UniMc
Facoltà Beni Culturali
a.a. 2011/12

Corso di progettazione e gestione di risorse digitali per i beni culturali

prof. Pierluigi Feliciati

Modulo 3

il ciclo di vita di un progetto di digitalizzazione in ambito culturale

Basato su *Linee guida tecniche per i programmi di creazione di contenuti culturali digitali*, ed. it. 2.0, MINERVA 2006 , in http://www.minervaeurope.org/publications/technicalguidelines.htm

2. cosa significa digitalizzare?

Le **risorse culturali digitali**, tanto born digital che prodotte da progetti e programmi di digitalizzazione, si inseriscono in un **contesto informativo ampio**, che comprende anzitutto le altre risorse culturali, della più varia tipologia.

Esse possono però essere integrate anche con strumenti e risorse per la formazione di base e permanente, anche a distanza (eLearning), con strumenti e risorse per il turismo, con strumenti e risorse culturali divulgative (edutainment), e impiegate per l'editoria, la ricerca ecc.

2. cosa significa digitalizzare?

Per l'utente è indispensabile avere un **accesso integrato** che aggreghi in modo trasparente e renda disponibile con facilità il più ampio spettro possibile di risorse informative e culturali, e che consenta inoltre di interagire con un **limitato numero di interfacce semplici e intuitive** che agevolino la ricerca.

Per offrire **servizi qualificati**, le istituzioni culturali devono dunque superare le attuali divisioni di competenze ora nettamente distinte tra archivi, biblioteche, musei e le altre tipologie istituzionali, non più comprensibili per un normale utente, per realizzare la **convergenza nel campo della rappresentazione digitale di oggetti culturali**.

2. cosa significa digitalizzare?

Il quadro di riferimento per la costruzione dell'ambiente culturale digitale è quello **della rete, o networking**, che significa lavorare in rete, con un'enfasi che non è nella rete, come tecnologia e canale di comunicazione, ma nel lavorare, o meglio, collaborare insieme.

L'ambiente digitale trasforma il flusso di lavoro nelle istituzioni culturali. Non somiglia più a una catena di montaggio, in cui ogni fase precede la successiva; in ogni fase, invece, si deve tener conto di tutte le altre: nella fase di creazione dei contenuti, si deve sapere come l'oggetto culturale sarà presentato e reso accessibile all'utente, e al tempo stesso occorre aver deciso quale strategia verrà attuata per assicurare la conservazione a lungo termine dell'oggetto digitale creato.

2. cosa significa digitalizzare?

Nell'ottica della rete, inoltre, l'organizzazione coinvolgerà più istituzioni alla volta: **non possono esistere biblioteche digitali o servizi online isolati**, pena la loro inefficacia.

Nell'apertura al *networking*, elementi essenzialmente **tecnici** andranno combinati con accordi **organizzativi** e **collaborativi**, insieme a un governo complessivo del sistema a livello locale o nazionale.

Una nuova organizzazione è legata quindi all'architettura dell'informazione in rete, con una corretta progettazione di **servizi condivisi e di servizi locali distribuiti**.

2. cosa significa digitalizzare?

Si presenteranno qui principi, criteri, conoscenze di base per chi voglia progettare e costruire una biblioteca digitale, un archivio digitale, un museo virtuale.

Un supporto per i responsabili, per garantire che i prodotti dei progetti avviati corrispondano ai requisiti essenziali dell'*interoperabilità, accessibilità, conservazione a lungo termine, sicurezza e autenticità*.

Tutti questi requisiti richiedono una organizzazione del lavoro e dei servizi e una infrastruttura tecnologica adeguate, ma anche conoscenze e competenze essenziali da parte del personale coinvolto, che sole possono garantire sia scelte appropriate, sia una corretta gestione e attuazione di tutto il flusso di lavoro.

2. il ciclo di vita di un progetto di digitalizzazione
le 10 fasi

1. **Progettazione**
2. **Preparazione per la digitalizzazione**
3. **Trattamento degli originali**
4. **Processo di digitalizzazione**
5. **Memorizzazione e conservazione del master digitale**
6. **Creazione dei metadati**
7. **Pubblicazione**
8. **Reperimento**
9. **Riuso**
10. **Diritti di proprietà intellettuale e copyright**

2. il ciclo di vita di un progetto di digitalizzazione
1. progettazione

La progettazione deve prevedere le seguenti attività:

- definizione degli obiettivi, inclusa la strategia di fruizione dei contenuti
- definizione dei criteri di selezione del materiale da digitalizzare (per esempio: rarità, conservazione, integrazione delle raccolte, richieste utenti)
- individuazione dei diritti di proprietà intellettuale e di riproduzione
- definizione del sistema di gestione degli oggetti digitali
- individuazione delle risorse (umane, tecnologiche, finanziarie), dei costi e della sostenibilità del progetto. Le risorse umane possono essere interne e/o esterne all'istituzione che sviluppa il progetto-
- identificazione del flusso di lavoro idoneo al perseguimento delle fasi del progetto.

2. il ciclo di vita di un progetto di digitalizzazione
1. progettazione

Particolare attenzione deve essere rivolta al flusso di lavoro (*workflow*) idoneo per perseguire le diverse fasi del progetto.

Il flusso di lavoro deve essere definito e valutato in rapporto alla specifica organizzazione del lavoro adottata; esso varia, ad esempio, a seconda che le operazioni si svolgano esclusivamente all'interno dell'istituzione, oppure siano affidate, come avviene frequentemente, in parte o interamente, a ditte esterne.

Vediamo ad esempio la successione delle fasi del flusso di lavoro (workflow) per uno specifico progetto (la *Biblioteca digitale galileiana* dell'Istituto e Museo di storia della scienza (IMSS) di Firenze).

2. il ciclo di vita di un progetto di digitalizzazione
1. progettazione

Il flusso di lavoro

1. catalogazione dell'oggetto da digitalizzare
2. acquisizione delle immagini digitali, produzione dei file con nomenclatura prestabilita, produzione dei metadati relativi alle immagini
3. primo controllo di qualità delle immagini
4. archiviazione dei master digitali delle immagini produzione dei metadati amministrativi, descrittivi e strutturali
5. generazione di formati per la pubblicazione in rete
6. pubblicazione nell’area di lavoro interna (Intranet) per il secondo controllo di qualità
7. pubblicazione sul Web
8. collaudo

un possibile flusso di lavoro in/out

2. il ciclo di vita di un progetto di digitalizzazione
1. progettazione

Identificazione delle risorse digitali

Le risorse digitalizzate dovrebbero essere identificate senza ambiguità e dovrebbero essere puntabili in maniera univoca direttamente dal browser web di un utente. È importante, infatti, che l'utente finale (umano o automatico) abbia la possibilità di identificre esattamente e in modo attendibile una singola risorsa, anziché dover indicare il collegamento al sito web di un intero progetto.

L'uso di tecnologie superate e ormai vietate come i *frames*, ad esempio, se apparentemente sembrano risolvere alcuni problemi nella progettazione delle interfacce utente, di fatto impediscono la localizzazione e contestualizzazione delle risorse, la cui visualizzazione è costituita da più files montati insieme, che sul web sono identificati sempre dallo stesso indirizzo: quello del *frameset*

2. il ciclo di vita di un progetto di digitalizzazione
2. *preparare la digitalizzazione*

Gli oggetti da avviare alla digitalizzazione devono essere catalogati/inventariati; in mancanza di risorse catalografiche o inventariali preesistenti, la catalogazione avverrà nell'ambito del progetto di digitalizzazione.

La descrizione conforme degli originali è di estrema importanza infatti per stabilire un legame con la collezione/archivio/catalogo esistente e per la conservazione dell'originale e della fruizione, reperimento e interpretazione sia dell'originale sia della risorsa digitale che lo rappresenta.

Da evitare assolutamente l'uso di standard di descrizione degli oggetti digitali (come Dublin Core) per sostituire una corretta catalogazione/inventariazione.

2. il ciclo di vita di un progetto di digitalizzazione
2. *preparare la digitalizzazione*

Nella fase di preparazione i progetti devono tener conto delle caratteristiche fisiche degli oggetti da digitalizzare e degli usi cui saranno destinate le risorse digitali che si creano.

