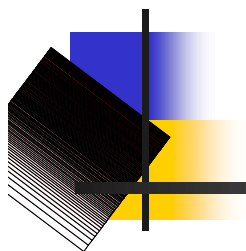


Metadati Per La Gestione Di Risorse Non Testuali

Di Nicola Benvenuti

Biblioteca Di Architettura – Università Di Firenze
(Benvenuti_n@unifi.it)



L'integrazione delle risorse eterogenee distribuite: problemi, strumenti e prospettive

Seminario a cura del Coordinamento nazionale biblioteche di architettura (CNBA)

27 ottobre 2004, Palazzo dei Congressi, Piazzale Kennedy 1, Roma

ore 14,30-17,30 - Sala Quirinale



Cosa definiscono gli schemi di metadati

Opera

- FRBR: creazione artistica, intellettuale; entità astratta, priva di oggetto materiale.
- CDWA: Opera è una entità fisica che esiste o è esistita o potrà esistere nel futuro. Può essere una creazione artistica (quadro o scultura), una performance, una composizione o un'opera letteraria; un edificio o un oggetto di cultura materiale. Può essere un singolo pezzo o consistere di molte parti

Immagine

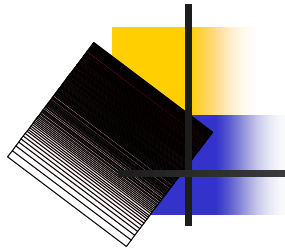
- Questo termine è ambiguo, serve per lasciare il campo delle opere "uniche" ed entrare in quello della "manifestazione" dell'opera caratterizzata dalla ripetibilità, almeno in linea di principio, dell'oggetto da descrivere, con un approccio più specificamente bibliotecario
 - ISBD(NBM)
 - VRA Core
 - Dublin Core

Collezione/archivio

- Nella pratica delle biblioteche l'immagine fa quasi sempre parte di un insieme più ampio (fondo d'archivio, raccolta in occasione di eventi, materiali di supporto alla didattica, etc.) che deve essere parimenti descritto
 - ISAD(G)
 - SEPIADES

Struttura del sistema informativo

- Nella progettazione di archivi digitali entrano in campo metadati di gestione e conservazione degli oggetti informativi
 - OAIS, METS, MAG



Parte I. Opera

- ✍ Metadati per gli oggetti d'arte:
 - ✍ [CDWA](#) Categories for Description of Work of Art
 - ✍ [CCO](#) (cataloguing Cultural Object)
 - ✍ [CIMI](#) Access Points
 - ✍ [FDA/ADAG](#) Guide to the Description of Architectural Drawings data categories
 - ✍ Art Museum Image Consortium ([AMICO](#)) data dictionary
 - ✍ [CIDOC](#)'s International Guidelines for Museum Object Information
 - ✍ [MDA's Spectrum](#)
 - ✍ [Object ID](#). pensato specificamente per aiutare le indagini di polizia sui furti d'arte

Schede per la catalogazione delle opere - ICCD

<http://www.iccd.beniculturali.it/standard/index.html>

Beni Mobili

- [Scheda RA](#): reperto archeologico
- [Scheda OA-D](#): opere e oggetti d'arte
- [Scheda OAC](#): opere d'arte contemporanea
- [Scheda S-MI](#): stampe e matrici d'incisione
- [Scheda F](#): fotografia
- [Scheda STS](#): beni storico scientifici
- [Scheda BDM](#) (FKO): beni demoetnoantropologici materiali
- [Scheda BDI](#): beni demoantropologici immateriali
- [Scheda TMA](#): tabella di materiale archeologico
- [Scheda SMO](#): strumenti musicali - organi
- [Scheda NU](#): numismatica

Beni Immobili

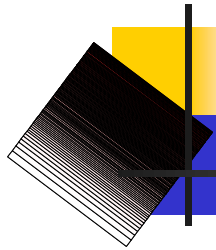
- [Scheda A](#): architettura
- [Scheda PG](#): parchi e giardini
- [Schede SU/TP](#): settore urbano/settore extraurbano
- [Scheda MA-CA](#): monumenti e complessi archeologici
- [Scheda SAS](#): saggio stratigrafico

Beni Urbanistico Territoriali

- [Scheda SI](#): sito archeologico
- [Scheda T](#): territorio comunale
- [Scheda CS](#): centro storico

Archivi

- [Scheda AUT/BIB](#): autore/bibliografia
- [Scheda DSC](#): scavo
- [Scheda RGC](#): ricognizione



La Scheda F

<http://www.iccd.beniculturali.it/download/schedaf.pdf>

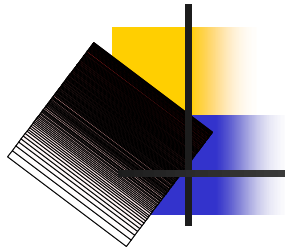
- ✍ 21 paragrafi o sezioni suddivisi in 79 campi
 - ✍ 23 sono campi semplici
 - ✍ 56 sono strutturati, cioè contengono sottocampi
 - ✍ In totale si tratta di 269 voci
- ✍ Identificazione del contesto culturale e delle fonti di riferimento.
Es.:
 - ✍ Sez.: Fonti e documenti di riferimento DO
 - ✍ Sez.: Riferimento ad altre schede SK
 - ✍ Campo: CRF-COPIE/REIMPIEGHI/ETC. permette di stabilire ulteriori rapporti e connessioni tra opere *qualitativamente* diverse, anche se riferibili e riconducibili ad uno stesso prototipo
- ✍ Non ha un "core set"



CDWA 2.0

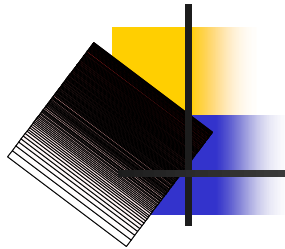
http://www.getty.edu/research/conducting_research/standards/cdwa/

- ✍ Categories for the Description of Works of Art (CDWA) del Getty Institute
 - ✍ Base per la descrizione di opere d'arte o oggetti museali
 - ✍ Formulate avendo in mente le esigenze di ricercatori ed accademici e con il riconoscimento della incertezza e soggettività dell'opera d'arte.
 - ✍ elementi *remarks* e *citation* per specificare il punto di vista assunto nella scelta delle informazioni inserite per l'identificazione e descrizione dell'oggetto
 - ✍ Distinzione tra
 - ✍ categorie intrinseche all'oggetto o all'immagine individuale
 - ✍ categorie estrinseche comuni a più opere, sono indicate una volta per tutte in *authority files*



