

Il quadro normativo nazionale e le sue criticità per
la conservazione a lungo termine

Lo standard UNI SInCRO

la conservazione “anticipata” e le sue criticità

- La normativa nazionale sulla conservazione sostitutiva ha portato notevole **confusione nelle procedure** di gestione documentaria
- Ha spinto le amministrazioni ad affidare precocemente all'esterno i propri archivi (documenti?) digitali senza chiarire la necessità e le modalità di una **sedimentazione archivistica** governata ma stabilendo obblighi di **interoperabilità dei servizi di firma e di marcatura**
- Le serie di documenti che si producono ‘fisicamente’ nei processi in corso di conservazione sostitutiva non corrispondono alle serie archivistiche che riflettono il flusso di lavoro e di produzione documentaria, ma riguardano i documenti sottoposti a digitalizzazione o i documenti firmati digitalmente a cura del responsabile della conservazione (**lotti o volumi di conservazione** gestiti con **indici di conservazione** riferiti alla procedura di chiusura dei documenti memorizzati a fini di conservazione sostitutiva e alla creazione e gestione di **file di chiusura** previsti dalle delibere Cnipa)

volumi o lotti di conservazione

- Nei sistemi di conservazione sostitutiva un volume o un lotto di conservazione è costituito **dall'insieme dei file** oggetto di conservazione e da un **file ausiliario (indice o file di chiusura)** che contiene informazioni identificative per ogni file conservato (almeno un codice identificativo e l'impronta).
- A fini di integrità l'indice deve essere firmato digitalmente dal responsabile della conservazione e corredato di un riferimento temporale opponibile ai terzi, che può essere una marca temporale.
- L'indice, con firma e riferimento temporale, è finalizzato ad assicurare **l'inalterabilità** di ogni file memorizzato nel sistema di conservazione (grazie all'impronta), la sua **identificazione** mediante un codice univoco e la **data certa** di memorizzazione nel sistema mediante la marcatura temporale.
- Si tratta di uno strumento definito dalla delibera Cnipa del 2004 e sviluppato dai fornitori dei servizi di conservazione sostitutiva al fine di assicurare **garanzie di controllo** al processo di acquisizione e tenuta di file digitali firmati.

consolidamento di documenti sottoscritti e conservazione

- Poiché un volume o un lotto di conservazione è obbligatoriamente corredato di un riferimento temporale opponibile ai terzi (marcatura temporale), tale procedura viene utilizzata non solo al fine di digitalizzare documenti analogici ma anche per risolvere il problema del **consolidamento** (processo di apposizione di una marca temporale a un documento dopo averne verificato la validità in termini di certificato non scaduto, non revocato o sospeso e integrità del contenuto) dei documenti informatici sottoscritti (considerata la difficoltà organizzativa e tecnologica delle amministrazioni e dei privati).
- “Tuttavia, poiché in un volume si accumulano un certo numero di documenti e solo al termine dell’accumulo si applicano la firma del RdC ed il riferimento temporale, può trascorrere un tempo significativo tra la sottoscrizione del documento ed il riferimento temporale. Se una revoca o una scadenza di certificato capita in quell’intervallo, si perde il valore probatorio” (Baudizzone)

l'indice di conservazione o file di chiusura: contenuti e formato

- Le disposizioni regolamentari (dpcm 31 marzo 2009) non danno indicazioni specifiche sulle modalità operative di tale processo né sul formato del cosiddetto “**file di chiusura**” (il termine **indice** è citato in forme scarsamente comprensibili tra i compiti del responsabile della conservazione ed è frutto della inconsapevolezza di chi ha scritto quelle norme).
- **L'interoperabilità** è quindi a rischio sia in relazione a servizi forniti da fornitori diversi per quanto riguarda l'apposizione di firme e di marcature temporali sia con riferimento a sistemi di indicizzazione utilizzati nel tempo anche da parte dello stesso fornitore.

