

SPECIFICHE TECNICHE DEL PACCHETTO DI ARCHIVIAZIONE

Indice

1	INTRODUZIONE
2	STRUTTURA DELL'INDICE DEL PACCHETTO DI ARCHIVIAZIONE
3	VOCABOLARIO

1 INTRODUZIONE

Il presente allegato illustra la struttura descrittiva dell'indice del pacchetto di archiviazione.

Tale struttura fa riferimento allo standard SInCRO - Supporto all'Interoperabilità nella Conservazione e nel Recupero degli Oggetti digitali (UNI 11386:2010), che è lo standard nazionale riguardante la struttura dell'insieme dei dati a supporto del processo di conservazione.

In analogia allo standard SInCRO, la struttura di seguito descritta prevede una specifica articolazione per mezzo del linguaggio formale XML, per la cui applicazione pratica si rimanda allo standard stesso.

Per completezza, si avverte che ciò che in questo documento è denominato IPdA (Indice del Pacchetto di Archiviazione) nello standard SInCRO è indicato come IdC (Indice di Conservazione) e, analogamente, PdA (Pacchetto di Archiviazione) è indicato come VdC (Volume di Conservazione).

L'IPdA è l'evidenza informatica associata ad ogni PdA, contenente un insieme di informazioni articolate come descritto nel seguito. Deve essere corredato da un riferimento temporale e dalla firma digitale o firma elettronica qualificata del soggetto che interviene nel processo di produzione del pacchetto di archiviazione.

Entrando nel dettaglio, all'interno dell'elemento IPdA si trovano le seguenti strutture:

- informazioni generali relative all'indice del pacchetto di archiviazione: un identificatore dell'IPdA, il riferimento all'applicazione che l'ha creato, eventuali riferimenti ad altri IPdA da cui deriva il presente, e un eventuale elemento "ExtraInfo" che consente di introdurre metadati soggettivi relativi all'IPdA liberamente definiti dall'utilizzatore con un proprio schema;
- informazioni inerenti il Pacchetto di Archiviazione, in particolare: un identificatore del PdA, eventuali riferimenti ad altri PdA da cui deriva il presente, informazioni relative a una eventuale tipologia/aggregazione (di natura logica o fisica) cui il PdA appartiene e infine un eventuale elemento "ExtraInfo" che consente di introdurre metadati soggettivi relativi al PdA;
- indicazione di uno o più raggruppamenti di uno o più file che sono contenuti nel PdA. È possibile raggruppare file sulla base di criteri di ordine logico o tipologico ed assegnare ad ogni raggruppamento / singolo file le informazioni di base e un eventuale elemento "ExtraInfo" che consente di introdurre metadati definiti dall'utilizzatore. Ogni elemento file contiene l'impronta attuale dello stesso, ottenuta con l'applicazione di un algoritmo di hash e un'eventuale impronta precedentemente associata ad esso: in questo modo è possibile ad esempio gestire il passaggio da un algoritmo di hash diventato non più sicuro ad uno più robusto;
- infine, informazioni relative al processo di produzione del PdA, come: l'indicazione del nome e del ruolo dei soggetti che intervengono nel processo di produzione del PdA (es. responsabile della conservazione, delegato, pubblico ufficiale ecc.), il riferimento temporale adottato (generico riferimento temporale o marca temporale), l'indicazione delle norme tecniche e giuridiche applicate per l'implementazione del processo di produzione del PdA ed, infine, anche per il processo, un elemento "ExtraInfo" che consente di aggiungere dati soggettivi relativi al processo.