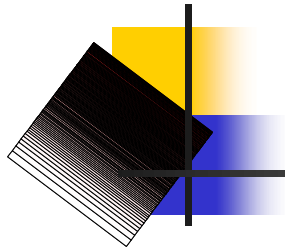
CDWA 2.0

- ☞ Categorie : 27 categorie principali suddivise in subcategorie per circa 300 elementi. Nelle categorie si possono distinguere informazioni intese per la visualizzazione e informazioni intese per il recupero. Le prime sono intese in un formato facilmente comprensibile per l'utente. Altre sono indicizzate e inseriti da catalogatori con vocabolari controllati.
- ☞ 4 authorities: *creator (ULAN)*, *place (TGN)*, *generic concept*, *subject (AAT)*
- ☞ 24 elementi sono indicati come "core" (essenziali)
- ☞ Per la sua complessità CDWA è raramente implementato per intero, ma la sua analisi è stata utilizzata per banche dati museali e per sviluppare metadati per numerose applicazioni
- ☞ Sul sito si trova un [crosswalk](#) tra CDWA, DC, VRA 3.0, USMARC, CCO, Object ID, FDA Guide, CIMI, EAD



Parte II. Immagine

- ✍ *Immagine*: una rappresentazione visiva di un'opera. E' il livello più simile alla "manifestazione" di FRBR , e se quindi l'opera d'arte è il campo del curatore di collezioni d'arte, l'immagine, intesa come "oggetto utilizzabile ai fini di consultazione, ricerca, informazione" è il campo del bibliotecario.
- ✍ Tipicamente una collezione di risorse visive è la riproduzione di un'opera posseduta dall'ente e si tratta generalmente di *slides*, fotografie o file digitali. Possono esserci più immagini della stessa opera e su diversi supporti
- ✍ L'approccio all'immagine è quindi diverso da quello per l'opera. "I bibliotecari tendono a catalogare disegni e fotografie di strutture; considerano il nome dell'edificio come soggetto, nome o nome dell'organizzazione: catalogano la rappresentazione invece dell'edificio" (Priscilla Caplan, *Metadata fundamentals for all librarians*. Chicago, ALA, 2003)
- ✍ Sono perciò necessari comportamenti diversi nella loro descrizione rispetto alle opere.



ISBD(G) e MARC

✍ Vantaggi:

- ✍ possibilità di avere tutto il materiale posseduto (libri, periodici, fotografie, risorse elettroniche ...) in un unico sistema automatico, con un unico modello di descrizione.

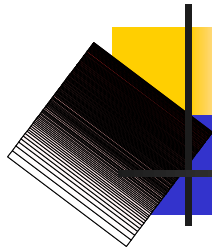
✍ Svantaggi:

- ✍ Pensato per descrivere unità individuali e difficoltà a dar conto di relazioni rilevanti tra set di fotografie (negativi o stampe)
- ✍ Difficoltà a stabilire i campi dove inserire alcuni dati
- ✍ Molte informazioni cruciali relative sia alle opere che alle rappresentazioni, devono essere inserite nell'area delle note
- ✍ Lo ICCD indica le corrispondenze tra la scheda F e UNIMARC



Visual Resource Association – VRA Core

- ✍ Inizialmente: 21 categorie in tre gruppi che descrivevano:
 - ✍ l'opera originale (l'oggetto),
 - ✍ il creatore
 - ✍ la rappresentazione (*visual document*);
- ✍ La versione 3.0 invece preferisce che per opera e rappresentazione/i (ora *image*) siano creati record diversi secondo la regola dell'1:1 (one2one) cioè, per ogni risorsa deve essere utilizzato un set di metadati
- ✍ È stato perciò introdotto l'elemento *type* che può contenere i valori: *work* o *image*.
- ✍ Deliberatamente modellato sul DC, è un "core set ", eventualmente integrato su base locale utilizzando *qualifier* per precisare i valori di alcuni elementi secondo la regola del *dumb down*;
- ✍ Come per il DC ogni elemento è opzionale e ripetibile



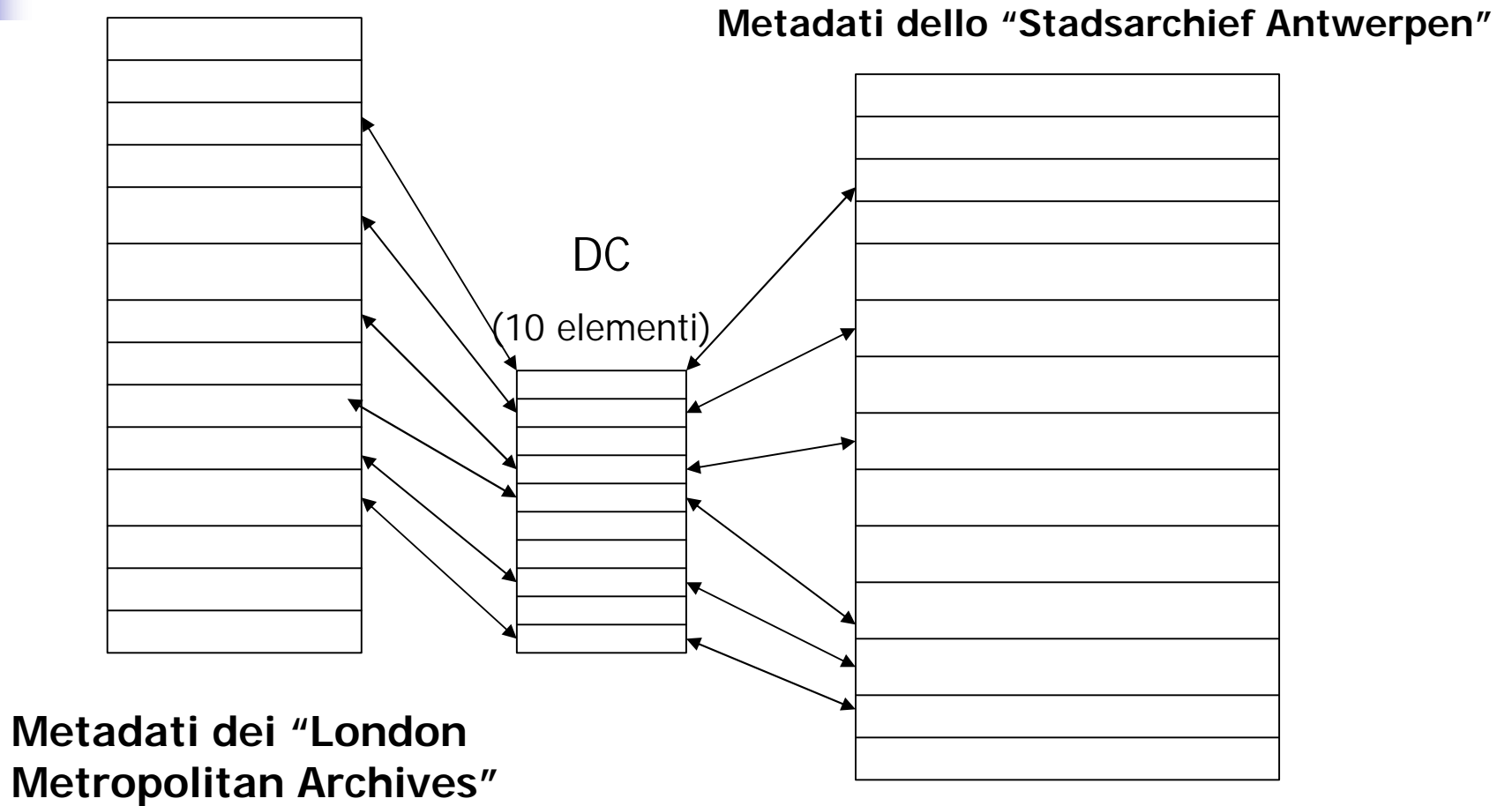
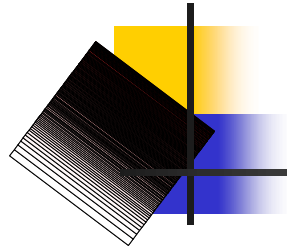
VRA Core (2)

