

CAPIRE PREMIS

Library of Congress Network Development and MARC Standards Office

Autore: Priscilla Caplan

Revisione: PREMIS Editorial Committee

Pubblicato originariamente nel 2009; rivisto nel 2017; rivisto 2021

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 (member of the PREMIS Editorial Committee)

Sapienza Università di Roma, Sistema Bibliotecario Sapienza

INDICE dei CONTENUTI

CAPIRE PREMIS.....	1
1. PREMIS IN CONTEXT.....	1
1.1. Cosa sono i metadati di conservazione?.....	1
1.2. Casa è PREMIS?.....	2
1.3. Cosa si trova nel Dizionario dei Dati di PREMIS?.....	2
1.4. Come PREMIS dovrebbe essere usato?.....	4
1.5. Dovresti usare PREMIS?.....	4
2. CONCETTI CHIAVE IN PREMIS.....	5
2.1. Unità Semantiche.....	5
2.2. Contenitori e Sottounità.....	6
2.3. Contenitori di Estensione.....	6
3. MODELLO di DATI PREMIS.....	7
3.1. Entità Oggetto.....	7
3.2. Eventi.....	10
3.3. Agenti.....	10
3.4. Diritti.....	11
4. IL DIZIONARIO dei DATI.....	11
4.1. Esempio di Unità Semantica Semplice come Voce del Dizionario dei Dati.....	11
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. La Conformità a PREMIS.....	16
6. MAGGIORI INFORMAZIONI.....	18
Appendice A: Esempi.....	19
A.1. Esempio di Oggetto.....	19
A.2. Esempio di Evento.....	22
A.3. Esempio di Agente.....	23
Appendice B: Glossario di 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 offre un'idea di cosa sia PREMIS e di quale sia il suo campo d'azione.

Per molti lettori, questa guida sarà sufficiente. Per coloro che hanno invece bisogno di avere una maggiore padronanza delle 250 e più, pagine del Dizionario dei Dati PREMIS per i Metadati di Conservazione, la presente guida potrebbe servire come introduzione che rende più facile l'approccio al documento più esteso.

I. CONTESTO APPLICATIVO DI PREMIS

I.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 in modo proprio. 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 per supportare il processo di conservazione digitale". Di seguito vengono esemplificate alcune attività di conservazione e come i metadati supportano tali attività:

- Una risorsa deve essere conservata abbassando la soglia di rischio ad un livello 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 difficoltoso 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.
- Nel lungo periodo 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 gestori della conservazione devono adottare le *strategie di conservazione* più opportune 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 strategie di conservazione potrebbero determinare il cambiamento delle risorse originarie (*migrazione*), o il cambiamento di come esse vengono supportate tecnologicamente per la fruizione (*emulazione*). 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.

I.2. Cosa 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. La Library of Congress ha poi pubblicato un insieme di schemi per rappresentare in XML gli elementi di metadati del Dizionario dei Dati. Una seconda edizione aggiornata del Dizionario dei Dati e un nuovo schema XML sono stati pubblicati nel Marzo 2008, seguiti da revisioni minori, versione 2.1 e 2.2. Nel giugno 2015 una revisione più consistente, la versione 3.0, è stata pubblicata. La Library of Congress sostiene l'attività di mantenimento di PREMIS.

Quest'ultima è denominata PREMIS Maintenance Activity. 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/v3/premis-3-0-final.pdf

Sito web PREMIS: www.loc.gov/premis

Lista di discussione del gruppo di implementatori PREMIS: pig@listserv.loc.gov

Per iscriversi mandare un'email a listserv@listserv.loc.gov con oggetto del messaggio:
subscribe pig [Tuo Nome]

I.3. Cos'è il Dizionario dei Dati PREMIS?

Il Dizionario dei Dati PREMIS definisce un insieme di base di unità semantiche (vedi sezione 2.1) che i depositi dovrebbero sapere per svolgere le loro funzioni di conservazione.

Le funzioni di conservazione possono essere diverse a seconda del tipo di sistema di deposito, ma generalmente includono azioni per assicurare che gli oggetti digitali rimangano accessibili per i media di supporto e possano essere visualizzati, eseguiti o altrimenti interpretati dalle applicazioni software, come anche 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, il più delle volte, la maggior parte dei depositi avrà bisogno di sapere. Diverse categorie di metadati sono escluse in quanto ritenute fuori obiettivo:

- 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.
- 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.
- Metadati descrittivi. Sebbene la descrizione delle risorse è ovviamente di pertinenza della conservazione, molti standard indipendenti possono essere usati a tale scopo, come il MARC, il MODS e il 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 aggiuntive sui diritti e sui permessi; il focus è incentrato su 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.

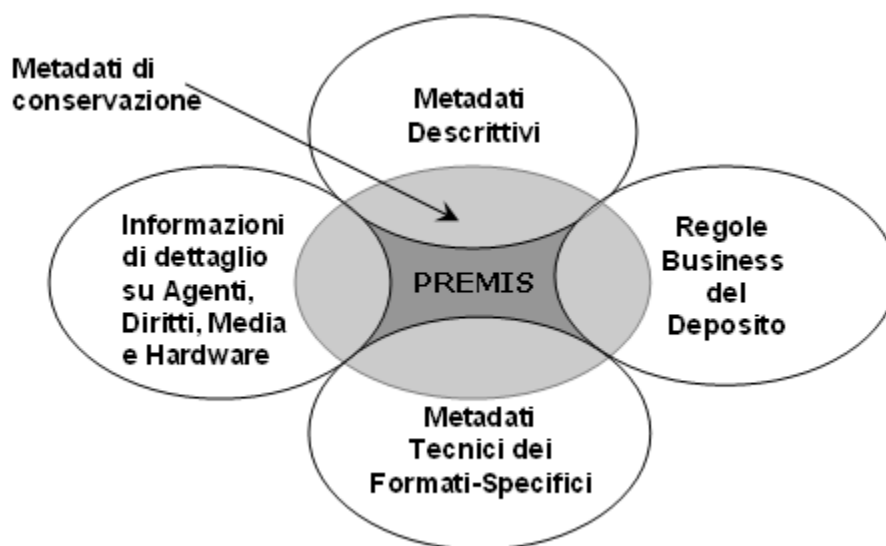


Figure 1: PREMIS come sottoinsieme di tutti i Metadati di Conservazione

La Figura 1 mostra tutti i metadati pertinenti alla conservazione come il cerchio 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.

