

Unsupervised modeling of mutational landscapes of adeno-associated viruses viability

*Original*

Unsupervised modeling of mutational landscapes of adeno-associated viruses viability / DE LEONARDIS, Matteo; Fernandez-de-Cossio-Diaz, Jorge; Uguzzoni, Guido; Pagnani, Andrea. - In: BMC BIOINFORMATICS. - ISSN 1471-2105. - 25:(2024). [10.1186/s12859-024-05823-5]

*Availability:*

This version is available at: 11583/2993182 since: 2024-10-08T12:49:16Z

*Publisher:*

BMC

*Published*

DOI:10.1186/s12859-024-05823-5

*Terms of use:*

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

*Publisher copyright*

(Article begins on next page)

## **Strutturazione dei dati di un GIS in ontologia: applicazione per la mappatura del degrado di un caso studio**

Francesca Noardo

Politecnico di Torino, DIATI, c.so Duca degli Abruzzi, 24 – 10129 Torino.  
email: francesca.noardo@polito.it

### **Riassunto**

Lo sviluppo delle ontologie applicate in ambito di Sistemi informativi e Computer Science per l'ottimizzazione della rappresentazione dei dati relativi a determinati settori di conoscenza si sta ampliando anche verso il campo dei Beni Culturali. La possibilità di porre vincoli all'interpretazione delle informazioni organizzate in data-base e sistemi informativi (GIS) offre notevoli vantaggi per lo sviluppo e la comunicazione della conoscenza relativa a questo tipo di oggetti.

Tale metodologia è stata applicata a un caso studio, nel corso della ricerca riguardante una possibile ontologia per i Beni Culturali per una tesi di dottorato: una facciata di un ex convento oggi in situazioni critiche di degrado, su cui era stata precedentemente realizzata un'analisi tematica dello stato di conservazione di materiali e paramenti murari superficiali, supportato dall'elaborato fotografico derivante da rilievo metrico e gestito tramite GIS. In questo primo caso i dati erano stati modellati utilizzando uno schema Entità-Relazione (ER) per determinare gli oggetti da rappresentare con le relative proprietà e le relazioni reciproche con i vincoli di cardinalità. Il grado di formalità del linguaggio utilizzato per la modellazione viene incrementato attraverso strumenti logici di livello superiore, che permettono la formulazione di un'ontologia di dati univocamente strutturati per ottenere una minore ambiguità nell'interpretazione dell'informazione contenuta.

### **Abstract**

*The development of applied ontologies in the field of Information Systems and Computer Science for the optimisation of the representation of data about some specific knowledge sector is reaching also the field of Cultural Heritage. The possibility to set constraints to the interpretation of information organized in data-base and Geographical Information Systems (GIS) offers noticeable advantages for the development and the communication of knowledge about this kind of objects.*

*This methodology was applied to a case study, during a PhD research about a possible ontology for Cultural Heritage: a façade of an ex convent, today in a critical degradation, on which a thematic analysis of the degradation state of materials and surfaces, on the base of a metric survey and managed in a GIS was previously realized. In this first study, the data were modelled using an Entity □ Relation (ER) schema in order to determine objects to represent with properties and mutual relations and cardinality constraints. The formality of the language used for the modelling is increased through upper-level logic instrument, which permit the formulation of an ontology of data uniquely structured for obtaining a less ambiguous interpretation on the information.*

### **Introduzione**

In ambito di rappresentazione e gestione della conoscenza in data-base e sistemi informativi è sempre più riconosciuta la necessità di avere come riferimento uno schema concettuale riconosciuto e condiviso che permette di arginare ogni ambiguità di interpretazione dei dati rappresentati. Questa necessità si risolve con la formulazione di ontologie applicate (Guizzardi, 2010), metodo già noto sia in ambito informatico sia in svariati altri campi specifici. Per quanto riguarda lo sviluppo di

ontologie in campo di Beni Culturali, sono disponibili alcuni modelli ontologici, che sarebbero però da integrare per riuscire a rappresentare tutta la conoscenza riguardante questi complessi oggetti appartenenti a categorie anche distanti tra loro, in un unico sistema (beni architettonici, artistici paesaggistici, archivistici ecc.).