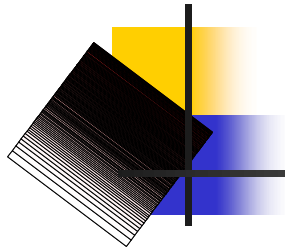
- ✍ Per quanto riguarda il *contesto* dell'unità documentaria, si distingue tra un legame essenziale e uno informativo:
 - ✍ *Essenziale* quando l'opera descritta include l'opera correlata o fisicamente o logicamente in un contesto più o meno ampio (si raccomanda in tali casi di relazione l'uso del *qualifier* "Title.LargerEntity")
 - ✍ *Informativo* quando la relazione tra due opere esiste ma le due opere correlate possono sussistere indipendentemente l'una dall'altra. Il link dovrebbe essere reciproco.



- ✍ Riconosciuto come schema chiave per lo scambio di informazioni tra sistemi informativi
- ✍ Utile come strumento di *resource discovery*
- ✍ Schema semplificato per la catalogazione delle risorse
- ✍ Facile da comprendere e da applicare
- ✍ Espresso in formato "medium independent": xml

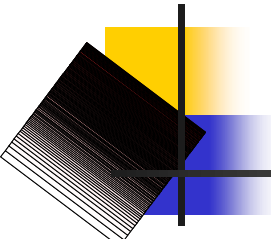
Ruolo del DC come strumento di interoperabilità (es. EVA project)





DC come schema semplificato

- ✍ Interpretazione del CIMI: Guide to best practice: Dublin Core, (1999):
- ✍ Interpretazione del progetto EVA
- ✍ Interpretazione di SEPIA
- ✍ Interpretazione del Getty Institute (ampio ricorso ai qualifier)
- ✍ ...



Catalogazione del contesto e la regola dello 1:1 (one2one)

- ✍ Per ogni risorsa un set di metadati, quindi ogni oggetto fisico è accompagnato dai suoi metadati. Es.:
 - ✍ 1 per la fotografia
 - ✍ 1 per il surrogato
 - ✍ 1 per l'immagine digitale
 - ✍ 1 per i derivati
- ✍ Permette di dar conto del contesto archivistico con l'uso dell'elemento DC.relation
- ✍ Ridondante ma compatibile con RDF



Parte III. Descrizione a livelli

ISAD(G)

- ✍ 26 elementi in 7 aree di informazioni descrittive
 1. Area dell'identificazione
 2. Area delle informazioni sul contesto
 3. Area delle informazioni relative al contenuto e alla struttura
 4. Area delle informazioni relative alle condizioni di accesso e utilizzazione
 5. Area delle informazioni relative a documentazione collegata
 6. Area delle note
 7. Area di controllo della descrizione
- ✍ 6 elementi sono essenziali per lo scambio di informazioni descrittive:
 - ✍ Segnatura e/o codice identificativo; titolo; soggetto produttore; data(e); consistenza dell'unità di descrizione; livello di descrizione

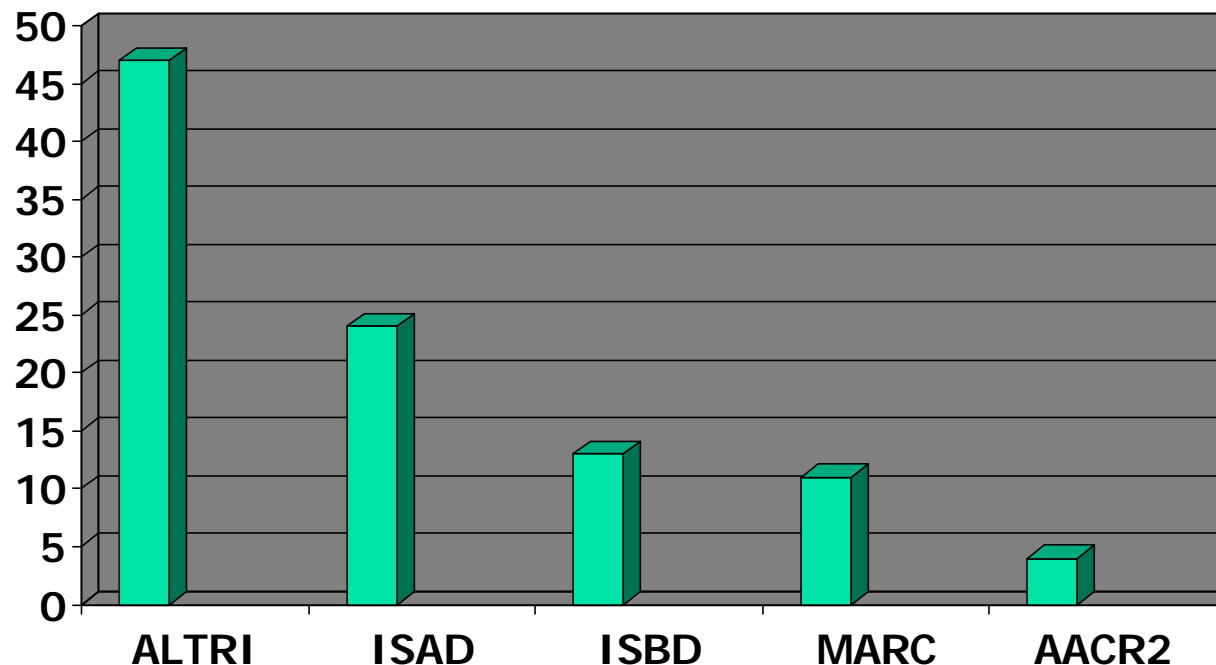


Isad(g)

