

I metadati per la gestione e conservazione dei documenti elettronici

Fiorella Foscarini
Firenze, 17 giugno 2007

Argomenti

- Che cosa sono i metadati
 - Criticità relative a definizione, indentificazione e classificazione dei metadati
- L'utilizzo dei metadati per la gestione delle risorse documentarie
- La funzione conservativa nell'ambiente digitale e il ruolo dei metadati
- I formati di interscambio
- Analisi dello standard *ISO 23081*

Identificazione dell'oggetto metadato: criticità

■ Definizione di metadato

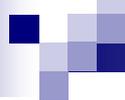
- Dati (strutturati) sui dati
- Dati che descrivono il contenuto, la struttura e il contesto dei documenti e la loro gestione nel tempo (*ISO 15489*)
- Rapporto dati-metadati: un *meta*-livello può divenire oggetto di contenuti per un *meta*-livello superiore

■ Classificazione dei metadati

- In base alle finalità operative: *descriptive / structural / administrative metadata*
- In base ai tipi di relazione con la risorsa: metadati interni / esterni
- In base alle responsabilità verso i documenti: metadati del produttore / metadati del conservatore

Identificazione dell'oggetto metadato: criticità (cont.)

- Analisi del **livello descrittivo** dei metadati rispetto alle informazioni da trattare
 - Utenti diversi sono interessati a livelli diversi di **granularità** e di **aggregazione**
- Individuazione di **tipologie funzionali** di oggetti informativi
 - **“Convergenza al digitale”**
- Identificazione dei **metadati descrittivi e gestionali** necessari per le attività di **conservazione**
 - Incertezza sull'oggetto informativo da conservare e su principi e contenuti della descrizione
- Proliferazione dei **formati di metadati**
 - Contrasta la finalità primaria dei metadati: **scambio delle informazioni per il recupero delle risorse e l'interoperabilità** dei sistemi informativi



Funzioni generali dei metadati

1. Descrizione delle risorse
2. Recupero e accesso alle informazioni
3. Gestione delle informazioni
4. Gestione dei diritti e delle informazioni relative alla provenienza
5. Protezione dell'autenticità degli oggetti informativi
6. Interoperabilità ed e-commerce
7. Conservazione delle risorse elettroniche

Obiettivi dei metadati archivistici

- Garantire l'identificazione permanente dei **singoli oggetti informativi**
- Garantire l'identificazione permanente delle **relazioni** tra gli oggetti informativi
- Conservare le informazioni che supportano l'**intellegibilità** degli oggetti informativi
- *Identificativo univoco (numero di protocollo, data, autore, ecc.)*
- *Indici di classificazione e fascicolazione*
- *Procedimento amministrativo cui il documento è connesso*

Obiettivi dei metadati archivistici (cont.)

- Mantenere le informazioni relative alle attività di tenuta rilevanti ai fini dell'**autenticità**
- Documentare gli specifici **processi conservativi** cui l'oggetto informativo è stato sottoposto
- *Disegno del sistema, controllo sugli accessi, responsabilità per la conservazione, dati della collocazione "fisica" nel deposito digitale, ecc.*
- *Migrazioni, copie, ecc.*

La funzione conservativa in ambiente digitale

- **Dematerializzazione** delle fonti documentarie prodotte e mantenute informaticamente
 - **Vantaggi**: accessibilità generalizzata, rapido e facile trasferimento su altri supporti ed in ambienti diversi, condivisione, ecc.
 - **Limiti**: necessità di una molteplicità di mediazioni, strumenti e risorse per poterle utilizzare (leggere e comprendere), obsolescenza tecnologica, facilità di manipolazione
- **La conservazione in ambiente digitale è una funzione dinamica e continua nel tempo**

La funzione conservativa in ambiente digitale (cont.)

- I **tempi di intervento** per il mantenimento delle memorie si sono accorciati
- La **diversificazione** dei prodotti non consente soluzioni univoche
- La **fragilità dei supporti** e la **facilità di manipolazione** richiedono investimenti significativi in termini di controllo e sicurezza
- **La conservazione digitale non è compatibile con la trascuratezza che ha caratterizzato il sistema conservativo tradizionale**

La funzione conservativa in ambiente digitale (cont.)

- La conservazione delle fonti digitali dipende in larga misura dalle condizioni della loro **formazione** e dalle condizioni previste nel **disegno del sistema**
- È indispensabile
 - definire **precocemente** le **responsabilità** per la tenuta dei documenti
 - utilizzare **formati standard** sia per la formazione dei documenti che per la predisposizione dei necessari metadati di contesto, ordinamento e gestione

Il ruolo dei metadati per la conservazione delle fonti digitali

- La semplice conservazione del **flusso di bit** non è sufficiente!
- Occorre anche mantenere le informazioni necessarie per assicurare la possibilità di **interpretazione futura** del flusso di bit → **metadati descrittivi e gestionali**
- I metadati sono fondamentali per mantenere i documenti digitali **accessibili, utilizzabili** (leggibili e comprensibili) e **autentici** (univocamente identificabili ed integri) nel medio e lungo periodo, **in un ambiente tecnologico certamente diverso da quello originario**

Il ruolo dei metadati per la conservazione delle fonti digitali (cont.)