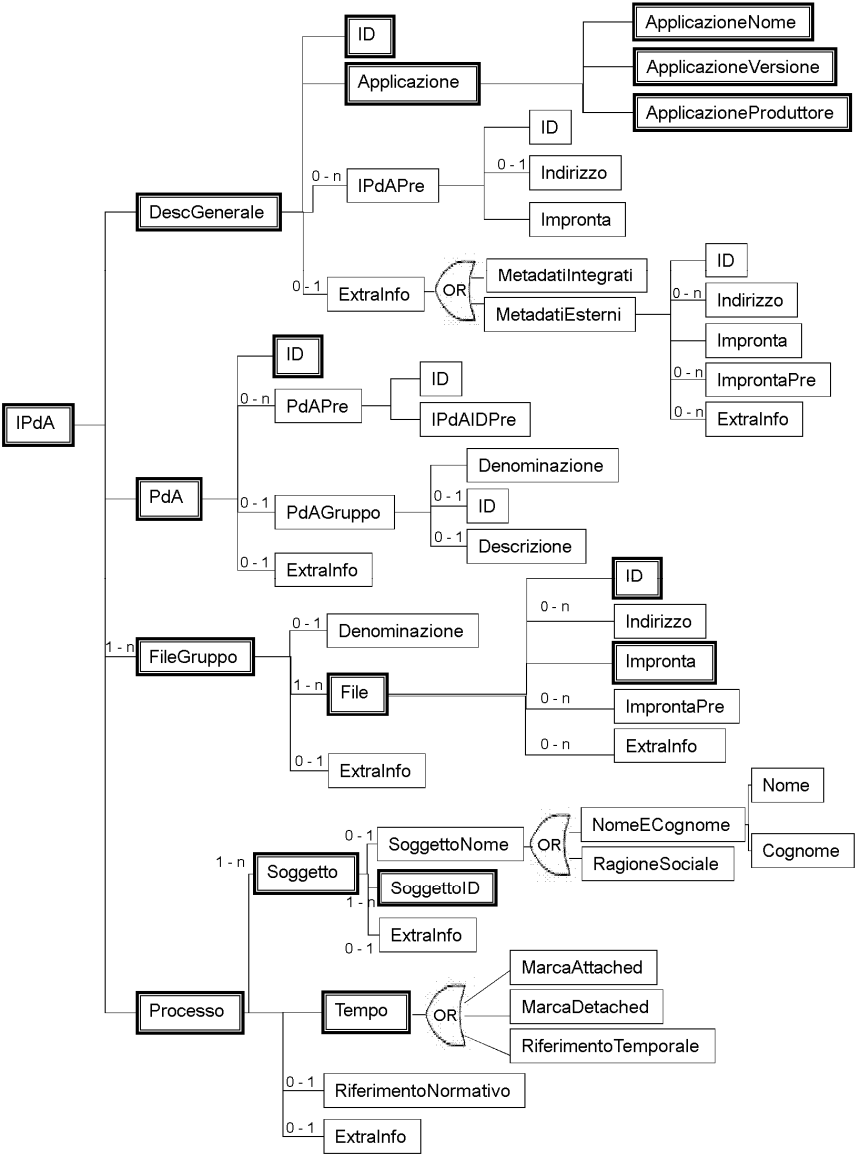
La flessibilità della struttura consente di gestire situazioni in cui è necessario ordinare in modo diverso gli indici creandone di nuovi, accorpando o frammentando le informazioni contenute negli IPdA precedenti, oppure generare uno nuovo IPdA facendo riferimento ad una precedente versione dello stesso: questo è il caso in cui si desidera effettuare migrazioni a causa di evoluzioni tecnologiche.

Infine, come accennato precedentemente nella specificazione delle varie strutture dell'indice del pacchetto di archiviazione, l'elemento "ExtraInfo" presente può essere oggetto di ulteriori specificazioni e deve essere inteso come una sorta di "plug-in" per strutture di metadati specialistiche, ove la specializzazione può essere relativa al dominio applicativo (sanità, banche, etc.) o alla tipologia documentaria (fatture, circolari, rapporti diagnostici, etc.).

Nei capitoli successivi sono riportati la rappresentazione grafica della struttura dell'indice del pacchetto di archiviazione e il relativo vocabolario.

STRUTTURA DELL'INDICE DEL PACCHETTO DI ARCHIVIAZIONE

Si riporta la rappresentazione grafica della struttura dell'indice del pacchetto di archiviazione.



Gli elementi racchiusi nella cornice in grassetto sono obbligatori.

2 Vocabolario

Nel vocabolario relativo alla struttura dell'indice del pacchetto di archiviazione sono riportati per ogni termine il nome, la descrizione, l'elemento da cui discende e l'elenco degli eventuali elementi di cui può essere l'origine.

