La revisione del Corpus di Conoscenze nel progetto Geographic Information – Need to Know

Laura Berardi (*), Sergio Farruggia (**), Beniamino Murgante (***),
Mauro Salvemini (**), Monica Sebillio(****)

(*) Università Roma -Sapienza, Piazza Borghese, 9 - 00186 Roma

Tel. +39.064991.8834/1, fax.+39.064991.8873, mail: laura.berardi@uniroma1.it

(**) AMFMGIS Italia, Piazza Borghese 9, 00186 Roma

Tel. +39.06.49918834, Fax: 178 2278320, mail: info@amfm.it

(***) Università della Basilicata, 10, Viale dell'Ateneo Lucano, 85100 - Potenza

Tel. +39-0971-205125, fax +39-0971-205185, mail: beniamino.murgante@unibas.it

(****) Università di Salerno, via Giovanni Paolo II n. 132, 84048 Fisciano (SA)

Tel. +39.0825.963375, mail: msebillio@unisa.it

Riassunto

Il progetto GI-N2K -- Geographic Information Need to Know - è ormai giunto nelle fasi operative più significative, quella di Revisione del Body of Knowledge (BoK) e quella di progettazione dell'ambiente Virtual Laboratory for the BoK (VirLaBok) atto a consentire una consultazione online della versione rivisitata del BoK. Dopo l'iniziale analisi integrata della domanda e dell'offerta di istruzione e di formazione geospaziale, i membri del Consortium sono ora impegnati nell'individuazione dell'opportuna strategia di revisione, sia in termini di contenuto che di fruibilità dello stesso. In questo lavoro sono introdotti i risultati iniziali di questa fase evidenziando le problematiche sorte riguardo le eventuali sovrapposizioni tra aree di conoscenza. Inoltre, sarà presentata una versione prototipale dell'ontologia alla base del VirLaBok.

Abstract

The European project GI-N2K -- Geographic Information Need to Know - started on October 2013. After an initial integrated analysis of the demand for and supply of geospatial education and training, the revision of the BoK and the design of the Virtual Laboratory for the BoK (VirLaBok) are currently under investigation. In particular, the Consortium is now involved into the recognition of a proper revision strategy, in terms of content and its usability through an e-platform. The goal of this paper is to discuss preliminary results of this phase and illustrate problems due to a possible overlapping among different knowledge areas. Finally, a prototyping version of the ontology underlying the VirLaBok is presented.

Introduzione

In un recente rapporto delle Nazioni Unite, lo *United Nations Committee of Experts on Global Geospatial Information Management - UNCE-GGIM* (<http://ggim.un.org>) ha affermato che l'informazione geografica (IG) giocherà un ruolo chiave per i governi e i decisori nella realizzazione di molte attività importanti, come ad esempio la misurazione e la valutazione del livello di sviluppo sostenibile raggiunto dai singoli Stati. Ovviamente, per poter concretizzare questa potenzialità, è importante mettere a fattor comune il patrimonio informativo costruito fino ad ora attraverso la condivisione di una conoscenza strutturata. A tale scopo risulta necessario realizzare all'interno di una infrastruttura *general-purpose* di dati spaziali una base di conoscenza amichevole per le informazioni geospaziali costruita a partire dalle basi di conoscenza esistenti create dalle organizzazioni professionali internazionali. Questa asserzione è ampiamente condivisa anche da

esperti di ambiti diversi a dimostrazione del fatto che ormai l'IG ha superato i confini che tipicamente la relegavano al territorio in quanto tale, ed ha assunto un ruolo più pervasivo. Anche in Italia, per agevolare l'accesso ai dati di interesse generale, disponibili presso le pubbliche amministrazioni a livello nazionale, regionale e locale, è stato istituito il Repertorio Nazionale dei dati territoriali (<http://www.rndt.gov.it>), dove "Per dato territoriale si intende qualunque informazione geograficamente localizzata", art. 59 del Codice dell'Amministrazione Digitale (<http://www.agid.gov.it>).

Sulla base di questi presupposti, appare allora evidente come sia importante progettare profili professionali sulla base dell'analisi della domanda attuale con riferimento alle conoscenze richieste e a quelle esistenti in ambito IG, in modo da poter puntare ad una facile integrazione degli esperti del settore nel mercato del lavoro.