Per il materiale archivistico e librario, ad esempio, le caratteristiche da prendere in considerazione includeranno:

- stato di conservazione e fragilità, maneggevolezza, legatura, supporto, dimensioni, cartulazione
- ordinamento
- presenza di altri elementi notevoli, quali sigilli, miniature, annotazioni, stemmi ecc.

2. il ciclo di vita di un progetto di digitalizzazione
2. *preparare la digitalizzazione*

Nella fase di preparazione, infine, sono da scegliere hardware e software da impiegare nel progetto.

Nella scelta, si deve tener conto delle caratteristiche degli oggetti originali e, in particolare, di formato, dimensioni, stato di conservazione, tenendo conto dell'importanza di una fedele registrazione di proprietà quali il colore.

- Per l'hardware, si deve garantire che le apparecchiature prescelte per la digitalizzazione generino oggetti digitali di qualità adeguata agli usi previsti, entro accettabili limiti di costo.
- Il software adottato deve avere i requisiti di interoperabilità, portabilità, aderenza agli standard, apertura.

2. il ciclo di vita di un progetto di digitalizzazione
3. *formazione del personale*

I progetti devono garantire che il personale coinvolto riceva una preparazione adeguata al trattamento corretto degli originali e all'uso dell'hardware e del software da impiegare per la digitalizzazione. Ciò garantisce l'efficacia del processo e riduce qualunque rischio per gli originali.

Le esigenze di formazione devono essere identificate rispetto a:

- tecnologie adottate
- trattamento degli originali
- operazioni di catalogazione
- gestione complessiva dei programmi di digitalizzazione.

2. il ciclo di vita di un progetto di digitalizzazione
4. *Trattamento degli originali*

I responsabili del progetto devono valutare i rischi cui vanno incontro gli oggetti originali in qualunque processo di digitalizzazione, a maggior ragione quando si tratta di pezzi unici, preziosi o fragili.

Si devono allora concordare le modalità di lavoro con i responsabili della conservazione degli originali.

Le problematiche della conservazione riguarderanno anche, una volta creato, l'oggetto digitale che rappresenta l'oggetto fisico.

2. il ciclo di vita di un progetto di digitalizzazione
5. il master digitale

I problemi posti dalla conservazione delle risorse digitali devono essere considerati come parte integrante del processo di produzione digitale.

La possibilità di conservare a lungo termine delle risorse digitali è agevolata dall'adozione di standard, ma dipende anche dalla documentazione dell'insieme dei procedimenti tecnici che hanno condotto alla creazione di un oggetto; molte informazioni essenziali possono, quindi, essere registrate solo al momento della creazione delle risorse da conservare.

2. il ciclo di vita di un progetto di digitalizzazione
5. il master digitale

La conservazione a lungo termine delle risorse digitali non dipende solo dalla scelta di opportuni formati o supporti. Deve essere considerata piuttosto una responsabilità gestionale essenziale per chi possieda e gestisca contenuti informativi digitali, poiché garantisce il loro uso e riuso a lungo termine.

La conservazione dipende da un gran numero di fattori, anche al di fuori del processo di digitalizzazione, come la stabilità istituzionale, la continuità dei finanziamenti e il possesso dei diritti di proprietà intellettuale.

È possibile tuttavia adottare strategie di tipo tecnico nel corso del processo di digitalizzazione allo scopo di agevolare la conservazione. Vediamole.

2. il ciclo di vita di un progetto di digitalizzazione
5. il master digitale

1. Formati dei file

Quando si producono delle risorse digitali si dovrebbero impiegare formati standard aperti, per massimizzarne l'accesso. L'impiego di formati aperti favorisce l'interoperabilità, garantendo la riusabilità delle risorse, che possono essere create e modificate da un gran numero di applicazioni. Si eviterà in tal modo anche la dipendenza da un unico fornitore.

Qualora non impieghino standard aperti, la normativa italiana prevede che i progetti pubblici giustifichino il ricorso a formati proprietari. Se si utilizzino formati proprietari, si deve prevedere una strategia di migrazione che consenta una successiva transizione verso standard aperti.

2. il ciclo di vita di un progetto di digitalizzazione
5. *il master digitale*

2. Digitalizzazione e memorizzazione del testo

a) Codifica dei caratteri

La codifica dei caratteri è un algoritmo che consente di rappresentare i caratteri in forma digitale, definendo una corrispondenza tra le sequenze di codice dei caratteri (i numeri interi corrispondendo ai caratteri in un repertorio) e sequenze di valori di 8 bit (byte o ottetti). Per poter interpretare correttamente i byte che compongono un oggetto digitale, un'applicazione richiede l'indicazione della codifica di carattere utilizzata nel documento.

La codifica dei caratteri usata nei documenti di testo dovrebbe essere quindi sempre segnalata esplicitamente.

2. il ciclo di vita di un progetto di digitalizzazione
5. il master digitale

2. Digitalizzazione e memorizzazione del testo

b) **Formati dei documenti**

Il contenuto testuale dovrebbe essere creato e gestito in un formato strutturato, adatto alla generazione di documenti **HTML o XHTML** per la pubblicazione. Nella maggior parte dei casi, l'opzione più opportuna sarà memorizzare il contenuto testuale in un formato a base SGML o **XML** che si uniformi a una DTD o a uno XMLschema ben formati, sulla base dei quali si validino i file.

In alcuni casi, si può scegliere di memorizzare contenuti a base esclusivamente testuale utilizzando il **Portable Document Format** (PDF) formato proprietario di Adobe. I file PDF sono compatti e possono essere visualizzati e stampati grazie al programma gratuito Acrobat Reader. Comunque, come accade per qualunque soluzione proprietaria, la sua adozione comporta dei rischi.

2. il ciclo di vita di un progetto di digitalizzazione
5. il master digitale

3. Digitalizzazione e memorizzazione di immagini fisse

Le immagini fisse digitali – come si è detto - si dividono in due categorie: le immagini a matrice di punti (RASTER) e le immagini vettoriali .

Il processo di digitalizzazione genera solitamente un'immagine a matrice di punti. hanno la forma di una griglia o matrice 2D, in cui ogni elemento dell'immagine (pixel) ha una collocazione unica all'interno della matrice e un valore (o un insieme di valori) di colore indipendente, che può essere modificato separatamente. Rientrano nella categoria delle immagini a matrice di punti anche le immagini create al computer (born digital), cosiddette grafiche non vettoriali, come loghi, icone, disegni al tratto ecc.

2. il ciclo di vita di un progetto di digitalizzazione
5. il master digitale

Immagini raster: parametri di acquisizione

La scelta dei parametri di acquisizione richiesti per digitalizzare utilmente una risorsa è determinata dalle dimensioni dell'originale, dalla quantità di dettagli in esso presenti, dai livelli di contrasto (dinamica) che si desidera rappresentare e dall'uso (conservazione, pubblicazione, editing ecc.) che si vuol fare dell'immagine digitale; infine, dalle caratteristiche tecniche del sistema di acquisizione.

Tra i molti fattori che determinano la qualità delle immagini prodotte da un determinato sistema di acquisizione, i due più importanti dipendono da:

2. il ciclo di vita di un progetto di digitalizzazione
5. *il master digitale*

(./.) parametri di acquisizione immagini raster

La densità di campionamento spaziale riferita anche come *risoluzione* spaziale: è la frequenza spaziale alla quale i campioni del documento originale sono acquisiti. Essa è legata alla capacità di rappresentare i dettagli e viene espressa come numero di campioni per pollice (spi: samples per inch), o più comunemente di pixel per inch (ppi).

La profondità di colore (o profondità di bit = bit depth) è il numero di bit a disposizione per rappresentare tutti i livelli di luminosità, utilizzabili per rappresentare i diversi colori (o toni di grigio) dell'originale. Tale numero si esprime in bit per pixel (bpp).

Le immagini dovrebbero essere acquisite alle maggiori densità spaziale e profondità di colore opportune ai fini del progetto, purché ne risulti un costo abbordabile e siano funzionali agli usi previsti.

2. il ciclo di vita di un progetto di digitalizzazione
5. *il master digitale*

Formato e compressione dei file

In stretta correlazione con i parametri di acquisizione è il formato di memorizzazione dei file immagine acquisiti, e in particolare il tipo di compressione che può essere operata sui dati digitali dell'immagine. Esistono due principali tipologie di compressione: quella reversibile o senza perdita (*lossless*) e quella con perdita (*lossy*).

