

Library of Congress Network Development and MARC Standards Office

Capire PREMIS

Autore: Priscilla Caplan

Data di pubblicazione: 1 Febbraio 2009

Copyright © 2009 The Library of Congress, eccetto che negli U.S.A.
I crediti devono essere riconosciuti negli stralci di questa pubblicazione.

Traduzione: Angela Di Iorio
Fondazione Rinascimento Digitale

CAPIRE PREMIS

INDICE DEI CONTENUTI

1. CONTESTO APPLICATIVO DI PREMIS.....	3
1.1. Cosa sono i metadati di conservazione?	3
1.2. Cos'è PREMIS?	4
1.3. Cosa c'è il dizionario dei dati PREMIS?	4
1.4. Come dovrebbe essere usato PREMIS?	5
1.5. Si dovrebbe usare PREMIS?.....	6
2. CONVENZIONI DEL DIZIONARIO DEI DATI.....	7
2.1. Unità Semantiche	7
2.2. Contenitori e Sottounità	7
2.3. Contenitori di estensione	8
3. MODELLO DEI DATI PREMIS	9
3.1. Entità Intellettuale	9
3.2. Entità Oggetto	10
3.3. Eventi	11
3.4. Agenti	11
3.5. Diritti	12
4. IL DIZIONARIO DEI DATI	13
4.1. Esempio di unità semantica semplice come voce del Dizionario dei dati.....	13
4.2. Esempio di unità contenitore come voce del Dizionario dei Dati	14
5. USO DI PREMIS	15
5.1. PREMIS in XML	15
5.2. Conformità PREMIS	16
6. PER MAGGIORI INFORMAZIONI	17
APPENDICI	
Appendice A: Lista di tutte le Unità Semantiche PREMIS	18
Appendice B: Esempio di Oggetto	21
Appendice C: Glossario dei termini	24

CAPIRE PREMIS

Questa guida è una panoramica relativamente breve sullo standard di metadati di conservazione PREMIS. Essa non copre tutte le informazioni per implementare PREMIS, ma da un'idea di cosa sia PREMIS e di quale sia il suo campo d'azione.

Per molti lettori, sarà sufficiente. Per coloro che hanno invece bisogno di avere una maggiore padronanza del Dizionario dei Dati PREMIS per i Metadati di Conservazione, la presente guida potrebbe servire come introduzione che rende più facile l'approccio a documenti più specifici. I termini che si presentano nel font VERDANA sono unità semantiche PREMIS. I termini presenti nel Glossario sono messi in evidenza dalla sottolineatura alla loro prima occorrenza.

1. CONTESTO APPLICATIVO DI PREMIS

1.1. Cosa sono i metadati di conservazione?

Chi lavora in una biblioteca o in un archivio ha buone possibilità di sapere almeno qualcosa sui metadati e sulla descrizione delle risorse. Probabilmente è al corrente del fatto che i metadati vengono suddivisi in categorie in base a ciò che si vuole ottenere: i metadati descrittivi aiutano nella ricerca e nell'identificazione delle risorse, i metadati amministrativi aiutano nella gestione e nel tracciamento e i metadati strutturali indicano come gli oggetti digitali complessi sono messi insieme per essere visualizzati e per qualsiasi altro uso. In modo simile, i metadati di conservazione supportano tutte quelle attività che servono ad assicurare, nel lungo periodo, l'usabilità della risorsa digitale.

Il Dizionario dei Dati PREMIS definisce i metadati di conservazione come "l'informazione usata da un deposito allo scopo di supportare il processo di conservazione digitale". Di seguito vengono esemplificate alcune attività di conservazione e come i metadati le supportano.

- Una risorsa deve essere conservata mantenendo un livello di sicurezza tale che nessuno possa modificarlo inavvertitamente (o intenzionalmente). L'informazione di *checksum* archiviata come metadato può essere usata per dire se un file archiviato è stato modificato in un determinato arco temporale.
- I file devono essere archiviati su supporti che possono essere letti dai computer attuali. Se i supporti vengono danneggiati o diventano obsoleti (come i *floppy disk* da 8 pollici usati negli anni '70) può essere difficile o impossibile recuperare i dati. I metadati possono essere utili alla gestione dei supporti di archiviazione, perché ne memorizzano ad esempio tipo ed età come anche le date in cui è stato eseguito l'ultima volta il *refresh* dei file.
- Nei lunghi periodi di tempo anche i formati dei file più diffusi possono diventare obsoleti, determinando il fatto che nessuna applicazione è più in grado di restituirne il contenuto. I manager della conservazione devono adottare delle strategie di conservazione per assicurare che le risorse rimangano usabili. Ciò potrebbe significare trasformare i vecchi formati negli equivalenti più aggiornati (migrazione), o ricreando lo stesso ambiente tecnologico di fruizione dei contenuti su nuove piattaforme hardware e software (emulazione). Entrambe le strategie, di migrazione e di emulazione richiedono metadati sui formati dei file originari e su hardware e software che li supportano.
- Le azioni di conservazione potrebbero determinare il cambiamento delle risorse originarie o il cambiamento di come esse vengono supportate tecnologicamente per la fruizione. Ciò può mettere in dubbio l'autenticità della risorsa. I metadati possono supportare l'autenticità documentando la provenienza digitale della risorsa – la sua catena di custodia e i cambi evolutivi autorizzati.

1.2. Cos'è PREMIS?

PREMIS sta per "PREservation Metadata: Implementation Strategies" che è il nome di un gruppo di lavoro internazionale sponsorizzato da OCLC e RLG dal 2003 al 2005. Il gruppo di lavoro ha prodotto un rapporto chiamato *PREMIS Data Dictionary for Preservation Metadata* (Dizionario dei Dati PREMIS per i Metadati di Conservazione) che include sia un dizionario dei dati che alcune indicazioni sui metadati di conservazione. Una seconda versione aggiornata è stata pubblicata nel Marzo 2008. La *Library of Congress* sostiene l'attività di mantenimento di uno schema XML per la rappresentazione di PREMIS.

Quest'ultima è chiamata *PREMIS Maintenance Activity* ed è sponsorizzata dalla *Library of Congress*. Essa include un sito web che rimanda a tutti i tipi di informazioni ufficiali e non ufficiali riguardanti il PREMIS, una lista di discussione, un *wiki* per gli implementatori di PREMIS e un Comitato Editoriale responsabile per le revisioni al dizionario dei dati e allo schema. La *Maintenance Activity* tenta anche di promuovere la conoscenza di PREMIS, sponsorizza seminari sull'uso di PREMIS e commissiona studi e pubblicazioni correlate al PREMIS, come questa guida.

Abitualmente, quando le persone fanno riferimento a "PREMIS" intendono il Dizionario dei Dati. Occasionalmente potrebbero fare riferimento allo schema XML, al gruppo di lavoro o all'intera iniziativa includendo la *Maintenance Activity*.

- Dizionario dei dati PREMIS: www.loc.gov/premis/v2/premis-2-0.pdf
- Sito web PREMIS: www.loc.gov/standards/premis/
- Lista di discussione del gruppo di implementatori PREMIS: listserv.loc.gov/listarch/pig.html

1.3. Cosa c'è nel Dizionario dei Dati PREMIS?

Il Dizionario dei Dati PREMIS definisce un insieme di base di unità semantiche (sezione 2.1) che i depositi dovrebbero avere per sopperire alle funzioni di conservazione .

Le funzioni di conservazione possono essere diverse a seconda del deposito, ma generalmente includono azioni per assicurare che gli oggetti digitali sopravvivano (p.e. possano essere letti dai media di supporto) e siano utilizzabili (p.e. possano essere visualizzati, eseguiti o altrimenti interpretati dalle applicazioni software), come devono assicurare che gli oggetti digitali nel deposito non vengano inavvertitamente alterati, e che i cambiamenti legittimi agli oggetti vengano opportunamente documentati.

Il Dizionario dei Dati non è stato concepito per definire tutti gli elementi di metadati di conservazione possibili, ma solo per definire quelli che la maggior parte dei depositi avrà bisogno di sapere la maggior parte delle volte. Diverse categorie di metadati sono escluse in quanto ritenute fuori obiettivo:

- o Metadati per formati specifici, ad esempio i metadati che riguardano solo un formato di file o classi di formati come audio, video o di grafica vettoriale.
- o Metadati per implementazioni e regole business specifiche, come per esempio i metadati che descrivono le politiche o le pratiche di un deposito su come esso fornisce accesso ai materiali.

Capire PREMIS

- Metadati descrittivi. Sebbene la descrizione delle risorse è ovviamente, di pertinenza della conservazione, molti standard indipendenti possono essere usati a tale scopo, come MARC21, MODS e Dublin Core.
- Informazioni dettagliate su media o su hardware. Di nuovo, sebbene chiaramente di pertinenza della conservazione, la definizione di tali metadati vengono lasciati alla competenza di altre comunità.
- Informazioni sugli agenti (persone, organizzazioni o software), tranne che per il minimo necessario all'identificazione.
- Informazioni sui diritti e sui permessi, eccezion fatta per quelli che hanno un impatto diretto sulle funzioni di conservazione.

Se si pensa che ogni metadato sia utile all'organizzazione che gestisce un deposito digitale, PREMIS può essere visto come definizione di un sottoinsieme che sta nel centro. Da una parte, non è coinvolto nella ricerca e nell'accesso, e dall'altra non definisce metadati dettagliati per formati specifici. Esso definisce solo i metadati comunemente necessari ad eseguire funzioni di conservazione su tutti i materiali.

