

Esperienze sull'interoperabilità tra conservatori accreditati

Gabriele Bezzi

Forum della Conservazione Agid
5° incontro Roma 12 dicembre 2018

Logica di “concentrazione” e di elevata specializzazione

Il Parer si configura come **un archivio unico** di tipo federato a cui gli enti produttori aderenti **conferiscono i proprio documenti digitali** usufruendo al contempo di un servizio di alto livello professionale, sia tecnologico che archivistico

Due fondamentali motivazioni:

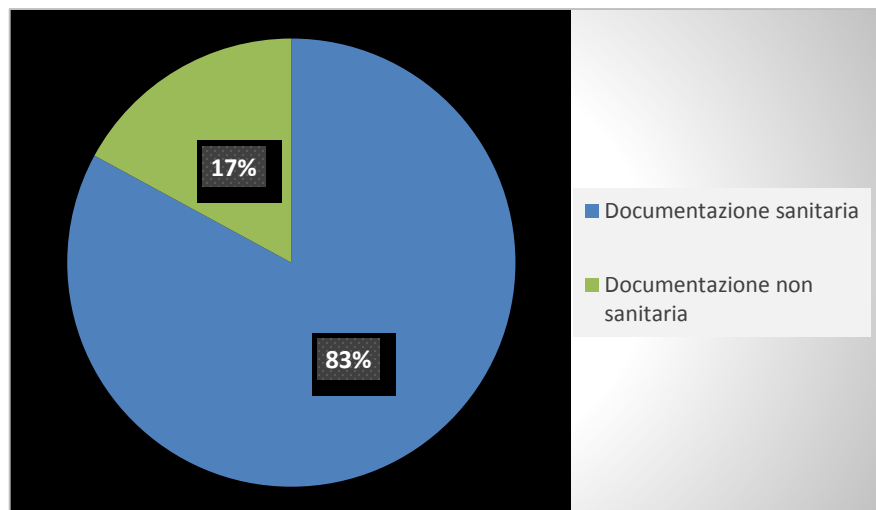
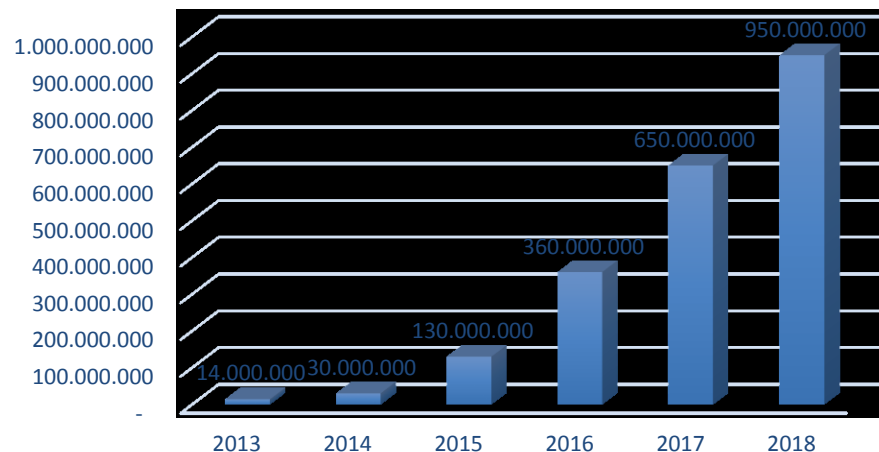
- **Il contenimento dei costi** per funzioni particolarmente onerose per l'ente produttore evitando il moltiplicarsi di strutture di conservazione
- **Il presidio archivistico tecnologico e organizzativo** su attività che implicano professionalità, risorse e strutture adeguatamente formate e gestite

Gli interventi necessari non sono oggi sostenibili da una singola amministrazione

