

Esperienze di utilizzo di XML in Italia: nuove prospettive di accesso e condivisione delle risorse archivistiche



Standard e formati di scambio per
l'interoperabilità dei sistemi archivistici

Bologna, 8-9 maggio 2008



Modalità di utilizzo di XML (1)

- recupero di dati esistenti su diverso formato
 - esempi:
 - Guida Generale degli Archivi di Stato italiani
 - Guida ai fondi e Casellario politico centrale dell'Archivio Centrale dello Stato
 - Descrizione dei complessi documentari (ex banca dati Anagrafe) dell'Archivio di Stato di Bologna
 - Descrizione dei complessi documentari (ex banca dati Anagrafe), gli inventari informatizzati e le informazioni sui soggetti produttori dell'Archivio di Stato di Napoli
 - Inventario analitico Datini dell'Archivio di Stato di Prato
 - Strumenti di ricerca della Camera dei deputati
 - Inventario dell'Archivio storico CGIL
 - Guida Unioncamere
 - Guida ai fondi dell'Archivio IRI
 - Inventari della Fondazione Feltrinelli
 - Strumenti degli archivi della Fondazione Gramsci di Roma
 - Inventari degli archivi storici comunali del Lazio - Progetto RInASCo
 - Archivi storici della Toscana - Progetto AST
 - caratteristiche:
 - recupero del preesistente all'interno dell'impianto EAD ed EAC
 - utilizzo di motori di ricerca per l'accesso all'informazione

Modalità di utilizzo di XML (2)

- formato di comunicazione e scambio di dati
 - esempi:
 - Software di descrizione archivistica: Sesamo, Arianna, GEA, GUARINI archivi
 - Sistemi informativi nazionali e territoriali: SIAS, SIUSA, SIASFI, PLAIN
 - caratteristiche:
 - database relazionali
 - impianto descrittivo ISAD e ISAAR
 - XML viene utilizzato come formato di esportazione, in qualche caso di importazione o di conservazione dei dati, o ancora come formato di dei dati forniti in risposta alle query impostate dall'utente

Modalità di utilizzo di XML (3)

- lavorare con XML in un ambiente integrato
 - adozione di un profilo applicativo XML, inteso non solo come linguaggio di scambio e di pubblicazione, ma anche come "strumento di lavoro" e di conservazione dei dati
 - entro un unico sistema convivono strutture descrittive diverse, autonome, flessibili
 - i vari elementi informativi sono ricondotti all'interno di un'unica sintassi, EAD, per la descrizione delle diverse tipologie di archivi, e di EAC, per la codifica degli authority file
 - ogni modalità di recupero e raccolta di dati viene realizzata senza dover strutturare le informazioni problemi tecnici di modifica di strutture preesistenti
 - piena indipendenza dalla piattaforma software e dagli strumenti applicativi

Un caso studio: gli archivi dell'Emilia-Romagna

- L'analisi degli elementi descrittivi di differenti archivi storici, locali e di aggregazione ha portato alla definizione di un tracciato comune di codifica delle informazioni preesistenti e nuove
- Nel corso della definizione del subset di elementi EAD sono emersi alcuni punti critici legati all'uso e alla codifica di
 - codice identificativo, segnature e altri sistemi di classificazione
 - elementi per la descrizione fisica
 - descrizione del contenuto
 - chiavi di accesso
 - altri elementi descrittivi: caratteristiche materiali e requisiti tecnici, strumenti di ricerca, bibliografia
 - il codice univoco dell'unità di descrizione

Codice identificativo e segnatura: definizioni

- Definizione ISAD 3.1.1 - codice identificativo/segnatura
 - *Scopo*: Identificare univocamente l'unità di descrizione e stabilire un collegamento con la descrizione che la rappresenta.
 - *Regola*: Indicare i seguenti elementi, secondo ciò che è necessario ai fini di una identificazione univoca:
 - il codice del paese secondo la versione più recente della norma *UNI EN 23166:1995, Codici per la rappresentazione dei nomi dei paesi*;
 - il codice dell'istituzione archivistica in conformità alle norme nazionali, o un altro elemento univoco di identificazione;
 - una specifica segnatura locale, un codice di controllo, o un altro elemento univoco di identificazione.
- Definizione EAD
 - <unitid>: «Qualunque stringa alfanumerica che funga da riferimento univoco o da numero di controllo per i materiali descritti [...] è in primo luogo un'indicazione logica [...]»
- Altri sistemi di numerazione: numero di ordinamento o di corda, posizione fisica, altre segnature

