

Innovazione tecnologica e Biblioteche

Inizia con questo fascicolo di Bibelot una nuova rubrica centrata sull'informazione su tecnologie e strumenti che possono migliorare e contribuire all'innovazione dei servizi bibliotecari. La Rubrica si propone di aiutare i bibliotecari che devono attuare scelte tecnologiche per migliorare l'impatto delle biblioteche.

Nota del Direttore

DSPACE-GLAM (Galleries, Libraries, Archives and Museums): un'estensione di DSpace per un Digital Library Management System innovativo e sostenibile

Andrea Bollini, Claudio Cortese, Emilia Groppo, Susanna Mornati

1. Introduzione

Nel corso degli ultimi venti anni, il settore delle Digital Library si è rivelato uno dei più interessanti e stimolanti per quanto concerne lo sviluppo di strumenti e funzionalità atti alla gestione, alla visualizzazione, alla cura e alla conservazione a lungo termine del patrimonio culturale digitale. Le Digital Library oggi non sono più chiamate a gestire solo testi o immagini digitali, ma anche una gran varietà di altri oggetti digitali come audio, video, blog, etc.¹ e si stanno ulteriormente evolvendo, nell'ottica di diventare, auspicabilmente nell'immediato futuro, una delle componenti dei Virtual Research Environment (VRE)², in grado di gestire in modo integrato tutti gli oggetti digitali che costituiscono il patrimonio delle istituzioni, ivi inclusa l'intera produzione scientifica delle istituzioni di ricerca (come ad esempio articoli, tesi, dataset). Di conseguenza la differenza tra Digital Library e *repository* si sta sempre più riducendo.

Il nostro gruppo di lavoro vanta una consolidata esperienza nel settore delle tecnologie digitali applicate ai beni culturali, della Digital Preservation e anche dei sistemi di fruizione basati su tecnologie Semantic web³. Dal 2004 siamo stati coinvolti in diversi progetti di creazione e gestione di Digital Library ed

¹ Candela L., Castelli D., Pagano P. (2011), "E-Publishing and Digital Libraries: Legal and Organizational Issues, chap. History, Evolution and Impact of Digital Libraries", Hershey USA, 1-30.

² Candela L., Castelli D., Manzi A., Pagano P. (2014), "Realising virtual research environments by hybrid data infrastructures: the D4science experience.", Proceedings of Science-International Symposium on Grids and Clouds (ISGC) 2014, 23-28 March, 2014, Academia Sinica, Taipei, Taiwan

Steiner C.M., Agosti M., Sweetnam M.S., Hilleman E.C., Orio N., Ponchia C., Hampson C., Munnely G., Nussbaumer A., Albert D., Conlan O. (2014), "Evaluating a digital humanities reserach environment: the CULTURA approach. International Journal on Digital Libraries (IJDL)", 15(1), 53-70

³ Barbera M., Cortese C., Zitarosa, R., Groppo E. (2009), "Building a Semantic Web Digital Library for the Municipality of Milan", Mornati S., Hedlund T. (eds.) Rethinking Electronic Publishing: Innovation in Communication Paradigms and Technologies - Proc. 13th International Conference on Electronic Publishing, Milan, 10-12, 133-154

Cortese C., Mantegari G. (2011), "Extending the digital archives of italian psychology with semantic data.", in Proceedings of the 1st International Workshop on Semantic Digital Archives, Berlin, 29 September 2011, 60-71

abbiamo sviluppato sistemi come 4S Digital Library⁴ e Codex[ml] utilizzati ancora oggi da molte prestigiose istituzioni italiane, che comprendono, tra le altre, la Biblioteca Nazionale di Napoli, la Veneranda Biblioteca Ambrosiana di Milano, la Fabbrica del Duomo di Milano, la Biblioteca del Conservatorio “Giuseppe Verdi” di Milano ed altre.