- ✗ ISAD(G) prevede la descrizione di un fondo nel suo complesso, ma può essere usato anche per descrivere le singole parti di cui è composto. L'insieme di tutte le descrizioni così ottenute, collegate gerarchicamente, costituisce la rappresentazione del fondo.
- ✗ La descrizione a livelli è guidata dalle seguenti regole:
 - ✗ Dal generale al particolare
 - ✗ Informazioni pertinenti al livello di descrizione
 - ✗ Collegamento tra le descrizioni (rende esplicita la posizione dell'unità di descrizione all'interno della struttura gerarchica)
 - ✗ Non ripetizione delle informazioni a livello di parti componenti
- ✗ I livelli di descrizione utilizzano il seguente vocabolario: *fondo* (livello obbligatorio), *sub-fondo*, *serie*, *sub-serie*, *unità archivistica* (file), *unità documentaria* (item)

SEPIA (Safeguarding European Photographic Images for access)

Raccoglie esperienze dei più significativi progetti europei



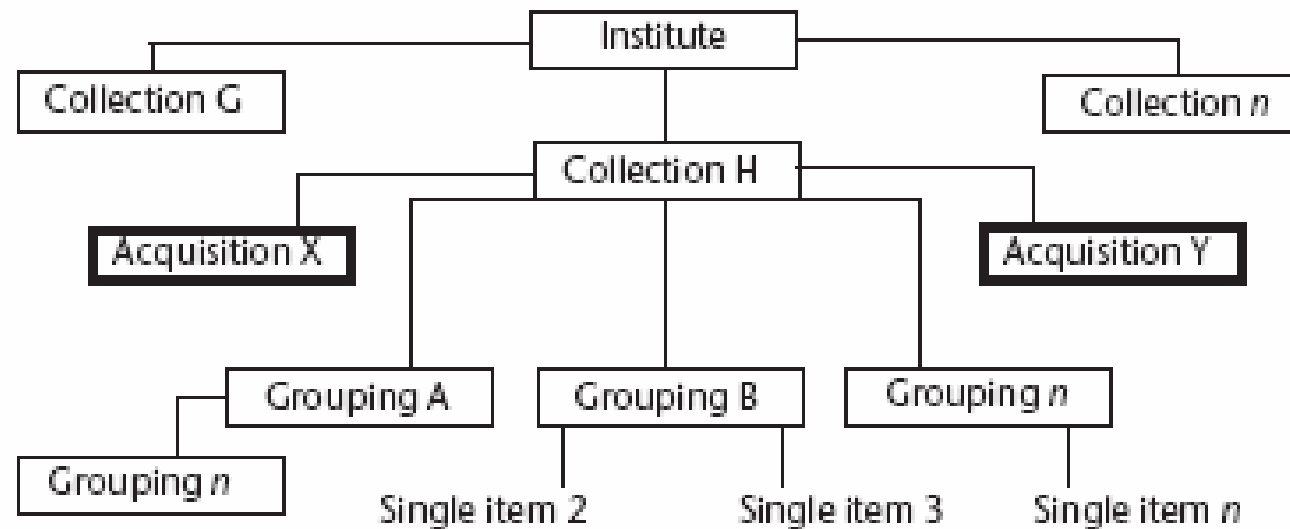
Con “Altri” si intendono schemi nazionali, personalizzati e comunque non standard.



- ✍ Più di 400 elementi per descrivere fotografie
- ✍ 21 elementi essenziali (core)
- ✍ Modello descrittivo multilivello
- ✍ Specifiche sull'uso del DC per descrivere collezioni fotografiche e corrispondenze tra il core set e DC

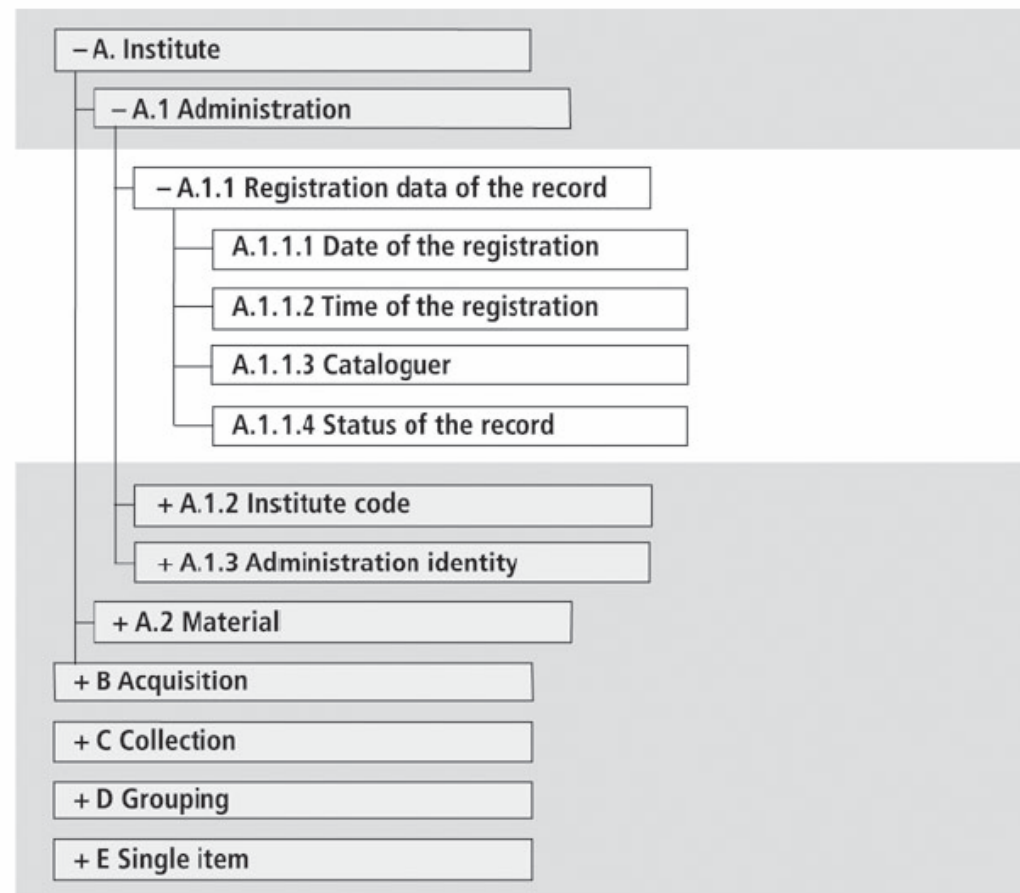
SEPIADES: struttura multilivello e registri di acquisizione

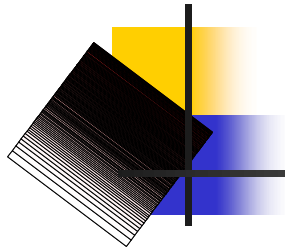
<http://www.knaw.nl/ecpa/publ/pdf/2710.pdf>



SEPIADES: struttura dello schema

A.1.1





nome, data, località

✍ Nome

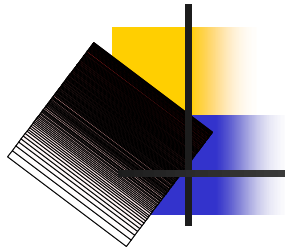
- ✍ Nome proprio
- ✍ Funzione (related, depicted)
- ✍ Tipo (person, dog, photographer ...)

