

# Un Modello di Sviluppo di Sistemi Informativi

Carlo.Meghini@isti.cnr.it  
Paola.Venerosi@isti.cnr.it

## Fase III – Sistemi informativi avanzati ed interoperabilita’: modelli per l’organizzazione spazio-temporale della conoscenza

### *Razionale*

Dopo aver preso in esame una visione filosofica della realta’ basata sui concetti di azioni ed eventi, ed una teoria che intende calcolarne gli effetti basandosi sulla nozione di *fluente* nel contesto spazio-temporale, avviamo ora la terza fase della ricerca.

Essa e’ dedicata a sistemi informativi avanzati, modellati sul concetto di evento, sull’introduzione di entita’ temporali e su ontologie finalizzate alla interoperabilita’. Vengono analizzati i modelli CIDOC/CRM e ABC considerati significativi per lo sviluppo di applicazioni informatiche.

### Parte I - CIDOC-CIRM, un approccio ontologico alla interoperabilita’ semantica dei metadati.

#### *CIDOC-CIRM caratteristiche ed obiettivi*

CIDOC-Conceptual Reference Model: un modello concettuale per la interoperabilita’ semantica tra sistemi informativi [1] che trattano problemi della conoscenza come processi temporali. Un modello che da’ la possibilita’ di registrare ogni cambiamento degli oggetti sotto osservazione e di ricostruirne la storia, capace inoltre di associare contemporaneamente alla proprieta’ di un oggetto valori diversi al fine di preservare opinioni contraddittorie.

Le funzioni di interoperabilita’ semantica sulle quali il modello si concentra derivano da considerazioni sulla specifica natura della conoscenza storica e culturale, ben espresse con esempi che si riferiscono all’area storica, archeologica e museale.

E’ da queste osservazioni dirette e minuziose che risalta per es. l’importanza di documentare la datazione, o l’attribuzione di un oggetto ad un specifico periodo storico come fatti su cui si ragiona, piu’ che come dati inconfutabili. E’ la logica del calcolo proposizionale ad essere contraddetta, perche’ i valori di “falso” e di “vero” non sono sempre attribuibili ai termini di una argomentazione.

Il modello si configura come strumento di conoscenza, quindi di supporto al ragionamento, “piu’ che strumento di controllo e predizione come avviene nell’ambito della conoscenza ingegneristica”.

Il modello che M. Doerr ci presenta e’ l’esperienza a piu’ elevato contenuto umanistico nella area della “cultural heritage”, data l’osservanza scrupolosa della metodologia usata nella ricerca storica. L’interoperabilita’ non e’ diretta a normalizzare la diversita’ delle strutture dati e dei sistemi terminologici ma, concepita come interoperabilita’ semantica mira a esplicitare quell’intreccio di relazioni che caratterizzano la parte piu’ ricca di semantica della struttura dati, cioe’ gli attributi e le proprieta’.

I tentativi finora fatti in tema di interoperabilita’ semantica si sono concentrati sullo sviluppo e standardizzazione di strutture dati, sistemi terminologici condivisi, e schemi comuni di metadati (Dublin Core Element Set [2], Library of Congress Subject Headings LCSH [3], Art & Architecture Thesaurus AAT [4]).

L’esito ha comportato lo sviluppo di un discreto numero di standard, unito ad una pletora di metadati proprietari. I modelli sono in genere singolarmente comprensibili, ma l’interoperabilita’ raggiunta ha richiesto semplificazioni che hanno causato o perdita di informazioni oppure un aumento dell’ambiguita’ semantica.

Questo esito non e’ pero’ in contraddizione con gli scopi che tali tentativi si prefiggevano. Infatti, la diversita’ delle strutture dati o metadati e’ stata funzionale al buon accesso all’informazione nei vari campi applicativi, ed e’ servita da guida a chi nel campo dell’archiviazione doveva selezionare

le informazioni da documentare; l'uso di schemi comuni ha poi facilitato l'accesso alla crescente quantità di informazione disponibile su Internet.

Il modello CIDOC-CRM, insieme ad ABC [5] segna una svolta rispetto al metodo della normalizzazione ed alla applicazione del *merging* usata nei comuni sistemi di metadati. Questi modelli si avvalgono di ontologie, basate sul concetto di temporalità, in grado di intercettare la maggior richiesta di semantica contenuta nelle query. Mentre ABC finalizza il suo intento al raggiungimento della capacità di modellare la creazione, evoluzione e transizione degli oggetti nel tempo, il modello CIDOC crea una rete semantica per arricchire di relazioni le informazioni contenute nelle varie risorse disponibili.

In particolare, il modello CIDOC-CRM sottolinea il fatto che un comune schema di metadati, indirizzato al recupero della informazione, non garantisce di per sé la compatibilità semantica delle informazioni ottenute, caratteristica quest'ultima ritenuta particolarmente utile per trattare temi attinenti alla "cultural heritage".

Per ottenere un simile esito, l'ottica e la ricerca si devono spostare dalle strutture dati, strumento consolidato per collezionare informazioni, a quello che gli autori del modello CIDOC-CRM chiamano "approccio semantico all'accesso integrato".

L'ambizioso target punta ad ottenere la visibilità del processo di integrazione delle informazioni (read-only integration), più che a normalizzare le strutture espressive. È infatti la visibilità dell'integrazione semantica a permettere l'uso corretto delle migrazioni di dati, della loro possibile fusione, ed a acquisire viste su le integrazioni virtuali dipendenti dal processo di mediazione delle query. Lo strumento principe del progetto è una *ontologia formale* ampiamente teorizzata nel campo della "knowledge reasoning and representation", ed ora proposta come "schema concettuale comune per l'integrazione della informazione".

L'ontologia si applica ogni qualvolta vi sia l'esigenza di analizzare il comportamento di proprietà e di attributi a livello di istanza o di servirsi di una comune base concettuale per la mappatura automatica di strutture dati e istanze. La sua fecondità sta nel consentire a comunità scientifiche diverse di seguire una ricerca terminologica guidata per lo sviluppo di vocabolari descrittivi.

L'uso dell'ontologia formale sembra costituire il passo successivo allo schema comune dei metadati o delle terminologie. Oltre al mero recupero dell'informazione, si aggiunge ora il recupero della sua rete semantica.

L'uso di schemi comuni di metadati e di terminologie ha permesso di sviluppare il concetto di "collezione virtuale"; con l'ontologia formale si aggiunge ora la griglia interpretativa dell'integrazione semantica ed il concetto di "integrazione virtuale".

### ***I principi informatori di CIDOC-CIRM***

Recuperare la semantica che sta dietro l'ampio insieme di strutture dati e metadati di un particolare dominio e delle sue molte sotto-discipline in una coerente ontologia formale ha richiesto la formulazione di una comune base concettuale. Questa è stata realizzata da esperti dei vari domini secondo i seguenti criteri informatori:

- attingere (approccio pragmatico) dalle strutture dati e metadati incontrate nei vari domini i concetti utili per creare un'ontologia di alto livello che esprimesse la conoscenza di base comune e che fosse di dimensioni tali da permetterne la gestione.
- evitare di imbrigliare le strutture dati in operazione di "merging", data la resistenza che queste oppongono ad un simile trattamento
- evitare di utilizzare in questa fase i sistemi terminologici incontrati
- distinguere la top-level ontology dai problemi connessi alla descrizione terminologica

Quest'ultimo punto sottolinea la contraddizione esistente tra criteri informatori dei comuni sistemi terminologici rispetto a quelli perseguiti nel modello CIDOC.

Nel primo caso infatti si punta a coprire gli aspetti del dominio con una ricca divisione di classi e sottoclassi disposte in gerarchie IsA, mentre poco spazio è concesso ad altri tipi di relazioni tutte

comprese nella generica relazione RT. Nel secondo, le classi rappresentano solo il dominio e il range delle relazioni piu' rilevanti, mentre ampio spazio viene dato alle relazioni create dalle *proprietà* ritenute il piu' grande veicolo del contenuto semantico.

In CIDOC, il *framework* integrativo rappresenta il raffinamento ontologico delle classi e viene attuato col recupero delle terminologie che nella fase di costruzione della top-level ontology era stato accantonato.

E' per questa ragione che il modello viene chiamato "ontologia incentrata sulle proprietà", tanto per evidenziare lo specifico carattere e funzionalità.

CIDOC e' dal 2003 candidato a divenire un ISO standard; contiene 80 classi e 130 proprietà.

### ***La normalizzazione: quel che il modello vuole evitare***

Ogni schema, da quelli per i data base a quelli ad etichette come l'SGLM/XML RDF, e' una struttura di compromesso tra la complessità dell'informazione che si vuole rendere accessibile alle query e quella che l'utente puo' effettivamente gestire, ed ancora tra il costo della implementazione della struttura e quello che serve per riempirla di contenuto, etc.

In questo compromesso rientra il processo di normalizzazione; in genere qualcosa si aggiunge, per es. una funzione in piu', ma qualcosa si perde, per lo piu' informazione.

CIDOC propone come standard una struttura dati comune, basata sul metodo della integrazione delle informazioni, in alternativa al piu' comune metodo della normalizzazione delle strutture.

Il metodo consiste in:

- analisi delle strutture dati esistenti in un certo dominio mediante l'utilizzo di tutte le risorse disponibili
- integrazione delle informazioni, in modo da evitare le semplificazioni ed esplicitare le costanti e le informazioni referenziali nascoste.

Il tentativo e' diretto a dare significato alle differenze di struttura, in modo da recuperare nell'ontologia la semantica che sta dietro di esse.

La convinzione e' che l'unificazione delle strutture e dei vocabolari non risolvono di per se il problema del recupero delle informazioni relative ad uno specifico soggetto, se non si procede prima all'integrazione delle informazioni provenienti da molteplici risorse. Ed e' in questo campo che l'ontologia da' i suoi frutti migliori.

### ***Analisi delle strutture dati esistenti: alcune osservazioni***

Tra le semplificazioni che violano i principi informatori del modello di CIDOC, la piu' comune riscontrata e' quella in cui le descrizioni di oggetti o episodi legati ad uno stesso evento o luogo etc., sono prive di informazione sull'evento o luogo medesimo. Questa lacuna impedisce di contestualizzare l'informazione, toglie una chiave essenziale di lettura, ed ostacola la creazione di relazioni tra le informazioni contenute nelle varie descrizioni.

In un record Dublin-Core infatti puo' avvenire che la descrizione di un fatto (per esempio il protocollo dei proceedings della conferenza di Crimea) manchi del riferimento all'evento storico piu' noto (la conferenza di Yalta). Questo avviene perche' le strutture dati riescono ad esprimere solo alcune relazioni tra *items* mentre altre rimangono celate.

Qui il processo di integrazione delle informazioni diventa risolutivo e l'ontologia efficace.

Altra annotazione sullo stesso tema riguarda la trattazione dell'attributo "data" (leggi "datazione") nel processo di documentazione. Se e' vero che "data" se riferita ad un pittore contemporaneo non puo' essere altro che le "date" che si riferiscono allo stesso, e' vero anche che per il cranio dell'uomo di Neanderthal, la "data" e' una complicata storia di esperimenti, risultati, sistemi di datazione, confutazioni e via dicendo.

La semplificazione che si effettua in questi casi nel processo di normalizzazione e' la traduzione della complicata storia di una "data" in una descrizione *free text*: il complesso di informazioni coinvolte in questa operazione non viene di solito considerato una ragione sufficiente per mutare lo schema descrittivo.

Un altro caso molto comune di semplificazione, quando si tratta di descrivere situazioni che coinvolgono persone in tempi diversi e in ruoli diversi, e' quello di usare proprieta' e relazioni di significato molto piatto e generico in modo da riuscire a raggruppare persone e circostanze. Riferito a persone, la relazione "coinvolte in" nell' intervallo di tempo X, pare una informazione sufficiente a descrivere un comune impegno; ma niente dice del ruolo che le persone svolgevano in quella situazione, e in quale arco di tempo effettivo. E' quello che avviene con i metadata record del Dublin Core, che accorpano persone e date senza creare specifiche interrelazioni temporali o di ruolo.