Inoltre, dal 2003 lavoriamo all’interno della community di DSpace⁵, dal 2006 anche come Committer e più recentemente nello Steering Group. Abbiamo dunque acquisito una consolidata esperienza nella creazione di repository istituzionali e di Research Information Systems per la governance della ricerca per prestigiose istituzioni internazionali, tra cui l’Università di Hong Kong, il Politecnico di Hong Kong, il National Institute of Education di Singapore, la Fraunhofer-Gesellschaft, il Consejo Superior de Investigaciones Científicas, l’Istituto Nazionale di Geofisica e Vulcanologia. Insieme con l’Università di Hong Kong⁶, abbiamo creato DSpace-CRIS⁷, l’unica soluzione open-source al mondo per la gestione delle informazioni di ricerca adottata anche dalla maggior parte delle università italiane come componente di IRIS, il sistema di gestione delle informazioni della ricerca integrato. Per aumentare l’interoperabilità tra i dati e i metadati relativi al patrimonio culturale e i prodotti e i dati della ricerca, abbiamo deciso di unire le nostre esperienze in entrambi i campi e abbiamo così esteso le funzionalità di DSpace per arrivare ad avere un Digital Library Management System ricco di funzionalità, ma, nello stesso tempo, facile da configurare e da utilizzare, sviluppato facendo tesoro a dell’esperienza e delle competenze della community internazionale.

2. DSpace

Nel concepire DSpace-GLAM siamo partiti dalla convinzione che, per gestire una Digital Library sostenibile nel tempo, si debbano oggi preferire soluzioni open source supportate da community internazionali affermate. Consideriamo infatti centrale dotare le istituzioni, soprattutto se medio-piccole, di strumenti sostenibili dal punto di vista economico, e nello stesso tempo atti a contrastare l’obsolescenza tecnologica e a permettere l’evoluzione dei servizi e delle funzionalità nel tempo. Proprio guardando innanzitutto alla sostenibilità abbiamo scelto DSpace, tra i sistemi open source più utilizzati, che permette, fin da subito, di caricare, archiviare, indicizzare, conservare e distribuire oggetti digitali in qualsiasi formato. DSpace è uno strumento che permette di gestire gli oggetti digitali in un repository strutturato e tecnologicamente molto avanzato, al fine di dare loro maggiore visibilità e garantirne l’accessibilità nel tempo. DSpace è specificamente apprezzato per la sua efficacia e per la semplicità

Boschini M.J., Cortese C. (2012), “The road (and the roadmap) for building trusted digital repositories within an interuniversity consortium”, in *Trusted Digital Repositories & Trusted Professionals*, Int. Conference, Firenze, 11-12 dicembre 2012

⁴ <http://dl.bnonline.it>

⁵ <http://www.dspace.org>

⁶ <http://hub.hku.hk/>

⁷ Palmer D.T., Bollini A., Mornati S., Mennielli M. (2014), *Dspace-cris@hku: Achieving visibility with a CERIF compliant open source system.*, in *Procedia Computer Science - 12th International Conference on Current Research Information Systems, CRIS 2014 - Managing data intensive science - The role of Research Information Systems in realising the digital agenda*, 118- 123

Bollini A., Mennielli M., Mornati S., Palmer D.T. (2016), “Iris: Supporting & managing the research life-cycle.”, in *Universal Journal of Educational Research*, 4, 738-743

delle procedure di installazione e di gestione. Esistono al mondo oltre 2200 repository digitali basati su DSpace e le istituzioni culturali e di ricerca di tutto il mondo lo utilizzano per una varietà di esigenze, tra cui anche quella di memorizzare, conservare e disseminare il patrimonio culturale digitale.

Le sue quattro macro-funzionalità principali sono:

1. Metadattare oggetti collegandoli ove possibile alla propria rappresentazione digitale;
2. Seguire tutte le fasi della Data Curation;
3. Permettere un facile accesso ai contenuti;
4. Garantire la conservazione a lungo termine degli oggetti digitali.

In quanto applicazione open source, DSpace è supportato da una community molto attiva di sviluppatori, di ricercatori e di utenti di tutto il mondo che contribuiscono, con la loro esperienza, a migliorare la piattaforma. Il Data Model è costruito intorno alla rappresentazione digitale dell'oggetto, in DSpace "item", costituito da metadati descrittivi e file correlati (bitstream). Ad ogni item viene assegnato un Persistent Identifier garantito dal Handle System. I file di metadati e i full-text sono indicizzati per essere reperibili anche mediante faceted browsing. Gli item sono organizzati in "collection", cioè in raggruppamenti di contenuti correlati tra loro secondo un qualsiasi criterio e definiti da uno specifico insieme di metadati, a loro volta raggruppate in community.