✍ Data

- ✍ Status (exact, approximately)
- ✍ Function (date of exposure, date of publication, date of processing)
- ✍ Role (depicted, related)

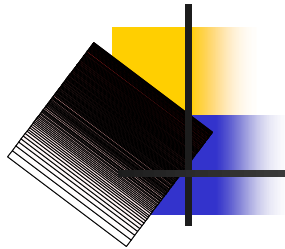
✍ Location

- ✍ Geographical location
- ✍ Specification (city, square, ...)
- ✍ Role (related, depicted)



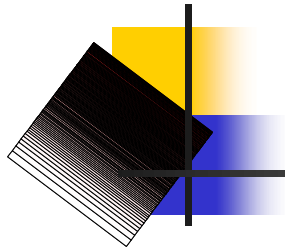
SEPIADES: core set e DC

<i>DC element</i>	<i>DC Definition</i>	<i>Qualifier</i>	<i>SEPIADES core element</i>
1. Title	A name given to the resource	Formal Invented	6. Title
2. Creator	An entity primarily responsible for making the content of the resource.		7. Creator
3. Subject	The topic of the content of the resource.		8. Descriptors/ subject headings/ classification
4. Description	An account of the content of the resource.		5. Description
5. Publisher	An entity responsible for making the resource available		2. Name of institute
6. Contributor	An entity responsible for making contributions to the content of the resource.		9. Names
7. Date	A date associated with an event in the life cycle of the resource.		Not applicable
8. Type	The nature or genre of the content of the resource.	Image	Not applicable



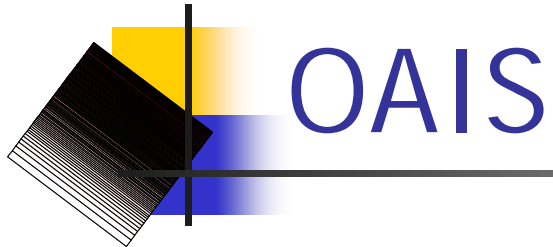
SEPIADES: core set e DC

9. Format	The physical or digital manifestation of the resource.		1. Main reference code 10. Date (of creation) 17. Photographic type 18. File format 16. Dimensions
10. Identifier	An unambiguous reference to the resource within a given context.		Not applicable, derives from the publication of the resource
11. Source	A Reference to a resource from which the present resource is derived.		6. Title (of collection) and/or 2. Name of institute
12. Language	A language of the intellectual content of the resource		Not applicable
13. Relation	A reference to a related resource.	Is part of/ Has parts Is based on/ is basis for Is referenced by/ references	13. Relationships 19. References
14. Coverage	The extent or scope of the content of the resource	spatial temporal	11. Geographical location 10. Date
15. Rights	Information about rights held in and over the resource.		12. Access restrictions/ copyright



Parte IV. Il sistema informativo

- ✍ Per gli oggetti digitali è molto più complicato assicurare accesso e autenticità rispetto alle controparti cartacee
- ✍ Sono perciò necessarie molte informazioni sulle caratteristiche tecniche e sulle procedure a cui sono sottoposti gli oggetti digitali
- ✍ Per identificare queste informazioni è stato ideato un modello di sistema informativo espresso nello standard ISO, *Open Archival Information System* (OAIS) - dove "open" sta per aperto al contributo di enti e istituzioni interessate



OAIS

- ✍ Identifica gli attori e le funzioni del sistema informativo (v. slide successiva)
- ✍ Stabilisce l'architettura dell'informazione elettronica cioè l'articolazione logica dei pacchetti informativi (*information packages*)
- ✍ Identifica diverse tipologie di pacchetti informativi per trasmettere le informazioni (SIP, AIP, DIP) e i loro componenti

