contesto di un particolare dominio di sicurezza). Il soggetto può essere un umano o una entità di diversa natura, come una società o un computer.

2.3 Interoperabilità relativa al dominio "funzionalità": buone pratiche e soluzioni

Gli approcci all'interoperabilità delle funzioni possono essere suddivisi in tre classi principali: (i) approcci orientati a risolvere i problemi di interoperabilità a livello dell'interfaccia della funzione; (ii) approcci orientati a risolvere i problemi di interoperabilità a livello delle modalità di comportamento della funzione; (iii) approcci orientati a risolvere i problemi di interoperabilità al livello della funzione stessa. Esistono inoltre protocolli e standard che sono stati concepiti per catturare il funziona-

²⁸ Kees Van der Sluijs e Geert-Jan Houben, *Towards a generic user model component*, in: *Proceedings of the Workshop on Decentralized, Agent Based and Special Approaches to User Modelling*, in: *International Conference on User Modelling*, Edinburgh, Scotland, 2005, p. 43-52.

²⁹ <http://openid.net/>.

³⁰ <http://docs.oasis-open.org/security/saml/v2.0/>.

Documenti e discussioni

mento di una funzione o servizio.

Gli approcci relativi alla riconciliazione dell'interfaccia si basano su descrizioni della funzione (di norma semanticamente arricchite) per specificarne importanti proprietà. Esistono numerosi approcci, con diversi gradi di automazione, che possono essere suddivisi in (a) approcci basati su standard, e (b) approcci dinamici, basati sulla mediazione. I primi sono di natura prevalentemente statica e consistono nell'esporre la specifica di interfacce predefinite per certi tipi di servizi. Appartengono a questa categoria primitive per la specificazione delle interfacce delle funzioni (ad esempio WSDL³¹, SAWDL³², OWL-S³³, WSMO³⁴), RosettaNet³⁵, e e-Framework³⁶. Per quanto riguarda i secondi, essi sono essenzialmente basati sull'utilizzazione – parzialmente o completamente automatica – di “adapters” che possono essere forniti automaticamente o manualmente. Tutti questi approcci fanno normalmente uso di appropriate specifiche di funzioni o servizi. È tuttavia importante notare che essi sono il risultato di attività di ricerca non ancora sperimentate su un prodotto o su scala industriale. Yelling and Storm³⁷ propongono un approccio capace di facilitare l'interoperazione di componenti a livello di interfaccia e di protocollo. Basandosi sull'uso di

una semantica appropriata e di un modello ispirato alle macchine a stati finiti, questi autori definiscono meccanismi capaci di sintetizzare “adapters” di componenti in modo (semi)automatico. Benatallah et al.³⁸ presentano un approccio semi-automatico che fa leva su schemi definiti manualmente per risolvere incompatibilità comportamentali e fra interfacce. Differenze presenti fra servizi sono rilevate usando modelli di discrepanza, utili anche per analizzarle e risolverle. Bordeaux et al.³⁹ forniscono un approccio formale per valutare e conciliare la compatibilità di funzionalità (servizi) rispetto a interfacce e modalità di funzionamento. Essi usano il Pi-calculus per rappresentare formalmente le proprietà dei servizi in termini di “protocolli” e algoritmi di confronto per valutare l'interoperabilità tra i servizi stessi. Ponnekanti e Fox⁴⁰ hanno anche presentato un approccio che fa leva su strumenti di analisi statica e dinamica per valutare la possibilità di rimpiazzare certi servizi con altri.

Gli approcci relativi alle problematiche di differenze nel funzionamento di funzioni (“behaviour”) sono in realtà complementari a quelli appena descritti, cioè l'interoperabilità a livello di funzioni è strettamente correlata con le problematiche legate a interfaccia, ‘pre-’ e ‘post-’ condizioni e comportamento.

³¹ D. Booth e K. C. Liu, *Web Services Description Language (WSDL) Version 2.0 Part 0: Primer*, 2007. <http://www.w3.org/TR/wsdl20-primer>.