Per affrontare inizialmente questa sfida, è stato preso come riferimento il GI S&T BoK (GI Science and Technology Body of Knowledge) del 2006 realizzato per lo University Consortium GIS degli USA (DiBiase et al., 2006). Il GI S&T BoK è lo strumento e il repository che contiene dati, indicazioni e informazioni su come sviluppare le conoscenze nella IG, ovviamente sia lato domanda che lato offerta. Esso ha l'obiettivo di dare maggiore coerenza ed efficacia alla formazione accademica offerta dalle università americane e rivolta a soddisfare le richieste degli utenti (enti pubblici, imprese e ONG). Questa versione del BoK è stata in grado di combinare la qualità dell'offerta accademica con le richieste del mercato del lavoro per lungo tempo. Tuttavia, nel 2014 tale documento è apparso obsoleto, in considerazione delle più recenti scoperte concettuali e tecnologiche in materia di IG provenienti sia dal mondo accademico che da quello professionale. Quindi è risultato fondamentale riformulare il BoK nel settore della GIS&T affinché la formazione accademica offerta dalle università potesse soddisfare le nuove richieste degli utenti.

È stato quindi proposto e finanziato il progetto europeo GI-N2K, "Geographic Information – Need to Know" (<http://www.gi-n2k.eu>). Questo progetto è coordinato dalla Katholieke Universiteit di Leuven e vede un consorzio internazionale di 31 partner di 25 paesi diversi (AMFM GIS Italia è un partner del consorzio). È finanziato dall'Erasmus Lifelong Learning Program ed ha come requisito principale l'esigenza di allineare i curriculum accademici in GI S&T con le esigenze del corrispondente mercato del lavoro ed imprenditoriali dell'Unione Europea.

L'obiettivo di questo lavoro è analizzare i risultati fino ad ora ottenuti e valutare la completezza degli obiettivi formativi nella definizione di profili per futuri esperti del settore che debbano operare a sostegno dello sviluppo locale e della crescita economica.

Geographic Information: Need To Know (GI-N2K)

La proposta del progetto GI-N2K nasce dalla consapevolezza che la richiesta di professionisti nel campo della IG è alta ma spesso disattesa. Le ragioni che potrebbero spiegare questa situazione sono molte e molte sono anche le soluzioni che sono state proposte per colmare questa mancanza. In particolare, è stato evidenziato che una delle cause maggiori è un significativo disallineamento tra la formazione nel campo della GI S&T e le attuali richieste del mercato del lavoro. Di fatto, molti corsi post-laurea sul tema GIS sono indirizzati soprattutto su argomenti di geografia, così come molti corsi di informatica non forniscono la necessaria conoscenza geografica.

Per affrontare questo aspetto rilevante, la soluzione proposta è la progettazione e la realizzazione di un sistema di supporto alla formazione di figure con competenze geospaziali, che tenga conto della domanda e dell'offerta in questo campo, e che inoltre, a causa della pervasività della IG, sia focalizzato su specifiche aree di conoscenza affinché si padroneggi in maniera proficua la disciplina sottostante.

Per raggiungere questo obiettivo, si è deciso di adottare la metodologia del *Body of Knowledge*, un *corpus delle conoscenze* che copre e descrive tutte le aree di conoscenza e di competenza che si devono gestire in un qualsiasi ambito scientifico / tecnologico. Una prima versione del GI S&T BoK è stata sviluppata nel 2006, presso lo University Consortium for Geographic Information Science (UCGIS). Il GI S&T BoK ha una organizzazione gerarchica ed include 10 Aree di

Conoscenza (KA), 73 unità (di cui 26 definite come unità "core"), 329 argomenti, e più di 1600 obiettivi educativi. Come esempio, la Figura 1 mostra la KA *Geospatial Data* (GD) insieme alle sue dodici unità (GD1 ÷ GD12) e agli argomenti associati.