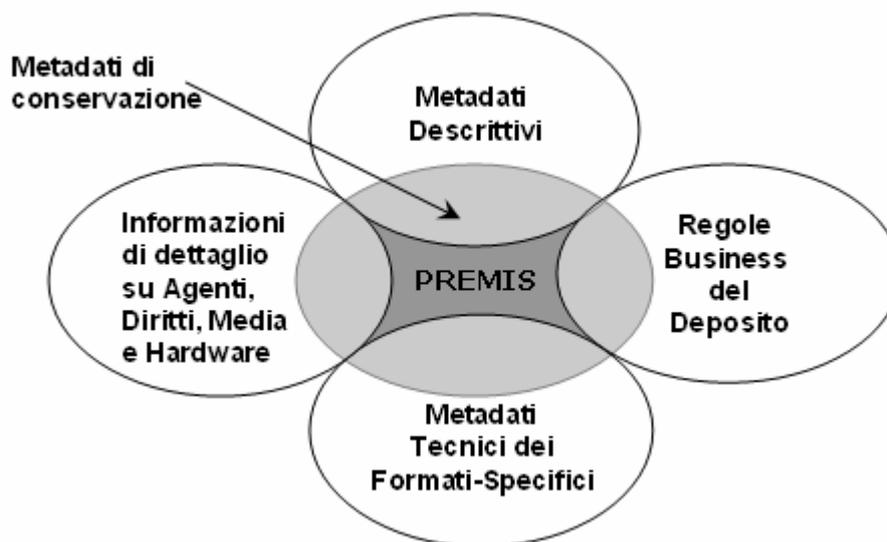


Figura 1. PREMIS come sottoinsieme di tutti i Metadati di Conservazione

La Figura 1 mostra tutti i metadati pertinenti alla conservazione come il cerchio in ombra posto al centro del diagramma. Il cerchio include alcuni metadati descrittivi, alcune regole business, alcuni metadati tecnici di dettaglio e alcune informazioni dettagliate su agenti, diritti, media e hardware. Il PREMIS è la piccola parte al centro dei metadati di conservazione, mostrato in grigio più scuro, che si differenzia dagli altri tipi.

1.4 Come dovrebbe essere usato PREMIS?

Il dizionario dei dati PREMIS definisce ciò che un deposito di conservazione ha bisogno di sapere. E' importante notare che l'attenzione è focalizzata sul sistema del deposito e sulla sua gestione e non sugli autori del suo contenuto digitale, o sulle persone che scannerizzano o altrimenti

Capire PREMIS

convertono il contenuto da analogico in digitale, o sullo *staff* che valuta e licenzia le risorse elettroniche commerciali. L'uso primario di PREMIS è per la progettazione e la valutazione dei depositi come anche per lo scambio di pacchetti di informazioni archiviate tra i depositi.

Coloro che progettano e/o sviluppano applicazioni software per i depositi di conservazione dovrebbero usare PREMIS come linea guida su quali informazioni dovrebbero essere ottenute e salvate dalle applicazioni o altrimenti essere presenti nella gestione del deposito.

Coloro che stanno pianificando di implementare un deposito di conservazione dovrebbero usare il PREMIS come una *checklist* per la valutazione dei software candidati. I sistemi che possono supportare il dizionario di dati PREMIS saranno maggiormente in grado di conservare le risorse informative nel lungo periodo.

Un deposito attivo potrebbe decidere ad un certo punto di esportare i pacchetti di informazione archiviati in un altro deposito. Per esempio, un'organizzazione di custodia potrebbe voler migrare da un deposito all'altro, o un cliente potrebbe voler cambiare l'azienda che eroga di servizio di conservazione. PREMIS fornisce un insieme comune di elementi di dati che può essere compreso sia dal deposito che esporta che da quello che importa, specialmente se viene usato lo schema XML PREMIS.

1.5. Si dovrebbe usare PREMIS?

Dipende. La maggior parte degli staff delle biblioteche, archivi, musei ed altre organizzazioni dei beni culturali non hanno un diretto coinvolgimento nella conservazione digitale. In tal caso, sarebbe abbastanza sapere cos'è PREMIS: un dizionario di dati per i metadati di conservazione.

Se si riveste una qualche responsabilità in qualsiasi aspetto della conservazione digitale probabilmente si troverà utile familiarizzare con PREMIS. Se si è coinvolti con la valutazione o l'implementazione di un deposito digitale o di un sistema di conservazione, si dovrebbe avere una buona comprensione di PREMIS. Si dovrebbe prendere in considerazione la frequenza di un seminario PREMIS che sono periodicamente tenuti dalla PREMIS Maintenance Activity.

Se si è coinvolti in progetti di digitalizzazione si dovrebbe capire se si stanno creando metadati PREMIS per usi successivi. La maggior parte degli elementi PREMIS sono stati pensati per essere forniti automaticamente dall'applicazione del deposito digitale. (Ovviamente questo non significa che le applicazioni attualmente disponibili li forniscano.) In ogni caso, ci sono alcune informazioni che, laddove possibile, dovrebbero essere salvate:

Inibitori. Gli inibitori vengono così definiti come quelle caratteristiche di un oggetto che ne inibiscono l'accesso, l'uso o la migrazione. Gli inibitori includono *password* di protezione e le criptazioni. E' difficile descrivere gli inibitori in modo automatico con un programma, perché potrebbero esserci protezioni anti-analisi dell'oggetto, così se si è a conoscenza della presenza di inibitori sui file, è importante che vengano salvati. PREMIS definisce le unità semantiche per il tipo di inibitore, le azioni che sono inibite e la chiave (*password* o altri meccanismi che scavalchino l'inibitore).

Provenienza. La Provenienza Digitale è la memoria della catena di custodia e i cambiamenti evolutivi di un oggetto digitale. Se un'istituzione crea l'oggetto, le circostanze in cui la creazione avviene è ovviamente un'importante informazione per la sua provenienza. Il nome e la versione dell'applicazione di creazione e la data di creazione può spesso essere estratta dall'intestazione del file (*header*), ma non sempre, pertanto è preferibile che queste informazioni vengano salvate.

Capire PREMIS

PREMIS permette di tenere memoria dei cambiamenti evolutivi nelle informazioni dell'entità Evento, che viene descritta successivamente.

In ogni caso, i tipi di evento PREMIS sono principalmente pensati per descrivere azioni che accadono dopo che qualcosa è stato immesso nel deposito. Per memorizzare gli eventi che accadono prima dell'immissione, come acquisizione e inclusione nel sistema di gestione, potrebbe essere necessario pensare ad appropriati tipi di evento.

Proprietà Significative. Le Proprietà Significative sono le caratteristiche di un oggetto che dovrebbero essere mantenute attraverso le azioni di conservazione. Per esempio, se si possiede un documento, sono soltanto le parole e le immagini che sono critiche, o sono ugualmente importanti i *fonts*, lo sfondo, la formattazione e altre caratteristiche di *'look and feel'*? L'idea di proprietà significative è uno dei concetti più importanti nella conservazione digitale e uno degli ultimi ad essere compreso. Esiste un numero di iniziative dedicate alla migliore modellazione e descrizione delle caratteristiche significative, ma queste sono ancora ad uno stadio prematuro. Cionondimeno, qualsivoglia istituzione che crei o acquisisca materiali digitali per una comunità utente dovrebbe pensare attentamente quali caratteristiche sono importanti per quella comunità e provare a memorizzare quelle informazioni per usi futuri.

Diritti. Le informazioni sui diritti non sono solo per la conservazione, ovviamente, ma sapere cosa si può fare con un oggetto è davvero importante per il processo di conservazione. Si dovrebbe fare attenzione a memorizzare qualsiasi informazione sui diritti, incluse quelle sullo stato del *copyright*, i termini di licenza e i permessi particolari.

2. CONVENZIONI DEL DIZIONARIO DEI DATI

2.1. Unità Semantiche

Il Dizionario dei Dati PREMIS definisce unità semantiche e non elementi di metadati. La distinzione è sottile ma importante. Un'unità semantica è un pezzo di informazione o di conoscenza. Un elemento di metadati è un modo definito di rappresentare quelle informazioni in un record di metadati, in uno schema o in un database. PREMIS non specifica come i metadati dovrebbero essere rappresentati in qualsivoglia sistema, definisce soltanto cosa è necessario sapere per poterli esportare su altri sistemi. Di conseguenza, per essere un purista di PREMIS, si deve pensare in termini di unità semantiche piuttosto astratte. Le unità semantiche PREMIS hanno un *mapping* diretto agli elementi di metadati definiti nello schema XML del PREMIS, e potrebbero avere un *mapping* meno diretto ai metadati di altri schemi.

I nomi delle unità semantiche PREMIS sono stringhe "camel case". Che sta per, parole non separate da spazi ma dalle lettere maiuscole: `objectIdentifier`, `relatedEventIdentification`.

In questo documento sono stampate nel font Verdana.

2.2. Contenitori e Sottounità

Alcune unità semantiche sono definite come contenitori, che sta a significare che non contengono un valore in modo diretto ma sono presenti in qualità di raggruppamento di unità semantiche correlate. Ad esempio, se si deve memorizzare un identificatore nel PREMIS è necessario

Capire PREMIS

specificare che tipo di identificatore è (p.e. "DOI", "ISBN", "assegnato dal sistema locale"). Il contenitore `objectIdentifier` viene usato per raggruppare le due sottounità `objectIdentifierType` e `objectIdentifierValue`.