Altri sistemi di numerazione: definizioni

- **numero di corda**
 - P. Carucci, *Le fonti archivistiche. Ordinamento e conservazione*: «la numerazione progressiva che collega le unità archivistiche di un archivio, o di un complesso archivistico[...]»
 - Definizione EAD di <num>: «Un elemento generico per informazioni di carattere numerico in qualunque forma. L'elemento Numero <num> è utilizzato solo ove necessario per presentare un numero in una maniera particolare o per identificarlo con l'attributo TYPE. [...]»
- **posizione fisica**
 - Definizione EAD di <container>: «[...] le informazioni che aiutano a localizzare i materiali descritti indicando i tipi di contenitore in cui sono conservati i materiali e individuando eventuali numeri seriali assegnati a tali contenitori»
- **altre segnature**
 - Definizione EAD di <odd> (Other Descriptive Data): elemento di codifica dell'informazione sul materiale descritto che non è possibile riportare in altri sottoelementi di <c>. [...] L'utilizzo dell'attributo TYPE consente di specificare la modalità di utilizzo di <odd>

Codici identificativi, segnature e altri sistemi di numerazione: esempio

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<c id="IT-ER.IBC-AS00081-0000224" level="file">
  <did>
    <unitid countrycode="IT" encodinganalog="ISAD 1 - 1 reference code" repositorycode="ER.IBC" identifier="AS00081" label="ASTCom
    Ozzano nell'Emilia">081.001.007.001 </unitid>
    <unittitle encodinganalog="ISAD 1 - 2 title">"Costruzione dell'acquedotto nella frazione di Villanova"
      <num type="ordinamento">368
        <emph>45</emph>
      </num>
      <unitdate encodinganalog="ISAD 1 - 3 date(s)" normal="19210208-19230220">08 febbraio 1921 - 20 febbraio 1923
        <emph>con antecedente del 1910</emph>
      </unitdate>
    </unittitle>
    <container label="posizione fisica" type="b.">3</container>
    <container label="posizione fisica" type="fasc.">1</container>
  </did>
  <descgrp type="codici e numerazioni">
    <odd type="segnatura precedente">
      <p>b. 89</p>
      <note>
        <p>[riordino 1960]</p>
      </note>
    </odd>
  </descgrp>
</c>
```

Descrizione fisica: definizioni

- Definizione ISAD 3.1.5 - consistenza e supporto dell'unità di descrizione (quantità, volume, dimensione fisica)
 - *Scopo*: Identificare e descrivere:
 - a) la consistenza materiale o logica;
 - b) il supporto dell'unità di descrizione.
 - *Regola*: Segnalare la consistenza dell'unità di descrizione, dando, in cifre, il numero totale delle unità materiali o logiche e l'unità di misura. Indicare lo/gli specifico/i supporto/i dell'unità di descrizione.[...]
- Definizione EAD
 - <physdesc>: «Un elemento di aggregazione per raggruppare le informazioni relative all'aspetto o alla struttura fisica dei materiali descritti, come ad esempio le dimensioni, la quantità, lo spazio occupato, il genere, la forma, le funzioni e ogni altro aspetto materiale, come il colore, la sostanza, lo stile, la tecnica e il metodo di produzione. Le informazioni possono essere presentate come testo semplice o distribuite all'interno dei sottoelementi Dimensioni <dimensions>, Consistenza <extent>, Tipologia e forma <genreform> e Aspetto fisico <physfacet>»

Descrizione fisica: esempio

```
<?xml version="1.0" encoding="ISO-8859-1"?>  
<c id="IT-ER.IBC-AS00081-0000224" level="file">  
  <did>
```

[...]

```
    <physdesc encodinganalog="ISAD 1 - 5 extent and medium of the unit of description" label="container">  
      <genreform>fascicolo</genreform>  
      <physfacet type="medium">cartaceo</physfacet>  
      <physfacet type="note">all'interno del fascicolo è presente l'elenco del contenuto</physfacet>  
    </physdesc>  
    <physdesc encodinganalog="ISAD 1 - 5 extent and medium of the unit of description" label="content">  
      <extent>47</extent>  
      <genreform>cc.</genreform>  
      <physfacet type="note">le carte sono dattiloscritte e numerate, manca l'ultima pagina</physfacet>  
    </physdesc>  
    <physdesc encodinganalog="ISAD 1 - 5 extent and medium of the unit of description" label="content">  
      <extent>1</extent>  
      <genreform>disegno</genreform>  
      <dimensions unit="mm">  
        <dimensions type="altezza">250</dimensions>  
        <dimensions type="base">210</dimensions>  
      </dimensions>  
      <physfacet type="scala">1:5000</physfacet>  
    </physdesc>  
    <physdesc encodinganalog="ISAD 1 - 5 extent and medium of the unit of description" label="content">  
      <extent>3</extent>  
      <genreform>ritagli di giornale</genreform>  
    </physdesc>
```

[...]

