

La metodologia e gli strumenti GeoUML per definire le specifiche e validare i contenuti dei database geotopografici

di Alberto Belussi, Jody Marca, Mauro Negri, Giuseppe Pelagatti

La metodologia GeoUML è stata sviluppata al fine di supportare la definizione di specifiche di contenuto per i dati geotopografici e la verifica di conformità di un database o dataset a tali specifiche. GeoUMLcatalogue and GeoUMLvalidator sono gli strumenti, conformi agli standard ISO 19100 e Open Geospatial Consortium (OGC), sviluppati per applicare la metodologia presentata.

La metodologia e gli strumenti GeoUML sono stati progettati e sviluppati per supportare la definizione delle Specifiche di un Database Geotopografico, il cui costituente principale è lo Schema Concettuale, e per eseguire una validazione automatica di un Data Product (con Data Product si intende un qualsiasi insieme strutturato di dati geotopografici, sia in forma di Files che di Database) relativamente a tale Schema Concettuale. I principi fondamentali che hanno guidato il progetto sono stati:

- l'adesione agli standard ISO 19100
- la capacità di essere utilizzata nel contesto della tecnologia attuale e del prossimo futuro
- l'indipendenza da ogni strumento GIS (commerciale o opensource)
- il mantenimento di una distinzione chiara tra il livello concettuale e il livello dell'implementazione fisica

L'architettura generale degli strumenti GeoUML è mostrata in figura 1. Lo strumento che supporta la definizione dello Schema Concettuale (SC) si chiama *GeoUML Catalogue* e il modello concettuale utilizzato per tale definizione è il *Modello GeoUML*. Catalogue e Modello sono descritti nel paragrafo *Modello GeoUML e GeoUML Catalogue* di questo articolo. Uno schema concettuale può essere esportato dal catalogue in un formato XML pubblicato e trasferito allo strumento che effettua la validazione dei dati rispetto a uno Schema Concettuale, che si chiama *GeoUML Validator* ed è descritto nel rispettivo paragrafo. I dati da validare devono essere fisicamente rappresentati secondo un mo-

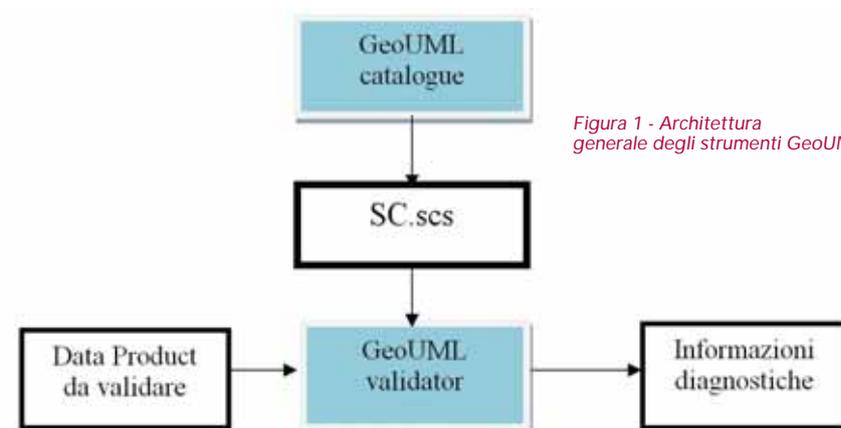


Figura 1 - Architettura generale degli strumenti GeoUML

dello fisico dipendente dalla tecnologia utilizzata (ad esempio, database SQL, GML, Shapefiles); le regole di trasposizione di uno schema concettuale in una particolare tecnologia sono dette *Modello Implementativo*. I Modelli Implementativi progettati e sviluppati finora sono descritti nell'omonimo paragrafo *I Modelli Implementativi*.

Infine, nel paragrafo *Esperienze ed evoluzioni* si riportano alcune considerazioni sulle esperienze realizzate e sulle possibili evoluzioni della metodologia e degli strumenti GeoUML.

Modello GeoUML e GeoUML Catalogue

Il modello GeoUML consiste dei seguenti elementi:

- tutti i costrutti dei "class diagrams" del linguaggio UML (conceptual schema language ISO 19100) con alcune restrizioni relativamente all'impiego dei DataTypes e alla ereditarietà multipla;

- un "profile" dello Spatial Schema ISO (ISO 19107) detto Extended Simple Feature (ESF)
- un insieme di "strutture prefabbricate" (template OCL) per la definizione di vincoli topologici;
- un insieme di costrutti opzionali aggiuntivi, che corrispondono a delle strutture di schemi UML, anch'esse "prefabbricate", per semplificare la rappresentazione di alcune situazioni molto comuni nei dati geografici.