I contenitori conferiscono al Dizionario dei Dati, una struttura gerarchica che nella versione 2.0 si riflette nella numerazione delle unità semantiche:

1.1 `objectIdentifier` (M, R)

1.1.1 `objectIdentifierType` (M, NR)

1.1.2 `objectIdentifierValue` (M, NR)

L'estratto dal Dizionario dei Dati mostra che l'unità semantica `objectIdentifier` è obbligatoria (M) e ripetibile (R). Dal momento che ci sono unità semantiche definite all'interno di essa, si può dedurre che `objectIdentifier` non contiene un valore in modo diretto ma serve come contenitore degli elementi che la compongono `objectIdentifierType` e `objectIdentifierValue`. Siccome `objectIdentifierType` e `objectIdentifierValue` non sono ripetibili (NR) all'interno del contenitore, è necessario ripetere l'intera struttura di contenimento per memorizzare due diversi tipi di identificatori.

2.3. Contenitori di estensione

Un contenitore di estensione è un tipo speciale di contenitore che non ha sottounità definite al suo interno. E' stato pensato per accogliere l'integrazione di metadati non-PREMIS. In tal modo, PREMIS può essere esteso per includere metadati che sono fuori portata o altrimenti non inclusi nel Dizionario dei Dati. I nomi dei contenitori di estensione hanno come suffisso il termine "Extension".

Per esempio, i metadati tecnici, che appartengono ad un formato specifico, non vengono inclusi nel PREMIS, ma sono informazioni molto importanti per la conservazione digitale. Il contenitore di estensione `objectCharacteristicsExtension` costituisce un luogo dove memorizzare i metadati tecnici definiti da altri dizionari dei dati, come lo standard Z39.87 per descrivere immagini *bitmap*.

Se si ha una qualche familiarità con l'XML sarà ovvio che il Dizionario dei Dati PREMIS è stato pensato per essere XML compatibile. Le unità semantiche PREMIS possono essere implementate come elementi XML; le unità contenitore sono elementi che come contenuto accolgono solo altri elementi, e le unità di estensione sono contenitori per gli elementi definiti in schemi esterni. Maggiori dettagli sulla relazione tra PREMIS e XML vengono forniti in seguito, al paragrafo 5.1.

3. MODELLO DEI DATI PREMIS

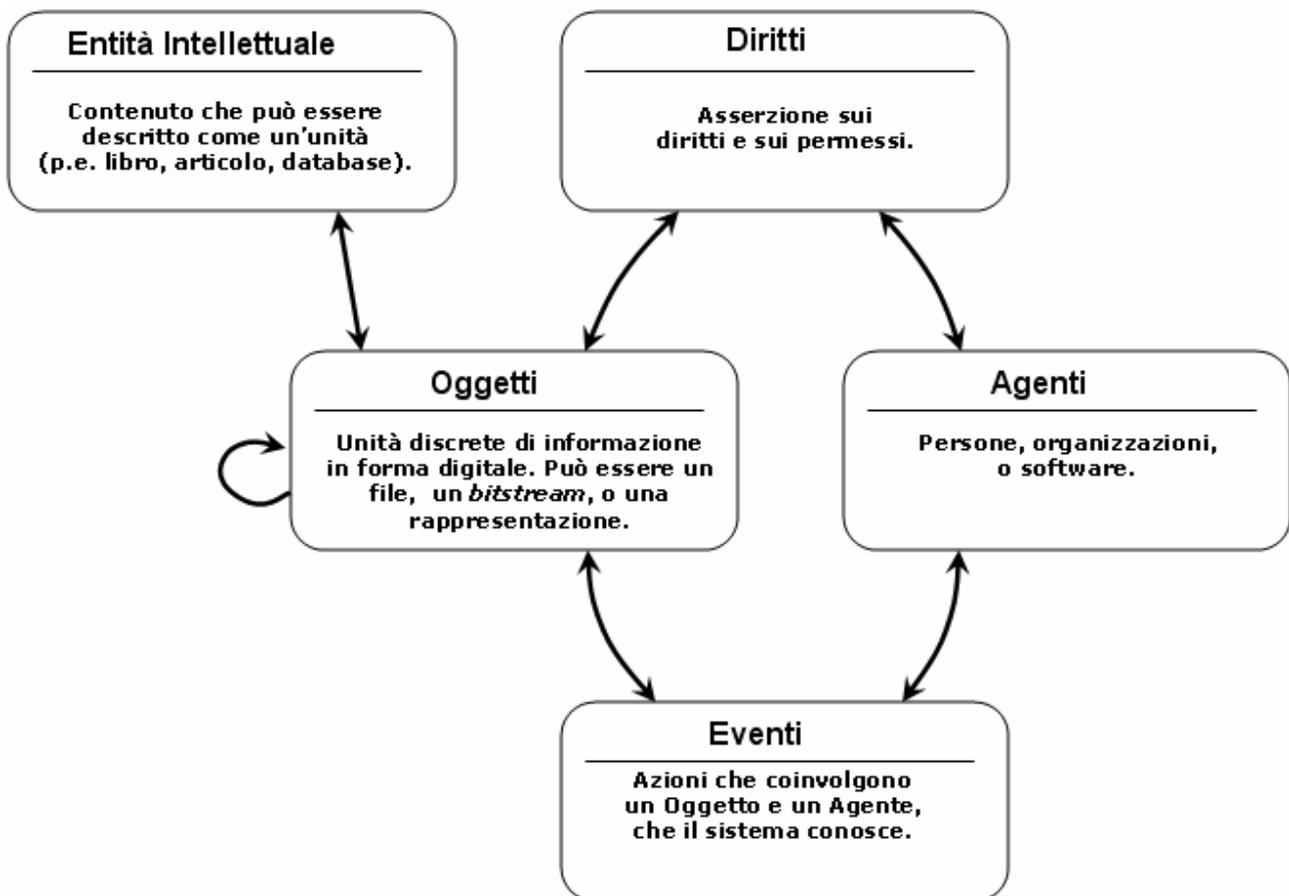


Figure 2. Modello dei dati PREMIS

Uno dei principi più importanti sottesi all'iniziativa PREMIS è l'esigenza di avere un'evidenza immediata di cosa viene descritto. PREMIS definisce cinque tipi di cose (chiamate entità) che possono essere descritte: Entità Intelletuali, Oggetti, Agenti, Eventi e Diritti. Queste definiscono ciò che è chiamato modello dei dati PREMIS e che è mostrato nella Figura 2.

3.1. Entità Intellettuale

Le Entità Intelletuali sono concettuali e potrebbero essere chiamate "entità bibliografiche". PREMIS definisce un'Entità Intellettuale come "un set di contenuto che è considerato un'unità intellettuale singola per scopi gestionali e descrittivi: per esempio, un particolare libro, una mappa, una fotografia o un database". PREMIS in realtà non definisce alcun metadato di riferimento alle Entità Intelletuali, in quanto esistono molti standard di metadati descrittivi tra cui poter scegliere.

PREMIS asserisce che un oggetto in un sistema di conservazione, dovrebbe essere associato, con l'entità intellettuale che rappresenta, includendo un identificatore dell'entità intellettuale nei metadati per l'oggetto. In tal modo, per esempio, se conserviamo una copia di "*Buddhism: The*

Ebook: an Online Introduction” si potrebbe usare l’ ISBN come collegamento all’ Entità Intellettuale nella descrizione PREMIS dell’ ebook.

3.2. Entità Oggetto

Gli Oggetti sono ciò che realmente viene archiviato e gestito nei depositi di conservazione. Buona parte del PREMIS è dedicata a descrivere gli oggetti digitali. Le informazioni che possono essere memorizzate includono:

- un identificatore univoco per l’ oggetto (tipo e valore),
- le informazioni di persistenza come il *checksum* e l’ algoritmo usato per derivarlo,
- la dimensione dell’ oggetto,
- il formato dell’ oggetto, che può essere specificato direttamente o mediante collegamento ad un registro di formati,
- il nome originale dell’ oggetto,
- le informazioni sulla sua creazione,
- le informazioni sui suoi inibitori,
- le informazioni sulle sue proprietà significative,
- le informazioni sul suo ambiente operativo (vedi di seguito),
- dove e su quale supporto è stato memorizzato,
- le informazioni sulla firma digitale,
- le informazioni sulle relazioni con gli altri oggetti e gli altri tipi di entità.

Diverse unità semantiche vengono definite per raccogliere le informazioni sull’ ambiente di un oggetto, che riguardano hardware e software necessario per il suo *rendering* e quali dipendenze esistano con altri oggetti. Per esempio, un file PDF può essere visualizzato da diverse versioni di Adobe Acrobat e Adobe Reader come da altri programmi *open source* e commerciali. Ognuno di questi, viene supportato in vari sistemi operativi e richiede alcuni requisiti minimi di hardware (velocità processore, memoria centrale e spazio disco). Dal momento che Adobe Reader non è un’ applicazione *standalone* ma un *plug-in* per *browser*, presenta anche delle dipendenze da alcune versioni di *browser*; ad esempio, Reader 9 per Mac OS richiede il *browser* Safari versione 2.0.4 o successiva.