Nome Elemento	Descrizione	Elemento Padre	Elementi Figli
Applicazione	Informazioni sull'applicazione che ha generato l'IPdA.	DescGenerale	ApplicazioneNome, ApplicazioneProdotto re, ApplicazioneVersione
ApplicazioneNome	Nome dell'applicazione che ha generato l'IPdA.	Applicazione	
ApplicazioneProdotto re	Nome del produttore dell'applicazione che ha generato l'IPdA.	Applicazione	
ApplicazioneVersione	Versione dell'applicazione che ha generato l'IPdA.	Applicazione	
Cognome	Cognome del soggetto che interviene nel processo di produzione del pacchetto di archiviazione.	NomeECognome	
Denominazione	Nome dell'eventuale tipologia o aggregazione a cui appartiene il File o il PdA.	FileGruppo, PdAGruppo	
Descrizione	Informazioni descrittive relative a una eventuale tipologia/aggregazione (di natura logica o fisica) cui il PdA appartiene.	PdAGruppo	
DescGenerale	Informazioni relative all'Indice del Pacchetto di Archiviazione, associate al pacchetto stesso.	IPdA	Applicazione, ExtraInfo, ID, IPdAPre
ExtraInfo	Ulteriori informazioni dell'elemento cui si riferisce, che non possono essere associate ad altri elementi, ad esempio per la definizione di strutture di metadati adeguate allo specifico contesto d'uso. Queste ulteriori informazioni	File, FileGruppo, DescGenerale, MetadatiEsterni , PdA, Processo, Soggetto	MetadatiEsterni, MetadatiIntegrati

	devono essere strutturate nel formato XML, utilizzando uno schema XML. L'insieme di queste informazioni può essere inserito direttamente all'interno o all'esterno dell'elemento come file avendo quindi la stessa struttura dell'elemento <File>.		
File	Informazioni relative al file contenuto nel pacchetto di archiviazione.	FileGruppo	ExtraInfo, ID, Impronta, ImprontaPre, Indirizzo
FileGruppo	Elemento di aggregazione di più file contenuti nel pacchetto di archiviazione. È funzionale alla creazione di insiemi di file sulla base di criteri logici o tipologici.	IPdA	Denominazione, ExtraInfo, File
ID	Identificativo univoco dell'elemento cui si riferisce.	File, DescGenerale, IPdAPre, MetadatiEsterni, PdA, PdAGruppo, PdAPre	
Impronta	Informazioni sull'impronta del file cui l'elemento si riferisce.	File, IPdAPre, MetadatiEsterni	
ImprontaPre	Informazioni relative a precedenti impronte del file contenuto nel pacchetto di archiviazione o del file di metadati (esterno all'IPdA) che contiene le informazioni dell'elemento <ExtraInfo>.	File, MetadatiEsterni	
Indirizzo	Informazioni relative all'indirizzo fisico del file dell'elemento cui si riferisce, espressa come indirizzo URI.	File, IPdAPre, MetadatiEsterni	
IPdA	Indice che contiene le informazioni relative al pacchetto di archiviazione prodotto.		FileGruppo, DescGenerale, PdA, Processo
IPdAIDPre	Identificativo univoco dell'indice del pacchetto di archiviazione associato al precedente pacchetto di archiviazione oggetto della descrizione. Il valore dell'identificativo deve	PdAPre	

	coincidere con il valore dell'elemento <ID> contenuto all'interno dell'elemento <IPdAPre>.		
IPdAPre	<p>Informazioni relative a uno o più indici dei pacchetti di archiviazione da cui è originato quello in oggetto. Tali informazioni sono fondamentali per ricostruire la storia degli oggetti posti in conservazione. L'IPdAPre può riferirsi a:</p> <ul style="list-style-type: none"> ▪ una precedente versione dell'IPdA attuale (ad esempio in caso di migrazione e/o modifiche del formato dei file, ove da un PdA si migri ad un nuovo PdA); ▪ più IPdA cronologicamente antecedenti che hanno generato per fusione l'IPdA attuale (ad esempio in caso di riorganizzazione della struttura dell'archivio, ove più PdA vengano aggregati in un singolo PdA); ▪ un IPdA cronologicamente antecedente che per frammentazione ha generato l'IdP attuale (ad esempio in caso di scarto di documenti da un PdA, ove a partire da un PdA si generino più PdA). 	DescGenerale	ID, Indirizzo, Impronta
MarcaAttached	Data e ora di produzione dell'indice del pacchetto di archiviazione, in forma normalizzata, nel caso in cui questa sia testimoniata con una marca temporale attached all'IPdA stesso. Al contrario dell'analogo elemento <MarcaDetached>, in questo caso non ha senso indicare l'URI della marca temporale.	Tempo	
MarcaDetached	Informazioni sulla localizzazione della marca temporale detached relativa a data e ora di produzione dell'indice del pacchetto di archiviazione. Il valore dell'elemento deve essere espresso nel formato URI.	Tempo	