Un'altra osservazione riguarda la descrizione delle istanze del dominio "cultural heritage". Il modello proposto e' diretto a trattare dati storici e quindi descrizioni fatte nel passato sul passato e che pertanto non possono essere mutate anche se incomplete. Tuttavia, a volte, il valore di un simile dato corrisponde ad una opinione piu' che ad una verita', e vi possono essere opinioni contrastanti e quindi valori diversi anche se contemporanei. L'ontologia deve porsi il problema dell'acquisizione di questo tipo di conoscenza ed essere attrezzata a compilare opinioni contraddittorie.

### ***Una architettura per il processo di integrazione***

Le osservazioni precedenti lasciano intendere quanto l'esame delle strutture e delle istanze sia fondamentale per il processo di integrazione della conoscenza.

CIDOC (il modello concettuale o ontologia) vuol essere una descrizione della conoscenza categoriale intorno a "possibili stati delle cose" piuttosto che intorno ad "uno stato delle cose"

L'architettura che integra l'ontologia ad alto livello e' composta (i) dalla *conoscenza categoriale* (analisi delle strutture dati, delle costanti nascoste e terminologie usate nei dati) che estende la semantica delle sue proprieta' ai sistemi terminologici subordinati; e (ii) da uno strato cosiddetto di *conoscenza fattuale* costruito a partire dai dati sorgivi, liste d'autorita', metadata etc.

L'assunzione e' che tutte le sorgenti abbiano un loro modello concettuale e che i dati sorgivi possano essere espressi senza perdita di significato nei termini del modello sorgente, e che questo sia basato sullo stesso formalismo del modello comune.

Nell'ontologia, la "conoscenza fattuale" deve provvedere:

- ad una interpretazione libera dal contesto ed una granularita' appropriata.  
Per es. una istanza della proprieta' come "data di nascita", non puo' essere interpretata senza aver esplicitato il concetto "nascita"; e la proprieta' "data di nascita dell'autore", non puo' essere interpretata senza che sia esplicitato il concetto "autore di cosa".
- ad ammettere, nel campo dei dati storici, la compilazione di *proposizioni alternative multiple* per consentire il ragionamento su un concetto. Per esempio quando si dibatte sulle ipotesi relative alla data di nascita di una persona.

Le attuali strutture di metadata violano di norma questi principi.

### ***Ontologia e "monotonicity"***

I cambiamenti monotoni sono quelli che non invalidano le istanze precedenti

Nel progettare una ontologia e' fondamentale poter distinguere cio' che e' in contraddizione da quello che non lo e'. Dato che la conoscenza cambia nel tempo, questa distinzione deve essere considerata ai fini del mantenimento della *monotonicity*.

L'impatto della *monotonicity* sull'ontologia si puo' notare in tre modi: *nella classificazione*, *nell' attribuzione di proprieta'* (specializzazione), *nella modellazione dei costrutti*. Simili cambiamenti in CIDOC sono considerati monotoni, nel senso che non invalidano istanze precedenti.

### ***La classificazione***

*Nessuna classe puo' essere complemento di un'altra.*

- Se una istanza e' classificata in un certo modo in conformita' con uno stato della conoscenza, l'aggiunta di altri apporti di conoscenza non deve invalidare la precedente classificazione, caso mai puo' comportare una seconda descrizione, o l'aggiunta di una classe.

Da un punto di vista epistemologico infatti non vi sarebbe contraddizione anche se dal punto di vista ontologico crea qualche problema.

Ci possono essere però dei cambiamenti di classificazione non-monotona: per es. il caso di un quadro attribuito ad una epoca in virtù delle tecnologie allora correnti, e attribuito successivamente ad un'altra epoca in base ad esami più sofisticati e probanti. Il cambiamento di classificazione che questo fatto comporta non può considerarsi monotono.

Le due classificazioni entrano in contraddizione. Se è vero che alla domanda “in quale epoca è stato dipinto quel vaso?” non si può rispondere se non con l'indicazione dell'epoca accertata in maniera inconfutabile, tuttavia l'impossibilità di rintracciare una precedente ipotesi di datazione determinerebbe una lacuna storica.

### ***La specializzazione***

L'estensione via specializzazione è prevista nella programmazione per oggetti, ma è poco seguita per estendere la granularità degli attributi (da una proprietà a una catena di proprietà e entità intermedie). L'operazione inversa, è ben compresa e corrisponde all'operazione *join* dell'Algebra Relazionale.

L'estensione della granularità degli attributi può conoscere cammini diversi ed infiniti e portare all'incompatibilità tra descrizioni semantiche che vengono via via a sovrapporsi. In molti casi queste estensioni o mancate estensioni sono giustificate perché funzionali a specifici obiettivi documentalistici.

In questi casi CIDOC modella due percorsi a livello di conoscenza fattuale. Quello diretto, considerato povero, viene chiamato *short cut*.

Perché il cammino indiretto rimanga monotono esso deve essere disciplinato in modo da non perdere la relazione con la informazione principale. Davidson [6] con “l'effetto fisarmonica” ne parla compiutamente.

### ***Modellazioni alternative***

La *monotonicity* può variare in dipendenza del tipo di modellazione scelto.

Per esempio, una storia può essere rappresentata come una catena di *stati* ed *eventi di transizione di stato*”, come avviene nel modello ABC [5].

Da un punto di vista ontologico la transizione può essere rappresentata fuori dalla descrizione dello stato o viceversa. Da un punto di vista epistemologico, la cosa è diversa: se la informazione è incompleta, gli stati e le transizioni non possono trasformarsi gli uni nelle altre, inoltre gli stati sono difficili ad essere osservati. Il fatto che una proprietà sia valida in un intervallo di tempo e non un momento prima o dopo richiede un'osservazione continua, difficilmente praticabile.

Uno *status* invece, cioè la validità di alcune proprietà in un certo *punto temporale* e un *evento di transizione* sono più facilmente osservabili.

CIDOC, dal momento che tratta della semantica di un insieme finito di strutture dati, propone di trasformare, ai fini della normalizzazione, *gli stati* in *eventi di transizione*, considerando la perdita di informazione molto bassa.

Se il modello si basasse sulla costruzione di insiemi di stati per qualsiasi lista di eventi diventerebbe inevitabilmente non monotono dato che l'informazione riguardo ad eventi addizionali richiederebbe la cancellazione degli stati preesistenti.

CIDOC, anche se non riesce a gestire il problema completamente, preferisce modellare i cambiamenti delle proprietà piuttosto che gli stati delle proprietà.

In ambito museale si nota che esistono eventi classificati sotto l'etichetta “accessione ai registri museali” o “de-accessione”, ritenuti non-monotoni se riferiti ad un trasferimento di un oggetto da un museo ad un altro: un oggetto può considerarsi *contemporaneamente* acquisito e scaricato nell'ambito di un trasferimento tra musei.

Se modelliamo il cambio di proprietà piuttosto che lo stato della proprietà ed usiamo descrittori *simmetrici* come “acquisizione” e “cambio di custodia”, non si incorre in questo problema ed il modello rimane monotono.

### ***Interpretazioni del modello***

Quando un modello viene proposto come standard devono essere soddisfatti alcuni attributi di validità'.

Uno di questi è l'indicazione di copertura. L'approccio proposto in CIDOC è di tipo *open-ended*. Durante l'opera di immissione di strutture dati eterogenee nello scopo di CIDOC, è stato osservato che l'ontologia di base si manteneva stabile e le strutture immesse producevano specializzazioni che restavano coperte dal modello senza che questo dovesse essere esteso orizzontalmente.

CIDOC è stato proposto per l'area della *documentazione scientifica ed eredità culturale* come modello comune, ed è il riferimento standard per una lista di strutture per le quali esiste una ben identificata copertura del modello, ed una chiara definizione delle restrizioni relative agli aspetti semantici del contenuto, i costrutti, la formulazione delle regole.

Nel modello concettuale, l'utenza considerata comprende gli studenti, i professionisti, gli organizzatori di cultura, gli insegnanti. Le attività previste sono quelle documentarie e di scambio delle informazioni nel settore delle biblioteche, degli archivi e musei, con esclusione di quelle rilevanti solo per il corrente management interno; gli oggetti presi in esame appartengono alle collezioni canoniche per le quali le strutture museali sono state edificate.

Nel processo di esecuzione del progetto diretto alle proprietà, le classi rappresentano parimenti il dominio e/o il raggio di estensione di alcune di esse, per similitudine con quello che avviene nelle tradizionali strutture dati dei musei. Infatti la granularità terminologica viene espressa come variabile tutte le volte che non si rende necessario rappresentarla in termini di proprietà'. L'idea forza del procedimento e dell'ontologia è rappresentata tuttavia dalle proprietà' piuttosto che dalle classi al contrario di quanto sostenuto da metodologie apprezzate e stabili come la Rational Rose di G.Booch et al.[7] oppure da altre, *object oriented*.

### ***I contenuti in CIDOC CRM***

CIDOC contiene *classi*, e gruppi logici di *proprietà'* che hanno a che fare con concetti come *partecipazione, partizione e struttura, locazione, identificazione, scopo, motivazione, uso, influenza*.

Queste proprietà' rendono centrali le *entità' temporali* e quindi gli *eventi*.

Infatti ogni relazione di una proprietà' con una "data" passa attraverso le entità' temporali; la relazione di una proprietà' con l'ubicazione, e anche in mancanza di riferimento ad una data viene intesa come *short cut* di entità' temporali.

Gli attori hanno sempre relazioni con cose materiali od immateriali mediate da entità' temporali. Ogni istanza di una classe è identificata con un appellativo (*appellation*), sia esso il nome, il titolo, una targhetta, a seconda del contesto storico. La relazione che lega al nome ed alla sua eventuale ambiguità' viene modellata come parte del processo di acquisizione della conoscenza storica.

Tutte le classi possono essere classificate in maggior dettaglio con i *tipi*. È così' che il tipo determina il *range* di valori delle proprietà'.

### ***Gruppi logici di proprietà': spazio, tempo e partecipazione***

La modellazione degli eventi permette di modellare i contenuti culturali in modo da poterli integrare meglio. Se un evento *eI* viene rappresentato come entità' temporale, la partecipazione o la presenza di alcune entità' non temporali in quell'evento servirà' ad arricchire le informazioni ed a permettere di articolare deduzioni.

Per esempio, si potrà' desumere che la partecipazione o la presenza di qualcuno o qualche cosa in un evento non si sia potuta verificare se non nello stesso intervallo di tempo e nello stesso luogo dell'evento *eI*, anche se il tempo ed il luogo non sono indicati. Sono esistiti a quel tempo, ed in quel tempo non potevano trovarsi altro che li'.

Se *e0* è l'evento di creazione di *i* (i partecipanti), questo non può' essere accaduto se non prima o nel medesimo istante dell'evento *eI* etc.

La modellazione dell'evento come entita' temporale permette un *ragionamento cronologico* su tutto quello che vi si riferisce, sia che siano oggetti, attori, o entita' non-temporali.

Questo tipo di ragionamento cronologico, meglio conosciuto come *termini postquem* e *termini antequem*, e' proprio della ricerca storica e CIDOC tende a conservarlo in quanto piu' efficace di quello meramente sequenziale basato su precise informazioni temporali.

In base a questi presupposti, un primo gruppo logico di proprieta' (relazioni) degli eventi viene focalizzato sulle specializzazioni delle proprieta' *avere partecipanti* e *avvenire in presenza di*.

*Avere partecipanti* significa che in quell'evento sono coinvolti attori passivi od attivi. Il dominio e' l'evento ed i partecipanti sono il range degli attori. *Avvenire in presenza di* significa che quell'evento si svolge in presenza di un insieme finito di oggetti. Il dominio e' sempre l'evento; il range e' riferito ad oggetti (concettuali e fisici).

Le specializzazioni che danno senso concreto alle proprieta' sono sottoproprieta' disposte in ordine gerarchico a livelli (dal piu' generale al piu' specifico) identificabili, nella presentazione della lista, dalla profondita' dell'indentatura..

