Potentials of data conflation

the example of webportals

Seminar: Time to take action!

Teachers: Melanie Eckle-Elze & Carolin Klonner

Presenter: Nikolaos Kolaxidis

University of Heidelberg

01.07.2022

CONTENTS

Data conflation

What is it?

How does it work?

What is conflation?

01

Conflate (Merriam-Webster):

1 a: to bring together, blend
b: confuse
2: to combine into a composite whole

Conflation (ESRI):

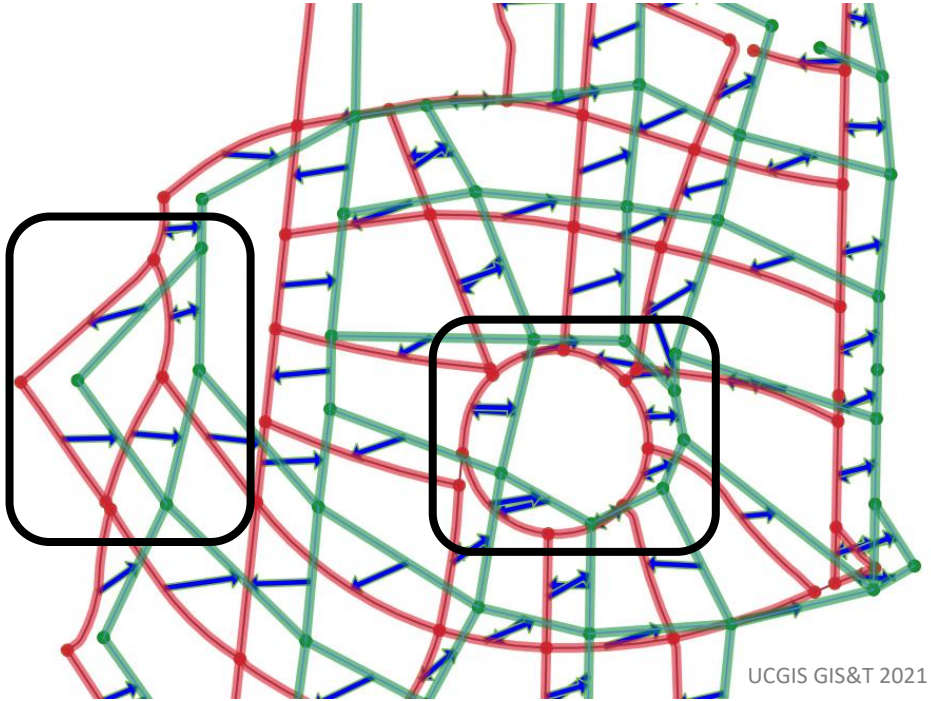
[data editing] A set of procedures that align the features of two geographic data layers and then transfers the attributes of one to the other

The goal of ***Conflation:***

"To combine the best-quality elements of [...] datasets to create a **composite dataset that is better than either of them**" (Chen/Knoblock 2008, p133ff.)

Conflation of GIS-data

01



Green: dataset A | Red: dataset B | Blue: spatial difference

Task: Estimation of next road repair works

Challenge:

- NOT a simple move & merge task
- Differently spatially transformed
- Slightly different features

Needed action:

- Assignment of same features to each other
- Supervised spatial transformation/alignment
- Merging of information

→ Enrichment of data

In Practice:

- Manually (labor-intensive, time-consuming)
- With GIS (no reliable automation yet)

Conflation toolset (ArcGIS/ArcPro)