```
  </did>
```

[...]

```
</c>
```

Descrizione del contenuto: definizioni

- Definizione ISAD 3.3.1 - ambiti e contenuto
 - *Scopo*: Permettere agli utenti di valutare la potenziale rilevanza dell'unità di descrizione
 - *Regola*: Illustrare sinteticamente gli ambiti (cronologici e geografici) e il contenuto (come la tipologia documentaria, le materie trattate, le procedure amministrative) dell'unità di descrizione, pertinenti al livello di descrizione.
- Definizione EAD
 - <scopecontent>: «Un testo che riassume in forma discorsiva gli ambiti e i soggetti coperti dai materiali descritti, spesso con la menzione della forma e dell'ordinamento dei materiali, nonché delle organizzazioni, degli individui, degli eventi, dei luoghi e dei soggetti di rilievo. L'obiettivo dell'elemento Ambiti e contenuto <scopecontent> è di aiutare i lettori a valutare la potenziale rilevanza dei materiali per la loro ricerca.[...]»

Descrizione del contenuto: esempio

```
<?xml version="1.0" encoding="ISO-8859-1"?>  
<c id="IT-ER.IBC-AS00081-0000224" level="file">
```

```
[...]
```

```
  <descgrp encodinganalog="ISAD 3 content and structure area">  
    <scopecontent encodinganalog="ISAD 3 - 1 scope and content">  
      <p>Carteggio relativo all'esproprio del terreno; progetto (relazione tecnica, computo metrico-estimativo, disegno del manufatto) redatto dall'ingegnere Aldo Donini; copie di deliberazioni consiliari sull'oggetto; articoli di giornale relativi al contenzioso sorto tra il Comune di Borgo San Renzo e Romeo Ferrari, proprietario del terreno espropriato.</p>  
      <list>  
        <head>Materiale a stampa allegato</head>  
        <item>  
          <bibref>  
            <persname>Silvio Fantini</persname>  
            <title>Sulla necessità di un acquedotto rurale</title>  
            <imprint>  
              <date>1910</date>  
              <geogname>Borgo San Renzo</geogname>  
              <publisher>Tipografia civica</publisher>  
            </imprint>  
            <num>p. 20</num>  
          </bibref>  
        </item>  
      </list>  
      <list>  
        <head>Materiale cartografico allegato</head>  
        <item>  
          <unittitle>"Torre dell'acquedotto"</unittitle>  
          <date>1921</date>  
        </item>  
      </list>  
      <list>  
        <head>Altri allegati</head>  
        <item>Articoli sul contenzioso relativo all'esproprio del terreno per la costruzione dell'acquedotto  
          <unittitle>"Lite sul podere Cincirella"</unittitle>  
          <date>1922</date>  
          <genreform>ritagli di giornale</genreform>  
          <num>3</num>  
          <note>  
            <p>Con annotazioni manoscritte.</p>  
          </note>  
        </item>  
      </list>  
    </scopecontent>  
  </descgrp>
```

```
[...]  
</c>
```

Chiavi di accesso: definizione EAD

- <controlaccess> (Controlled Access Headings):
 - «Un elemento di aggregazione che individua delle chiavi d'accesso per i materiali descritti e consente la ricerca per termini controllati su mezzi di corredo distribuiti su una rete informatica»[...], codificabili nei sottoelementi <corpname>, <famname>, <function>, <genreform>, <geogname>, <occupation>, <persname>, <subject>, ecc.[...]
 - L'attributo AUTHFILENUMBER può essere utilizzato per identificare un record di un authority file, che presenta informazioni aggiuntive sul termine codificato. L'attributo ROLE può essere invece usato per fornire informazioni specifiche sulle voci codificate, come la tipologia o una specifica qualifica, anche nei confronti del materiale descritto

Chiavi di accesso: esempio

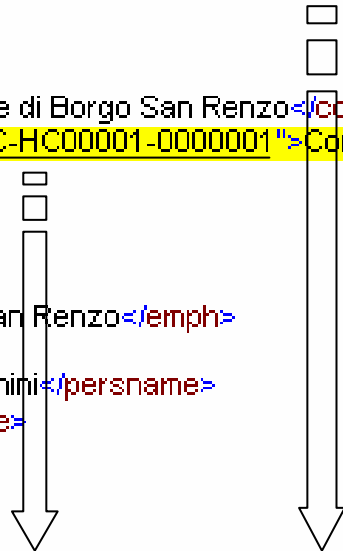
```
<?xml version="1.0" encoding="ISO-8859-1"?>  
<c id="IT-ER.IBC-AS00081-0000224" level="file">  
  <did>
```

[...]