lo standard UNI SInCRO

- Lo standard SInCRO è stato elaborato dalla **Commissione UNI-DIAM** e in particolare dal comitato sulla gestione dei documenti e degli archivi per iniziativa di fornitori preoccupati per la qualità nel tempo dei servizi offerti (Ferri e Baudizzone)
- Ha la finalità di definire una **struttura di dati in XML** che consenta di predisporre sia le informazioni identificative minime (previste dal legislatore) che una infrastruttura generale in grado di gestire tutte le informazioni archivistiche necessarie al processo di formazione e tenuta dei documenti digitali in modo da assicurare l'interoperabilità tra sistemi e la conservazione a lungo termine
- **Presuppone un corretto processo a monte di analisi archivistica dei documenti da digitalizzare e conservare e un adeguato sistema di tenuta dei documenti negli archivi correnti**

lo standard UNI SInCRO: le ragioni

- Le regole tecniche per la conservazione sostitutiva descrivono gli aspetti procedurali e indicano le responsabilità degli attori di questo specifico processo, ma non forniscono **dettagli tecnici** sulle **modalità di rappresentazione dei dati e documenti oggetto di conservazione**, e non contengono alcuna specifica disposizione mirata a conseguire o a promuovere forme d'interoperabilità.
- Gli articoli 3 e 4 in particolare (dedicati alla conservazione sostitutiva di documenti informatici ed analogici) **si limitano a prescrivere**
 - l'uso della firma digitale e di un riferimento temporale per perfezionare il processo,
 - con l'intervento del responsabile della conservazione eventualmente integrato o sostituito da quello di un pubblico ufficiale.
 - Tecnicamente l'obbligo è limitato alla apposizione della firma digitale e del riferimento temporale *“sull'insieme dei documenti o su una evidenza informatica contenente una o più impronte dei documenti o di insiemi di essi”*.
- Libertà troppo ampia; rischi di auto-referenzialità e perdita degli investimenti già nel medio periodo

lo standard UNI SInCRO: gli scenari applicativi

evoluzione dei sistemi

- l'assenza d'interoperabilità rende difficile e costosa la migrazione verso nuovi sistemi e soluzioni tecnologiche, vincolando di fatto i soggetti possessori di archivi digitali ai fornitori di servizi e alle loro scelte, con ciò limitando l'auspicabile fluidità delle dinamiche di mercato

accesso ai documenti

- necessità di poter esibire (ad esempio in sede giudiziale, ma anche a mero scopo di consultazione da parte degli aventi diritto) i documenti conservati: in assenza d'interoperabilità e di standard di riferimento si richiedono laboriose perizie per leggere i dati generati da applicazioni *ad hoc*.

lo standard UNI SInCRO – obiettivi specifici

- La norma definisce la struttura dell'insieme di dati a **supporto del processo di conservazione sostitutiva**: in particolare, precisa e integra alcune disposizioni contenute nella deliberazione CNIPA 19 febbraio 2004, n. 11, individuando gli **elementi informativi necessari alla creazione dell'Indice di Conservazione** (il cosiddetto "file di chiusura") e descrivendone sia la semantica sia l'articolazione per mezzo del linguaggio formale XML.
- L'obiettivo della norma è di consentire agli operatori del settore di utilizzare una **struttura-dati condivisa al fine di raggiungere un soddisfacente grado d'interoperabilità nei processi di migrazione**, grazie all'adozione dello Schema XML appositamente elaborato.

lo standard UNI SInCRO: i vantaggi

- L'adozione di questo standard nello sviluppo di sistemi di conservazione “anticipata” è finalizzato a favorire:
 - **l'interoperabilità** nei sistemi attivi tra produttori diversi e a fini di esibizione
 - **l'aggiornamento** grazie alla presenza di uno schema di metadati aperto alle integrazioni nelle fasi successive di gestione
- E' **indipendente dal fornitore e dall'applicazione**

lo standard UNI SInCRO: definizioni tecniche

- **Volume di Conservazione (VdC):** Unità logica elementare risultato finale di un processo di conservazione sostitutiva. Il VdC è composto logicamente da:uno o più *file* ai quali si applica unitariamente il processo di conservazione sostitutiva, l'Indice di Conservazione (IdC), gli Indici di Conservazione antecedenti, se l'Indice di Conservazione attuale è stato originato da questi.
In aggiunta ai precedenti elementi, il VdC può contenere ulteriori componenti, per lo più con finalità di carattere gestionale
- **Indice di Conservazione (IdC):** Evidenza informatica associata ad ogni Volume di Conservazione (VdC), contenente un insieme di informazioni articolate come descritto dallo Schema XML fornito nello standard.
L'IdC deve essere corredato da:riferimento temporale firma digitale dei soggetti intitolati a effettuare il processo di conservazione sostitutiva, coerentemente con le disposizioni della normativa vigente.
In pratica, l'IdC coincide con lo Schema XML descritto nello standard, istanziato secondo le specifiche esigenze di contesto e provvisto di riferimento temporale e firma digitale.
- **Riferimento temporale dell'IdC:** Informazioni contenenti data e ora di realizzazione dell'IdC.

