

## Metadati per l'e-Learning: una introduzione

dispensa a cura di Gino Roncaglia

per il Master Universitario di I livello in e-learning – Università degli studi della Tuscia

© Gino Roncaglia, 2008

Questo contenuto è reso disponibile con licenza [ccLearn](#) Attribution-NonCommercial-ShareAlike nell'ambito del progetto OpenCourseWare promosso dal Master in e-learning dell'Università della Tuscia.

### 1. Cosa sono i metadati

È arrivato il momento di parlare un po' più a fondo di metadati: un concetto che ha un ruolo importantissimo non solo nel campo dell'e-Learning, ma per la gestione di qualunque tipo di contenuto informativo o documentale, soprattutto quando l'informazione disponibile è molta e deve essere dunque selezionata e organizzata per facilitarne il reperimento e l'uso. Non a caso, nonostante si tratti di un concetto perfettamente applicabile in tante situazioni che non hanno nulla a che fare con il mondo della rete e delle nuove tecnologie, i metadati sono uno dei 'mattoni' fondamentali del nuovo web. E, come vedremo, hanno un ruolo centrale anche nell'organizzazione di attività di e-Learning e nella creazione e distribuzione di materiali di apprendimento.

Ma cosa sono i metadati? Etimologicamente, il termine richiama l'idea di 'dati di secondo livello', dati utilizzati per descrivere e classificare altri dati. L'esempio tipico, e quello più comunemente utilizzato per spiegare il concetto, è il catalogo di una biblioteca. Gli scaffali di una biblioteca raccolgono libri (i nostri dati di primo livello, o 'informazione primaria'), mentre il catalogo raccoglie schede che *descrivono* i libri, sulla base di una serie di caratteristiche prefissate: il nome dell'autore, il titolo, il luogo e l'anno di pubblicazione, il numero di pagine, la posizione negli scaffali della biblioteca, e così via. Le schede di un catalogo contengono dunque dei dati di secondo livello, dei *metadati* appunto, che descrivono l'informazione primaria e aiutano a gestirla e reperirla. Dal canto suo, la biblioteca comprende *sia* i libri *sia* il catalogo: una buona raccolta di risorse informative, infatti, comprende sempre anche informazione secondaria, metadati, che permettono di descrivere, organizzare e reperire l'informazione primaria (una buona biblioteca, peraltro, offre al proprio utente anche molte altre cose, e in particolare un insieme assai ampio di strumenti e servizi che aiutano a lavorare sia sull'informazione primaria sia sui metadati).

Se proviamo a pensarci, ci accorgiamo che situazioni simili – in cui utilizziamo dei metadati per identificare e descrivere alcune caratteristiche del mondo informativo che ci circonda, in modo da muoverci al suo interno nel modo più efficace – sono molto comuni: decidiamo quale programma guardare in televisione o

## Metadati per l'e-Learning: una introduzione

---

in quale cinema andare la sera utilizzando le guide alla programmazione pubblicate dal giornale (e cosa sono queste guide se non raccolte organizzate di metadati relativi ai programmi televisivi o cinematografici?), ordiniamo un pasto al ristorante dopo aver valutato sul menu i piatti disponibili e il loro prezzo, facciamo la spesa al supermercato consultando le etichette dei prodotti per verificarne gli ingredienti, il prezzo o la data di scadenza, decidiamo se comprare un libro dando un'occhiata all'indice e alla descrizione in quarta di copertina, e così via.

Tutto molto semplice, sembrerebbe. Ma in realtà anche i metadati nascondono le loro insidie (o il loro fascino, se amate l'avventura intellettuale!). Ad esempio: dove finiscono i dati e dove cominciano i metadati? Il titolo di un libro, stampato a chiare lettere sulla sua copertina (e ripreso nella scheda della nostra ipotetica biblioteca) è un dato (dopotutto è parte del testo del libro...) o un metadato? Il mio nome e il mio cognome servono a individuarmi... sono forse metadati anch'essi? E se lo sono, cosa descrivono? Il nostro linguaggio non è forse costituito in primo luogo di nomi utilizzati per riferirci a oggetti o concetti? Ma allora *tutto* il linguaggio è fatto di metadati?

Insomma, come vedete è facilissimo scivolare dai metadati alla filosofia. E non è certo un caso che, come vedremo, nel parlare di metadati intervenga spesso un termine nato proprio in ambito filosofico: il termine *ontologia*. Ma procediamo con ordine, cercando – a rischio di qualche semplificazione e di qualche ingenuità – di chiarirci un po' le idee.

Partiamo da quello che è probabilmente l'aspetto fondamentale dei metadati: abbiamo a che fare con metadati quando, volendo descrivere o gestire una risorsa informativa o un certo insieme di risorse informative, ne individuiamo e selezioniamo alcune caratteristiche al fine di costruirne un modello semplificato, funzionale ai nostri scopi e facilmente condivisibile con altri. Questo avviene, di norma, individuando all'interno dell'universo informativo di partenza (l'informazione primaria) proprietà, tratti comuni e differenze che ci aiutino a organizzarlo. Così, davanti ai libri della mia libreria, potrei scegliere come caratteristiche rilevanti per descriverli e gestirli l'autore e il titolo di ciascuno di essi, ma anche, ad esempio, la loro dimensione, il colore della copertina, il tipo di carattere tipografico utilizzato, il numero di pagine, l'anno di pubblicazione, e ancora la data in cui l'ho comprato, se l'ho prestato o no a qualcuno, se il libro ha delle sottolineature, e così via. E dovendo organizzare il sito web di una scuola o di una università, potrei organizzare l'offerta di corsi e di attività didattiche (curricolari o integrative) per materia, per docenti, per orario, per fasce di età dei destinatari...

Ciascuna di queste caratteristiche – e potrei immaginarne un'infinità di altre – corrisponde a un possibile tipo di metadati, e ciascuna potrebbe essere utile in certi contesti. Selezionare un certo insieme di metadati (e scartarne molti altri possibili) dipende sempre dai nostri scopi, e implica la scelta di un particolare modello, di una particolare rappresentazione dell'informazione primaria. In un certo senso, equivale a scegliere una particolare 'visione del mondo', definendo le categorie (e le relazioni) in base alle quali vogliamo organizzare l'universo informativo che ci interessa. Per questo, come si è accennato, a proposito di metadati si parla spesso di *ontologie*: un'ontologia ci dice 'cosa è reale' per noi, quali sono le caratteristiche che scegliamo di prendere in considerazione nel nostro dominio di riferimento.

## Metadati per l'e-Learning: una introduzione

Così, se il nostro scopo è organizzare i libri della libreria in base al contenuto informativo del testo che contengono, il colore della copertina ci interesserà relativamente poco, mentre ci interesseranno di più il titolo, l'autore, gli argomenti trattati. Se prestiamo molti libri (e magari alcuni non tornano indietro) diventa rilevante tener traccia delle persone alle quali li abbiamo prestati; e se la libreria non è la nostra ma quella, poniamo, di un grande scrittore, può essere importante sapere se e quali libri hanno delle sottolineature o delle note manoscritte, che potrebbe essere interessante studiare. In ciascuno di questi casi, selezioniamo caratteristiche per noi rilevanti all'interno dell'universo informativo di riferimento, e le utilizziamo per organizzarlo in base a uno fra i tanti modelli possibili<sup>1</sup>.



Figura 1 - Metadati: esempio di metadati relativi a una curiosa scultura cinese (da <http://www.lib.ku.edu/imagegateway/help/workspace.cfm>)

Abbiamo detto che esistono molti tipi possibili di metadati. Non stupirà dunque che ne siano state tentate diverse classificazioni (il che equivale un po' a individuare e selezionare dei meta-metadati...), anche con l'obiettivo di aiutare a strutturare l'insieme di metadati che utilizziamo, distinguendo al suo interno caratteristiche e tipologie diverse.

Una distinzione abbastanza diffusa è quella fra metadati *descrittivi*, *gestionali-amministrativi* e *strutturali*: come suggeriscono i nomi, i metadati descrittivi (ad esempio, il titolo di un corso o un insieme di parole chiave che ne caratterizzano il contenuto) hanno lo scopo fondamentale di descrivere l'informazione

<sup>1</sup> Il gran numero di questi criteri permette addirittura di dar vita a veri e propri paradossi logici, come la curiosa versione del paradosso di Russell denominata 'paradosso del bibliotecario'. I lettori che volessero sapere qualcosa di più su questo tema (di cui si è occupato anche Umberto Eco nel suo *De Bibliotheca*, consultabile in rete all'indirizzo [http://www.liberliber.it/biblioteca/e/eco/de\\_bibliotheca/html/testo.htm](http://www.liberliber.it/biblioteca/e/eco/de_bibliotheca/html/testo.htm)) possono leggere la voce che ad esso dedica la versione italiana di Wikipedia ([http://it.wikipedia.org/wiki/Paradosso\\_del\\_bibliotecario](http://it.wikipedia.org/wiki/Paradosso_del_bibliotecario), consultata nella stesura del 5 dicembre 2007).

## Metadati per l'e-Learning: una introduzione

---

primaria; quelli amministrativo-gestionali (ad esempio, la scadenza per il pagamento della quota di iscrizione) hanno lo scopo principale di semplificarne la gestione, e quelli strutturali (ad esempio, la suddivisione del corso in unità didattiche) aiutano a strutturarla. Ma questa classificazione non è né l'unica possibile, né necessariamente esaustiva o sempre soddisfacente: proprio come la scelta dei metadati stessi, la scelta di particolari classificazioni o organizzazioni dei metadati è legata al tipo di informazione primaria di cui ci stiamo occupando, ai nostri obiettivi, alla nostra 'visione del mondo'.

Così, ad esempio, in certi contesti possono avere rilievo (mentre in altri contesti può essere preferibile evitarli) metadati *valutativi*, come le 'stelle' attribuite ad un film dalla recensione di un giornale. E in effetti fra le varie distinzioni proposte in quest'ambito potrà capitarvi di incontrare proprio quella fra metadati *oggettivi* (ricavabili direttamente dall'analisi dell'informazione primaria, in maniera tendenzialmente indipendente dalle opinioni o valutazioni di chi classifica) e metadati *soggettivi* (come appunto i metadati valutativi, in cui l'elemento di valutazione soggettiva è preponderante). Così come vi potrà capitare di trovare la distinzione, relativa alla responsabilità della metadattazione, fra metadati *primari* (forniti dall'autore della risorsa informativa), *secondari* (forniti dagli utenti) e *terziari* (forniti da servizi o organizzazioni indipendenti)<sup>2</sup>. O ancora la distinzione fra metadati inclusi nella risorsa informativa (*embedded metadata*) e metadati esterni alla risorsa informativa (*external metadata*). Insomma, così come nel descrivere, strutturare o gestire una risorsa informativa possiamo scegliere metadati diversi, anche nel guardare ai metadati (che sono, ricordiamolo sempre, a loro volta informazione, dunque anch'essi dei dati) possiamo adottare prospettive diverse, e proporre diverse distinzioni e classificazioni.

Un altro aspetto su cui richiamare l'attenzione è la *natura testuale o numerica* della maggior parte dei metadati. Anche quando l'informazione primaria che intendiamo descrivere e gestire non è testuale, ad esempio nel caso di brani musicali o di fotografie, i metadati che utilizziamo per descriverla e gestirla sono quasi sempre testuali o numerici (autore del brano, titolo, esecutore, ecc. nel caso della musica; autore della fotografia, luogo, macchina fotografica utilizzata, valore ISO ecc. nel caso di fotografie). Ne risulta, in questi casi, una sorta di 'salto ontologico' fra l'informazione primaria e i suoi metadati: utilizziamo in sostanza il linguaggio naturale (o delle sue porzioni rigorose e formalizzate) per 'parlare' di oggetti che sono però per loro natura qualcosa d'altro, e dei quali il campo del linguaggio verbale non può esaurire la complessità.

In alcuni di questi casi, si è cercato di sperimentare anche l'uso di metadati non testuali: un brano musicale potrebbe così essere descritto anche attraverso una rappresentazione semplificata della sua linea melodica, fra i descrittori di una fotografia potrebbe comparire l'istogramma che rappresenta la gamma dinamica dei colori utilizzati e la loro luminosità, oppure uno schizzo che ne rappresenta una versione semplificata. In un certo senso, anche le cosiddette immagini di anteprima, o 'thumbnail', che in molti programmi, sistemi operativi e siti di archiviazione offrono una anteprima in scala ridotta dell'immagine originale, sono un tipo di metadati non testuali. Ma utilizzare metadati non testuali e non numerici (non esprimibili dunque in forma alfanumerica) per effettuare ricerche e classificazioni non è banale, e abbiamo probabilmente appena cominciato a esplorare le caratteristiche di questo problema e le sue possibili soluzioni.

---

<sup>2</sup> Su queste distinzioni e il loro ruolo nell'ambito dell'e-Learning si veda A. Fini e L. Vanni, *Learning object e metadati. Quando, come e perché avvalersene*, Erickson, Trento 2004, pp. 63-65.

## Metadati per l'e-Learning: una introduzione

---

Metadati (e meta-metadati) diversi per scopi diversi, dunque: lo vedremo ancor meglio più avanti, quando dovremo discutere della organizzazione dei sistemi di metadati utilizzati per classificare contenuti e risorse di apprendimento.

Prima di procedere, però, è importante a questo punto approfondire un po' la questione, ovviamente essenziale, di cosa si intenda per metadati, soffermandoci su un aspetto di particolare importanza. I metadati non nascono, come Minerva, perfettamente compiuti e indipendenti dalla mente di Giove: sono invece il risultato di un lavoro di elaborazione, che richiede siano presi in considerazione tre elementi, tutti e tre indispensabili: 1) la risorsa informativa che rappresenta la nostra informazione primaria (ad esempio, una copia della Divina Commedia); 2) le proprietà, categorie o relazioni attraverso cui abbiamo deciso, per i nostri scopi specifici e in accordo con la nostra 'visione del mondo', di organizzare l'universo informativo al quale quella risorsa appartiene (ad esempio 'autore', 'titolo', ecc.) e 3) i valori che a tali proprietà vengono attribuiti nel caso della particolare risorsa informativa che stiamo considerando (ad esempio, la proprietà 'autore' ha, nel caso di quel libro, il valore 'Dante Alighieri').

Volendo adottare un linguaggio appena un po' più preciso (ma senza inoltrarci in un formalismo veramente rigoroso, che richiederebbe ulteriori e ben più complesse precisazioni), possiamo dire che l'assegnazione di un metadato a una risorsa informativa corrisponde a una *tripla*  $(a_i, p_j, v_k)$  il cui primo elemento è appunto la risorsa informativa, cioè un oggetto del dominio composto dalle nostre risorse informative primarie  $(a_i \in \mathcal{D})$ , dove  $\mathcal{D} = \{a_1, a_2, a_3 \dots\}$ , e  $a_1, a_2 \dots$  sono le risorse informative che ci interessano, ad esempio i libri di una biblioteca); il cui secondo elemento è una particolare proprietà tratta dall'insieme di proprietà utilizzate per descrivere quel tipo di risorse primarie  $(p_j \in \mathcal{P})$ , dove  $\mathcal{P} = \{p_1, p_2, p_3 \dots p_n\}$  e  $p_1 \dots p_n$  sono le proprietà che ci interessano, ad esempio 'autore', 'titolo' ecc.); e il cui terzo elemento è un valore, scelto fra quelli ammissibili per la proprietà che stiamo considerando  $(v_k \in \mathcal{V}^{p_j})$ , dove  $\mathcal{V}^{p_j} = \{v_1, v_2, v_3 \dots\}$  e  $v_1, v_2 \dots$  sono i valori ammissibili per la proprietà  $p_j$ ; così, se ad esempio  $p_j$  è la proprietà 'autore',  $v_1, v_2 \dots$  potranno essere 'Dante Alighieri', 'Alessandro Manzoni' ecc.).