- Per assicurare la **completezza** dei metadati raccolti, occorre anticipare l'intervento conservativo
- Metadati di **qualità**
 - Esaustivi, accurati, catturati in modo il più possibile automatico, leggibili e intellegibili nel tempo, interscambiabili
- **Complessità** dei metadati per la conservazione dei documenti digitali
 - Ogni componente informativa (sia interna che esterna al documento) deve essere resa **esplicita**

Il ruolo dei metadati per la conservazione delle fonti digitali (cont.)

■ I metadati vanno gestiti!

- Necessità di documentare struttura e procedure di gestione degli **schemi di metadati** adottati, le loro reciproche relazioni e l'evoluzione degli schemi nel corso del tempo
- Occorre garantire sia l'**autonomia** dei singoli domini che l'**interoperabilità** tra gli schemi

La separazione di forma e contenuto nei documenti digitali

- Caratteristica dei documenti digitali: indipendenza di **struttura fisica** e **struttura logica**
- Quanto più è complessa e ricca la natura intrinseca di un documento, tanto più la struttura logica deve essere resa **esplicita** mediante accorgimenti **'estrinseci'**
- Al fine di **esplicitare la struttura logica** è centrale l'**individuazione coerente e sistematica delle componenti informative essenziali del e sul documento**:
 - Contenuto, elementi e attributi della forma estrinseca ed intrinseca, contesti di produzione e d'uso (contesto giuridico, amministrativo, documentario, tecnologico), ecc.

L'esplicitazione della struttura e i formati di interscambio

■ SGML

- Linguaggio di codifica generale o metalinguaggio

■ XML (*eXtensible Markup Language*)

- Sottoinsieme di codifiche SGML
- Metodo **diffuso**, a **basso costo** e **scalabile** per una gestione **indipendente dal software** di documenti e metadati
- Codifica comprensibile sia agli uomini che alle macchine
- Obiettivi: disseminazione, scambio e ricerca di informazioni (incl. documenti ipermediali) nel web; conservazione

■ DTD (*Document Type Definition*)

- Grammatica di SGML/XML

L'esplicitazione della struttura e i formati di interscambio (cont.)

- I linguaggi di marcatura
 - supportano il trattamento della struttura logica e semantica dei contenuti
 - consentono la gestione dei riferimenti esterni al documento
 - facilitano il riconoscimento e riuso a fini molteplici di componenti informative interne
- I metadati prodotti e gestiti tramite tali formati orientati ai dati (o *'formati standard'*) sono allo stesso tempo contenuto e informazione sul contenuto
- Vantaggi:
 - **Riuso, interoperabilità, leggibilità e intellegibilità** nel tempo dei documenti e dei metadati associati



Linee guide per la gestione dei metadati. Lo standard ISO 23081-1

ISO 23081-1

Information and documentation – Records management processes – Metadata for records

- **Part 1: Principles** (2006, già pubblicato come TS nel 2004)
 - Si basa sulle specifiche **ISO 15489**
 - Definisce le caratteristiche generali di un **quadro di riferimento** per lo sviluppo e l'applicazione di un sistema di metadati per la gestione documentaria
- **Part 2** (2007)
 - **Guida pratica** all'implementazione dello standard
- **Part 3** (di prossima pubblicazione)
 - Strumento per la **valutazione degli schemi di metadati** esistenti, sulla base dei requisiti dello standard

ISO 23081-1

Finalità generali

- Identificare **responsabilità** e **modalità** per gestire e coordinare creazione, uso e mantenimento degli **schemi di metadati**
- Definire il ruolo dei metadati all'interno del **sistema informativo** di un ente – e del sottostante quadro normativo
- Stabilire meccanismi di **controllo** ed **aggiornamento** tali da garantire la **continuità** del sistema

ISO 23081-1

Finalità generali (cont.)

- Garantire la **sostenibilità** di un sistema di metadati in relazione alle continue trasformazioni amministrative e organizzative
- Prestare particolare attenzione ai problemi della **conservazione digitale** e alle strategie specifiche necessarie a catturare e mantenere i metadati necessari
- Incorporare i **metadati di gestione** dei documenti nei **sistemi archivistici**

ISO 23081-1

Procedure contenute nelle linee guida

- **Condivisione** dei metadati tra domini diversi
- **Documentazione** degli schemi di metadati sviluppati mediante:
 - Sistemi di controllo delle versioni
 - Registri degli schemi all'interno di ogni comunità di settore
 - Procedure definite di archiviazione
 - Procedure di monitoraggio e mantenimento dei collegamenti tra gli schemi

ISO 23081-1

Procedure contenute nelle linee guida (cont.)