lo standard UNI SInCRO: l'indice

SCOPO E CAMPO DI APPLICAZIONE
RIFERIMENTI NORMATIVI
TERMINI E DEFINIZIONI
L'INDICE DI CONSERVAZIONE: LA STRUTTURA
La rappresentazione grafica
La rappresentazione formale: DTD XML
La rappresentazione formale: Schema XML
L'INDICE DI CONSERVAZIONE: LA TAG LIBRARY
APPENDICI: ESEMPI, FONTI LEGISLATIVE, BIBLIOGRAFIA

lo standard UNI SInCRO: gli elementi

1. Agent
2. Agent_ID
3. AgentName
4. AttachedTimeStamp
5. CreatingApplication
6. Description
7. DetachedTimeStamp
8. EmbeddedMetadata
9. ExternalMetadata
10. File
11. FileGroup
12. FirstName
13. FormalName
14. Hash
15. ID
16. IdC
17. IdC_ID
18. Label
19. LastName
20. LawAndRegulations
21. MoreInfo
22. Name
23. NameAndSurname
24. Path
25. PreviousHash
26. Process
27. Producer
28. SelfDescription
29. SourceIdC
30. SourceVdC
31. TimeInfo
32. TimeReference
33. VdC
34. VdCGroup
35. Version

lo standard UNI SInCRO: gli attributi

1. CanonicalXML
2. Encoding
3. Extension
4. Format
5. Function
6. Language
7. Normal
8. OtherRole
9. OtherScheme
10. RelatedIdC
11. RoleScheme
12. Scheme
13. Type
14. Url
15. Version
16. XMLScheme

lo standard UNI SInCRO: qualche elemento in dettaglio - Agent

Agent: informazioni relative ai soggetti che intervengono nel processo di conservazione sostitutiva.

Attributi: *role*, *otherRole*, *type*

- *type* (person | organization) #REQUIRED

- *role* (Delegate | PreservationManager | Operator | PublicOfficer | OtherRole) #REQUIRED

- *otherRole* CDATA #IMPLIED

Elementi subordinati:

1. **Agent_ID:** codice per l'identificazione univoca (l'attributo di riferimento è *scheme* costituita da una lista semi-aperta che include il codice fiscale, la partita Iva ma che può essere ampliato con l'utilizzo di altri schemi *otherScheme*)
2. **AgentName:** elementi subordinati *FormalName*, *NameAndSurname*
3. **MoreInfo:** informazioni ulteriori, che non possono essere associate ad altri elementi; devono essere strutturate nel formato XML debitamente localizzata; le informazioni possono essere inserite direttamente all'interno dell'IDC nel sub-elemento `<EmbeddedMetadata>` o all'esterno `<ExternalMetadata>`, ovvero un altro file

lo standard UNI SInCRO: qualche elemento in dettaglio – AttachedTimeStamp

AttachedTimeStamp: Data e ora di realizzazione dell'Indice di Conservazione, in forma normalizzata, nel caso in cui questa sia testimoniata con una marca temporale *attached* all'IdC stesso.

Attributi: *normal* (per esprimere data e ora in forma normalizzata)

Elementi subordinati: [nessuno]

lo standard UNI SInCRO: qualche elemento in dettaglio – Description

Description: informazioni descrittive relative a una eventuale tipologia/aggregazione (di natura logica o fisica) cui il VdC appartiene

Attributi: *language* (per identificare il linguaggio in cui sono espresse le informazioni)

Elementi sovraordinati: VdCGroup

Elementi subordinati: [nessuno]

lo standard UNI SInCRO: qualche elemento in dettaglio – EmbeddedMetadata

- **Description:** Le informazioni dell'elemento <MoreInfo>, integrate all'interno dell'IdC e strutturate nel formato XML