Abbiamo detto che questa descrizione, pur costituendo una approssimazione accettabile per i nostri scopi<sup>3</sup>, è in realtà ancora tutt'altro che rigorosa. In effetti, per renderla tale dovremmo considerare il fatto che oltre alle proprietà di una singola risorsa informativa ci può interessare anche prendere in considerazione *relazioni* fra risorse informative diverse; dovremmo inoltre disporre di strumenti per esprimere in maniera chiara quali sono i vincoli posti sui valori che una certa proprietà può assumere; potremmo poi voler definire all'interno del dominio di riferimento dei sottoinsiemi che ammettono proprietà diverse, fare scelte migliori per quel che riguarda il formalismo utilizzato (ad esempio distinguendone esplicitamente gli aspetti sintattici e gli aspetti semantici), introdurre regole e magari addirittura assiomatizzazioni che permettano di operare deduzioni (ad esempio, permettano di dedurre l'assegnazione di particolari valori a determinate proprietà in conseguenza di assegnazioni o scelte fatte in precedenza), e così via. La definizione rigorosa di tutte queste caratteristiche 'astratte' del nostro sistema di dati e di metadati, in

---

<sup>3</sup> I lettori che hanno qualche familiarità con i formalismi del web semantico avranno notato che la nostra descrizione si avvicina molto al modo in cui il data model RDF (Resource Description Framework) considera gli *statements*, che corrispondono appunto a triple composte, fondamentalmente, da risorse, proprietà e valori. Non a caso, RDF è uno degli strumenti essenziali per esprimere in maniera formalmente rigorosa e per condividere metadati strutturati

particolare per quanto riguarda il livello semantico, corrisponde alla individuazione di un'*ontologia*: una specificazione formale, esplicita e condivisibile con altri (siano essi persone o agenti software), di quali siano le tipologie di oggetti che compongono il nostro dominio di riferimento, di quali siano le proprietà e relazioni che possono essere stabilite fra di essi, e di quali siano le regole che ne governano il funzionamento.

Torneremo fra breve su alcuni di questi problemi: per concludere questa discussione introduttiva, e alla luce delle considerazioni fin qui esposte, limitiamoci per ora a dire, in maniera abbastanza generale, *che i metadati sono (meta)informazione, di norma altamente strutturata, utilizzata per descrivere, strutturare o gestire una risorsa informativa o un insieme di risorse informative (la nostra informazione primaria) attraverso l'identificazione di alcune sue proprietà e l'assegnazione ad esse di specifici valori.*

## 2. Interoperabilità, standardizzazione

L'insieme di metadati che si sceglie di utilizzare per descrivere un certo insieme di risorse informative (ad esempio i libri di una biblioteca, o i contenuti di un corso) corrisponde, come si è detto, a delle scelte consapevoli, che portano comunque a privilegiare alcune caratteristiche e proprietà dell'informazione primaria a scapito di altre. Questo spiega perché, in situazioni diverse, si possa scegliere di utilizzare, con riferimento alla stessa informazione primaria, insiemi diversi di metadati. Ma se organizzazioni che devono gestire in maniera fondamentalmente analoga lo stesso tipo di informazioni primarie facessero scelte troppo diverse relativamente ai metadati utilizzati, diventerebbe molto difficile, se non impossibile, rendere interoperabili i rispettivi sistemi di gestione informativa.

Così, ad esempio, se due musei utilizzassero insiemi di metadati molto diversi per descrivere i quadri esposti, diventerebbe molto difficile realizzare cataloghi collettivi, o gestire in maniera efficace l'informazione connessa al prestito di un quadro da un museo all'altro. E se biblioteche diverse schedassero i propri libri in maniera completamente diversa, il povero utente che ne dovesse consultare i cataloghi, magari per trovare un libro di difficile reperibilità, dovrebbe ogni volta imparare e adottare nuove convenzioni. L'esigenza dell'*interoperabilità* porta dunque a una naturale richiesta di *standardizzazione*: adottare insiemi standardizzati di metadati permette non solo di semplificarsi la vita (e di semplificarla ai propri utenti), ma anche di scambiare informazioni e 'schede', di utilizzare gli stessi programmi per la loro gestione, di realizzare servizi di ricerca integrata, e così via. E in effetti, come è facile capire, queste esigenze diventano ancor più stringenti quando – come accade ormai quasi nella totalità dei casi – riguardano sistemi di gestione dell'informazione basati sull'uso di computer e agenti software.

La standardizzazione, naturalmente, è possibile finché gli scopi alla base della nostra scelta di metadati, e la 'visione del mondo' che adottiamo davanti a un certo tipo di informazione primaria, sono ragionevolmente uniformi. In caso contrario, può essere del tutto ragionevole utilizzare per la stessa informazione primaria sistemi diversi di metadati (e dunque ontologie diverse). Così, ad esempio, un ufficio postale smisterà la corrispondenza sulla base dell'indirizzo e del codice postale scritti sulla busta, senza occuparsi – di norma – del contenuto e del mittente delle lettere, mentre il destinatario della corrispondenza potrà organizzarla nel suo archivio proprio sulla base di questi ultimi elementi. Le due prospettive, quella dell'ufficio postale

che deve smistare la corrispondenza indipendentemente dal suo contenuto informativo e quella del destinatario umano che deve leggerla, interpretarla e gestirla proprio in funzione del suo contenuto, sono così lontane fra loro da rendere difficile l'identificazione di una 'ontologia' comune, se non a un livello molto generale ed astratto (compito, quest'ultimo, che può comunque rivelarsi interessante: non a caso, chi si occupa di ontologie ha un'innata passione per l'astrazione!).

Comunque, in tutti i casi in cui è sensato farlo, è più che ragionevole cercare di utilizzare sistemi di metadati standardizzati, o almeno sistemi nei quali sia standardizzata la gestione dei metadati comuni (una libreria e una biblioteca hanno entrambi bisogno di sistemi di metadati per la gestione dei libri, e questi sistemi potranno coincidere per una larga parte, ma la biblioteca non avrà ad esempio bisogno dei metadati relativi agli sconti applicabili sul prezzo di vendita, mentre la libreria non avrà bisogno dei metadati relativi al prestito dei libri).

La ricerca della standardizzazione è dunque una caratteristica essenziale nella costruzione di sistemi di metadati: non a caso, il lavoro di definizione di uno standard è in genere lungo e complesso, ed è spesso affidato a commissioni di esperti, non di rado con la supervisione di istituzioni sovranazionali. Ma la ricerca della standardizzazione ha un prezzo, per quanto necessario: per cercare di adattare un sistema di metadati al maggior numero possibile di situazioni, si finisce spesso per allargarlo e renderlo più complesso di quanto ciascuna delle singole situazioni richiederebbe. Capita dunque in molti casi e per molti progetti specifici di utilizzare *sottoinsiemi* di sistemi di metadati standard: il riferimento allo standard garantisce l'interoperabilità (anche se solo rispetto ai metadati effettivamente utilizzati), senza sacrificare eccessivamente la semplicità delle applicazioni.

### 3. Organizzare i metadati: dalla strutturazione al rigore formale

Fin qui, abbiamo parlato di metadati immaginandoli un po' come una successione lineare di descrittori. Ma è importante ricordare che, nella maggior parte dei casi, i sistemi di metadati non sono 'piatti' ma sono invece fortemente strutturati. Per capire questo concetto, torniamo per un istante all'esempio di un libro e a uno dei suoi metadati più tradizionali, l'autore. Chi è l'autore di un libro? La risposta è a prima vista assai semplice: chi lo ha scritto. Ma... e se si tratta ad esempio di un libro illustrato? Chi è l'autore dei primi fumetti di Asterix: l'autore dei testi, Goscinny, o il disegnatore, Uderzo? Tutti e due, risponderemmo probabilmente, ma con ruoli diversi. E l'autore della prefazione di un libro, o il responsabile di una edizione critica, o il curatore di una antologia, che ruolo hanno, e come devono figurare nella nostra ipotetica 'scheda' dell'opera? Si tratta evidentemente di figure diverse con ruoli diversi, ma di figure che hanno tutte una parte nella costruzione dei contenuti del libro. Può essere utile, allora, prevedere dei metadati più generali (ad esempio 'creator' e 'contributor'), articolandoli poi in ruoli più specifici: autore principale, illustratore, prefatore, curatore, e così via. In una parola: i sistemi di metadati sono normalmente *strutturati*, non consistono in una serie piatta di descrittori ma in gerarchie organizzate.

Inoltre, anche se molti metadati riguardano quelle che sembrano essere intuitivamente delle *proprietà* di un oggetto (così il titolo o l'autore possono sembrare intuitivamente delle 'proprietà' di un libro), abbiamo già avuto occasione di accennare al fatto che molti metadati hanno una natura *relazionale*. Così, ad

## Metadati per l'e-Learning: una introduzione

---

esempio, fra le informazioni relative a un libro possiamo trovare “traduzione inglese”. Essere una traduzione può sembrare a prima vista una proprietà del libro (e in un certo senso lo è), ma... quel libro è una traduzione *di che cosa*? In effetti, per essere completa la nostra indicazione deve fare riferimento a un'altra risorsa informativa: l'informazione ha in questo caso anche una natura relazionale, ci dice che un certo libro è una traduzione di un altro libro, e nel farlo stabilisce una relazione di un certo tipo fra due oggetti del nostro dominio.

Insomma, i sistemi di metadati hanno di norma una natura complessa, altamente strutturata, spesso gerarchica, e comprendono non solo metadati diversi ma, come abbiamo visto, anche metadati *di tipo* diverso o con *funzioni* diverse. Una considerazione che del resto non dovrebbe sorprenderci troppo, dato che, per quanto semplificati, i sistemi di metadati si propongono di offrire modelli per la gestione di una informazione di primo livello che è a sua volta in genere complessa, variamente articolata e strutturata.

Nelle intenzioni, si è detto, i sistemi di metadati dovrebbero aiutarci a gestire l'informazione primaria in maniera funzionale e standardizzata. Ma... funzionale *per chi*? Ricordiamo di nuovo che sempre più spesso a gestire questa informazione non sono esseri umani ma programmi (quelli che tecnicamente si chiamano 'agenti software'). Ai programmi non basta la standardizzazione, serve anche una estrema attenzione al rigore formale. Ad esempio, l'anno di pubblicazione e il numero di pagine di un libro sono tutti e due espressi attraverso numeri, ma si tratta di numeri di tipo ben diverso! Per capirlo, basta pensare che in certe circostanze può aver senso sommare dei numeri di pagina (ad esempio per avere la paginazione complessiva di un libro in più tomi), mentre non ha senso sommare due anni di pubblicazione. L'anno di pubblicazione è una *data*, e le date vanno gestite in modo da permettere di riconoscerle e tenere conto della loro particolare natura. Un essere umano riconosce in genere a colpo d'occhio una data e capisce che si tratta di un dato numerico di tipo molto particolare, mentre il computer lo sa fare solo se abbiamo *esplicitato* le differenze esistenti fra i due tipi di dati.

Una buona formalizzazione di un sistema di metadati richiede che sia prestata specifica attenzione a tutti questi aspetti. Innanzitutto, dovrà essere esplicitato il dominio di riferimento: a quali classi di dati si applicano i nostri metadati? Definire il dominio di riferimento può sembrare banale, ma spesso non lo è affatto. Se vogliamo descrivere dei libri, dovremo capire (o stabilire attraverso qualche convenzione) *cosa* esattamente si intenda per libro; se vogliamo descrivere dei contenuti per la didattica a distanza, dovremo capire *quali tipi di contenuti* ci interessa considerare e classificare. Una operazione che spesso ci porta a individuare non solo una classe di dati, ma anche la presenza al suo interno di sottoclassi rilevanti ai fini della nostra classificazione. Ad esempio, se vogliamo un sistema di metadati che si applichi sia ai libri sia alle riviste, ma che distingua queste due classi di oggetti relativamente ad alcune proprietà, potremmo voler distinguere all'interno del nostro dominio di riferimento due sottoclassi specifiche.

Dovremo poi specificare quali proprietà vogliamo considerare, e a quali classi o sottoclassi di oggetti queste proprietà si possono applicare. Molto spesso, fra proprietà diverse esistono relazioni di dipendenza (ad esempio, se un libro ha la proprietà di essere diviso in più volumi, ciascuno di questi volumi avrà a sua volta la proprietà di essere parte del libro in questione): se vogliamo che il sistema sia capace di gestirle, magari ricavando autonomamente alcune delle conclusioni deducibili a partire dalla (meta)informazione già in nostro possesso, anche queste relazioni dovranno essere riconosciute ed espresse formalmente. Dovremo

specificare quali tipi di valore possano assumere le proprietà che consideriamo (l'anno di pubblicazione sarà ad esempio una data, ma potrebbe anche essere espresso attraverso un punto interrogativo o attraverso un intervallo di date, se non siamo sicuri di quella giusta...). Se i nostri metadati sono strutturati in forma gerarchica, dovremmo esprimere formalmente anche le relative dipendenze. Insomma, la costruzione di un buon sistema di metadati implica, come si era già accennato, la creazione di una vera e propria *ontologia*, descritta con il maggiore rigore formale possibile. E sono stati elaborati linguaggi specifici che hanno proprio lo scopo di descrivere tali ontologie in modo rigoroso e formale (il più noto è un linguaggio che si chiama OWL, Web Ontology Language; ma non preoccupatevi, non ce ne occuperemo in questa sede!).

In generale, comunque, i sistemi di metadati legati al mondo della rete hanno ormai trovato nel (meta)linguaggio di marcatura XML (eXtensible Mark-up Language) la loro 'lingua franca'. XML, infatti, offre uno strumento estremamente potente per includere metadati in un oggetto informativo, distinguendoli in maniera esplicita dall'informazione primaria nonché dichiarandoli e organizzandoli in forma rigorosa. A XML è dedicato, nell'ambito del master, un approfondimento specifico, che dovrebbe aiutarvi a capire come mai questo linguaggio sia così efficace nel permettere di definire metadati e di utilizzarli.

#### 4. Ontologie rigorose o classificazioni libere?

Tutto questo lavoro può certo intimidire. Pensavamo di dover descrivere la nostra informazione primaria (ad esempio i materiali di un corso a distanza) usando un po' di parole chiave e di descrittori intuitivi, e ci troviamo impegnati in un lavoro di costruzione di complicati sistemi formali, che useremo poi magari per un decimo delle loro potenzialità ma che renderanno comunque il nostro lavoro di descrizione e classificazione assai più complesso e faticoso. È davvero necessaria tutta questa fatica?

Per un verso, la risposta è sicuramente affermativa. Anche se complesso e faticoso, il lavoro di costruzione di sistemi di metadati 'solidi', condivisi e ben strutturati garantisce la compatibilità fra le nostre descrizioni e quelle date da altri, permette la gestione automatica della nostra informazione primaria (e degli stessi metadati), risponde insomma all'esigenza già richiamata di *interoperabilità* dei sistemi di gestione informativa permettendoci in molti casi di automatizzare il loro funzionamento. Inoltre la discussione che accompagna la definizione di uno standard, e la sua sperimentazione in molte situazioni diverse, aiutano a migliorare e a *validare* il nostro sistema di metadati, e spesso anche a capire meglio la natura della nostra informazione primaria: un risultato importante, giacché – come abbiamo ricordato – un sistema di metadati rappresenta sempre una particolare 'visione del mondo' (anche se limitata alla specifica classe di dati che ci interessa descrivere), con tutti i rischi di arbitrarietà che questo può comportare.

D'altro canto, la descrizione rigorosa dell'informazione ha un serio problema: è difficile! E la conseguenza di scelte sbagliate, o di applicazioni sbagliate di sistemi di classificazione magari validissimi, può essere in questo caso assai pesante, in termini di costi e di risultati.