Un altro gruppo di proprieta' inserisce la nozione di *limite* nella comune percezione dell'evento come processo continuo: portato all'esistenza, tolto dall'esistenza. Anche queste proprieta' consentono un ragionamento temporale in assenza di un esplicito riferimento al tempo.

E' significativo notare come una volta assegnati alcuni valori alla proprieta' "date" si possano desumere eventi e circostanze temporalmente vere dal punto di vista logico, anche se questi valori non corrispondono a verita'.

#### **La locazione**

Le proprieta' preposte alla locazione si riferiscono a due tipi di entita': *aree geografiche* ed *oggetti* (le coordinate geografiche di una specifica mappa; oppure "sul Monte Amiata"; etc.). Le coordinate geografiche individuano punti all'interno di grandi aree geografiche, intese come oggetti.. Nell'uno e nell'altro caso l'interesse in CIDOC e' posto sulla relazione che si viene a stabilire tra aree geografiche e gli oggetti di riferimento (per es. la Queen Mary) che possono cambiare di luogo. Piu' che le coordinate assolute per normalizzare il concetto di luogo, nella ricerca storica serve l'informazione relativa.

I luoghi hanno un nome, gli avverbi di luogo sono definiti, gli indirizzi non sono considerati luoghi, ma punti di contatto per persone e organizzazioni.

Esiste inoltre un'altra proprieta' (*forms part of*) che e' una normale relazione part-of riferita alle aree.

#### **L'influenza**

La nozione di *influenza* e' rilevante per capire le motivazioni che guidano l'attivita' umana, sia essa fisica che intellettuale. Certe attivita' possono influenzare altre attivita'. Se queste influenze esistono, si possono anche dedurre sequenze temporali.

Le proprieta' che connotano una influenza sono generalmente legate ai concetti di *modificazione*, *produzione e creazione concettuale*, se sono prototipi o oggetti fisici ad influenzare le attivita'.

Altre proprieta' descrivono l'influenza che si e' verificata nel prodotto senza che vi sia stato un coinvolgimento nella conoscenza del processo. In questo caso si tratta di descrizioni di persone, oggetti, eventi, periodi ben identificati che hanno influenzato il prodotto dall'esterno.

Altre ancora indicano un tipo di influenza che deriva dal carattere stesso dell'attivita' umana, come ordini, impressioni, emozioni (ordino' lo sbarco in Normandia; la collera lo spinse ad anticipare la carica).

La relazione, *had a specific purpose*, gioca un ruolo rilevante nel ragionamento storico.

In CIDOC, il ruolo della *motivazione* nelle attivita' e' riconosciuta, anche se appena accennata, e si ricava dalla logica della *consecutio temporum*.

#### **Commenti**

Il ragionamento cronologico e temporale viene introdotto a partire dalla classe *entita' temporali* nella quale confluiscono sia gli eventi che le azioni. Questa classe permette, mediante relazioni

temporali, di collegare il processo temporale degli eventi agli attori e cose che vi partecipano. La temporalità è il concetto base che dà significato alla rete di relazioni tra eventi, attori e cose; resta invece opaca la possibilità di rappresentare una istantanea del processo.

Va compreso se il mancato riconoscimento da parte di CIDOC degli istanti o punti di tempo (CIDOC riconosce solo gli intervalli di tempo) pregiudichi la disponibilità di qualche importante funzione.

La rappresentazione dei fenomeni che usualmente vengono percepiti come una catena di stati ed eventi di transizione di stati (ABC) viene semplificata in CIDOC con la contrazione degli stati negli *eventi di transizione di stato*. Le ragioni sono pragmatiche e legate all'osservazione dei fenomeni e puntano a privilegiare la modellazione del mutamento delle proprietà per seguire compiutamente una transizione piuttosto che quella degli stati (difficilmente osservabili).

Questa impostazione è confrontabile con quanto Shanahan [8] ottiene con il concetto di *fluente* che rappresenta qualsiasi cosa il cui valore è soggetto a cambiare nel tempo siano esse proprietà dipendenti dal tempo e oggetti che cambiano parzialmente.

L'oggetto di interesse è analogo: tuttavia in Shanahan la chiave di lettura è deterministica e finalizzata al calcolo, e richiede pertanto che si precisi la causa e l'inizio del mutamento (*what happens when*) ed il suo effetto (*what actions do*); in Doerr la chiave di lettura è storicistica e intesa a ragionare su ogni possibile concausa di mutamento lasciando alla teoria probabilistica la verifica ed il calcolo delle condizioni per il verificarsi di un effetto.

Riguardo alla struttura dell'ontologia, M. Doerr sostiene che ai fini interpretativi della realtà, la copiosità delle strutture dati prese in esame non porta necessariamente ad aumentare la complessità della ontologia (diretta ad integrare le informazioni); tale ontologia potrebbe invece aiutare a rendere le strutture dati interpretabili. L'obiettivo infatti non è tanto la ricostruzione di una qualche verità, quanto piuttosto disporre di tutte le informazioni possibili per sviluppare un ragionamento.

Singolare quanto sia sfocato l'approccio tecnologico nel processo di integrazione; appare fondamentale infatti l'impegno concettuale (manuale) nell'interpretazione delle strutture dati, dal quale si fa dipendere la possibilità di usare automatismi di mappatura.

La mancata indicazione dello strumento di ricerca (query language) genera perplessità sulla effettiva interrogabilità della ontologia nel suo complesso.

Se l'ontologia di base, ossia "comune" non necessita, come viene dichiarato, di significative estensioni orizzontali (completezza) mentre si prevedono complessi cammini verticali gerarchizzati, ne consegue che l'ontologia direttamente interrogabile è quella "comune", limitata.

### ***Il ragionamento cronologico in archeologia***

In uno studio successivo "*Supporting Chronological Reasoning in Archeology, 2004*" [9], M. Doerr et al. perfezionano il ragionamento cronologico in ambito archeologico. Nella indagine archeologica la realtà del passato non è accessibile con osservazioni dirette ed il sistema di datazione si deve basare su una conoscenza generale della realtà e su strumenti e metodi di evidenza storica che permettono inferenze cronologiche.

Scopo dichiarato di questa ricerca è quello di (i) analizzare l'applicabilità di ogni metodo di ragionamento sviluppato per singoli casi di datazione, (ii) di creare sistemi di ragionamento per casi specifici, (iii) di assistere nella selezione di sistemi di ragionamento efficienti, e di (iv) raggiungere una interoperabilità semantica tra la documentazione culturale e sistemi di ragionamento cronologico.

L'ambizioso *target* è diretto a costruire una teoria generale basata su: (i) un modello degli elementi cronologicamente più rilevanti della realtà così come si manifestano nei discorsi cronologici, (ii) un modello di relazioni tra le date ed i fenomeni del mondo reale, (iii) una classificazione delle argomentazioni cronologiche, (iv) uno schema formale per risolvere problemi cronologici.



Dopo un esame delle teorie e metodi per risolvere problemi di ragionamento temporale nell'ambito dell'intelligenza artificiale M. Doerr et al. analizzano gli approcci tecnico-formali nel campo della scienza (radiocarbonio, stratigrafia). In particolare studiano l'approccio sistematico di Harris, "Harris-Matrix" sul sistema di datazione relativa nella stratigrafia, ed al lavoro di M.K. Holst "Complicated relations and blind dating: formal analysis of relative chronological structures". Quest'ultimo lavoro suggerisce di creare un *framework* in cui combinare le informazioni stratigrafiche con gli effetti temporali derivati dalle osservazioni sulle relazioni strutturali verificate in altri siti scavati nelle vicinanze.

Con un ragionamento simile M. Doerr assume che l'ontologia possa essere considerata come un modello di possibili *stati delle cose*, dove *stato delle cose* viene definito come una distribuzione di item potenzialmente osservabili insieme alle loro relazioni e qualità, in una dimensione spazio-temporale.

Questo indirizzo altro non è che una interpretazione e una estensione di CIDOC-CRM, in quanto dal modello dei possibili stati delle cose del mondo reale si astraggono i fenomeni storici ed archeologici come *rete di item persistenti* che interagiscono nello spazio e nel tempo. L'approccio richiede che la nozione di *periodo* non abbia solo una valenza temporale, ma acquisisca il significato di *un insieme coerente di fenomeni legati insieme nello spazio e nel tempo* e che "l'evento" venga considerato un caso speciale di "periodo", in un *volume di coerenza* dove i partecipanti (being, ending ones) sono, secondo un concetto fisico, nella posizione di interagire (*interacting position*).

Viene introdotto il concetto di *meeting* che favorisce straordinarie possibilità di interpretazione di un evento. (Se trattiamo la "datazione di un oggetto" come "meeting" significa che cerchiamo di approssimarci al volume di coerenza di qualche evento nel quale l'oggetto era presente).

Sia gli eventi storici che quelli geologici si prestano a questa interpretazione. Per esempio la morte di Cesare è l'incontro con Bruto e la sua spada. Tuttavia l'evento può essere visto anche come un'aggregazione di sotto-eventi: l'incontro di Cesare con i senatori prima, ed il corpo di Cesare arso nei pressi del Foro dopo (l'effetto fisarmonica di Davidson, [6]).

Anche lo scambio di informazioni può essere concepito come incontro in un certo senso. Il messaggio è qualcosa di immateriale (*conceptual object* in CIDOC) che può coesistere in punti geografici diversi.

È per dare la possibilità di creare cronologie relative tra i partecipanti di un evento, indipendentemente dalla natura dell'evento e dei suoi partecipanti che CIDOC si avvale di un insieme di proprietà come *occurred in the presence of, brought into existence, took out of existence*.

### **Un modello formale per il ragionamento temporale**

Il modello si basa sull'assunto che la vera estensione di un evento non si possa osservare, ma che tuttavia si possano identificare *date* che sono antecedenti o posteriori all'*end-point* di un evento.

Il formalismo si basa su una struttura Evento/Tempo definita come  $ETS = (E, TM, h, \pi)$  dove

- $E$  rappresenta un insieme di eventi o periodi
- $TM$  è un modello di tempo lineare dove:
  - $D$  è un insieme di date del calendario giuliano  $d$  rappresentate da numeri reali
  - $T \subset (D \times D)$  è un insieme di intervalli di tempo (convessi) specificati dai loro *end-points*
  - $u(t), t \in T$  è una funzione che accoppia il più alto *end-point* di un intervallo ad un elemento di  $D$
  - $l(t), t \in T$  è una funzione che accoppia il più basso *end-point* ad un elemento di  $D$
  - $\leq$  è l'ordine temporale completo in  $D$
- $h$  è una funzione che accoppia ciascun elemento  $e \in E$  ad un elemento  $t \in T$  e vuole rappresentare il tempo nel quale l'evento stava realmente svolgendosi.
- $\pi$  è la funzione che accoppia ciascun elemento  $e \in E$  e  $d \in D$  ad una funzione  $f$  (di distribuzione di probabilità) che fornisce la probabilità che ha l'evento od il periodo considerato di essere in corso alla data  $d$ .

Gli intervalli di tempo sono sia i veri intervalli degli eventi (quantunque indeterminati o sconosciuti) sia la loro approssimazione.

Il modello permette di ragionare in una condizione di conoscenza temporale imprecisa ed evidenzia le relazioni che intercorrono tra un intervallo di tempo  $t \in T$  ed un evento  $e$ :

- (D1) *indeterminazione*  $i(t, e)$  si verifica se e solo se la funzione che mappa ogni elemento  $e$  di  $E$  con un elemento  $t$  di  $T$  e' inclusa in  $t$ :  $i(t, e) \Leftrightarrow h(e) \subset t$
- (D2) *determinazione*,  $d(e, t)$  si ha solo se la funzione che mappa ogni elemento  $e$  di  $E$  con un elemento  $t$  di  $T$  include  $t$ :  $d(e, t) \Leftrightarrow h(e) \supset t$
- (D3) *indeterminazione dell'inizio*  $b(t, e) \Leftrightarrow l(h(e)) \in t$
- (D4) *indeterminazione della fine*  $e(t, e) \Leftrightarrow u(h(e)) \in t$

Questo modello usato come base di un'algebra dell'intervallo serve a formalizzare un ragionamento cronologico. Una estensione del modello CIDOC potrebbe riuscire a sostituire i limiti temporali dell'intervallo con i limiti spazio-temporali dei *volumi di coerenza*.