Tra i modelli riconosciuti come standard internazionale troviamo il CIDOC-CRM (*Conceptual Reference Model*), definito a partire dal 1996 dalla commissione CIDOC (International Committee for Documentation) dell'ICOM (International Council of Museums) e diventato standard ISO21127:2006 a settembre 2006 (Doerr, Ore, Stead, 2007). Il modello è rivolto in particolar modo alla rappresentazione della conoscenza relativa alle collezioni museali, con l'obiettivo di permettere lo scambio e l'integrazione delle informazioni da fonti differenti secondo uno schema univoco. Il modello comprende entità descrittive dell'informazione temporale, caratterizzata da diversi livelli di precisione/incertezza come necessario per tali oggetti, e spaziale. Quest'ultima viene descritta da informazioni alfanumeriche, senza la componente geometrica, che è invece oggetto di un secondo standard OGC: CityGML. Questo standard prevede un modello semantico comune per la rappresentazione 3D condivisibile della cartografia numerica urbana. Alcune ricerche hanno già tentato di integrare i due modelli (Eide et al., 2008) e altre hanno espanso CityGML permettendo di arrivare a un maggiore livello di dettaglio, indispensabile per la rappresentazione 3D dei Beni Architettonici (Costamagna, 2012). Studi in questa direzione sono stati fatti nella unione dello standard OGC GeoSPARQL, per la rappresentazione ed interrogazione dei dati spaziali utilizzando i linguaggi W3C RDF-OWL e SPARQL, con il CIDOC-CRM (Doerr, Hiebel, 2013), unitamente ad un raffinamento della gestione dei dati temporali, trattando la rappresentazione dei dati su 4 dimensioni.

Un ulteriore modello ontologico utile è stato elaborato da (Cacciotti et al., 2013), MONDIS (*Monument Damage Information System*). Quest'ultimo modello parte dal riferimento del CIDOC-CRM per arrivare alla rappresentazione di un settore applicativo più specialistico: l'analisi del degrado e la pianificazione degli interventi sugli edifici. Oltre a includere il degrado superficiale dei materiali, è particolarmente orientato alla rappresentazione dei danni e rischi strutturali, rispondendo alle esigenze di gestione e documentazione di un aspetto fondamentale degli edifici storici.

L'Obiettivo del presente studio è la traduzione di un precedente sistema informativo spaziale, costruito su un modello entità-relazioni i cui contenuti fanno riferimento a linee guida e standard nazionali (uni 11182), in un linguaggio formalizzato nelle ontologie formali esistenti.

La formulazione di uno standard si propone di assolvere alle esigenze di interoperabilità, scambiabilità del dato e univocità dell'informazione e della sua interpretazione. Perciò il rispetto di codici formalizzati offre vantaggi per la definizione e la condivisione dei contenuti rappresentati, insieme alla possibilità di estendere il modello e la conoscenza del sistema secondo strutture predisposte e collaudate.

### **Il caso studio**

L'oggetto rappresentato è una facciata particolarmente degradata dell'ex convento di Carmelitani del "Colletto", comune di Roletto (TO). Nel precedente studio il fronte era stato mappato, in ambiente ArcGIS, utilizzando come riferimento un modello concettuale strutturato come in figura 1.

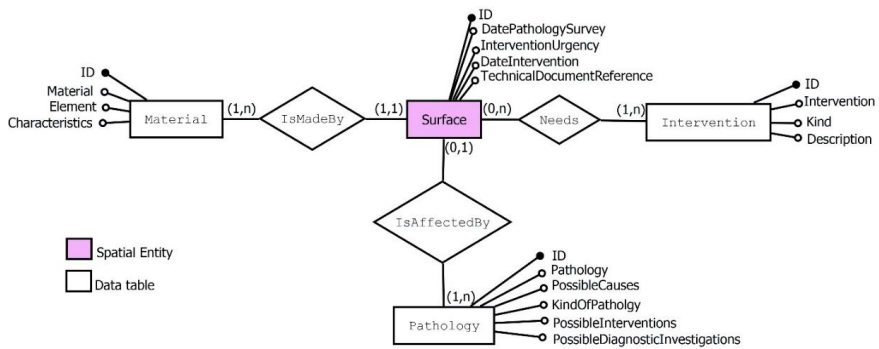


Figura 1 - Modello ER per la strutturazione dei dati sul degrado in GIS.