Si noti che l'impiego di strutture prefabbricate i cui componenti fondamentali appartengono agli standard citati permette, se necessario, di decomporle ottenendo esattamente strutture standard. Il profilo ESF è una estensione del modello *Simple Feature Access* definito nello standard ISO 19125 adottato dalla maggior parte dei sistemi GIS attuali; le estensioni mirano a trattare alcune delle evoluzioni della tecnologia rispetto al modello *Simple Feature Access* che purtroppo non sono ancora state accolte nello standard, e cioè:

- punti e curve 3D, incluse le relazioni spaziali tra queste geometrie; si noti che in GeoUML le relazioni spaziali tra le primitive 3D sono interpretate in 3D (ad esempio, 2 curve 3D si intersecano solo se la loro intersezione esiste in 3D), mentre molti sistemi attuali consentono di rappresentare curve 3D ma interpretano le relazioni spaziali sulla loro proiezione planare;
- superfici 2D con una frontiera 3D: questo tipo di componente spaziale mira a fornire una interpretazione rigorosa di alcune implementazioni fornite dai sistemi GIS correnti (ad esempio i cosiddetti "polygonZ" degli Shapefiles)

L'espressione dei vincoli fa riferimento ai tipi 3D indicati e quindi permette di esprimere vincoli in 3D. Gli stessi vincoli sono estremamente difficili da scrivere e interpretare utilizzando direttamente il linguaggio UML/OCL (OCL = *Object Constraint Language* è la componente di UML utilizzata per esprimere vincoli su uno schema di classi UML) e la loro validazione automatica è ancora parziale e limitata a prototipi di ricerca.

Il seguente esempio mostra l'applicazione dei vincoli GeoUML a un vincolo espresso in forma testuale nel documento INSPIRE "D2.8.1.7 *Data Specification on Transport Networks*"; il vincolo in forma testuale è il seguente:

Requirement 10: In a Transport Networks data set which contains nodes, these nodes shall only be present where Transport Links connect or end.

In GeoUML lo stesso vincolo verrebbe espresso nel modo seguente e sarebbe quindi validabile automaticamente dal GeoUML Validator:

```
TransportNode.Geometry
(TC) exists : (InNetwork
= TransportNode.
InNetwork) TransportLink.
CenterlineGeometry
```

Lo strumento GeoUML Catalogue permette di creare uno Schema GeoUML svolgendo le seguenti funzioni:

- controlla che lo schema sia sintatticamente corretto direttamente nella fase di editing
- permette di aggiungere informazioni descrittive allo Schema Concettuale

- produce una ricca documentazione che può essere utilizzata in situazioni che richiedono una specifica con valenza legale (ad esempio, nell'ambito di un capitolato di produzione)
- può produrre un Application Schema (AS) conforme allo standard ISO 19109 tramite la decomposizione dei costrutti prefabbricati citati sopra
- può esportare e importare l'intera specifica (Schema Concettuale + parti descrittive e struttura dei documenti) in un formato XML pubblicato

GeoUML Validator

Questo strumento è in grado di leggere un Data Product e di controllarne la conformità a una specifica prodotta dal GeoUML Catalogue. Dato che per leggere un Data Product è necessario conoscerne la struttura fisica, la corrispondenza tra lo schema concettuale e la struttura fisica deve essere nota al Validator; questo aspetto verrà trattato nella prossima sezione.

I controlli eseguiti dal Validator riguardano i seguenti aspetti:

- la struttura del Data Product: ad esempio, l'esistenza delle classi e degli attributi, l'ammissibilità dei valori dei domini, l'integrità referenziale, ecc...
- la validità di ogni geometria rispetto alle regole dei tipi geometrici del modello ESF (questo controllo include i controlli 3D, ad esempio che un anello 3D sia chiuso in 3D; questo tipo di controlli non viene normalmente eseguito dai sistemi attuali)
- il soddisfacimento di tutti i vincoli espressi nello schema concettuale

Un aspetto critico dell'implementazione del Validator è costituito dalla gestione degli aspetti metrici e della risoluzione delle coordinate, perché nella tecnologia corrente è possibile che 2 sistemi valutino in maniera diversa le proprietà topologiche degli stessi dati a causa degli arrotondamenti.

Il GeoUML Validator produce una diagnostica che viene memorizzata in un Database Derby dal quale può essere interrogata, prodotta in forma di report analitici e sintetici, ecc... Un aspetto complesso che ha dovuto essere affrontato nella produzione della diagnostica è il seguente: gli errori vengono ovviamente rilevati sui dati, quindi a livello fisico, ma la generazione di una diagnostica riferita alla sola struttura fisica è stata ritenuta insufficiente, in quanto alcuni errori devono poter essere interpretati rispetto allo Schema Concettuale. Pertanto il Validator ricostruisce all'indietro il percorso di trasformazione concettuale-fisico segnalando anche la struttura

concettuale responsabile dell'errore.