Ambiente e Componenti Funzionali OAIS

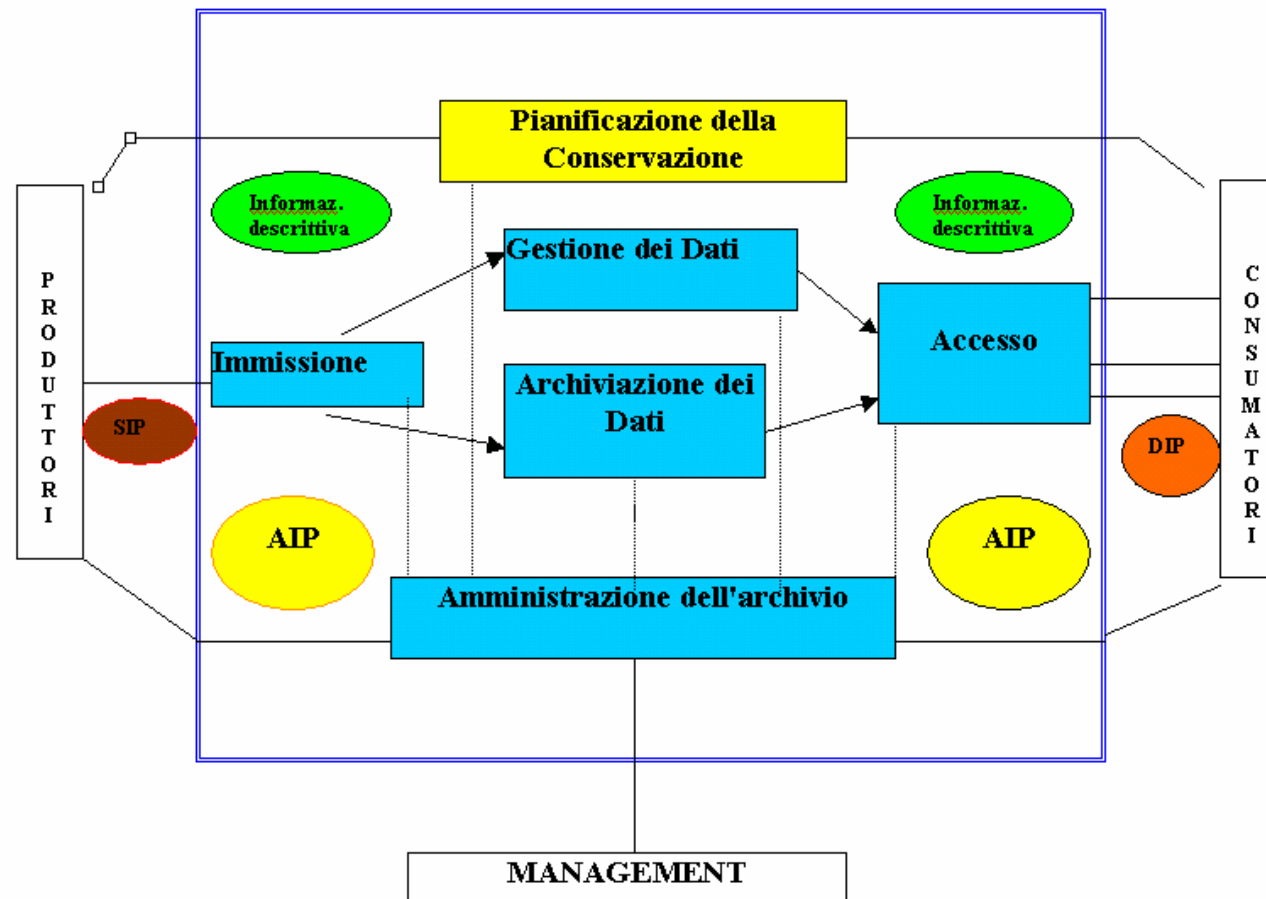
attori

funzioni

Sistema informativo

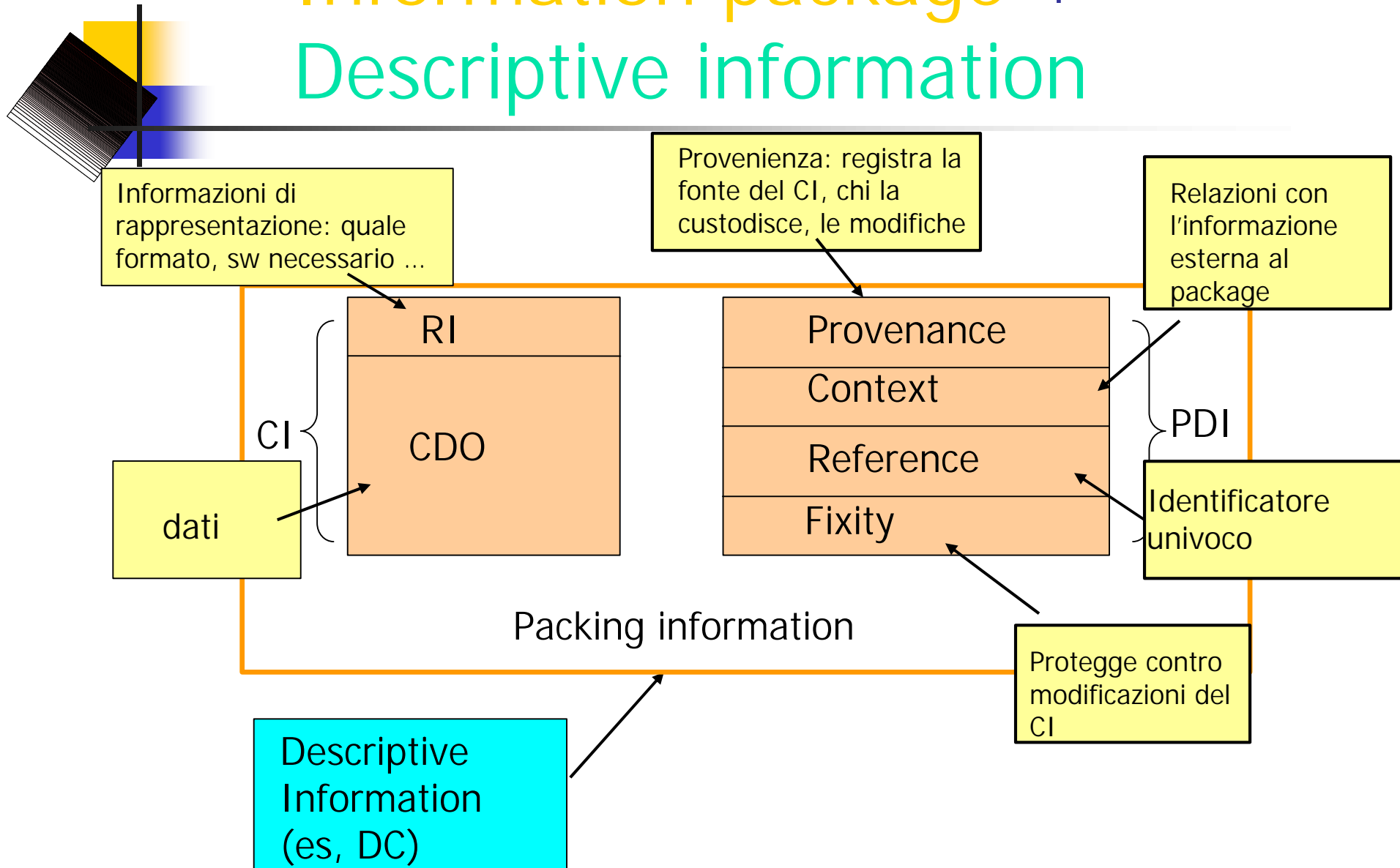


Flussi e relazioni



Da: Gloria Cirocchi, Il modello OAIS: <http://www.iccu.sbn.it/PDF/Cirocchi.pdf>

Information package + Descriptive information



Architettura dei metadati sdecondo la Library of Congress

<http://www.loc.gov/standards/metadata.html>

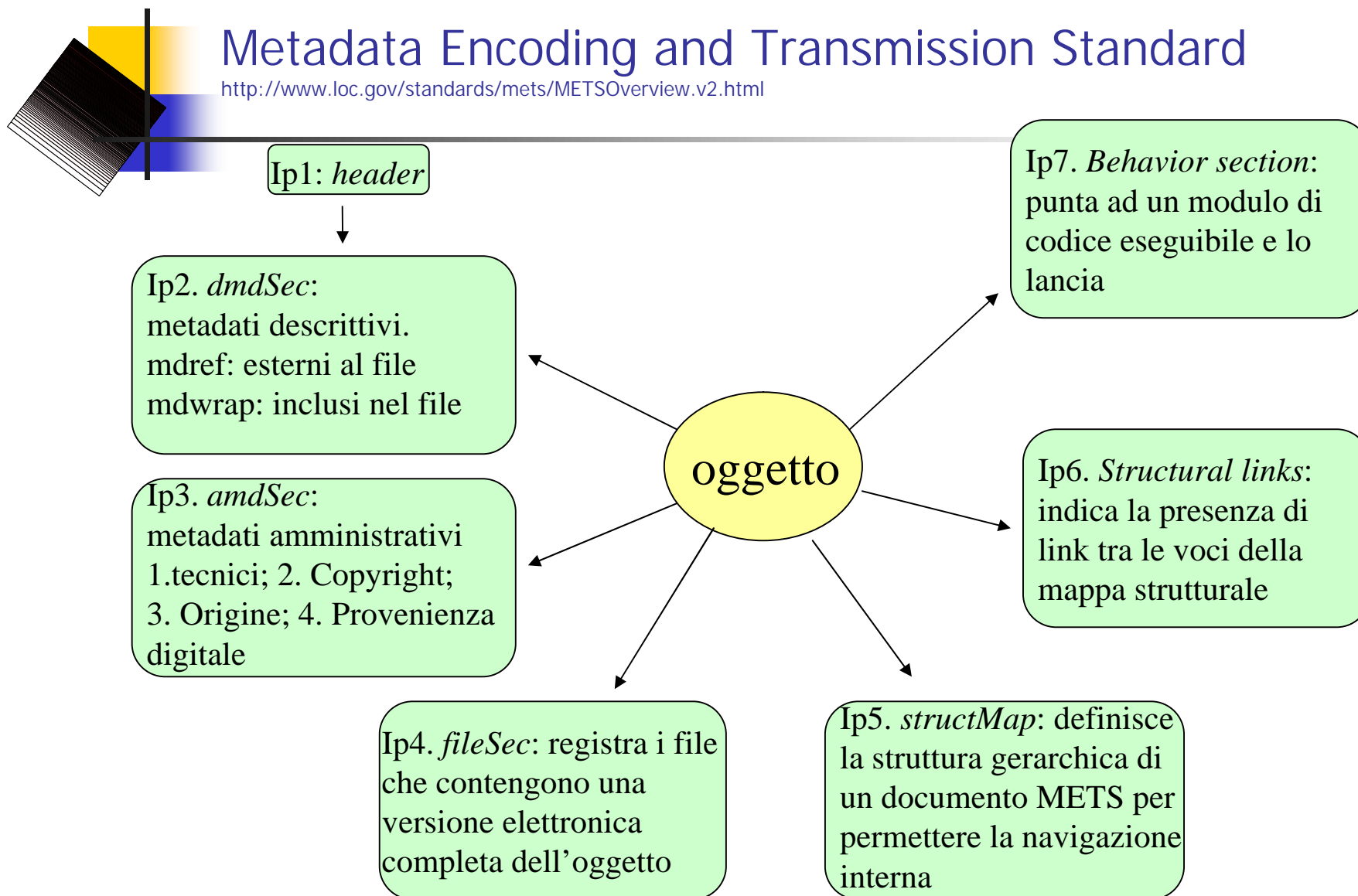