DSpace supporta l'idea di "community" organizzative (in genere mappate sulle sotto-unità di un'istituzione) per permettere di soddisfare esigenze distinte di gestione delle informazioni. Ogni community può, infatti, adattare il sistema a rispondere alle proprie particolari esigenze e configurare autonomamente il workflow di inserimento dei dati.

DSpace è in grado di gestire diverse categorie di metadati relativi ai propri contenuti digitali.

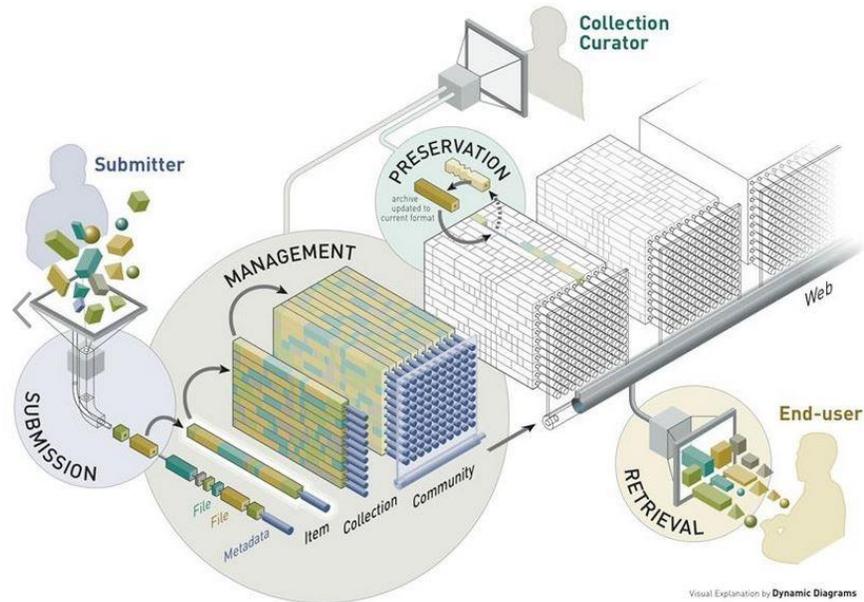
1. Metadati descrittivi

Di default sono gestiti i metadati contenuti nel DCMI (Dublin Core Metadata Initiative) Metadata Terms, ma può essere supportato qualunque schema di metadati "piatto", possono inoltre essere utilizzati contemporaneamente più standard, ad esempio per la descrizione di risorse di tipo diverso, o essere selezionati campi da schemi diversi.

2. Metadati amministrativi

La maggior parte di questi sono gestiti all'interno del database di DSpace, e dove significativi (ad esempio, le dimensioni byte dei bitstream e MIME type) vengono replicati anche all'interno dell'AIP (Archival Information Package) per essere facilmente accessibili all'esterno di DSpace.

Figura 1: il workflow di DSpace



3. Metadati strutturali

Questa tipologia di metadati contiene le informazioni relative alle relazioni tra le parti che costituiscono un item. La gestione nativa dei metadati strutturali in DSpace è attualmente limitata e permette solo di creare semplici sequenze di bitstream.

Il processo di inserimento dei dati può essere manuale o batch. DSpace è corredato di diversi strumenti per gestire i SIP (Submission Information Package, costituiti normalmente da metadati strutturati accompagnati da file): sono supportati molti formati di importazione, come il METS, e sono a disposizione strumenti di crosswalk. DSpace è corredato anche di strumenti batch per esportare oggetti e supporta protocolli, come OAI-PMH (Open Access Initiative Protocol for Metadata Harvesting), SWORD (Simple Web-service Offering Repository Deposit), e OpenURL. Consente inoltre di pubblicare contenuti memorizzati secondo lo standard Linked Open Data, convertendo i metadati in formato RDF. I dati sono memorizzati in un Triple Store, corredato di un end-point, per rendere i dati così trasformati accessibili via SPARQL. Può essere utilizzato qualsiasi Triple Store che supporti SPARQL Query Language 1.1 e lo SPARQL 1.1 Graph Store HTTP Protocol.

A fini di Digital Preservation, la piattaforma produce AIP (Archival Information Package) completi di metadati per la conservazione. Come già detto in precedenza, inoltre, il sistema attribuisce dei Persistent Identifier ad ogni item, e cattura e registra ogni evento significativo della vita dello stesso. Il sistema è dotato inoltre di una routine per checksum check, per il controllo periodico dell'integrità dei file conservati.

