

Integrazione di dati spaziali ufficiali e da OpenStreetMap a sostegno della Strategia europea per i dati

Alessandro Sarretta & Marco Minghini

FOSS4G-IT 2021 - 21 settembre 2021





Strategia europea per i dati

- Lanciata a febbraio 2020 nell'ambito della priorità A Europe Fit for the Digital Age.
- Mira alla creazione di un mercato unico europeo per i dati.
- Sottolinea i problemi da risolvere:
 - disponibilità dei dati (licenze),
 condivisione, interoperabilità, qualità,
 governance, infrastrutture & tecnologie,
 competenze, sicurezza informatica



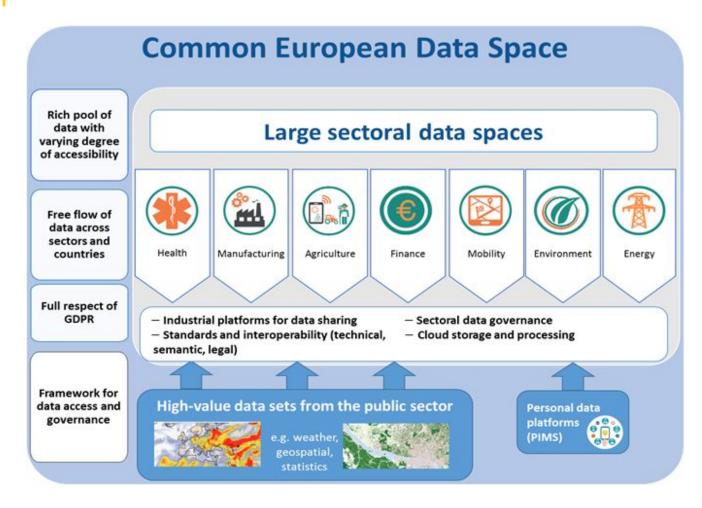
Brussels, 19.2.2020 COM(2020) 66 final

COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS

A European strategy for data



Strategia europea per i dati



- Immagina la creazione di:
 - uno spazio di dati comune europeo
 - spazi di dati settoriali
- che giovino all'economia e alla società
- attraverso la combinazione di dati tra attori e settori
 - tra cui i dati generati dai cittadini!

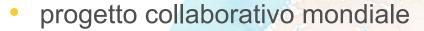


Obiettivo dello studio

- Zoom sul dominio dei dati spaziali e su OpenStreetMap (OSM) come progetto più importante di informazione geografica generata da cittadini.
- Scopo: esplorare fattori abilitanti & ostacoli nell'integrare OSM con dati ufficiali dagli enti cartografici ufficiali in UE, con un focus su:
 - produzione di dataset integrati (non solo confronto/valutazione della qualità di OSM)
 - scala nazionale (non locale)
- Output: fornire raccomandazioni sull'interoperabilità tecnica, semantica, organizzativa e legale per supportare la creazione degli spazi di dati.
- Primo esperimento: integrazione tra il dataset nazionale ufficiale degli indirizzi finlandesi (prodotto dalla National Land Survey of Finland) e il dataset degli indirizzi in OSM.

Fonti dei dati

OpenStreetMap (OSM)



dati aperti (licenza ODbL)

 modello dati semplice e piatto, basato su tag

 download dell'intero database (Planet)



National Land Survey of Finland (NLS)

dataset nazionale ufficiale

dati aperti (licenza CC-BY 4.0)

conforme al tema Indirizzi di INSPIRE

OGC API - Features





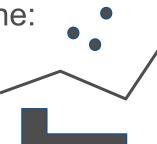


Modello dati: OSM



3 primitive geometriche:

- nodi
- way
- relazioni



Attributi:

- tag
 - <chiave> = <valore>
 - <chiave2> = <valore>
 - •

Indirizzi

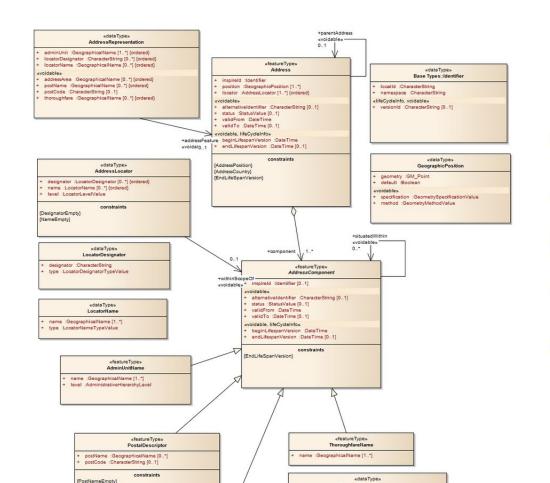
OSM tag	Description
addr:country	country code of the address
addr:city	name of the city of the address
addr:street	name of the street of the address
addr:housenumber	building number of the address