### ***Il ragionamento cronologico***

Il formalismo presentato si basa sull'assunzione che la datazione significhi approssimazione ai limiti temporali del *volumi di coerenza* di qualche evento e che datare un oggetto significhi datare un evento nel quale l'oggetto era presente per qualche ragione (produzione, uso, etc.).

Lo scopo della cronologia, per M. Doerr et al., infatti e' la determinazione dell'intervallo di tempo (con la minima indeterminazione dell'inizio e della fine) coperto dall'evento preso in considerazione.

Un ulteriore raffinamento del metodo consiste nel determinare la probabilita' che ha ciascun evento o periodo di cominciare o di finire ad un certo tempo.

Il processo cronologico quindi e' o la piu' piccola concentrazione temporale di tutti i possibili stati delle cose temporalmente rilevanti e consistenti con una data evidenza, oppure la determinazione dei piu' probabili stati delle cose consistenti con una data evidenza. Gli autori considerano il primo caso come pre-requisito vincolante per la riuscita del secondo (una valida teoria probabilistica). Sui due fronti la ricerca e' aperta.

Il contenuto dell'*evidenza* viene classificato in base alle conseguenze cronologiche che esso produce ed alla complessita' del ragionamento cronologico che puo' essere coinvolto.

Viene considerata *evidenza primaria* di eventi del passato (i) quello che questi hanno prodotto, (ii) le tracce permanenti, (iii) quello che risulta nei rapporti orali o scritti ed (iv) anche alcune caratteristiche degli oggetti come lo stato di decadimento (chimica) o le alterazioni chimiche che forniscono evidenza ai fini del computo cronologico.

Si distinguono:

#### *- La cronologia assoluta*

Essa proviene da tre tipi di fonti: (i) record storici che si riferiscono ad osservazioni legate ad un calendario conosciuto oppure eventi astronomici od altri per i quali sia possibile una datazione certa. (ii) confronto tra un campione di una sequenza parziale di tracce ed una sequenza completa e conosciuta (campioni di ghiaccio polare, per es.). (iii) stato di invecchiamento di un oggetto rilevato in base agli effetti noti che su esso hanno prodotto agenti e processi (radiocarbonio, potassio-argon, termoluminescenza).

#### *- La cronologia relativa basata sull'ordine degli eventi*

Riguarda casi di evidenza diretta dell'ordine temporale di eventi multipli. Tre le sorgenti: (i) record storici che legano temporalmente una osservazione su un evento ad un altro evento (lista cronologica degli imperatori cinesi), (ii) ordine delle tracce di eventi differenti che mostrano un evidente ordine temporale (stratigrafia), (iii) relazioni causali tra eventi: eventi la cui esistenza rappresenta il requisito necessario perche' altri eventi si verificano (il trasferimento di una informazione avviene al momento o dopo la creazione della informazione ed i mezzi per trasmetterla)

L'informazione sull'ordine degli eventi non dà indicazioni sulla durata di un evento ma permette di creare network temporali che possono essere combinati con elementi di cronologia assoluta, in modo che intervalli considerati relativi possano mutarsi in assoluti.

- *La cronologia relativa basata sull'inclusione di un evento*

E' il caso di grandi processi cronologici (super-eventi) che includono altri processi (sub-eventi) che possono essere datati singolarmente. Ci sono tre tipi di fonti informative: (i) record storici di osservazioni attuali, (ii) inclusione di tracce, (iii) relazioni causali.

L'informazione data dall'inclusione provoca l'inserimento di un nuovo vincolo nel ragionamento temporale e una migliore approssimazione della durata del sub-evento.

- *La cronologia relativa basata sulle distanze temporali*

Si hanno tre tipi di fonti: (i) record storici di osservazioni attuali, (ii) calcolo della distanza temporale: la distanza temporale nell'accadimento di un effetto si calcola mettendo in relazione l'ampiezza dell'effetto con il ritmo del cambiamento (stima ponderata) con cui l'effetto si produce (velocità di mutamento di uno stile, per esempio), (iii) conoscenza pregressa della durata media di un fenomeno (la gestazione, per esempio).

- *Datazioni categoriali o tipologiche*

Le forme di datazione incontrate fino ad ora si riferiscono a fatti (material facts); tuttavia esempi di datazione possono riguardare anche concetti astratti o categorie. In questo caso il processo di datazione avviene in due passi successivi: l'attribuzione degli oggetti ad una categoria; la categoria è associata ad alcuni eventi, il cui insieme viene datato entro i limiti esterni che divengono i limiti della datazione della categoria.

Si assume che i particolari oggetti di una categoria condividano l'ambito di datazione della stessa.

Tre sono le forme possibili di datazione categoriale: (i) gli eventi di creazione  $p(o_i)$  di tutti gli oggetti  $o_i$  di una categoria  $C$  cadono entro l'estensione spazio-temporale di uno specifico periodo minimo  $P(C)$ :

$$P(C) := \inf \{ t \in T : \forall o_i \in C \Rightarrow h(p(o_i)) \subset t \}$$

(ii) gli eventi di produzione di tutti gli oggetti di un tipo sono tutti posteriori agli eventi di produzione di un altro tipo di oggetti (iii) gli eventi di produzione di un certo tipo di cose hanno una specifica *distanza temporale* dagli eventi di produzione di un altro tipo di oggetti; *the speed of evolution* per esempio.

### **Commenti**

Il formalismo basato sugli intervalli, suggerito nell'articolo, richiede (i) una nuova definizione di *evento* che diviene *complesso, non-istantaneo e con limiti temporali confusi*, (ii) l'introduzione della nozione di *proprietà osservabili*. Queste esigenze richiedono la messa a punto di alcune estensioni di CIDOC.

E' da notare come la suddivisione in *cinque classi dell'evidenza archeologica* e della conoscenza del passato, effettuata in base alle conseguenze temporali che esse producono, semplifichi il ragionamento cronologico e permetta di individuare forme matematiche adeguate alla loro rappresentazione.

La tipologia delle fonti individuata per ciascuna classe, appaiono sufficienti a descrivere i principali generi dell'osservazione elementare condotta sui fenomeni.

Come dichiarato dall'autore, lo studio si pone come preliminare ad una più generale teoria del ragionamento cronologico e richiede una verifica sulla effettiva copertura dell'evidenza archeologica e sulla compatibilità tra i formalismi basati sull'intervallo e quelli probabilistici.

E' da rilevare che la nozione di tempo, non è intesa a stabilire, come avviene in Shanahan [7], l'inizio e la fine di un evento o di una azione e quindi la loro durata. Nei lunghi periodi archeologici, la durata effettiva di un evento sembra perdere d'importanza, mentre si afferma la necessità di stabilire il suo posizionamento temporale (cronologico) in un insieme di intervalli.

Negli enunciati di Shanahan l'indicazione dell'istante in cui ha luogo una azione fa parte della stessa descrizione dell'azione e determina il suo essere causa di un effetto; nel formalismo adottato per la datazione archeologica la rappresentazione dell'istante non è significativo e la temporalità

degli eventi non e' altro che la conseguenza temporale (l'esito logico) dell'evidenza presa in considerazione, sia essa evidenza storica, legge chimico-fisica, confronto, deduzione.

L'algebra degli intervalli come metodo di trattazione degli aspetti cronologici degli eventi e delle loro relazioni si dimostra utile la' dove le relazioni cronologiche tra periodi (antequem, postquem) riescono a caratterizzare altri periodi (eventi).

La ricomposizione cronologica degli eventi nonostante l'indeterminatezza della loro durata e' significativa nei lunghi periodi storici dell'archeologia in quanto permette di datare relativamente frammenti di un passato che ha lasciato poche tracce di evidenza dietro di se.

Le stesse ragioni sembrano non valere nel caso della ricostruzione del percorso degli avvenimenti che caratterizzano il *life time* di un oggetto. Quando, dal mondo popolato dalla storia di eventi ed azioni si passa alla considerazione di *un oggetto e della storia degli eventi che lo hanno riguardato*, come avviene in ambito museale, la datazione certa degli eventi diviene una esigenza (anche lo statuto dei diritti proprietari lo impone) e si riafferma anche l'importanza di disporre di una rappresentazione istantanea della processualita' di un evento.

## **Parte II – ABC Model: a metadata model and ontology for Web and digital libraries by J. Hunter and C. Lagoze**

### ***Scopo e obiettivi del Modello ABC***

Lo scopo del progetto ABC [5] [10] e' mirato a modellare la creazione, evoluzione e trasformazione di un oggetto *nel tempo* con un numero di entita' base e relazioni, comuni a piu' vocabolari di metadati, in modo da rendere interoperabile la ricerca.

Le esigenze da soddisfare sono (i) Il superamento di una descrizione catalografica statica dell'oggetto (assegnata una volta per tutte) e (ii) la costruzione di una base concettuale comune tale da consentire l'interoperabilita' dei linguaggi.

Nel contesto delle collaborazioni essenziali per la riuscita del progetto, figurano le comunita' di Dublin Core Initiative[11], CIMI [12], IFLA FRBR[13], INDECS [14], CIDOC-CRM, RDF [15] ed altre.

Significativo l'apporto di IFLA per la definizione del concetto di *life-time* di un oggetto e di INDECS che per primo ha puntato sull'*evento* come entita' base del modello.

L'obiettivo si e' focalizzato sulla messa a punto di metodologie adeguate per descrivere l'oggetto (in una varieta' di domini) e modellarne il cambiamento in assoluta trasparenza.

### ***Overview dell'ontologia ABC***

ABC e' stato progettato come ontologia primitiva in modo da permettere alla utenza destinataria (comunita' che opera su metadati individuali e progettisti di sistemi) di potere costruire in testa adattandovi le proprie esigenze.

Gli elementi dell'ontologia ABC sono divisi in classi corredate di proprieta'.

La categoria primitiva e' **Entity** che ha come sottoclassi:

- **Temporality** che provvede a dare un contesto esistenziale e temporale alle entita' con le sottoclassi: *evento, situazione, azione*;
- **Actuality** che si riferisce a entita' che hanno una esistenza tangibile e duale (*esistenziale e universale*) con le sottoclassi: *artefatto (manifestazione, item), agente*;
- **Abstraction** che si riferisce ad entita' che sono idee o concetti con la sottoclasse: *work*

Significativo l'aggancio all'interpretazione FRBR della storia di un oggetto (dall'idea al manufatto) che indica l'impegno mirato ad integrare i modelli.

Mentre la gerarchia delle classi, arricchita dalle proprieta' che associano istanze delle classi, costituisce la struttura del modello e determina la sua espressivita', l'ambito semantico attribuito alle classi e sottoclassi e la scelta delle proprieta' e del loro significato permettono di descrivere una infinita' di eventi e situazioni.

Le classi selezionate colgono gli aspetti essenziali della realta' e le proprieta' descrivono il loro essere e divenire.

La categoria **Temporality** con le sottoclassi *evento, situazione, azione* e' preposta a modellare il tempo, l'evoluzione dei valori delle proprieta' di un oggetto nel tempo ed a soddisfare domande sul *quando, come, chi e dove* di una entita'.

Le entita' temporali che la compongono, *situation, event, action* sono entita' ontologiche che (diversamente da quelle di Dublin Core) consentono di associare le responsabilita' dell'agente ed i cambiamenti di stato della risorsa.

- *Situation* e' il contesto in cui si fanno asserzioni esistenziali sulle *Actuality*.

E' il contesto in cui si asserisce cosa precede (*precedes*) o cosa segue (*follows*) un evento ed il legame esistente tra una *Actuality* ed una *Situation*.

*Situation* ha la durata di tempo determinata dagli eventi di confine.