Parer: i numeri a dicembre 2018

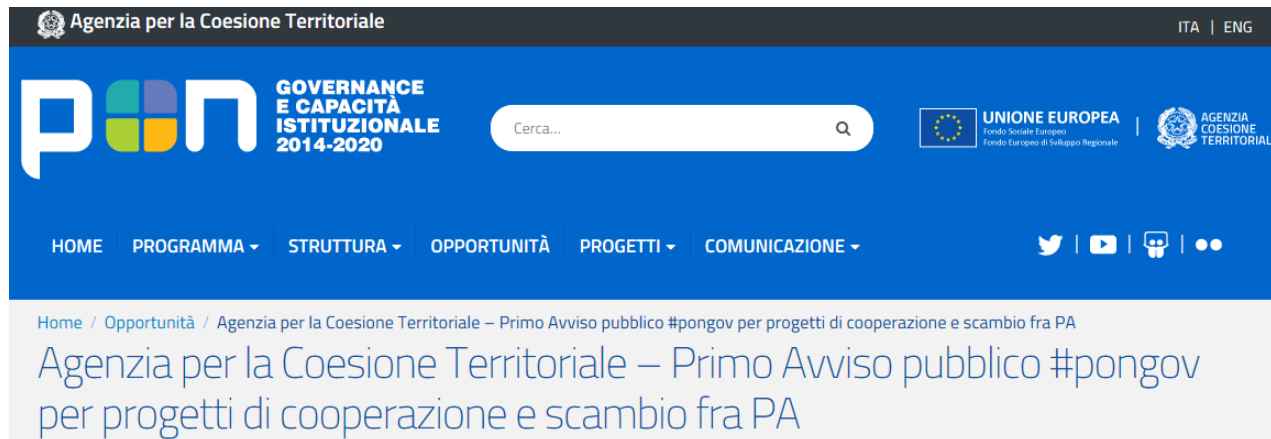
Trend di crescita

Documenti in conservazione



Prevalenza della
documentazione sanitaria

Partecipazione di Parer al bando PAOC 2020 dell'agenzia della coesione territoriale



Con **Provincia autonoma di Trento, Regione Val D'Aosta, Regione Puglia e Comune di Padova** assieme alla **Regione Emilia-Romagna** è stato presentato il 7 giugno 2017 **un progetto** in risposta all' Avviso per il finanziamento di interventi volti al trasferimento, all'evoluzione e alla diffusione di buone pratiche attraverso Open Community PA 2020 (**Bando PAOC 2020**) della Agenzia per la Coesione Territoriale

L' 8 febbraio 2018 è giunta comunicazione che Il progetto, denominata **RICORDI (Riuso Conservazione Record Digitali)**, presentato al BANDO PAOC2020 è arrivato **terzo tra tutti i progetti presentati e sarà finanziato.**

Il progetto Ricordi è attualmente in corso

Rapporti con altri conservatori

- Pacchetti di Archiviazione Unisincro trattati
 - 6 conservatori accreditati
 - 15 produttori
 - 9 tipologie documentali: Registri giornalieri di protocollo, Fatture, Documenti protocollati, Contratti, Determine, Atti, Disposizioni

- Pacchetto di archiviazione
- File Indice
- Soggetto produttore e Conservatore
- Informazioni sull'impacchettamento
- Contenuto pacchetto di archiviazione
- Composizione documenti
- File Schema XML di MoreInfo
- MoreInfo
- Errori vari

Sono diverse le composizioni del pacchetto che può riferirsi sia ad un singolo file che a numerosi documenti di diverse tipologie.

Non vi sono elementi nell'indice o documentazione a supporto che dichiarino tale composizione: pochissimo usato VdcGroup (description) e anche FileGroup.

Ciò comporta una difficoltà iniziale anche solo a capire cosa contiene il pacchetto di archiviazione e con quali criteri è stato costruito, rendendo molto onerosa la gestione della procedura di acquisizione.

Il Pacchetto di archiviazione dovrebbe essere (secondo OAIS) l'unità elementare (AIU) con cui è costituito l'archivio, e anche Unisincro definisce il volume di conservazione (VdC) «unità logica elementare», ma senza una relazione con una articolazione d'archivio.