3. Funzionalità principali di una Digital Library

Come abbiamo visto nella sezione precedente, DSpace offre già “out-of-the-box” molte delle funzionalità richieste dai Digital Library Management System, e in questo senso viene utilizzato in molti progetti in

tutto il mondo. Tuttavia, sulla base della nostra esperienza, ci sono contesti in cui sono necessarie ulteriori funzionalità unite ad una maggiore granularità.

3.1 Gestione dei metadati

La gestione dei metadati è uno dei temi più importanti per un Digital Library Management System, soprattutto se il sistema è orientato alla conservazione a lungo termine. Infatti è necessario avere la sicurezza di importare SIP completi e affidabili, comprendenti metadati legati a standard diversi, sia di tipo “piatto” sia di tipo gerarchico, come l’EAD (Encoded Archival Description). Quindi i processi batch di inserimento devono essere corredati di sistemi di validazione dei metadati. Inoltre per facilitare le istituzioni che non sono in grado di fornire metadati standard, il sistema deve essere in grado di estrarre automaticamente i metadati amministrativi e tecnici dai file che vengono importati.

Nella nostra esperienza infatti, la creazione di SIP gestibili è una delle questioni più complesse ed impegnative quando si imposta un nuovo progetto di Digital Library. Anche per quanto concerne la creazione dei DIP (Dissemination Information Package) devono essere supportati i più importanti schemi di metadati per ciascun dominio (come MODS, METS, EAD e PREMIS). Per quanto riguarda i metadati strutturali, il sistema deve essere in grado di descrivere anche le strutture più complesse a livello di documento, caratterizzate da molti livelli e sottolivelli, come quelli dei manoscritti antichi o dei documenti musicali.

3.2 Gestione del contesto

Gestire le relazioni tra oggetti digitali e tra essi e altre entità (quali persone, eventi, luoghi, ecc.) deve essere considerata una funzionalità fondamentale per contestualizzare l’informazione culturale. Nell’ambito delle Digital Humanities, infatti, contestualizzare i documenti, i libri, i quadri o i resti archeologici è un’operazione della massima importanza per acquisire nuove conoscenze e gettare nuova luce sul patrimonio culturale. In quest’ottica, con il termine “contesto” non va considerato solo l’ambiente fisico (l’archivio, lo scavo, la galleria o la collezione), ma anche l’ambiente culturale geografico o storico od ogni altro criterio rilevante per rispondere alle domande della ricerca. Le scienze umane infatti spesso sono in grado di valutare il valore di un contenuto culturale solo contestualizzandolo e analizzando tutte le sue relazioni mediante un approccio globale e multidimensionale. Di conseguenza un Digital Library Management System deve permettere questo tipo di approccio finalizzato a descrivere i contesti culturali nella loro totalità.

3.3 Gestione dei bitstream

La fruizione di immagini, audio e video di alta qualità è fondamentale per lo studio e la valorizzazione del patrimonio culturale. Immagini, riviste, libri, giornali, manoscritti, mappe, materiali d’archivio e molti altri tipi di contenuti sono continuamente digitalizzati, ma spesso non è possibile accedervi o riutilizzarli a causa del mancato utilizzo degli standard di interoperabilità o delle limitazioni di banda della rete.

I ricercatori hanno invece bisogno di poter visualizzare velocemente immagini ad alta risoluzione, adattandole alle dimensioni dello schermo e al livello dello zoom, senza doverle necessariamente scaricare completamente o visualizzarne porzioni non necessarie. Per quanto riguarda la gestione delle risorse audio e video, è necessario ricorrere all’”adaptive streaming”, in modo da ridurre al massimo

l'utilizzo della banda, ma fornendo contemporaneamente a ogni utente la migliore risoluzione possibile dello stream audio/video, in base al dispositivo in uso ed alla velocità di connessione disponibile. Nel momento in cui, durante la riproduzione dello stream, variasse la velocità di connessione, il sistema deve essere in grado di modificare lo stream, in modo da migliorare la qualità della fruizione o la fluidità di riproduzione in caso di diminuzione della qualità del collegamento.