I.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 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 linee guida secondo le quali le informazioni dovrebbero essere ottenute e memorizzate 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 lista di controllo 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.

I.5. Dovresti 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 tenuto nell'ambito della 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, non li forniscano già.) In ogni caso, ci sono alcune informazioni che, laddove possibile, dovrebbero essere salvate:

Inhibitors. Gli inibitori vengono così definiti come quelle caratteristiche di un oggetto che ne inibiscono l'accesso, l'uso o la migrazione. Gli inibitori includono le 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).

Provenance. 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. PREMIS permette di tenere memoria dei cambiamenti evolutivi nelle informazioni dell'entità Evento, che viene descritta successivamente. Un vocabolario controllato per i tipi di evento è disponibile, al seguente indirizzo: id.loc.gov/vocabulary/preservation/eventType.

Molti dei tipo di evento PREMIS sono stati previsti per descrivere azioni che accadono dopo che qualche oggetto viene acquisito da un deposito, sebbene qualcuno di essi è utile per eventi che accadono anche prima dell'acquisizione (ingestion), come la cattura e l'accessione (assegnazione dell'inventario all'oggetto).

Significant Properties. Le Proprietà significative sono le specifiche 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 importanti, 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. Nonostante ciò, 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 come proprietà significative.

Rights. 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 conosciuta sui diritti, incluse quelle sullo stato del copyright, i termini di licenza e i permessi particolari.

2. CONCETTI CHIAVE IN PREMIS

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. Del resto, gli elementi di metadati sono abbastanza simili.

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

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 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 si riflette nella numerazione delle unità semantiche:

I.1 <code>objectIdentifier</code> (M, R) I.1.1 <code>objectIdentifierType</code> (M, NR) I.1.2 <code>objectIdentifierValue</code> (M, NR)

Questo estratto del 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 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 della sua competenza 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, perché considerati fuori dall'area di competenza descrittiva del 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 (il suo schema XML viene denominato come “NISO Metadata for Images in XML Schema--MIX”).

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, nel paragrafo 5.1.

Una specifica iniziativa, intrapresa in seno al PREMIS, ha sviluppato un Ontologia PREMIS codificata in OWL, allo scopo di facilitare anche le implementazioni come i Linked Data. La prima versione dell'ontologia è compatibile con la versione 2.2. del PREMIS, e una versione è in corso di definizione per la versione 3.0. Le informazioni relative alla PREMIS OWL Ontology sono disponibili all'indirizzo: www.loc.gov/standards/premis/ontology

3. 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 Intellettuali, Oggetti, Agenti, Eventi e Diritti.

Nelle versioni 1 e 2 di PREMIS noi abbiamo definito 5 tipi di cose (chiamate *Entities*) di cui si può parlare: Entità Intellettuale, Oggetto, Agente, Evento e Diritti (*Il diagramma del modello dei dati per le versioni 1 e 2 è reperibile nel sito web di PREMIS.*). Nella versione 3, che è stata pubblicata nel Giugno 2015, in Modello di datai PREMIS è stato rivisto rendendo l'Entità Intellettuale, una categoria aggiuntiva dell'Oggetto.