Le **immagini raster (a matrice di punti) destinate al master digitale** dovrebbero, in linea di principio, essere memorizzate nella forma non compressa generata dal processo di digitalizzazione, senza subire trattamenti successivi; tuttavia l'utilizzo di metodi di compressione senza perdita consente di salvare i dati utilizzando una minore quantità di spazio sui supporti di memorizzazione.

2. il ciclo di vita di un progetto di digitalizzazione
5. il master digitale

Immagini vettoriali

Le immagini vettoriali sono costituite da molteplici oggetti geometrici (linee, ellissi, poligoni e altre forme) costruiti attraverso una sequenza di comandi o matematiche per tracciare linee e forme.

Le immagini grafiche vettoriali dovrebbero essere create e memorizzate in un formato aperto come Scalable Vector Graphics (SVG), un linguaggio XML per la descrizione di grafici di questo tipo. I disegni SVG possono essere interattivi e dinamici e sono scalabili secondo lo schermo di visualizzazione e le diverse risoluzioni della stampante.

Nel caso si usi il formato proprietario Macromedia Flash, nel progetto dovrebbe essere prevista una strategia di migrazione verso formati aperti.

2. il ciclo di vita di un progetto di digitalizzazione
5. il master digitale

5. Digitalizzazione e memorizzazione di **documenti video**

I documenti video dovrebbero per consuetudine essere memorizzati nel **formato non compresso** derivante dal dispositivo di registrazione, senza ulteriori trattamenti.

Il documento video dovrebbe essere realizzato con i valori più alti di densità spaziale, di profondità di bit e di frequenza di quadro, che siano finanziariamente accessibili e funzionali rispetto agli usi previsti.

Ogni progetto deve identificare il livello minimo di qualità proprio del progetto stesso.

2. il ciclo di vita di un progetto di digitalizzazione
5. il master digitale

6. Digitalizzazione e memorizzazione di **documenti audio**

I documenti audio dovrebbero per consuetudine essere memorizzati nel **formato non compresso** derivato direttamente dai dispositivi di registrazione, senza ulteriori trattamenti, come ad esempio la riduzione del rumore.

I trattamenti successivi (compressione) devono essere considerati nella fase di pubblicazione.

2. il ciclo di vita di un progetto di digitalizzazione
5. *il master digitale*

7. scelta dei **supporti di memorizzazione**

I diversi supporti di memorizzazione digitale richiedono per l'accesso differenti requisiti hardware e software e supporti diversi pongono sfide diverse in ordine al processo di memorizzazione e alla gestione e conservazione. Vi sono due tipi di minacce alla continuità di accesso ai supporti digitali:

- il **deterioramento** fisico o il **danneggiamento** del supporto
- **cambiamenti tecnologici** che portano all'obsolescenza dell'infrastruttura hardware e software necessaria per accedere al supporto.

Le risorse generate nei progetti di digitalizzazione generalmente vengono memorizzate sui dischi rigidi di uno o più server e/o su supporti portatili (DVD etc.).

2. il ciclo di vita di un progetto di digitalizzazione
5. *il master digitale*

La scelta dei **supporti di memorizzazione** (2)

I responsabili dei progetti dovrebbero riflettere sulla possibilità di realizzare copie di tutte le risorse digitali – registrazioni dei metadati e oggetti digitali – su due tipi diversi di supporti per la memorizzìazione (backup differenziati).

Una copia deve essere conservata in un luogo distinto da quello principale per garantirne la salvaguardia nel caso in cui una calamità coinvolga il luogo di conservazione principale. Tutti i trasferimenti verso supporti portatili dovrebbero essere registrati.

I supporti devono essere regolarmente "**rinfrescati**" (*refreshment*, ovvero, i dati devono essere copiati su un nuovo supporto).

2. il ciclo di vita di un progetto di digitalizzazione
5. il master digitale

8.Strategie per la **conservazione a lungo termine**

Esistono tre approcci tecnici principali alla conservazione digitale:

- la conservazione della tecnologia,
- l'emulazione della tecnologia
- la migrazione dei dati.

I primi due approcci si concentrano sulla tecnologia utilizzata per accedere all'oggetto, o conservando l'hardware e il software originali, o utilizzando la tecnologia attuale per simulare l'ambiente originale.

Le strategie di migrazione si concentrano sul mantenimento degli oggetti digitali in un formato accessibile tramite tecnologia corrente.

2. il ciclo di vita di un progetto di digitalizzazione
5. *il master digitale*

8.la **conservazione a lungo termine e i metadati**

La cattura dei metadati rappresenta un aspetto fondamentale della strategia di conservazione basata sulla migrazione.

I metadati sono necessari per **supportare la gestione dell'oggetto e del procedimento di migrazione**, ma, in più, la migrazione conduce inevitabilmente, almeno a lungo termine, ad alcuni cambiamenti o a perdite delle funzionalità originarie. Quando queste sono significative per l'interpretazione dell'oggetto, si ricorrerà ai metadati relativi al processo di migrazione e a quelli relativi all'oggetto originario e alle sue trasformazioni per cercare di comprendere la funzionalità presente nell'ambiente tecnologico originario.

2. il ciclo di vita di un progetto di digitalizzazione
6. *metadati*

Le operazioni supportate dai metadati possono comprendere **individuazione e recupero, gestione e controllo** (incluso il rights management) **e conservazione a lungo termine** delle risorse.

Per supportare queste differenti funzioni possono essere necessari metadati diversi. Nella pratica, però, singoli schemi di metadati spesso supportano molteplici funzioni e comprendono più d'una delle categorie che si elencano.

2. il ciclo di vita di un progetto di digitalizzazione
6. *metadati*

Ambito di applicazione dei metadati

Può essere necessario fornire metadati per varie classi di risorse, in particolare per:

- l'oggetto fisico digitalizzato
- gli oggetti digitali creati durante il processo di digitalizzazione e memorizzati come master digitali
- gli oggetti digitali derivati da tali master per la diffusione in ambiente di rete
- le nuove risorse create impiegando tali oggetti digitali
- collezioni di ogni sorta di oggetti sopra citati.

2. il ciclo di vita di un progetto di digitalizzazione
6. *metadati*

I metadati possono essere distinti in tre categorie funzionali:

- Descrittivi: per l'identificazione e il recupero degli oggetti digitali; sono costituiti da descrizioni normalizzate dei documenti fonte (o dei documenti digitali nativi), risiedono generalmente nelle basi dati dei sistemi di information retrieval all'esterno dei depositi degli oggetti digitali e sono collegati a questi ultimi tramite appositi link.
- Amministrativi e gestionali / di conservazione: per le svariate operazioni di gestione degli oggetti digitali all'interno dell'archivio.
- Strutturali: per descrivere la struttura interna dei documenti (ad esempio, introduzione, capitoli, indice di un libro) e gestire le relazioni fra le varie parti componenti gli oggetti digitali.

2. il ciclo di vita di un progetto di digitalizzazione
6. *metadati*

i metadati descrittivi

Servono per il reperimento e l'interpretazione dell'oggetto digitale. Per favorire il reperimento delle loro risorse da parte di un'ampia gamma di applicazioni e servizi, i progetti devono catturare e immagazzinare metadati descrittivi sufficienti a generare per ciascun oggetto una descrizione attraverso il **set di metadati Dublin Core** (DCMES, Dublin Core Metadata Element Set) nella sua formulazione semplice/non qualificata.

Il DCMES è uno schema di metadati descrittivi molto semplice, prodotto di un'iniziativa interdisciplinare e progettato per supportare il reperimento di risorse attraverso una vasta gamma di domini.

2. il ciclo di vita di un progetto di digitalizzazione
6. *metadati descrittivi*

Il Dublin Core Metadata Element Set

Il DCMES definisce quindici elementi che consentono di reperire risorse appartenenti a domini diversi:

Titolo, Creatore, Soggetto, Descrizione, Editore, Autore di contributo subordinato, Data, Tipo, Formato, Identificatore, Fonte, Lingua, Relazione, Copertura, Gestione dei diritti.

Non si deve intendere che per ogni oggetto vada registrato soltanto il set elementare di metadati DC: la capacità di fornire i metadati DC semplici è, piuttosto, il requisito minimo indispensabile per consentire il reperimento della risorsa. Nella prassi, il set base di metadati DC è generalmente un sottoinsieme di un più ricco corredo di metadati *item level*.

