

## Armonizzazione dei dizionari tematici e razionalizzazione delle corrispondenti tassonomie attraverso algoritmi Neural Network

Claudia Pegoraro <sup>(a)</sup>

<sup>(a)</sup> Professionista indipendente, via Richelmy 25/A 10126 Torino, 335 1395923, [claudia\\_pegoraro@yahoo.it](mailto:claudia_pegoraro@yahoo.it)

### Riassunto

Di dizionari tematici (si veda l'esempio classico della Copertura e dell'Uso del Suolo) ne esistono molteplici in qualsiasi contesto, numerosi dei quali accreditati come standard lessicali a livello internazionale o nazionale. Armonizzare questi vocabolari non è semplice, per via della necessità di consolidate competenze tematiche oltre che di altrettanto consolidate competenze tecniche su come gestire l'interoperabilità semantica. Inoltre le voci tematiche sono spesso organizzate in corpose strutture di classificazione multi-livello che pongono termini e relativi significati in dipendenza gerarchica tra loro, rendendo laborioso orientarsi in ottica comparativa ed arrivare all'individuazione di voci equivalenti.

Si rivela in tali casi vincente una stretta collaborazione tra le due suddette competenze professionali che ne amalgama le visioni, fermo restando che ad una compete l'esposizione dei requisiti, all'altra la struttura organizzativa ottimale dell'informazione. Le *Neural Network* sono uno strumento perfetto per favorire simili processi di interazione, specie quando le comunità di esperti tematici sono, oltre che particolarmente vaste, legate a differenti scuole di pensiero fortemente consolidate e tra loro antagoniste, come ad esempio occorre per i vari sistemi di classificazione internazionali per Copertura ed Uso del Suolo: questi ultimi tendono ad essere considerati di qualità equivalente o quantomeno complementare, ma restano di fatto antagonisti e mai sono stati formalmente armonizzati in un unico standard universale.

Una sperimentazione su tale tematica è partita recentemente, con l'obiettivo preciso di operare una traslitterazione onnicomprensiva delle voci dei dizionari esistenti, apparentando terminologie dai significati omogenei ed approdando infine ad una tassonomia possibilmente unica, e, ove non possibile, ad un sistema di classificazione comparativo e ragionato, ove siano evidenziate le diramazioni tassonomiche in maniera chiara ed intuitiva, accompagnate da un sistema di raffronto ed equivalenza lessicale. In un tale contesto sperimentale gli algoritmi *Neural Network* intervengono nella fase "intermedia" di razionalizzazione della tassonomia, celati da un "cruscotto di lavoro" da mettere a disposizione degli esperti tematici quale collettore semi-automatico dei requisiti, in quanto comparatore dei significati e suggeritore di ritrovate corrispondenze semantiche (processo che risulterebbe altrimenti difficilmente controllabile). Il metodo ha il pregio di essere incrementale e per sua natura iterativo, quindi risulta riusabile in qualsiasi circostanza di ripensamento, affinamento o ottimizzazione dei requisiti, dunque applicabile in qualsiasi fase di avanzamento della progettazione di una struttura classificatoria o tassonomica. È inoltre replicabile, previo adattamento del contenuto informativo, in qualsiasi contesto tematico in cui la pluralità di strutture semantiche, codificate ed accreditate come standard tra gli esperti di settore, renda complesso il raffronto critico e la conciliazione dei punti di vista degli addetti ai lavori.

### Abstract

There are lots of thematic dictionaries (as for Land Cover and Land Use) in place today for one specific theme and several of them are recognized as theme standards at a national or international

level. Harmonization of these vocabularies is not an easy task, since strong thematic knowledge should be integrated with real technical experience on semantic interoperability. Moreover terms and their meanings are often included in big and complex taxonomies with many hierarchical levels and equivalent terms are difficult to be compared and found out, since coherence is not immediately clear among different taxonomies.

A successful strategy to overcome these difficulties is the collaboration between two professional families, thematic experts and IT analysts, where a good synergy is created during collection of requirements and technical formalization for data management. Neural Network are an additional perfect tool to support such a professional interaction, almost because expert communities are so wide and devoted to specific, different, not harmonized standards, as well known for the Land Cover and Land Use themes.

An applied research session is now ongoing with the goal to reach a coherent and universal taxonomy or almost a comparative tree-structure for Land Use and Land Cover classification systems, where dictionary terms have to be compared and equivalent ones to be detected in an easy way and in natural language. Neural Network are artificial intelligence tools employed to guide thematic experts to explore equivalent items and suggest potential meaning and term match, this way allowing IT technicians to discover and collect substantial requirements. This is an incremental and iterative strategy; it is able to be re-used, repeated, just for a first requirement collection phase or for any following optimization phase. It is also an adaptable method to any thematic context and any content where multiple dictionaries, taxonomies and classification systems should be arranged and harmonized in a coherent vision.

### **Riferimenti bibliografici**

- Fisher P., Comber A., Wadsworth R. (2006), "Fundamentals of Spatial Data Quality", Cap. 3 "Approaches to Uncertainty in Spatial Data", *Rodolphe Devillers & Robert Jeansoulir Editions*
- Fisher P., Comber A., Wadsworth R. (2001), "What is land cover?", *Environment and Planning B*
- Fisher, P., Comber A., Wadsworth R. (2001), "Re-Presenting GIS", Cap. "Land use and Land cover: Contradiction or Complement. ", *Peter Fisher and David Unwin Editions*
- Jha G. K. (2007), "Artificial Neural Networks and its applications", *Prentice Hall*