Le informazioni sull’ ambiente operativo sono critiche per alcune strategie di conservazione. Semplici informazioni sull’ ambiente, come quelle relative ad un ambiente in cui è risaputo che il contenuto dell’ oggetto sia accessibile, può essere facilmente memorizzato in PREMIS, ma informazioni d’ ambiente più estese possono essere difficili ed onerose da determinare. Quando informazioni dettagliate vengono applicate a classi di oggetti piuttosto che alle stanze locali di oggetti, come nel caso delle informazioni d’ ambiente, esse possono essere utilmente aggregate in registri e condivise da tutti i depositi. PREMIS permette ai depositi di collegare informazioni contenute in registri esterni, in quei casi in cui il riferimento esterno sia preferibile all’ archiviazione a livello locale.

PREMIS in realtà definisce tre diversi tipi di oggetti e richiede che gli implementatori facciano una distinzione tra essi. Ci sono oggetti file, oggetti rappresentazione e oggetti bitstream.

Un oggetto file è quello che sembra essere ovvero un file per computer, come un PDF o un file JPEG.

Un oggetto rappresentazione è l’ insieme di tutti gli oggetti file necessari a fare il *rendering* di un’ Entità Intellettuale. Per esempio, si vuole conservare un pagina Web, diciamo la *home page* di un’ istituzione ad una certa data. E’ probabile che la *home page* sia composta da diversi file: uno o

più file HTML, più qualche immagine GIF o JPEG e probabilmente un piccolo file audio o un'animazione Flash. E' probabile che venga usato anche un foglio di stile per la visualizzazione della *home page*, così se un deposito vuole conservare una *home page* visualizzabile, deve raccogliere le informazioni su tutti questi file e come vengono messi insieme. L'oggetto di rappresentazione permette al deposito non solo di identificare l'insieme dei file correlati, ma anche di descrivere qualsiasi caratteristica della totalità (ad esempio, la pagina Web nella sua interezza) che potrebbe essere diversa da quella delle sue parti.

Gli oggetti *bitstream* sono sottoinsiemi di file. Un oggetto *bitstream* viene definito come dati (*bit*) all'interno di un file che a) per gli obiettivi di conservazione hanno proprietà comuni, e b) non possono stare da soli senza aggiungere un'intestazione file (*header*) o un'altra struttura. Così per esempio, se si ha un file in formato AVI (audio-video *interleaved*), si potrebbe voler distinguere il *bitstream* audio da quello video, descrivendoli come oggetti *bitstream* separati. Alcune unità semantiche definite nel Dizionario dei dati PREMIS sono applicabili a tutti i tre tipi di oggetto. Avere diversi tipi di oggetti costringe a pensare realmente cosa si sta descrivendo e ad essere precisi quanto più possibile, il che è importante per l'elaborazione-automatica.

3.3. Eventi

L'entità EVENTI aggrega informazioni sulle azioni che riguardano gli oggetti del deposito. Un record di eventi accurato ed affidabile è critico per mantenere la provenienza digitale di un oggetto, che è importante per dimostrare l'autenticità dell'oggetto.

Le informazioni che possono essere raccolte sugli eventi includono:

- un identificatore univoco per l'evento (tipo e valore),
- il tipo di evento (creazione, immissione, migrazione, ecc.),
- la data e l'ora in cui è avvenuto un evento,
- una descrizione dettagliata dell'evento,
- l'esito dell'evento in forma codificata,
- una descrizione più dettagliata dell'evento,
- gli agenti coinvolti nell'evento ed il loro ruolo rispetto ad esso,
- gli oggetti coinvolti nell'evento ed il loro ruolo rispetto ad esso.

Ogni sistema di deposito deve prendere decisioni proprie ed appropriate di quali eventi mantenere memoria come parte permanente della storia di un oggetto. PREMIS raccomanda che le azioni che cambiano un oggetto dovrebbero sempre essere memorizzate, e le voci del dizionario dei dati per `eventType` forniscono una "lista iniziale" dei tipi di evento importanti, per incoraggiare i depositi a memorizzare questi eventi in modo consistente.

3.4. Agenti

Gli Agenti sono attori che hanno dei ruoli negli eventi e nelle dichiarazioni sui diritti (3.5 Diritti). Gli Agenti possono essere persone, organizzazioni o applicazioni software. PREMIS definisce solo un numero minimo di unità semantiche necessarie a identificare gli agenti, dal momento che ci sono diversi standard esterni che possono essere usati per memorizzare informazioni più dettagliate. (Come esemplificazione consultare "*Metadata standards and specifications for describing people and their interests*" www.ukoln.ac.uk/metadata/resources/people/.) Un deposito potrebbe scegliere di usare uno standard a parte per memorizzare informazioni aggiuntive sugli agenti, o potrebbe usare l'identificatore dell'agente per puntare alle informazioni memorizzate esternamente.

Capire PREMIS

Il Dizionario dei Dati include:

- un identificatore univoco per l'agente (tipo e valore),
- il nome dell'agente,
- la designazione del tipo di agente (persona, organizzazione, software).

Quando viene fatto riferimento ad un agente, in relazione ad un evento o ad una dichiarazione sui diritti, il ruolo dell'agente dovrebbe essere memorizzato. Ogni agente potrebbe avere diversi ruoli. Per esempio, io potrei essere l'autore ed il detentore dei diritti di un lavoro, l'autore (ma non il detentore dei diritti) di un secondo lavoro e il depositante di un terzo lavoro. Nel modello PREMIS un deposito assegnerebbe un identificatore univoco a me e farebbe riferimento a quell'identificatore in qualsiasi evento o dichiarazione dei diritti in cui io sono agente, insieme al mio ruolo in quel particolare contesto. Un ruolo di agente in relazione ad un evento o di dichiarazione di diritto viene considerata una proprietà dell'entità evento o dell'entità diritti, e non dell'agente stesso

3.5. Diritti

La maggior parte delle strategie di conservazione comporta l'esecuzione di copie identiche e di versioni di derivazione degli oggetti digitali, azioni che per la legge sul *copyright* sono permesse solo ai proprietari dei diritti. L'Entità Diritti aggrega informazioni sui diritti e permessi che sono direttamente pertinenti agli oggetti nel deposito. Ogni dichiarazione sui diritti asserisce due cose: le azioni che il deposito ha diritto ad intraprendere e in base a quale diritto può farle.

Per esempio, un deposito potrebbe avere una versione scannerizzata di un libro che è stato pubblicato nel 1848 e che di conseguenza è di pubblico dominio. Il deposito può fare qualsiasi cosa con la sua versione digitale in base allo stato del *copyright* della risorsa. Un altro deposito potrebbe avere un oggetto copiato da un CD pubblicato, dove la "licenza a strappo" (*shrink-wrap*) permette di fare copie ma ne restringe l'accesso e l'uso.

Le informazioni che possono essere memorizzate in una dichiarazione sui diritti include:

- un identificatore univoco per la dichiarazione sui diritti (tipo e valore),
- se la base su cui rivendicare il diritto è il *copyright*, la licenza o le disposizioni statutarie,
- informazioni più dettagliate sullo stato di *copyright*, termini di licenza o statuto, su come è applicabile
- le azioni che la dichiarazione sui diritti permette,
- ogni restrizione sulle azioni,
- il termine di garanzia, o il periodo di tempo in cui rimane attiva tale dichiarazione,
- gli oggetti a cui si applica la dichiarazione,
- gli agenti coinvolti nella dichiarazione sui diritti e i loro ruoli.

La maggior parte delle informazioni è stata pensata per essere azionabile (che sta per, memorizzata in una forma controllata che può essere azionata da un programma di computer). La dichiarazione sui diritti di PREMIS è un'asserzione sui diritti, non un *record* di informazioni dal quale è possibile determinare i diritti. Che significa che PREMIS non definisce il tipo di informazioni di dettaglio su autori, data e luogo di pubblicazione, e notifica del *copyright* che vengono definite nella specifica di metadati di *Copyright (copyrightMD)* della Biblioteca Digitale della California (www.cdlib.org/inside/projects/rights/schema/). Lo scopo del *copyrightMD* è di aiutare gli esseri umani a fare delle determinazioni sui diritti su base evolutiva, mentre lo scopo dell'entità diritti di PREMIS è quello di fornire informazioni azionabili per i sistemi di conservazione digitale.

4. II DIZIONARIO DEI DATI

4.1. Esempio di Unità Semantica Semplice come voce del Dizionario dei Dati

La Figura 3 mostra la voce del Dizionario dei Dati per l'unità semantica *size* (dimensione), che è un componente o sottounità del contenitore chiamato *objectCharacteristics*. La dimensione non ha sottounità. La voce del Dizionario dei Dati include la definizione dell'elemento e la motivazione per cui è stata inclusa nel *core* dei metadati PREMIS, come anche esempi e note su come il valore potrebbe essere ottenuto ed usato. Questi sono tutti pensati per aiutare gli implementatori ad usare l'elemento in modo proprio.

Le due righe categoria dell'oggetto e applicabilità sono usate insieme per mostrare quando le unità semantiche sono appropriate per la descrizione delle rappresentazioni, dei file e/o dei *bitstream*. Qui la dimensione viene mostrata di sola pertinenza dei file e dei *bitstream*. Infine, ci sono un insieme di regole per l'uso: vincoli sui dati, ripetibilità e obbligatorietà.

I vincoli sui dati specificano restrizioni sui valori che l'unità semantica può assumere. In questo esempio, il valore della dimensione deve essere un intero. Un altro vincolo comune sui dati è che il valore deve essere preso da un vocabolario controllato. Qualche volta i termini nel vocabolario vengono specificati nel Dizionario dei Dati e qualche volta no, ma in ogni caso il nome del vocabolario usato deve essere memorizzato. Non esistono unità semantiche nel Dizionario dei Dati per i nomi dei vocabolari, ma la *PREMIS Maintenance Activity* sta sviluppando una soluzione attraverso uno schema XML.