- *Event* marca la transizione tra una situazione ed una altra. Le proprieta' *precedes* e *followes* associano ad un evento la *Situation* precedente e quella seguente: Sono proprieta' che indicano un effetto di entrata e di uscita, ma non sono relazioni causali; queste vengono espresse solo dalla relazione *hasResult* che nel dominio *Event e Action* determina la modifica di una *Actuality* nel suo aspetto esistenziale.

Solo *creates* (sottoproprieta' di *hasResult*) e *destroys* (sottoproprieta' di *hasPatient*) determinano rispettivamente la creazione o la scomparsa di una *Actuality*.

La granularita' della transizione di un evento e' variabile (istantanea od intervallo) ed e' espressa tramite la proprieta' *atTime* che associa l'entita' *Time* alle sottocategorie di *Temporality*, mentre *atPlace* associa l'entita' *Place* alle sottocategorie sia di *Temporality* che di *Actuality*.

- *Action* e' un verbo, una attivita' nel contesto di un evento compiuta da un agente o piu' agenti nelle loro responsabilita' individuali o corporate. L'agente, in un ruolo meno attivo, puo' essere presente (*hasPresence*) o partecipare (*hasParticipant*, sottoproprieta' di *hasPresence*) ad una azione od evento.

L'azione, nel contesto dell'evento, puo' coinvolgere per compiersi (*involves*) delle *Actuality*.

*HasPatient* e *usesTools* sono le specializzazioni di *involves* che evidenziano il come del coinvolgimento.

### **ABC: altri aspetti del modello**

In ABC, una parte importante della realta' viene sistemata, come abbiamo gia' detto, nella sottoclasse di Entity **Actuality**, che racchiude la classe di entita' percepibili con i sensi.

Queste, abbiamo visto, possono avere sia aspetti permanenti, *i. e.* non dipendenti dal tempo, che aspetti dipendenti dal tempo, e relazioni che permettono di collegare l'aspetto esistenziale di una entita' a quello universale e questo ad una specifica situazione.

Il meccanismo, come vedremo nella narrativa, che si integra con le funzioni di *Temporality* e' preposto ad esprimere e modellare il cambiamento di un oggetto nel tempo, obiettivo target del modello.

Lo sviluppo di *Actuality* nelle sottoclassi *Artifact* (con le articolazioni *Manifestation, Item*) ed *Agent* permette inoltre di evidenziare come:

- una tangibile realizzazione (*Artifact*) possa manifestarsi in un certo numero di modi (versioni) a loro volta associabili tramite la proprieta' *has copy* alle rispettive copie (*items*), che vengono a formare una classe di equivalenza esatta;

e come

- *Agent* (sia esso persona, strumento, organizzazione) possa partecipare con una azione ad un evento in quanto tipo di *Actuality*.

Nel modello, come abbiamo osservato, e' inserita un'altra categoria primaria, **Abstraction**, che permette di ottenere sotto-classi di pura informazione o concetti. (Questa categoria contrasta con *Actuality* che si riferisce a sotto-classi di entita' che hanno una esistenza tangibile e la cui identita' e' rappresentata da proprieta' sia *universali* che *esistenziali*).

La super-categoria *Abstraction* da' corso alla processualita' temporale prevista nello schema IFLA RFBR per la descrizione della storia di un oggetto. Il processo inizia con *Work*, che nella filosofia IFLA designa la proprieta' intellettuale; ma *Work* e' solo l'idea di qualcosa ed esiste nel modello solo in quanto si attualizza, si manifesta. Questa eventualita' e' rappresentata dalla proprieta' *has*

*relation* che associa *Work*, in una relazione uno-molti, ad una o più espressioni (*manifestation*). Queste a loro volta possono avere copie *has copy* ed esistere come *item*.

Notiamo che *Actuality* e *Abstraction* condividono un percorso mutuato da IFLA (*manifestation* ← *item*). Questo significa che:

- nel mondo delle astrazioni vi sono idee (*Work*) che si attualizzano (*has realization*) in qualcosa di manifesto (*manifestation*) e che
- questo qualcosa che si manifesta come realizzazione tangibile (*manifestation*) non è che una forma di *Artifact* (sottoclasse di *Actuality*) la cui caratteristica è quella di potersi riprodurre in un certo numero di modi e copie.

### ***Gli elementi dell'ontologia della transizione***

La logica che sovrintende e coordina gli elementi dell'ontologia diretta a descrivere il cambiamento si basa sulle seguenti assunzioni:

- l'evento (istantaneo o processuale) è l'entità di cui si parla; ha un nome ed un codice identificativo, una data ed un luogo in cui si compie ed ha sempre proprietà temporali.
- l'evento è il dominio delle azioni eseguite (e le azioni del mondo naturale?) dagli agenti, (persone, strumenti, organizzazioni) (*has agent*) che hanno un nome ed esercitano un ruolo (*has role*).
- l'evento, sia che sia processuale od istantaneo, è un mutamento che indica (*marks*) la transizione tra una situazione ed un'altra di un qualcosa che è una *Actuality* nel suo aspetto esistenziale ed universale, dunque qualcosa che è già manifesto. L'azione e gli agenti contribuiscono a determinare una nuova situazione.
- La transizione è codificata dalle proprietà *precedes* e *followes* che associano all'evento la situazione antecedente e quella successiva.
- il concetto di *situazione* e di *evento* sono poste in contrasto. Il primo indica intervalli o istanti di stabilità nel cui contesto vengono asserite alcune proprietà esistenziali di una sottocategoria di *Actuality*. Il secondo marca gli scenari del cambiamento, la cui granularità è variabile.
- La proprietà *isPartOf* nel dominio della categoria *Entity* ed il suo inverso *contains* permettono di raffinare la modellazione di una *Actuality* che non avrebbe senso senza quella parte. (Museo di archeologia e reperti archeologici)
- *IsSubEventOf* è una sottoproprietà di *isPartOf* nel dominio dell'evento che stabilisce una relazione tra due eventi (il sacco di Roma e la fuga di Clemente VIII). Questa sottoproprietà, a differenza delle proprietà *contains* e *isPartOf* non implica il mantenimento dei vincoli semantici.
- Ciascuno degli eventi legati da una relazione di *isSubEventOf* agisce come punto di transizione tra situazioni precedenti e successive. La proprietà *hasAction* (dominio *event*) comporta invece che si parli di un singolo evento, della sua situazione precedente e seguente e delle azioni o verbi che agiscono individualmente all'interno di quell'evento.
- La proprietà *inContext* lega una *Actuality* (nel suo aspetto esistenziale, cioè il suo set di proprietà esistenziali) al contesto della *Situation*.
- *isPhaseOf* (dominio di *Actuality*) è il legame tra le faccette esistenziali di una *Actuality* (nel suo stato manifesto) ed il suo unico aspetto universale. Le fasi (di successione nel tempo) corrispondono ai diversi valori della faccetta esistenziale legata ai valori immutabili dell'aspetto universale.

### **Modellazione dei cambiamenti: la narrativa**

Se volessimo ricostruire la storia di un oggetto, poniamo un vaso, usando il modello ABC (un vaso che sappiamo modellato da qualcuno, decorato da altri, collocato in un museo, prestato per una mostra, e successivamente fotografato), dovremmo ragionare sulle tre traiettorie logiche proposte da ABC in maniera distinta e cooperante:

- (i) una traiettoria inizia nel mondo delle astrazioni (*Abstraction*), prosegue in quello delle idee di qualcosa, in questo caso *del* vaso (*Work*) e quindi giunge all'idea *di un* vaso particolare (*Wk0*); questa idea viene legata ad una sua realizzazione (*has Realization*) che e' una delle manifestazioni possibili *Manifestation (MN0)*
- (ii) l'altra si avvia, a partire dalla super-categoria *Entity*, verso il mondo delle concretizzazioni sensibili (*Actuality*) e prende forma nella concretizzazione di quel vaso particolare (*Artifact*) che puo' manifestarsi (*MN0*) con le sue caratteristiche *universali* in un certo numero di modi e copie.

Con questo percorso si riesce a rappresentare la creazione del vaso.

- (iii) un'altra traiettoria che prende avvio dalla entita' *Actuality* prosegue il percorso antecedente. A partire da (*MN0*), scandisce il collegamento (*phaseOf*) tra una fase di evoluzione degli aspetti esistenziali e l'aspetto universale del vaso. I mutamenti nella faccetta esistenziale in dipendenza dagli eventi, sono collegati tra loro nel contesto (*inContext*) della *Situation* (sottoclasse di *Temporality*). Attraverso i valori della faccetta esistenziale di *Situation* si esprime lo stato delle cose riguardante il vaso prima e dopo ciascun evento.

Con la successione delle *Situation* si intende disegnare la storia del vaso nelle sue tappe e garantire (i) la presentazione di un panorama dinamico dei mutamenti del vaso e di quel che gli e' occorso nel tempo (life-time), (ii) la possibilita' di accesso al contesto dei valori della faccetta esistenziale presente in ogni *Situation* (iii) e la produzione di istantanee

### **Commenti**

Emerge nell'esposizione del modello un notevole impegno di sintesi dei sistemi presi in esame (INDECS: la centralita' dell'evento, CIDOC: le categorie temporali, IFLA: lo schema logico, etc.). E' inoltre sorprendente come l'esiguita' degli elementi dell'ontologia (rispetto a CIDOC per es.) oltretutto rendere maneggevole il modello riesca a far fronte alle esigenze di modellazione della realta'.

Di ABC convince la modellazione della creazione dell'oggetto e la distinzione, resa visibile negli esempi, delle proprieta' la cui presenza e valori sono dipendenti dal tempo da quelle che ne sono del tutto indipendenti, distinzione su cui si forgia la modellazione dei cambiamenti e accadimenti che riguardano l'oggetto nel tempo.

Il concetto di temporalita' in ABC non rimanda alla creazione di una vera e propria ontologia dei tempi, ma e' diretta a rappresentare temporalmente (tramite entita' temporali come *Event*, *Action* e *Situation*) i mutamenti subiti dall'oggetto e facilitarne la lettura e la ricerca.

In particolare, l'evento ha un inizio nel tempo (*atTime*) ed una estensione (*duration*) ed e' la vera entita' temporale su cui si ragiona e dalla quale viene desunta la temporalita' delle altre sottocategorie quali *azioni* e *situazioni*

- l'evento si svolge sempre nel tempo (istante o intervallo); l'evento marca le transizioni tra una situazione ed un'altra;
- la situazione ha una durata determinata dai tempi di inizio e fine degli eventi di confine.
- l'azione viene sempre concepita nel contesto di un evento; non si da il caso che un'azione umana o naturale provochi un evento.
- l'evento puo' avere una o piu' azioni *has Action* eseguite da agenti. Si usa l'azione per modellare la figura dell'agente (*agent*) e le sue responsabilita' nell'evento

In questo disegno l'entità *situation* sottocategoria temporale che modella i mutamenti nello stato delle cose tra un evento ed un altro in un succedersi di eventi non fornisce informazioni suppletive di interesse rispetto a quelle ottenute con una modellazione incentrata sui mutamenti delle proprietà dell'oggetto

Anche la sottocategoria *action* in ABC merita qualche osservazione.

Per quel che riguarda la temporalità e' da rilevare come essa sia confinata entro la durata temporale di un evento *has action*, talche' non si distingue un evento *nascita* da un evento *creazione*.

La nascita infatti presuppone una azione, *parenting*, che e' antecedente e causante l'evento *nascita* e quindi non complemento oggetto della sua proprietà *has action*, come invece potrebbe avvenire nel caso si trattasse dell'evento *creazione*.

E' da notare inoltre che nella narrativa la rappresentazione dei diversi significati che *action* può assumere in rapporto all'evento, causale, di presenza, di coinvolgimento etc. viene soddisfatta da verbi come *involves*, *hasPatient*, *hasParticipant* che rendono opaca, in certi casi, la distinzione nell'uso di *has action* rispetto a *isSubEventOf*.

In effetti nell'evento "creazione" di un vaso si possono distinguere fasi di lavoro come modellazione, decorazione, incisione, pittura che sono a tutti gli effetti sub-eventi ma allo stesso tempo sono anche azioni temporalmente distinte.