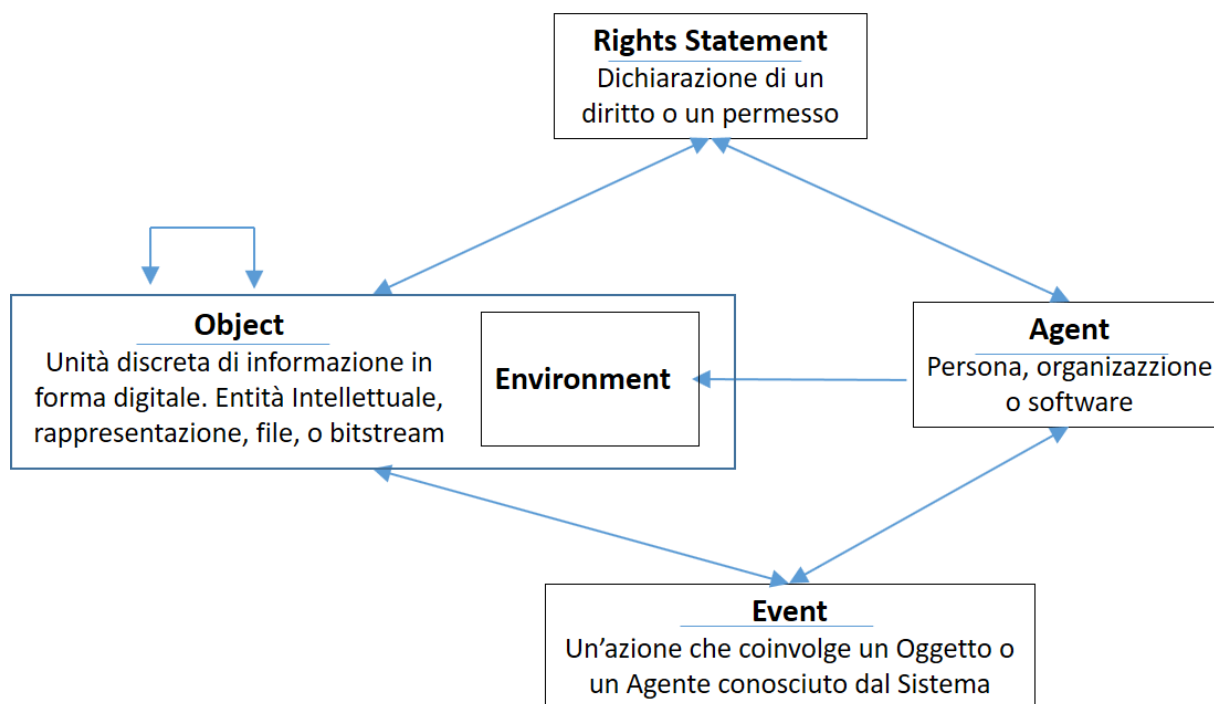


Figure 2: PREMIS Data Model, version 3

3.1. Entità Oggetto

Gli Oggetti sono ciò che realmente viene archiviato e gestito nei depositi di conservazione. Buona parte del PREMIS è dedicata a descrivere 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* (message digest) e l'algoritmo usato per calcolarlo,

- la dimensione dell'Oggetto,
- il formato dell'Oggetto, che può essere specificato direttamente o mediante collegamento ad un registro dei 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 archiviato,
- le informazioni sulla firma digitale,
- le informazioni sulle relazioni con gli altri Oggetti e gli altri tipi di entità.

PREMIS in realtà definisce quattro diversi tipi di Oggetti e richiede agli implementatori di fare delle distinzioni tra di essi. Le quattro tipologie sono i *Bitstream*, i *File*, le *Rappresentazioni*, e le *Entità Intellettuali*.

Un **Oggetto File** è ciò che sembra essere, ovvero un file per computer come un PDF o un JPEG.

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.

Un **Oggetto Rappresentazione** l'insieme di tutti gli Oggetti File necessari a restituire il contenuto (*rendering*) di un'Entità Intellettuale. Per esempio, si vuole conservare una pagina Web, diciamo le *home page* di alcune date. E' probabile che la *home page*, visualizzata nel browser, 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.

Un **Oggetto Entità Intellettuale** viene definito come un insieme di contenuto che è considerato una singola unità intellettuale 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 Intellettuali, in quanto esistono molti standard di metadati descrittivi tra cui poter scegliere. Nella versione 1 e 2 le Entità Intellettuali potevano solo essere identificate in quanto considerate solo concettuali. Ma nella versione 3 un'Entità Intellettuale può essere descritta sia da metadati descrittivi fuori dal PREMIS, **sia da** metadati di conservazione come un Oggetto interno a PREMIS. In molti casi le unità semantiche usate per le Entità Intellettuali sono le stesse utilizzate per le Rappresentazioni.

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.

Alcune unità semantiche definite nel Dizionario dei Dati di PREMIS sono applicabili a tutti i quattro tipi di Oggetto, mentre altre sono applicabili solo ad uno o due tipi di Oggetti.

Diverse unità semantiche sono definite per memorizzare l'ambiente tecnologico (*environment*) di un Oggetto, specificamente l'*hardware* ed il *software* necessari a restituirne il contenuto (*render*) e quali dipendenze esistono rispetto ad altri Oggetti. Nelle versioni 1 e 2 queste informazioni sono parte delle descrizioni di un Oggetto. Nella versione 3 l'Environment, viene definito come un tipo specifico di Oggetto Entità Intellettuale, e potrebbe essere collegato a e da File, Rappresentazioni e Bitstream che li usano. Per esempio, un file PDF può essere visualizzato da diverse versioni di Adobe Acrobat e Adobe Reader come da altri programmi commerciali o *open source*. Ognuno di questi, è supportato a sua volta da vari sistemi operativi che richiedono dei requisiti minimi di *hardware* (velocità processore, memoria centrale e memoria disco). Se Adobe Reader è in versione plug-in per browser, esso dipende anche da alcune versioni di alcuni *web browser*; per esempio la versione Reader 9 per il Mac OS richiede il browser Safari versione 2.0.4 o successive. Le informazioni sull'Environment sono critiche per alcune strategie di conservazione, ma è difficile e oneroso raccoglierle, e potrebbe essere meglio che fossero raccolte in registri centralizzati come il registro PRONOM per i formati, mantenuto dal National Archives of the UK. PREMIS permette ai sistemi di deposito di collegare le informazioni mantenute da registri esterni se preferibile al gestirle localmente.

Per illustrare le relazioni tra gli Oggetti, consideriamo di nuovo la pagina Web menzionata precedentemente. La pagina Web è una *Entità Intellettuale* che può essere espressa in un numero di diverse Rappresentazioni. Una Rappresentazione, come dettagliato precedentemente, consiste di un numero di Oggetti File separati (pagine HTML, immagini, fogli di stile.). Comunque, la pagina Web potrebbe anche essere archiviata come un singolo file Web Archive (WARC) (vedi specifiche: <https://www.digitalpreservation.gov/formats/fdd/fdd000236.shtml>). Questa è una diversa Rappresentazione della stessa Entità Intellettuale, che contiene un solo file, come viene di seguito illustrato:

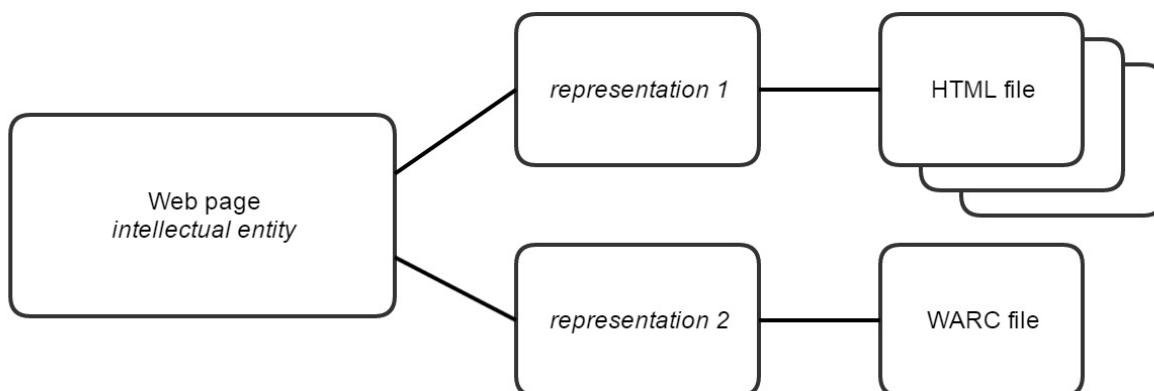


Figura 3: Esempio di relazioni dell'Oggetto PREMIS

3.2. Eventi

L'entità Evento 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'esito 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 come parte permanente della storia di un Oggetto. PREMIS raccomanda che le azioni che cambiano un oggetto dovrebbero sempre essere memorizzate, e fornisce un vocabolario controllato di tipi di Evento di importanza rilevante, per incoraggiare i depositi a memorizzare questi Eventi in modo consistente.

Il vocabolario è disponibile al seguente indirizzo: <http://id.loc.gov/vocabulary/preservation/eventType>.

3.3. Agenti

Gli Agenti sono attori che hanno dei ruoli negli Eventi e nelle dichiarazioni sui Diritti (vedi 3.4 Diritti), e negli Oggetti Ambiente.

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. 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.

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, hardware),
- Versione dell'Agente (per software o hardware),
- Una nota generale sugli Agenti,
- Eventi associati all'Agente,
- Dichiarazione di Diritti associati all'Agente,
- Oggetti Ambiente associati all'Agente.

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'opera, l'autore (ma non il detentore

dei diritti) di una seconda opera e il depositante di una terza opera. Nel modello PREMIS un deposito mi assegnerebbe un identificatore univoco 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.

3.4. 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 di conservazione, in modo tale che il deposito possa fare ciò che necessita, per conservarli. 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, o altri diritti,
- informazioni più dettagliate sullo stato del 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 restrizione, 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 e 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, per esempio, nella specifica di metadati di Copyright (copyrightMD) della Biblioteca Digitale della California (<https://cdlib.org/groups/rights-management-group-copyrightmd/>). 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. IL DIZIONARIO DEI DATI

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

La Tabella 1 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

(rationale) 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 attributi sono tutti pensati per aiutare gli implementatori ad usare l'elemento in modo proprio.

Le due righe categoria dell'oggetto (Object Category) e applicabilità (Applicability) sono usate insieme per mostrare quando le unità semantiche sono appropriate per la descrizione delle Entità Intellettuali, delle Rappresentazioni, dei File e/o dei Bitstream. Qui la dimensione viene mostrata di sola pertinenza dei File e dei Bitstream. Infine, esiste un insieme di regole per l'uso: vincoli sui dati (Data constraint), ripetibilità (Repeatability) e obbligatorietà (Obligation).

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. Quando un vocabolario controllato viene indicato come vincolo del dato, qualche termine del vocabolario viene specificato a titolo di esempio nel Dizionario dei Dati. L'insieme dei vocabolari controllati principalmente utilizzati sono consultabili al seguente indirizzo:

<https://id.loc.gov/preservationdescriptions/>

Altri vocabolari potrebbero essere usati, in tal caso il nome del vocabolario controllato usato deve essere indicato. Non ci sono unità semantiche definite nel Dizionario dei Dati per i nomi del vocabolario, ma lo schema XML di PREMIS permette di indicarne la sua identificazione attraverso un URI, che generalmente ne riflette il nome.

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 potrebbe non 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 in cui si renda necessario. (vedi Sezione 5.2. Conformità PREMIS.). Alcune unità semantiche sono obbligatorie all'interno di un contenitore; se un contenitore non è obbligatorio l'unità semantica può non essere usata se in contenitore non viene usato.

Table 1: Estratto dal Dizionario dei Dati per l'unità semantica size

Semantic unit	I.5.3 size		
Semantic components	None		
Definition	La dimensione in byte di un file o un bitstream in possesso del deposito.		
Razionale	La dimensione è utile per assicurare che si possa riottenere il corretto numero di byte dallo storage (unità generica di memorizzazione) e che un'applicazione abbia abbastanza spazio per spostare o elaborare i file. Potrebbe inoltre essere usato per contabilizzare lo spazio di storage.		
Data constraint	Intero		
Object category	Intellectual Entity / Representation	File	Bitstream
Applicability	Non applicabile	Applicabile	Applicabile
Examples		2038937	2038937
Repeatability		Non ripetibile	Non ripetibile
Obligation		Opzionale	Opzionale
Creation / Maintenance notes	Ottenuto automaticamente dal deposito.		
Usage notes	Definire questa unità semantica come dimensione in byte, 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 per entrambi i partner.		

4.2. Esempio di Unità Contenitore come Voce del Dizionario dei Dati

La Tabella 2 mostra l'inizio della voce del Dizionario dei Dati `objectCharacteristics`, l'unità contenitore di `size`. Si può dire che è un contenitore avendo dei componenti semantici e la specifica "contenitore" come restrizione sui dati (Data Constraint). Si noti che i componenti semantici inclusi possono essere unitari come `size`, o altri contenitori, come nel caso di `format`.