La ripetibilità indica se un'unità semantica può essere ripetuta o meno.

L'obbligatorietà indica se il valore dell'unità semantica è obbligatorio o opzionale. L'obbligatorietà potenzialmente può far confondere, in quanto PREMIS stabilisce chiaramente che non viene richiesto ad un deposito di memorizzare un'informazione in particolare. Un'unità semantica che è obbligatoria non deve essere memorizzata ed archiviata all'interno del deposito. Ma, il deposito deve essere in grado di generare il valore dell'unità semantica quando necessario, come nello scambio con altri depositi. Per esempio, nell'improbabile eventualità che un deposito abbia archiviato solo immagini in formato TIFF 6.0, esso non dovrebbe aver memorizzato le informazioni sul formato per ogni singolo oggetto. Pur tuttavia, la gestione del deposito dovrebbe essere a conoscenza che i propri oggetti file sono TIFF 6.0, e potrebbe fornire quelle informazioni, nel caso si renda necessario. (sezione 5.2. Conformità PREMIS.)

Semantic unit (Unità semantica)	1.5.3 size		
Semantic components (Unità semantiche componenti)	None		
Definition (Definizione)	La dimensione in <i>byte</i> di un file o un bitstream in possesso del deposito.		
Razionale (Motivazioni)	La dimensione è utile per assicurare che si possa risalire ad un corretto numero di <i>byte</i> dallo <i>storage</i> (unità generica di memorizzazione) e che un applicazione abbia abbastanza spazio per spostare o elaborare i file. Potrebbe inoltre essere usato per fatturare lo spazio di storage.		
Data constraint (Vincolo sui dati)	Intero		
Object category (Categoria dell'Oggetto)	Representation	File	Bitstream
Applicability (Applicabilità)	Non applicabile	Applicabile	Applicabile
Examples	2038937		
Repeatability (Ripetibilità)	Non ripetibile		Non ripetibile
Obligation (Obbligatorietà)	Opzionale		Opzionale
Creation / Maintenance notes (Note sulla creazione e manutenzione)	Ottenuto automaticamente dal deposito.		
Usage notes (Note sull'uso)	Definire questa unità semantica come dimensione in <i>byte</i> , rende inutile la memorizzazione dell'unità di misura. In ogni caso, a scopo di scambio l'unità di misura dovrebbe essere stabilita o resa intelligibile a entrambi i partner.		

Figura 3: Estratto dal Dizionario dei Dati dell'unità semantica size.

4.2. Esempio di unità contenitore come voce del Dizionario dei Dati

La Figura 4 mostra l'inizio della voce del Dizionario dei Dati `objectCharacteristics`, l'unità contenitore di `size`. Si può dire che è un contenitore perché ha componenti semantici e la restrizione sui dati (`Data Constraint`) indica che è un "contenitore". Si noti che i componenti semantici inclusi possono essere unitari come `size`, o contenitori essi stessi, come `format`.

Semantic unit (Unità semantica)	1.5 objectCharacteristics		
Semantic components (Unità semantiche componenti)	1.5.1 compositionLevel 1.5.2 fixity 1.5.3 size 1.5.4 format 1.5.5 creatingApplication 1.5.6 inhibitors 1.5.7 objectCharacteristicsExtension		
Definition (Definizione)	Proprietà tecniche di un file o di un <i>bitstream</i> che sono applicabili a tutti o alla maggior parte dei formati.		
Razionale (Motivazioni)	Esistono alcune importanti proprietà tecniche che si applicano agli oggetti di qualsiasi formato. Definizioni dettagliate sulle proprietà specifiche di un formato non sono di competenza di questo Dizionario dei Dati, sebbene alcune di queste proprietà potrebbero essere incluse in objectCharacteristicsExtension		
Data constraint (Vincoli sui dati)	Contenitore		
Object category (Categoria dell'Oggetto)	Representation	File	Bitstream
Applicability (Applicabilità)	Non applicabile	Applicabile	Applicabile
Repeatability (Ripetibilità)	Ripetibile		Ripetibile
Obligation (Obbligatorietà)	Obbligatorio		Obbligatorio
Usage notes (Note sull'uso)	Le unità semantiche incluse in <i>objectCharacteristics</i> dovrebbero essere considerate con un insieme di informazioni che riguardano un oggetto singolo e ad un singolo <i>compositionLevel</i> . Le caratteristiche dell'oggetto potrebbero essere ripetute		

Figura 4: Estratto dal Dizionario dei Dati dell'unità semantica objectCharacteristics.

5. USO DI PREMIS

5.1. PREMIS in XML

Ci si aspetta che PREMIS venga rappresentato in XML (sebbene non sia un requisito), nel caso in cui venga usato per lo scambio tra depositi. La *PREMIS Maintenance Activity* fornisce uno schema XML che corrisponde direttamente al Dizionario dei Dati al fine di fornire una descrizione diretta degli oggetti, degli eventi, degli agenti e dei diritti. La Figura 5 mostra un estratto dei metadati PREMIS che usano lo schema XML.

In realtà, molti sistemi di conservazione usano formati XML per l'importazione e l'esportazione dei dati. Molti usano METS (Metadata Encoding and Transmission Standard), altro standard sostenuto dalla Library of Congress, come contenitore XML per mettere insieme diversi tipi di metadati. E' possibile inserire PREMIS all'interno di METS, ma questo non è interamente praticabile per due motivi. In primo luogo, METS, suddivide le informazioni in diverse sezioni in base al tipo di metadati se essi sono tecnici, se riguardano i diritti o descrivono la provenienza. Lo schema PREMIS, che rispetta il Dizionario dei Dati, comprende sezioni per oggetti, diritti, eventi e agenti. Esistono una certa corrispondenza tra le due strutture ma non è perfetta, specialmente per le informazioni sull'agente. In secondo luogo, PREMIS e METS hanno alcune sovrapposizioni; per esempio, ognuno di essi definisce un elemento per memorizzare l'impronta dei file. Se vengono usati insieme, è necessario decidere se memorizzare tali elementi di sovrapposizione nelle sezioni di PREMIS, in quelle di METS o in entrambe.

Ovviamente, se ogni deposito di conservazione prende una decisione propria ci potrebbe essere una grande varietà nel come i dati vengono rappresentati, ostacolandone l'interoperabilità. Inoltre, esistono diversi tentativi per definire le buone pratiche nell'uso di PREMIS e METS insieme. Per documentarsi sugli sviluppi consultare il sito web della *PREMIS Maintenance Activity*.

```
<event>
  <eventIdentifier>
    <eventIdentifierType>DAITSS</eventIdentifierType>
    <eventIdentifierValue>10012</eventIdentifierValue>
  </eventIdentifier>
  <eventType>Validation</eventType>
  <eventDateTime>2008-05-06T10:40:22-04:00</eventDateTime>
  <eventOutcomeInformation>
    <eventOutcome>Invalid</eventOutcome>
    <eventOutcomeDetail>ill-formed DateTime
      value</eventOutcomeDetail>
  </eventOutcomeInformation>
</event>
```

Figura 5: Frammento di PREMIS codificato in XML.

5.2. Conformità PREMIS

Il Dizionario dei dati PREMIS contiene una sezione su cosa significhi per un deposito essere conforme al PREMIS. Essenzialmente, esistono tre requisiti da acquisire:

- 1) Se il deposito implementa (archivia o esporta) un elemento di dati che si presuppone sia una unità semantica PREMIS, l'elemento di dati dovrebbe avere la stessa definizione, gli stessi vincoli sui dati e la stessa applicabilità, come unità semantica definita in PREMIS.
- 2) Se il deposito implementa un'unità semantica PREMIS, la sua ripetibilità e obbligatorietà può essere più stringente ma non più flessibile rispetto a ciò che PREMIS richiede. Che sta a significare che un'unità semantica ripetibile può essere implementata come non-ripetibile ma non vice-versa, e un elemento obbligatorio non può essere reso opzionale.
- 3) Se il deposito esporta informazioni affinché un altro deposito le usi, deve fornire le informazioni per tutte quelle unità semantiche che sono obbligatorie nel Dizionario dei Dati. Esiste comunque una qualche flessibilità in questo, dal momento che ai depositi non viene richiesto di supportare unità semantiche obbligatorie per quei tipi di entità che essi non gestiscono. In altre parole, un

Capire PREMIS

deposito è libero di gestire o meno l'entità Agenti di PREMIS, ma se la gestisce di conseguenza agentIdentifier è obbligatorio. In modo simile, un deposito in particolare potrebbe non gestire gli oggetti *bitstream*, in tal caso non deve fornire gli identificatori obbligatori del *bitstream*.

Tali requisiti dovrebbero essere considerati nel contesto di un numero di cose che NON sono richieste per la conformità. Come notato sopra, ad un deposito non viene richiesto di supportare tutti i tipi di entità definite nel modello dei dati PREMIS. Come anche non viene richiesto di archiviare i metadati internamente per usare i nomi delle unità semantiche PREMIS, o per usare i valori che rispettano i vincoli PREMIS sui dati. In altre parole, non ha importanza come un deposito "è a conoscenza" di un valore PREMIS – mediante il suo immagazzinamento con lo stesso nome o con un nome diverso, mappandolo da un altro valore, puntando ad un registro, per deduzione, per *default*, o per altro mezzo. Così se il deposito, quando si renda necessario, può fornire un PREMIS di buona qualità, esso è conforme.