Modello dati: NLS



Basato sul modello dati UML per il tema INSPIRE Indirizzi



NLS attribute	Description	
component_ThoroughfareName_name fin	name of the street of the address in Finnish	
component_ThoroughfareName_name swe	name of the street of the address in Swedish	
component_ThoroughfareName_name sme	name of the street of the address in Sami	
locator_designator_addressNumber	building number of the address	
component_AdminUnitName_4	code of the city of the address	
component_AdminUnitName_1	country name of the address	



Integrazione dei modelli dati

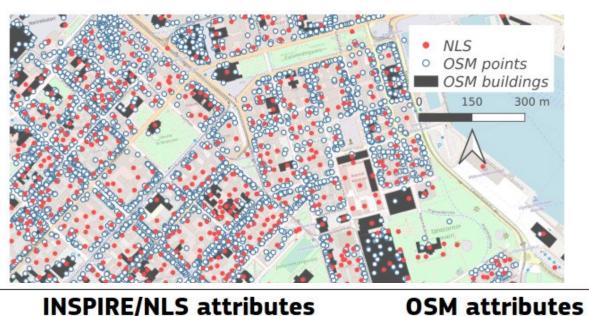


NLS

NLS attribute	Description	
component_ThoroughfareName_name fin	name of the street of the address in Finnish	
component_ThoroughfareName_name swe	name of the street of the address in Swedish	
component_ThoroughfareName_name sme	name of the street of the address in Sami	
locator_designator_addressNumber	building number of the address	
component_AdminUnitName_4	code of the city of the address	
component_AdminUnitName_1	country name of the address	

OSM

OSM tag	Description
addr:country	country code of the address
addr:city	name of the city of the address
addr:street	name of the street of the address
addr:housenumber	building number of the address



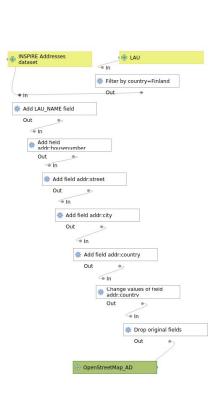
INSPIRE/NLS attributes

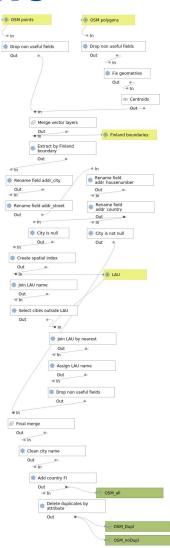
component_ThoroughfareName_name fin addr:street addr:housenumber locator_designator_addressNumber component_AdminUnitName_4 addr:city component_AdminUnitName_1 addr:country

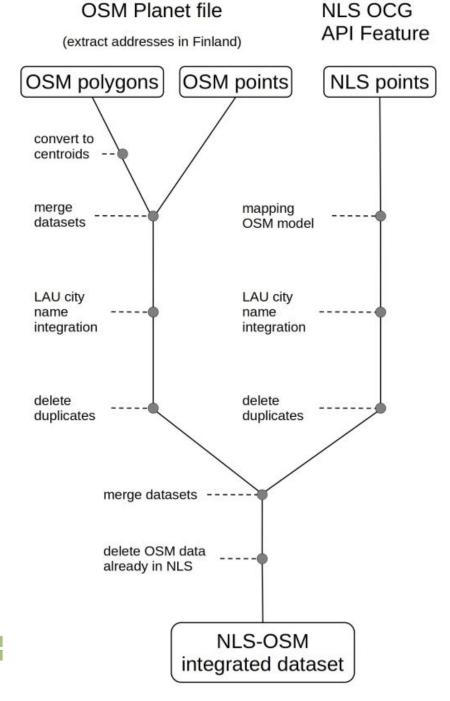


Processo di integrazione

- Workflow
 riproducibile passo
 dopo passo
- Implementazione tramite FOSS4G: QGIS Graphical Modeler
- Repository GitHub