La differenza nell'uso di queste proprietà viene evidenziata dal fatto che *isSubEventOf* lega gli eventi in modo che ciascuno di questi possa agire come momento di transizione tra una situazione precedente ed una seguente, mentre con *hasAction* c'è un unico evento con una unica situazione precedente ed una seguente e una o più azioni che agiscono come verbi entro quell'evento.

ABC nel caratterizzare le conseguenze delle due diverse applicazioni di fatto sembra lasciare all'utente la scelta.

A livello più generale, colpisce la nomenclatura delle entità del modello perché poco intuitive a partire da *Existential* o *Universal Actuality*, o addirittura devianti come avviene per *Situation*.

Risultano poco chiare anche le ragioni di alcuni percorsi che sembrano doppi tracciati: *Work*, *Manifestation*, *Item* e *Actuality*, *Artifact*, *Manifestation*, *Item*.

Infatti rimarcare la distinzione tra entità concrete ed astratte non aiuta a rendere fluido il modello e non aggiunge possibilità di espressione più di quanto non possa consentire lo schema consolidato di IFLA. da solo.

Tuttavia e' proprio all'inizio, all'affaccio sull'impalcatura del modello che sorgono le prime perplessità circa le distinzioni o accorpamenti concettuali. Alcuni sembrano derivare più che da una visione organica della realtà da frettolose esigenze di natura informatica.

Se ammettessimo invece, insieme a Davidson, che il motore dell'universo e' consegnato all'interagire di eventi ed azioni (legate ad agenti) ed ai loro effetti più o meno causali (come mutamenti nel pensiero, nel mondo naturale, negli artefatti), ci ritroveremmo forse più prossimi al concetto di *fluente* di Shanahan [8] piuttosto che a quello di *Situation di ABC*.

### **Parte III - La rappresentazione ontologica e multimediale della realtà museale: con CIDOC CRM + MPEG7 un nuovo approccio all'interoperabilità**

#### ***L'espressione multimediale***

L'ontologia CIDOC intende catturare la semantica del complesso di informazioni inerenti la realtà museale.

Se queste informazioni sono espresse in forma multimediale (MPEG7) [15] di quale semantica si parla e come la si può indagare?

A questa domanda risponde J. Hunter in "Combining the CIDOC CRM and MPEG7 to describe Multimedia in Museums", 2002 [16].

La multimedialità richiede ontologie specifiche a seconda del linguaggio usato.



Tra gli aspetti che in tale contesto vanno evidenziati, distinguiamo quelli che si riferiscono al contenuto da quelli che si riferiscono al contesto.

Le informazioni sul contenuto multimediale (*what about*) sono necessarie per conoscere quello che si è voluto registrare in forma multimediale: eventi, scene, persone, località, monumenti, oggetti, etc.

Sono informazioni strutturali, che si riferiscono alle specificità fisiche dei media, al loro essere statiche (still image) oppure in movimento (film); in questi casi sono necessarie altre informazioni sui singoli segmenti temporali, spaziali e spazio/temporali (scenes, shots, frames, image regions) e sulle loro relazioni.

Le informazioni sul contesto sono utili, in quanto soddisfano quesiti di interesse storico: come conoscere la persona che ha prodotto l'oggetto multimediale se di rilievo; la tecnica di ripresa, se particolare; oppure sapere se l'oggetto multimediale fa parte di una collezione pregiata, oppure rappresenta un raro documento storico etc.

Le informazioni di tipo  $p^3$  (*provenance, proprietary rights, preservation*) sono da considerarsi prescrittive. Esse riguardano la storia degli eventi che hanno interessato la realtà museale nel suo complesso e quelle relative ai diritti che si sono maturati di conseguenza.

Tra i metodi di preservation ricordiamo che ogni oggetto può conservare memoria di sé attraverso una copia. Oggi le risorse multimediali forniscono una rappresentazione visiva dell'oggetto, alternativa alla copia.

Importanti, poi, le informazioni sul tipo di format, codifica e requisiti di sistema.

Perché tutti questi aspetti possano essere catturati è necessario che le risorse multimediali siano rappresentate oltre che dalle tradizionali informazioni bibliografiche comprese quelle sul contenuto, anche da informazioni specifiche per i singoli media.

Da un punto di vista semantico è da notare come le *textual search query* e quelle sviluppate in base a specifiche fisiche *search by features* possano essere sia concorrenziali che complementari ai fini del risultato della ricerca.

### ***L'interoperabilità tra CIDOC e ABC***

Questi aspetti, fondamentali per descrivere compiutamente risorse multimediali, portano Hunter a ritenere incompleta l'ontologia di CIDOC ed a proporre una sua integrazione con una ontologia che rappresenti le caratteristiche dei mezzi espressivi (multimediali) che CIDOC non ricopre.

Dato che nella proposta Doerr l'ontologia è essenziale ai fini della interoperabilità semantica, sembra necessario riflettere su cosa significhi "semantica" in un tipico linguaggio multimediale.

La metodologia che Hunter segue per definire una ontologia composita, adatta a interpretare la descrizione multimediale della realtà museale è essenzialmente di tipo pragmatico.

Dato un esempio, tratto dall'ambito museale fattuale, Hunter analizza in successione ciò che non è possibile descrivere usando l'ontologia CIDOC e di seguito quello che si ottiene utilizzando MPEG7.

Per riempire le lacune evidenziate dall'analisi, Hunter suggerisce un'integrazione aggiuntiva rispetto a quelle previste dalla metodologia CIDOC.

In CIDOC, la fase di integrazione delle informazioni consiste nella costruzione di una rete informativa e in una fusione della conoscenza testuale proveniente dalle varie risorse prese in considerazione ed in una ricomposizione gerarchica a partire da una limitata ontologia di base. Per Hunter si tratta di giustapporre le faccette che esprimono la semantica delle varie versioni mediali.

### ***Rappresentare cosa: i livelli rappresentativi***

Per apprezzare il discorso di Hunter ed ammettere la possibilità di una vera integrazione tra informazioni strutturalmente eterogenee è necessario risalire ai contesti multimediali nei quali si sviluppano le varie semantiche e verificare se si possa parlare di integrazione o solo della introduzione di un nuovo punto di vista conoscitivo.

L'approfondimento di quest'ultimo aspetto ci porta ad aprire una parentesi ed osservare i vari significati del termine "rappresentazione".

Notiamo per esempio che, una cosa e' avere una visione (rappresentazione filosofica) del mondo e tratterla, altro e' rappresentare un aspetto particolare della realta' che ci circonda: un terremoto, un sasso che si incontra per la strada, un manufatto, oppure una storia di fatti accaduti dove si incontrano anche terremoti, sassi e quant'altro.

Cosa diversa ancora e' rappresentare la stessa realta' che ci interessa con linguaggi diversi. Se cambiamo linguaggio cambia anche la semantica. Pensiamo per es. ad un film ed a un racconto che trattino dello stesso episodio.

La rappresentazione di una determinata realta' e' un processo e si compie con una serie di passaggi. Se prendiamo per esempio la parola evento e la riempiamo di contenuto: il terremoto di Messina del..... possiamo fare ricerche storiche sul suo conto, ricostruire l'evento, e parlarne. Se le varie descrizioni su quell'evento collimano in alcuni punti, questi costituiranno la base per imbastire intorno l'idea (rappresentazione) di quell'evento che potremo poi, in seguito, comunicare.

Se di fronte a noi c'e' un manufatto e ne vogliamo parlare dobbiamo indicare quali aspetti lo caratterizzano, e dare loro un peso rappresentativo; nel fare questo prendiamo spunto dalla conoscenza di altri manufatti simili per evidenziare le somiglianze e le differenze dei vari aspetti.

Il procedimento ha come risultato una rappresentazione del manufatto.

Questo sforzo di analisi non e' meramente speculativo e fine a se stesso, perche' finalizzato alla possibilita' di comunicare e confrontarsi.

La comunicazione di cui i linguaggi sono tramite si realizza con il linguaggio parlato, con la scrittura, con un disegno, una stampa, e se l'evento e' attuale, con la fotografia, un film, o una realta' virtuale etc.; cioe' con tutti i linguaggi a nostra disposizione compresi quelli artificiali che dotati della capacita' di rappresentare un pensiero, una idea una immagine etc. siano funzionali alla comunicazione.

Nell'universo dei sistemi informativi, l'oggetto dell'indagine pero' non e' la rappresentazione concettuale e filosofica del mondo ma quella caratterizzante uno specifico dominio, interpretato secondo le finalita' che si vogliono raggiungere ed il linguaggio di rappresentazione adottato.

Infatti, per uno stesso dominio la rappresentazione risultera' diversa a seconda degli aspetti che si vogliono mettere in risalto (entita', relazioni) e del linguaggio di comunicazione usato.

Per es. il funzionamento di un sistema puo' essere descritto in un rapporto tecnico e/o in una grafica; se spiegato ad un operatore gli aspetti messi in rilievo risulteranno diversi da quelli che si evidenziano per rivolgersi ad un utente.

Ancora, uno stesso evento potra' avere interpretazioni diverse dipendenti dal linguaggio espressivo usato: un poeta, un musicista, un narratore, un documentarista usano linguaggi espressivi diversi. Quindi i vari linguaggi non solo hanno sintassi e grammatiche diverse, ma svolgono funzioni diverse, in quanto evidenziano aspetti significativi ognuno nell'ambito in cui si trova ad operare.

I linguaggi espressivi dell'uomo sono tra i piu' versatili nel catturare le intenzioni ed emozioni individuali. Tra questi, il linguaggio naturale parlato o scritto e' il piu' diffuso; ma uguale potenzialita' espressiva hanno altri linguaggi legati a manifestazioni come le "fine arts".

I linguaggi tecnici, specialistici, nati da una negoziazione del significato scientifico-tecnico dei termini (siano essi concetti oppure relazioni), sono per loro natura piu' rigidi.

I linguaggi di rappresentazione basati sulla logica, la matematica, i linguaggi documentari codificati sono un'altra cosa. Essi costituiscono il tentativo di restituire in una visione sintetica e formale la realta' di interesse; sono linguaggi preminentemente simbolici.

Il desiderio di rappresentare la realta' con metodologie formali e' primordiale: anche il linguaggio parlato una volta consolidato e' divenuto oggetto di rappresentazione simbolica mediante l'uso e la combinazione di caratteri (alfabeto), ideogrammi, etc.

Con mezzi diversi l'uomo si e' impegnato per ottenere una visione del mondo: con l'osservazione, con l'ideazione, con l'argomentazione (filosofia) ed ha creato i linguaggi per poterla trasmettere.

Si e' anche servito di strutture linguistiche per imprimere veridicit  alle proprie tesi (retorica). Fin dall'origine i linguaggi di rappresentazione hanno svolto una funzione probante: vari formalismi sono stati creati per provare la verita' di un assunto (logica, matematica). Dunque linguaggi vincolati ad una struttura, ad una logica, a leggi, per interpretare formalmente il significato degli enunciati.

In un secolo particolare, il XVIII<sup>o</sup>, e' prevalso l'interesse per l'aspetto gnoseologico. La ricerca ha puntato alla creazione di metodi per rappresentare in modo sintetico la conoscenza acquisita sul mondo circostante. Per ottenere questa rappresentazione, il sapere doveva essere categorizzato, settorializzato, sistematizzato, e per averne una visione di insieme doveva potersi codificare. In questo secolo, detto Illuminismo, si svilupparono le enciclopedie, le terminologie, gli schemi di classificazione.

Quando parliamo di classificazione e di codifica, parliamo di sistemi di rappresentazione espressi in simboli con i quali identificare gli oggetti di interesse. Se in un sistema di classificazione sciogliamo un codice identificativo troviamo un concetto, e se il sistema e' gerarchico (ad albero) avremo una diramazione di codici e di termini con i quali articolare il concetto.

Tale indirizzo diretto a "organizzare la conoscenza" ha dato forma ad una varieta' di linguaggi di rappresentazione (soggettari, tassonomie, thesauri, ontologie, etc.) e si e' evoluto con l'accrescersi di strutture dedicate alla raccolta di materiali (documenti ed oggetti) come biblioteche, archivi, e musei.