In futuro la *PREMIS Maintenance Activity* potrebbe riconsiderare i requisiti di conformità e renderli più stringenti, ma con la versione 2.0 è abbastanza facile per qualsiasi deposito essere "conforme al PREMIS". D'altra parte, più unità semantiche gestisce, e maggiore è il valore che acquisisce nell'uso di PREMIS. Il Dizionario dei Dati PREMIS è stato sviluppato per identificare il *core* delle informazioni, di cui la maggior parte dei depositi avrà bisogno per conservare il contenuto digitale a lungo termine. Il responsabile di un deposito di conservazione dovrebbe guardare con attenzione al PREMIS e avere buoni motivi per non implementare qualsiasi parte del Dizionario dei Dati.

6. PER MAGGIORI INFORMAZIONI

Il sito Web della *PREMIS Maintenance Activity* (www.loc.gov/standards/premis/) contiene informazioni per le diverse esigenze, come i collegamenti al PREMIS Implementers Group (PIG), ai tutorial, agli schemi, agli strumenti e alle notizie. Contiene anche una sezione "Resources: Articles and Presentations" che rimanda alla letteratura riguardante PREMIS e gli argomenti correlati. Alcune di quelle più utili ad un pubblico generico vengono qui elencate:

Sui metadati di conservazione in generale:

"Preservation Metadata" (*PDF:209K/21pp.*)

Brian Lavoie (OCLC) and Richard Gartner (Oxford)

Joint report of OCLC, Oxford Library Services, and the Digital Preservation Coalition. Published as DPC Technology Watch Report No. 05-01: September 2005. <http://www.dpconline.org/docs/reports/dpctw05-01.pdf>

Sulla differenza tra la versione 1.0 e la versione 2.0 PREMIS:

"PREMIS with a Fresh Coat of Paint: Highlights from the Revision of the PREMIS Data Dictionary for Preservation Metadata"

Brian Lavoie, *D-Lib Magazine*, May/June 2008.

<http://www.dlib.org/dlib/may08/lavoie/05lavoie.html>

Sull'uso di PREMIS e METS insieme:

"Battle of the Buzzwords: Flexibility vs. Interoperability When Implementing PREMIS with METS"

Rebecca Guenther, *D-Lib Magazine*, July/August 2008.

<http://www.dlib.org/dlib/july08/guenther/07guenther.html>

Appendix A: Lista di tutte le Unità Semantiche PREMIS

Dal *Dizionario dei dati PREMIS per i Metadati di Conservazione* versione 2.0.

Come leggere questa lista

1.2 objectCategory (M, NR)

L'unità semantica objectCategory è Obbligatoria (M) e Non Ripetibile(NR) . Si applica a tutti i tipi di oggetti (file, rappresentazione, *bitstream*).

1.3 preservationLevel (O, R) [representation, file]

1.3.1 preservationLevelValue (M, NR) [representation, file]

L'unità semantica preservationLevel è Opzionale(O) e Ripetibile(R) e può essere usata soltanto per gli oggetti rappresentazione e per gli oggetti file. E' un'unità contenitore perchè contiene almeno un'unità semantica subordinata, preservationLevelValue, che pure si applica soltanto agli oggetti rappresentazione e oggetti file.

Unità Semantiche per l'Entità Oggetto

1.1 objectIdentifier (identificatore oggetto) (M, R)

1.1.1 objectIdentifierType (tipo identificatore oggetto) (M, NR)

1.1.2 objectIdentifierValue (valore identificatore oggetto) (M, NR)

1.2 objectCategory (categoria oggetto) (M, NR)

1.3 preservationLevel (livello di conservazione) (O, R) [representation, file]

1.3.1 preservationLevelValue (valore livello di conservazione) (M, NR) [representation, file]

1.3.2 preservationLevelRole (O, NR) [representation, file]

1.3.3 preservationLevelRationale (O, R) [representation, file]

1.3.4 preservationLevelDateAssigned (O, NR) [representation, file]

1.4 significantProperties (proprietà significative) (O, R)

1.4.1 significantPropertiesType (tipo proprietà significative) (O, NR)

1.4.2 significantPropertiesValue (valore proprietà significative) (O, NR)

1.4.3 significantPropertiesExtension (estensione proprietà significative) (O, R)

1.5 objectCharacteristics (caratteristiche oggetto) (M, R) [file, bitstream]

1.5.1 compositionLevel (livello di composizione) (M, NR) [file, bitstream]

1.5.2 fixity (persistenza) (O, R) [file, bitstream]

1.5.2.1 messageDigestAlgorithm (algoritmo impronta) (M, NR) [file, bitstream]

1.5.2.2 messageDigest (impronta) (M, NR) [file, bitstream]

1.5.2.3 messageDigestOriginator (originatore impronta) (O, NR) [file, bitstream]

1.5.3 size (dimensione) (O, NR) [file, bitstream]

1.5.4 format (formato) (M, R) [file, bitstream]

1.5.4.1 formatDesignation (designazione formato) (O, NR) [file, bitstream]

1.5.4.1.1 formatName (nome formato) (M, NR) [file, bitstream]

1.5.4.1.2 formatVersion (versione formato) (O, NR) [file, bitstream]

1.5.4.2 formatRegistry (registro formato) (O, NR) [file, bitstream]

1.5.4.2.1 formatRegistryName (nome registro formato) (M, NR) [file, bitstream]

1.5.4.2.2 formatRegistryKey (chiave registro formato) (M, NR) [file, bitstream]

1.5.4.2.3 formatRegistryRole (ruolo registro formato) (O, NR) [file, bitstream]

1.5.4.3 formatNote (note formato) (O, R) [file, bitstream]

1.5.5 creatingApplication (applicazione di creazione) (O, R) [file, bitstream]

1.5.5.1 creatingApplicationName (nome applicazione di creazione) (O, NR) [file, bitstream]

1.5.5.2 creatingApplicationVersion (versione applicazione di creazione) (O, NR) [file, bitstream]

1.5.5.3 dateCreatedByApplication (data in cui è stato creato dall'applicazione) (O, NR) [file, bitstream]

1.5.5.4 creatingApplicationExtension (estensione applicazione di creazione) (O, R) [file, bitstream]

1.5.6 inhibitors (inibitori) (O, R) [file, bitstream]

1.5.6.1 inhibitorType (tipo inibitore) (M, NR) [file, bitstream]

1.5.6.2 inhibitorTarget (obiettivo inibitore) (O, R) [file, bitstream]

1.5.6.3 inhibitorKey (chiave inibitore) (O, NR) [file, bitstream]

1.5.7 objectCharacteristicsExtension (estensione caratteristiche oggetto) (O, R) [file, bitstream]

1.6 originalName (nome originario) (O, NR) [representation, file]

1.7 storage (storage) (M, R) [file, bitstream]

1.7.1 contentLocation (collocazione contenuto) (O, NR) [file, bitstream]

Capire PREMIS

- 1.7.1.1 contentLocationType (tipo collocazione contenuto) (M, NR) [file, bitstream]
- 1.7.1.2 contentLocationValue (valore collocazione contenuto) (M, NR) [file, bitstream]
- 1.7.2 storageMedium (supporto storage) (O, NR) [file, bitstream]
- 1.8 environment (ambiente) (O, R)
 - 1.8.1 environmentCharacteristic (caratteristica ambiente) (O, NR)
 - 1.8.2 environmentPurpose (finalità ambiente) (O, R)
 - 1.8.3 environmentNote (note ambiente) (O, R)
 - 1.8.4 dependency (dipendenza) (O, R)
 - 1.8.4.1 dependencyName (nome dipendenza) (O, R)
 - 1.8.4.2 dependencyIdentifier (identificatore dipendenza) (O, R)
 - 1.8.4.2.1 dependencyIdentifierType (tipo identificatore dipendenza) (M, NR)
 - 1.8.4.2.2 dependencyIdentifierValue (valore identificatore dipendenza) (M, NR)
 - 1.8.5 software (O, R)
 - 1.8.5.1 swName (nome sw) (M, NR)
 - 1.8.5.2 swVersion (versione sw) (O, NR)
 - 1.8.5.3 swType (tipo sw) (M, NR)
 - 1.8.5.4 swOtherInformation (altre informazioni sw) (O, R)
 - 1.8.5.5 swDependency (dipendenza sw) (O, R)
 - 1.8.6 hardware (O, R)
 - 1.8.6.1 hwName (nome hw) (M, NR)
 - 1.8.6.2 hwType (tipo hw) (M, NR)
 - 1.8.6.3 hwOtherInformation (altre informazioni hw) (O, R)
 - 1.8.7 environmentExtension (estensione ambiente) (O, R)
- 1.9 signatureInformation (informazione firma) (O, R) [file, bitstream]
 - 1.9.1 signature (firma) (O, R)
 - 1.9.1.1 signatureEncoding (codifica firma) (M, NR) [file, bitstream]
 - 1.9.1.2 signer (firmatario) (O, NR) [file, bitstream]
 - 1.9.1.3 signatureMethod (metodo firma) (M, NR) [file, bitstream]
 - 1.9.1.4 signatureValue (valore firma) (M, NR) [file, bitstream]
 - 1.9.1.5 signatureValidationRules (regole di validazione firma) (M, NR) [file, bitstream]
 - 1.9.1.6 signatureProperties (proprietà firma) (O, R) [file, bitstream]
 - 1.9.1.7 keyInformation (informazioni chiave) (O, NR) [file, bitstream]
 - 1.9.2 signatureInformationExtension (estensione informazioni firma) (O, R) [file, bitstream]
- 1.10 relationship (relazione) (O, R)
 - 1.10.1 relationshipType (tipo relazione) (M, NR)
 - 1.10.2 relationshipSubType (sottotipo relazione) (M, NR)
 - 1.10.3 relatedObjectIdentification (identificazione oggetto correlato) (M, R)
 - 1.10.3.1 relatedObjectIdentifierType (tipo identificatore oggetto correlato) (M, NR)
 - 1.10.3.2 relatedObjectIdentifierValue (valore identificatore oggetto correlato) (M, NR)
 - 1.10.3.3 relatedObjectSequence (sequenza oggetto correlato) (O, NR)
 - 1.10.4 relatedEventIdentification (identificazione evento correlato) (O, R)
 - 1.10.4.1 relatedEventIdentifierType (tipo identificatore evento correlato) (M, NR)
 - 1.10.4.2 relatedEventIdentifierValue (valore identificatore evento correlato) (M, NR)
 - 1.10.4.3 relatedEventSequence (sequenza evento correlato) (O, NR)
- 1.11 linkingEventIdentifier (collegamento identificatore evento) (O, R)
 - 1.11.1 linkingEventIdentifierType (collegamento tipo identificatore evento) (M, NR)
 - 1.11.2 linkingEventIdentifierValue (collegamento valore identificatore evento) (M, NR)
- 1.12 linkingIntellectualEntityIdentifier (collegamento identificatore entità intellettuale) (O, R)
 - 1.12.1 linkingIntellectualEntityIdentifierType (collegamento tipo identificatore entità intellettuale) (M, NR)
 - 1.12.2 linkingIntellectualEntityIdentifierValue (collegamento valore identificatore entità intellettuale) (M, NR)
- 1.13 linkingRightsStatementIdentifier (collegamento identificatore dichiarazione diritti) (O, R)
 - 1.13.1 linkingRightsStatementIdentifierType (collegamento tipo identificatore dichiarazione diritti) (M, NR)
 - 1.13.2 linkingRightsStatementIdentifierValue (collegamento valore identificatore dichiarazione diritti) (M, NR)