Table 2: Estratto del Dizionario dei Dati per l'unità semantica `objectCharacteristics`

Semantic unit	1.5 <code>objectCharacteristics</code>		
Semantic components	1.5.1 <code>compositionLevel</code> 1.5.2 <code>fixity</code> 1.5.3 <code>size</code> 1.5.4 <code>format</code> 1.5.5 <code>creatingApplication</code> 1.5.6 <code>inhibitors</code> 1.5.7 <code>objectCharacteristicsExtension</code>		
Definition	Proprietà tecniche di un file o di un bitstream che sono applicabili a tutti o alla maggior parte dei formati.		
Rationale	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	Container		
Object category	Intellectual Entity / Representation	File	Bitstream
Applicability	Non applicabile	Applicabile	Applicabile
Repeatability		Ripetibile	Ripetibile
Obligation		Obbligatorio	Obbligatorio
Usage notes	Le unità semantiche incluse in <code>objectCharacteristics</code> dovrebbero essere considerate come un insieme di informazioni riguardanti un oggetto singolo e ad un singolo livello di composizione (<code>compositionLevel</code>). Le caratteristiche dell'oggetto potrebbero essere ripetute quando un oggetto viene creato applicando due o più codifiche, come la compressione e la criptazione. In questo caso la ripetizione di <code>objectCharacteristics</code> dovrebbe avere <code>compositionLevel</code> incrementalmente più alto. Quando viene applicata la criptazione, il blocco di <code>objectCharacteristics</code> deve includere una unità semantica <code>inhibitor</code> . Un bitstream incluso in un file potrebbe avere caratteristiche diverse dal file. Quando queste caratteristiche sono pertinenti alla conservazione dovrebbero essere memorizzate.		

5. PREMIS IN USE

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 sistemi di deposito. 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. Esiste una certa corrispondenza tra le due strutture sebbene non sia 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 - checksum - (message digest) 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. Di conseguenza, gli implementatori hanno sviluppato buone pratiche nell'uso di PREMIS e METS insieme. Le linee guida per usare PREMIS con METS sono disponibili nel sito web della PREMIS Maintenance Activity, al seguente indirizzo:

<https://www.loc.gov/standards/premis/guidelines2017-premismets.pdf>

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

Figura 4: Un frammento di PREMIS in XML

5.2. La Conformità a PREMIS

La specifica PREMIS contiene una sezione su cosa significhi per un deposito essere conforme al PREMIS. Il *conformance statement* (la condizione in cui si trova un sistema di deposito e in base alla quale può affermare se è conforme o meno al PREMIS), che il PREMIS Editorial Committee ha pubblicato per la prima volta nell'Ottobre 2010 e revisionato nel 2015, specifica i principi di uso, livelli di conformità, e guida le implementazioni. Il documento del *conformance statement* è disponibile al seguente indirizzo:

<https://www.loc.gov/standards/premis/premis-conformance-20150429.pdf>.

Principi di uso

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 viene definita nell'unità semantica del PREMIS. Se un elemento dei metadati condivide la definizione di un'unità semantica del PREMIS ma non condivide lo stesso nome, il sistema di deposito deve stabilire un mapping tra l'elemento di metadati e la sua corrispondente unità semantica 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) Un'implementazione deve includere le unità semantiche obbligatorie per qualsiasi Entità del Modello dei Dati (Object, Event, Agents o Rights) supportate dal deposito. Si noti che un componente semantico obbligatorio viene solo incluso se il contenitore viene implementato.

4) 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 deposito è libero di gestire o meno l'entità Agente di PREMIS, ma se la gestisce, allora *agentIdentifier*, essendo obbligatorio, DEVE essere incluso. In modo simile, un deposito in particolare potrebbe non gestire gli Oggetti Bitstream, in tal caso non deve fornire gli identificatori obbligatori del Bitstream.

Livelli di conformità

I livelli specificano tre modi di affermare la conformità con PREMIS in un sistema di deposito:

- **Livello 1.** Essere in grado di mappare i metadati di deposito verso PREMIS,
- **Livello 2.** Essere in grado di esportare i metadati di deposito come PREMIS, e
- **Livello 3.** Usare PREMIS come schema interno in modo da non richiedere nessun ulteriore mapping o conversione.

Questi livelli sono ulteriormente divisi in categorie: implementazioni della sola Entità Oggetto o implementazione dell'Entità Oggetto più una o altre Entità. In altre parole, un deposito deve aver almeno

implementato informazioni riguardanti l'Oggetto, ma non è richiesto di supportare tutti gli altri tipi di Entità definite nel PREMIS Data Model. Come non viene richiesto di memorizzare i metadati internamente, usando i nomi delle unità semantiche del PREMIS, o di usare valori che seguano le restrizioni sui dati del PREMIS. In altre parole, non è importante come il deposito sia a conoscenza di un valore PREMIS – se lo gestisce con lo stesso nome o con un nome diverso, se lo mappa da un altro valore, se punta ad un registro, se lo deduce o se per default, o attraverso altri mezzi. Di conseguenza se il deposito può fornire un valore accettabile per PREMIS, è conforme.

D'altra parte, maggiore è il numero di unità semantiche supportate dal deposito, e maggiore è il valore acquisito dall'uso del PREMIS. Il Dizionario dei Dati PREMIS è stato sviluppato per identificare le informazioni più importanti, che la maggior parte dei depositi necessita di avere per conservare contenuti digitali nel lungo periodo. Un responsabile del deposito di conservazione dovrebbe considerare attentamente il PREMIS e avere buone ragioni per commettere errori nell'implementare qualsiasi parte del Dizionario dei Dati.