3.4 Supporto OCR

La possibilità di supportare l'OCR (Optical Character Recognition) è una funzionalità essenziale, dal momento che costituisce il primo passo verso l'analisi dei testi con i metodi della Data Science. Di conseguenza, un Digital Library Management System dev'essere integrato con sistemi OCR. Non va tuttavia dimenticato che anche i più recenti studi sottolineano la difficoltà di lettura di tali software per quanto riguarda i testi antichi, sia manoscritti che a stampa. Per questo motivo il sistema deve essere in grado di integrare diversi software OCR, al fine di soddisfare le esigenze di ogni particolare classe di documenti, tipologia di caratteri o tecnica di stampa. Inoltre è necessario includere il supporto specifico per il formato hOCR (lo standard aperto per la rappresentazione dei risultati della lettura OCR come sottoinsieme di HTML) per rendere più facile memorizzare, condividere, elaborare e visualizzare i risultati ottenuti con le tecnologie OCR.

4. DSpace-GLAM e i nuovi moduli per DSpace

Per soddisfare le esigenze delineate nella sezione precedente, il team di 4Science ha sviluppato diverse migliorie e funzionalità: è stato migliorato il processo di ingest di DSpace al fine di validare e gestire diversi standard di metadati, ed è stato esteso il Data Model, non solo per gestire i metadati strutturali, ma anche per descrivere gli oggetti digitali e le loro relazioni contestuali. Inoltre, sono stati sviluppati alcuni moduli per DSpace: l'IIIF (International Image Interoperability Framework) Image Viewer, il modulo per lo streaming audio/video e il modulo OCR.

4.1 Migliorie nella gestione dei metadati

Il processo batch di DSpace per l'inserimento dei dati è stato potenziato per permettere una corretta validazione dei metadati (in formato .csv) secondo lo standard prescelto e l'estrazione automatica dei metadati amministrativi, ove necessario. Inoltre, è possibile utilizzare file XML strutturati secondo i formati Dublin Core, MAG, METS, MODS, EAD, e altri ampiamente utilizzati, per modifiche massive e per compilare diversi tipi di metadati (descrittivi, amministrativi, tecnici e strutturali).

Al fine di garantire l'interoperabilità tramite il protocollo OAI-PMH, i formati METS, MODS, EAD e Dublin Core sono per ora i formati implementati. Per quanto riguarda lo standard PREMIS, esso viene utilizzato per esportare le informazioni relative alla storia degli oggetti digitali (controllo delle versioni, operazioni di controllo dell'integrità, ecc).

4.2 Estensione del Data Model

DSpace-GLAM consente di estendere il modello dei dati di DSpace definendo nuove entità (Persone, Eventi, Luoghi, Concetti, ecc.) e modificare quelle esistenti creando nuovi attributi o nuove relazioni. In questo modo DSpace-GLAM permette l'implementazione di qualunque modello concettuale atto a contestualizzare l'informazione culturale secondo diversi livelli di granularità.

Questa flessibilità viene sfruttata al fine di gestire le strutture di metadati complesse e multidimensionali richieste oggi alle Digital Library [CCSDS: Reference Model for an OAIS, 2012].

Innanzitutto è stato implementato un supporto più raffinato alla gestione dei metadati strutturali. In questo modo, durante il processo di ingest, il sistema è in grado di ricavare automaticamente una Table of Contents (TOC) predefinita, utilizzando come riferimento il nome del file e l'articolazione delle cartelle. L'operatore potrà poi modificare la TOC proposta introducendo livelli nidificati e/o riordinando le pagine. Una TOC più raffinata può essere creata, durante il processo di caricamento, utilizzando la sezione "structMap" di un file METS o la sezione "stru" di un file MAG, se fornita.

Inoltre è stato aggiunto il supporto per la descrizione gerarchica delle risorse archivistiche, ad esempio secondo lo standard Encoded Archival Description (EAD), e per gestirle nel loro contesto, utilizzando, per esempio, i concetti definiti nell'ambito dello standard EAC-CPF (Encoded Archival Context – Corporate Bodies, Persons and Families). Grazie all'implementazione di tali standard, DSpace-GLAM è quindi in grado di descrivere ed esporre strutture specifiche come quelle relative ai materiali archivistici, per poi fornire la fruizione delle risorse digitali correlate grazie al modulo IIIF (vedere la sezione successiva, 4.3).