2. il ciclo di vita di un progetto di digitalizzazione
6. metadati descrittivi: DCMES

Nell'insieme dei **dieci attributi** , **sei sono comuni** a tutti gli elementi DC:

- Versione — 1.1
- Registrazione di autorità — Dublin Core Metadata Initiative
- Lingua — en, it, fr …
- Obbligatorietà — opzionale
- Tipo di dato — stringa di caratteri
- Occorrenza massima — illimitata

gli **altri quattro** sono:

- Nome, Identificatore, Definizione, Commento

2. il ciclo di vita di un progetto di digitalizzazione
6. metadati descrittivi: DCMES

gli elementi di Dublin Core (1-8)

nome	*identificatore*	*definizione*
Titolo	**DC.Title**	Un nome dato alla risorsa
Creatore	**DC.Creator**	Un'entità che ha la responsabilità principale della produzione del contenuto della risorsa
Soggetto	**DC.Subject**	L'argomento della risorsa.
Descrizione	**DC.Description**	Una spiegazione del contenuto della risorsa
Editore	**DC.Publisher**	Un'entità responsabile della produzione della risorsa, disponibile nella sua forma presente
Autore di contr. subordinato	**DC.Contributor**	Un'entità responsabile della produzione di un contributo al contenuto della risorsa
Data	**DC.Date**	Una data associata a un evento del ciclo di vita della risorsa, nella forma YYYY-MM-DD
Tipo	**DC.Type**	La natura o il genere del contenuto della risorsa, ad es. la lista dei Dublin Core Types

2. il ciclo di vita di un progetto di digitalizzazione
6. *metadati descrittivi: DCMES*

gli elementi di Dublin Core (9-15)

nome	*identificatore*	*definizione*
Formato (cfr. MIME)	**DC.Format**	La manifestazione fisica o digitale della risorsa
Identificatore	**DC.Identifier**	Un riferimento univoco alla risorsa nell'ambito di un dato contesto (URL, DOI, ISBN…)
Fonte	**DC.Source**	Un riferimento a una risorsa dalla quale è derivata la risorsa in oggetto
Lingua	**DC.Language**	La lingua del contenuto intellettuale della risorsa
Relazione	**DC.Relation**	Un riferimento alla risorsa correlata
Copertura	**DC.Coverage**	L'estensione o scopo del contenuto della risorsa (sua estensione spaziale, temporale o giurisdiz.)
Gestione dei diritti	**DC.Rights**	Informazione sui diritti esercitati sulla risorsa

2. il ciclo di vita di un progetto di digitalizzazione
6. *metadati*

I metadati amministrativi e gestionali, più che quelli descrittivi, assumono una importanza preponderante ai fini del mantenimento e dell'accessibilità a lungo termine dell'eredità culturale digitale.

Essi documentano le modalità di generazione, immissione, archiviazione e manutenzione degli oggetti digitali; forniscono inoltre specifiche formali per le fasi di raccolta e archiviazione dei metadati.

Si prestano particolarmente bene a essere utilizzati all'interno di modelli logico-funzionali dell'archivio degli oggetti digitali come l'Open Archival Information System (OAIS), divenuto nel 2003 standard ISO 14721.

2. il ciclo di vita di un progetto di digitalizzazione
6. *metadati amministrativi gestionali*

I metadati amministrativi vengono usati per la gestione dell'oggetto digitale e per offrire maggiori informazioni sulla sua creazione e su ogni restrizione d'uso. La categoria può comprendere:

- metadati **tecnici**, che descrivono le caratteristiche tecniche della risorsa digitale
- metadati **per la conservazione**:
  - metadati relativi alla **fonte**, che descrivono l'oggetto dal quale è derivata la risorsa digitale
  - metadati relativi alla **provenienza**, che descrivono la storia delle operazioni effettuate su un oggetto digitale
- metadati per la **gestione dei diritti**
- metadati per la **conservazione a lungo termine**

2. il ciclo di vita di un progetto di digitalizzazione
6. *metadati amministrativi gestionali*

I metadati tecnici

comprendono informazioni che possono essere raccolte efficacemente solo nel corso dello stesso processo di digitalizzazione; ad esempio, informazioni sulla natura del materiale fonte, sulla strumentazione usata per la digitalizzazione e sui parametri adottati (formati, tipi di compressione ecc.) e sugli agenti responsabili per il processo di digitalizzazione. E' possibile generare alcuni di questi metadati attraverso il software di digitalizzazione.

Per le **immagini**, un comitato dello statunitense NISO (National Information Standard Organisation) ha prodotto un *data dictionary* di metadati tecnici per le immagini digitali raster fisse (NISO MIX).

2. il ciclo di vita di un progetto di digitalizzazione
6. *metadati amministrativi gestionali*

I metadati per la conservazione

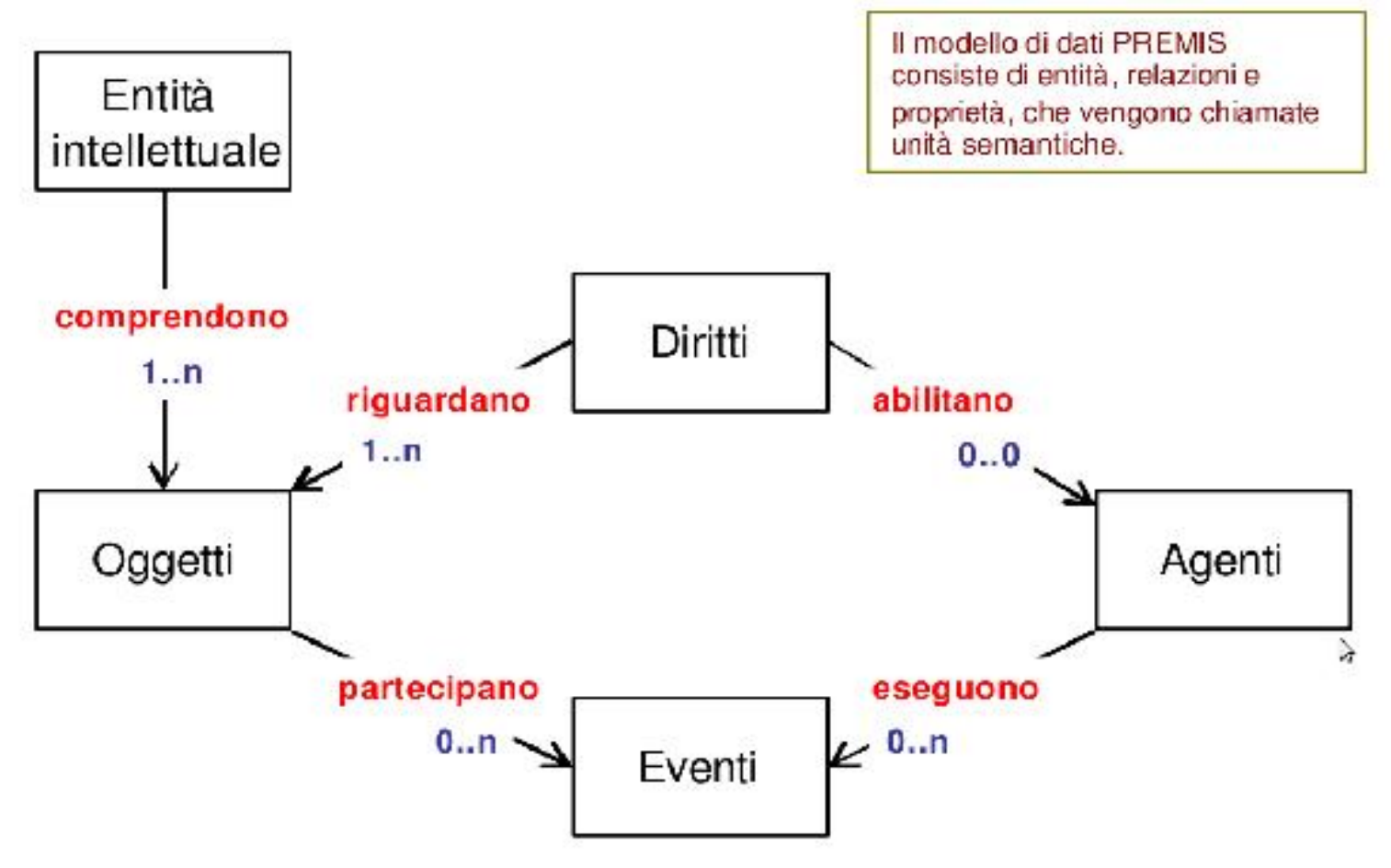
Un set di sedici elementi base di metadati per la conservazione a lungo termine è stato pubblicato nel 1998 da un gruppo di lavoro sui metadati per preservation, istituito dal Research Libraries Group (RLG).