³² J. Farrell e H. Lausen, *Semantic Annotations for WSDL and XML Schema (W3C Recommendation)*, 2007. <http://www.w3.org/TR/sawSDL/>.

³³ D. Martin, M. Burstein, J. Hobbs, O. Lassila, D. McDermott, S. McIlraith - et al., *OWL-S: Semantic Markup for Web Services*, W3C, 2004. <http://www.w3.org/Submission/OWL-S/>.

³⁴ D. Fensel e C. Bussler, *The Web Service Modeling Framework WSMF*, “Electronic Commerce Research and Applications”, Vol. 1, n. 2, 2002.

³⁵ RosettaNet Community. (2010). RosettaNet Web site: <http://www.rosettanet.org>

³⁶ <http://www.e-framework.org/>.

³⁷ D. Yellin e R.E. Strom, *Protocol Specifications and Component Adaptors*, “ACM Transactions on Programming Languages and Systems (TOPLAS)”, vol. 19, n. 2, 1997, p. 292–333.

³⁸ B. Benatallah, F. Casati, D. Grigori, H. R. Nezhad, F. Toumani, *Developing Adaptors for Web Services Integration*, in: *CAISE 2005, The 17th Conference on Advanced Information Systems Engineering*, Porto, Portugal, 2005.

³⁹ L. Bordeaux, G. Salaun, D. Berardi, M. Mecella, *When are two Web Services Compatible?* in: *VLDB TES'04*, Toronto, Canada, 2004.

⁴⁰ S. R. Ponnekanti e A. Fox, *Interoperability among Independently Evolving Web Services*, in: *Middlewar'04*. Toronto, Canada, 2004.

Documenti e discussioni

Perciò, gli approcci sia statici che dinamici sopra accennati risolvono anche quanto riguarda gli aspetti del comportamento. Esistono inoltre standard per rappresentare il comportamento funzionale, ad esempio, WS-CDL⁴¹, WSCL⁴², WS-BPEL⁴³. Approcci basati sulla mediazione si avvalgono di meccanismi di specifica formale del comportamento delle funzioni, ad esempio modelli a stati finiti e lambda-calculus, e algoritmi appropriati. Questi includono: (i) meccanismi basati sull'intelligenza artificiale, che usano modelli di macchine a stati finiti e tecniche di pianificazione tipiche dell'intelligenza artificiale. Dalle descrizioni dei servizi, di norma arricchite semanticamente, sono estratti sistemi a transizione di stati e, in dipendenza dei vincoli introdotti da ciascun approccio, sono utilizzati algoritmi di pianificazione deterministici o non deterministici per far valere e fornire l'integrazione dei servizi; (ii) Deng et al.⁴⁴ fanno uso del Pi-calculus per modellare formalmente il comportamento e l'interazione dei servizi. Essi propongono anche un metodo basato sulla semantica operativa del Pi-calculus per rendere automatica la verifica della compatibilità fra due servizi, e un algoritmo per misurare quantitativamente il grado di compatibilità; (iii) Peng et al.⁴⁵ usano un modello di comportamento dei servizi basato sulle reti di Petri. Propongono inoltre un metodo formale per verificare e calcolare la compatibilità del com-