Per minimizzare la dipendenza del Validator dalle diverse strutture fisiche dell'input, la sua architettura è basata su un database normalizzato, nel quale i dati vengono trasferiti dal processo di lettura che li trasforma dalla struttura fisica di input nella struttura fisica normalizzata. Il database normalizzato è implementato utilizzando il database opensource PostGIS.

I Modelli Implementativi

Il concetto di Modello Implementativo permette di superare la distanza tra i livelli concettuale e fisico. Ogni *Data Product* ha una struttura fisica che è descritta dal suo Schema Fisico; il modo in cui è definito lo schema fisico dipende dalla tecnologia utilizzata per il Data Product, ad esempio:

- SQL Data Definition Language per database georelazionali
- XML schema definition language (XSD) per file GML
- un insieme di shapefile vuoti per Shapefiles © ESRI

La definizione dettagliata delle corrispondenze tra uno schema concettuale e il corrispondente schema fisico viene definita in una parte della specifica detta *Data Product Specification*, la cui componente più importante è il cosiddetto "Mapping Concettuale-Fisico".

Un Modello Implementativo è un insieme di regole che permette di generare automaticamente lo schema fisico e il mapping concettuale-fisico di un *Data Product*.

La nozione di Modello Implementativo si inserisce nell'architettura degli strumenti GeoUML come mostrato in figura 2. Per ogni modello implementativo MI_x che si vuole poter utilizzare è necessario creare due componenti software:

1. il *Mapping Generator*, vale a dire il modulo che genera a partire dallo schema concettuale e seguendo le regole del modello implementativo lo schema fisico corrispondente
2. il *Reader*, vale a dire il modulo che nel Validator è in grado di trasferire un *Data Product* con una struttura aderente a un Modello Implementativo nel database normalizzato interno.

I modelli implementativi si possono classificare in base a diversi criteri. Prima di tutto i modelli implementativi dipendono dalla tecnologia alla quale fanno riferimento; al momento le tecnologie di riferimento sono principalmente gli *Shapefile*, il GML e i Database Georelazionali (spatial SQL).

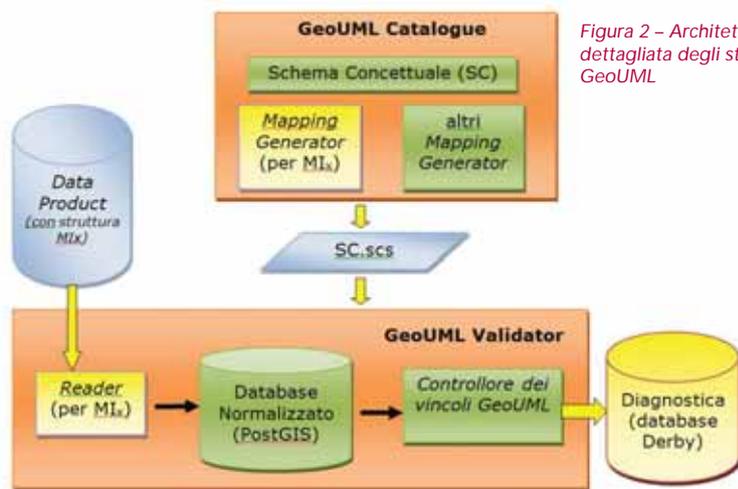


Figura 2 – Architettura dettagliata degli strumenti GeoUML

I MI disponibili negli strumenti sono distinti in base all'uso che viene fatto dei Data Product che li utilizzano:

- di TRASFERIMENTO, per i Data Product utilizzati per lo scambio di dati. Per i modelli di trasferimento è opportuno utilizzare un MI rigido non adattabile, in quanto il destinatario del trasferimento deve essere in grado di conoscere come il Data Product è strutturato. Attualmente sono disponibili:
 - il modello per il GML;
 - il modello Shape_flat (a oggetti) basato su shapefile nel quale ogni oggetto possiede la propria geometria, e non c'è condivisione di primitive geometriche tra oggetti diversi;
 - il modello Shape_topo (topologico) basato su shapefile nel quale le geometrie che devono rispettare vincoli di condivisione sono raggruppate in insiemi topologici che memorizzano le componenti elementari condivise e la geometria degli oggetti non è rappresentata, ma è derivabile a partire dagli insiemi topologici.
- per DATABASE, per i Data Product utilizzati da un Ente per svolgere le sue attività. Per i modelli per database è opportuno utilizzare un MI adattabile, consentendo di personalizzare alcuni aspetti dello schema fisico generato ai requisiti specifici dell'ente che deve utilizzarlo. I MI disponibili si diversificano nella tecnologia adottata e nei vincoli imposti sulla geometria sono:
 - i modelli Oracle e Postgis monogeometria che non ammettono più di una geometria in ogni tabella;
 - il modello Oracle multigeometria che permette di avere più geometrie in ogni tabella.