Negli ultimi due o tre anni, si è parlato molto dell'uso di un paradigma del tutto diverso nel campo della descrizione dei contenuti informativi: la descrizione 'dal basso', da parte degli utenti, attraverso 'etichette'

('tag') libere. È la strada del cosiddetto 'social tagging', adottato da molti sistemi di gestione dell'informazione 'user generated' presenti in rete: meccanismi di social tagging sono presenti ad esempio in molti blog, in sistemi di gestione di immagini come Flickr o di video come YouTube, e in sistemi di descrizione condivisa di risorse come del.icio.us. In questi casi gli utenti descrivono i loro contenuti (e quelli degli altri) usando le etichette che preferiscono. Il sistema può poi usare meccanismi di filtraggio collaborativo ('collaborative filtering') per fare emergere le etichette e le classificazioni più usate o considerate più utili da parte degli utenti stessi.

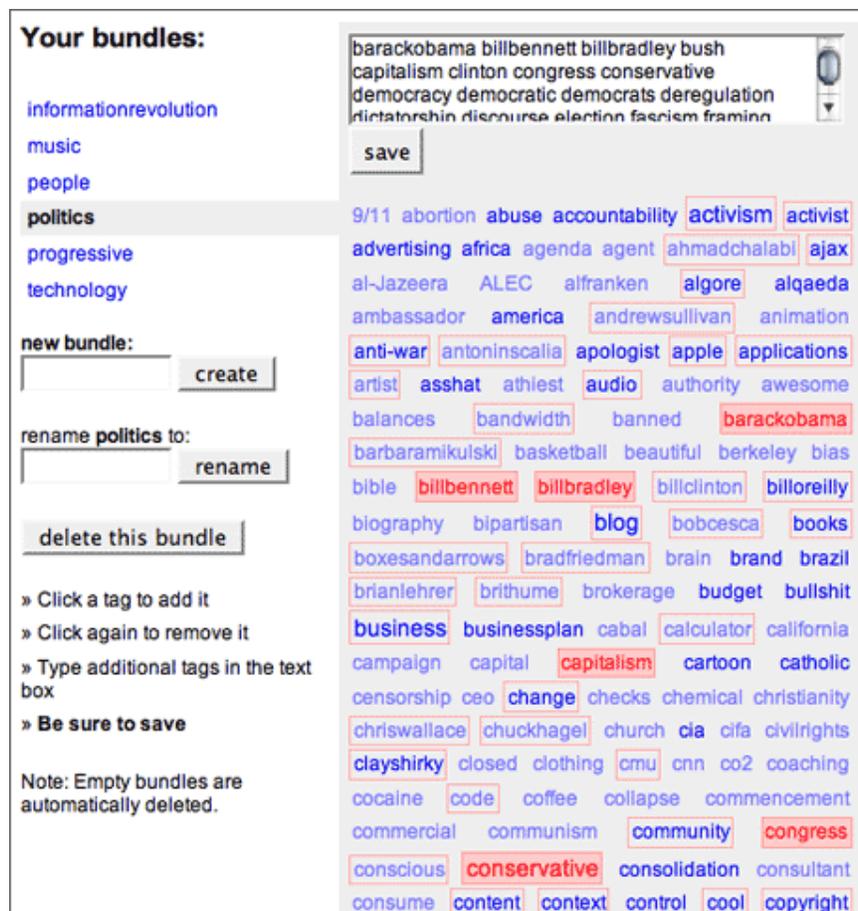


Figura 2 - Social tagging: un esempio particolarmente evoluto, la funzione di raggruppamento dei tag in del.icio.us

Ci occupiamo più a fondo di questi sistemi (spesso chiamati anche 'folksonomy') in altra sede, anche perché il social tagging rappresenta una delle tendenze distintive del cosiddetto 'Web 2.0'. Qui può essere il caso di ricordare che, se il social tagging può essere effettivamente utile a garantire un primo livello di gestione 'facile' e immediata di grandi quantità di contenuti generati dagli utenti, funziona assai meno bene quando l'obiettivo è garantire una buona reperibilità e interoperabilità di informazione validata e dalle caratteristiche un po' più uniformi. Va infatti considerato da un lato che le 'etichette' attribuite dagli utenti

sono di norma piatte, non strutturate<sup>4</sup> (e questo impedisce di rendere conto di molte forme di dipendenza e strutturazione della nostra informazione primaria), dall'altro che l'assenza di normalizzazione, anche solo nel lessico utilizzato, impedisce in molti casi il reperimento affidabile e funzionale dell'informazione che ci interessa.

Questo non toglie, però, che l'uso di sistemi di social tagging possa spesso utilmente *accompagnare* (anziché sostituire) una descrizione più rigorosa e standardizzata dei contenuti. In questo caso, la sua funzione diventa quella di esprimere descrizioni alternative, percorsi trasversali, permettendo di dar voce anche al punto di vista dell'utente (o meglio, ad alcuni fra i molti punti di vista possibili di molti tipi diversi di utenza). E indubbiamente il social tagging può rappresentare l'unica strada sensatamente percorribile quando l'informazione da descrivere è in quantità o di tipologia tale da non permettere, o da rendere troppo oneroso, l'intervento di classificatori esperti. Come accade spesso, appunto, nel mondo dell'informazione 'user generated' resa disponibile attraverso la rete. In questi casi, però, è bene che anche i sistemi di social tagging messi a disposizione degli utenti siano 'pensati' in maniera attenta e consapevole, in modo da cercare di guidare gli utenti (ad esempio – come già accennato – attraverso l'impiego di strumenti di collaborative filtering) a una scelta di etichette il meno arbitraria e il più informativa e sensata possibile. Questo potrà permettere, in alcuni casi, di strutturare e organizzare a posteriori, anche in maniera automatica, i metadati immessi (quasi) liberamente dagli utenti.

### 5. Metadati e repository; l'harvesting

Abbiamo visto che lo scopo principale dei metadati è di norma descrivere l'informazione primaria, organizzarla, semplificarne e automatizzarne la gestione. Non stupirà dunque che uno dei più importanti contesti d'uso di sistemi di metadati sia rappresentato dai depositi organizzati (*repository*) di contenuti, siano essi digitali o no. Anzi, possiamo dire che per poter parlare di deposito *organizzato* di contenuti, di repository appunto, la presenza di metadati rappresenta una condizione imprescindibile. Un repository è dunque uno strumento di raccolta, conservazione e (in genere) distribuzione di contenuti informativi descritti, gestiti e organizzati attraverso opportuni insiemi di metadati.

In questa sede, il nostro interesse si concentra soprattutto sui repository di contenuti in formato digitale, accessibili via rete. In genere, essi sono gestiti attraverso programmi che permettono l'archiviazione o l'autoarchiviazione di contenuti, la loro identificazione univoca, la loro metadattazione, la loro organizzazione in raccolte o collezioni, la loro conservazione anche di lungo termine, la loro ricerca e il loro recupero. Di norma, in particolare per l'accesso alle funzionalità di archiviazione dei contenuti, è prevista un'identificazione degli utenti attraverso sistemi di autenticazione. Talvolta i repository sono dotati anche di strumenti aggiuntivi, ad esempio per la gestione dei diritti relativi ai contenuti archiviati, per la conversione automatica di formati, per la creazione di pacchetti di contenuto, per l'esposizione dei metadati, per l'aggiunta ai contenuti di commenti, giudizi o etichette da parte degli utenti, per la

---

<sup>4</sup> Anche se cominciano ad apparire strumenti, come quelli che consentono di raggruppare ('bundle') le etichette all'interno di del.icio.us (<http://del.icio.us>), che cercano proprio di superare questo problema.

## Metadati per l'e-Learning: una introduzione

---

distribuzione di feed RSS relativi ai nuovi contenuti immessi o resi disponibili, ecc.: ci occuperemo in seguito di alcune di queste funzionalità, mentre di altre abbiamo parlato o parleremo in altre occasioni.

I LCMS (Learning Content Management System) sono repository di questo genere, destinati alla raccolta, descrizione, gestione e distribuzione di contenuti e risorse di apprendimento. Alcuni LCMS consentono anche la produzione di contenuti, attraverso appositi moduli per l'editing di learning content; a volte sono presenti anche moduli player, che consentono la fruizione di contenuti all'interno dello stesso ambiente di distribuzione. Ma gli LCMS non sono certo il solo modello di repository esistente: fra le tipologie di repository più diffuse e importanti (a loro volta variamente articolate) ricordiamo le Digital Library (biblioteche digitali), gli Open Archive (archivi aperti), le piattaforme per la raccolta, gestione e condivisione di contenuti sonori o visivi (immagini e filmati), i sistemi per la gestione e l'accesso a basi dati di riviste e libri elettronici realizzati da molte case editrici, e così via. Spesso i confini fra queste diverse tipologie sono labili e dipendono, più che dallo specifico strumento informatico utilizzato, dall'uso che ne viene fatto, o da considerazioni di mercato. In tutti questi casi, comunque, la gestione dei metadati gioca un ruolo fondamentale, e una delle caratteristiche distintive dei vari repository è proprio legata ai sistemi di metadati che essi riconoscono e consentono di utilizzare.

Per rendere disponibili e ricercabili i propri contenuti, un repository deve dunque innanzitutto rendere disponibili e ricercabili i propri metadati. Ed è bene ricordare che, anche se frequentemente pensiamo ai metadati come a una sorta di 'involucro' che circonda l'informazione primaria, non è affatto detto che i metadati debbano sempre essere utilizzati 'insieme' ai relativi contenuti. Proprio come le schede del catalogo di una biblioteca possono essere utilizzate anche indipendentemente dai libri che descrivono (e magari essere raccolte in pubblicazioni autonome, utili a controllare anche a distanza il patrimonio di una biblioteca o a recuperare specifiche informazioni bibliografiche), anche i metadati in formato elettronico possono avere una vita autonoma, circolare ed essere distribuiti per disseminare le informazioni relative ai contenuti di un certo repository e facilitarne il reperimento.

Una delle modalità più interessanti di questa circolazione 'autonoma' dei metadati è rappresentata dal concetto di *harvesting*. Vediamo di capire di cosa si tratta.

Intuitivamente, saremmo portati a pensare che il modo migliore per semplificare il reperimento e il recupero di grandi quantità di contenuti informativi sia quello di non disperdere i contenuti stessi e realizzare invece repository il più larghi e centralizzati possibile, abbastanza ricchi e completi da non richiedere all'utente eccessive 'peregrinazioni' da un repository all'altro per trovare il contenuto che lo interessa.

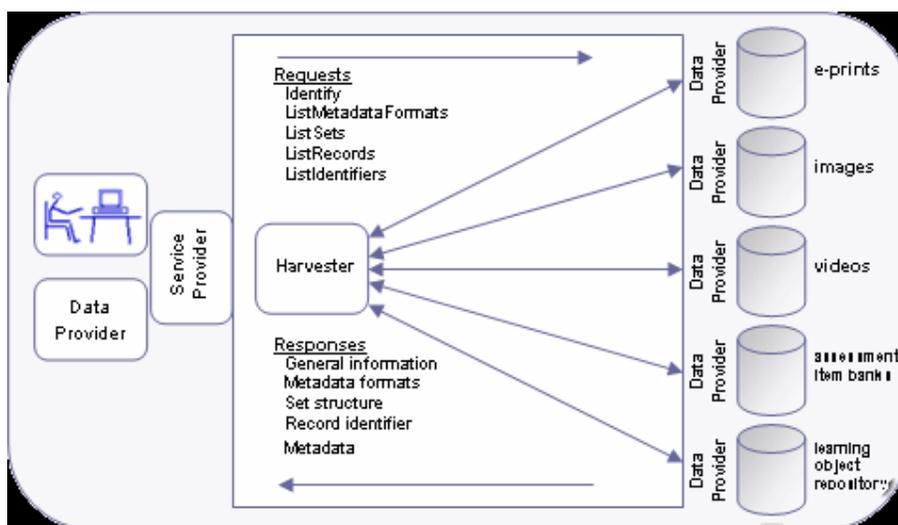
A ben vedere, però, per garantire semplicità ed efficienza non occorre necessariamente utilizzare pochi grandi repository: è anche possibile (e in molti casi è più semplice ed efficace) realizzare più depositi, magari settoriali o corrispondenti a istituzioni diverse, che però condividano l'uso di sistemi standard di metadati e di recupero dei contenuti, e rendano disponibili i loro metadati per forme di indicizzazione integrata e cooperativa. In sostanza, non occorre centralizzare i contenuti: basta standardizzare i metadati e centralizzare i servizi per la loro ricerca e indicizzazione (tenendo anche presente che in genere i

## Metadati per l'e-Learning: una introduzione

metadati 'pesano' in termini di bit molto meno delle relative risorse informative primarie, e possono dunque viaggiare con estrema facilità).

È chiaro che questo modello richiede l'adozione generalizzata di standard e protocolli condivisi, sia per quello che riguarda i sistemi di metadati, sia per quello che riguarda le procedure per facilitarne la circolazione e la consultazione da parte di servizi esterni, sia per quello che riguarda la richiesta da parte dell'utente e l'invio dei contenuti da parte del repository.

Un meccanismo di questo tipo si baserà dunque su una pluralità di repository che fungano da *data provider*, che cioè rendano disponibili i contenuti, e che *espongano* i relativi metadati in un formato standard e riconoscibile (basato in genere sull'uso di file XML). Alla pluralità di data provider corrisponderanno pochi *service provider*, che saranno in grado di raccogliere automaticamente (*harvesting*) i metadati esposti dai vari data provider, di integrarli e di offrire servizi centralizzati di interrogazione e ricerca (ma anche, ad esempio, servizi statistici o di altro tipo).



**Figura 3 - Harvesting dei metadati: i diversi data provider espongono i loro metadati, raccolti dal service provider attraverso l'harvester. Il service provider permette poi all'utente di svolgere ricerche sull'intero insieme di metadati raccolti.**

Questo modello, nato nel mondo degli Open Archive (la tipologia di repository utilizzata sempre più spesso da istituzioni accademiche e di ricerca per raccogliere e distribuire in forma aperta i prodotti delle attività di ricerca: articoli, pre-print, post-print ecc.), è evidentemente applicabile a moltissimi altri tipi di repository, e l'uso del protocollo OAI-PMH (Open Archives Initiative – Protocol for Metadata Harvesting), un insieme di convenzioni che disciplinano l'esposizione dei metadati e il loro harvesting<sup>5</sup>, si sta estendendo dal mondo Open Archive a repository di altro genere, compresi quelli che raccolgono Learning Object o altri contenuti di apprendimento.

<sup>5</sup> Su OAI-PMH il sito di riferimento è <http://www.openarchives.org/pmh/>; per una discussione più approfondita sull'uso di repository OAI nel campo dell'e-learning si veda G. Roncaglia, *Quali repository per l'e-Learning*, in "Il Giornale dell'e-Learning" a. 2 n. 1 (gennaio 2008), in rete alla pagina <http://www.wbt.it/index.php?pagina=440>.

E' possibile che in futuro uno sviluppo analogo attenda anche un altro protocollo nato nel mondo Open Archive, il protocollo OAI-ORE (la sigla ORE è un acronimo di Object Reuse and Exchange), che consente lo scambio non solo dei metadati ma anche direttamente degli oggetti digitali, prefigurando una nuova generazione di servizi capaci di valorizzare i contenuti di un repository creando, ad esempio, raccolte e collezioni composte da oggetti che si trovano in repository diversi<sup>6</sup>. Anche in questo caso, le possibili applicazioni al campo dell'e-learning sono di evidente interesse.

### 6. Quali metadati per l'e-Learning?

Fin qui, abbiamo cercato di fornire una introduzione abbastanza generale al concetto di metadati e ai principali strumenti ad esso associati. È arrivato il momento di applicare un po' più direttamente questi concetti al campo dell'e-learning.