portamento dei servizi; (iv) Stollberg et al.⁴⁶ prendono spunto dal WSMO per creare un modello di mediazione capace di trattare e risolvere l'eterogeneità che può manifestarsi nel dominio dei servizi nel Web semantico. Gli approcci relativi alla modellazione delle condizioni di funzionamento sono basati sulla specificazione delle condizioni che devono verificarsi perché possa essere attivata una funzione (pre-condizioni) e di quelle che devono verificarsi dopo la sua esecuzione (post-condizioni). Meccanismi ed approcci che sono stati usati per descrivere tali condizioni includono: (i) primitive per la specifica delle pre-/post- condizioni di funzionamento. Tali condizioni permettono di attivare metodi e strumenti formali capaci di validare sistemi sia al momento del progetto sia al momento dell'esecuzione. Le asserzioni costituiscono l'approccio più comune all'implementazione di tali meccanismi in numerosi linguaggi di programmazione, ad esempio Eiffel, Java e C/C++; e (ii) approcci di riconciliazione delle pre-/post- condizioni di funzionamento. Approcci relativi alle pre-/post- condizioni sono sempre stati confrontati con temi relativi alle interfacce e alle modalità di funzionamento. Così gli approcci sopra accennati sono utilizzati per affrontare anche questo problema. Un loro requisito è l'uso di rappresentazioni formalizzate delle pre-/post-condizioni, come quelle usate da WSMO e OWL-S.

⁴¹ N. Kavantzaz, D. Burdett, G. Ritzinger, T. Fletcher, Y. Lafon, *Web Services Choreography Description Language Version 1.0. (W3C)*, 2004. <http://www.w3.org/TR/ws-cdl-10>.

⁴² A. Banerji, C. Bartolini, D. Beringer, V. Chopella, K. Govindarajan, A. Karp, et al., *Web Services Conversation Language (WSCL) 1.0. (W3C)*, 2002. <http://www.w3.org/TR/2002/NOTE-wscl10-20020314/>.

⁴³ A. Alves, A. Arkin, S. Askary, C. Barreto, B. Bloch, F. Curbera, et al., *Web Services Business Process Execution Language, Version 2.0. (OASIS)*, 2007. <http://docs.oasis-open.org/wsbpel/2.0/wsbpel-v2.0.html>.

⁴⁴ S. Deng, Z. Wu, M. Zhou, Y. Li, J. Wu, *Modeling Service Compatibility with Pi-calculus for Choreography*, in: *Proceedings of 25th International Conference on Conceptual Modeling*, Tucson, AZ, USA, 2006.

⁴⁵ Y. Peng, Z. Zheng, J. Xiang, J. Gao, J. Ai, Z. Lu, et al. *A Model of Service Behavior Based on Petri Nets with Weights*, in: *World Congress on Software Engineering*, 2009, p. 3-6.

⁴⁶ M. Stollberg, E. Cimpian, A. Mocan, D. Fensel, A. Semantic Web Mediation Architecture, in: *Proceedings of 1st Canadian Semantic Web Working Symposium (CSWWS)*, 2006.

2.4 Interoperabilità relativa al dominio “politiche di governo”: buone pratiche e soluzioni

Nel dominio delle “politiche di governo” l’interoperabilità è praticamente un terreno inesplorato, perciò per i suoi problemi sarebbe più appropriato parlare di “desideri che la comunità delle biblioteche digitali vorrebbe vedere realizzati” piuttosto che di “soluzioni”. Tuttavia, ai fini dell’interoperabilità delle politiche di governo sono stati identificati i seguenti gruppi di approcci: (i) interoperabilità delle politiche di accesso: XML; EML-Election Markup Language⁴⁷; METS⁴⁸; DOI⁴⁹; COUNTER

3 Code of Practice⁵⁰; OpenURL Framework Standard⁵¹; W3C WAI WCAG - Web Content Accessibility Guidelines⁵²; W3C Markup Validation Service⁵³; US Federal Government Section 508 Guidelines⁵⁴; DLF ERM Initiative⁵⁵; (ii) interoperabilità delle politiche di preservazione: PRONOM⁵⁶; DROID⁵⁷; JHOVE⁵⁸; UDFR⁵⁹; Global GDFR⁶⁰; Planets Testbed Beta⁶¹; OAIS⁶²; TRAC⁶³; DRAMBORA Interactive toolkit⁶⁴; LOCKSS⁶⁵; Portico’s Digital Preservation Service⁶⁶; EAD⁶⁷; METS⁶⁸; OAI-PMH⁶⁹; XML⁷⁰; PREMIS⁷¹; DIDL⁷²; DCMI⁷³; MARC⁷⁴; ONIX⁷⁵; (iii) Interoperabilità delle politiche di governo dei servizi: iRODS⁷⁶; WSDL⁷⁷; XACML⁷⁸; (iv) intero-

⁴⁷ OASIS, *OASIS Election Markup Language (EML). Version 5.0 Process and Data Requirements*, OASIS, 2007. <http://docs.oasis-open.org/election/eml/v5.0/os/EML-Process-Data-Requirements-v5.0.htm>.