Una delle ragioni del successo di questo indirizzo prende avvio dalla convinzione che la classificazione del sapere, oltre a fornire un quadro sinottico delle varie branche della scienza, possa facilitare il recupero delle informazioni ad essa sottesa.

Se alla categorizzazione di un sapere specialistico fosse corrisposta anche una precisa collocazione logica e fisica del materiale di riferimento, il lavoro di ricerca delle informazioni sarebbe risultato avvantaggiato. Inoltre la rete di relazioni tra codici che si potevano imbastire avrebbero arricchito il loro contenuto semantico e permesso un recupero della informazione piu' preciso.

Nell'era informatica, nella quale l'obiettivo del recupero differenziato della informazione e' divenuto preminente, le relazioni /conoscenza della realta' di interesse/selezione e organizzazione dei dati/ loro codifica/ sono individuate in funzione del tipo di informazione da recuperare e delle modalita' di recupero.

Queste prime tappe dell'evolversi dell'organizzazione della conoscenza sono ormai storia; i risultati pero' sono presenti nelle tecnologie e metodologie correnti di rappresentazione e recupero della informazione.

### ***Hunter-Doerr: quali premesse per l'interoperabilita'?***

Riprendiamo ora il discorso della interoperabilita' semantica e dell'integrazione delle informazioni che sono alla base dei lavori di M. Doerr e J. Hunter.

La realta' presa in esame e' in ambedue quella museale. L'obiettivo e' il recupero della informazione che vi si riferisce, qualsiasi siano le risorse disponibili, dovunque si trovino ed in qualsiasi modo siano esse espresse.

La semantica che si vuole catturare e' il contesto informativo che contribuisce ad arricchire di significato l'informazione richiesta.

I punti di partenza dei due autori sono tuttavia diversi. Lo schema concettuale di Hunter e' in certo qual modo una visione generale, perche' adattabile a diverse interpretazioni della realta'. Infatti Hunter non rinuncia ad una propria argomentazione sulla rappresentazione dell'esistente e tenta di farvi convergere un modello funzionale.

Doerr non presenta un proprio schema concettuale comprensivo di una interpretazione della realta'; ne costruisce uno (ontologia) ad hoc (dedicata ai musei) estraendo concetti basilari dagli schemi concettuali oggi in uso per descrivere quella realta'. L'intenzione che persegue e' la costruzione di un linguaggio (ontologia) che sia ad un tempo l'integrazione di tutti i linguaggi documentari possibili che insistono sullo stesso dominio e dall'altro capace di colmare le lacune esistenti in ciascuno di essi.

Questo approccio pragmatico pare scontrarsi col modello di rappresentazione seguito da Hunter che si avvale di un approccio intellettualistico nella scelta delle classi, sottoclassi e relazioni e di un certo eclettismo nella appropriazione di concetti e metodi estrapolati dalle *Foundations*.

Questi due tentativi in grande misura dissimili, sottendono tuttavia la comune convinzione che l'interoperabilit  per essere effettiva richiede un comune ed unico intendimento concettuale e metodologico sul come rappresentare il mondo che ci circonda (modello), oppure, a livello pragmatico, una operazione di integrazione e sintesi dei linguaggi documentari esistenti, dominio per dominio.

In conclusione:

- ambedue gli indirizzi si iscrivono nell'ambito della ricerca documentaria (concettuale o manuale che dir si voglia), sensibile al tema della semantica e quindi vigile sull'aspetto del contesto in cui l'informazione si colloca;
- in ambedue gli indirizzi si ritiene che la possibilit  di un'effettiva interoperabilit  nello sviluppo di sistemi informativi dipenda dall'introduzione, fin dalle premesse, di accorgimenti metodologici tali da uniformare i criteri di rappresentazione del dominio (modello universale, integrazione di schemi o ontologie).

### ***Hunter-Doerr: quale il livello prescelto per rappresentare la realt  museale?***

Nel caso specifico del modello ABC (J.Hunter, C. Lagoze) vi   una interpretazione della realt  incentrata sul concetto di *evento* rivisitato ed adattato all'uso informatico sulla base di estrapolazioni dottrinali.

In CIDOC (M. Doerr et al.) l'attenzione   rivolta alla costruzione di una complessa rete semantica capace di cogliere nel tempo le variazioni dei valori delle propriet  degli oggetti e delle loro relazioni e pone in primo piano le entit  *temporali* e con queste *gli eventi*.

Hunter e Doerr nel condividere l'esperienza CIDOC non fanno riferimento ad una particolare visione del mondo, ad una rappresentazione di come esso si muova, ad una filosofia.

Si pongono sul piano analitico di un sistema informativo settoriale, e nonostante alcune rimarchevoli diversit  d'impostazione, individuano nell'*evento* l'epicentro informativo, foriero di trasformazioni descrivibili in una dimensione temporale.

Tuttavia, sul piano della sperimentazione, CIDOC, come modello di dati estensibile orientato all'oggetto ed espresso nello schema RDF si presentava in una fase evolutiva pi  matura rispetto ad ABC, ed offriva una ontologia (con un corredo di specifiche dettagliate gi  disponibili, comprensive delle descrizioni riguardanti le definizioni di classe, propriet , gerarchie e relazioni) tale da costituire una base appetibile per ulteriori estensioni.

### ***L'ontologia CIDOC versus l'ontologia MPEG***

L'estensione che Hunter persegue, come si   detto all'inizio,   la verifica della capacit  di CIDOC CRM di descrivere risorse multimediali e di integrare quando necessario le superclassi e le super-propriet  di CIDOC con le specifiche sotto-classi e sotto-propriet  multimediali di MPEG-7.

Questa operazione di grande interesse, che mette a confronto due ontologie *domain-specific* sviluppate per standardizzare la semantica associata alla descrizione degli oggetti museali l'una, ed alla descrizione del contenuto multimediale l'altra, viene condotta a livello pragmatico nell'intento di elaborare uno standard atto a descrivere il contenuto multimediale museale.

### ***Le carenze di CIDOC***

Il pragmatismo analitico, che si riflette in tutto lo sviluppo di CIDOC, permette agli autori di asserire che le capacit  descrittive di CIDOC sono deboli o assenti per i seguenti tipi di informazioni sugli oggetti:

- informazioni per l'identificazione, per l'acquisizione e la propriet 
- informazioni relative ai movimenti ed alla locazione fisica, agli attributi ed alle caratteristiche fisiche
- descrizione di oggetti digitali e contenuti digitali multimediali

- definizione di attributi di formattazione per immagini digitali, audio, video
- caratteristiche fisiche audio-visive come istogrammi, regioni, tessitura, volume, forma
- sommari sequenziali o gerarchici di contenuto audio-visivo come fotogrammi, e cambiamenti di scena

Il *documento* in CIDOC, infatti, può essere una registrazione /visiva/audio/audio-visiva; tuttavia non esiste un supporto descrittivo per i singoli media.

### **MPEG-7**

MPEG-7, è un gruppo di lavoro del ISO/IEC [15] dal quale si attende lo standard finale per generare e comprendere le descrizioni audiovisive.

MPEG-7 attualmente dispone di

- un nucleo di descrittori (Ds) da usare per descrivere le caratteristiche del contenuto multimediale
- strutture pre-definite di descrittori e loro relazioni, Schemi di descrizione (DSs)

Il target che il gruppo si prefiggeva era ottenere (i) un efficiente recupero della informazione dagli archivi digitali e (ii) filtrare su Internet il flusso audiovisivo delle emittenti radiotelevisive.

MPEG-7 descrive l'informazione audiovisiva senza badare ai problemi di memoria, codifica, trasmissione, mezzi e tecnologie. Si rivolge ad una varietà di tipi di media quali le immagini fisse, la grafica, i modelli tridimensionali, l'audio, video, e le loro combinazioni.

Le definizioni MPEG-7 erano espresse inizialmente in XML.

La necessità di assicurare l'interoperabilità semantica sufficiente a permettere lo scambio dei dati tra sistemi, la comprensione del significato dei dati, e la loro traduzione e integrazione ha spinto il gruppo MPEG-7 a definire formalmente la semantica dei termini MPEG-7 ed a esprimere le definizioni nel linguaggio RDF [17], ritenuto il più valido a rappresentare la semantica e le relazioni semantiche tra le classi e le gerarchie delle proprietà.

Di seguito, una tipica collezione multimediale all'interno di un museo:

immagini	fotografie, stampe, mappe, manoscritti, documenti, schizzi, disegni, poster
audio	canzoni, musiche, interviste, racconti orali, programmi radio, parlato, letture, registrazioni
video/film	film, documentari, clip, riprese antropologiche e/o naturalistiche, animazioni
grafica	modelli tridimensionali, percorsi simulati, siti archeologici, VRLM
multimedia	presentazioni, slide, file

E di seguito un elenco di categorie di metadati per descrivere le risorse multimediali:

- *metadati bibliografici* - sono le informazioni sulla creazione/produzione della risorsa (data, luogo), gli individui e le organizzazioni coinvolte, e la composizione catalografica della risorsa
- *metadati di formattazione* - includono le informazioni sui formati, sulla codifica, sulla registrazione ed i requisiti del sistema
- *metadati strutturali* - si riferiscono alla scomposizione strutturale della risorsa multimediale in segmenti temporali, spaziali, spazio-temporali ed alle relazioni tra questi
- *metadati per indicizzare il contenuto* - possono essere descrizioni in linguaggio naturale o caratteristiche audio o video come istogrammi di colori e volume.
- *metadati per descrivere il tempo di vita di una risorsa* - includono informazioni relative all'acquisizione e ricollocazione di una risorsa, ad eventi di riformattazione, repackaging, distribuzione, ad eventi relativi all'uso di una risorsa, ad agreement sui diritti di riproduzione e permessi

### L'ontologia MPEG-7

Il gruppo *ad hoc* MPEG-7 per l'interoperabilità ha sviluppato, in seguito, sulla base di queste definizioni, l'ontologia MPEG-7 esprimendola nello schema RDF e DAML + le estensioni OIL, necessarie per soddisfare requisiti quali il bisogno di campi multipli, ed i vincoli specifici per le sotto-classi [18].

L'ontologia MPEG-7 si compone di due classi fondamentali (entità) il *Contenuto multimediale* ed i *Segmenti* che si sviluppano entrambi in relazioni *part of*. Le parti dei segmenti costituiscono un livello ulteriore di *part of* del contenuto multimediale.

Vale a dire: se *Audio* e *Video* sono parti della classe *Contenuto multimediale*, *Video segment* ed *Audio segment* sono rispettivamente sotto-classi di *Audio* e *Video* e contemporaneamente sotto-classi di *Segment*

In questo modello di tipo *entity-relationship*, le relazioni che si stabiliscono tra i diversi tipi di segmento e le entità multimediali sono dipendenti dai seguenti tipi di decomposizione.

- Decomposizione spaziale (per es. regioni spaziali all'interno di una immagine)
- Decomposizione temporale (per es. segmenti temporali di un video)
- Decomposizione spazio-temporale (il movimento di regioni all'interno di un video)
- Decomposizione della sorgente mediale (per es. le differenti bande sonore dentro un file audio)

A ciascuna delle sotto-classi menzionate sono associate varie proprietà che definiscono le relazioni permesse tra le classi del *Segmento*, (vale a dire dipendenti dagli schemi di descrizione usati) e gli attributi audio, video, e audiovideo permessi, associabili ai diversi tipi di segmenti multimediali.