- ✍ Descrittivi (MODS)
- ✍ Amministrativi
- ✍ Strutturali

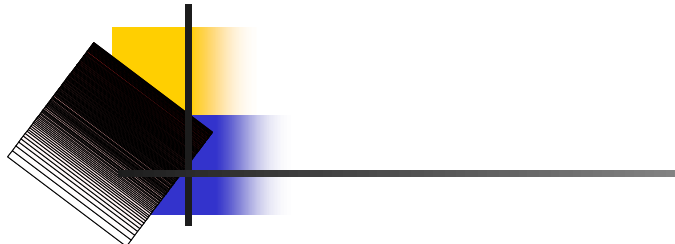
- ✍ Struttura dei metadati
 - ✍ Set (Collezione digitale)
 - ✍ Aggregate (aggregato)
 - ✍ Primary object (oggetto primario)
 - ✍ Intermediary object (oggetto intermedio)
 - ✍ Digital entity (Entità digitale)

METS

Metadata Encoding and Transmission Standard

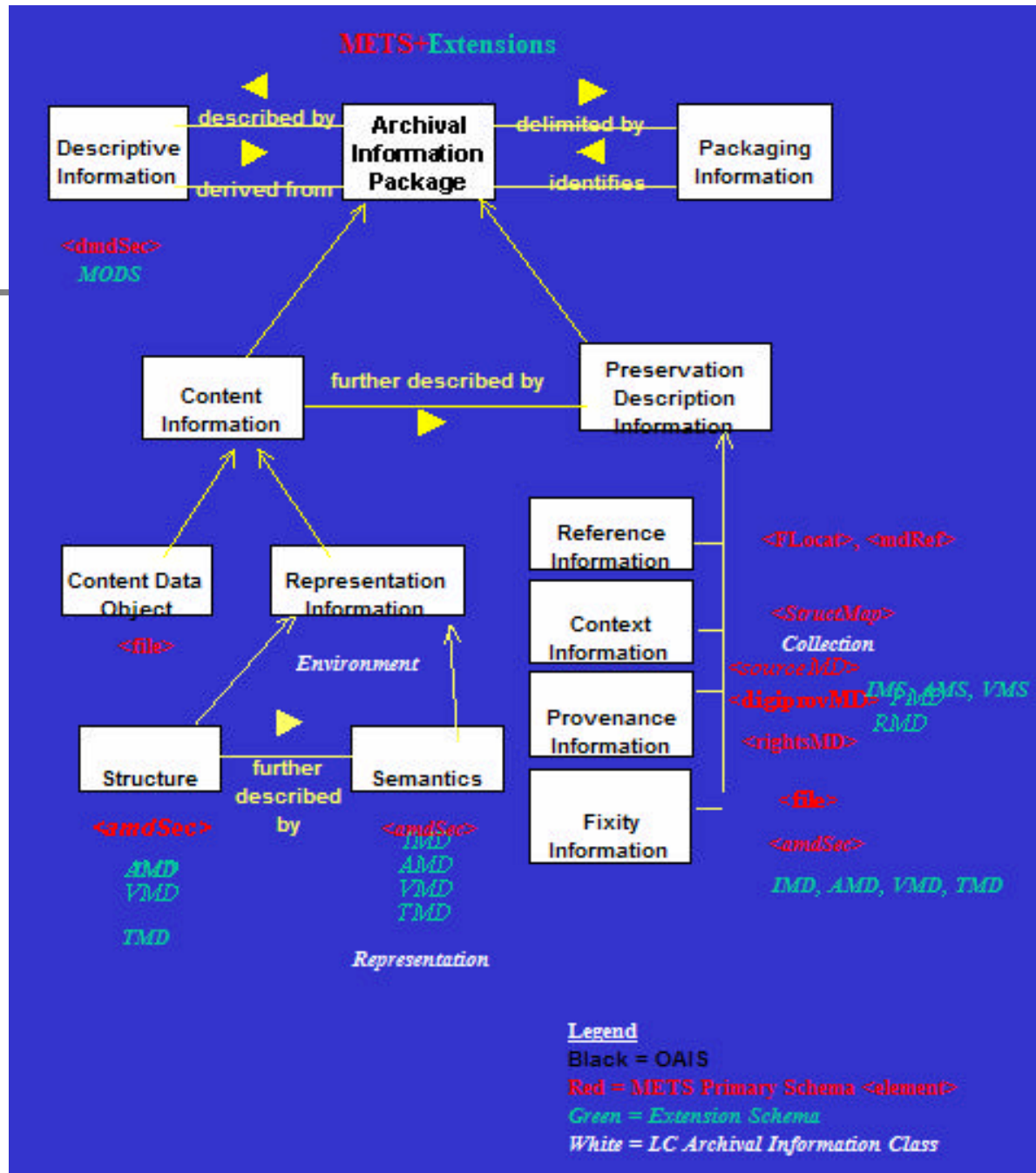
<http://www.loc.gov/standards/mets/METSOverview.v2.html>