⁴⁸ Library of Congress. (n.d.). Metadata Encoding and Transmission Standard (METS) Official Web Site. <http://www.loc.gov/standards/mets/>.

⁴⁹ <http://www.doi.org/>.

⁵⁰ http://www.projectcounter.org/code_practice.html.

⁵¹ http://www.niso.org/kst/reports/standards?step=2&project_key=d5320409c5160be4697dc046613f71b9a773cd9e.

⁵² <http://www.w3.org/TR/WCAG20/>.

⁵³ <http://validator.w3.org/>.

⁵⁴ <http://www.section508.gov/>.

⁵⁵ <http://www.clir.org/dlf.html>.

⁵⁶ <http://www.nationalarchives.gov.uk/PRONOM/Default.aspx>.

⁵⁷ <http://droid.sourceforge.net/>.

⁵⁸ <http://hul.harvard.edu/jhove/>.

⁵⁹ <http://www.udfr.org/>.

⁶⁰ <http://www.gdfr.info/>.

⁶¹ <http://testbed.planets-project.eu/testbed/>.

⁶² <http://public.ccsds.org/publications/archive/650x0b1.pdf>.

⁶³ <http://www.crl.edu/archiving-preservation/digital-archives/metrics-assessing-and-certifying>.

⁶⁴ <http://www.repositoryaudit.eu/>.

⁶⁵ <http://lockss.stanford.edu/lockss/Home>.

⁶⁶ <http://www.portico.org/digital-preservation/>.

⁶⁷ <http://www.loc.gov/ead/>.

⁶⁸ <http://www.loc.gov/standards/mets/>.

⁶⁹ <http://www.openarchives.org/OAI/openarchivesprotocol.html>.

⁷⁰ <http://www.w3.org/XML/>.

⁷¹ <http://www.loc.gov/standards/premis/>.

⁷² <http://xml.coverpages.org/mpeg21-didl.html>.

⁷³ <http://dublincore.org/>.

⁷⁴ <http://www.loc.gov/marc/>.

⁷⁵ <http://www.editeur.org/15/Previous-Releases/>.

⁷⁶ https://www.irods.org/index.php/IRODS:Data_Grids,_Digital_Libraries,_Persistent_Archives,_and_Real-time_Data_Systems.

⁷⁷ <http://www.w3.org/TR/wsdl>.

⁷⁸ <http://xml.coverpages.org/nlmJournals.html>.

Documenti e discussioni

perabilità delle politiche per la proprietà intellettuale: METS⁷²; NLM XML DTDs for Journal Publishing, Archiving and Interchange⁷³; PREMIS⁸⁰; CC-Creative Commons licences⁸¹; (v) interoperabilità delle politiche di autenticazione: XACML⁸²; Shibboleth⁸³; Athens⁸⁴; (vi) interoperabilità delle politiche di valutazione e accertamento: DRIVER Guidelines 2.0 for content providers; SHAMAN Assessment Framework⁸⁵; (vii) interoperabilità delle politiche per la rappresentazione delle politiche stesse e della loro obbligatorietà: PLEDGE project⁸⁶; AIR Policy Language⁸⁷; iRODS rules⁸⁸; SWRL⁸⁹; Turtle RDF Triples⁹⁰; REVERSE Policy Language⁹¹; OWL⁹²; KAoS⁹³; Web Services Policy Framework (WS-Policy)⁹⁴; Web Services Policy 1.5⁹⁵; WSPL⁹⁶; XACML⁹⁷; Rei⁹⁸.