4.3 IIIF Image Viewer

Per garantire una navigazione efficace delle immagini ad alta risoluzione, una community di esperti di biblioteche e musei e altre istituzioni ha deciso di sviluppare l'International Image Interoperability Framework (IIIF); il nostro gruppo di lavoro ha quindi sviluppato un nuovo modulo per DSpace basato su di esso.

Il modulo permette l'integrazione di DSpace con qualsiasi server compatibile con le API IIIF⁸ come ad esempio Digilib⁹, IIPImage¹⁰ o altri¹¹. Il modulo include una funzionalità per rendere i disponibili i file di immagine depositati in DSpace su Digilib facilmente adattabile ad altri image server. Al fine di rendere disponibili i metadati e la struttura dei contenuti (i metadati dei bitstream e le relative relazioni), il modulo utilizza l'IIIF Presentation API¹². Nella sua configurazione di default basata su Digilib, il modulo IIIF Viewer supporta una vasta gamma di formati di immagini: TIFF, JPG, JPG2000, PNG, BMP, RAW, ecc.

Lato client, viene utilizzato Universal Viewer¹³, un visualizzatore open-source compatibile con IIIF in grado di offrire una navigazione di alta qualità degli item con immagini multiple, permettendo una fruizione a singola pagina, a doppia pagina, mediante galleria di miniature. La vista corrente, il livello di zoom o la rotazione possono essere facilmente condivisi come URL, al fine di facilitare la condivisione e di inserire collegamenti precisi all'interno dei contributi scientifici. E'anche possibile copiare e incollare il codice per integrare il visualizzatore all'interno di siti web esterni (ad es.: blog e siti web

8 <http://iiif.io/api/image/2.1/>

9 digilib.sourceforge.net

10 iipimage.sourceforge.net

11 <http://iiif.io/app-demo/>

12 <http://iiif.io/api/presentation/2.1>

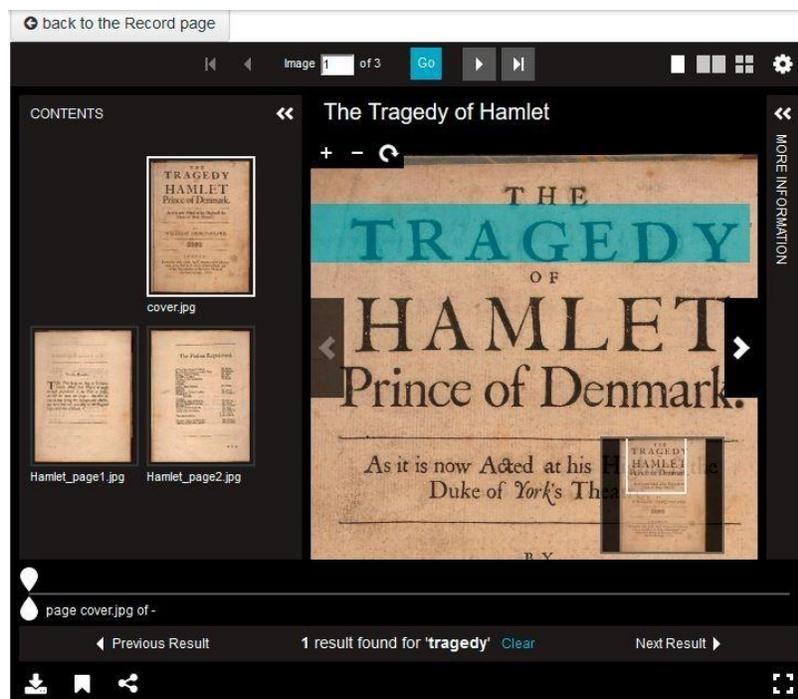
13 <https://github.com/UniversalViewer/universalviewer>

istituzionali) senza rinunciare ad una corretta attribuzione della risorsa. Il modulo rende possibile lavorare sulle immagini senza dover aspettare il tempo necessario a scaricarle interamente; inoltre apre alla possibilità di fornire a categorie di utenti prestabilite l'accesso a risoluzioni diverse dell'immagine, consentendo la protezione di immagini ad alta risoluzione o l'adozione di politiche di sfruttamento anche economico dei contenuti.