Da dove partire? Come abbiamo visto, uno dei problemi fondamentali è capire con quale tipo di contenuti abbiamo a che fare. Cos'è – esattamente – che vogliamo classificare?

Una prima risposta, abbastanza generale, è che al centro della nostra attenzione ci sono *risorse e contenuti di apprendimento che siano a) validati o validabili, e b) distribuibili attraverso strumenti informatici e telematici*. A questo primo livello non ci interessano solo o necessariamente contenuti fortemente strutturati, o destinati alla fruizione all'interno di corsi formali di e-Learning, ma qualunque tipo di contenuto possa rappresentare una risorsa utile all'apprendimento, purché suscettibile di validazione (provenga cioè da una fonte affidabile, o possa essere in qualche modo certificato o validato) e di distribuzione in formato digitale. Un campo ovviamente vastissimo, che comprende tipologie assai eterogenee di risorse.

All'interno di questo campo, ci sono due aspetti (fra loro interconnessi) che ci interessano più direttamente: 1) la *riusabilità* dei contenuti, e 2) la loro *modularità*.

Non tutti i contenuti validabili e distribuibili in formato digitale sono necessariamente riusabili: alcuni possono essere stati costruiti in funzione di una singola, specifica situazione di fruizione, e risultare quindi talmente dipendenti dal contesto da non essere di fatto riutilizzabili in situazioni diverse. Altri possono non essere riusabili per problemi di diritti, o di dipendenza da tecnologie non disponibili altrove (ad esempio dipendenza da una piattaforma strettamente proprietaria). Dal nostro punto di vista hanno naturalmente un particolare rilievo i contenuti riusabili, che può essere sensato distribuire anche all'esterno del contesto specifico in cui sono stati prodotti. Il rilievo attribuito alla caratteristica della riusabilità suggerisce evidentemente la necessità di metadati specifici che indichino in che modo e a quali condizioni la risorsa possa essere riutilizzata, e quali ne siano i potenziali destinatari.

La discussione sulla modularità dei contenuti di apprendimento è complessa e non è certo il caso di affrontarla in questa sede; ci interessa però ricordare che, anche in questo caso, saranno necessari

---

<sup>6</sup> Su OAI-ORE si veda <http://www.openarchives.org/ore/> e, per una presentazione sintetica, <http://www.ukoln.ac.uk/repositories/digirep/index/OAI-ORE>.

## Metadati per l'e-Learning: una introduzione

---

metadati specifici che indichino l'eventuale inclusione di ogni contenuto distribuito all'interno di unità di macrolivello (corsi, serie ecc.), nonché le eventuali dipendenze gerarchiche, funzionali o di altro genere di un contenuto rispetto ad altri. Anche se l'idea di *autoconsistenza* e quella di *granularità* sono spesso richiamate nel definire le caratteristiche di una delle più diffuse tipologie di learning content, i Learning Object, bisogna osservare che la totale autosufficienza di un contenuto di apprendimento è l'eccezione e non certo la regola. E sarebbe del resto assai strano se la complessità della nostra esperienza e del contesto sociale e culturale che ci circonda si lasciasse esprimere – a livello di formazione e di apprendimento – attraverso contenuti completamente chiusi, privi di legami fra loro e con l'universo informativo esterno.

Ciò significa, dal punto di vista che più direttamente ci interessa in questa sede, che il nostro insieme di metadati, oltre a specificare la natura, il livello di complessità, i destinatari previsti di una risorsa di apprendimento, dovrà avere anch'esso la possibilità di 'guardare all'esterno', collocando il contenuto nel suo contesto di produzione (sempre rilevante per valutarne la validità e la rilevanza rispetto alle possibili situazioni di riuso), specificando se e a quale tipo di processo di validazione il contenuto è stato sottoposto, indicando se, in quale misura e in quali situazioni il contenuto stesso può essere eventualmente riutilizzato, e permettendo di esprimere la rete di possibili dipendenze gerarchiche e funzionali che lo caratterizzano.

Come ricorderete, la scelta di un insieme di metadati, e le scelte fatte nella sua strutturazione e organizzazione, corrispondono al modo in cui scegliamo di guardare la specifica porzione del mondo al quale quei metadati si applicano. Nel nostro caso, dunque, al modo in cui interpretiamo il concetto di 'risorsa di apprendimento'.

Ma il mondo dell'e-learning non è fatto solo di risorse di apprendimento: è fatto anche degli strumenti tecnici utilizzati per preparare e per fruire quelle risorse, di piattaforme attraverso cui archivarle e distribuirle, di corsi che le utilizzano, di piattaforme attraverso cui i corsi sono erogati, dei diversi ruoli e figure coinvolti in questo processo (a cominciare da docenti, tutor, corsisti...), di istituzioni che a vario livello e con vari ruoli intervengono a stimolarlo, sostenerlo e finanziarlo... e ciascuno di questi elementi dovrà essere adeguatamente descritto, permettendo di esprimerne la natura e le dipendenze reciproche. Insomma, come avrete capito molto difficile, se non impossibile, esprimere la complessità del mondo dell'e-learning attraverso pochi insiemi chiusi e uniformi di metadati. E tuttavia per interpretare e gestire questa realtà complessa abbiamo comunque bisogno di modelli (non a caso, i sistemi di metadati vengono di norma espressi attraverso quelli che tecnicamente si chiamano *data model*), che in qualche modo la rendano più semplice e manipolabile. È in quest'ottica che dobbiamo guardare agli standard e agli insiemi di metadati che ci vengono proposti (o che vogliamo costruire): non nel tentativo di trovare il 'vero' standard, perfetto, unico e definitivo, ma nel tentativo di individuare quello più efficace, rigoroso e condiviso in funzione della nostra particolare visione del mondo e dei nostri particolari obiettivi.

In questa sede, introduciamo brevemente solo alcuni di questi standard: non intendiamo descriverli in maniera completa ed esaustiva, ma solo darne una prima idea, accennando alle principali situazioni e funzioni per le quali sono stati costruiti, ma anche rilevando alcuni fra gli scopi e le situazioni in cui si possono invece dimostrare meno specificamente indicati.

Come si è accennato, il primo problema è sempre e comunque capire a quale ambito ci stiamo riferendo, quale tipo di 'oggetti' vogliamo descrivere.

### 7. I metadati LOM

Le risorse di apprendimento sono senz'altro un primo tipo di oggetti informativi che nel campo dell'e-learning è importante poter descrivere e caratterizzare attraverso metadati adeguati. Quali metadati utilizzare? Fra le risorse di apprendimento, i Learning Object – della cui natura e caratteristiche ci occupiamo a fondo in altra sede – rappresentano sicuramente una delle tipologie più diffuse (e più discusse). Per gestirli, è stato proposto un insieme di metadati specifici (dunque, un *data model*), denominato LOM (Learning Object Metadata).

In realtà, il nome 'ufficiale' dello standard è più complesso: IEEE 1484.12.1 – 2002. La sigla IEEE si riferisce all'*Institute of Electrical and Electronics Engineers*, una organizzazione no-profit nata nel 1963 per promuovere lo sviluppo scientifico nel campo dell'elettricità, dell'elettronica, dell'informatica e delle telecomunicazioni. L'IEEE si è imposta rapidamente come una delle organizzazioni più impegnate nella definizione di standard in questi settori, principalmente attraverso l'attività della *Standards Association* operante al suo interno (IEEE-SA). LOM è solo uno dei molti (circa 900!) standard proposti da questa associazione, anche se ovviamente è quello che in questa sede ci interessa in maniera più diretta.

In molti di questi casi, l'IEEE si avvale anche di consulenze e indicazioni provenienti da altri soggetti qualificati: in particolare, un ruolo determinante nello sviluppo dei metadati LOM lo ha avuto l'IMS Learning Consortium, una associazione nata nel 1997 per iniziativa di *Educause*, l'organizzazione no-profit responsabile della gestione dei domini .edu statunitensi e avente per obiettivo il progresso dell'educazione superiore attraverso l'uso intelligente delle tecnologie dell'educazione. L'IMS (acronimo di *Instructional Management Systems*) Learning Consortium è anch'esso una organizzazione no-profit, che ha come obiettivo specifico proprio la definizione di standard in grado di garantire l'interoperabilità fra sistemi di apprendimento e contenuti di apprendimento. Lo ritroveremo dunque all'opera nella definizione di molti degli standard di cui ci occupiamo in questa sede, e incontreremo spesso la sigla 'IMS' in tutti gli ambiti dell'e-Learning nei quali sono in corso sforzi di standardizzazione.

L'IMS Learning Consortium – avendo partecipato alla loro definizione – ha cominciato subito ad adottare le prime versioni dello standard LOM, incorporandole in un sistema denominato IMS Learning Resource Metadata specification (o, in breve, IMS LRM). Lo sviluppo di questi due sistemi (LOM e IMS LRM) non è stato sempre perfettamente allineato, ma la versione 1.3 di IMS LRM ha rimesso le cose a posto, e il termine 'LOM' può oggi essere usato indifferentemente per lo standard IEE e per la versione 1.3 di IMS LRM.

Ci siamo soffermati con qualche dettaglio su questo complicato arcipelago di sigle – che ricorda un po' il meccanismo di 'scatole cinesi' utilizzato per mascherare le vere proprietà di certe società off-shore – sia per cercare di renderle un po' meno 'aliene' (e combattere la pessima abitudine di introdurre sigle ed acronimi senza spiegarne il significato e la storia), sia perché, come spesso accade, l'origine di una proposta aiuta a capirne il significato e le caratteristiche. Così, ad esempio, il fatto che IMS sia nato in un contesto legato in

primo luogo all'istruzione superiore e universitaria ha reso necessari specifici aggiustamenti per garantire la funzionalità dello standard LOM anche per risorse legate all'apprendimento infantile e ai primi gradi scolastici, mentre in generale il forte collegamento con contesti di formazione formale rende i metadati LOM talvolta non pienamente funzionali nella descrizione di risorse di apprendimento legate alla formazione informale o al lifelong learning.

Abbiamo detto che la sigla LOM sta per 'Learning Object Metadata'. Ma cosa si intende qui per 'Learning Object', un termine la cui connotazione sappiamo bene essere tutt'altro che chiara e uniforme? Gli obiettivi dello standard sono ambiziosi: fornire un sistema di metadati applicabili a "qualunque tipo di entità, digitale o non digitale, che possa essere utilizzata per l'apprendimento, l'educazione o la formazione". Questa definizione di Learning Object è indubbiamente assai ampia, e la discuteremo in altra sede. In che misura i metadati LOM siano realmente in grado di rispondere all'esigenza di descrivere e gestire in maniera funzionale un dominio così ampio di oggetti informativi è questione assai dibattuta (e assai interessante), ma che esula anch'essa dai nostri obiettivi immediati. Per il momento, ci interessa un obiettivo più specifico: farci un'idea di quali siano, e come siano strutturati, i metadati LOM.

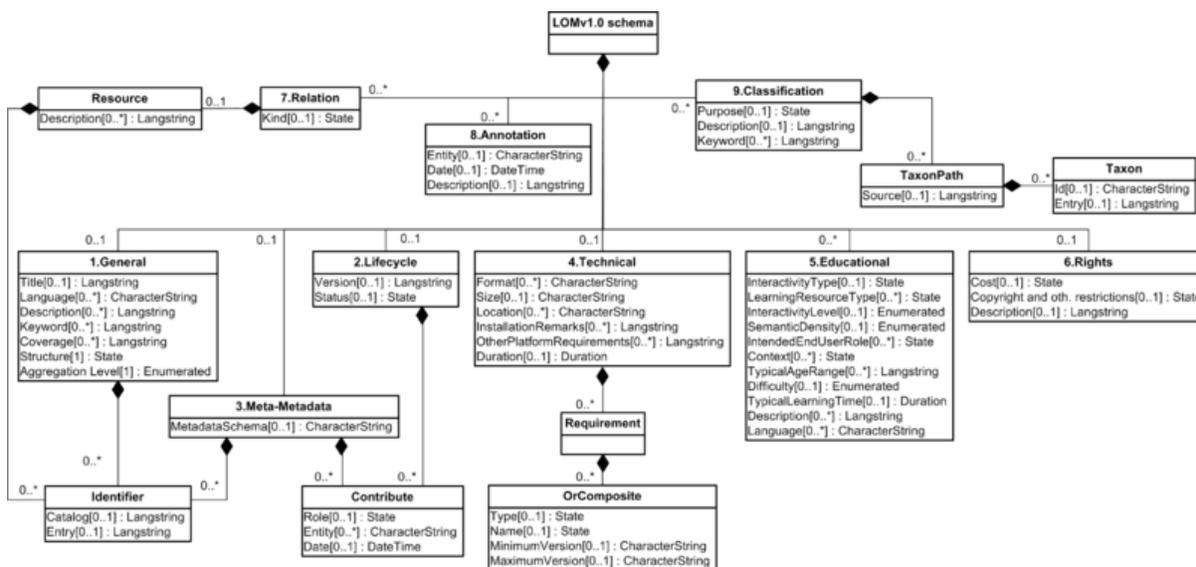


Figura 4 - Schema gerarchico di strutturazione dei metadati LOM

Come si può vedere dalla figura precedente, si tratta di una strutturazione fortemente gerarchica: i metadati LOM non sono un insieme piatto di categorie, ma sono organizzati in una gerarchia di elementi di diverso livello. Il livello più alto è composto da nove categorie, che è il caso di ricordare brevemente:

1. **General**: questo gruppo di elementi raccoglie descrittori generali (e di grande importanza) come il titolo della risorsa di apprendimento, la sua descrizione, le parole chiave che desideriamo associarvi, l'ambito spaziale o temporale coperto, il livello di aggregazione (singola lezione, intero corso ecc.).

## Metadati per l'e-Learning: una introduzione

---

2. **Lifecycle:** è utilizzato per fornire indicazioni sul ciclo di vita della risorsa: ad esempio il numero di versione, o lo stato di completamento.
3. **Meta-metadata:** come suggerisce il nome, si tratta di descrittori relativi all'insieme di metadati utilizzati, ad esempio la versione di LOM utilizzata, in genere espressa attraverso un riferimento al relativo identificativo univoco (e cioè al particolare *metadata schema* utilizzato).
4. **Technical:** le caratteristiche tecniche della risorsa (formato, dimensione ecc.) e i requisiti tecnici richiesti per fruirne (software necessari, sistemi operativi compatibili ecc.).
5. **Educational:** si tratta del sottoinsieme di descrittori specificamente destinati a dar conto della natura di risorsa educativa propria dell'oggetto informativo che stiamo descrivendo. È dunque l'insieme di caratteristiche che più direttamente distingue i metadati LOM rispetto ad altri schemi di metadati, e anche il più 'rivelatore' della concezione dell'apprendimento alla base del sistema di descrizione (non a caso, i metadati 'educational' sono fra i più discussi...). Alcuni di questi metadati sono fortemente soggettivi (ad esempio il livello di difficoltà della risorsa, il suo livello di interattività, la sua densità semantica), altri un po' meno (contesto di apprendimento, ruolo dell'utente-tipo, età tipica del fruitore, tempo di fruizione previsto...). Ma certo la compilazione dei metadati 'educational' di una risorsa di apprendimento richiede particolare attenzione, e spesso anche l'adozione di alcune convenzioni comuni, ad esempio, alle risorse di apprendimento legate a uno stesso progetto o ospitate da uno stesso repository.
6. **Rights:** Tutte le informazioni che riguardano la gestione dei diritti relativi alla risorsa: la tipologia del copyright, chi lo detiene, quali autorizzazioni sono eventualmente richieste per utilizzare la risorsa (ad esempio, quanto costa acquistarne una licenza d'uso), ecc.
7. **Relation:** sull'importanza del poter esprimere non solo una descrizione 'isolata' di una risorsa, ma anche le relazioni che essa ha con l'universo informativo circostante ci siamo già soffermati. I metadati di tipo 'relation' permetteranno dunque di esprimere, ad esempio, relazioni di dipendenza fra una risorsa di apprendimento e un'altra (prerequisiti ecc.), o relazioni di inclusione e contenimento (per specificare, ad esempio, che una particolare risorsa di apprendimento è parte di un corso), ecc.
8. **Annotation:** permette di rendere più flessibile la descrizione di una risorsa aggiungendovi indicazioni e commenti specifici (ad esempio, commenti degli utenti). Anche in questo caso, è spesso utile adottare convenzioni comuni per l'uso di questi metadati, in modo da sfruttarne le potenzialità in relazione alle particolari caratteristiche del progetto e delle risorse di apprendimento utilizzate.
9. **Classification:** questo insieme di metadati permette di far riferimento a qualsiasi standard o sistema di classificazione esterno, in modo da consentire una descrizione delle nostre risorse di apprendimento anche utilizzando sistemi di classificazione specifici.