2.5 Interoperabilità relativa al dominio “qualità del servizio”: buone pratiche e soluzioni

Nelle biblioteche digitali uno dei maggiori ostacoli all’identificazione di soluzioni per

l’interoperabilità nel dominio “qualità del servizio” è che spesso la “qualità” non è formalmente descritta, ma implicata o “nascosta” dietro indicatori di eccellenza quali “rispetto degli standard,” “efficienza”, “performance”, etc. Approcci e buone pratiche sono organizzati in (i) interoperabilità riguardante la qualità dei dati, (ii) soluzioni identificate nel mondo del Web, e (iii) linee guida e liste di controllo.

Nel campo delle ricerche sulla qualità dei dati sono stati costruiti ambienti specifici di interoperabilità per sistemi informativi cooperativi, cioè sistemi informativi costruiti su larga scala per interconnettere sistemi operanti in diverse organizzazioni autonome che condividono obiettivi comuni. Questi ambienti includono: DaQuinCIS Framework⁹⁹ – che copre un’ampia fascia di problemi nell’area della valutazione, della correzione dei dati, dell’identificazione degli oggetti, dell’affidabilità delle fonti e dell’integrazione dei dati, offrendo un insieme di servizi orientati alla qualità; e Fusionplex

⁷⁹ <http://xml.coverpages.org/nlmJournals.html>.

⁸⁰ <http://www.loc.gov/standards/premis/>.

⁸¹ <http://creativecommons.org/>.

⁸² <http://xml.coverpages.org/nlmJournals.html>.

⁸³ <http://shibboleth.internet2.edu/>.

⁸⁴ <http://www.athens.ac.uk/>.

⁸⁵ <http://portal.acm.org/citation.cfm?id=1643823.1643899&coll=GUIDE&dl=GUIDE&CFID=63081623&CFTOKEN=61810568>.

⁸⁶ http://pledge.mit.edu/index.php/Main_Page#PLEDGE_Project_Wiki.

⁸⁷ <http://dig.csail.mit.edu/TAMI/2007/amord/air-specs.html>.

⁸⁸ <https://www.irods.org>.

⁸⁹ <http://www.w3.org/Submission/SWRL/>.

⁹⁰ <http://www.w3.org/TeamSubmission/turtle/>.

http://en.wikipedia.org/wiki/Resource_Description_Framework.

⁹¹ <http://reverse.net/>.

⁹² <http://www.w3.org/TR/owl-features/>.

⁹³ <http://www.w3.org/2004/08/ws-cc/kaos-20040904>.

⁹⁴ <http://www.ibm.com/developerworks/library/specification/ws-polfram/>.

<http://msdn.microsoft.com/library/en-us/dnglobspec/html/ws-policy.asp>.

⁹⁵ <http://www.w3.org/TR/2007/REC-ws-policy-20070904/>.

⁹⁶ <http://www-106.ibm.com/developerworks/library/ws-polas/>.

⁹⁷ <http://en.wikipedia.org/wiki/XACML>.

<http://www.oasis-open.org/committees/xacml>.

⁹⁸ <http://rei.umbc.edu/>.

⁹⁹ *DaQuinCIS: Methodologies and Tools for Data Quality inside Cooperative Information Systems*. <http://www.dis.uniroma1.it/dq/>.

Documenti e discussioni

Framework¹⁰⁰ – un sistema per integrare una molteplicità di fonti di informazione eterogenee che usa la fusione dei dati per risolvere inconsistenze fattuali fra le varie fonti. Questo sistema caratterizza ciascuna sorgente di dati e usa tali caratterizzazioni per risolvere i conflitti. Le soluzioni identificate nel mondo del Web sono di solito basate su ontologie. Nei campi QoS (Quality of Service) e SLA (Service Level Agreement) sono state sviluppate numerose ontologie per servizi Web che definiscono i livelli che devono raggiungere, per essere accettabili, fattori quali disponibilità, accessibilità, integrità, prestazioni, etc.