Il modulo di visualizzazione IIIF può interagire con altri moduli per fornire funzionalità aggiuntive. A titolo di esempio, può essere combinato con il modulo Document Viewer per convertire automaticamente file PDF ed in futuro ulteriori formati di e-book quali ad esempio EPUB , in una sequenza di immagini che può essere visualizzata all'interno del browser, utilizzando solo le funzioni web standard, senza la necessità di plug-in esterni, come Flash Player, e riducendo al minimo l'uso della banda.

Combinato con il modulo OCR (vedi paragrafo 4.5), può fornire la possibilità di ricercare un testo all'interno di qualsiasi tipo di documento e permettere il confronto tra la lettura OCR e l'immagine presente sullo schermo.

Figura 2 Il modulo Image Viewer integrato con il modulo OCR



4.4 Il modulo streaming Audio/ Video

Il modulo di streaming audio/video permette di convertire automaticamente i più diffusi formati audio/video in un formato e in una codifica appropriata per il media server adottato. Il modulo si occupa della registrazione del file convertito sul media server, locale o remoto, se necessario, con la creazione di diversi streaming con bitrate crescenti, segmentazione e diversificazione dei canali (audio, video, sottotitoli) per abilitare lo streaming adattivo. Il player, sulla base di VideoJS, permette l'integrazione

con il media server tramite i protocolli standard RTMP, HLS e DASH. In particolare, il media server proposto implementato utilizzando solo componenti open-source utilizza il protocollo standard DASH .

4.5 Il modulo OCR

DSPACE è stato integrato con un software OCR esterno al fine di fornire le basi per l'analisi del testo. Di default il modulo supporta il software open-source Tesseract¹⁴, ma la piattaforma prevede la possibilità di integrare anche altri diffusi motori OCR, come OCRopus o Abbyy FineReader. Utilizzando la rappresentazione del testo estratto in formato standard hOCR è possibile l'indicizzazione full-text in SOLR. Il modulo OCR fornisce anche supporto per le IIIF Search API attraverso una componente server che si preoccupa di presentare i risultati delle ricerche sui file hOCR in formato web annotation. Le IIIF Search API, utilizzabili con l'Image Viewer, consentono l'attivazione della funzionalità di ricerca all'interno del visualizzatore IIIF, dando la possibilità di effettuare ricerche all'interno delle immagini ed evidenziando i termini della ricerca sull'immagine stessa.

Tesseract supporta una varietà molto ampia di lingue e caratteri, tra cui: alfabeti latini (italiano, francese, spagnolo, tedesco), la scrittura da destra a sinistra (arabo), ideografica (cinese) e molte altre. Il motore OCR può anche essere istruito con i file di training personalizzati per riconoscere caratteri e lingue specifiche.

5. Conclusioni

L'estensione di DSPACE mediante moduli aggiuntivi specifici per le Digital Library rappresenta un'importante opportunità per la creazione di una nuova generazione di Digital Library Management System dotati di funzionalità avanzate per la gestione dei dati e dei metadati. L'integrazione di DSPACE con i moduli per la visualizzazione delle immagini e per la gestione dello streaming audio/video, l'aderenza agli standard di interoperabilità e l'estensione del Data Model alzano decisamente il livello della piattaforma.

Non si tratta più solo di un sistema per la gestione e la conservazione di file, ma di uno strumento per la fruizione, la condivisione e il riuso del patrimonio culturale e scientifico. La disponibilità e la qualità dei dati sul web è in continua crescita e c'è bisogno di strumenti che permettano di utilizzarli al meglio. Un Digital Library Management System come quello qui presentato costituisce un buon punto di partenza. Inoltre siamo pieni di nuove idee per migliorare ulteriormente le funzionalità della piattaforma. In particolare vorremmo dotare il sistema di funzionalità di annotation per sostenere forme di trascrizione collaborativa, la revisione OCR e per soddisfare le nuove esigenze dei ricercatori che studiano il patrimonio culturale.

Per il prossimo futuro, e per meglio identificare le linee di sviluppo e le priorità avremo, comunque, bisogno di un confronto serrato con la community di DSPACE e con le istituzioni che hanno adottato ed adotteranno DSPACE-GLAM. Come per tutto ciò che facciamo in 4Science, anche per tutti i nostri moduli l'obiettivo finale è quello di rilasciarli in open source il più presto possibile. E quelli che abbiamo

¹⁴ <https://github.com/tesseract-ocr>

descritto finora non costituiscono un'eccezione.