MetadatiEsterni	Le informazioni dell'elemento <ExtraInfo>, contenute all'esterno dell'IPdA in un file XML le cui caratteristiche sono descritte nei subelementi. Trattandosi di un file, questo elemento ha la stessa struttura dell'elemento <File>. Tale file pur essendo esterno all'IPdA è comunque contenuto nel PdA.	ExtraInfo	ExtraInfo, ID, Impronta, ImprontaPre, Indirizzo
MetadatiIntegrati	Le informazioni dell'elemento <ExtraInfo>, integrate all'interno dell'IPdA e strutturate nel formato XML.	ExtraInfo	
Nome	Nome del soggetto che interviene nel processo di produzione del pacchetto di archiviazione.	NomeECognome	
NomeECognome	Nome e cognome del soggetto che interviene nel processo di produzione del pacchetto di archiviazione. Tale elemento deve essere valorizzato nel caso in cui il soggetto sia persona fisica.	SoggettoNome	Nome, Cognome
PdA	Informazioni relative al pacchetto di archiviazione.	IPdA	ExtraInfo, ID, PdAGruppo, PdAPre
PdAGruppo	Informazioni relative a una eventuale tipologia o aggregazione (di natura logica o fisica) cui il PdA appartiene.	PdA	Denominazione, Descrizione, ID
PdAPre	Informazioni relative a uno o più pacchetti di archiviazione da cui è originato quello in oggetto (ad esempio per migrazione di un pacchetto o per aggregazione di più pacchetti).	PdA	ID, IPdAIDPre
Processo	Informazioni relative alle modalità di svolgimento del processo di produzione del pacchetto di archiviazione.	IPdA	ExtraInfo, RiferimentoNormativo, Soggetto, Tempo
RagioneSociale	Ragione sociale del soggetto che interviene nel processo di produzione del pacchetto di archiviazione. Tale elemento deve essere valorizzato nel caso	SoggettoNome	

	in cui il soggetto sia persona giuridica.		
RiferimentoNormativo	Informazioni su norme, regolamenti e standard che regolano il processo di produzione del pacchetto di archiviazione.	Processo	
RiferimentoTemporale	Informazioni relative a data e ora di produzione dell'indice del pacchetto di archiviazione, nel caso in cui non venga apposta 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).	Tempo	
Soggetto	Informazioni relative ai soggetti che intervengono nel processo di produzione del pacchetto di archiviazione.	Processo	ExtraInfo, SoggettoID, SoggettoNome
SoggettoID	Identificativo univoco del soggetto che interviene nel processo di produzione del pacchetto di archiviazione. Se l'identificativo è un codice con ambito nazionale, a tale codice deve essere premesso il codice Paese definito da ISO 3166 seguito dal carattere “.”. Se il soggetto è colui che appone la firma digitale all'IPdA è da privilegiare l'uso di un codice identificativo presente in un campo del suo certificato digitale.	Soggetto	
SoggettoNome	Nome o denominazione sociale del soggetto che interviene nel processo di produzione del pacchetto di archiviazione.	Soggetto	NomeECognome, RagioneSociale
Tempo	Informazioni relative a data e ora di produzione dell'indice del pacchetto di archiviazione. Tale elemento è necessario a distinguere i seguenti casi: <ul style="list-style-type: none"> ▪ riferimento temporale (l'elemento <RiferimentoTemporale>) 	Processo	MarcaAttached, MarcaDetached, RiferimentoTemporal e

	<ul style="list-style-type: none">▪ marca temporale detached (il cui indirizzo URI valorizza l'elemento <MarcaDetached>)▪ marca temporale attached (all'elemento vuoto <MarcaAttached> è associata la data in forma normale).		
--	--	--	--

Nella tabella seguente vengono riportati per ogni attributo il nome, la descrizione, gli elementi a cui può essere associato e le caratteristiche.