Nel quadro del modello OAIS, un gruppo di lavoro OCLC/RLG sui metadati per la conservazione a lungo termine ha sviluppato proposte per due componenti del modello informativo OAIS pertinenti direttamente ai metadati per la preservation.

Il modello più attuale è quello di PREMIS (2008), che comprende entità (oggetti, rappresentazioni,agenti ed eventi), loro relazioni e proprietà delle entità per documentare interamente, nel tempo, la vita di ogni oggetto nel/i deposito/i dove viene conservato.

2. il ciclo di vita di un progetto di digitalizzazione
6. *metadati amministrativi gestionali*

Lo schema delle entità di PREMIS



2. il ciclo di vita di un progetto di digitalizzazione
6. *metadati amministrativi gestionali*

I MAG: lo standard italiano

Con l'acronimo MAG - Metadati Amministrativi e Gestionali – è stato proposto un *application profile* con l'obiettivo di fornire le specifiche formali per la fase di raccolta, di trasferimento e disseminazione dei metadati e dei dati digitali nei rispettivi archivi. Lo Schema MAG è realizzato e mantenuto dal Comitato Mag.

L'ICCU ha costituito nel 2000 un *Gruppo di studio sugli standard e le applicazioni di metadati nei beni culturali* a cui hanno partecipato rappresentanti di biblioteche, musei e archivi.

Da luglio 2003 si è formalmente costituito un gruppo di lavoro permanente – il Comitato Mag.

2. il ciclo di vita di un progetto di digitalizzazione
6. *metadati amministrativi gestionali*

I MAG: lo standard italiano

lo schema MAG fornisce delle **specifiche formali per le fasi di raccolta e archiviazione dei metadati** e fornisce elementi per:

- identificare in maniera univoca gli oggetti digitali
- certificare l'autenticità e l'integrità dei contenuti
- documentare la catena di custodia degli oggetti digitali
- documentare i processi tecnici eseguiti per la conservazione permanente degli oggetti digitali
- informare sulle condizioni e i diritti di accesso agli oggetti digitali da parte degli utenti.

2. il ciclo di vita di un progetto di digitalizzazione
6. *metadati amministrativi gestionali*

Sui MAG occorre specificare che:

- più che per tipologie di documenti fonte (es. periodici, musica a stampa o manoscritta, carte geografiche, fotografie etc.), MAG viene definito per tipologie di oggetti digitali (es. immagini statiche, testi prodotti con tecnologia OCR, suono, audiovisivi etc.);
- i MAG sono espressi nella sintassi XML;
- i MAG si presentano come uno standard aperto, documentato, liberamente disponibile e del tutto indipendente da specifiche piattaforme hardware e software;
- di norma i MAG vanno associati alla risorsa digitale nel suo complesso, e comprendono informazioni generali sul progetto e sul metodo di digitalizzazione ed altre informazioni relative all'unità elementare dell'archivio digitale, cioè al singolo oggetto digitale;

2. il ciclo di vita di un progetto di digitalizzazione
6. *metadati amministrativi gestionali*

Sui MAG occorre specificare che:

- i MAG comprendono una quantità limitata di metadati descrittivi, ereditati dagli applicativi di catalogazione ed espressi in formato Dublin Core. Di norma tali metadati descrittivi sono funzionali a operazioni gestionali interne all'archivio degli oggetti digitali;
- di norma, con poche eccezioni, i MAG vanno raccolti o prodotti con procedure automatizzate nel corso del processo di digitalizzazione o trattamento (post-processing) dell'oggetto digitale.

2. il ciclo di vita di un progetto di digitalizzazione
6. *metadati amministrativi gestionali*

Lo schema risulta composto di diverse sezioni, utilizzabili a seconda del contenuto digitale e dell'impiego dello stesso:

- GEN: informazioni generali sul progetto e sul tipo di digitalizzazione
- BIB: metadati descrittivi dell’oggetto analogico
- STRU: metadati strutturali dell’oggetto digitale
- IMG: metadati specifici per le immagini fisse
- OCR: metadati specifici relativi al riconoscimento ottico del testo
- DOC: metadati specifici per oggetti digitali in formato testo che possono essere derivati o born digital
- AUDIO: metadati specifici per file audio
- VIDEO: metadati specifici per file video
- DIS: metadati specifici per la distribuzione di oggetti digitali.

2. il ciclo di vita di un progetto di digitalizzazione
6. *metadati strutturali*

I metadati strutturali descrivono le **relazioni logiche o fisiche che collegano le parti di un oggetto composito**.

Un libro “fisico”, ad esempio, consiste in una sequenza di pagine. Un documento archivistico può consistere in una sequenza di carte di un fascicolo inserito in una busta, a sua volta parte di una serie inserita in un fondo particolare.

Il processo di digitalizzazione può generare un certo numero di risorse digitali distinte, ad esempio una immagine per ogni pagina, ma il fatto che queste risorse formino una sequenza e che quella sequenza costituisca un oggetto composito o faccia parte di una struttura complessa multilivellare è evidentemente essenziale per il loro uso e la loro interpretazione.

2. il ciclo di vita di un progetto di digitalizzazione
6. *metadati strutturali*

Il Metadata Encoding and Transmission Standard (METS) fornisce un formato di codifica per metadati descrittivi, amministrativi e strutturali, ed è progettato per supportare sia la gestione di oggetti digitali che la distribuzione e lo scambio di oggetti digitali fra i diversi sistemi.

Il MAG include una sezione dedicata alla codifica di metadati strutturali.

Le IMS Content Packaging Specification (specifiche IMS per la confezione del contenuto) indicano un modo per descrivere la struttura e organizzare le risorse composite per l'apprendimento.

2. il ciclo di vita di un progetto di digitalizzazione
6. *metadati strutturali*

Descrizione delle collezioni

Una risorsa digitale **non è mai creata isolatamente**, ma come parte di una collezione digitale, e dovrebbe essere presa in considerazione **nel contesto di quella collezione** e del suo sviluppo.

Le stesse collezioni possono poi essere viste come componenti intorno ai quali è possibile costruire molti tipi diversi di servizi digitali.

Le collezioni dovrebbero essere descritte in modo tale da consentire a un utente di identificarne le caratteristiche salienti e in modo tale da poterle integrare nel più ampio novero delle collezioni digitali esistenti e all'interno di servizi digitali operanti attraverso queste collezioni.

2. il ciclo di vita di un progetto di digitalizzazione
6. *metadati strutturali*

Descrizione delle collezioni

Per questo livello di metadati esistono attualmente:

- lo schema per la descrizione delle collezioni del Research Support Library Programme (RSLP)
  http://www.ukoln.ac.uk/metadata/rslp/
- il set per la descrizione delle collezioni definito da MINERVA
  http://www.minervaeurope.org/intranet/documents/specinv0311.pdf
- il Dublin Core Collection Description Application Profile

Inoltre, è in corso il progetto MICHAEL, che sta costituendo un inventario europeo di collezioni culturali digitali basato sugli standard Dublin Core e le specifiche MINERVA.

Cfr. http://www.michael-culture.it/ e http://www.michael-culture.eu/

2. il ciclo di vita di un progetto di digitalizzazione
7. pubblicazione

Prevedibilmente, **l'accesso alle risorse digitali da parte dell'utente finale** avverrà anzitutto attraverso l'uso di **reti locali** e/o di **protocolli Internet**.

La preparazione per la pubblicazione richiede che il master digitale venga elaborato per generare **oggetti digitali adatti** a essere usati nel contesto di Internet, tipicamente attraverso meccanismi di scalabilità, come la riduzione della qualità e la generazione di file di dimensioni adeguate al trasferimento attraverso la rete, oppure attraverso metodi di trasmissione progressiva, che consentono una fruizione graduale e controllabile a seconda di parametri collegati con il profilo utente.

2. il ciclo di vita di un progetto di digitalizzazione
7. pubblicazione

Inoltre, **i file audio e le sequenze video** possono essere resi disponibili sia per lo scaricamento (download) che per lo streaming.