Il sistema si proponeva di archiviare dati variabili nel tempo in un unico data-base spaziale, per estrarre informazione utile dai dati inseriti e tematizzazioni del fronte secondo le esigenze di un progetto di restauro, arricchite dalla possibilità di visualizzare i dati relativi ad ogni porzione di superficie in un'unica tabella (Figura 2). Le tabelle di dati erano state collegate a ogni porzione di superficie con operatori "Join" e "Relate" di ArcMap, senza prevedere collegamenti molti - a - molti, come invece può essere utile in diversi casi (es. collegamento degradi - superfici).

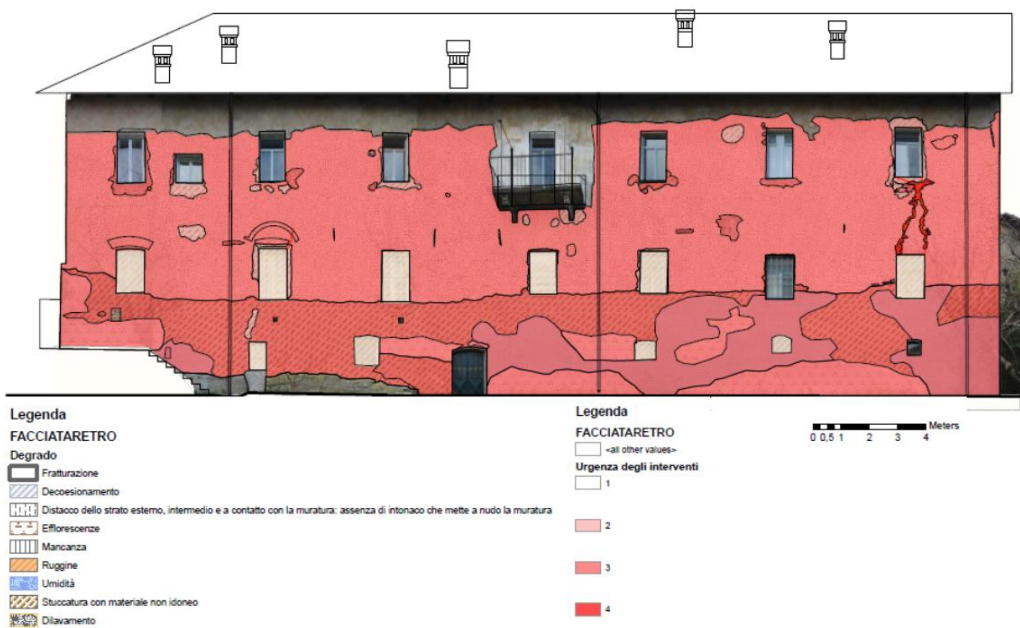


Figura 2 - Tematizzazione della facciata in base al degrado delle superfici e all'urgenza degli interventi.

### L'integrazione del CIDOC-CRM e MONDIS

Nel CIDOC-CRM lo stato di degrado di un oggetto rappresentato è esprimibile in modo generico tramite l'entità *E3\_ConditionState*, ma per la formulazione di un progetto di conservazione di un manufatto complesso, come un'architettura storica, è necessario trattare il fenomeno in modo più articolato, prendendo in considerazione diversi fattori, come si propone di fare il progetto MONDIS. Entrambe le ontologie sono disponibili on line in formato .owl (Standard W3C per la formulazione di ontologie) e si possono leggere con il software per la formulazione di ontologie

*Protégé*. Questo ha permesso di integrare i due schemi estendendo il CIDOC-CRM con le entità e le proprietà definite da MONDIS (Figure 3).