Spesso è concepito come lotti di versamento nell'arco di un determinato periodo di tempo.

È possibile che all'interno di un pacchetto di archiviazione sia molto difficoltoso individuare il file dell'Indice (IDpA o IdC) perché posizionato assieme ai file degli oggetti conservati e con il nome dello stesso non ricavabile mediante espressione regolari o procedimenti simili.

In questi casi occorre implementare delle logiche ad hoc, non sempre agevoli, che rendono ulteriormente onerosa la procedura di acquisizione del PdA (o VdC).

Mentre nell'Indice (IPdA o IdC) è definito con chiarezza (semantica) chi ha prodotto il PdA, l'identificazione di Conservatore e Soggetto produttore potrebbe non essere esplicitata.

Non è definita la figura del firmatario come appartenente al Conservatore o Responsabile della Conservazione del soggetto produttore

Essendo questi i due soggetti sicuramente presenti nel processo di conservazione, Conservatore e Produttore dovrebbero essere chiaramente (e obbligatoriamente) identificati o identificabili nell'Indice stesso per garantire la corretta collocazione dei documenti conservati nell'archivio del Produttore, vero titolare dei documenti.

Informazioni sull'impacchettamento

I pacchetti sono strutturati in modi anche molto diversi tra loro e le informazioni contenute nell'Indice non sempre consentono di ricavare le informazioni sull'impacchettamento.

Questo è particolarmente rischioso quando i pacchetti di archiviazione sono logici e non fisici, ad esempio sono veicolati in un pacchetto di distribuzione che prevede tutti gli Indici in una cartella e tutti i file dei documenti in un'altra cartella.

Lo standard UNISInCRO mette a disposizione varie alternative per associare il file oggetto di conservazione al relativo pacchetto (e al relativo set di metadati nell'Indice): l'ID del file e il path.

Qui sotto un estratto da uno degli esempi proposti nel documento UNI:

```
<sincro:File sincro:encoding="binary" sincro:format="application/pkcs7-mime">  
  <sincro:ID>12345678</sincro:ID>  
  <sincro:Path>file://REF_ID_12345678/REF_ID_12345678.bin</sincro:Path>  
  [...]  
</sincro:File>
```

Informazioni sull'impacchettamento

Purtroppo nella maggior parte dei casi trattati l'ID è valorizzato con quelli che presumibilmente sono gli identificativi dei file nel sistema di conservazione, senza una chiara specificazione del sistema di riferimento mentre il path, essendo facoltativo, spesso non è presente nell'Indice.

Qui un esempio:

```
<sincro:File sincro:encoding="binary" sincro:format="application/pdf">  
<sincro:ID sincro:scheme="Ref">342c923a-50db-48e9-bdc2-  
5601ce54ace6</sincro:ID>  
<sincro:Hash sincro:function="SHA-  
256">Pbl4Ew70PCUitP3F01Cx4JQile0l2FUgw=</sincro:Hash>  
<sincro:MoreInfo sincro:XMLScheme="http://xsd.it">  
[...]  
</sincro:File>
```

Informazioni sull'impacchettamento

Per cui l'associazione tra file e Indice PdA deve essere ricostruita in altri modi, ad esempio mediante l'hash (ma questo è praticabile solo se i file da sottoporre a hash sono di un numero ragionevole), oppure, come avviene più frequentemente, utilizzando un metadato contenuto nelle MoreInfo del File in cui si può empiricamente presumere che contenga il nome del file stesso (con tutte le incertezze del caso).

Ad esempio:

```
<sincro:MoreInfo sincro:XMLScheme="http://xsd.it">  
<sincro:EmbeddedMetadata>  
<doc>  
<Data_Registrazione>20160525</ Data_Registrazione >  
<File_Name>1245_2424_125.xml</File_Name>  
[...]
```

Spesso le unità documentarie contenute nei PdA (o VdC) sono composte da più file (es.: Fattura e Ricevute SDI) e il meccanismo previsto in Unisincro per trattare questi casi è l'utilizzo del *FileGroup*.