Questa breve presentazione dovrebbe darvi un'idea di cosa siano i metadati LOM, ma naturalmente usarli in pratica è tutt'altra cosa! Nel provarci incontrerete sicuramente dubbi, situazioni problematiche, casi in

cui i metadati disponibili non vi sembrano completamente adatti o in cui vi sembra mancare qualcosa di necessario per il vostro progetto. Per questo, ci sentiamo di raccomandarvi due cose: *studiare* come sono state classificate da altri risorse di apprendimento abbastanza simili a quelle con le quali lavorate voi, magari affidandovi anche ai molti esempi e tutorial disponibili in rete, e *collaborare*: la metadattazione, per sua natura, è una operazione che si svolge sempre in un contesto sociale, deve risultare funzionale per gli utenti e per le altre persone con cui lavoriamo. È dunque di norma una attività collaborativa, e non solitaria e personale.

Come abbiamo accennato, peraltro, l'uso di uno standard di metadattazione come LOM avviene spesso o facendo riferimento a insiemi ridotti di metadati (scelti fra quelli previsti dallo standard, in funzione delle particolari esigenze, ad esempio, di un corso o di un progetto specifico), o, al contrario, 'incorporando' i metadati suggeriti dallo standard (o un loro sottoinsieme) all'interno di una cornice più generale, che permetta di descrivere anche altri aspetti (o altre tipologie di oggetti) del nostro dominio di lavoro. L'importante è che in questi casi le scelte fatte vengano sempre esplicitamente *dichiarate* (possibilmente in forma rigorosa), *motivate* e *documentate*. Nel farlo, è sicuramente consigliabile – se appena possibile – rifarsi ad altre esperienze 'autorevoli' con caratteristiche ragionevolmente simili alla nostra.

Un esempio della prima possibilità – il ricorso a insiemi ridotti e semplificati di metadati, spesso identificati attraverso l'uso di termini come 'core' o 'lite' – è offerto dal progetto canadese CanCore (<http://www.cancore.ca/>), il cui sito web costituisce una risorsa preziosa anche per i molti esempi e tutorial che contiene, e per le raccomandazioni (*guidelines*), utilissime, sul modo migliore per utilizzare i vari insiemi di metadati.

L'esempio per noi più importante della seconda possibilità – l'uso di metadati LOM all'interno di una 'cornice' descrittiva e operativa più generale – è invece dato dallo standard SCORM, sul quale ci soffermeremo brevemente nel prossimo paragrafo.

### 8. SCORM: una breve introduzione

Lo standard (o meglio, l'insieme di standard) SCORM (*Sharable Content Object Reference Model*) è nato nel gennaio del 2000, nell'ambito della *Advanced Distributed Learning (ADL) Initiative*, sviluppata dal dipartimento della difesa statunitense. Si tratta dunque di standard nati in ambito governativo, legati alle specifiche caratteristiche e necessità della formazione militare e industriale: creare procedure semplici e standardizzate per 'somministrare' pacchetti di formazione, in genere legati ad attività tecniche o operative a una vasta base di utenza dalle caratteristiche assai differenziate.

E in effetti l'aspetto più importante di SCORM (che nel corso del tempo è passato attraverso versioni diverse: le più recenti e diffuse sono la 1.2 e le SCORM 2004, la cui ultima versione è in realtà del 2006) è proprio il fatto di fornire un modello di riferimento (*reference model*) per la creazione, la descrizione e la gestione di 'pacchetti' di contenuti formativi, e per l'interazione fra tali pacchetti e le piattaforme di distribuzione. SCORM non si limita dunque a *descrivere* risorse di apprendimento, ma permette di rappresentare la loro *organizzazione e concatenazione* (senza entrare naturalmente nel merito dei

## Metadati per l'e-Learning: una introduzione

---

contenuti, ma fornendo un modello per la loro aggregazione in pacchetti), e fornisce strumenti per la gestione e distribuzione di tali pacchetti da parte di una piattaforma.

Questo significa che SCORM ha come obiettivo l'*interoperabilità* fra contenuti di apprendimento e piattaforme web: una piattaforma in grado di riconoscere lo standard SCORM sarà comunque in grado di gestire e distribuire un pacchetto di contenuti conforme a tale standard, indipendentemente dall'origine dell'una e dell'altro. Un po' come un programma creato per un certo sistema operativo è in grado di funzionare su qualunque computer dotato di quel sistema operativo, indipendentemente da chi ha fabbricato il computer e da chi ha creato il programma<sup>7</sup>.

Le specifiche tecniche del modello consentono infatti ai pacchetti SCORM di essere 'avviati' all'interno di appositi ambienti esecuzione. Ciò significa che è possibile creare veri e propri 'player SCORM', in grado di 'eseguire' (e dunque di proporre all'utente) qualunque pacchetto costruito in modo conforme allo standard. Ovviamente, questi player SCORM troveranno di norma posto all'interno di piattaforme didattiche (LMS) o di piattaforme per la distribuzione di contenuti di apprendimento (LCMS); ma possono essere implementati anche come strumenti stand-alone (un po' come un programma 'player' per video o musica, solo che in questo caso a essere visualizzati sono i contenuti del pacchetto SCORM, nell'ordine previsto da chi li ha assemblati), e, in prospettiva, anche come 'widgets' o browser application per siti web.

E' importante notare che SCORM è sempre e comunque orientato al web, e nasce per la gestione di contenuti di apprendimento destinati alla distribuzione su web: in sostanza, la cornice per le risorse di apprendimento e le piattaforme di distribuzione che SCORM si propone di rendere interoperabili, è rappresentata proprio dal web.

Sullo standard SCORM, e sulla creazione di pacchetti SCORM, torneremo più volte nel corso del master, anche con tutoriali specifici. Lo scopo di questo paragrafo è solo quello di fornire una prima introduzione alle caratteristiche generali del modello. Esaminiamolo dunque un po' più da vicino.

Il *reference model* SCORM è composto – nella sua versione più recente – da tre diverse componenti, che corrispondono a loro volta a modelli in grado di svilupparsi anche in maniera indipendente:

- il modello relativo alla organizzazione e aggregazione dei contenuti (*content aggregation model*);
- le specifiche relative all'ambiente di esecuzione (*run-time environment*);
- le specifiche relative all'ordinamento (sequenziazione) dei contenuti e alla loro navigazione (*sequencing and navigation*).

---

<sup>7</sup> Forse qualcuno obietterà che questo non è sempre del tutto vero: sappiamo fin troppo bene che le caratteristiche hardware del computer e la particolare configurazione del suo sistema operativo possono impedire la piena fruizione di alcuni programmi, anche se tali programmi sono stati creati proprio per quel sistema operativo. In un certo senso, lo stesso vale per SCORM: il rispetto degli standard dovrebbe sempre garantire una interoperabilità di base, ma alcuni contenuti del pacchetto possono poi richiedere risorse specifiche per essere fruiti in maniera pienamente soddisfacente, e non tutte le piattaforme di distribuzione (o i programmi client – a partire dai browser web – necessari per utilizzarle) possono effettivamente offrire quelle risorse.

## Metadati per l'e-Learning: una introduzione

---

Vediamo di capire di cosa si tratta.

In SCORM, il modello di organizzazione dei contenuti assomiglia molto a un insieme di scatole cinesi, o, per usare una metafora frequente, a una 'libreria' composta di insiemi organizzati di volumi. Il livello più alto è appunto quello della *content organization* (che corrisponde un po' alla libreria). Una *content organization* SCORM può corrispondere a un corso, a un insieme di corsi, a una parte di un corso... si tratta comunque di una macro-aggregazione organizzata e strutturata di contenuti e risorse, per i quali si ritiene utile costruire una sorta di 'mappa' comune.

Una *content organization* SCORM comprende uno o più SCO. Gli SCO, ovvero *Sharable Content Object*, sono i contenuti organizzati di apprendimento, e dunque sostanzialmente proprio quei Learning Object (in un senso abbastanza ampio del termine) che LOM permette di descrivere. Gli SCO sono a loro volta costituiti da uno o più *asset*, risorse 'atomiche' che vanno a costituire il contenuto di apprendimento (ad esempio, le slide di una presentazione, le singole dispense, ecc.), e sono avvolti – o meglio, accompagnati – da un involucro, o *wrapper*, che fornisce al player SCORM istruzioni su come i contenuti devono essere presentati all'utente.

Come vedremo più avanti, i vari SCO e asset possono poi essere organizzati in attività (*activity*) di diverso livello: al livello più alto, la stessa *content organization* può essere pensata come una *activity*, organizzata a sua volta in sotto-attività<sup>8</sup>, con una struttura anche molto complessa.

Come avrete capito, il modello di aggregazione dei contenuti proprio di SCORM è basato su una struttura articolata, ma nel contempo fortemente gerarchica. E a ogni livello della gerarchia sono presenti gli opportuni metadati. Lo SCO vero e proprio è descritto attraverso metadati LOM, mentre le aggregazioni o concatenamenti di più SCO che costituiscono una *content organization* sono descritti attraverso metadati *esterni* ai singoli SCO, i *Content Aggregation Metadata* o CAM. I CAM diranno dunque in quale sequenza devono essere utilizzati gli SCO, quali sono le loro relazioni di dipendenza reciproca, ecc.

Fin qui, abbiamo parlato della aggregazione di contenuti in termini ancora abbastanza astratti. Ma cosa corrisponde, concretamente, a una *content organization*? Dal punto di vista pratico, una *content organization* si traduce in un pacchetto (*content package*) SCORM, che comprende tutti i file necessari a permetterne il funzionamento all'interno di un ambiente di esecuzione che rispetti lo standard. Questi file sono opportunamente organizzati e descritti attraverso un documento XML specifico, denominato *IMS manifest* (il riferimento a IMS, come avrete sospettato, è legato al ruolo avuto da questa organizzazione nella definizione dei metadati inclusi nel manifest). Il manifest di un pacchetto SCORM è dunque una sorta di 'indice' del pacchetto, che ne descrive analiticamente il contenuto e l'organizzazione, attraverso gli opportuni metadati, sia relativi ai singoli SCO (e si tratterà in questo caso principalmente di metadati LOM) sia relativi alla loro aggregazione.

Il *content package* comprende dunque il manifest e i file che fanno parte del progetto; per tenere insieme il tutto viene usato il formato di compressione più popolare e diffuso, ZIP, ma per semplificarne il

---

<sup>8</sup> Nel linguaggio SCORM, le attività composte di sotto-attività sono dette anche *cluster*, mentre le attività 'atomiche', che non prevedono sotto-attività, sono dette *leaf activities* (perché corrispondono alle 'foglie' dell'albero gerarchico delle attività).

riconoscimento il file che ne risulta usa l'estensione PIF, acronimo di 'Package Interchange File'. Un file PIF è dunque un file zippato, che contiene risorse di apprendimento e comprende un manifest e un insieme di file di contenuto.

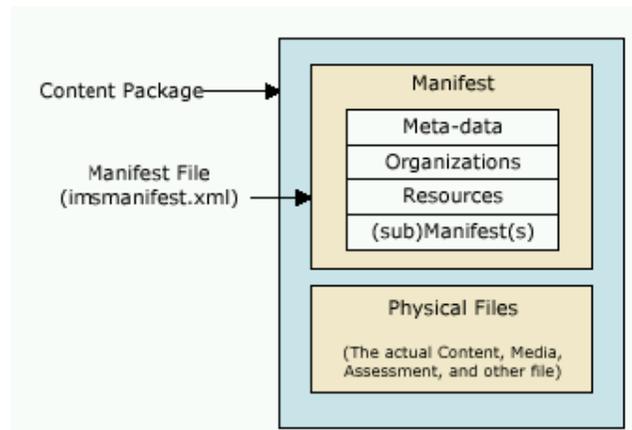


Figura 5 - Struttura di un content package SCORM (dalla versione 1.3 del CAM SCORM, ©ADL 2004)

Lo scopo del secondo elemento essenziale di SCORM, il run-time environment o ambiente di esecuzione, è quello di permettere alla piattaforma LMS di riconoscere i pacchetti SCORM, 'aprirli' e gestire l'interazione fra l'utente e i singoli SCO compresi nel pacchetto (SCO che l'utente dovrà utilizzare uno alla volta, sulla base delle regole di sequenziazione contenute nel manifest del pacchetto).

La comunicazione fra gli SCO e la piattaforma LMS avviene attraverso le API (Application Programming Interface) SCORM, ma una loro discussione richiederebbe competenze di programmazione in cui non ci inoltreremo in questa sede. Basti sapere che le API SCORM mettono a disposizione dei due protagonisti di questo dialogo – lo SCO e la piattaforma – delle funzioni o metodi che consentono lo scambio di dati e dunque un tracciamento di base della fruizione dello SCO da parte dei singoli utenti. SCORM comunque è uno standard abbastanza 'spartano' dal punto di vista del tracciamento, dato che la piattaforma può dialogare con lo SCO ma non con i singoli asset che lo compongono (lo SCO può comunque scambiare con la piattaforma valori relativi, ad esempio, ai risultati di test di autovalutazione somministrati all'utente, o bookmark che permettono di ritornare a punti specifici del contenuto).

Infine, la componente di *sequencing and navigation*, basata anch'essa sul lavoro di IMS, permette di organizzare i contenuti e le attività di apprendimento in alberi di attività, o *activity tree*. Abbiamo già accennato al ruolo delle attività, utilizzate per costruire un modello astratto del lavoro che ci si aspetta sia compiuto dal fruitore del pacchetto SCORM. E in effetti il modello di sequencing and navigation proposto da SCORM è – almeno per ora – totalmente orientato al fruitore, senza prendere in considerazione il possibile ruolo di altre figure (tutor, docenti ecc.). Attraverso l'applicazione delle funzioni di tracciamento già ricordate all'*activity tree* che viene creato per ogni utente, il sistema è in grado di tenere traccia delle attività avviate (in gergo SCORM, *attempted*) dall'utente, e di quelle che l'utente stesso ha concluso. Il modello permette inoltre di associare alle attività anche degli obiettivi di apprendimento (*learning objectives*).

Le regole per il sequencing delle attività, e per la navigazione al loro interno, sono piuttosto complesse, e non ce ne occuperemo in questa sede, anche perché avremo occasione nel corso del master di parlare specificamente di un modello di organizzazione e pianificazione dell'interazione fra agenti, attività di apprendimento e risorse di apprendimento per certi versi più potente di quello previsto da SCORM (il modello *Learning Design*, sviluppato anch'esso – potevate dubitarne? – da IMS).

### 9. SCORM: qualche riflessione

Come si vede, il modello SCORM è notevolmente complesso e strutturato, e ha il suo punto di forza nell'interoperabilità fra pacchetti di contenuti e piattaforme di gestione, garantita dal rispetto degli standard, nonché nella comodità pratica dei pacchetti di contenuto, che semplificano la distribuzione di materiali di apprendimento complessi e strutturati e ne permettono la fruizione all'interno di una piattaforma LMS, rendendo del tutto trasparente all'utente il livello della loro organizzazione interna e della loro interazione con la piattaforma.

