

# Dig *Italia*

Anno IV, Numero 1 - **2009**

Rivista del digitale nei beni culturali

ICCU-ROMA

# Seminario “CASPAR Training Day for the Cultural and Scientific Domains”

Roma 12 gennaio 2009

**Giovanni Michetti**

*Università di Urbino*

Il 12 gennaio 2009 si è svolto a Roma, presso l'Istituto centrale per il catalogo unico delle biblioteche italiane e per le informazioni bibliografiche (ICCU), il “CASPAR Training Day for the Cultural and Scientific Domains”, una giornata di formazione e aggiornamento dedicata al tema della conservazione nell'ambito scientifico e dei beni culturali. L'evento è stato organizzato dal gruppo di ricerca del progetto CASPAR (Cultural, Artistic and Scientific knowledge for Preservation, Access and Retrieval)<sup>1</sup>, un'iniziativa co-finanziata dall'Unione Europea all'interno del Sesto Programma Quadro (IST-2005-2.5.10 “Access to and preservation of cultural and scientific resources”) con l'obiettivo di:

- implementare, estendere e sperimentare il modello OAIS (ISO 14721:2003);
- consolidare le tecniche per l'acquisizione delle informazioni di rilievo per la conser-

- vazione di oggetti digitali;
- progettare servizi e sistemi di virtualizzazione a supporto della conservazione a lungo termine, in grado di affrontare gli inevitabili cambiamenti dell'hardware, del software e della comunità di fruitori delle risorse digitali;
- integrare i moduli per la gestione dell'accesso, dell'autenticità e dei diritti digitali;
- sviluppare casi di studio all'interno di specifici domini e ambienti tecnologici, individuare le condizioni per replicare le esperienze positive;
- contribuire attivamente alle attività di elaborazione e sviluppo di standard di settore;
- promuovere la conoscenza in materia di conservazione digitale e favorire in questo ambito la diffusione di sistemi, strategie, modelli e servizi diversificati, onde garantire una migliore adattabilità alle esigenze dei fruitori<sup>2</sup>.

<sup>1</sup> Il sito ufficiale del progetto è all'indirizzo <http://www.casparpreserves.eu>.

<sup>2</sup> Il progetto CASPAR coinvolge importanti enti di ricerca scientifica, istituzioni specializzate nella conservazione digitale e soggetti del mondo imprenditoriale. Il consorzio è composto da:

- Advanced Computer Systems S.p.A. (IT);
- Agenzia spaziale europea;
- Asemantics S.r.L. (IT);
- CNR. Istituto di scienza e tecnologie dell'informazione (IT);
- Centre National de la Recherche Scientifique (FR);
- Engineering Ingegneria Informatica (IT);
- FORTH - Foundation for Research and Technology (GR);
- IBM. Laboratori di ricerca di Haifa;
- Institut National de l'Audiovisuel (FR);
- International Centre for Art and New Technologies (CZ);
- IRCAM – Institut de Recherche et Coordination Acoustique/Musique (FR);
- Metaware S.p.A. (IT);
- Science and Technology Facilities Council (UK);

Il seminario organizzato presso l'ICCU deve essere quindi contestualizzato all'interno di un più ampio processo teso a diffondere una *cultura* della conservazione in ambiente digitale, con un *focus* particolare sui risultati conseguiti dal gruppo di ricerca.

La giornata è stata articolata in una sessione mattutina, dedicata principalmente all'esplorazione delle principali criticità che affliggono la conservazione in ambiente digitale e alla presentazione dei fondamentali concetti e nodi tematici che costituiscono il bagaglio indispensabile per affrontare il tema con un linguaggio comune. Di seguito alla parte teorica, nel pomeriggio sono state illustrate delle esperienze concrete di analisi e applicazione condotte nell'ambito del progetto CASPAR, articolando i lavori in due sessioni parallele, l'una concentrata sul dominio dei beni culturali, l'altra sul dominio scientifico<sup>3</sup>.

Rossella Caffo, in qualità di direttrice dell'ICCU, ha dato il benvenuto ai partecipanti presentando brevemente le attività dell'istituto e i maggiori progetti di ricerca internazionali nei quali è coinvolto.

David Giaretta, direttore del progetto CASPAR, ha significativamente proposto un intervento dal titolo "Preservation Threats" con l'obiettivo di inquadrare sommariamente i nuclei critici del problema conservativo, focalizzando la seconda parte della sua relazione sui risultati del censimento condotto all'interno del progetto PARSE.Insight (Permanent Access to the Records of Science in Europe), un'iniziativa cofinanziata dall'Unione Europea nell'ambito del

Settimo Programma Quadro e dedicata in maniera puntuale al tema della conservazione delle informazioni digitali prodotte in ambito scientifico<sup>4</sup>. Le indagini condotte su un campione significativo di più di duemila soggetti evidenziano chiaramente l'urgenza del tema e confermano la natura metodologica – più che tecnologica – del problema: le risposte fornite dai partecipanti al censimento hanno evidenziato che i pericoli di maggiore rilievo nei processi conservativi digitali riguardano la rottura della catena di conservazione, la labilità del concetto di autenticità, la dipendenza degli oggetti da una fitta rete di sistemi, oggetti e protocolli. La soluzione del problema conservativo risiede dunque in un buona miscela di strategie, politiche, scelte tecnologiche, organizzative e procedurali che complessivamente e in maniera integrata consentano di continuare a far vivere tanto gli oggetti quanto il contesto che contribuisce in maniera essenziale a definirne il loro *significato*.