Tutti i modelli permettono di arricchire il database con vincoli di chiave primaria e

di chiave foreign al fine di garantire l'integrità dei dati.

Infine si sottolinea che in tutti i MI realizzati ad eccezione del GML (nel quale la struttura è nidificata) si è adottata una struttura tabellare non nidificata (simile alla struttura piatta degli shapefile); nei MI per database non sono state quindi utilizzate le tabelle innestate per gestire gli attributi multivalore.

Esperienze ed evoluzioni

La metodologia GeoUML e i suoi strumenti hanno due ambiti di applicazione immediati: la produzione di database geotopografici e il supporto alla loro gestione e alla loro integrazione in infrastrutture di dati territoriali (IDT). Nella produzione di database geotopografici i punti di intervento sono essenzialmente due:

- l'inclusione, nell'ambito del capitolato, dello schema concettuale e del relativo schema fisico, ottenendo un capitolato che prescrive in maniera assolutamente precisa e dettagliata la struttura dei dati di fornitura; inoltre, il GeoUML Catalogue supporta la realizzazione di uno Schema Concettuale conforme al cosiddetto National Core, che è stato realizzato anch'esso con questo strumento
- la validazione della conformità intrinseca dei dati alla specifica concettuale tramite il GeoUML Validator

Nel supporto alla realizzazione dei database geotopografici e delle IDT le aree di intervento della metodologia GeoUML sono ancora in fase di approfondimento e sperimentazione; alcuni aspetti sicuramente rilevanti sono i seguenti:

- la possibilità di definire uno schema concettuale condiviso per l'intera IDT, supportando la realizzazione di database organizzati con strutture fisiche e

tecnologie diverse da parte dei diversi enti partecipanti alla IDT

- la possibilità di validare i dati che vengono trasferiti sulla IDT, permettendo a un ente di controllare i dati che importa da altri enti

Gli strumenti GeoUML costituiscono un esempio di architettura "Schema Driven", i cui strumenti operano in base allo schema concettuale; inoltre si tratta di un'architettura estremamente modulare ed estendibile. Questa architettura permette di progettare nuovi componenti che sfruttino la conoscenza dello schema concettuale dei dati e quindi supportino alcune operazioni di rilevante importanza nella creazione delle IDT, quali, ad esempio, convertitori di modello implementativo e strumenti per il confronto tra schemi diversi, includendo, se possibile, la definizione di corrispondenze tra di loro. In particolare, si sta analizzando la possibilità di supportare il processo di fornitura di dati conformi alla direttiva INSPIRE.

Parole chiave

GeoUML, INSPIRE, GEODATABASE, VALIDAZIONE

Abstract

The GeoUML methodology has been developed in order to support the definition of specifications of geotopographic data and the validation of a database or dataset with respect to such specifications. The GeoUMLcatalogue and GeoUMLvalidator tools have been implemented in order to apply the methodology; they are conformant to the relevant ISO 19100 and Open Geospatial Consortium (OGC) standards.

Autori

ALBERTO BELUSSI
ALBERTO.BELUSSI@UNIVR.IT
DIPARTIMENTO DI INFORMATICA, UNIVERSITÀ DI VERONA, VERONA, ITALY

JODY MARCA
JODY.MARCA@POLIMI.IT
MAURO NEGRI
MAURO.NEGRI@POLIMI.IT
GIUSEPPE PELAGATTI
GIUSEPPE.PELAGATTI@POLIMI.IT

DEI, POLITECNICO DI MILANO, MILAN, ITALY



**VENTO FREDDO.
TERRENO SCIVOLOSO.
UN SALTO DI 30 METRI.
IN POSTI COME QUESTI
SIAMO TOTALMENTE A
NOSTRO AGIO.**

Non importa dove si sta lavorando, il nuovo ricevitore Trimble R10 GNSS misura più velocemente e più facilmente che mai. Trimble R10 utilizza una bolla elettronica per una maggiore precisione e garanzia di qualità, mentre il suo più leggero, design più ergonomico rende quei lunghi giorni sul campo molto più facile. Inoltre, offre una lunga serie di nuove tecnologie delle quali presto non potrete più fare a meno. Il nuovo ricevitore Trimble R10. Perché non sei solo fuori a misurare confini, tu li stai spingendo. Guard al nuovo Trimble R10 su trimble.com/R10showcase.

 **Trimble.**