6. MAGGIORI INFORMAZIONI

Il sito Web della PREMIS Maintenance Activity (<https://www.loc.gov/standards/premis/premis-conformance-20150429.pdf>) contiene strumenti di documentazione per chiunque ne abbia la necessità, come: il PREMIS Implementers Group (PIG), PREMIS *implementation fairs* e tutorial, schemi, strumenti e notizie varie.

Contiene, inoltre, una sezione di "Risorse PREMIS" che collegano alla letteratura del PREMIS e ai relativi argomenti (<https://www.loc.gov/standards/premis/bibliography.html>).

Alcune delle risorse più utili per gli utenti interessati al PREMIS, in modo più generale possono essere consultate ai seguenti indirizzi:

Sui metadati di conservazione in generale:

“Preservation Metadata” 2nd edition (PDF:986KB/36pp.)

Brian Lavoie (OCLC) and Richard Gartner (Oxford)

Published by the Digital Preservation Coalition as DPC Technology Watch Report No. 13-03:
May 2013.

dx.doi.org/10.7207/twr13-03

Sull'implementazione di PREMIS:

Digital Preservation Metadata for Practitioners: Implementing PREMIS.

Angela Dappert, Rebecca Squire Guenther, Sébastien Peyrard, Editors. Springer, 2016.

dx.doi.org/10.1007/978-3-319-43763-7

Sui cambiamenti della versione 3.0 (Webinar):

“Digital Preservation Metadata and Improvements to PREMIS in Version 3.0: A DCMI/ASIST Joint Webinar Presented by Angela Dappert” (Wednesday, May 27, 2015)

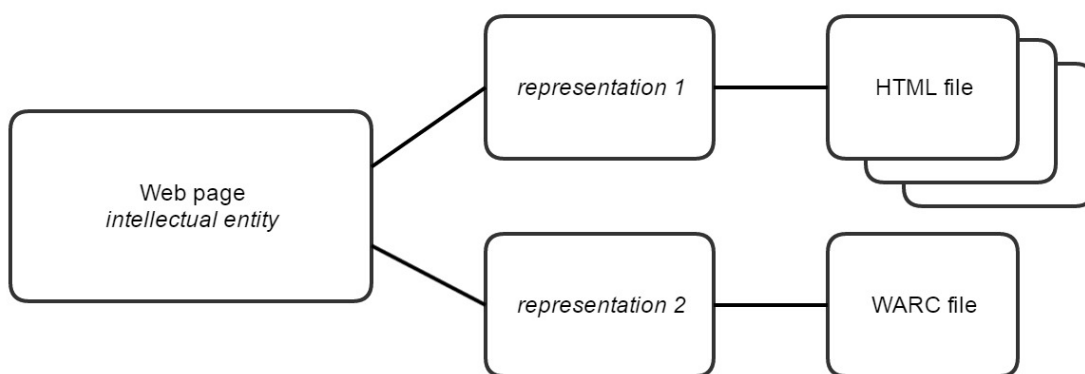
<https://www.loc.gov/standards/premis/v3/tutorial.html>

Appendice A: Esempi

A.1. Esempio di Oggetto

Questo esempio descrive come un sito web potrebbe essere modellato in un ipotetico sistema di conservazione. L'esempio scorre attraverso un diagramma di alto livello, alcuni dettagli su come le unità semantiche possono essere popolate, e a titolo di esempio alcuni estratti del codice XML vengono mostrati alla fine, mettendo in evidenza come le unità semantiche potrebbero essere implementate nel deposito. Gli esempi non mostrano tutte le unità semantiche che sono rilevanti per ogni livello dell'Oggetto.

In questo esempio, il sito web viene chiamato "PREMIS Website." Questa è l'Entità Intellettuale che viene preservata. Esistono due Rappresentazioni di questo sito. Nell'esempio, una è un master di conservazione e l'altra è un master di tipo accesso.



Entità Intellettuale

Il livello di Entità Intellettuale (IE) in questo esempio configura una proprietà significativa del sito web. Questa governa entrambe le Rappresentazioni. E' ovviamente molto lontano da qualsiasi proprietà significativa che potrebbe essere più granulare, ma l'esempio serve a mostrare come possano essere utilizzate a livello di IE.

PREMIS Semantic Unit	IE I
1.1 objectIdentifier	
1.1.1 objectIdentifierType	IEPID
1.1.2 objectIdentifierValue	17415491
1.4 significantProperties	
1.4.1 significantPropertiesType	behavior
1.4.2 significantPropertiesValue	External links open in new browser window

Rappresentazione

Il livello di conservazione descrive la politica istituzionale per questo Oggetto. Il livello di conservazione può essere applicato a livello di IE, Rappresentazione o File. In questo caso è stato applicato al livello Rappresentazione, come l'istituzione vuole che ogni Rappresentazione sia trattata in modo specifico.

L'istituzione in questo esempio ha deciso che la Rappresentazione comprende un file WARC (Representation 2) che è quello che considera come master di conservazione; quello che verrà mantenuto in modo perpetuo e da cui i master di accesso vengono ricavati. *Representation 1* è il master di accesso che è la Rappresentazione che viene usata dal pubblico come vista. Il suo valore non è per il lungo termine, e viene mantenuta a lungo quanto l'attuale meccanismo di accesso lo permetta.

Rappresentazione 1 viene assegnato il valore "bit preservation". La politica istituzionale definisce che per questo valore vengano fatte solo copie di backup senza analisi del rischio con requisiti di risultante emulazione o migrazione.

Rappresentazione 2 viene assegnato il valore "logical preservation" con intenzione completa (full intention). Ciò viene definito dall'istituzione come specifica Rappresentazione da sottoporre all'analisi del rischio di pertinenza e che richiede che azioni di emulazione o migrazione vengano intraprese allo scopo di assicurare, nel futuro, la conservazione.