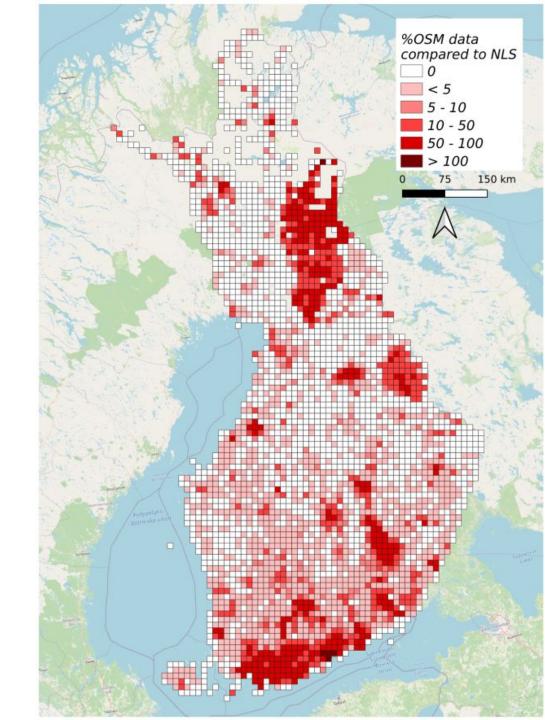


Risultati

- Nei dati originali, maggior numero di indirizzi in NLS rispetto a OSM: 3.3M vs 0.5M
- Distribuzione geografica non uniforme in OSM data: import localizzati e mappatura disomogenea
- Alta densità di dati OSM principalmente nelle aree urbane: elevata densità abitativa → numero maggiore di mappatori → più elementi mappati

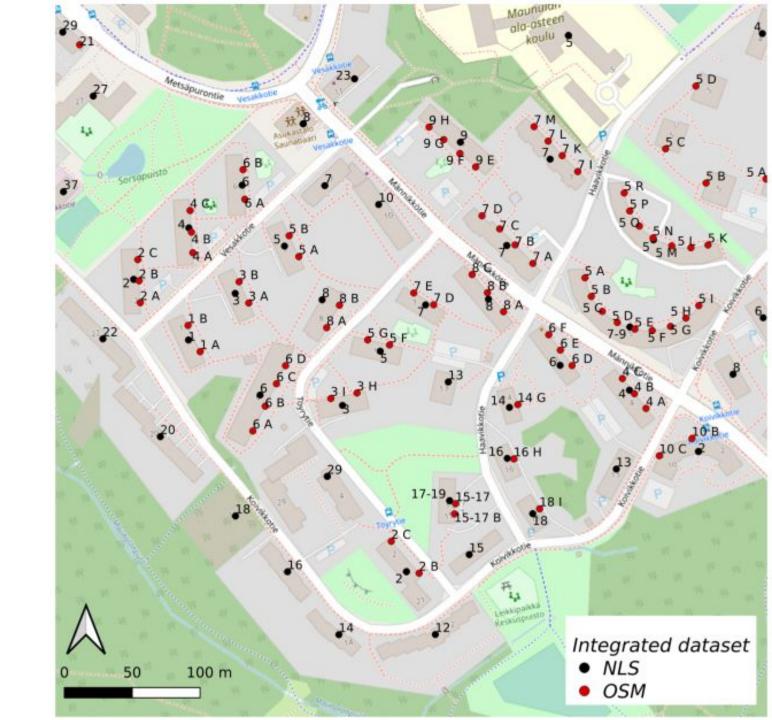
griglia di riferimento EEA 10x10km

- celle vuote: nessun indirizzo né in NLS né in OSM
- celle bianche: nessun indirizzo in OSM
- celle nere: indirizzi in OSM > indirizzi in NLS



Risultati

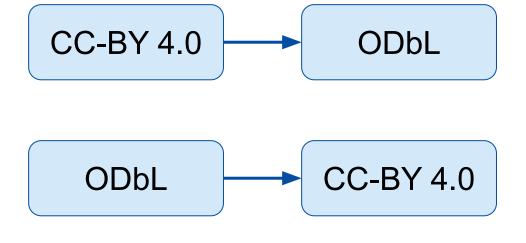
- dati integrati: 96% NLS (solo ~80.000 OSM)
- alcuni dati in OSM sono errori di scrittura di nomi di strade e città
- casi in cui OSM include dati più dettagliati o aggiornati



Incompatibilità tra licenze

La maggiorparte dei dati governativi (anche geografici) è rilasciata con licenza Creative Commons Attribuzione 4.0 International (<u>CC-BY 4.0</u>).