Carlo Meghini, ricercatore dell'Istituto di scienza e tecnologia dell'informazione del Consiglio nazionale delle ricerche (CNR), ha dedicato lo spazio del suo ampio intervento all'illustrazione del modello OAIS (Open Archival Information System), un'architettura metodologico-concettuale ormai assunta a fondamento della ricerca scientifica di settore<sup>5</sup>. L'introduzione ai principali concetti e definizioni è stata accompagnata dalla presentazione di alcune concrete iniziative di implementazione del modello: le sperimentazioni del British Atmospheric Data Centre e dell'Atlas Petabyte Storage Service risalgono al 2004<sup>6</sup> e dimostrano

- UNESCO;
- Università di Urbino. Istituto di studi per la tutela dei beni archivistici e librari (IT);
- University of Glasgow. HATII - Humanities Advanced Technology and Information Institute (UK);
- University of Leeds. Interdisciplinary Centre for Scientific Research in Music (UK).

<sup>3</sup> I lucidi delle presentazioni sono disponibili all'indirizzo <http://www.iccu.sbn.it/genera.jsp?id=359&l=it>.

<sup>4</sup> Il sito ufficiale del progetto è all'indirizzo <http://www.parse-insight.eu/index.php>.

<sup>5</sup> Le principali caratteristiche del modello OAIS sono state già descritte sulle pagine di questa rivista: cfr. G. Michetti, *Il modello OAIS*, «Digitalia», III (2008), n. 1, p. 32-49, 4: [http://digitalia.sbn.it/upload/documenti/digitalia20081\\_MICHETTI.pdf](http://digitalia.sbn.it/upload/documenti/digitalia20081_MICHETTI.pdf).

<sup>6</sup> Cfr. *Applying the OAIS standard to CCLRC's British Atmospheric Data Centre and the Atlas Petabyte Storage Service*, Proceedings of the UK e-Science All Hands Meeting 2004, [http://ndg.nerc.ac.uk/public\\_docs/Giaretta\\_AHM.pdf](http://ndg.nerc.ac.uk/public_docs/Giaretta_AHM.pdf).

il ruolo strategico dello standard OAIS quale strumento di verifica della solidità delle scelte operative poste in essere; allo stesso periodo risale l'iniziativa congiunta del National Archives britannico e dello UK Data Archives, volta a disegnare una mappa di corrispondenze fra le attività dei due istituti e il modello funzionale proposto da OAIS<sup>7</sup>; numerosi altri progetti già realizzati o in corso di sviluppo hanno approfondito specifici aspetti del modello OAIS (come ad esempio i processi di acquisizione e di archiviazione, o la funzione di pianificazione), contribuendo a consolidarne l'impianto o a raffinarne il dettaglio.

Yannis Marketakis, ricercatore del FORTH (Foundation for Research and Technology), ha presentato il modello concettuale per la gestione e la conservazione della conoscenza elaborato dal gruppo di ricerca greco: l'approccio specialistico fondato sull'utilizzo di linguaggi formali ha permesso di evidenziare il sostrato tecnico che funge da fondamento per la definizione di modelli concettuali e più in generale per l'individuazione di un approccio al problema conservativo. Il nucleo fondamentale della riflessione svolta dai ricercatori greci è rappresentato dal concetto di *dipendenza*, definito per individuare un vincolo fra due oggetti, una relazione di subordinazione che impone la conoscenza, l'utilizzo o più genericamente la presenza di un oggetto *y* per comprendere un dato oggetto *x*: la nozione non è in sé originale, poiché il valore del *contesto* per una corretta interpretazione di una fonte informativa è acquisizione ormai consolidata nel settore, ma la novità risiede nella formalizzazione di quest'idea, che trova nelle descrizioni del linguaggio RDF (Resource

Description Framework) e delle ontologie una nuova linfa. Di qui l'introduzione del concetto di *chiusura*, inteso come insieme degli oggetti – e più in generale come *corpus* di conoscenze – necessario per comprendere una data risorsa; di qui la nozione di *divario di conoscenza* (*intelligibility gap*), con cui misurare la *distanza* fra un oggetto ed un suo eventuale fruitore; di qui l'innesto con il Web semantico e la previsione di possibili evoluzioni per un disegno concettuale denso di sviluppi.