- Definizione di regole per l'**identificazione logica** o l'**esportazione fisica** dei metadati associati ai documenti
- Identificazione di **strategie di memorizzazione** dei metadati coerenti con le strategie per la conservazione permanente dei documenti medesimi
- Adozione di procedure di **conservazione** nel tempo dei metadati con particolare attenzione alle esigenze di **autenticità**, **affidabilità** e **usabilità**

ISO 23081-1

Definizioni

■ Definizione di metadato:

- Informazioni strutturate o semi-strutturate che consentono la formazione, registrazione, classificazione, accesso, conservazione e selezione di documenti nel tempo e nell'ambito di diversi domini

■ In particolare, i *records management metadata* consentono di

- identificare, autenticare e contestualizzare i documenti
- ... nonché le persone, le regole, i processi ed i sistemi coinvolti nella creazione, gestione e conservazione dei documenti

ISO 23081-1

Caratteristiche dei *RM metadata*

- I metadati hanno **natura dinamica**
 - Le informazioni relative al **contesto documentario** e alla gestione delle **processi lavorativi** si accrescono nel tempo
 - I metadati rilevanti nella **fase attiva** possono continuare ad applicarsi nelle fasi successive di gestione
- I metadati per la gestione dei documenti digitali presentano **livelli diversi di rilevanza**
- I set di metadati vanno adeguatamente gestiti

ISO 23081-1

Caratteristiche dei *RM metadata* (cont.)

Distinzione fondamentale tra:

- **Metadati catturati nel momento della formazione dei documenti**
 - Contesto di produzione (soggetto produttore, competenze, organizzazione, ecc.)
 - Contesto amministrativo (processi di lavoro, responsabilità, ecc.)
 - Contesto documentario, incl. attributi tecnici (contenuto, forma fisica, struttura, requisiti di visualizzazione, ecc.)
- **Metadati catturati nelle fasi di gestione successive**
 - Processi di *RM* e processi amministrativi cui il sistema documentario e i singoli documenti sono sottoposti durante la loro esistenza, incl. ogni cambiamento di contenuto, struttura e comportamento

ISO 23081-1

Gestione dei metadati

- Identificazione di **ruoli e responsabilità** specifiche per la gestione dei metadati, distinguendo tra:
 - funzioni di **creazione e gestione dei metadati**
e
 - funzioni di **creazione, applicazione e gestione delle regole** che ne governano i processi e le strutture di gestione (DTD, data dictionaries, schemi di metadati, ecc.)
- L'insieme dei metadati costituisce a sua volta un documento (***metadata record***) che richiede di essere adeguatamente gestito e conservato ... almeno per tutta l'esistenza dell'oggetto informativo a cui si riferisce

ISO 23081-1

Gestione dei metadati (cont.)

- Occorre predisporre e conservare le strutture di metadati sotto forma di veri e propri **schemi**, in grado di:
 - documentare le relazioni tra i singoli elementi e renderle comprensibili e ricostruibili
 - sostenere esigenze di integrazione, interoperabilità, uniformità semantiche e sintattiche e modularità dei sistemi informativi
- Esempi di *metadata schema*
 - Schemi per la classificazione dei documenti correlati alle funzioni amministrative
 - Schemi DTD XML per la definizione di tipologie documentarie, strutture di database e altri oggetti
 - Schemi logici di database
 - Schemi relativi ai diritti di accesso e alla sicurezza

ISO 23081-1

Tipologie di metadati

1. Metadati relativi ai documenti

- Estremi cronologici, persone, struttura dei documenti e degli archivi, formato, supporto, caratteristiche tecniche, relazioni interne tra gli elementi costitutivi, requisiti di presentazione, relazioni coi processi amministrativi, aggregazioni documentarie, sicurezza e accesso, sistemi di recupero, informazioni su eventuali processi di migrazione o emulazione

2. Metadati relativi alle regole e politiche dell'ordinamento giuridico-amministrativo

- Regolamenti di organizzazione, norme relative al sistema documentario, regole di accesso e sicurezza

3. Metadati relativi ai soggetti produttori

- Denominazioni, quadro delle strutture interne, responsabilità e profili utenti

ISO 23081-1

Tipologie di metadati (cont.)

4. Metadati relativi ai processi di lavoro

- Funzionigramma, organigramma, quadro dei processi amministrativi in relazione a iter e responsabilità

5. Metadati relativi ai processi di gestione documentaria

- Sistemi di registrazione, piani di classificazione, piani di conservazione, classificazione degli accessi, piano della sicurezza, documentazione dei processi conservativi

6. Metadati relativi al documento sui metadati

- Regole e responsabilità per produzione, sviluppo e modifica dei metadati, procedure di controllo versioni degli schemi di metadati, procedure di mantenimento dei collegamenti interni/esterni agli schemi, strategie di memorizzazione dei metadati

Riferimenti bibliografici

- International Standard Organization, *ISO 23081: Information and documentation. Records management processes. Metadata for records. Part 1: Principles* (2006).
- Maria Guercio, “Principi, metodi e procedure per la conservazione a lungo termine dei documenti informatici”, dispensa conservazione (Urbino, maggio 2005).
- Kate Cummings, “Metadata Matters”, in Julie MacLeod e Catherine Hare (a cura di), *Managing Electronic Records*, Facet Publishing, Londra, 2005.
- Archivi Nazionali Australiani (NAA), “Describing Records Using Metadata“. Disponibile online al sito: <http://www.naa.gov.au/records-management/create-capture-describe/describe/index.aspx>.