PREMIS Semantic Unit	Representation 1	Representation 2
I.1 objectIdentifier		
I.1.1 objectIdentifierType	REPPID	REPPID
I.1.2 objectIdentifierValue	17415492	17415495
I.3 preservationLevel		
I.3.1 preservationLevelType	bit preservation	logical preservation
I.3.2 preservationLevelValue	low	full
I.3.3 preservationLevelRole	capability	intention
I.3.4 preservationLevelRationale	institutional policy	institutional policy
I.3.5 preservationLevelDateAssigned	2015-02-23	2015-02-23
I.13 relationship		
I.13.1 relationshipType	derivation	derivation
I.13.2 relationshipSubType	has source	Is source of
I.13.3 relatedObjectIdentifier		
I.13.3.1 relatedObjectIdentifierType	REPPID	REPPID
I.13.3.2 relatedObjectIdentifierValue	17415495	17415492

Esempio XML di come il livello di conservazione (metadato: `preservationLevel`) potrebbe essere codificato:

```
<premis:preservationLevel>
  <premis:preservationLevelType>logical preservation</premis:preservationLevelType>
  <premis:preservationLevelValue>full</premis:preservationLevelValue>
  <premis:preservationLevelRole authority="preservationLevelRole"
    authorityURI="http://id.loc.gov/vocabulary/preservation/preservationLevelRole"
    valueURI="http://id.loc.gov/vocabulary/preservation/preservationLevelRole/int">intention
  </premis:preservationLevelRole>
  <premis:preservationLevelRationale>institutional policy</premis:preservationLevelRationale>
  <premis:preservationLevelDateAssigned>2015-02-23</premis:preservationLevelDateAssigned>
</premis:preservationLevel>
```


File

Rappresentazione I include un numero discreto di file, ognuno dei quali viene descritto come un Oggetto File con i propri insiemi di elementi di metadato.

PREMIS Semantic Unit	File 1	File 2
1.1 objectIdentifier		
1.1.1 objectIdentifierType	FILEPID	FILEPID
1.1.2 objectIdentifierValue	17415493	17415494
1.5 objectCharacteristics		
1.5.1 compositionLevel	0	0
1.5.2 fixity		
1.5.2.1 messageDigestAlgorithm	SHA256	SHA256
1.5.2.2 messageDigest	d2bed92b73c7090bb30a0b30016882e7069c437488e1513e9deaacbe29d38d92	074862dbfa0806ef5a26c3ff748d394e79728e9c957ff8c198ae13214c9cfec0
1.5.2.3 messageDigestOriginator	NRI	NRI
1.5.3 size	4859385	123451
1.5.4 format		
1.5.4.1 formatDesignation		
1.5.4.1.1 formatName	Extensible Hypertext Markup Language	JPEG File Interchange Format
1.5.4.1.2 formatVersion	1	1.02
1.5.4.2 formatRegistry		
1.5.4.2.1 formatRegistryName	PRONOM	PRONOM
1.5.4.2.2 formatRegistryKey	fmt/102	fmt/44
1.5.4.2.3 formatRegistryRole	specification	specification
1.13 relationship		
1.13.1 relationshipType	structural	structural
1.13.2 relationshipSubType	Is included in	Is included in
1.13.3 relatedObjectIdentifier		
1.13.3.1 relatedObjectIdentifierType	REPPID	REPID
1.13.3.2 relatedObjectIdentifierValue	17415492	17415492

Esempio XML di come le caratteristiche dell'entità Oggetto (metadato: objectCharacteristics) potrebbero essere codificate:

```
<premis:objectCharacteristics>
  <premis:compositionLevel>0</premis:compositionLevel>
  <premis:fixity>
    <premis:messageDigestAlgorithm>SHA256</premis:messageDigestAlgorithm>
    <premis:messageDigest>d2bed92b73c7090bb30a0b30016882e7069c437488e1513e9deaacbe29d38d92
    </premis:messageDigest>
    <premis:messageDigestOriginator>NRI</premis:messageDigestOriginator>
  </premis:fixity>
  <premis:size>4859385</premis:size>
  <premis:format>
    <premis:formatDesignation>
      <premis:formatName>Extensible Hypertext Markup Language</premis:formatName>
      <premis:formatVersion>1</premis:formatVersion>
```

```

</premis:formatDesignation>
<premis:formatRegistry>
  <premis:formatRegistryName>PRONOM</premis:formatRegistryName>
  <premis:formatRegistryKey>fmt/102</premis:formatRegistryKey>
  <premis:formatRegistryRole>specification</premis:formatRegistryRole>
</premis:formatRegistry>
</premis:format>
</premis:objectCharacteristics>

```

A.2. Esempio di Evento

Questo esempio descrive un'azione che il deposito applica ad un oggetto. Una funzione comune di un deposito è identificare il formato del file di un oggetto e verificare che sia conforme allo standard di quel formato, notificando compatibilità o eccezioni. L'esempio illustra la validazione del File I del precedente esempio di Oggetto.