Geospatial Data	
GD1 Earth geometry 1-1 History of understanding Earth's shape 1-2 Geoids 1-3 Spheroids and ellipsoids	GD8 Digitizing 8-1 Tablet digitizing 8-2 On-screen digitizing 8-3 Scanning and automated vectorization
GD2 Land partitioning systems 2-1 Unsystematic methods 2-2 Systematic methods	GD9 Field data collection 9-1 Sample size selection 9-2 Spatial sample types 9-3 Sample intervals 9-4 Field data technologies
GD3 Georeferencing systems 3-1 Geographic coordinate system 3-2 Plane coordinate systems 3-3 Tessellated referencing systems 3-4 Linear referencing systems	GD10 Aerial imaging and photogrammetry 10-1 Nature of aerial image data 10-2 Platforms and sensors 10-3 Aerial image interpretation 10-4 Stereoscopy and orthoimagery 10-5 Vector data extraction 10-6 Mission planning
GD4 Datums 4-1 Horizontal datums 4-2 Vertical datums	GD11 Satellite and shipboard remote sensing 11-1 Nature of multispectral image data 11-2 Platforms and sensors 11-3 Algorithms and processing 11-4 Ground verification and accuracy assessment 11-5 Applications and settings
GD5 Map projections 5-1 Map projection properties 5-2 Map projection classes 5-3 Map projection parameters 5-4 Georegistration	GD12 Metadata, standards, and infrastructures 12-1 Metadata 12-2 Content standards 12-3 Data warehouses 12-4 Exchange specifications 12-5 Transport protocols 12-6 Spatial Data Infrastructures
GD6 Data quality 6-1 Geometric accuracy 6-2 Thematic accuracy 6-3 Resolution 6-4 Precision 6-5 Primary and secondary sources	
GD7 Land surveying and GPS 7-1 Survey theory and electro-optical methods 7-2 Land records 7-3 Global Positioning System	

Figura 1 - Area di Conoscenza *Geospatial Data* e le sue unità.

Successivamente, ogni KA è introdotta brevemente e ogni unità è dettagliata in termini di argomenti e obiettivi formativi. Come esempio, la Tabella 1 riporta la descrizione iniziale di GD, mentre gli obiettivi formativi del singolo argomento GD3-1 (*Geographic coordinate system*) sono illustrati nella Tabella 2.

Tabella 1 - La KA *Geospatial Data* (in inglese).

Knowledge Area: Geospatial Data (GD)
Geospatial data represent measurements of the locations and attributes of phenomena at or near Earth's surface. Information is data made meaningful in the context of a question or problem. Information is rendered from data by analytical methods. Information quality and value depends to a large extent on the quality and currency of data (though historical data are valuable for many applications). Geospatial data may have spatial, temporal, and attribute (descriptive) components, as well as associated metadata. Data may be acquired from primary or secondary data sources. Examples of primary data sources include surveying, remote sensing (including aerial and satellite imaging), the global positioning system (GPS), work logs (e.g., police traffic crash reports), environmental monitoring stations, and field surveys. Secondary geospatial or geospatial-temporal data can be acquired by digitizing and scanning analog maps, as well as from other sources, such as governmental agencies.
[...]

Tabella 2 - La GD3 *Georeferencing systems* e il topic GD3-1 (in inglese).

<i>GD3 - Georeferencing systems (core unit)</i>
GD3-1 (Geographic coordinate system)
Distinguish between various latitude definitions (e.g., geocentric, geodetic, astronomic latitudes)
Explain the angular measurements represented by latitude and longitude coordinates
Locate on a globe the positions represented by latitude and longitude coordinates
Write an algorithm that converts geographic coordinates from decimal degrees (DD) to degrees, minutes, seconds (DMS) format

Calculate the latitude and longitude coordinates of a given location on the map using the coordinate grid ticks in the collar of a topographic map and the appropriate interpolation formula
Mathematically express the relationship between Cartesian coordinates and polar coordinates
Calculate the uncertainty of a ground position defined by latitude and longitude coordinates specified in decimal degrees to a given number of decimal places
Use GIS software and base data encoded as geographic coordinates to geocode a list of address-referenced locations

Dal 2006, il BoK è stato utilizzato per diverse iniziative, ognuna delle quali ha messo in evidenza le potenzialità di questo strumento nel processo di creazione ed esplorazione di contenuti GI S&T BoK (Painho and Curvelo, 2015), (F. Rip, 2011), (Hossain and W. Reinhardt, 2012).