Nel campo delle biblioteche digitali sono state prodotte numerose linee-guida, liste di controllo e metodi di certificazione aventi per scopo la soluzione di problemi di eterogeneità nella interoperazione. Lo stabilire regole comuni, l'uso di standard e buone pratiche può avere scopi diversi (integrazione dei dati, preservazione digitale di lungo termine, etc.) e diversi settori di applicazione (contenuto informativo, funzionalità, politiche di governo, etc.); tuttavia obiettivo generale di questi strumenti è facilitare la cooperazione e lo sviluppo all'interno di reti e infrastrutture di biblioteche digitali. Fra tali strumenti possiamo citare: (i) DRIVER Guidelines 2.0¹⁰¹ che costituiscono uno strumento per garantire che gli archivi che partecipano alla federazione DRIVER rispettino certi canoni, ad esempio rispetto al protocollo OAI-PMH, rispetto ai metadati da esporre. Tali linee-guida possono però anche dare ai responsabili di nuovi archivi un utile orientamento alla politica di ge-

stione dei loro dati locali, agli amministratori di archivi già esistenti l'indicazione dei passi da fare per migliorare i servizi, e, agli sviluppatori di piattaforme per l'archiviazione, l'indicazione delle funzionalità che è utile aggiungere nelle versioni successive; (ii) DINI Certificate¹⁰² – sviluppato nel contesto dell'iniziativa tedesca per l'informazione in rete (Deutsche Initiative für Netzwerk Information). Fra gli scopi che tale iniziativa si propone c'è quello di rendere evidente ad utenti ed operatori come un archivio debba osservare standard e raccomandazioni; (iii) Data Seal of Approved Guidelines¹⁰³ – un'attestazione di affidabilità stabilito da un certo numero di istituzioni interessate alla durata dei dati di ricerca memorizzati negli archivi. Con l'attribuzione di tale attestazione, queste istituzioni vogliono non solo assicurare la persistenza dei dati in questione, ma anche promuovere, in generale, il mantenimento degli archivi di dati nel lungo periodo; (iv) DRAMBORA (Digital Repository Audit Method Based on Risk Assessment)¹⁰⁴ – una metodologia di autovalutazione degli archivi digitali, che incoraggia le organizzazioni a raggiungere la piena consapevolezza dei loro obiettivi, delle loro attività e risorse prima di identificare, valutare e gestire i rischi impliciti nella loro organizzazione; (v) Trustworthy Repository Audit & Certification (TRAC)¹⁰⁵ – che pone le basi per una collaborazione internazionale per il controllo e la certificazione degli archivi digitali, promossa da DCC, RLG (ora OCLC-RLG), NARA, NESTOR, e US Center for Research Libraries.

¹⁰⁰ A. Motro e P. Anokhin, *Fusionplex: resolution of data inconsistencies in the integration of heterogeneous information sources*, "Information Fusion", vol. 7, n. 2, 2006, p. 176-196.

¹⁰¹ DRIVER Project, *DRIVER Guidelines 2.0*. DRIVER Project, 2008.

¹⁰² German Initiative for Networked Information. <http://www.dini.de/>.

¹⁰³ The Data Seal of Approval Organisation. <http://www.datasealofapproval.org/>.

¹⁰⁴ P. Innocenti, S. Ross, E. Maceciuvite, T. Wilson, J. Ludwig, W. Pempe, *Assessing Digital Preservation Frameworks: the approach of the SHAMAN project*, in: *Proceedings of the International Conference on Management of Emergent Digital EcoSystems*, 2009, p. 412-416.

¹⁰⁵ B. Ambacher, K. Ashley, J. Berry, C. Brooks, R. L. Dale, D. Flecker, et al., *Trustworthy Repositories Audit & Certification: Criteria and Checklist*, CRL, The Center for Research Libraries, 2007.