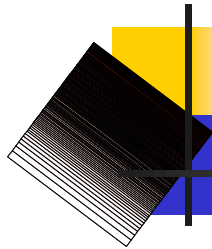




OAIS AIP to METS + Extensions Map (and the OCLC/RLG Metadata Framework)

Library of Congress Audio-Visual Prototyping Project (AV Project):






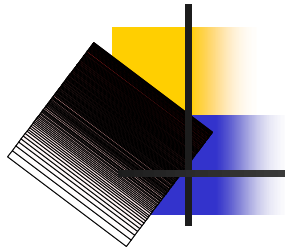


ARSBNI

<http://www.iccu.sbn.it/metaAG1.pdf>

Tre archivi

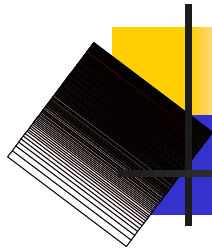
-  **L'archivio bibliografico.** Si ritiene infatti che al di là di esigenze particolari o di pura sperimentazione l'accesso dell'utente agli oggetti digitali sia sempre mediato dall'archivio bibliografico (visto dall'utente come OPAC).
-  **L'archivio dei *metadati*.** Le informazioni necessarie a gestire l'acquisizione, l'archiviazione, l'accesso e l'amministrazione degli oggetti e delle entità digitali (in altre parole i *metadati* oggetto della presente relazione)
-  **L'archivio dei dati digitali.** I file risultanti dal progetto di digitalizzazione.



ARSBNI



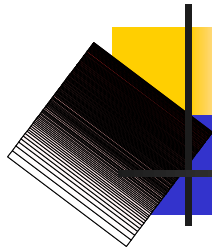
- ✍ Il SIP è composto da un insieme di file correlati.
Un file di tipo *xml*:
 - ✍ raccoglie tutte le informazioni sull'oggetto digitalizzato
 - ✍ contiene la mappa di tutti i file relativi all'oggetto digitalizzato
 - ✍ contiene anche informazioni bibliografiche e gestionali derivate dall'archivio bibliografico finalizzate a permettere il link tra l'archivio bibliografico e l'archivio dei *metadati*
 - ✍ guida il caricamento e l'aggiornamento degli archivi descritti in 4.1
- ✍ L'AIP è il risultato della collaborazione di tutti e tre gli archivi
- ✍ Il DIP viene creato su richiesta dell'utente grazie alla etichetta UNIMARC 956



RLG and preservation I

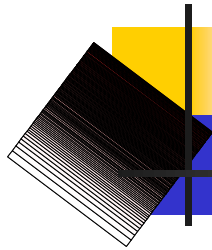
✚ Working Group on Preservation Issues of Metadata, Final Report, May, 1998: 16 elementi

- ✚ Date: data creazione del file
- ✚ Transcriber: (o) nome dell'agenzia responsabile per l'attribuzione dei metadati
- ✚ Producer: (o) agenzia responsabile della creazione fisica del file.
- ✚ Capture device: indica marca e funzionamento della videocamera o scanner utilizzato
- ✚ Capture details: (dipende dal tipo di capture device)
- ✚ Change history: documenta le modifiche subite dal file e le versioni generate, identificando chi le ha fatte e quando
- ✚ Validation key: conto delle righe o algoritmo basato sui bits che compongono il file al fine di attestarne l'integrità.



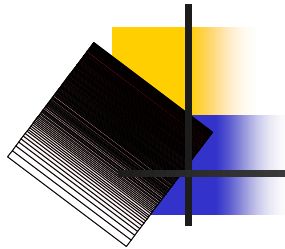
RLG and preservation II

- ✗ Encryption: tipo di criptazione utilizzata
- ✗ Watermark: indica se c'è un'impronta digitale
- ✗ Resolution: (es: dpi, ppi)
- ✗ Compression: indica se il file è compresso e il metodo e livello di compressione
- ✗ Source: descrive le caratteristiche fisiche della fonte e il suo ruolo nella catena (originale, copia, altro formato ..)
- ✗ Color: risoluzione dei pixel
- ✗ Color management: sistema usato per aumentare coerenza del colore nella cattura, visualizzazione e disseminazione di una immagine
- ✗ Color bar/gray scale bar: indicare se c'è o meno, e se c'è il tipo.
- ✗ Control targets: informazioni sugli indicatori inclusi nel file scansito a scopo di controllo di qualità, calibratura, verifica, etc.



SCHEMA MAG

- ✍ **gen** informazioni generali sul progetto e sul tipo di digitalizzazione;
- ✍ **bib** descriptive metadata on digitized object;
- ✍ **stru** structural metadata;
- ✍ **img** specific metadata for fixed images
- ✍ **ocr** specific metadata for optical character recognition
- ✍ **doc** section for pdf or for example rtf format files description



BDI: specifiche di interoperabilità

Set minimo degli elementi Dublin Core – Confronto DC-Unimarc ISBD/SBN-Scheda F

Dublin Core	UNIMARC	ISBD/SBN	Scheda F
Title	200\$a	Titolo proprio	SGLT
Title alternative	540\$a	Titolo attribuito	SGLA
Creator	700\$a	Autore	AUFN, AUFB
Subject	610\$a	Soggetto	SGTI
Identifier	001	Bid	NCTR, NCTN, NCTS
DateCreated	210 \$h	Data di esecuzione	LRD
DateIssue	210\$d	Data di pubbl.	DTSI, DTSF, DTSL
Coverage	300	Note	LRCS, LRCC, LRA

Fonte: ICCU, Linee di indirizzo per i progetti di digitalizzazione del materiale fotografico,
http://www.iccu.sbn.it/PDF/Linee_guida_fotografie.pdf