L'intervento di Luigi Briguglio (Engineering Ingegneria Informatica) ha ricondotto i principali interrogativi teorici sul terreno dell'implementazione: come consentire l'accessibilità nel tempo? come assicurare la comprensibilità nel futuro? come garantire l'affidabilità degli oggetti? e che dire del recupero? L'architettura modulare elaborata nell'ambito del progetto CASPAR, fondata sul modello OAIS, è una possibile soluzione: ogni modulo, in autonomia o in associazione con altri, *controlla* un problema, e l'impianto complessivo coordina i diversi profili con l'obiettivo di proporre una soluzione coordinata al problema conservativo.

Giovanni Michetti, membro del gruppo di ricerca dell'Università di Urbino, ha esplorato il tema dell'autenticità nella conservazione in ambiente digitale: pur essendo ormai numerose le iniziative di settore sulla *digital preservation*, occorre sottolineare come – ad eccezione di casi isolati, seppure qualitativamente rilevanti<sup>8</sup> – solo recentemente la comunità scientifica abbia introdotto l'autenticità nell'agenda dei lavori, riconoscendo priorità a questo argomento, sempre più spesso declinato come *provenienza*.

<sup>7</sup> Cfr. Hilary Beedham – Julie Missen – Matt Palmer – Raivo Ruusalepp, *Assessment of UKDA and TNA Compliance with OAIS and METS Standards*, 2005, <http://www.data-archive.ac.uk/news/publications/oaismets.pdf>.

<sup>8</sup> I principali interventi scientifici sul tema dell'autenticità in ambiente digitale sono firmati da Luciana Duranti e risalgono alla metà degli anni Novanta. Il primo e – ancora oggi – più significativo progetto di analisi del tema in una cornice internazionale è dovuto a Luciana Duranti: il progetto InterPARES (International Research on Permanent Authentic Records in Electronic Systems) è tuttora in corso e costituisce un fondamentale punto di riferimento per gli studi di settore (cfr. il sito ufficiale all'indirizzo <http://www.interpares.org>).

L'autenticità degli oggetti digitali è fondata sulla loro *integrità* e sulla loro *identità*, ove la prima si riferisce alla completezza degli oggetti e all'assenza di corruzioni relative agli aspetti significativi, mentre la seconda si riferisce all'insieme delle caratteristiche che univocamente identificano una risorsa distinguendola dalle altre. Tali nozioni sono quindi più complesse di quanto l'accezione comune lasci intendere: l'integrità di un oggetto digitale – per quanto paradossale – non implica la conservazione del suo flusso di bit; e la sua identità non coincide banalmente con un qualsivoglia forma di identificazione, bensì è estesa a ricomprendere l'intero contesto entro cui l'oggetto è immerso. Nell'ambito del progetto CASPAR, il gruppo di ricerca di Urbino ha quindi concentrato l'attenzione su questo profilo, arrivando a proporre un modello astratto di gestione dell'autenticità, finalizzato allo sviluppo di un modulo software dedicato a questo aspetto. Come già accennato nel caso dell'intervento di Marketakis, le definizioni e i concetti non sono originali, anzi affondano le radici in lavori e riflessioni consolidate dal vaglio scientifico della comunità internazionale: l'elemento di rilievo è piuttosto la formalizzazione di un modello astratto mediante l'utilizzo di diagrammi UML, a supporto della definizione di un modello complessivo composto da moduli specializzati ma integrati. L'intervento di Mariella Guercio, coordinatrice del gruppo di ricerca urbinato, ha chiuso i lavori della mattinata fornendo degli spunti di riflessione sul tema della fruizione e dei bisogni degli utenti.

La sezione pomeridiana, articolata in due sessioni parallele, è stata dedicata alle esperienze di dominio: da una parte Mario Hernandez (UNESCO), con la collaborazione di Davide Palmisano (Asemantics) e di Marco Callieri (Istituto di scienza e tecnologie dell'informazione – ISTI-CNR), ha illustrato le criticità della conservazione dei dati afferenti al dominio dei beni culturali, focalizzando l'attenzione sul problema specifico della virtualizzazione, considerato il ruolo sempre più rilevante che questa tecnica assume nella valorizzazione dei beni culturali, soprattutto nell'ambito museale; dall'altra parte Luigi Fusco (European Space Agency – ESA), con la collaborazione di Sergio Albani (ESA), di Fulvio Marelli (Advanced Computer Systems – ACS) e di Yannis Marketakis, ha ripreso il tema dei fabbisogni degli utenti, specializzandolo all'interno del dominio scientifico e più in particolare del settore delle osservazioni geospaziali, e ha poi illustrato il caso di studio dell'Agenzia spaziale europea. In entrambe le sessioni le esperienze sono state presentate con l'obiettivo di mettere in luce le criticità, di illustrare metodologie, di evidenziare problemi applicativi, e non con l'intento di promuovere a paradigma un caso di studio, considerata la necessità di consolidare risultati e riflessioni.

A conclusione della giornata David Giaretta ha sintetizzato i temi e i concetti fondamentali che hanno percorso il seminario, ha raccolto gli spunti di riflessioni promossi dai relatori e avanzati dalla platea di partecipanti, e ha incoraggiato la diffusione dei temi oggetto della giornata di studio, auspicando l'intensificarsi del dibattito e la realizzazione di iniziative analoghe nel futuro.