Unità Semantiche per l'Entità Evento

- 2.1 eventIdentifier (identificatore evento) (M, NR)
 - 2.1.1 eventIdentifierType (tipo identificatore evento) (M, NR)
 - 2.1.2 eventIdentifierValue (valore identificatore evento) (M, NR)
- 2.2 eventType (tipo evento) (M, NR)
- 2.3 eventDateTime (data e ora evento) (M, NR)
- 2.4 eventDetail (dettaglio evento) (O, NR)
- 2.5 eventOutcomeInformation (informazioni risultato evento) (O, R)
 - 2.5.1 eventOutcome (risultato evento) (O, NR)
 - 2.5.2 eventOutcomeDetail (dettaglio risultato evento) (O, R)
 - 2.5.2.1 eventOutcomeDetailNote (note dettaglio risultato evento) (O, NR)
 - 2.5.2.2 eventOutcomeDetailExtension (estensione dettaglio risultato evento) (O, R)
- 2.6 linkingAgentIdentifier (collegamento identificatore agente) (O, R)
 - 2.6.1 linkingAgentIdentifierType (tipo collegamento identificatore agente) (M, NR)
 - 2.6.2 linkingAgentIdentifierValue (valore collegamento identificatore agente) (M, NR)
 - 2.6.3 linkingAgentRole (collegamento ruolo agente) (O, R)
- 2.7 linkingObjectIdentifier (collegamento identificatore oggetto) (O, R)
 - 2.7.1 linkingObjectIdentifierType (tipo collegamento identificatore oggetto) (M, NR)
 - 2.7.2 linkingObjectIdentifierValue (valore collegamento identificatore oggetto) (M, NR)
 - 2.7.3 linkingObjectRole (collegamento funzione oggetto) (O, R)

Unità Semantiche per l'Entità Agente

- 3.1 agentIdentifier (identificatore agente) (R, M)
 - 3.1.1 agentIdentifierType (tipo identificatore agente) (M, NR)
 - 3.1.2 agentIdentifierValue (valore identificatore agente) (M, NR)
- 3.2 agentName (nome agente) (O, R)
- 3.3 agentType (tipo agente) (O, NR)

Unità Semantiche per l'Entità Diritti

- 4.1 rightsStatement (dichiarazione diritti) (O, R)
 - 4.1.1 rightsStatementIdentifier (identificatore dichiarazione diritti) (M, NR)
 - 4.1.1.1 rightsStatementIdentifierType (tipo identificatore dichiarazione diritti) (M, NR)
 - 4.1.1.2 rightsStatementIdentifierValue (valore identificatore dichiarazione diritti) (M, NR)
 - 4.1.2 rightsBasis (base dei diritti) (M, NR)
 - 4.1.3 copyrightInformation (informazioni copyright) (O, NR)
 - 4.1.3.1 copyrightStatus (status copyright) (M, NR)
 - 4.1.3.2 copyrightJurisdiction (giurisdizione copyright) (M, NR)
 - 4.1.3.3 copyrightStatusDeterminationDate (data di determinazione status copyright) (O, NR)
 - 4.1.3.4 copyrightNote (note copyright) (O, R)
 - 4.1.4 licenseInformation (informazioni licenza) (O, NR)
 - 4.1.4.1 licenseIdentifier (identificatore licenza) (O, NR)
 - 4.1.4.1.1 licenseIdentifierType (tipo identificatore licenza) (M, NR)
 - 4.1.4.1.2 licenseIdentifierValue (valore identificatore licenza) (M, NR)
 - 4.1.4.2 licenseTerms (termini di licenza) (O, NR)
 - 4.1.4.3 licenseNote (note licenza) (O, R)
 - 4.1.5 statuteInformation (informazioni statuto) (O, R)
 - 4.1.5.1 statuteJurisdiction (giurisdizione statuto) (M, NR)
 - 4.1.5.2 statuteCitation (citazione statuto) (M, NR)
 - 4.1.5.3 statuteInformationDeterminationDate (data di determinazione delle informazioni statuto) (O, NR)
 - 4.1.5.4 statuteNote (note statuto) (O, R)
 - 4.1.6 rightsGranted (diritti garantiti) (O, R)
 - 4.1.6.1 act (atto) (M, NR)
 - 4.1.6.2 restriction (restrizioni) (O, R)
 - 4.1.6.3 termOfGrant (termini di garanzia) (M, NR)
 - 4.1.6.3.1 startDate (data di inizio) (M, NR)
 - 4.1.6.3.2 endDate (data di fine) (O, NR)
 - 4.1.6.4 rightsGrantedNote (note diritti garantiti) (O, R)
 - 4.1.7 linkingObjectIdentifier (collegamento identificatore oggetto) (O, R)
 - 4.1.7.1 linkingObjectIdentifierType (collegamento tipo identificatore oggetto) (M, NR)
 - 4.1.7.2 linkingObjectIdentifierValue (collegamento valore identificatore oggetto) (M, NR)
 - 4.1.8 linkingAgentIdentifier (collegamento identificatore agente) (O, R)
 - 4.1.8.1 linkingAgentIdentifierType (collegamento tipo identificatore agente) (M, NR)
 - 4.1.8.2 linkingAgentIdentifierValue (collegamento valore identificatore agente) (M, NR)
 - 4.1.8.3 linkingAgentRole (collegamento ruolo agente) (M, NR)
- 4.2 rightsExtension (estensione diritti) (O, R)

Appendice B: Esempio di Oggetto

Questo esempio è una versione semplificata di quella preparata dalla Library of Congress (LC) per un seminario PREMIS. Mostra le unità semantiche PREMIS ed i valori usati per descrivere un'immagine TIFF.

La LC usa degli *handles* (identificatori creati e gestiti da un'applicazione chiamata *The Handle System*) per gli oggetti file contenuti nel deposito (par. 1.1 *objectIdentifier* e 1.10.3 *relatedObjectIdentification*). L'oggetto da descrivere viene individuato all'interno del deposito dal suo identificatore univoco (1.1 *objectIdentifier*) ma il suo nome originario prima dell'immissione era "001h.tif" (1.6 *originalName*). E' noto che esso è un file grazie alla sua categorizzazione come oggetto (1.2 *objectCategory*).

Il file ha un *checksum* di tipo MD5 calcolato dal sistema di gestione dei contenuti digitali della LC (1.5.2 *fixity*). Le informazioni sul suo formato vengono fornite in due modi, mediante l'indicazione come tipo MIME e la sua versione (1.5.4.1 *formatDesignation*) e puntando al registro dei formati PRONOM che mantiene una copia della specifica TIFF.

La LC prende in considerazione il fatto che ci siano due applicazioni legate all'azione di creazione. Qualcosa che probabilmente ha creato un file RAW senza nome con uno scanner usando l'applicazione ScandAll e che lo ha convertito in TIFF usando Adobe Photoshop (1.5.5 *creatingApplication*). Non esistono criptazioni o altre forme di inibitori, che LC ha ritenuto fossero importanti da indicare esplicitamente (1.5.6 *inhibitors*).

Il file risiede su un disco alla posizione indicata mediante il percorso delle *directory* (1.7.1 *contentLocation*); l'unità disco viene identificata molto specificamente, in modo tale che presumibilmente, il gestore sia a conoscenza di quando i dati richiedano la migrazione su un'unità più recente (1.7.2 *storageMedium*).