Type	Feature	Descriptors
Visual	<b>Color</b>	DominantColor
		ScalableColor
		ColorLayout
		ColorStructure
		Extension ofColorStructure
	<b>Texture</b>	Homogeneous Texture
		TextureBrowsing
		EdgeHistogram
	<b>Shape</b>	RegionShape
		ContourShape
		Shape3D
	<b>Motion</b>	CameraMotion
		MotionTrajectory
		ParametricMotion
MotionActivity		
Audio	<b>Silence</b>	Silence
	<b>Timbre</b>	InstrumentTimbre
		HarmonicInstrumentTimbre
		PercussiveInstrumentTimbre
	<b>Speech</b>	Phoneme
		Articulation
		Language
	<b>Musical Structure</b>	MelodicContour
		Rhythm
	<b>SoundEffects</b>	Reverberation, Pitch, Contour, Noise

Tavola 1 – Caratteristiche Audio e Video MPEG7 ed i corrispondenti descrittori

L'organizzazione per aree delle strutture di descrittori (DSs) Multimediali MPEG7 permette di descrivere un contenuto multimediale comprensivo di:

- Informazioni sulla creazione e processo di produzione; sull'impiego del contenuto; sulle caratteristiche di memorizzazione
- Informazioni strutturali sulle componenti temporali, spaziali, spazio-temporali del contenuto; sulle caratteristiche di basso livello del contenuto; informazioni concettuali e semantiche della realta' espressa nel contenuto
- Informazioni sulle modalita' di *browse*; sull'organizzazione delle collezioni di oggetti, e dei modelli
- Informazioni sull'interazione degli utenti con il contenuto

### ***Il confronto tra le due ontologie***

Dal confronto delle ontologie CIDOC e MPEG-7 viene rilevato che:

- l'ontologia CIDOC e' meglio organizzata per descrivere artefatti fisici museali ed eventi del mondo reale e la base di conoscenza ad essa sottesa e' piu' facilmente scomponibile in proposizioni atomiche di senso compiuto.

La terminologia CIDOC infatti, basata su un modello orientato all'oggetto di tipo gerarchico dove le entita' *eventi* e *attivita'* sono centrali, crea una impalcatura particolarmente adatta a registrare eventi, il cambiamento dei valori degli attributi, e le relazioni dinamiche che caratterizzano le risorse.

- l'ontologia MPEG-7 e' meglio attrezzata per le descrizioni granulari del contenuto multimediale, in particolare quello digitale.

La standardizzazione delle descrizioni favorisce inoltre la ricerca, il recupero ed il filtering di una rilevante quantita' di contenuto multimediale.

### ***La fusione delle ontologie***

La struttura dell'ontologia CIDOC, costituita a livello piu' alto da un set circoscritto di classi e proprieta' dalle quali diramano ontologie di metadati di domino piu' specifico, e' parsa la piu' confacente ad ospitare l'estensione diretta all'inserimento dei metadati multimediali relativi alle componenti specifiche MPEG-7.

In particolare, la classe *document* dell'ontologia CIDOC e' sembrata la piu' adatta a servire come punto di innesto dell'estensione, mentre una sua proprieta' *is composed of* viene estesa per definire metadati strutturali o di segmentazione associati con le risorse multimediali.

Questo approccio permette di fornire tutte le decomposizioni spaziali, temporali e spazio-temporali presenti in MPEG-7.

Le nozioni MPEG-7 *mediaTime* e *media Place*, necessarie a specificare le collocazioni spaziali e temporali all'interno di una risorsa multimediale, diventano sotto-classi delle classi CIDOC *Time-Space* e *Place*.

Altre estensioni di CIDOC riguardano l'inserimento delle proprieta' di formattazione ed i descrittori delle caratteristiche audio e video associabili ai contenuti multimediali ed ai tipi di segmento. Alcuni metadati relativi alla formattazione possono trovare ospitalita' nella classe CIDOC *dimension*, ma altre informazioni devono trovare sistemazione in nuove classi e relative proprieta'.

Nello specifico, i descrittori audio/video previsti in MPEG7 dovrebbero essere raccolti come sotto-classi in due nuove classi CIDOC *visual features* e *audio features*. Inoltre le proprieta' *ha caratteristiche audio* ed *ha caratteristiche video* dovrebbe permettere di associare le nuove classi e sotto-classi ai tipi del documento multimediale.

### ***Commenti***

L'esperienza intrapresa per unificare sistemi di metadati distinti per funzione, anche se di comune dominio di interesse, pone questioni diverse da quelle dell'*efficienza del metodo* sempre presente quando si affronti il tema dell'interoperabilita'.

E' interessante notare infatti che l'approccio di Hunter al problema dell'interoperabilita' si focalizza sulla completezza degli aspetti da prendere in esame nel processo di archiviazione.

Non si tratta infatti di stabilire la bonta' di un modello nei confronti di un altro e di indicare quale dei due garantisca una effettiva interoperabilita', ma di allargare il concetto di interoperabilita' estendendo il contesto della sua applicabilita', (da un contesto mono-semantic ad uno multi-semantic).

Tra i progettisti dello schema MPEG7 si usa distinguere, in tema di descrizione del contenuto, gli aspetti strutturali da quelli semantici. Marcare questa differenza fa risaltare la diversità dei metodi di ricerca ed il loro oggetto, ma non aiuta a precisare l'articolazione complessiva dell'aspetto semantico: aspetto che emerge quando l'uso dei descrittori preposti ad indagare la struttura di un video o di un audio, indica e mette in risalto la semantica che l'autore o gli autori vi hanno voluto imprimere.

L'intersezione tra le diverse tipologie di metadati crea infatti un luogo (volume) dagli inaspettati esiti semantici. Non solo perché soddisfa le esigenze semantiche sia di una utenza interessata ad informazioni bibliografiche che a quella interessata ad informazioni strutturali, ma anche perché permette di espandere l'indagine e precisarla da qualsiasi punto di vista si inizi a costruire la query. Pensiamo alle potenzialità cognitive dell'*overlapping* audio/video/testo.

## Conclusioni

Il calcolo degli eventi di Shanahan è il metodo formale con il quale vorremmo poter rappresentare ogni aspetto della realtà

L'impianto deterministico non lo consente e lo relega all'uso ingegneristico e predittivo per il quale è nato. Le estensioni ad altre realtà fino ad ora sperimentate (linguistica) richiedono grossi investimenti con esiti ancora incerti.

La nozione di *temporalità* nel calcolo degli eventi viene esaltata dalla necessità di esprimere, nella relazione causa/effetto, il binomio temporale dell'inizio/fine causa, inizio/fine effetto; tuttavia è proprio la relazione causale a rilevarsi limitativa nelle applicazioni.

Resta il fascino del concetto di *fluente* che riesce a rappresentare in modo intuitivo ed esemplare qualsiasi cosa il cui valore sia soggetto a cambiamento, sia essa *proprietà* che muti di valore od *oggetto* che si modifichi parzialmente.

Questa caratteristica del linguaggio di Shanahan fa pensare ad un suo possibile uso nella valutazione dei sistemi informativi (dinamica dei flussi informativi, modifica del loro contenuto, tempistica, etc.).

ABC e CIDOC sono due ipotesi di modello riferite a comunità archivistiche. L'una protesa verso la generalizzazione del modello; l'altra concepita a misura delle esigenze archivistiche, storiche e archeologiche. Vengono analizzate assieme perché contemporanee e ambedue mirate alla soluzione dei problemi della *temporalità* e della *interoperabilità*

ABC ha il merito di essere stato uno dei primi modelli a dare risposta alle richieste di "temporalità" provenienti dall'ambiente archivistico ed a utilizzare il concetto di *evento* come base del modello, sulle orme di quanto avvenuto in INDECS per l'e-commerce

Nella ragionevole convinzione che "l'interoperabilità" si dovesse basare soprattutto su di una comune intesa sui criteri di rappresentazione dell'oggetto e dei suoi cambiamenti, le varie versioni di ABC rappresentano l'esito di altrettante negoziazioni effettuate e perseguite all'interno di un concerto di Istituzioni e Progetti (Harmony).

Con propositi di generalizzazione ambiziosi, la versione qui esaminata offre spunti notevoli sul versante della *modellazione del "cambiamento" di un oggetto nel tempo* e su quello *delle funzionalità interoperative*, impostate queste ultime sulla scelta di entità e relazioni ad impatto non sempre intuitivo.

La categoria *Temporality* composta da *azioni*, *eventi*, *situazioni*, ne modella i tempi ed indica l'evoluzione dei valori delle proprietà. *Event* marca scenari di mutamento, la cui granularità è variabile, *Situation* indica situazioni di stabilità tra un evento e l'altro nelle quali asserire le proprietà esistenziali dell'oggetto.

La necessità di una sottocategoria come *Situation* che esprime lo stato delle cose dell'oggetto prima e dopo un evento, non appare chiara, così come l'interesse da attribuire ai valori della faccetta esistenziale.

Spiace che l'interruzione dell'iter delle revisioni non abbia consentito di semplificare e perfezionare l'ontologia del modello.



CIDOC e' la testimonianza di una felice combinazione tra una visione informatica ed una pragmatica: la' dove la scienza non arriva una buona dose di manualita' fa il resto.

Colpisce infatti il lavoro ostinato e paziente teso a generare l'ontologia formale top-level (livello delle astrazioni) distillando informazioni dalle strutture dati esistenti e dall'esperienza degli esperti; colpisce inoltre l'impegno a garantire una compatibilita' (integrazione) semantica delle informazioni ai fini di una *interoperabilita'* possibile, e la determinazione a seguire metodi di normalizzazione delle strutture dati senza perdita di informazioni.

La nozione di *temporalita'* e' esaltata nel modello dalla presenza di *gruppi logici di proprieta'* che permettono, una volta modellato un evento come entita' temporale, ragionamenti cronologici su tutto quello che vi si riferisce, siano essi oggetti, attori, o entita' non temporali.

Ne segue che il recupero dell'informazione e' inteso in una accezione in cui la informazione e' estesa al suo contesto semantico. Questo ultimo e' reso visibile da un sistema a rete dove ogni faccetta di un argomento, di un concetto o di un evento, seppure espresso in strutture distinte, diventa raggiungibile.

Un invito quindi alla ricerca (search) all'interno di un sistema sofisticato dove si considera, per esempio, il termine "autore" privo di significato se non viene espresso "di cosa" sia autore; dove il termine "data" e' articolato anche nella sua accezione di "datazione"; e dove, per definire la temporalita' si scelgono metodi di rappresentazione distinti a seconda del dominio di interesse: talche', in campo archeologico, la modellazione dell'istante puo' ritenersi superflua, mentre il calcolo basato su concetti temporali come "postquem" e "antequem" diviene necessario.

### **References**

- [1]CIDOC Conceptual Reference Model  
<http://cidoc.ics.forth.gr/>
- [2]Dublin Core Metadata Element Set  
<http://www.dublincore.org/documemts/dces/>
- [3]Library of Congress Subject Headings  
<http://www.loc.gov/>
- [4]Art & Architecture Thesaurus  
<http://www.getty.edu/research/>
- [5] ABC Model: C. Lagoze, J. Hunter, *The ABC Ontology and Model*, Journal of Digital Information, v. 2, issue 2, november 2001
- [6] Davidson, D., *Azioni ed Eventi*, Bologna, Il Mulino, 1992
- [7] G.Booch, J. Rumbaugh, and I. Jacobson, *The unified modelling language user guide*. Reading Mass., Addison-Wesley, 1999
- [8] Shanahan, M. *The Event calculus explained*, in Artificial Intelligence Today, ed. M.J. Wooldridge and M. Veloso, Lecture Notes in Artificial Intelligence 1600, 1990
- [9]M. Doerr et al., *Supporting Chronological Reasoning in Archeology*, CAA 2004, Prato, Italy
- [10] C.Lagoze, J.Hunter, and D.Brickley, *An Event-Aware Model for Metadata Interoperability*, ECDL, 2000, LNCS 1823
- [11] Dublin Core Metadata Initiative  
<http://dublincore.org>
- [12] CIMI Consortium  
<http://www.cimi.org>
- [13]IFLA FRBR  
<http://www.ifla.org/>
- [14]INDECS Home Page  
<http://www.indecs.org/>
- [15]MPEG-7 Overview  
<http://mpeg.telecomitalia.com/standards/mpeg-7.htm>
- [16]J. Hunter, *Combining the CIDOC CRM and MPEG-7 to describe Multimedia in Museum*, Museums and the Web 2002
- [17]Semantic Web: Resource Description Framework (RDF)  
<http://www.w3.org/RDF>
- [18] DAML + OIL  
<http://www.daml.org/>