```
    <origination encodinganalog="ISAD 2 - 1 name of creator(s)">  
      <corpname authfilenumber="IT-ER.IBC-HC00001-0000432">Comune di Borgo San Renzo</corpname>  
    </origination>  
  </did>  
  <controlaccess>  
    <corpname role="ente locale">Comune di Borgo San Renzo</corpname>  
    <corpname authfilenumber="IT-ER.IBC-HC00001-0000001">Comune di Fontanelice</corpname>  
    <geogname>Cincirella  
      <emph>podere</emph>  
    </geogname>  
    <geogname>Villanova  
      <emph>frazione di Borgo San Renzo</emph>  
    </geogname>  
    <persname role="ingegnere">Aldo Donini</persname>  
    <persname>Romeo Ferrari</persname>  
    <persname>Silvio Fantini</persname>  
  </controlaccess>
```

[...]

```
</c>
```



Il codice riportato nell'attributo AUTHFILENUMBER richiama la banca dati separata dell'authority file (in questo caso degli enti) condiviso da tutti gli archivi presenti nel sistema informativo

Chiavi di accesso e authority files

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<eac type="corpname">
  <eachheader detaillevel="minimal" status="draft">
    <eacid countrycode="IT" ownercode="ER.IBC">IT-ER.IBC-HC00001-0000001 </eacid>
    <mainhist>
      <mainevent maintype="create">
        <maindesc>Comune</maindesc>
        <maindate>02/04/2007</maindate>
        <name>Administrator Regesta.exe</name>
      </mainevent>
    </mainhist>
    <languagedecl>
      <language>italiano</language>
    </languagedecl>
  </eachheader>
  <condesc>
    <identity>
      <conhead normal="Comune di Fontanelice" type="authorized">
        <part type="ord">Fontanelice</part>
        <part type="add">Comune di</part>
      </conhead>
    </identity>
    <desc>
      <bioghist>
        <p>Il luogo in cui sorge Fontanelice risulta abitato fin dall'epoca preistorica ed esistono reperti archeologici che testimoniano la presenza umana nei periodi villanoviano, etrusco, celtico e romano. Tuttavia per avere una prima menzione certa della località nei documenti [...]</p>
      </bioghist>
    </desc>
    <resourcerels>
      <resourcerel reltype="origination">
        <archunit id="IT-ER.IBC-AS00081-0000224">
          <unittitle>"Costruzione dell'acquedotto nella frazione di Villanova"</unittitle>
        </archunit>
      </resourcerel>
    </resourcerels>
  </condesc>
</eac>
```

Caratteristiche materiali e requisiti tecnici: definizioni

- Definizione ISAD 3.4.4
 - *Scopo*: Fornire informazioni su tutte le caratteristiche materiali e i requisiti tecnici che siano rilevanti ai fini dell'utilizzazione dell'unità di descrizione.
 - *Regola*: Indicare tutte le condizioni materiali, che influiscano sull'utilizzazione dell'unità di descrizione, quali ad esempio esigenze di conservazione. [...]
- Definizione EAD:
 - <phystech>: «Una descrizione di rilevanti condizioni o caratteristiche fisiche che interessano l'archiviazione, la conservazione o l'uso dei materiali descritti. Comprende dettagli sulla loro composizione fisica o sulla necessità di specifici strumenti hardware o software per la conservazione o l'accesso ai materiali»