CC-BY 4.0 e ODbL non sono pienamente compatibili: è uno dei maggiori problemi evidenziati durante alcune interviste con autorità cartografiche nazionali e regionali.



L'importazione di dati CC-BY 4.0 in OSM richiede un permesso esplicito da parte dei licenziatari (<u>indicazioni OSMF</u>).

L'inclusione di dati OSM in altri dataset richiede il rilascio dei dati integrati con ODbL.

Conclusioni e prospettive

- Qualsiasi processo di integrazione dei dati dovrebbe essere preparato con cura.
- Da un punto di vista tecnico, gli strumenti FOSS4G permettono l'integrazione tra i dataset degli indirizzi OSM e NLS e permettono un miglioramento di entrambi i dataset.
- Specificità nazionali (oltre alla disomogeneità di OSM) possono complicare l'integrazione: l'interoperabilità di INSPIRE può facilitare l'applicazione in altri paesi.
- L'incompatibilità delle licenze è un problema rilevante per lo scambio bidirezionale di informazioni tra istituzioni e OpenStreetMap
- OSM sta diventando un ecosistema complesso (cittadini, volontari, governi, aziende private/commerciali): tale complessità deve essere affrontata e gestita.
- Primo passo per una discussione più ampia su una collaborazione permanente tra enti cartografici istituzionali e OpenStreetMap.

Articolo scientifico

The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLVI-4/W2-2021 FOSS4G 2021 – Academic Track, 27 September–2 October 2021, Buenos Aires, Argentina

TOWARDS THE INTEGRATION OF AUTHORITATIVE AND OPENSTREETMAP GEOSPATIAL DATASETS IN SUPPORT OF THE EUROPEAN STRATEGY FOR DATA

A. Sarrettaa, *, M. Minghinib

^a National Research Council, Research Institute for Geo-hydrological Protection, Padua, Italy - alessandro.sarretta@irpi.cnr.it ^b European Commission, Joint Research Centre (JRC), Ispra, Italy - marco.minghini@ec.europa.eu

Commission IV, WG IV/4

KEY WORDS: Citizen-generated data, Europe, Interoperability, National Mapping Agencies, OpenStreetMap, Spatial Data Infrastructures

ABSTRACT:

Digital transformation is at core of Europe's future and the importance of data is well highlighted by the recently published European strategy for data, which envisions the establishment of so-called European data spaces enabling seamless data flows across actors and sectors to ultimately boost the economy and generate innovation. Integrating datasets produced by multiple actors, including citizen-generated data, is a key objective of the strategy. This study focuses on OpenStreetMap (OSM), the most popular crowdsourced geographic information project, and is the first step towards an exploration of pros and cons of integrating its open-licensed data with authoritative geospatial datasets from European National Mapping Agencies. In contrast to previous work, which has only tested data integration at the local or regional level, an experiment was presented to integrate the national address dataset published by the National Land Survey (NLS) of Finland with the corresponding dataset from OSM. The process included the analysis of the two datasets, a mapping between their data models and a set of processing steps—performed using the open source QGIS software—to transform and finally combine their content. The resulting dataset confirms that, while addresses from the NLS are in general more complete across Finland, in some areas OSM addresses provide a higher detail and more up-to-date information to usefully complement the authoritative one. Whilst the analysis confirms that an integration between OSM and authoritative geospatial datasets is technically and semantically feasible, future work is needed to evaluate enablers and barriers that also exist at the legal and organisational level.



Thank you!



alessandro.sarretta@irpi.cnr.it marco.minghini@ec.europa.eu



The authors acknowledge the support of the European Commission - Joint Research Centre (JRC) through contract number CT-EX2013D133588-101, entitled "Evaluation of Novel approaches for governing (location) data and technology. Combined use of public sector and citizen-generated data".

© European Union 2020



Keep in touch



EU Science Hub: ec.europa.eu/jrc



@EU_ScienceHub



EU Science Hub – Joint Research Centre



EU Science, Research and Innovation



Eu Science Hub