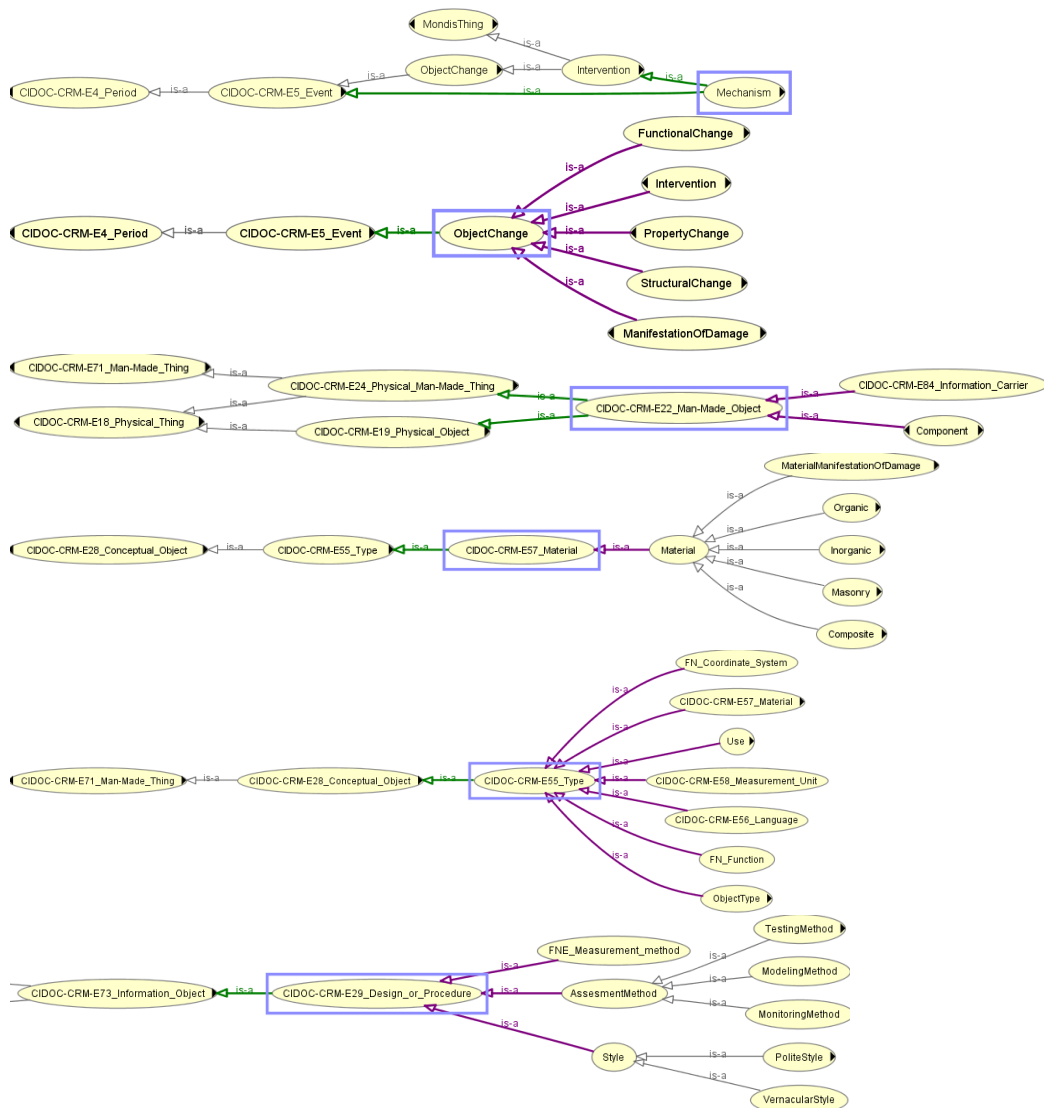


Figure 3 - Visualizzazione di parti dell'ontologia CIDOC estesa con MONDIS.

Altre entità introdotte al primo grado della gerarchia sono “*ComponentValue*” e “*Risk*”, fattori importati nella rappresentazione dei Beni architettonici. Inoltre MONDIS include una tassonomia tratta da bibliografia o precedenti standard di settore, che specifica alcune entità quali “*Intervention*”, “*Mechanism*”, “*ManifestationOfDamage*”, “*Component*”, che a sua volta contiene “*Construction*”, “*Connection*” e “*Surfaces*” specificate in una ulteriore tassonomia, “*Use*”, “*ObjectType*”, diviso a sua volta in “*FunctionalType*” e “*StructuralType*”.

Una volta definita la nuova gerarchia è stato possibile ridefinire le proprietà (relazioni) su tutto il nuovo modello.