2.6 Interoperabilità nel dominio “architettura”: buone pratiche e soluzioni

Nel dominio “architettura” il problema dell’interoperabilità è un problema variegato e tuttavia molto concreto che sorge quando due entità, in realtà due sistemi rispettivamente nel ruolo di *fornitore* e *consumatore*, desiderano condividere componenti architetturali inizialmente posseduti solo dal *fornitore*. Questo problema sembra ben conosciuto, ma in pratica non è così, perché esso viene spesso confuso con tutti gli altri problemi di interoperabilità. Inoltre esso si presenta ben al di là del dominio delle biblioteche digitali; ad esempio, gli approcci basati su metodi e pratiche di programmazione includono la programmazione orientata agli aspetti¹⁰⁶ e quella orientata al soggetto¹⁰⁷, la separazione multi-dimensionale degli oggetti di interesse¹⁰⁸, tanto per fare alcune citazioni.

Nel campo delle biblioteche digitali, gli approcci all’interoperabilità nel dominio “Architettura” sono indirizzati a: (i) esporre una caratterizzazione di un componente architetturale (profilo del componente) per permetterne l’uso da parte del *consumatore*; (ii) sfruttare, con l’uso di standard, un componente architetturale messo a disposizione da una terza parte; (iii) usare, attraverso un servizio di mediazione, un componente messo a disposizione da una terza parte.

Gli approcci relativi al profilo del componente

hanno per scopo la condivisione delle informazioni sulle caratteristiche del componente stesso in modo da poter dare al *consumatore* le conoscenze necessarie ad usarlo. Attualmente esistono due importanti approcci alla modellazione del profilo dei componenti architetturali. Il primo ha per scopo la definizione di un profilo che sia specifico del sistema che lo definisce. Il secondo è rivolto a modellare, attraverso una struttura del profilo che abbia valenza generale, scenari possibili di interoperabilità correlati con architetture orientate ai servizi. Un esempio di quest’ultimo approccio è costituito dal profilo definito nel contesto della iniziativa *e-Framework*, che cerca di promuovere l’uso dell’approccio orientato al servizio nell’analisi e nella progettazione di software per l’istruzione e la ricerca. Un esempio concreto di questa specie di soluzioni per l’interoperabilità è WS-I Basic Profile¹⁰⁹, un insieme di specifiche di servizi Web non proprietari, specifiche che, accompagnate da spiegazioni, chiarimenti, interpretazioni ed ampliamenti, hanno lo scopo di promuovere l’interoperabilità dei servizi Web. Gli approcci per sfruttare, attraverso standard, un componente architetturale messo a disposizione da una terza parte includono SRU (search/Retrieval via URL)¹¹⁰, OpenSearch¹¹¹, e SWORD (Simple Web Service Offering Repository Deposit)¹¹². Questi sono standard che vengono implementati da certi componenti architetturali per rendere disponibili (dal

¹⁰⁶ G. Kiczales, J. Lamping, A. Mendhekar, C. Maeda, C. Lopes, J. M. Loingtier, *et al.*, *Aspect-oriented programming*, in *Proceedings ECOOP’97 — Object-Oriented Programming*, Springer, 1997, p. 220-242.

¹⁰⁷ W. Harrison e H. Ossher, *Subject-oriented programming: a critique of pure objects*, “ACM SIGPLAN Notices”, vol. 28, n. 10, 1993.

¹⁰⁸ P. Tarr, H. Ossher, W. Harrison, S. M. Sutton, *N degrees of separation: multi-dimensional separation of concerns*, in: *Proceedings of the 21st ACM international conference on Software engineering (ICSE ’99)*, 1999, p. 107-119.

¹⁰⁹ R. Chumbley, J. Durand, G. Pilz, T. Rutt, *WS-I Basic Profile Version 2.0*, 2010. <http://www.ws-i.org/Profiles/BasicProfile-2.0.html>.