SCORM è nato con sei obiettivi fondamentali:

- **accessibilità** dei contenuti di apprendimento, anche a distanza (garantita dall'impostazione strettamente web-oriented del modello)<sup>9</sup>;
- **adattabilità**, nel permettere la realizzazione di risorse e attività di apprendimento rispondenti alle necessità dei singoli fruitori e dei singoli progetti didattici (garantita dall'organizzazione fortemente modulare dei contenuti);
- **sostenibilità** (attraverso la riduzione dei tempi e dei costi di sviluppo e di fruizione, garantita dall'adozione di standard interoperabili e dalla conseguente possibilità di realizzare e utilizzare strumenti di sviluppo, nonché di operare una selezione fra numerose piattaforme didattiche compatibili);
- **durabilità** (garantita dall'adozione di modelli standard di organizzazione dei contenuti che consentono, ad esempio, di effettuare eventuali aggiornamenti senza doverne ridisegnare l'intera organizzazione interna);
- **interoperabilità** (possibilità di utilizzare qualunque pacchetto di contenuto e qualunque SCO conformi agli standard attraverso una pluralità di piattaforme, progettate indipendentemente dai contenuti e realizzate da altri soggetti);
- **riusabilità** (gli SCO e talvolta anche i singoli asset possono essere riutilizzati in situazioni diverse e in contesti diversi, conservandone l'aderenza agli standard).

---

<sup>9</sup> Va detto tuttavia che SCORM è stato criticato dal punto di vista dell'accessibilità dei contenuti, in particolare per il ricorso, nella definizione del run-time environment e delle relative API, al linguaggio di scripting *Java Script*, la cui interpretazione pone alcuni vincoli agli ambienti di fruizione effettivamente utilizzabili.

## Metadati per l'e-Learning: una introduzione

---

Tuttavia, nonostante i suoi indubbi meriti, SCORM è ben lontano dal rappresentare un modello davvero universale per l'organizzazione e fruizione via web di qualunque tipo di contenuti di apprendimento (uno scopo, a onor del vero, che lo standard non si è comunque mai posto). In particolare, SCORM è

- totalmente centrato sul fruitore, con pochissimi spazi per gestire all'interno del modello e in maniera organizzata l'intervento di figure come tutor e docenti nelle varie fasi della somministrazione dei contenuti, o per gestire forme di lavoro collaborativo sui contenuti stessi; da questo punto di vista, SCORM corrisponde a un paradigma didattico fortemente trasmissivo e lontano dalle prospettive della didattica collaborativa<sup>10</sup>;
- legato a una concezione assai chiusa dei contenuti di apprendimento distribuiti, con forti limiti nella gestione di forme di interazione davvero avanzate fra questi e l'utente finale da un lato, e fra i contenuti e il resto dell'universo informativo di rete dall'altro;
- legato a un modello fortemente strutturato e gerarchico di organizzazione dei contenuti, che può rivelarsi inadatto – ad esempio – in situazioni di apprendimento informale;
- fortemente dipendente dall'interazione con un LMS, un'altra caratteristica che può pregiudicarne l'uso in situazioni di apprendimento informale.

SCORM, insomma, dà il meglio di sé quando ci sono contenuti organizzati in forma relativamente chiusa e stabile, da trasmettere all'esterno fornendo agli utenti una piattaforma di distribuzione ma relativamente poca assistenza nella fruizione, che si prevede largamente autonoma e autogestita dal discente. L'importanza di queste situazioni, tuttavia, non va sottovalutata: può trattarsi infatti di un buon modello per il riuso esterno di contenuti di apprendimento prodotti da istituzioni formative che offrono al loro interno strategie didattiche e strumenti più complessi e partecipativi, ma che in questo modo possono aprire almeno una parte dei materiali prodotti anche alla fruizione autonoma da parte di utenti esterni (una situazione tipica, ad esempio, delle iniziative di Open Courseware, di cui parleremo in altra sede<sup>11</sup>).

### 10. Lo standard Dublin Core

Fin qui, il nostro percorso ci ha portato a esplorare sistemi di metadati nati e destinati specificamente al mondo dell'e-learning. Questa presentazione, però, non sarebbe completa senza dedicare qualche attenzione a uno standard di metadato nato in ambito bibliotecario, ma che ha influenzato ampiamente l'intero campo delle semantiche di rete (e non solo), come vedremo anche per quel che riguarda l'ambito della formazione. Lo standard al quale ci riferiamo si chiama *Dublin Core*, e propone un sistema di metadati che ha l'obiettivo di essere semplice e nel contempo sufficientemente generale (e se necessario espandibile), tanto da coprire un vasto insieme di situazioni e necessità.

---

<sup>10</sup> Si vedano al riguardo le considerazioni di A Fini e L. Vani in *Learning object e metadati. Quando, come e perché avvalersene*, Erickson, Trento 2004, pp. 146-151.

<sup>11</sup> Anche sulla base di queste considerazioni, il nostro master usa il formato SCORM per la distribuzione aperta all'esterno di parte dei propri contenuti didattici.

# Metadati per l'e-Learning: una introduzione

---

La *Dublin Core Metadata Initiative* (DCMI) è una delle prime iniziative organizzate di definizione di metadati nata per rispondere anche alle esigenze di standardizzazione e gestione automatica dei metadati legate allo sviluppo degli strumenti di rete e alla diffusione delle risorse informative in formato digitale. Il nome non deriva, come si potrebbe pensare, dalla capitale irlandese ma dalla città di Dublin nell'Ohio, dove ha sede il grande consorzio bibliotecario statunitense OCLC (Online Computer Library Center) e dove nel 1995 si è tenuta una conferenza che proprio l'OCLC ha promosso su questi temi.

L'esigenza principale di garantire la semplicità del sistema ha portato inizialmente all'identificazione di un insieme di quindici metadati principali, che costituiscono il *Simple Dublin Core Metadata Element Set* (DCMES). Li riportiamo di seguito accompagnati, con poche modifiche, dalla descrizione che ne viene data dalla voce italiana di Wikipedia 'Dublin Core'<sup>12</sup>, a sua volta ripresa in gran parte dalla descrizione italiana dello standard, curata dall'ICCU (l'Istituto centrale per il catalogo unico delle biblioteche italiane e per le informazioni bibliografiche)

- **Titolo (Title)**. Nome dato alla risorsa. In particolare, un Titolo sarà un termine con il quale la risorsa è formalmente conosciuta.
- **Autore (Creator)**. Entità che ha la responsabilità principale della produzione del contenuto della risorsa. Esempi di Autore possono essere una persona, un'organizzazione o un servizio responsabili del contenuto intellettuale della risorsa.
- **Soggetto (Subject)**. Argomento principale della risorsa. In particolare un Soggetto può essere espresso da parole o frasi chiave, o da codici di classificazione che descrivono l'argomento della risorsa. Solitamente questi termini vengono scelti tra i valori di un vocabolario controllato o di uno schema di classificazione formale.
- **Descrizione (Description)**. Descrizione del contenuto della risorsa. Testo descrittivo libero che può includere un riassunto analitico, un indice, o una rappresentazione grafica del contenuto.
- **Editore (Publisher)**. Entità responsabile della pubblicazione della risorsa. Esempi di Editore possono essere una persona, un'organizzazione o un servizio che si occupa di rendere disponibile la risorsa nella sua forma attuale.
- **Autore di contributo subordinato (Contributor)**. Entità responsabile della produzione di un contributo al contenuto della risorsa. Esempi di Autore secondario includono una persona, un'organizzazione o un servizio che contribuiscono alla produzione della risorsa.
- **Data (Date)**. Data associata ad un evento del ciclo di vita della risorsa. Normalmente la data è associata al momento di creazione o di disponibilità della risorsa e viene indicata attraverso una stringa di 8 caratteri nella forma YYYY-MM-DD, come definita nel profilo dello standard ISO 860190. In questo schema l'elemento data 1994-11-05 corrisponde al 5 novembre 1994.

---

<sup>12</sup> [http://it.wikipedia.org/wiki/Dublin\\_Core](http://it.wikipedia.org/wiki/Dublin_Core); consultazione, 14 aprile 2008.

## Metadati per l'e-Learning: una introduzione

---

- **Tipo (Type).** Natura o genere del contenuto della risorsa. L'elemento "Tipo" include termini che descrivono categorie generali, funzioni, generi, o livelli di aggregazione per contenuto presi generalmente da un vocabolario controllato.
- **Formato (Format).** Manifestazione fisica o digitale della risorsa. Normalmente l'elemento "Formato" può includere il tipo di supporto o le dimensioni, ossia grandezza e durata, della risorsa. Format può essere usato per determinare il software o l'hardware necessari alla visualizzazione o all'elaborazione della risorsa.
- **Identificatore (Identifier).** Riferimento univoco alla risorsa nell'ambito di un dato contesto. Solitamente le risorse vengono identificate per mezzo di una sequenza di caratteri alfabetici o numerici secondo un sistema di identificazione formalmente definito. Esempi di tali sistemi di identificazione includono l'Uniform Resource Identifier (URI) (incluso l'Uniform Resource Locator (URL)), il Digital Object Identifier (DOI) e l'International Standard Book Number (ISBN).
- **Fonte (Source).** Riferimento a una risorsa dalla quale è derivata, in tutto o in parte, la risorsa in oggetto.
- **Lingua (Language).** Lingua del contenuto intellettuale della risorsa. Per i valori dell'elemento Lingua si utilizza un codice di linguaggio, seguito opzionalmente da un codice di paese, entrambi su due caratteri. Ad esempio "it" per l'italiano o "en-uk" per l'inglese usato nel Regno Unito.
- **Relazione (Relation).** Riferimento ad una risorsa correlata.
- **Copertura (Coverage).** Estensione o scopo del contenuto della risorsa. Normalmente Copertura include la localizzazione spaziale (il nome o le coordinate geografiche di un luogo), il periodo temporale (l'indicazione di un periodo, una data o una serie di date) o una giurisdizione (ad esempio il nome di un'entità amministrativa).
- **Gestione dei diritti (Rights).** Informazione sui diritti esercitati sulla risorsa. Normalmente un elemento "Diritti" contiene un'indicazione sulla gestione dei diritti sulla risorsa, o un riferimento al servizio che fornisce questa informazione. Questo campo comprende gli Intellectual Property Rights (IPR), il copyright, e vari diritti di proprietà. Se l'elemento Rights è assente, non si può fare alcuna ipotesi sui diritti della risorsa.

Questo insieme di metadati (ciascuno dei quali è opzionale, ripetibile e ulteriormente specificabile attraverso attributi) è stato poi via via raffinato, cercando di conciliare l'esigenza di semplicità con quella di una maggiore potenza espressiva. Il risultato è il Dublin Core Qualificato (*Qualified Dublin Core*), che aggiunge ai quindici elementi originati alcuni nuovi elementi (al momento sette, uno solo dei quali tuttavia – l'elemento 'Audience' – è considerato elemento 'raccomandato', mentre gli altri sei sono considerati 'conformi', status attribuito a elementi suggeriti dalla comunità degli implementatori e di cui viene riconosciuta la conformità al modello astratto DCMI). Quattro di questi nuovi elementi sono rilevanti ai nostri scopi ( gli altri tre, *accrualMethod*, *accrualPeriodicity*, *accrualPolicy*, sono per noi meno direttamente rilevanti e si riferiscono alle modalità di accrescimento della risorsa nel tempo):

# Metadati per l'e-Learning: una introduzione

---

- **Audience.** Una classe di utenti – o, in generale, di entità – a i quali la risorsa è destinata o per i quali può essere utile.
- **Provenance.** Indicazione di cambiamenti significativi nella proprietà o custodia della risorsa dopo la sua creazione.
- **RightsHolder.** La persona o organizzazione che detiene o gestisce i diritti sulla risorsa.
- **InstructionalMethod.** Un processo che si propone di generare conoscenza, capacità e competenze, e che la risorsa è destinata a supportare. L'instructional method comprende tipicamente modelli di presentazione di materiali formativi, schemi di interazione fra discente e discente e fra discente e docente, e meccanismi per la misurazione di livelli di apprendimento individuali o di gruppo. Può includere tutti gli elementi di un processo didattico o di apprendimento, dalla pianificazione all'implementazione, fino alla valutazione e al feedback.

In particolare, come avrete visto, l'elemento InstructionalMethod è il solo che fa riferimento esplicito al carattere di risorsa di apprendimento che una risorsa informativa può possedere (non a caso, la sua introduzione è stata proposta – nel 2004 – dal Dublin Core Education Working Group, il gruppo di lavoro che si occupa specificamente delle applicazioni di Dublin Core all'ambito educativo). In effetti Dublin Core nasce come strumento per la descrizione generale di risorse informative, e non per la descrizione specifica di particolari tipologie di risorse (come potrebbero essere appunto le risorse legate all'apprendimento). La proposta del nuovo elemento, tuttavia, intende proprio sottolineare il ruolo e rilievo particolari che le risorse di apprendimento possono avere nell'ambito del campo più generale delle risorse informative, e l'importanza del segnalarne esplicitamente tale natura. Va considerato comunque che, pur essendo 'conforme' allo standard, l'elemento InstructionalMethod non è supportato da tutti i sistemi di metadattazione Dublin Core, ed è del resto piuttosto povero nel momento in cui si volesse operare una metadattazione specifica di risorse di apprendimento.

L'importanza del Dublin Core Qualificato non è però tanto negli elementi aggiuntivi, quanto nella possibilità di raffinare l'uso di ciascuno degli elementi attraverso *qualificatori*, che aumentano considerevolmente la potenza espressiva del linguaggio, permettendo di specificare in maniera più precisa la natura dei metadati utilizzati. Così, ad esempio, l'elemento 'relation' può essere specificato indicando se la relazione a cui facciamo riferimento è una relazione di inclusione (*IsPartOf*), di versione (*IsVersionOf*), di riferimento (*IsReferencedBy*), e così via.

## 11. Dublin Core e il mondo dell'e-learning. Il progetto GEM

Da quanto abbiamo esposto finora, sarà risultato chiaro che i metadati Dublin Core, pur offrendo una cornice generale di riferimento per la metadattazione di un larghissimo campo di risorse informative, non si propongono come strumento specifico per la metadattazione strutturata di risorse legate al campo dell'e-learning. In effetti, i metadati DC – pur potendo essere specificati attraverso qualificatori – non hanno né la specificità né la strutturazione gerarchica che abbiamo invece riscontrato nei metadati LOM.

## Metadati per l'e-Learning: una introduzione

---

I due insiemi di metadati hanno dunque scopi diversi. Ma sarebbe assolutamente sbagliato sottovalutare il rilievo che i metadati Dublin Core possono avere anche nel campo dell'e-learning. Infatti, le risorse di apprendimento non costituiscono (e non devono costituire) un recinto chiuso e separato rispetto ad altre tipologie di risorse informative. Piuttosto, per loro stessa natura, debbono poter fare riferimento, integrarsi, interagire con l'intero universo informativo che le circonda, e dunque con molte altre tipologie di risorse di rete. In molte situazioni, i metadati DC possono dunque costituire un elemento essenziale per garantire l'apertura e l'interoperabilità anche delle risorse di apprendimento.

È inutile nascondere, tuttavia, che questa situazione crea un problema di non poco conto: dato che la metadattazione LOM e quella Dublin Core sono *strutturalmente* diverse, come garantirne la necessaria interoperabilità? La soluzione più semplice è quella di 'mappare' – per quanto possibile – i metadati utilizzati da uno standard su quelli utilizzati dall'altro. Ma proprio per la differenza strutturale fra i due modelli questo non è sempre facile, e i risultati non sono sempre soddisfacenti. Il lavoro sull'interoperabilità di questi due standard dovrà dunque indubbiamente continuare in futuro.