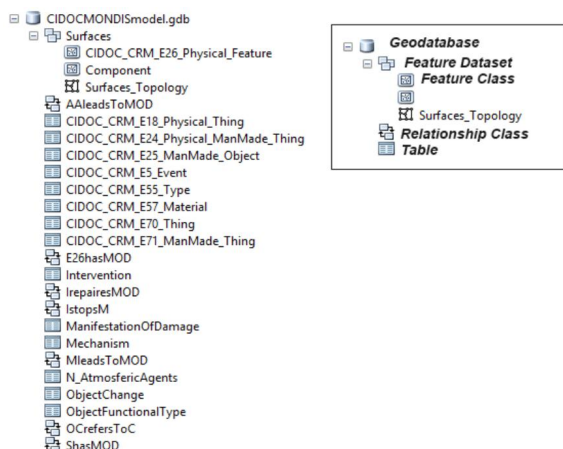


Figura 5 □ Struttura del geodatabase costruito.

Una volta definito in modo univoco il *design* del *geodatabase* è possibile compilare le tabelle ed editare le geometrie secondo necessità, sottostando ai vincoli imposti precedentemente. Tra questi grande importanza per l'integrità dei dati hanno la definizione di *subtypes* e valori di *domain*, che permettono di utilizzare sottoclassi e tassonomie preconfezionate, raggiungendo una maggiore standardizzazione dei dati prodotti (Figura 6) e rispettando le terminologie dettate dalle ontologie di riferimento.

OBJECTID	superclass	Value	subtype
1	<Null>	ArtificialStone	Composite
2	<Null>	Concrete	Composite
3	<Null>	ReinforcedConcrete	Composite
4	<Null>	CementMortar	Composite
5	<Null>	HydraulicLimeMortar	Composite
6	<Null>	LimeMortar	Composite
7	<Null>	<Null>	Inorganic
8	<Null>	ArtificialStone	Inorganic
9	<Null>	Concrete	Inorganic
10	<Null>	ReinforcedConcrete	Inorganic
11	<Null>	CementMortar	Masonry
12	<Null>	HydraulicLimeMortar	Composite

Figura 6 □ Utilizzo di subtype e attribute domain per la compilazione delle tabelle.

Le relazioni prestabilite permettono di interrogare trasversalmente tutte le tabelle collegate (Figura 7), traendo immediato vantaggio dalla lettura di valori prestabiliti da standard ufficiali, norme e bibliografia affermata. Gli operatori del settore possono trarre grande vantaggio dall'aver a immediata disposizione i link, i valori dettati dalle norme, dettagli tecnici relativi a diversi settori (materiali, diagnostica, restauro, storia dell'architettura...) e i dati pregressi. In alcuni casi, si potrebbero anche offrire suggerimenti per la tutela dei manufatti e la buona riuscita degli interventi (possibili indagini diagnostiche, operazioni sconsigliate dagli organi di tutela o dalla natura dei materiali, prodotti a disposizione sul mercato per determinati interventi, etc.), tentando di arginare gli interventi scorretti.