L'obiettivo dell'attuale progetto GI-N2K è ora costruire una versione aggiornata del GI S&T BoK sulla base della discussione in corso in Europa sulle nuove tendenze nella IG in termini di scienze e tecnologie. In particolare, lo scopo è definire un sistema per la formazione che meglio incontri le attuali esigenze del settore geospaziale, contribuendo a costruire un know-how più adeguato agli sviluppi tecnologici di questo settore.

Per raggiungere questo scopo, GI-N2K trae beneficio dai risultati del network creato per il progetto in VESTA-GIS (www.vesta-gis.eu), i cui partner stanno contribuendo a sviluppare un repository digitale dei concetti di base nel campo della IG. Inoltre, il progetto mira anche alla costruzione di una versione digitale su base ontologica dei contenuti affinché questi siano disponibili per un uso più comodo e proficuo a dimensione europea, sulle base delle ultime innovazioni tecnologiche.

In sintesi, questo obiettivo può essere raggiunto realizzando alcune specifiche attività descritte nei relativi *workpackage*:

- In WP1 sono previsti *survey* ed interviste per identificare le KA, le unità, gli argomenti ed i concetti che devono essere inclusi nella versione aggiornata del GI S&T BoK. Inoltre, è prevista un'analisi dell'attuale offerta di istruzione e formazione nell'ambito della GI S&T sulla base degli approcci e dei risultati di iniziative precedenti.
- WP2 si focalizza sull'esame e il miglioramento del GI S&T BoK esistente. Dopo la definizione della struttura del BoK, i.e. le KA che ne faranno parte, un team di esperti ha il compito di sviluppare i concetti, l'usabilità e la rilevanza, e di aggiornare il contenuto.
- WP3 riguarda la progettazione e lo sviluppo di un BoK repository online, corredato di diversi tool per l'aggiornamento e l'utilizzo. Al fine di confrontare i requisiti iniziali e le funzionalità dei toolset, vanno condotte delle interviste all'interno di un workshop. Un ulteriore workshop è da organizzare anche per mostrare e testare il repository e i toolset.
- In WP4 il nuovo BoK e i tool sviluppati verranno applicati a vari casi reali per ottenere indicazioni su come gli output di WP2 e WP3 possano essere modificati e migliorati. La selezione degli use case è basata sui risultati di un survey e di interviste. L'insieme selezionato va presentato durante una serie di workshop organizzati in diversi stati.

I risultati del processo di valutazione del BoK

Il WP1 si è concluso nel Maggio 2014. Le attività che sono state svolte hanno incluso il monitoraggio della domanda e dell'offerta di formazione e training nel settore della IG attraverso la diffusione di survey su scala europea presso università, imprese pubbliche e private, istituzioni e ONG, e l'analisi del rapporto tra domanda professionale e offerta accademica attraverso la comparazione dei risultati ottenuti per i due settori. A valle di queste attività è stato redatto un interessante rapporto (www.gi-n2k.eu/surveys-results/), in cui sono stati presentati e discussi i risultati dei questionari e delle interviste, utilizzati successivamente per effettuare la revisione del BoK.

In generale la valutazione è stata focalizzata su:

- la consapevolezza e l'uso del GIS&T BoK da parte di comunità di esperti,
- il teaching gap: il gap tra la Domanda e l'Offerta delle competenze,
- il content gap: l'incompletezza del GIS&T BoK.

Dal lato domanda, l'analisi effettuata ha evidenziato tre richieste fondamentali. Innanzi tutto occorre spostare l'attenzione dall'ambito dell'acquisizione dei dati primari alla gestione di grandi quantità di dati spaziali. E' da tenere in conto poi la mancanza di competenze nella programmazione e nello sviluppo di applicazioni. Ed infine è importante accrescere il ruolo del Web in tutte le sue forme, divulgativa, di supporto alla formazione, di training, ecc.

Per quanto riguarda invece l'importanza del contenuto del BoK rispetto alla richiesta di competenze, le diverse KA sono state valutate in modo differente. Il rating medio più alto è stato raggiunto dalla KA Geospatial Data, seguita da Cartography and Visualization e da Design Aspects. A livello più basso si trova invece la KA Geocomputation. In modo consistente anche le unità all'interno di una stessa KA sono state valutate generalmente con rating diversi. Per la KA Analytical Methods ad esempio, mentre le operazioni di base come la misurazione delle proprietà geometriche e l'esecuzione delle interrogazioni sono state ritenute "molto rilevanti", i metodi avanzati come la regressione spaziale o l'ottimizzazione matematica sono state valutate "meno rilevanti". Di conseguenza la KA Geocomputational che è interamente composta da tali unità avanzate ha ottenuto in generale il rating più basso.