Va ricordato, a questo riguardo, che esistono proposte di definizione di standard specifici di metadati legati al campo delle risorse di apprendimento nati anche nel 'mondo' Dublin Core, e che adottano soluzioni e formalismi più vicini a quelli di Dublin Core. In particolare, già nel 1996 – e dunque prima ancora dello sviluppo di LOM – il Dipartimento dell'educazione statunitense aveva avviato un progetto di repository di metadati relativo a risorse educative distribuite via web. Il progetto, denominato GEM (*Gateway to Educational Materials*), prevedeva un proprio insieme di metadati, allo scopo di facilitare la descrizione e il recupero di tali risorse. Di fatto, la sigla GEM identificava sia il Consorzio impegnato a lavorare al progetto, sia il repository, sia l'insieme di metadati proposto e le specifiche per la sua applicazione.

Già nella sua prima versione, e ancor più chiaramente nella successiva versione 2.0, le specifiche GEM si collegavano in maniera diretta al modello Dublin Core. L'obiettivo era infatti quello di creare un insieme di metadati ricco e specificamente rivolto alla gestione di risorse di apprendimento, ma pienamente allineato ai principi generali della metadattazione Dublin Core, dalla quale erano ripresi meccanismi come quello dei qualificatori.

In particolare, GEM 1.0 aggiungeva ai quindici elementi di base da cui era partito Dublin Core un insieme di otto nuovi elementi: *audience*, *catalogingAgency*, *duration*, *essentialResources*, *educationalLevel*, *pedagogy*, *qualityAssessment* (in seguito abbandonato per il suo carattere fortemente soggettivo), ed *educationalStandards*. Non entriamo qui nei dettagli relativi alle caratteristiche di questi metadati, limitandoci a ricordare come la già ricordata accettazione di *audience* e (pur se attraverso la sola dichiarazione di conformità) di *educationalLevel* fra gli elementi Dublin Core, e l'abbandono di *qualityAssessment*, abbiano portato a prevedere nella versione 2.0 delle specifiche solo cinque elementi aggiuntivi a quelli Dublin Core, trasferendo dunque anche al campo delle risorse educative l'impostazione 'essenziale' e la ricerca di semplicità che caratterizzano tale modello.

Il progetto GEM – e altri progetti legati al mondo Dublin Core, ma non necessariamente compatibili fra loro, come quello EdNA dell'Educational Network australiano – offrono il vantaggio di una migliore interoperabilità con il mondo delle risorse informative esterne al campo specifico dell'e-Learning (GEM è

anche stato fra i primi progetti a riconoscere le potenzialità del meccanismo dell'harvesting, del quale abbiamo già parlato nel discutere le caratteristiche dei repository), ma di fatto hanno avuto un successo abbastanza limitato. Nel settore dell'e-Learning, l'uso di metadati LOM sembra al momento assai più diffuso, probabilmente anche perché la loro inclusione nelle specifiche SCORM ne ha facilitato l'adozione al momento della effettiva costruzione di pacchetti di contenuti.

L'esistenza di troppe specifiche diverse relative ai metadati da utilizzare – nessuna delle quali in fondo pienamente soddisfacente – pone comunque sicuramente un problema che il mondo dell'e-Learning dovrà affrontare e cercare di risolvere, se si vuole passare dallo stadio delle sperimentazioni pionieristiche a quello di un settore pienamente maturo, in cui l'idea di interoperabilità e accessibilità diffusa dei contenuti si trasformi da uno slogan in una pratica davvero generalizzata.

### 12. Metadati in pratica

Fin qui, abbiamo cercato di capire insieme cosa siano i metadati, perché siano tanto importanti, e quali siano i principali standard e modelli esistenti in questo settore, in particolare per quanto riguarda il campo dell'e-learning. Immagino però che a questo punto molti di voi si saranno posti un interrogativo perfettamente naturale: ma cosa possiamo o dobbiamo fare, concretamente, con i metadati? Ci sono degli strumenti 'facili' per lavorare con standard che sembrano comunque complessi e poco amichevoli?

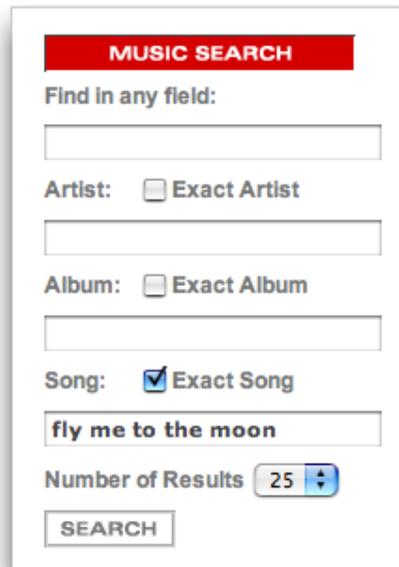
Dal punto di vista pratico, lavorare con i metadati vuol dire saperli usare al meglio a) nella *ricerca* di risorse informative (ad esempio, all'interno di un repository); b) nella *gestione* di risorse informative (ad esempio, nell'amministrazione di un LMS); c) nella *creazione* di risorse informative (ad esempio, nella creazione di pacchetti di risorse di apprendimento).

Per quanto riguarda la *ricerca*, l'uso esplicito di metadati riguarda normalmente quattro tipologie di strumenti: 1. le funzioni di ricerca avanzata di motori di ricerca generalisti, come Google; 2. i motori di ricerca settoriali o tematici che lavorano su basi di (meta)dati organizzate (a cominciare dagli OPAC, i cataloghi on-line delle biblioteche); 3. i repository<sup>13</sup>; 4. i siti e strumenti che ricorrono a etichettature libere (social tagging).

Nei primi tre casi, vengono tipicamente utilizzate maschere organizzate secondo il principio della cosiddetta 'Query By Example': i campi presenti nella maschera corrispondono ai principali metadati presenti nel sistema. In questo modo, l'utente è portato a utilizzare nella ricerca, in maniera guidata e intuitiva, alcune fra le tipologie di metadati utilizzate nella descrizione delle risorse indicizzate (quelle considerate più utili per il reperimento efficace del documento, alla luce delle necessità e tipologie di ricerca più comuni).

---

<sup>13</sup> Le tipologie 2 e 3 potrebbero essere considerate insieme; le abbiamo distinte soprattutto per praticità, dato che la gestione dei metadati all'interno di un repository che mette a disposizione anche le relative risorse informative primarie presenta problemi specifici rispetto alla gestione di una base dati composta unicamente da metadati, come accade, ad esempio, nel caso dell'OPAC di una biblioteca.



**MUSIC SEARCH**

Find in any field:

Artist:  Exact Artist

Album:  Exact Album

Song:  Exact Song

fly me to the moon

Number of Results 25

SEARCH

Figura 6 - QueryBy Example: un esempio (Gracenote)

Provate, come esercizio, a esaminare le interfacce di ricerca avanzata di alcuni fra gli strumenti che usate più frequentemente, e a ricavarne indicazioni sui metadati utilizzati. Domandatevi anche, di volta in volta, se si tratti di metadati immessi direttamente dall'autore della risorsa informativa, di metadati immessi dal responsabile della sua indicizzazione, o di metadati estratti automaticamente dal motore di ricerca analizzando alcune caratteristiche della risorsa informativa stessa. Scoprirete che tutte e tre le tipologie sono abbastanza frequenti.

Un altro esercizio utile è esaminare le caratteristiche di interfaccia delle maschere di ricerca: in che modo vengono proposti i metadati su cui effettuare la ricerca? Si tratta di metadati 'piatti', o in qualche modo strutturati (ad esempio in un sistema di categorie e sottocategorie)? Le scelte dei valori da utilizzare nella ricerca sono libere o guidate?<sup>14</sup>

Infine, consigliamo di provare il funzionamento degli strumenti di ricerca presenti in siti che fanno uso di social tagging: un buon punto di partenza è il già ricordato [del.icio.us](http://del.icio.us) (<http://del.icio.us>), che come si è accennato permette anche una organizzazione minimale dei tag in cluster. Vedrete anche che [del.icio.us](http://del.icio.us) – come molti altri siti web 2.0 – fa uso delle cosiddette *tag cloud*, che rappresentano visivamente le etichette utilizzate per classificare un certo contenuto o insieme di contenuti, o le etichette più frequentemente associate all'etichetta considerata (le dimensioni del carattere utilizzato per ogni etichetta ne esprimono la frequenza relativa)<sup>15</sup>.

<sup>14</sup> Per alcuni esempi, si veda la pagina <http://www.searchtools.com/info/faceted-metadata.html>.

<sup>15</sup> Per un approfondimento su social tagging e tag cloud si veda A. Falossi, *Tagging: una questione di etichette*, all'interno del sito *Il mestiere di scrivere*: <http://www.mestierediscrivere.com/testi/tagging.htm> (ultima consultazione 14 aprile 2008)



Figura 7 – Un altro esempio di tag cloud strutturata in del.icio.us

Come abbiamo visto, una seconda situazione di uso concreto dei metadati riguarda chi ha la responsabilità della *gestione* di risorse informative complesse (ad esempio, un LMS o un repository di qualsiasi genere). In questo caso è difficile dare ricette specifiche, perché molto dipende dallo strumento concretamente utilizzato e dalle sue caratteristiche. Le indicazioni che possiamo dare sono dunque solo piuttosto generali, e si traducono in alcune raccomandazioni. In primo luogo, al momento di scegliere uno strumento – di qualsiasi natura – destinato alla gestione di risorse informative, chiedetevi sempre:

- quali tipologie di metadati è opportuno poter impiegare, relativamente alle risorse informative che si devono gestire?
- lo strumento che stiamo considerando permette effettivamente di utilizzare i metadati necessari? In base a quali standard, e attraverso quale interfaccia?
- è facile aggiungere o modificare i metadati immessi? Chi può farlo, e come? Sono disponibili strumenti di controllo delle autorizzazioni relative alla modifica e gestione dei metadati?
- i metadati sono utilizzati efficacemente dalle funzionalità di ricerca disponibili?
- lo strumento permette di scambiare (importazione e/o esportazione) i metadati con altri strumenti dello stesso tipo (ad esempio attraverso meccanismi di harvesting)?
- è possibile utilizzare forme di social tagging? In caso affermativo, sono alternative o complementari rispetto a metadazioni strutturate?

È inoltre sempre opportuno verificare quali scelte abbiano fatto, relativamente alla gestione dei metadati, altri progetti che utilizzino i nostri stessi strumenti, così come è sempre opportuno verificare se e come

## Metadati per l'e-Learning: una introduzione

---

funzionino altri progetti che – pur avendo a che fare con la gestione di risorse informative non troppo lontane da quelle con le quali dobbiamo lavorare noi – abbiano scelto di utilizzare strumenti diversi dai nostri.

Infine, due parole sulla *creazione* di contenuti accompagnati da metadati. Come si fa per aggiungere metadati a una risorsa informativa? In linea generale, abbiamo tre possibilità, che possono anche essere utilizzate insieme: 1. associare i metadati alla nostra risorsa informativa all'interno di un repository, o comunque attraverso strumenti esterni alla risorsa informativa stessa (pensate nuovamente al modello rappresentato dalle schede di una biblioteca); 2. 'avvolgere' i metadati attorno alla risorsa informativa stessa creando un 'pacchetto' (*content package*, o *archive file*) che ingloba contenuto e metadati, e 3. integrare i metadati all'interno della risorsa informativa primaria (come avviene ad esempio nel caso dei metadati inseriti all'interno di una pagina HTML).

In questa sede ci interessa soprattutto la seconda possibilità, che riconoscerete essere quella alla base dei content package SCORM e che corrisponde di norma alla creazione di un file (*manifest*) XML contenente i metadati, che viene poi archiviato assieme alla risorsa informativa utilizzando uno strumento di impacchettamento e compressione come il diffuso formato ZIP<sup>16</sup>. In realtà, comunque, la differenza fra la seconda e la terza possibilità è più sfumata di quanto non potrebbe sembrare: molto spesso, infatti, quello che a noi sembra un file unico corrisponde in realtà a un pacchetto; solo che aprire quel pacchetto non è compito nostro ma dei programmi che gestiscono quel particolare tipo di file, e quindi noi siamo portati a considerare automaticamente il file stesso come un oggetto informativo unitario.

In ogni caso, sia essa realizzata all'interno di un singolo file o utilizzando un formato di pacchetto, la stretta associazione fra contenuti informativi e metadati è molto più diffusa di quanto non si possa pensare. Ad esempio, anche una normale fotografia digitale 'ingloba' all'interno del file una serie di metadati informativi (data e ora dello scatto, il tipo di macchina fotografica utilizzata, l'ISO, e, volendo, perfino le coordinate geografiche esatte del luogo in cui la foto è stata scattata... provate a cercare queste informazioni navigando in un album fotografico su Flickr). Metadati, 'interni' o 'di pacchetto'<sup>17</sup>, sono poi contenuti nei file creati dai principali programmi di videoscrittura (dopo aver aperto un documento, provate a esaminare la voce 'proprietà' presente in genere nel menu 'file': quali metadati permette di consultare o di aggiungere?), e naturalmente, come abbiamo ricordato, metadati interni sono contenuti nei file creati attraverso linguaggi di marcatura come HTML.

In questo contesto, l'uso dei formati di pacchetto ha comunque un enorme vantaggio: permette ai metadati di accompagnare l'informazione primaria (all'interno del pacchetto) restandone tuttavia distinti, e disponibili quindi anche per un uso indipendente.

Ma come si costruisce, concretamente, un pacchetto di contenuti (ad esempio un pacchetto SCORM)?

---

<sup>16</sup> Inizialmente, l'uso del formato ZIP è stato abbastanza criticato, perché si trattava di uno standard proprietario; la sua diffusione e la nascita di algoritmi di compressione pubblici compatibili con il formato ha in parte superato il problema, anche se il lavoro per la creazione di uno standard ISO relativo al formato ZIP è ancora in corso.

<sup>17</sup> Il diffuso software aperto *Open Office* è stato fra i primi programmi ad utilizzare un formato di pacchetto per la gestione dei propri documenti.

## Metadati per l'e-Learning: una introduzione

In genere, per la costruzione di pacchetti che uniscono risorse informative primarie e metadati si utilizzano appositi programmi (*authoring tool*), che consentono sia di assemblare i contenuti, strutturandoli e permettendo di indicare, ad esempio, l'ordine nel quale devono essere presentati all'utente, sia di aggiungere i metadati. Alcuni dei metadati inseriti saranno il risultato della modalità di organizzazione scelta per i contenuti, e saranno dunque creati automaticamente dal programma, mentre altri saranno inseriti esplicitamente dal creatore del pacchetto, in genere attraverso apposite maschere di immissione.

Fra gli esempi di authoring tool più utilizzati nell'ambito dell'e-learning, meritano di essere ricordati due programmi disponibili gratuitamente e distribuiti in forma aperta: *eXeLearning* e *Reload*.

eXeLearning, o più semplicemente eXe (acronimo di 'e-Learning XHTML editor'), è un programma open source realizzato in Nuova Zelanda e divenuto ora parte delle attività del progetto CORE Education (<http://core-ed.net/>), e può essere scaricato dall'indirizzo <http://exelearning.org/>. È, insieme, un editor HTML (o meglio, XHTML: la versione di HTML pienamente compatibile con XML), che permette di creare in maniera assistita pagine di contenuto (ad esempio, un tutorial da inserire in rete), e uno strumento di metadatazione che permette di aggiungere ai contenuti creati tutti i metadati necessari e di esportarli sotto forma di pacchetto SCORM. In sostanza, eXeLearning permette di creare contenuti di apprendimento che hanno la forma di un mini-sito web autosufficiente: sarà poi l'utente a decidere se il risultato dovrà essere pubblicato in rete per la navigazione attraverso un normale browser, o 'impacchettato' in formato SCORM e distribuito attraverso una piattaforma, ad esempio all'interno di un LMS.