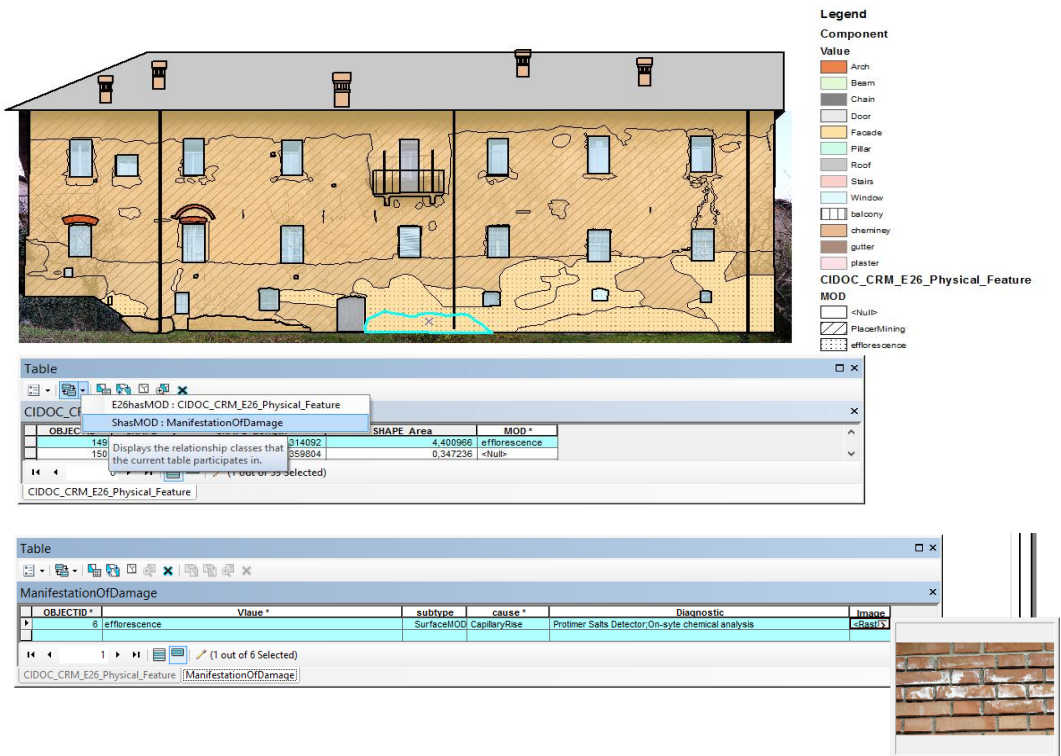


Figura 7 □ Visualizzazione in ArcGIS delle tabelle correlate.

La caratteristica che più rende efficace l'organizzazione di un *geodatabase ESRI* per la gestione della conoscenza secondo le strutture ontologiche è la possibilità di esportarne lo schema (il *design*) in formato XML, per poterlo utilizzare invariato nella gestione di ulteriori GIS senza dover reimpostare le classi, i vincoli e i valori predeterminati delle tassonomie, ma importandone semplicemente la struttura codificata in XML e mantenendo perciò la scambiabilità e l'interoperabilità dei dati.

### Considerazioni finali

Le ontologie applicate rappresentano uno strumento molto efficace per gestire conoscenza complessa in modo completo, operando inferenze derivanti dalla stessa organizzazione dei dati. Anche in campo di beni architettonici questo strumento offre grandi potenzialità, ulteriormente ampliabili una volta risolta l'integrazione con i dati spaziali e l'incorporazione ufficiale dei segmenti di conoscenza ulteriormente specializzabili (come propone MONDIS) da parte del CIDOC-CRM, che per il momento rimane la *core ontology* di riferimento.

Con gli strumenti proposti è possibile definire una struttura costante nel tempo coerente con le strutture definite dagli standard di rappresentazione e gestione dati. Il vantaggio di questo approccio al problema, utilizzando software proprietari largamente diffusi, consiste nel poter utilizzare processi noti e disponibili e *dataset* precedentemente confezionati con tali software in modo più semplice. Un futuro sviluppo della ricerca sarà necessariamente verso l'utilizzo di strumenti *open-source*, per aumentare ulteriormente la possibilità di scambiare dati e di utilizzare strutture comuni definite in linguaggi "open", come sono quelli definiti dagli standard ufficiali (owl, rdf, xml, gml, sparql, geosparql...).

**Riferimenti bibliografici**

Cacciotti R., Valach J., Kuneš P., Čerňanský M., Blaško M., Křemen P. (2013), “Monument Damage Information System (Mondis): an ontological approach to Cultural Heritage documentation”, *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, II-5/W1.

Costamagna E. (2012), *GIS 3D: Studio e applicazione alla documentazione dei beni culturali*. PhD Thesis, Dottorato in Beni Culturali XXIV, Politecnico di Torino, Italy.

Doerr M., Ore Ch.E., Stead S. (2007), “The CIDOC Conceptual Reference Model - A New Standard for Knowledge Sharing”. *Tutorials, posters, panels and industrial contributions at the 26th International Conference on Conceptual Modeling - ER 2007; CRPIT*. Auckland, New Zealand: ACS, 83, 51-56.

Eide Ø., Felicetti A., Ore Ch-E., D’Andrea A., Holmen J. (2008), “Encoding Cultural Heritage Information for the SemanticWeb. Procedures for Data Integration through CIDOC-CRM Mapping”, In *EPOCH Conference on Open Digital Cultural Heritage System*, 2008, 1-7.

Guizzardi G. (2010), “Theoretical Foundations and Engineering Tools for Building Ontologies as Reference Conceptual Models”, *Semantic Web Journal*, IOS Press, Amsterdam.

<http://www.opengeospatial.org/standards/citygml>

<http://www.opengeospatial.org/standards/geosparql>

<http://www.w3.org/2001/sw/wiki/OWL>

World Wide Web Consortium (W3C), *OWL Web Ontology Language Guide*, 2004.  
<http://www.w3.org/TR/2004/REC-owl-guide-20040210/>