La LC ha ritenuto opportuno indicare una raccomandazione per l'ambiente operativo per la visualizzazione di detto file: usare Adobe Acrobat version 5.0 per computer Intel x86 con Windows XP. La velocità del processore deve essere almeno di 60 MHz e dovrebbe avere almeno 32 MB di memoria, ma si raccomandano 64 MB (1.8 *environment*).

Il file TIFF è parte di una rappresentazione con l'identificatore *R200802948* assegnato internamente.

Questa informazione viene fornita nel terzo blocco delle informazioni relazionali (1.10 *relationship*). Esistono altri due file in questa rappresentazione, identificati attraverso *handles*.

Siamo a conoscenza del fatto che sono oggetti file perchè il tipo di relazione viene descritta come "*sibling*" ovvero allo stesso livello. E' consigliabile consultare la descrizione dell'oggetto in *loc.music/gottlieb.09602* e *loc.music/gottlieb.mets* per maggiori informazioni, per esempio su che tipo di file sono.

Non abbiamo nessuna informazione descrittiva sull'entità intellettuale a cui si ha accesso attraverso questa rappresentazione. Tuttavia, è possibile risalire ad essa da un link che è stato inserito come suo identificatore, un LCCN (1.12 *linkingIntellectualEntityIdentifier*). Presumibilmente, in corrispondenza di tale LCCN, si troverà una descrizione bibliografica dell'entità intellettuale (sebbene non in questo esempio, perchè LCCN è inventato).

Unità semantica	Valore
1.1 objectIdentifier (identificatore oggetto)	
1.1.1 objectIdentifierType (tipo identificatore oggetto)	hdl
1.1.2 objectIdentifierValue (valore identificatore oggetto)	loc.music/gottlieb.09601
1.2 objectCategory (categoria oggetto)	file
1.3 preservationLevel (livello di conservazione)	
1.3.1 preservationLevelValue (valore del livello di conservazione)	Completa
1.5 objectCharacteristics (caratteristiche dell'oggetto)	
1.5.1 compositionLevel (livello di composizione)	0
1.5.2 fixity (persistenza)	
1.5.2.1 messageDigestAlgorithm (algoritmo dell'impronta)	MD5
1.5.2.2 messageDigest (impronta)	36b03197ad066cd719906c55eb68ab8d
1.5.2.3 messageDigestOriginator (creatore dell'impronta)	LocalDCMS
1.5.3 size (dimensione)	20800896
1.5.4 format (formato)	
1.5.4.1 formatDesignation (designazione del formato)	
1.5.4.1.1 formatName (nome del formato)	image/tiff
1.5.4.1.2 formatVersion (versione del formato)	6.0
1.5.4.2 formatRegistry (registro dei formati)	
1.5.4.2.1 formatRegistryName (nome del registro dei formati)	PRONOM
1.5.4.2.2 formatRegistryKey (chiave del registro dei formati)	fmt/10
1.5.4.2.3 formatRegistryRole (funzione del registro dei formati)	specificata
1.5.5 creatingApplication (applicazione di creazione)	
1.5.5.1 creatingApplicationName (nome applicazione di creazione)	ScandAll 21
1.5.5.2 creatingApplicationVersion (versione applicazione di creazione)	4.1.4
1.5.5.3 dateCreatedByApplication (data in cui l'applicazione l'ha creato)	1998-10-30
1.5.5 creatingApplication (applicazione di creazione)	
1.5.5.1 creatingApplicationName (nome applicazione di creazione)	Adobe Photoshop
1.5.5.2 creatingApplicationVersion (versione applicazione di creazione)	CS2
1.5.5.3 dateCreatedByApplication (data di creazione dall'applicazione)	1998-10-30T08:29:02
1.5.6 inhibitors (inibitori)	
1.5.6.1 inhibitorType (tipo di inibitore)	nessuno
1.6 originalName (nome originale)	001h.tif
1.7 storage (storage)	
1.7.1 contentLocation (localizzazione del contenuto)	
1.7.1.1 contentLocationType (tipo localizzazione del contenuto)	percorso del file
1.7.1.2 contentLocationValue (valore localizzazione del contenuto)	amserver/
1.7.2 storageMedium (supporto di memorizzazione)	IBM DS4000 System 1740-5208
1.8 environment (ambiente)	
1.8.1 environmentCharacteristic (caratteristiche dell'ambiente)	raccomandato
1.8.2 environmentPurpose (finalità ambiente)	recupero
1.8.2 environmentPurpose (finalità ambiente)	editare
1.8.5 software	
1.8.5.1 swName (nome software)	Windows
1.8.5.2 swVersion (versione software)	XP
1.8.5.3 swType (tipo software)	sistema operativo
1.8.6 hardware	
1.8.6.1 hwName (nome hardware)	Intel x86
1.8.6.2 hwType (tipo hardware)	processore
1.8.6.3 hwOtherInformation (altre informazioni hardware)	minimo 60 Mhz
1.8.6 hardware	
1.8.6.1 hwName (nome hardware)	64MB RAM
1.8.6.2 hwType (tipo hardware)	memoria
1.8.6.3 hwOtherInformation (altre informazioni hardware)	minimo 32 MB
1.10 relationship (relazione)	
1.10.1 relationshipType (tipo relazione)	strutturale
1.10.2 relationshipSubType (sottotipo di relazione)	elementi allo stesso livello
1.10.3 relatedObjectIdentification (identificazione oggetto correlato)	
1.10.3.1 relatedObjectIdentifierType (tipo identificatore oggetto correlato)	hdl
1.10.3.2 relatedObjectIdentifierValue (valore identificatore oggetto correlato)	loc.music/gottlieb.09602

Capire PREMIS

1.10.3.3 relatedObjectSequence (sequenza identificatore oggetto correlato)	0
1.10 relationship (relazione)	
1.10.1 relationshipType (tipo relazione)	strutturale
1.10.2 relationshipSubType (sottotipo di relazione)	è incluso in:
1.10.3 relatedObjectIdentification (identificazione oggetto correlato)	
1.10.3.1 relatedObjectIdentifierType (tipo identificatore oggetto correlato)	deposito locale
1.10.3.2 relatedObjectIdentifierValue (valore identificatore oggetto correlato)	R200802948
1.10.3.3 relatedObjectSequence (sequenza oggetto correlato)	0
1.12. linkingIntellectualEntityIdentifier (identificatore entità correlata)	
1.12.1 linkingIntellectualEntityIdentifierType (tipo identificatore entità correlata)	LCCN
1.12.2 linkingIntellectualEntityIdentifierValue (valore identificatore entità correlata)	2007-86121

Appendice C: Glossario dei termini

Questo glossario mette insieme le definizioni che sono presenti nel testo precedente.

Le definizioni potrebbero essere meno formali di quelle che appaiono nel Dizionario dei Dati PREMIS.

azionabile: La qualità di essere memorizzato in una forma controllata che può essere azionata da un programma di computer.

agente: Una persona, una organizzazione o un programma per computer che ha un ruolo in un evento o una dichiarazione sui diritti.

oggetto *bitstream*: Un tipo di oggetto PREMIS; i dati all'interno di un file che hanno proprietà comuni ai fini della conservazione e non possono stare da soli.

unità contenitore: Unità semantiche che non hanno valore per se stesse ma esistono per raggruppare le sottounità correlate.

provenienza digitale: Documentazione della catena di custodia e dei cambi evolutivi di una risorsa digitale.

emulazione: Una strategia di conservazione che implica il riprodurre un vecchio ambiente tecnologico di fruizione su hardware e/o software di nuova generazione.

entità: In PREMIS è un tipo di cose su cui si possa argomentare. I tipi di entità PREMIS sono le Entità Intellettuali, gli Oggetti, gli Agenti, gli Eventi, i Diritti.

ambiente: Hardware, software e altri oggetti richiesti per permettere la fruizione dei contenuti di un oggetto.

entità evento: Una entità PREMIS che aggrega informazioni sulle azioni che hanno una qualche influenza sugli oggetti di un deposito.

contenitore di estensione: Un tipo speciale di unità contenitore PREMIS che non ha sottounità definite ma viene definita come segnaposto per i metadati non-PREMIS.

oggetto file: Un tipo di oggetto PREMIS; un file di computer, come un PDF o un JPEG.

inibitore: Caratteristica di un oggetto digitale che ne restringe l'accesso, l'uso o la migrazione.

entità intellettuale: Un insieme di contenuto che viene considerato come un'unità per obiettivi di gestione e descrizione; simile a una "entità bibliografica" nella scienza biblioteconomica.

migrazione: Una strategia di conservazione che implica la creazione di una versione diversa di un file digitale in un nuovo formato.

oggetti: Gli oggetti digitali che sono attualmente archiviati e gestiti in un deposito di conservazione: file, bitstream e rappresentazioni.

metadati di conservazione: I metadati che supportano le attività che servono ad assicurare l'usabilità nel lungo periodo di una risorsa digitale.

Capire PREMIS

strategie di conservazione : Le tecniche impiegate per assicurare che le risorse digitali rimangano usabili nel lungo periodo; due strategie comuni sono la migrazione e l'emulazione.

rappresentazione : Un tipo di oggetto PREMIS; l'insieme di tutti gli oggetti file necessari alla fruizione di un'*entità intellettuale*.

entità diritti : Un'entità PREMIS che aggrega informazioni sui diritti e sui permessi che riguardano gli oggetti che sono nel deposito di conservazione.

unità semantiche : Pezzi di informazione o di conoscenza.

proprietà significative : Caratteristiche di un oggetto che dovrebbero essere mantenute attraverso le azioni di conservazione.