Nome Attributo	Descrizione	Elementi	Caratteristiche
altroruolo	Valorizzazione del ruolo rivestito dal soggetto nell'ambito del processo di produzione del pacchetto di archiviazione, nel caso in cui risultino non adeguati i valori previsti dall'attributo Ruolo.	Soggetto	attributo opzionale di tipo CDATA (Character data)
altroschemarif	Valorizzazione del sistema di riferimento utilizzato per identificare il soggetto nel caso in cui risultino non adeguati i valori previsti dall'attributo schemarif.	SoggettoID	attributo opzionale di tipo CDATA (Character data)
codifica	Valorizzazione del tipo di codifica utilizzato nella scrittura del file.	File, MetadatiEsterni	attributo obbligatorio, Valori ammessi: 7bit 8 bit base64 binario quotedprintable xtoken
estensione	Estensione che caratterizza il nome del file.	File, MetadatiEsterni	attributo opzionale di tipo NMTOKEN (ovvero esprimibile con caratteri alfanumerici, punti, trattino, due punti o underscore)
formato	Informazioni sulla struttura dati del file a cui si riferisce.	File, MarcaDetached, MetadatiEsterni	attributo obbligatorio di tipo NMTOKEN (ovvero esprimibile con caratteri alfanumerici, punti, trattino, due punti o underscore)

funzione	Specificazione della funzione di hash utilizzata.	Impronta, ImprontaPre	attributo obbligatorio di tipo NMTOKEN (ovvero esprimibile con caratteri alfanumerici, punti, trattino e underscore). Valore di default: "SHA-256"
IPdAcorrelato	Identificatore univoco dell'indice del pacchetto di archiviazione contenente la precedente impronta del file contenuto nel pacchetto di archiviazione o del file di metadati (esterno all'IPdA) che contiene le informazioni dell'elemento <ExtraInfo>.	ImprontaPre	attributo obbligatorio di tipo NMTOKEN (ovvero esprimibile con caratteri alfanumerici, punti, trattino, due punti o underscore)
lingua	Lingua in cui sono espresse le informazioni.	Descrizione, RiferimentoNormativo	attributo opzionale di tipo NMTOKEN (ovvero esprimibile con caratteri alfanumerici, punti, trattino, due punti o underscore). Deve essere espresso con un codice a due caratteri, coerentemente con lo standard ISO 639-1:2002. Valore di default: "it"
normal	Indicazione della data e dell'ora di produzione dell'indice del pacchetto di archiviazione, espressa in forma normalizzata. 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).	MarcaAttached, MarcaDetached	attributo obbligatorio di tipo CDATA (Character data)
ruolo	Valorizzazione del ruolo rivestito dal soggetto nell'ambito del processo di produzione del pacchetto di archiviazione.	Soggetto	attributo obbligatorio. Valori ammessi: Delegato, Responsabile della conservazione, Pubblico ufficiale, Altro ruolo
schema	Eventuali informazioni relative al sistema di riferimento nel quale assume significato il	ID, IPdA_IDPre	attributo opzionale di tipo CDATA (Character data).

	valore dell'identificativo univoco.		Valore di default: "local"
schemarif	Valorizzazione del sistema di riferimento utilizzato per identificare il soggetto.	SoggettoID	attributo obbligatorio. Valori ammessi: codice fiscale, partita IVA, codice del Servizio Sanitario Nazionale, altroschemarif
schemaxml	Indirizzo URL dove è presente lo schema XML dei metadati utilizzato per descrivere le ExtraInfo.	ExtraInfo	attributo obbligatorio di tipo NMTOKEN (ovvero esprimibile con caratteri alfanumerici, punti, trattino, due punti o underscore); Deve assumere la forma di URL.
tipo	Indicazione della natura del soggetto.	Soggetto	attributo obbligatorio. Valori ammessi: denominazione, ragione sociale
url	Indirizzo URL dove è presente lo schema XML dell'indice del pacchetto di archiviazione.	IPdA	attributo obbligatorio di tipo NMTOKEN (ovvero esprimibile con caratteri alfanumerici, punti, trattino, due punti o underscore); Deve assumere la forma di URL. Valore di default: "www.uni.com/U3011/sincro"
versione	Indicazione della versione dello schema XML dell'indice del pacchetto di archiviazione al fine di gestire l'evoluzione dello standard.	IPdA	attributo obbligatorio di tipo NMTOKEN (ovvero esprimibile con caratteri alfanumerici, punti, trattino, due punti o underscore). Valore di default fisso: "1.0"
xml canonico	Indicazione se l'eventuale file in formato xml è trasformato in forma canonica.	Impronta, ImprontaPre	attributo opzionale. Valori ammessi: SI NO