In quest'ultima modalità, anziché trasferire l'intero file prima che possa avviarsi la sua esecuzione, viene creato uno spazio di buffer (memoria tampone o di transito) sul computer dell'utente e i dati vengono trasmessi nel buffer. Non appena il buffer è pieno, il file streaming comincia a essere eseguito, mentre i dati continuano ad essere trasmessi.

2. il ciclo di vita di un progetto di digitalizzazione
7. pubblicazione

Deve esser tenuto in considerazione che è nella logica delle reti che possano variare anche sensibilmente:

- il tipo di hardware e il tipo di software impiegati dagli utenti
- i livelli di larghezza di banda entro i limiti dei quali gli utenti operano.

Per raggiungere il più vasto pubblico possibile, i progetti dovrebbero rendere le risorse disponibili i**n più dimensioni, formati, risoluzioni o velocità di bit alternativi**, scegliendo, quando possibile, formati che offrano scalabilità di risoluzione e di qualità delle immagini.

2. il ciclo di vita di un progetto di digitalizzazione
7. pubblicazione

1. Procedure per la pubblicazione

1 Pubblicazione del testo

La codifica dei caratteri adottata nei documenti di testo dovrebbe essere trasmessa nello header HTTP e anche registrata all'interno dei documenti come opportuno (UTF-8, Unicode etc.)).

Il contenuto a base testuale dovrebbe essere distribuito come XHTML 1.0 oppure HTML 4 (o versioni successive), sebbene possa talora essere adeguato l'impiego di formati SGML o XML conformi ad altre DTD o schemi.

In alcuni casi, la distribuzione in formati proprietari come PDF (Portable Document Format), RTF (Rich Text Format) o Microsoft Word può essere opportuna, come formato aggiuntivo rispetto a XHTML/HTML, ma i progetti devono garantire che i problemi relativi all'accessibilità vengano affrontati.

2. il ciclo di vita di un progetto di digitalizzazione
7. pubblicazione

1. Procedure per la pubblicazione

1 Pubblicazione di immagini fisse

Immagini a matrice di punti (raster) - Dovrebbe essere presa in considerazione la possibilità di offrire la stessa immagine in diversi formati, dimensioni, qualità per consentirne una fruizione adeguata al contesto d'uso. Alle decisioni in merito a dimensioni e qualità dell'immagine possono concorrere le considerazioni riguardanti i diritti di proprietà intellettuale.

Le tipologie dimensionali per le immagini raster sul Web sono:

- **immagini francobollo** (icone, o thumbnail): con un lato di 100-200 pixel ca.
- **immagini cartolina**, derivate da un originale riducendone le dimensioni (600 pixel ca. per la dimensione maggiore)
- **immagini full-screen** (un massimo di 1024-1280 pixel per la dimensione maggiore).

2. il ciclo di vita di un progetto di digitalizzazione
7. pubblicazione

1. Procedure per la pubblicazione

1 Pubblicazione di immagini fisse

Immagini a matrice di punti (raster) - Formato dei file

Le immagini raster derivanti dalla digitalizzazione di documenti o fotografie dovrebbero essere messe a disposizione sul Web in formato JPEG/SPIFF.

Le immagini grafiche non vettoriali dovrebbero essere distribuite sul Web usando il Graphical Interchange Format (GIF) o il formato Portable Network Graphics (PNG).

Tra i formati di compressione emergenti c'è anche il JPEG 2000, concepito per risolvere gran parte delle esigenze di un progetto di digitalizzazione e pubblicazione via Web, ma ancora non diffuso a livello mondiale a livello tale da imprne l'uso.

2. il ciclo di vita di un progetto di digitalizzazione
7. *pubblicazione*

1. Procedure per la pubblicazione

1 Pubblicazione di immagini fisse

Immagini a matrice di punti (raster) - Tempi di trasmissione

Si dovrebbe tenere in considerazione che la larghezza di banda a disposizione di un utente qualsiasi può essere tale (modem a 56 Kbps) da richiedere un tempo notevole per la visualizzazione completa di una immagine trasmessa via Web; anche in questo caso è opportuna una scelta oculata tra formati, dimensioni e qualità dell'immagine.

A titolo d'esempio, i tempi di scaricamento per un file compresso di circa 150 KByte sono questi:

Scaricamento (download) di un file di ~150 Kbyte	
Alla velocità di (bps)	Tempo impiegato (s.)
56 K	~ 28
128 K	~ 12
384 K	~ 4
512 K	~ 3
1 M	~ 1,4

2. il ciclo di vita di un progetto di digitalizzazione
7. pubblicazione

1. Procedure per la pubblicazione

1 Pubblicazione di immagini fisse

Immagini grafiche vettoriali

Le immagini grafiche vettoriali dovrebbero essere distribuite sul Web tramite l'uso di formati Scalable Vector Graphics (SVG).

In alternativa si può prendere in considerazione la possibilità di rasterizzare le immagini grafiche vettoriali, ossia di trasformarle in immagini bitmap, in modo da trasmetterle secondo le modalità già viste per le immagini raster o fotografiche.

La scelta del software di rasterizzazione è importante in quanto è responsabile della qualità finale dell'immagine. Con questo metodo è possibile evitare l'invio su Web di file vettoriali originali coperti da copyright.

2. il ciclo di vita di un progetto di digitalizzazione
7. pubblicazione

1. Procedure per la pubblicazione

2 Pubblicazione di documenti video

Si dovrebbe tenere in considerazione il fatto che l'accesso degli utenti al video può essere limitato da restrizioni poste dall'ampiezza di banda: può pertanto essere opportuno fornire una serie di file o stream ("flussi") di qualità differente.

Per lo scaricamento (downloading), i video dovrebbero essere distribuiti in rete per il download (scaricamento) attraverso il formato MPEG-1 o MPEG-4. In alternativa, posso essere adottati i formati proprietari Microsoft Audio Video Interleave (AVI), Windows Media Video (WMV) o Apple Quicktime.

I video da trasmettere (streaming) dovrebbero essere distribuiti sul Web nei formati Microsoft Advanced Streaming Format (ASF), Windows Media Video (WMV) o Apple Quicktime.

2. il ciclo di vita di un progetto di digitalizzazione
7. pubblicazione

1. Procedure per la pubblicazione

3 3D e realtà virtuale

I progetti che fanno uso di *fly through* e di modelli di realtà virtuale tridimensionale devono prendere in considerazione le esigenze degli utenti che accedono al loro sito usando comuni computer e connessioni modem a 56k.

I modelli di realtà virtuale vengono usati tipicamente per la ricostruzione di edifici e di altre strutture o nella simulazione di intere aree di un paesaggio. I modelli sono tradizionalmente costruiti e mostrati usando potenti workstation, e questo è tuttora il caso dei modelli più dettagliati. Questi non sono adatti ai progetti cui è richiesto di distribuire in Internet i propri risultati al più vasto pubblico possibile. Tuttavia, modelli meno complessi possono utilmente essere inseriti nei siti web a disposizione degli utenti.

2. il ciclo di vita di un progetto di digitalizzazione
7. pubblicazione

1. Procedure per la pubblicazione

4 Sistemi Informativi Geografici (GIS)

Molti beni e contenuti culturali, essendo fondati sul territorio, posseggono un indirizzo geografico. Questa loro caratteristica offre la possibilità di applicare ad essi specifici strumenti di conoscenza e di analisi.

Tali strumenti sono generalmente individuati con il nome di SIT (Sistemi Informativi Territoriali), in inglese Geographic Information Systems (GIS). Essi si definiscono come l'insieme di strumenti, apparati, metodi e dati in grado di analizzare, progettare, controllare e gestire l'ambiente e il territorio e nel caso in ispecie i beni culturali.

Le informazioni territoriali possono essere memorizzate sia in modalità raster che vettoriale, e possedere diversi tipi di indirizzo geografico, da quello alfanumerico (come l'indirizzo stradale) a quello cartografico (quali le coordinate geografiche).

2. il ciclo di vita di un progetto di digitalizzazione
7. pubblicazione

1. Procedure per la pubblicazione

4 Siti web

Le risorse dei progetti devono essere **accessibili tramite un browser web**.

Questo si ottiene normalmente usando il linguaggio HTML o XHTML e il protocollo HTTP 1.1.