Attributi: [nessuno]

Elementi sovraordinati: MoreInfo

Elementi subordinati: [nessuno]

lo standard UNI SInCRO: qualche elemento in dettaglio – ExternalMetadata

- **Description:** Le informazioni dell'elemento <MoreInfo>, contenute all'esterno dell'IdC in un *file* XML le cui caratteristiche sono descritte nei subelementi di <ExternalMetadata>. Trattandosi concretamente di un *file*, questo elemento ha la stessa struttura dell'elemento <File>, eventualmente sovrabbondante. Si noti che tale *file* può essere sì *esterno* all'IdC, ma deve comunque essere *interno* al VdC

Attributi:

- *encoding* (per fornire informazioni relative al tipo di codifica)
- *extension* (per indicare l'estensione del *filename*)
- *format* (per fornire informazioni sulla struttura-dati)

Elementi sovraordinati: MoreInfo

Elementi subordinati: Hash, ID, MoreInfo, Path, PreviousHash

lo standard UNI SInCRO: qualche elemento in dettaglio – File

- **Description:** Informazioni relative al file oggetto di conservazione sostitutiva

Attributi:

- *encoding* (per fornire informazioni relative al tipo di codifica)
- *extension* (per indicare l'estensione del *filename*)
- *format* (per fornire informazioni sulla struttura-dati)

Elementi sovraordinati: FileGroup

Elementi subordinati: Hash, ID, MoreInfo, Path, PreviousHash

lo standard UNI SInCRO: qualche elemento in dettaglio – FileGroup

- **Description:** Elemento di aggregazione di più *file* oggetto di conservazione sostituiva. Tale elemento è funzionale alla creazione di insiemi di *file* sulla base di criteri di ordine logico o tipologico

Attributi: [nessuno]

Elementi sovraordinati: IdC

Elementi subordinati: File, Label, MoreInfo

lo standard UNI SInCRO: qualche elemento in dettaglio – ID

- **Description:** identificatore univoco dell'entità descritta

Attributi: *scheme* (per specificare lo schema, la sintassi o in generale il sistema di riferimento all'interno del quale assume significato il valore dell'elemento)

Elementi sovraordinati: ExternalMetadata, File, SelfDescription, SourceIdC, SourceVdC, VdC, VdCGroup (ovvero tutte le entità che devono essere identificate univocamente)

Elementi subordinati: [nessuno]

lo standard UNI SInCRO: qualche elemento in dettaglio – IdC

- **Description:** informazioni relative ai *file* oggetto di conservazione sostitutiva e al processo di conservazione sostitutiva. Tali informazioni sono considerate stabili e indipendenti dai processi di trasferimento fra sistemi: come tali non possono essere modificate e sono ricomprese nel calcolo della firma digitale

Attributi:

- *url* (per localizzare lo Schema di riferimento dell'IdC)
- *version* (per gestire l'evoluzione dello standard e coadiuvare la compatibilità puntuale dei sistemi con le versioni precedenti)

Elementi sovraordinati: [nessuno]

Elementi subordinati: FileGroup, Process, SelfDescription, VdC

lo standard UNI SInCRO: qualche elemento in dettaglio – IdC_ID

- **Description:** identificatore univoco dell'Indice di Conservazione associato al **precedente** Volume di Conservazione oggetto della descrizione. Il valore dell'identificatore deve coincidere con il valore dell'elemento <ID> contenuto all'interno dell'elemento <SourceIdC> (è utilizzato per ricostruire la storia degli oggetti sottoposti al processi di conservazione)

Attributi:

- *scheme* (per specificare lo schema, la sintassi o in generale il sistema di riferimento all'interno del quale assume significato il valore dell'elemento)

Elementi sovraordinati: SourceVdC

Elementi subordinati: [nessuno]

lo standard UNI SInCRO: qualche elemento in dettaglio – Label

- **Description:** Nome dell'eventuale tipologia/aggregazione cui appartiene il File o il VdC

Attributi:

- *scheme* (per specificare lo schema, la sintassi o in generale il sistema di riferimento all'interno del quale assume significato il valore dell'elemento)

Elementi sovraordinati: FileGroup, VdCGroup

Elementi subordinati: [nessuno]

lo standard UNI SInCRO: qualche elemento in dettaglio – LawAndRegulations

- **Description:** Informazioni su norme, regolamenti e standard che guidano il processo di conservazione sostitutiva

Attributi:

- *language*

Elementi sovraordinati: Process

Elementi subordinati: [nessuno]

lo standard UNI SInCRO: qualche elemento in dettaglio – MoreInfo

- **Description:** Informazioni ulteriori, relative all'elemento padre di <MoreInfo>, che non è possibile associare ad altri elementi.