Caratteristiche materiali e requisiti tecnici: esempio

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<c id="IT-ER.IBC-AS00081-0000224" level="file">
[...
  <descgrp encodinganalog="ISAD 4 conditions of access and use area">
    <phystech encodinganalog="ISAD 4 - 4 physical characteristics and technical requirements">
      <p>stato di conservazione discreto</p>
      <phystech type="causa">
        <p>fuoco</p>
      </phystech>
      <phystech type="danno">
        <p>macchie</p>
      </phystech>
    </phystech>
  </descgrp>
[...
</c>
```

Strumenti ricerca: definizioni

- Definizione ISAD 3.4.5
 - *Scopo*: Segnalare tutti gli strumenti di ricerca relativi all'unità di descrizione.
 - *Regola*: Dare informazioni su tutti gli strumenti di ricerca esistenti presso l'istituto archivistico o presso il soggetto produttore che possano fornire informazioni sul contesto e il contenuto dell'unità di descrizione. Se opportuno, specificare a chi rivolgersi per ottenere copia di tali strumenti.
- Definizione EAD
 - <otherfindaid> (Other finding Aid): « Informazioni su guide supplementari o alternative ai materiali descritti, come ad esempio schedari, cataloghi di vendita o elenchi creati dal soggetto che ha prodotto o redatto i documenti. Tale elemento è utilizzato per indicare l'esistenza di ulteriori mezzi di corredo, ma non è adatto per codificare il contenuto di tali guide. [...] Si può utilizzare l'elemento Riferimento archivistico <archref> per fornire una citazione formale di un altro mezzo di corredo o per fare un collegamento a una sua versione online»

Strumenti ricerca: esempio

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<c id="IT-ER.IBC-AS00081-0000224" level="file">
[...
  <descgrp encodinganalog="ISAD 4 conditions of access and use area">
    <otherfindaid encodinganalog="ISAD 4 - 5 finding aids">
      <p>[testo libero]</p>
      <bibref>
        <persname>[autore]</persname>
        <title>[titolo citazione]</title>
        <imprint>
          <geogname>[luogo di pubblicazione]</geogname>
          <date>[data]</date>
          <publisher>[editore]</publisher>
        </imprint>
        <num>[pagina/e]</num>
      </bibref>
      <archref href="URL">
        <unittitle>[titolo della fonte web]</unittitle>
        <abstract>[descrizione]</abstract>
      </archref>
    </otherfindaid>
  </descgrp>
[...
</c>
```

Bibliografia: definizioni

- Definizione ISAD 3.4.5
 - *Scopo*: Identificare tutte le pubblicazioni che si riferiscano all'unità di descrizione o siano basate sullo studio, l'analisi od una qualche altra forma di utilizzazione di essa.
 - *Regola*: Citare e/o fornire informazioni sulle pubblicazioni che si riferiscono all'unità di descrizione o sono basate sullo studio, l'analisi od una qualche altra forma di utilizzazione di essa. Includere riferimenti a facsimili o a trascrizioni edite.
- Definizione EAD
 - <bibliography>: «Citazioni di lavori che si basano o che riguardano i materiali descritti; o che assumono particolare rilievo ai fini del loro utilizzo; o nelle quali è presente un riferimento o una breve descrizione dei materiali stessi. [...]. L'elemento Bibliografia <bibliography> può consistere in una lista semplice (utilizzando l'elemento Lista <List>) o in una lista di riferimenti bibliografici e archivistici (utilizzando rispettivamente gli elementi Riferimento bibliografico <bibref> e Riferimento archivistico <archref>), oppure in una serie di periodi (elemento <p>)»

Bibliografia: esempio

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<c id="IT-ER.IBC-AS00081-0000224" level="file">
[...]
  <descgrp encodinganalog="ISAD 5 allied materials area">
    <bibliography encodinganalog="ISAD 5 - 4 publication note">
      <p>[testo libero]</p>
      <bibref>
        <persname>Amedeo Morini</persname>
        <title>Storia delle frazioni di Borgo San Renzo: Villanova</title>
        <edition>Corriere di Borgo San Renzo</edition>
        <imprint>
          <geogname>Borgo San Renzo</geogname>
          <date>1988</date>
          <publisher>Tipografia civica</publisher>
        </imprint>
        <num>pp. 10-12</num>
      </bibref>
    </bibliography>
  </descgrp>
</c>
[...]
```

Codice univoco dell'unità di descrizione

```
<c id="IT-ER.IBC-AS00081-0000224" level="file">
```

- il codice alfanumerico univoco viene attribuito automaticamente dal sistema informatico a ciascuna unità di descrizione
- consente il recupero in fase di ricerca, anche di record cancellati, ed il collegamento con gli authority files e con altri sistemi informativi, anche non XML
- Esempi: ASBO-SIAS, IBC-CAStER