¹¹⁰ The Library of Congress, *SRU: Search/Retrieval via URL*. <http://www.loc.gov/standards/sru/index.html>.

¹¹¹ [opensearch.org](http://www.opensearch.org). <http://www.opensearch.org/Home>.

¹¹² <http://www.swordapp.org/>.

Documenti e discussioni

punto di vista del *fornitore*) e/o sfruttare (dal punto di vista del *consumatore*) una funzione ben nota, ad esempio un servizio di ricerca nel caso di SRU e OpenSearch, o un servizio di deposito nel caso di SWORD.

Gli approcci relativi ai servizi di mediazione sono complementari a quelli basati su standard; questi ultimi rappresentano, idealmente, la migliore soluzione quando è in gioco una nuova implementazione di una certa funzionalità. Il loro uso tuttavia non sempre è possibile, o desiderabile, poiché non possono essere applicati ai sistemi esistenti non costruiti in base a standard e possono inoltre precludere implementazioni in qualche modo ottimizzate. Oltre a ciò, il processo di standardizzazione stesso spesso richiede perfino più tempo dell'implementazione degli standard. I mediatori sono componenti specificamente pensati per ospitare i meccanismi di interoperabilità. Questi componenti realizzano approcci di vario genere, ad esempio approcci basati: (i) su *blackboard* – cioè su componenti che permettono una comunicazione asincrona tra componenti di un sistema; (ii) su *connettori/adattatori* – cioè su componenti che trasformano l'interfaccia di un componente in una interfaccia compatibile; (iii) su *Proxy* – cioè su componenti che forniscono un'interfaccia per un altro componente; (iv) su *mediatori* – cioè basati su componenti che forniscono una interfaccia unica ad un insieme di interfacce di altri componenti incapsulando le modalità di interazione di questo insieme di oggetti; (v)

su *broker* – cioè basati su componenti aventi la responsabilità di coordinare la comunicazione, ad esempio inviando richieste, trasmettendo risultati ed anomalie; (vi) su *registri* – cioè componenti usati per garantire l'accesso ad altri componenti.

Note finali

L'interoperabilità è una delle sfide da affrontare nella realizzazione di sistemi informativi e biblioteche digitali che devono essere costruiti "assemblando" entità esistenti sviluppate indipendentemente (sistemi essi stessi) ma destinate a cooperare e fare affidamento l'uno su l'altro per poter eseguire compiti più complessi. Il progetto DL.org ha affrontato questa sfida con il contributo di più di 50 esperti internazionali che, riuniti in una serie di gruppi di lavoro tematici, hanno investigato le problematiche dell'interoperabilità da prospettive diverse ma complementari. Il risultato di questa attività è documentato nel rapporto *elaborato dal progetto DL.org, già citato*, e riassunto in questo breve saggio. Tuttavia l'interoperabilità è un tema "caldo", per il quale vengono continuamente progettate e sviluppate nuove soluzioni. Perciò il rapporto è considerato dagli autori un "manufatto" dinamico, che può essere oggetto di ampliamenti ed evoluzioni grazie all'apporto dell'intera comunità delle biblioteche digitali.

Per contattare gli estensori del rapporto ed offrire contributi, si prega di inviare una email a cookbook@dlorg.eu.

Although the importance of interoperability is well known and many attempts have been made in the past to give solutions to interoperability problems in the Digital Library domain, the results obtained are, in general, poor. The main reasons for this are a lack of systematic approaches, and, on average, a scarce knowledge of what has already been produced by various initiatives.

The DL.org project is the first initiative to examine interoperability from an all-encompassing perspective by harnessing leading figures in the Digital Library space globally. The output is an innovative Digital Library Technological and Methodological Cookbook with a portfolio of best practices and pattern solutions to common issues faced when developing interoperable digital library systems. This document is abstracted and abridged from "The Digital Technology and Methodology Cookbook", April 2011, <http://www.dlorg.eu/index.php/outcomes/dl-org-cookbook>.