PREMIS Semantic Unit	File I
2.1 eventIdentifier	
2.1.1 eventIdentifierType	DPS
2.1.2 eventIdentifierValue	25
2.2 eventType	validation
2.3 eventDateTime	2013-07-01T17:23:25Z
2.4 eventDetailInformation	
2.4.1 eventDetail	Format identification performed on file
2.5 eventOutcomeInformation	
2.5.1 eventOutcome	success
2.5.2 eventOutcomeDetail	
2.5.2.1 eventOutcomeDetailNote	FORMAT_ID=fmt/102;IDENTIFICATION_METHOD=SIGNATURE;FILE_EXTENSION=html;DEPOSIT_ACTIVITY_ID=659202;PID=FL26274026;SIP_ID=444381;PRODUCER_ID=29408010705;TASK_ID=48;PROCESS_ID=;MF_ID=8
2.6 linkingAgentIdentifier	
2.6.1 linkingAgentIdentifierType	local
2.6.2 linkingAgentIdentifierValue	DROID6_63_I
2.6.3 linkingAgentIdentifierRole	implementer
2.7 linkingObjectIdentifier	
2.7.1 linkingObjectIdentifierType	FILEPID
2.7.2 linkingObjectIdentifierValue	I74I5493

Esempio XML di come i metadati relativi ad un Evento di identificazione di formato potrebbero essere codificati:

```

<premis:event>
  <premis:eventIdentifier>
    <premis:eventIdentifierType>DPS</premis:eventIdentifierType>
    <premis:eventIdentifierValue>25</premis:eventIdentifierValue>
  </premis:eventIdentifier>
  <premis:eventType authority="event type"
    authorityURI="http://id.loc.gov/vocabulary/preservation/eventType"
    valueURI="http://id.loc.gov/vocabulary/preservation/eventType/val">
    validation</premis:eventType>
  <premis:eventDateTime>2013-07-01T17:23:25Z</premis:eventDateTime>

```

```

<premis:eventDetailInformation>
  <premis:eventDetail>Format identification performed on file</premis:eventDetail>
</premis:eventDetailInformation>
<premis:eventOutcomeInformation>
  <premis:eventOutcome>success</premis:eventOutcome>
  <premis:eventOutcomeDetail>
    <premis:eventOutcomeDetailNote>
      FORMAT_ID=fmt/102;IDENTIFICATION_METHOD=SIGNATURE;FILE_EXTENSION=html;
      DEPOSIT_ACTIVITY_ID=659202;PID=FL26274026;SIP_ID=444381;
      PRODUCER_ID=29408010705;TASK_ID=48;PROCESS_ID=;MF_ID=8</premis:eventOutcomeDetailNote>
    </premis:eventOutcomeDetail>
  </premis:eventOutcomeInformation>
<premis:linkingAgentIdentifier>
  <premis:linkingAgentIdentifierType>LOCAL</premis:linkingAgentIdentifierType>
  <premis:linkingAgentIdentifierValue>DROID6_63_1</premis:linkingAgentIdentifierValue>
  <premis:linkingAgentRole authority="eventRelatedAgentRole"
    authorityURI="http://id.loc.gov/vocabulary/preservation/eventRelatedAgentRole"
    valueURI="http://id.loc.gov/vocabulary/preservation/eventRelatedAgentRole/imp">
    implementer</premis:linkingAgentRole>
</premis:linkingAgentIdentifier>
<premis:linkingObjectIdentifier>
  <premis:linkingObjectIdentifierType>FILEPID</premis:linkingObjectIdentifierType>
  <premis:linkingObjectIdentifierValue>17415493</premis:linkingObjectIdentifierValue>
</premis:linkingObjectIdentifier>
</premis:event>

```

A.3. Esempio di Agente

Questo esempio descrive l'agente associato all'evento sopradescritto. In questo caso l'agente è un software che esegue l'azione sull'oggetto.

PREMIS Semantic Unit	
3.1 agentIdentifier	
3.1.1 agentIdentifierType	local
3.1.2 agentIdentifierValue	DROID6_63_1
3.2 agentName	REG_SA_DROID
3.3 agentType	software
3.4 agentVersion	Version 6.01
3.5 agentNote	Signature version Binary SF v.63/ Container SF v.1

Esempio XML di come i metadati dell'entità Agente potrebbero essere codificati:

```

<premis:agent>
  <premis:agentIdentifier>
    <premis:agentIdentifierType>local</premis:agentIdentifierType>
    <premis:agentIdentifierValue>DROID6_63_1</premis:agentIdentifierValue>
  </premis:agentIdentifier>
  <premis:agentName>REG_SA_DROID</premis:agentName>
  <premis:agentType>software</premis:agentType>
  <premis:agentVersion>Version 6.01</premis:agentVersion>
  <premis:agentNote>Signature version Binary SF v.63/ Container SF v.1</premis:agentNote>
</premis:agent>

```

Appendice B: Glossario dei termini

Questo glossario mette insieme le definizioni che sono presenti nel testo che precede. 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 tale che possa essere azionato da un programma di computer.

Entità Agente: Una persona, una organizzazione o un programma per computer che ha un ruolo in un evento o in 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 comporta la riproduzione di un vecchio ambiente tecnologico di fruizione su hardware e/o software di nuova generazione.

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

Oggetto Ambiente: Un Oggetto che fa parte della stratificazione di software, hardware ed altre dipendenze necessarie ad interpretare correttamente le Rappresentazioni, i File e i Bitstream.

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 del *Dizionario dei Dati PREMIS* che non ha sottounità definite, ma che 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 gli obiettivi di gestione e descrizione; simile a una "entità bibliografica" nella scienza biblioteconomica.

Migrazione: Una *strategia di conservazione* che comporta la creazione di una versione diversa di un file digitale in un nuovo formato.

Oggetti: Oggetti digitali che sono realmente archiviati e gestiti in un deposito di conservazione. PREMIS definisce quattro tipi di Oggetti: *File*, *Bitstream*, *Rappresentazioni* e *Entità Intelletuali*.

Metadati di conservazione: I metadati che supportano le attività che assicurano l'usabilità nel lungo periodo di una risorsa digitale.

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*.

Oggetto 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 del deposito di conservazione.

Unità semantiche: Pezzi di informazione o di conoscenza.

Proprietà significative: Caratteristiche di un oggetto che dovrebbero essere mantenute dalle azioni di conservazione.