01

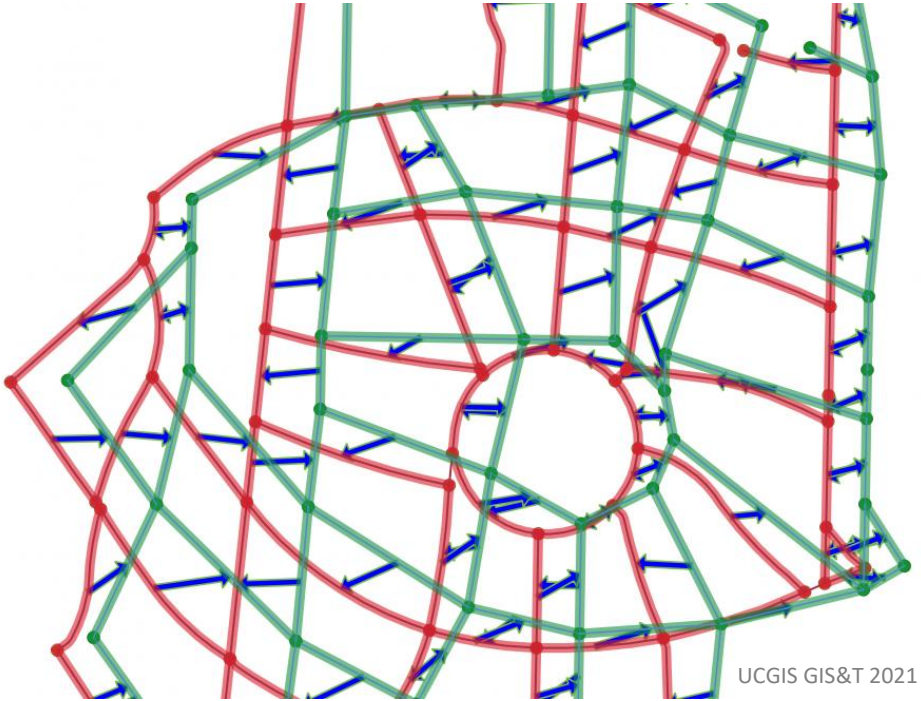
Tool	Description
Align Features	Identifies inconsistent portions of the input features against target features within a search distance and aligns them with the target features.
Calculate Transformation Errors	Calculates residue errors and root mean square error (RMSE) based on the coordinates of the input links between known control points to be used for spatial data transformation.
Edgematch Features	Modifies input line features by spatially adjusting their shapes, guided by the specified edgematch links, so they become connected with the lines in the adjacent dataset.
Generate Edgematch Links	Finds matching but disconnected line features along the edges of the source data's area and its adjacent data's area, and generates edgematch links from the source lines to the matched adjacent lines.
Generate Rubbersheet Links	Finds where the source line features spatially match the target line features and generates lines representing links from source locations to corresponding target locations for rubbersheeting.
Rubbersheet Features	Modifies input features by spatially adjusting them through rubbersheeting, using the specified rubbersheet links, so they are better aligned with the intended target features.
Split Line By Match	Splits input features based on matching relationships to obtain better corresponding line segmentation.
Transfer Attributes	Finds where the source line features spatially match the target line features and transfers specified attributes from source features to matched target features.
Transform Features	Converts the coordinates of input features from one location to another through scaling, shifting, and rotating based on the transformation links between known corresponding control points.

ESRI 2022

Potentials

Humanitarian GIS data and why conflation is needed

Conflation of GIS-data



Green: dataset A | Red: dataset B | Blue: spatial difference

Task: Estimation of next road repair works

Do I need the info?
Can I aquire it without conflation?

Potentials:

- **Fusion** of different data (sets/types/info etc.)
- **Enrichment** of spatial data
- Data is more **accessible**

→ **Cost-benefit evaluation is important**

Conflation in the humanitarian context

Different sources/types/scopes/formats of data:

VGI

- **Local knowledge and data acquisition** (Passive sensing, Participatory Sensing/Mapping, Volunteer Thinking, Environmental and Ecological Observation, Civic/Community science etc.)
- **Indigenous knowledge and data acquisition** (Focus groups, Word of mouth etc.)

- **Humanitarian Aid Organizations & Movements**
- **Official public service data**
- **Scientific data**

"**Often the best information comes from those who are closest to it**" (J. McGlade, 2008)

Conflation in the humanitarian context

Potentials:

Public is Prosumer

- Ensure **data quality** (regardless of source)
- **Building resilience** of locals to risks by:
  - providing **professionally edited** data from different sources (regardless of data quality)
  - giving the **option to participate** (any data is welcome)
  - therefore **rising awareness** of local key issues
  - commenting the data and **giving courses of action** (strengthen adaptation)
- Better GIS data for **future analyses** (giving it back to the community)

Challenges:

- Need to **aquire** all data
- **No worldwide standards** for spatial data acquisition (CRS, notations, formats etc.)
- Conflating data **needs professionals**
- **Not automated** yet (time-consuming)
- Conflated data **needs a suitable platform** to be provided back to locals (not scientific papers)
- All challenges of VGI in general

Webportals

Types, potentials, challenges

What are web portals?