Infine, altre KA, come GIS&T and Society, hanno ottenuto un numero significativo di "somehow relevant" il che sta ad indicare che il GIS&T è ancora visto soprattutto come una disciplina tecnica. Per quanto riguarda in generale le tendenze ed i bisogni, le keyword che primeggiano nel dominio GIS&T, sia in termini di tendenze che di esigenze (obiettivi formativi) o mancanze sono GIS, Data e Analysis. In generale, le competenze attese sono orientate soprattutto alle applicazioni web e mobile, alla gestione dei big- e open data, e alla programmazione. Relativamente invece ai concetti del tutto mancanti nel BoK, ci si riferisce in generale allo sviluppo di applicazioni (e.g., API, geoJSON, python, javascript), al WebGIS (e.g., HTML5, smartpone, mobile, GPRS, RESTful, semantic web), alle SDI - Spatial Data Infrastructure (e.g., INSPIRE, harmonization, geoportal), all'acquisizione dei dati (e.g., UAV, VGI, crowdsourcing), e altri argomenti come big data, realtà aumentata e standard per la modellazione 3D come CityGML.

Infine, sono state raccolte delle interessanti osservazioni da parte di esperti di IG riguardanti le competenze future attese da parte dei professionisti. In sintesi, c'è ampio consenso sulla necessità di personale qualificato in ambito sensori e applicazioni mobili, così come un aspetto rilevante è già da ora l'integrazione di masse enormi di dati e l'uso di database noSQL. Inoltre, è sempre più forte la tendenza a distinguere due diverse tipologie di ruoli professionali: lavori tecnicamente orientati alla realizzazione di servizi di IG, e lavori orientati alla progettualità che richiedono una comprensione più profonda dei concetti fondanti e delle esigenze specifiche del dominio.

Un'ontologia alla base del BoK rivisitato

Lo scorso Maggio si è tenuto a Lisbona un workshop di due giorni in cui si è lavorato alla definizione della strategia di revisione per il BoK. A valle dei lavori, i partner del consorzio hanno concordato un metodo di revisione in cui il processo utilizzato come canvas prevede di partire dalle 10 KA del BoK originale insieme ad una KA aggiuntiva che concerne gli sviluppi tecnologici recenti e del tutto assenti nella versione 2006 del BoK.

Contestualmente, sono in via di definizione 11 gruppi di esperti, provenienti sia dai partner del consorzio che da altri esperti di campi specifici. In questa prima fase, ogni gruppo lavorerà su diversi argomenti ed unità per decidere quali nel frattempo siano diventati obsoleti, quali dovranno essere modificati, ed infine quali concetti dovranno essere introdotti come nuovi argomenti o unità. Intanto è stata effettuata una prima trasposizione del contenuto del BoK in una struttura ontologica come mostrato in Figura 2. Ogni KA è gerarchicamente collegata alle corrispondenti unità e quindi agli argomenti. Alla luce dei risultati che verranno dalle prossime attività del WP2, si potrà arricchire l'ontologia con le relazioni semantiche tra i concetti utili anche a garantire l'interoperabilità tra gli stessi. In particolare, si dovrà costruire un network di concetti legati da diversi tipi di relazioni, incluse: Gerarchie, Super-concetti & sub-concetti, Similarità e Dipendenza

(pre-requisiti). Inoltre, ogni concetto avrà degli attributi (proprietà) - e.g. un nome ed una descrizione.

In conclusione, a valle del WP 2 sarà possibile avere una versione aggiornata del BoK da testare su use case reali per arrivare alla definizione di un *corpus di conoscenze* nell'ambito della IG realmente aderente alle esigenze del mercato del lavoro, in grado di supportare la definizione di profili che abbiano la giusta risposta metodologica e tecnologica alle richieste del settore.