Se vengono usati altri protocolli (ad esempio, Z39.50) devono essere disponibili dei gateway per consentire l'accesso tramite un browser web. I progetti dovrebbero garantire la massima disponibilità del proprio sito web. Il programma di finanziamento dovrebbe richiedere una giustificazione di periodi significativi di indisponibilità del sito.

Sui **criteri per assicurare la necessaria qualità del contesto web** nel quale si pubblicano le risorse digitali culturali, si dirà più avanti.

2. il ciclo di vita di un progetto di digitalizzazione
8. *reperimento*

Le collezioni create da un progetto di digitalizzazione fanno parte di un più ampio insieme di risorse. Per agevolare la **localizzazione delle risorse all'interno del corpus**, i progetti devono pensare a **esporre metadati per ciascuna collezione** di risorse da essi sviluppata, in modo tale che essa possa essere utilizzata da altre applicazioni e altri servizi attraverso uno o più dei protocolli o interfacce standard.

L'esatta identificazione di quali metadati debbano essere offerti e del modo in cui esporli dipenderà dalla natura delle risorse create e dalle applicazioni e servizi con i quali tali metadati vanno condivisi.

I progetti dovrebbero esporre una o più registrazioni di metadati a livello della collezione, che cioè descrivano le loro collezioni come un unico insieme. I progetti possono esporre in aggiunta registrazioni di metadati descrittivi delle singole risorse digitali nel contesto delle loro collezioni.

2. il ciclo di vita di un progetto di digitalizzazione
8. *reperimento*

1 La raccolta automatica (harvesting) dei metadati

La raccolta o harvesting dei metadati rende disponibili nel Web i metadati così da consentire la ricerca contestuale tra più depositi (repository) di metadati.

L'harvesting è reso possibile da un protocollo tecnico; lo standard più noto e diffuso è l'OAI-PMH. La Open Archives Initiative ha elaborato un quadro di riferimento che, inizialmente progettato per gli archivi aperti di tipo *eprints*, si è poi affermato come modello di riferimento per l'architettura della biblioteca digitale.

I componenti del modello OAI sono:

- i **data provider** che gestiscono i digital repository di metadati e dati
- i **service provider** che effettuano la cattura dei metadati
- il **Protocol for Metadata Harvesting** (OAI-PMH)

2. il ciclo di vita di un progetto di digitalizzazione
8. *reperimento*

1 La raccolta automatica (harvesting) dei metadati

Il protocollo OAI-PMH è stato adottato per l'harvesting dei metadati e l'interoperabilità di depositi (repository) tanto dal progetto internazionale MICHAEL (Multilingual Inventory of Cultural Heritage in Europe) che dal progetto nazionale di Portale della cultura italiana (PICO).

I progetti possono voler rendere i propri metadati disponibili per l'harvesting costituendo depositi di metadati conformi allo standard OAI. I progetti che istituiscono tali depositi dovrebbero valutare l'opportunità di includervi una dichiarazione dei diritti detenuti sui propri metadati per garantirsi la titolarità dei diritti di proprietà su di essi.

<http://www.openarchives.org/OAI/openarchivesprotocol.html>

2. il ciclo di vita di un progetto di digitalizzazione
8. *reperimento*

2 La ricerca in architettura distribuita

I progetti possono aver bisogno di conoscere lo standard Z39.50, un protocollo di rete che consente la ricerca su database eterogenei, generalmente remoti, e il recupero dei dati tramite un'unica interfaccia utente.

Z39.50 è prevalentemente impiegato per ricercare schede bibliografiche, ma esistono anche implementazioni non bibliografiche. I progetti che adottano il protocollo Z39.50 devono essere a conoscenza del Profilo di Bath (*Bath Profile*) in rapporto all'interoperabilità fra i diversi settori. I progetti possono anche aver bisogno di conoscere il protocollo dei servizi web di ricerca e recupero (SRW/SRU: Search/Retrieve Web Service) che si basa sulla semantica di Z39.50 per offrire funzionalità simili impiegando tecnologie Web Service.

<http://www.loc.gov/z3950/agency/>

2. il ciclo di vita di un progetto di digitalizzazione
8. *reperimento*

3 Alerting

Può essere utile ai progetti conoscere la famiglia di specifiche per i **web feed** RSS (acronimo diversamente sciolto in Really Simple Syndication, Rich Site Summary, RDF Site Summary). I web feed possono essere adottati per trasmettere automaticamente aggiornamenti e nuovi contenuti confezionati in un formato standard ad altri siti web o a utenti che abbiano richiesto al sito il servizio di alerting. RSS fornisce un meccanismo semplice basato su XML per la condivisione dei metadati descrittivi, generalmente sotto forma di una lista di elementi, ciascuno dei quali contenente una breve descrizione testuale con un collegamento (link) alla fonte dei dati.

Un'altra famiglia di standard per i web feed basati su XML è quella di Atom (Atom Syndication Foprmat, Atom Publishing Protocol). L'Atom Syndication Format è stato adottato dalla Internet Engineering Task Force.

2. il ciclo di vita di un progetto di digitalizzazione
8. *reperimento*

4 RDF e ontologie web

I progetti possono far uso di ontologie per il Web.

Un'ontologia è una definizione esplicita, condivisa e socialmente accettata di una porzione della realtà (dominio di conoscenza) per mezzo di un modello concettuale. Questo modello può essere immerso in un software o sistema informativo, in un oggetto o in un processo. Ciò permette che il modello sia recepito e utilizzato da un'ampia gamma di utenti potenziali, siano essi persone, organizzazioni o agenti software.

RDF è un linguaggio di ontologia rilasciato come raccomandazione dal World Wide Web Consortium e usato per rappresentare l'informazione e scambiare la conoscenza in sul Web.

Abbiamo parlato di questi temi trattando di web semantico.

2. il ciclo di vita di un progetto di digitalizzazione
9. *riuso*

In molti casi, gli utenti vorranno riorganizzare e riutilizzare materiali sviluppati dai progetti di digitalizzazione, impiegandoli per altri scopi. L'adozione di standard è importante per facilitare tale riuso.

I progetti dovrebbero tener conto del potenziale riuso delle risorse da essi create e prendere atto del fatto che gli utenti finali o terzi possano voler estrarre degli elementi da una data risorsa e riconfezionarli insieme a parti di altre risorse provenienti da proprie collezioni o da altre fonti.

La possibilità da parte degli utenti di riutilizzare le risorse prodotte dai progetti dovrà essere resa esplicita (vedi oltre, sulla gestione dei diritti).

2. il ciclo di vita di un progetto di digitalizzazione
9. *riuso*

L'esigenza di riuso è particolarmente frequente nel settore della formazione. Sono in corso numerose iniziative che hanno lo scopo di creare strumenti per la gestione di risorse digitali educative, attraverso l'impiego di metadati specifici per le risorse destinate all'apprendimento.

Lo schema **Dublin Core** può essere usato per descrivere le risorse per la didattica, per le quali è stato creato un Gruppo di lavoro (DC-Education) dedicato all'adattamento dello standard alle specifiche esigenze della comunità e-learning.

L'IEEE ha creato uno standard di **Learning Object Metadata (LOM)**, con l'insieme minimo di attributi necessari per gestire, localizzare e valutare le risorse didattiche. Il progetto più importante basato su IEEE/LOM è Instructional Management Systems (IMS), per lo sviluppo di specifiche per l'interoperabilità fra sistemi tecnologici per l'apprendimento (e-learning).

2. il ciclo di vita di un progetto di digitalizzazione
10. *diritti di proprietà intellettuale*

L'IPR (Intellectual Property Rights) comprende la proprietà letteraria e scientifica (il diritto d'autore e i diritti connessi, o, in ambito anglosassone, il *copyright*) e la proprietà industriale (i brevetti, i marchi, i disegni e modelli industriali) ma sono soprattutto i diritti d'autore e di riproduzione che interessano quando si digitalizzano contenuti culturali.

I progetti devono essere consapevoli dei diritti che entrano in gioco nella creazione, riproduzione, negoziazione, diffusione e "repackaging" dei contenuti digitali culturali, preoccupandosi soprattutto di due questioni:

a) **identificare i diritti** sul materiale che intendono digitalizzare

b) **agevolare il riuso** e come evitare usi **non autorizzati**

2. il ciclo di vita di un progetto di digitalizzazione
10. *diritti di proprietà intellettuale*

I progetti devono rispettare i diritti di proprietà intellettuale e di riproduzione detenuti sui materiali con i quali lavorano, e in particolare:

- i diritti detenuti sui **materiali originali digitalizzati**
- i diritti detenuti sulle **risorse digitali**
- i diritti o le concessioni **assegnate ai service provider** per mettere a disposizione le risorse digitali
- i diritti o le concessioni **assegnati all'utente** delle risorse digitali.