Tali informazioni devono essere strutturate nel formato XML, utilizzando uno Schema XML la cui localizzazione è specificata nell'attributo XMLSchema dell'elemento. L'insieme di queste informazioni costituisce un corpo che può essere inserito direttamente all'interno dell'IdC (nel subelemento <EmbeddedMetadata>) o all'esterno (nel subelemento <ExternalMetadata>). Nel secondo caso, tale insieme individua concretamente un *file* e pertanto l'elemento <ExternalMetadata> presenta la stessa articolazione dell'elemento <File>

All'interno dell'elemento <File>, il sub-elemento <MoreInfo> può essere utilizzato per descrivere i *file* oggetto di conservazione sostitutiva coerentemente con uno schema che utilizzi un insieme di metadati specializzato nel dominio: quindi l'elemento <MoreInfo>, qui definito il più genericamente possibile ai fini dell'interoperabilità, offre ampi margini d'intervento per l'elaborazione di *raffinamenti* ovvero per la definizione di strutture di metadati adeguate allo specifico contesto d'uso

Attributi:

- *XMLScheme* (per localizzare lo Schema XML della struttura di metadati. Il valore deve essere espresso nella forma di URL)

Elementi sovraordinati: Agent, ExternalMetadata, File, FileGroup, Process, SelfDescription, VdC (ovvero tutte le entità per le quali si prevedono informazioni identificative e descrittive specifiche)

Elementi subordinati: EmbeddedMetadata, ExternalMetadata

lo standard UNI SInCRO: qualche elemento in dettaglio – Name

- **Description:** Nome dell'applicazione che ha generato l'IdC
Attributi :[nessuno]
Elementi sovraordinati: CreatingApplication
Elementi subordinati: :[nessuno]

lo standard UNI SInCRO: qualche elemento in dettaglio – Path

- **Description:** Informazioni relative alla localizzazione del *file* cui l'elemento si riferisce, espressa come indirizzo URI. Il valore dell'elemento coincide con il path *relativo* del *file* rispetto alla posizione dell'IdC. **Tale elemento può essere valorizzato e assume significato solo nel caso in cui, nei processi di esportazione, i *file* siano memorizzati all'interno di un *file system*:** in presenza di altre architetture di sistema (ad esempio un database), il valore dell'elemento <ID> è ritenuto sufficiente all'identificazione puntuale del *file*

Attributi :[nessuno]

Elementi sovraordinati: ExternalMetadata, File, SourceIdC

Elementi subordinati: :[nessuno]

lo standard UNI SInCRO: qualche elemento in dettaglio – PreviousHash

- **Description:** Informazioni su precedenti impronte del *file* oggetto di conservazione sostitutiva (riferite a Indici di Conservazione precedenti) o del *file* di metadati (esterno all'IdC) che contiene le informazioni dell'elemento <MoreInfo>

Attributi:

- canonicalXML (per determinare, in caso di *file* XML, se questo sia stato trasformato nella forma canonica prima di essere sottoposto a hash)
- function (per identificare la funzione di hash utilizzata)
- relatedIdC (per identificare l'Indice di Conservazione associato alla precedente impronta)

Elementi sovraordinati: ExternalMetadata, File

Elementi subordinati: :[nessuno]

lo standard UNI SInCRO: qualche elemento in dettaglio – Process

- **Description:** Informazioni relative alle modalità di svolgimento del processo di conservazione sostitutiva

Attributi :[nessuno]

Elementi sovraordinati: IdC

Elementi subordinati: Agent, LawAndRegulations, MoreInfo, TimeReference

lo standard UNI SInCRO: qualche elemento in dettaglio – SelfDescription

- **Description:** Informazioni relative all'Indice di Conservazione stesso, associato al Volume di Conservazione

Attributi :[nessuno]