Però *FileGroup* è usato pochissimo: nella maggior parte dei casi la struttura degli oggetti conservati viene appiattita a livello di File (con la regola «un file, un documento»).

Ciò rende non sempre agevole ricondurre i singoli file alla loro struttura documentale effettiva, se non utilizzando i metadati di MoreInfo, ma anche in questo caso con tutte le aleatorietà del caso.

Descrizione Oggetti conservati

La mancanza di una vera e propria semantica condivisa che descriva gli oggetti conservati costituisce una delle principali difficoltà ad acquisire in modo adeguato i pacchetti di archiviazione.

L'Indice potrebbe (e in alcuni casi, effettivamente, lo fa) non fornire informazioni essenziali sugli oggetti conservati: tipologia, identificativi, vincoli archivistici e via dicendo.

Il risultato, paradossale, è che, per via della mancanza di una semantica condivisa (ontologia), le informazioni descrittive di un oggetto perdono via via di qualità a ogni passaggio da un sistema di conservazione all'altro.

File dello Schema XML di MoreInfo

I metadati di MoreInfo spesso si presentano come unica fonte per ricostruire le informazioni sull'impacchettamento o per determinare il contesto documentale in cui il file deve essere considerato.

L'utilizzo di MoreInfo a questi scopi, che già di per sé presenta profili evidenti di criticità, è reso ancora più complesso dal fatto che, nell'Indice, gli Schema XML delle MoreInfo sono individuati/individuabili da un attributo che contiene l'URL.

Nella quasi totalità degli esempi trattati, l'URL riconduceva a una pagina inesistente o comunque non contenente lo Schema XML cui si riferiva.

File dello Schema XML di MoreInfo

Una soluzione che alcuni hanno adottato è quella di ricomprendere gli Schema XML nel pacchetto stesso, soluzione che sembrerebbe particolarmente robusta in quanto renderebbe il pacchetto autoconsistente anche in funzione delle strutture dei metadati MoreInfo utilizzate.

In questi casi, normalmente, l'attributo viene valorizzato non come URL ma come path, soluzione non esplicitamente prevista nello standard

<sincro:MoreInfo sincro:XMLScheme="./XSD/Documento.xsd">


Si registrano negli Indici diversi errori, soprattutto nel definire i Path e nell'utilizzo dei Namespace che rendono necessari introdurre parametri per correggere tali errori e consentire in tale modo alla procedura di acquisizione di leggere correttamente questi valori.

Sito internet PArER: <http://parer.ibc.regione.emilia-romagna.it>


Regione Emilia-Romagna
ENG


ParER
Polo archivistico dell'Emilia-Romagna

Seguici su








Il Polo archivistico ▾
Conservazione ▾
Contatti ▾

Conservazione delle fatture elettroniche

Comunicazione agli enti che intendono conservare i documenti a rilevanza fiscale o li stiano già conservando, ma non hanno ancora completato le attività di versamento relative...




Vai a:
[Archivio notizie in evidenza](#)

In evidenza




Sistema di conservazione



Help Desk



Documentazione



Conservazione fatture

I numeri del ParER


940.360.085

Documenti conservati da ParER al 10/12/2018

ParER news



Una sentenza sulla PEC



Blockchain in ambito sanitario



Gli archivi fotografici

[Archivio](#)


Newsletter e Comunità web sugli archivi digitali



Notizie e approfondimenti dal mondo
degli archivi e della conservazione digitale

numero 330 - 6/12/2018



[Online gli inventari di tre fondi dedicati a Pasolini](#)

Sul portale IBC Archivi l'Archivio dell'Associazione Fondo Pasolini, e le Carte Laura Betti e Sergio Citti



[L'archivista digitale è un middle manager](#)

Un "identikit" della figura professionale, a cura di Donatella Mazzetto, per Agenda Digitale

Ci si può iscrivere dal sito internet del Parer



Grazie per l'attenzione