I progetti devono inoltre rispettare **ogni diritto discendente** dai termini e dalle condizioni dei programmi di digitalizzazione nel cui quadro essi operano.

2. il ciclo di vita di un progetto di digitalizzazione
10. *diritti di proprietà intellettuale*

Vediamo alcune circostanze che si verificano frequentemente nell'ambito delle attività di digitalizzazione delle istituzioni pubbliche:

Materiali pubblicati. Gli editori solitamente non sono inclini ad autorizzare la digitalizzazione di materiali su cui ancora detengono i diritti, a meno che essa non risulti per loro di qualche vantaggio. I diritti morali dell'autore (paternità e integrità dell'opera, inedito) invece non hanno scadenza.

Materiali di produzione interna. I diritti di utilizzazione economica su ogni opera prodotta da dipendenti di un'istituzione nell'esercizio delle loro funzioni appartengono all'istituzione. Chi presta la propria opera a titolo volontario è titolare dei diritti di utilizzazione economica sul proprio lavoro, a meno che non li ceda per iscritto.

2. il ciclo di vita di un progetto di digitalizzazione
10. *diritti di proprietà intellettuale*

Opere commissionate da istituzioni. Su tali opere, ad esempio fotografie, sono normalmente assicurati alle istituzioni i diritti di riproduzione, ma tali diritti possono non estendersi alla digitalizzazione, a meno che questo non sia espressamente previsto dagli accordi stipulati.

Doni, lasciti e prestiti. Vi possono essere collegate clausole che condizionano la disponibilità per la digitalizzazione.

Infine, bisogna tener conto anche di **altre normative** quali quelle normative a tutela della **riservatezza** (privacy) o di particolari soggetti (ad esempio, i bambini), le leggi di **tutela del patrimonio culturale**, quelle che definiscono **oscenità e indecenza** o pongono restrizioni alla diffusione di particolari categorie di contenuti (ad es. la propaganda nazista).

2. il ciclo di vita di un progetto di digitalizzazione
10. *diritti di proprietà intellettuale*

Per gestire correttamente i diritti detenuti sulle risorse culturali, i progetti devono anzitutto identificare e registrare quali siano i diritti esistenti sui materiali. Quando necessario, i progetti devono negoziare con i detentori dei diritti per ottenere l'autorizzazione a usare i materiali (**liberatoria di digitalizzazione** della risorsa). Si devono registrare nell'ambito del progetto le liberatorie, le assegnazioni, le licenze d'uso, che specificano la natura e l'ambito del contenuto, i modi in cui può essere utilizzato, l'estensione geografica della licenza, la durata della licenza e, se necessario, il pagamento dovuto. Inoltre, si deve presentare un avviso o dichiarazione di copyright sul sito web, che chiarisca la titolarità dei diritti e i termini di utilizzazione delle risorse cui si dà accesso sul sito stesso.

2. il ciclo di vita di un progetto di digitalizzazione
10. *diritti di proprietà intellettuale*

Una volta identificati i diritti di proprietà intellettuale e negoziate le licenze, i progetti devono assicurare la protezione dei diritti di cui sono titolari e di quelli di terzi, prendendo misure che impediscano l'uso non autorizzato dei materiali.

Una **licenza**, ovvero l'atto con cui il titolare dei diritti di proprietà intellettuale conferisce ad altri il permesso di utilizzare un particolare contenuto per determinati usi ben specificati, **è un contratto a tutti gli effetti giuridici**.

Una o più **licenze-quadro** accuratamente negoziate che coprano tutte le attività on-line di una determinata istituzione culturale, eventualmente rivedibili, possono facilitare notevolmente il lavoro dell'istituzione culturale, riducendo gli oneri di amministrazione delle liberatorie e ampliando al tempo stesso il mercato per le associazioni di titolari di diritti.

2. il ciclo di vita di un progetto di digitalizzazione
10. *diritti di proprietà intellettuale*

L'espressione dei **Termini di disponibilità** o delle **Regole commerciali** per le risorse è demandata ai metadati per la gestione dei diritti. Per chiarezza sia degli utenti sia dei fornitori di servizi è raccomandato distinguere i metadati relativi ai diritti di proprietà intellettuale dalla risorsa.

Per esprimere i termini delle licenze senza ambiguità negli schemi di metadati amministrativi si possono usare gli standard REL (Rights Expression Language). Lo schema ODRL (Open Digital Rights Language) è un'applicazione open source di REL in XML; XrML è invece uno standard proprietario.

I progetti dovrebbero mantenere dati sui diritti che detengono o acquisiscono in modo coerente, così da poterli condividere con sistemi esterni in un formato standard.

2. il ciclo di vita di un progetto di digitalizzazione
10. *diritti di proprietà intellettuale*

Accesso aperto – Open access

Movimento internazionale, in ambito di letteratura scientifica, per l'accesso alla produzione scientifica. Alla sua base la **Dichiarazione di Berlino** (20 - 22 Oct 2003)

vers.it. <http://www.zim.mpg.de/openaccess-berlin/ BerlinDeclaration_it.pdf>:

- l'informazione deve essere resa largamente e prontamente disponibile alla società

- si devono sostenere nuove possibilità di disseminazione della conoscenza attraverso **l'accesso aperto**: una fonte estesa del sapere umano e del patrimonio culturale che siano stati validati dalla comunità scientifica

- il Web del futuro dovrà essere sostenibile, interattivo e trasparente

- i contenuti e i mezzi di fruizione compatibili e ad accesso aperto

2. il ciclo di vita di un progetto di digitalizzazione
10. *diritti di proprietà intellettuale*

Creative Commons

Si collega al modello Open Access e all'Open Use l'iniziativa *Creative Commons*, nata nel 2001, che ha rilasciato un insieme di **sei licenze standard** a livello internazionale che offrono uno spettro di iniziative intermedie fra un regime full copyright (tutti i diritti riservati) e il dominio pubblico. Licenze e condizioni d'uso sono esplicitate in RDF (Resource Description Framework).

Le licenze Creative Commons combinano fra loro **quattro diversi elementi base**: Attribution (Attribuzione), Non-Commercial (Non commerciale), Non-derivative (Non opere derivate), Share-alike (Condividi allo stesso modo).

Ogni licenza è disponibile in tre versioni: per i legali; per gli utenti; in formato leggibile dalla macchina.

2. il ciclo di vita di un progetto di digitalizzazione
10. *diritti di proprietà intellettuale*

Creative Commons

<http://www.creativecommons.it/>

Le licenze Creative Commons **consentono**:

- la visualizzazione pubblica; la distribuzione; la copia; l'integrazione dell'opera in una o più opere collettive, obbligando sempre e comunque a citare la fonte del materiale riutilizzato.

Fra le altre possibili **restrizioni e condizioni**:

- l'utilizzazione per scopi non commerciali; la limitazione del riuso per opere derivate.

Prima di poter accedere a un qualsiasi contenuto, l'utente dovrà esplicitamente accettare i termini e le condizioni d'uso. Le licenze Creative Commons sono state utilizzate dall'OAI Open Archives Initiative e dall'iniziativa Dublin Core.

2. il ciclo di vita di un progetto di digitalizzazione
10. *diritti di proprietà intellettuale*

E-Commerce

È comune per i programmi di creazione di contenuti culturali digitali del settore pubblico specificare che i contenuti creati devono essere messi gratuitamente a disposizione degli utenti ai punti di accesso, almeno a fini educativi.

In alcuni casi però i programmi incoraggiano o espressamente richiedono ai progetti di produrre reddito con i materiali creati.

Marcatura e impronte digitali

I progetti possono considerare la possibilità di apporre una filigrana o un'impronta digitale sui materiali che producono allo scopo di salvaguardare i diritti su di essi detenuti.

La filigrana digitale, o marchio elettronico (watermarking), è un segnale (marchio) incorporato permanentemente nei dati digitali. È essenziale che l'utenza sia informata dell'esistenza del watermark sulle immagini distribuite, come *warning* agli usi non consentiti.