Web portals:

- No unique definition
- Efficient and fast search for information
- Horizontal vs. vertical portals:
  - starting point for browsing the web
  - corporate portals
  - industry portals
  - customer portals

Geoportal/Web GIS portals:

- Central point to find and access spatial data
- GIS that uses web technologies
- Types:
  - Collection of GIS data
  - Visualization & web mapping
  - Printing maps/customization
  - Geospatial analysis
  - Professional conflation

sentinelhub Playground

esa opernicus

seasketch

HOT Tasking Manager

EO Browser

OpenStreetMap

USGS science for a changing world

EarthExplorer

OpenAerialMap

SketchMapTool

HDX

Types of GIS webportals

Viewing/Navigating/Downloading

EO Browser

sentinelhub Playground

esa opernicus

USGS science for a changing world
EarthExplorer

Viewing/Navigating/Downloading/Contributing/Communicating

OpenAerialMap

SketchMapTool

HDX

HOT Tasking Manager

seasketch

OpenStreetMap

directly in the browser

Advantages of GIS webportals

- GIS functionality regardless of location (**global reachability**)
- **Many users** concurrently
- **Interoperability** (cross-platform)
- **Flexibility** (can provide only specific tools or add some)
- **Low cost** (for users)
- **Ease of use** (no advanced GIS knowledge needed)
- **Currentness** (automatic updates)
- Humanitarian context:
  - **Open Access**
  - Conflation of **spatial data and additional information**
  - **Fast and easy acquisition** of different data from different sources
  - **Building resilience** of locals to risks with an **accessible** and **customizable** platform

Challenges of GIS webportals in the humanitarian context

- Open Access not always possible due to:
  - **Governmental restrictions**
  - **Access** to internet (fast and reliable Internet connection required)
  - **Limited devices** suitable
- **Language** barrier
- **Security** of provided data (who is responsible?)
- **Copyright infringement**
- **Communication** with government, locals and other organizations
- **Data quality** of VGI, **conflation** or editing needed beforehand

Possible solutions:

- Using the public and VGI to peer-review own VGI data
- Strict framework or manuals to collect consistent data
- Using the same basemap (OSM) and defining it as standard
- Using standards for data and tags (Interoperability)
- your ideas...

Literature

Bećirspahić, L./Karabegović, A. (2015): Web Portals for Visualizing and Searching Spatial Data. - In: 38th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO) & Distributed Computing, Visualization and Biomedical Engineering Conference, Croatia.

Chen, CC./Knoblock, C. (2008): Conflation of Geospatial Data. - In: Shekhar, S./Xiong, H. (eds): Encyclopedia of GIS. Springer, Boston, MA.

Citygate GIS (2022): Conflation. - URL: https://www.citygategis.com/conflation [as of 01.07.2022].

ESRI (2022:) An overview of the Conflation toolset. - URL: https://pro.arcgis.com/en/pro-app/latest/tool-reference/editing/an-overview-of-the-conflation-toolset.htm [as of 01.07.2022].

Gazi, T. (2020): Data to the rescue: how humanitarian aid NGOs should collect information based on the GDPR. - In: Journal of International Humanitarian Action, 5:9.

Grajales, D. P., Degrossi, L. C., Barros, D. D. R., Khan, M. R., Silva, F. L. E., Cunha, M. A., Trajber, R. de Albuquerque, J. P. (accepted): Enabling Participatory Flood Monitoring Through Cloud Services. - In: Proceedings of the 19th ISCRAM Conference, Tarbes, France, May 2022.

Ivanova, I./Him Fa, J.S./McMeekin, D.A./Arnold, L.M./Deakin, R./Wilson, M. (2020): From spatial data to spatial knowledge infrastructure. A proposed architecture. - In: Transactions in GIS, 24:p1526–1558.

McGlade, J., 2008. Environmental information and public participation. - URL: http://www.eea.europa.eu/media/speeches/environmentalinformation-and-public-participation [as of 27.06.2022].

Sharma, M. (2022): Web 1.0, Web 2.0 and Web 3.0 with their difference. - URL: https://www.geeksforgeeks.org/web-1-0-web-2-0-and-web-3-0-with-their-difference/ [as of 01.07.2022].

Tenney, C./Ziakas, J. (n.y.): Conflation of a gas utility's data. The challenge and reward. Presentation, Questar Gas Company, Utah.

Ulrichs, M./Slater, R./Costella, C. (2019): Building resilience to climate risks through social protection: from individualised models to systemic transformation. - In: Disasters, 43: p368-387.

United Nations Office for the Coordination of Humanitarian Affairs (2022): The Humanitarian Exchange Language. - URL: https://hxlstandard.org [as of 01.07.2022].

University Consortium for Geographic Information Science (2021:) Conflation. - URL: https://gistbok.ucgis.org/topic-keywords/conflation [as of 01.07.2022].

Thank you very much

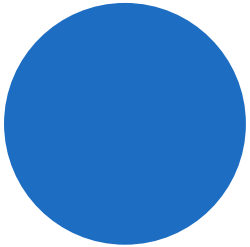
for your attention!

Seminar: Time to take action!

Presenter: Nikolaos Kolaxidis

University of Heidelberg

01.07.2022



Discussion

Generate your own take-home messages (see Recap in Moodle)

Lightning quick history of the WWW (extra)