

26/10/2017

Llinatge geoespacial en entorns distribuïts



En l'àmbit de les Tecnologies d'Informació Geoespacial (TIG) es treballa amb grans quantitats de dades provinents de diferents orígens i que es processen en entorns distribuïts, al núvol. Investigadors de la UAB i del CREA F proposen nous mecanismes per tal de descriure i retenir, en aquest tipus d'entorns, el llinatge de les dades, és a dir, les metadades que descriuen les seves fonts i els processos que donen lloc a cada capa d'informació.

La irrupció de la web, va suposar una revolució en l'accés i l'intercanvi de dades i serveis en la majoria de camps científics. En l'àmbit de les Tecnologies d'Informació Geoespacial (TIG), aquest fet ha comportat la possibilitat d'usar i treballar amb dades i eines de geoprocessament en entorns distribuïts ubicats al núvol. Aquest canvi de paradigma ha contribuït a augmentar considerablement els estudis que combinen dades provinents de diferents disciplines i orígens, possibilitant la generació de nous productes o serveis cartogràfics, p. e. descriure un aspecte dels possibles impactes del canvi climàtic en la societat. Malgrat els beneficis, també s'ha incrementat la dificultat d'ús i d'interpretació de les dades degut a la manca d'informació auxiliar sobre el context de generació de les pròpies dades (el que anomenem metadades).

Així doncs, necessitem un mitjà per transportar les metadades i disposar informació sobre els autors, l'escala, les unitats, el sistema de projecció, el format i la qualitat de

dades i/o serveis geoespaciales. A més, en un context científic caracteritzat pel rigor, determinar la qualitat i la veracitat de les dades és capital. En aquest sentit, disposar d'informació de llinatge a les metadades geoespaciales pot ajudar a discernir la idoneïtat d'ús de determinats productes geogràfics en estudis científics concrets. El llinatge és definit com aquella part de les metadades que descriu les fonts i els processos utilitzats en l'elaboració d'un producte geoespacial. Tot i la reconeguda importància i rellevància del llinatge geoespacial, la seva presència a les metadades geoespaciales és en general encara escassa, i en molts casos fragmentada, en part degut a la falta d'eines per retenir aquesta informació.

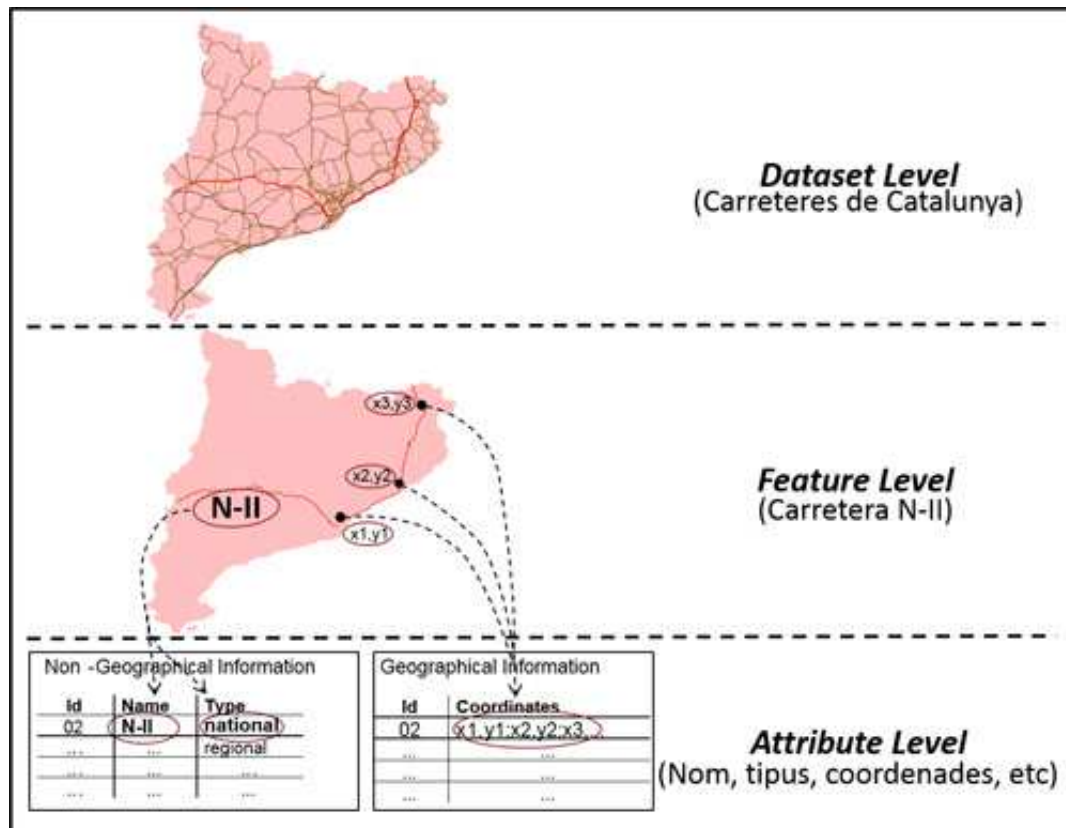


Figura 1 – Esquema conceptual representant els tres nivells de granularitat de la informació geoespacial.

L'objectiu del treball presentat és el de proposar nous mecanismes per descriure i retenir el llinatge geoespacial en un paradigma distribuït. Per fer-ho, s'ha usat el model proposat per un dels organismes que vetllen pel bon funcionament a la web: el W3C, però s'ha adaptat perquè sigui aplicable a les particularitats de la informació geoespacial. L'estructura modular del W3C PROV i la capacitat de definició de les relacions entre diferents elements, permet descriure el llinatge geoespacial als diferents nivells de granularitat de la informació geoespacial (Figura 1). Així doncs, podem descriure les fonts de dades i els processos que han donat lloc a una capa d'informació de manera general (p. e. totes les carreteres de Catalunya), però també, element a element (i així saber l'origen i el dia d'actualització de la descripció de la carretera N-II). Aquesta descripció pot ser exportada per poder ser usada en la Infraestructura de Dades Espaciales (IDE) Europea (INSPIRE) però també es pot integrar en el model semàntic global de les dades enllaçades de la web (linked data), incrementant la interoperabilitat i interconnexió entre diferents models científics.

Guillem Closa

Guillem.closa@uab.cat

Grumets Research Group, Dep de Geografia
Universitat Autònoma de Barcelona

Joan Masó

Joan.maso@uab.cat

Grumets Research Group, CREAM
Universitat Autònoma de Barcelona

Referències

[View low-bandwidth version](#)