Elementi sovraordinati: IdC

Elementi subordinati: CreatingApplication, ID, MoreInfo, SourceIdC

lo standard UNI SInCRO: qualche elemento in dettaglio – SourceIdC

- **Description:** Informazioni relative a uno o più Indici di Conservazione da cui è originato quello in oggetto. Tali informazioni assumono rilievo per ricostruire la storia degli oggetti digitali sottoposti a conservazione. Il SourceIdC può riferirsi a:
 - una precedente versione dell'IdC attuale (ad esempio in caso di migrazione e/o modifiche del formato dei *file*, ove da un VdC si migri ad un nuovo VdC)
 - più IdC cronologicamente antecedenti che hanno generato per fusione l'IdC attuale (ad esempio in caso di riorganizzazione della struttura dell'archivio, ove più VdC vengano aggregati in un singolo VdC)
 - un IdC cronologicamente antecedente che per frammentazione ha generato l'IdC attuale (ad esempio in caso di scarto di documenti da un VdC, ove a partire da un VdC si generino più VdC).

Attributi :[nessuno]

Elementi sovraordinati: SelfDescription

Elementi subordinati: Hash, ID, Path

lo standard UNI SInCRO: qualche elemento in dettaglio – SourceVdC

- **Description:** Informazioni relative a uno o più Volumi di Conservazione da cui è originato quello in oggetto (ad esempio per migrazione di un Volume o per aggregazione di più Volumi).

Attributi :[nessuno]

Elementi sovraordinati: VdC

Elementi subordinati: ID, IdC_ID

lo standard UNI SInCRO: qualche elemento in dettaglio – TimeInfo

- **Description:** Informazioni relative a data e ora di realizzazione dell'Indice di Conservazione, nel caso in cui queste non assumano la forma di una marca temporale. Il valore dell'elemento deve essere nel formato ISO 8601, e più precisamente nella forma YYYY-MM-DDT00:00:00±00 (per l'Italia è di default +01)

Attributi:[nessuno]

Elementi sovraordinati: TimeReference

Elementi subordinati: [nessuno]

lo standard UNI SInCRO: qualche elemento in dettaglio – TimeReference

- **Description:** Informazioni relative a data e ora di realizzazione dell'Indice di Conservazione. Tale elemento funge da mero contenitore di elementi subordinati, per distinguere i seguenti casi:
 - generico riferimento temporale (usato per valorizzare l'elemento <TimeInfo>)
 - marca temporale *detached* (il cui indirizzo URI valorizza l'elemento <DetachedTimeStamp>)
 - marca temporale *attached* (all'elemento vuoto <AttachedTimeStamp> è associata la data in forma normale)

Attributi:[nessuno]

Elementi sovraordinati: Process

Elementi subordinati: AttachedTimeStamp, DetachedTimeStamp, TimeInfo

lo standard UNI SInCRO: qualche elemento in dettaglio – VdC

- **Description:** Informazioni relative al Volume di Conservazione
Attributi:[nessuno]
Elementi sovraordinati: IdC
Elementi subordinati: ID, MoreInfo, SourceVdC, VdCGroup

lo standard UNI SInCRO: qualche elemento in dettaglio – VdCGroup

- **Description:** Informazioni relative a una eventuale tipologia/aggregazione (di natura logica o fisica) cui il Volume di Conservazione appartiene

Attributi:[nessuno]

Elementi sovraordinati: VdC

Elementi subordinati: Description, ID, Label

a proposito del piano operativo della conservazione

- I manuali utilizzati dai fornitori di servizi si concentrano su alcuni nodi tecnologici previsti dalla normativa:
 - identificazione dei lotti firmati (VdC)
 - definizione di responsabilità interne (anche in termini di tracciamento): chi e compiti
 - documentazione di processo limitatamente ai processi interni di riversamento diretto (refreshment)
 - sicurezza fisica e logica del sistema e delle copie di sicurezza
 - gestione della leggibilità (ma non dell'intelligibilità)
- E' del tutto assente l'attenzione ai nodi dell'obsolescenza tecnologica che implica il riversamento indiretto o 'sostitutivo' (ad esempio in caso di obsolescenza dei formati)
- Manca qualunque riferimento alla gestione conservativa dei sistemi di gestione originaria dei documenti