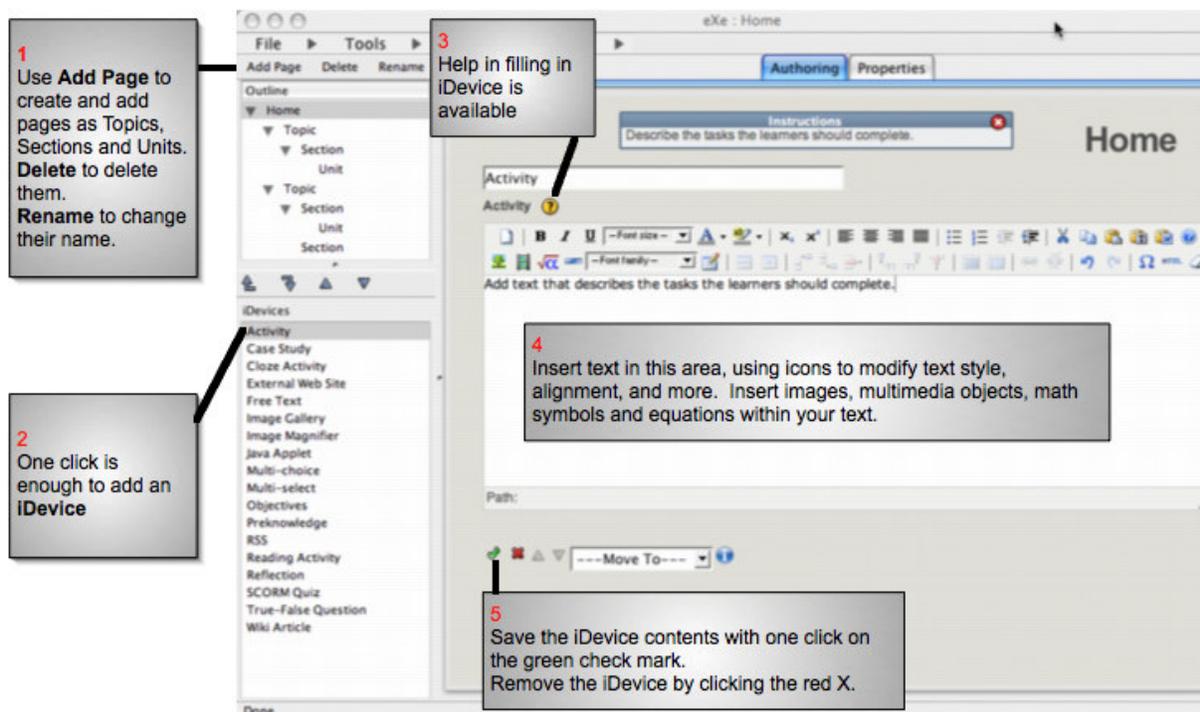


Figura 8 - eXeLearning: l'interfaccia dell'editor XHTML (da <http://exelearning.org/AuthoringScreenshots>)

## Metadati per l'e-Learning: una introduzione

---

Da segnalare anche che eXeLearning ha sperimentato in passato la possibilità di creare pacchetti di contenuto utilizzabili anche su dispositivi mobili (per l'esattezza, sull'iPod<sup>18</sup>): una funzionalità sempre più importante per gli authoring tool nel campo dell'e-Learning, sulla quale avremo occasione di tornare in altra sede.

Oltre che nel sito del progetto (che fornisce anche diversi tutorial video, in lingua inglese, molti dei quali disponibili anche su YouTube), una buona raccolta di risorse relative a eXe è disponibile sul sito Wikieducator, all'indirizzo [http://www.wikieducator.org/Training\\_resource\\_development\\_for\\_eXe](http://www.wikieducator.org/Training_resource_development_for_eXe); una traduzione italiana del manuale è disponibile, grazie all'eccellente lavoro svolto nell'ambito dell'edizione 2006-2007 del master dal nostro ex-corsista Onorio Zaralli, all'indirizzo [http://www.wikieducator.org/images/5/5d/EXe\\_Manual\\_Italian.pdf](http://www.wikieducator.org/images/5/5d/EXe_Manual_Italian.pdf).

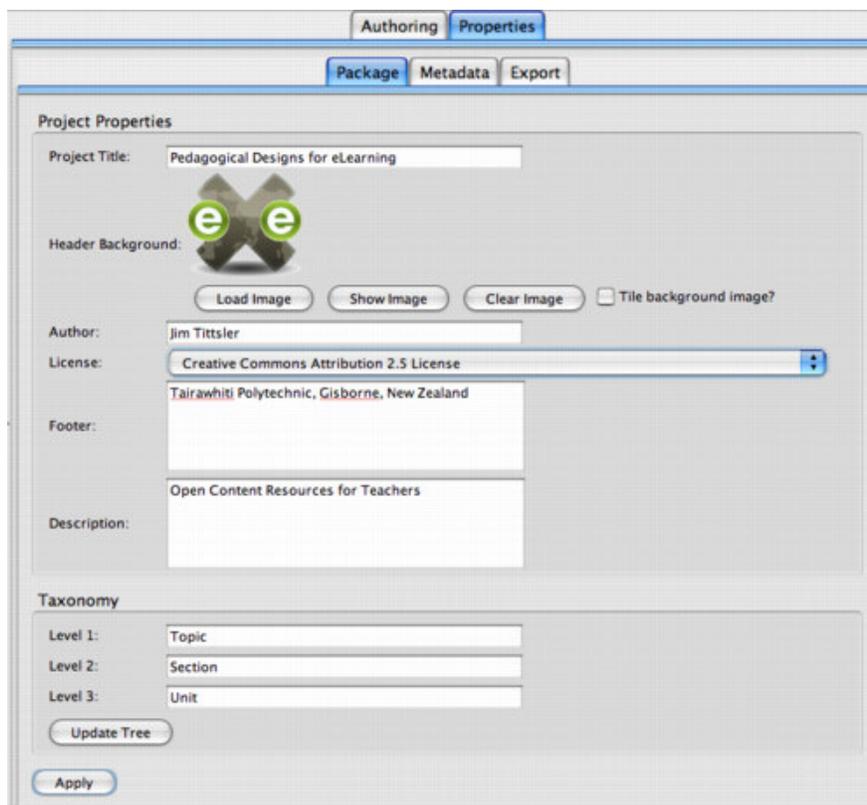
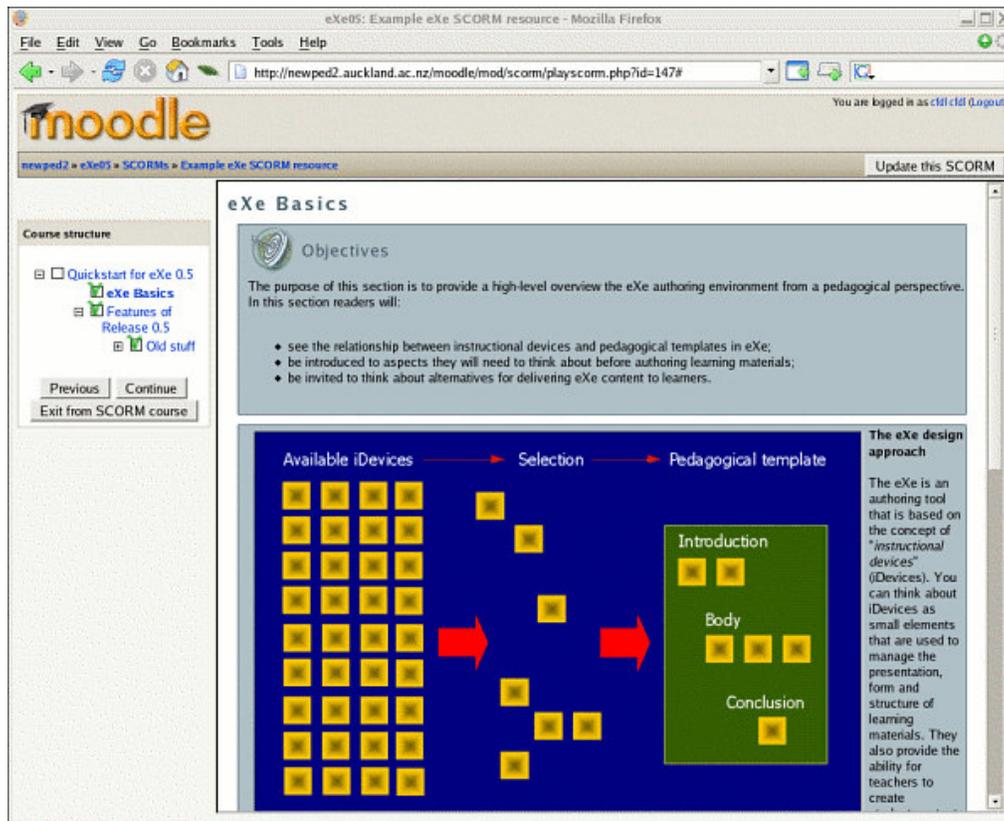


Figura 9 - eXeLearning: maschera per l'aggiunta di metadati a un pacchetto SCORM (da <http://exelearning.org/AuthoringScreenshots>)

<sup>18</sup> Il format utilizzato è quello iPod Notes, che presenta tuttavia notevoli limitazioni. Cfr. al riguardo <http://exelearning.blogspot.com/2007/04/exe-for-mobile-content.html>.



**Figura 10 - Un content package SCOIRM realizzato in eXeLearning e visualizzato all'interno di una piattaforma Moodle (da [http://exelearning.org/ContentPackaging/SCORM\\_Import\\_Screenshots](http://exelearning.org/ContentPackaging/SCORM_Import_Screenshots))**

Mentre eXe permette sia di creare contenuti sia di raccogliarli in un pacchetto SCORM, Reload è solo uno strumento di impacchettamento e aggiunta di metadati: i contenuti devono prima essere creati indipendentemente, e Reload permette poi di organizzarli, descriverli attraverso i metadati necessari e esportarli in forma di pacchetto. Questo consente a Reload di creare pacchetti a partire da qualsiasi tipologia di contenuti.

Il programma – che lavora in ambiente Java e necessita dunque di un computer su cui tale ambiente sia installato – può essere scaricato dal sito <http://www.reload.ac.uk/>, e allo stesso indirizzo è disponibile anche un 'language pack' per utilizzarne l'interfaccia in una versione italiana.

Il sito di distribuzione mette a disposizione anche diversi altri strumenti correlati: un player SCORM indipendente, che può essere utilizzato per testare il funzionamento del pacchetto creato anche al di fuori da una piattaforma LMS, nonché un editor e un player destinati specificamente alla creazione di progetti di Learning Design.

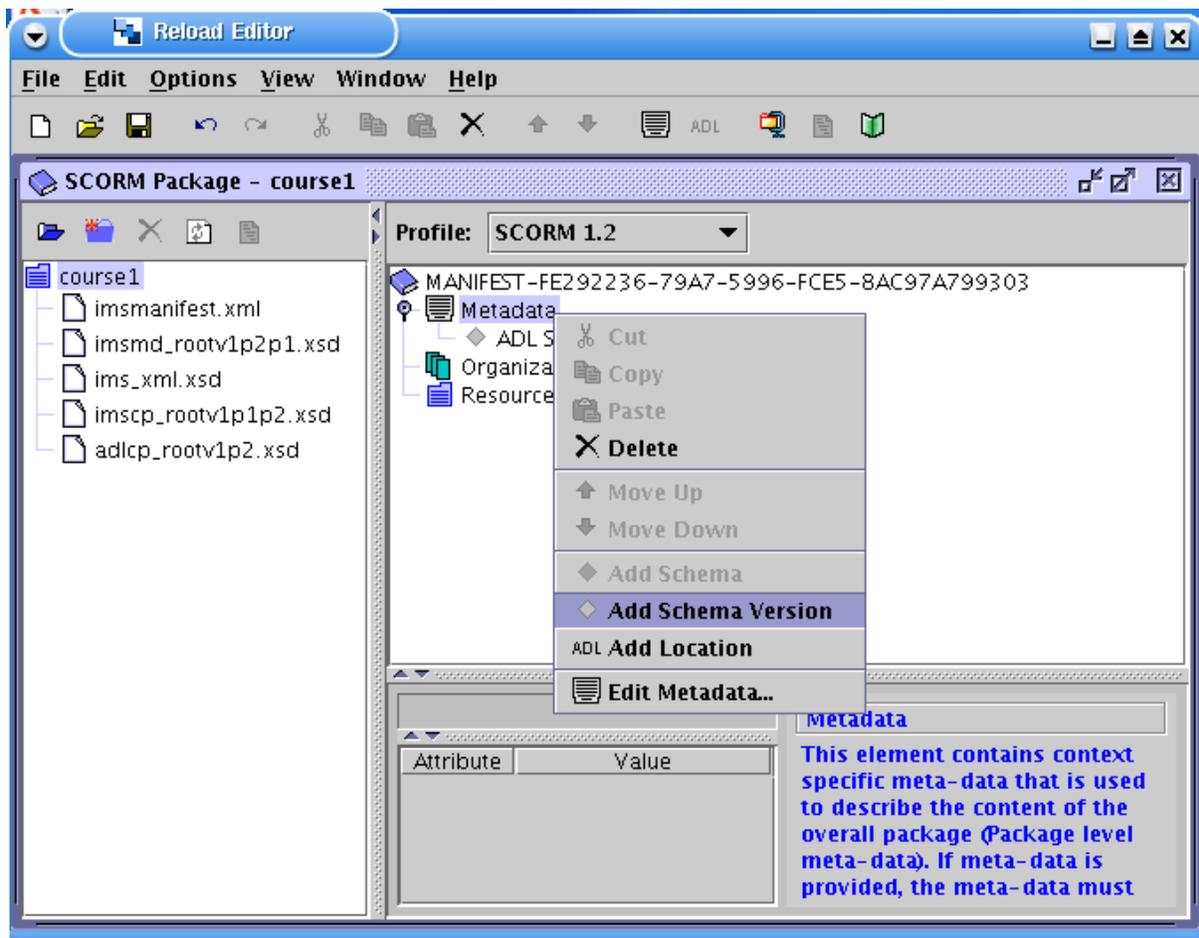


Figura 11 - Reload: l'interfaccia del programma

L'uso di Reload (come del resto quello di molti strumenti di questo genere) non è proprio immediato, ma con un po' di buona volontà e l'aiuto della documentazione disponibile non si tratta davvero di un compito proibitivo. Tenete conto che la creazione dei primi pacchetti porta sempre via parecchio tempo, ma che dopo aver acquistato un po' di familiarità con lo strumento tutto diventa più semplice.

Una raccolta di buoni tutorial animati sull'uso di Reload è comunque disponibile alla pagina <http://www.reload.ac.uk/links.html> - e nel nostro piccolo intendiamo anche noi mettere a disposizione presto un tutorial sull'uso di questo strumento, interamente in italiano.

Infine, una citazione merita anche un prodotto che è il risultato di un progetto tutto italiano: *FreeLOms* (FreeLearning Object Management System). FreeLOms – realizzato nell'ambito del Progetto SLOOP – è innanzitutto un ambiente di scambio e di produzione collaborativa di Learning Object, ma fra gli strumenti messi a disposizione ci sono anche procedure semi-automatiche per la creazione di un pacchetto SCORM, utilizzando risorse già presenti in FreeLOms, presentazioni video in formato PPT o pagine web. Permette inoltre di creare modelli (template) di metadati da applicare a risorse didattiche. Il sito di riferimento (che richiede una registrazione gratuita) è <http://www.freeloms.org/>.