Geospatial Data
 concept created 2006-09-01 by GIS&T BoK 2006ed.
 Should this be part of the Body of Knowledge?
 • true
 • true
 Proposed short codes:
 • GD
 Proposed names:
 • Geospatial Data
 Proposed descriptions:
 • Geospatial data represent measurements of the locations and attributes of phenomena at or near Earth's surface. Information is data made meaningful in the context of a question or problem. Information is rendered from data by analytical methods. Information quality and value depends to a large extent on the quality and currency of data (though historical data are valuable for many applications). Geospatial data may have spatial, temporal, and attribute (descriptive) components as well as associated metadata. Data may be acquired from primary or secondary data sources. Examples of primary data sources include surveying, remote sensing (including aerial and satellite imaging), the global positioning system (GPS), work logs (e.g., police traffic crash reports), environmental monitoring stations and field surveys. Secondary geospatial or geospatial-temporal data can be acquired by digitizing and scanning analog maps, as well as from other sources, such as governmental agencies. The legitimacy of geographic information data as a discrete field has been claimed in terms of the unique properties of geospatial data. In a paper in which he coined the term GIScience, Goodchild (1992) identified several such properties, including: 1. Geospatial data represent spatial locations and non-spatial attributes measured at certain times. 2. The Earth's surface is highly complex in shape and continuous in extent. 3. Geospatial data tend to be spatially autocorrelated. It has long been said that data account for the largest portion of geospatial project costs. While this maxim remains true for many projects, practitioners and their clients now can reasonably expect certain kinds of data to be freely or cheaply available via the World Wide Web. Federal, state, regional, and local government agencies, as well as commercial geospatial data producers, operate clearinghouses that provide access to geospatial data. The U.S.

Super-concepts
 Geographic Information Science and Technology

Sub-concepts
 Digitizing metadata, standards, and infrastructures
 Aerial imaging, tand surveying and photogrammetry and GIS
 Land partitioning systems
 Datasets
 Georeferencing systems

Post-requisites
Prerequisites
Similar concepts

The GI-N2K Body of Knowledge Wiki
 Welcome to the visual Wiki interface to the GI-N2K Body of Knowledge, hosted by BigKnowledge™. This site presents the BoK as a network of related concepts. You can use our graph or text interfaces to familiarize yourself with the structure of the BoK. Then, you may use it to help the GI-N2K community produce a BoK that serves its stated purposes. Please send your comments and suggestions to the [GI-N2K team](#).
 In this site, every subject is called a *concept*. With the knowledge areas, units, and topics of the 2006 GIS&T BoK as a starting point, the interface allows to propose modifications to existing concepts as well as proposed the creation of new concepts or the deletion of outdated concepts (via asserting its non-existence). The same can be done to relationships between concepts. What was referred to as "learning objectives" in the 2006 GIS&T BoK, is here referred to as *skills* because they are practical abilities that demonstrate competence in relation to a concept. In addition to the hierarchical

Figura 2 - La struttura ontologica del BoK.

Riferimenti bibliografici

DiBiase D W, DeMers M N, Johnson A J, Kemp K K, Taylor-Luck A, Plewe B S, and Wentz E A (eds) (2006), *Geographic Information Science and Technology Body of Knowledge* (First Edition). Washington, D.C., University Consortium for Geographic Information Science and Association of American Geographers.

Hossain I. and Reinhardt W., (2012) "*Curriculum Design Based on the UC GI S&T Body of Knowledge Supported by a Software Tool* " in Proceedings of the 8th European GIS Education Seminar - EUGISES 'GIS-education: Where are the boundaries?' 6-9 September, 2012, Leuven, BE. pp. 19-26.

Painho, M., Curvelo, P. (2008) "*BoK e-Tool Prototype An ontological-based approach to the exploration of Geographic Information Science & Technology Body of Knowledge*" in online Proceedings of the 6th European GIS Education Seminar - EUGISES 2008, Cirencester, UK. pp. 1-8.

Rip, F., (2011), "*GI-education: the impact of EduMapping*" in Proceedings of the Workshop 'GIS-education in a changing academic environment - LeGIO', 18 November 2011, Leuven, BE. pp. 3-12.

Wallentin, G., Hofer B., and Traun C., (2001) "*Assessment of Workforce Demands to Shape GIS&T Education*" in Transaction in GIS A.K. Dey, "Understanding and using context," in Personal Ubiquitous Comput. 5, 1, 2001.