



INSTITUT NATIONAL
DE L'INFORMATION
GÉOGRAPHIQUE
ET FORESTIÈRE



Collaborative Data Production: Research Perspectives

Guillaume Touya

EuroSDR Meeting – May 25th 2016

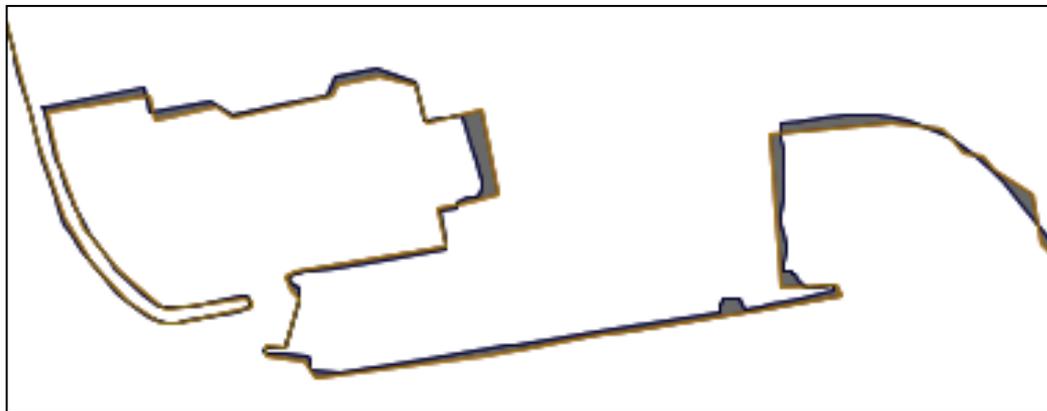
Outline of the presentation

- Quality Assurance with Crowdsourced Data
 - With reference data
 - Intrinsic assessment
- Map Making with Crowdsourced Data
 - Level of Detail heterogeneity
 - Large scale maps
 - Small scale maps
- National Mapping Agencies (NMA) and Crowdsourcing

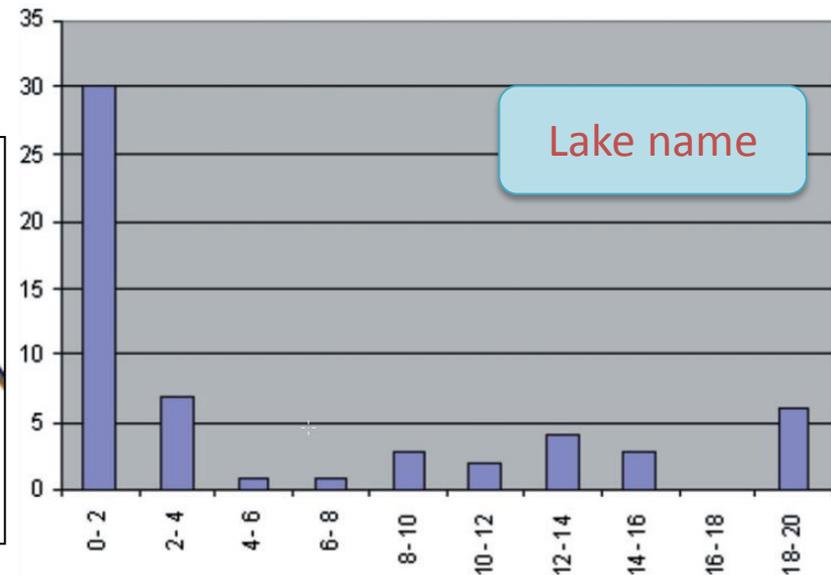
Comparison with reference data

Geometrical precision

Tag quality



Shape similarity metrics



Levenshtein distance

$$D(\text{Nantes}, \text{Nante}) = 1$$

Outline of the presentation

- Quality Assurance with Crowdsourced Data
 - With reference data
 - **Intrinsic assessment**
- Map Making with Crowdsourced Data
 - Level of Detail heterogeneity
 - Large scale maps
 - Small scale maps
- National Mapping Agencies (NMA) and Crowdsourcing

What is the best location for a Point of Interest?



Proposed methodology



Results for amenities



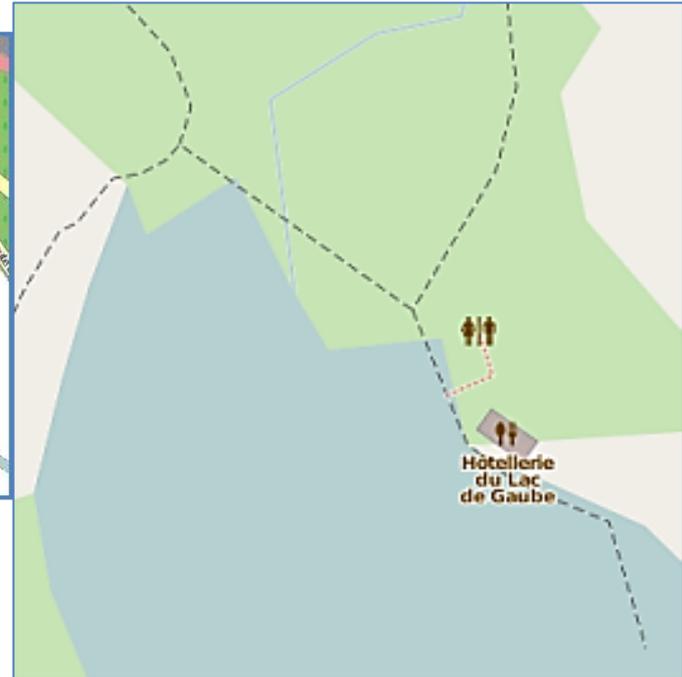
- ~6000 POIs in Paris
- 3 to 5,5% outside buildings: **bad quality**
- Great heterogeneity of locations inside buildings
- Bigger amenities -> **bigger errors and heterogeneity**

| | % of POIs outside buildings | Mean distance to centroid (DC) | Standard deviation of DC | Mean distance to walls (DW) | Standard deviation of DW |
|--------------------------|-----------------------------|--------------------------------|--------------------------|-----------------------------|--------------------------|
| Gift shops | 0 | 12,9 | 11,5 | 2,3 | 3 |
| Bars, cafes, restaurants | 3 | 10,8 | 10,5 | 2,3 | 2,1 |
| Cinemas | 5,5 | 16,9 | 18,4 | 4,9 | 4,9 |

Outline of the presentation

- Quality Assurance with Crowdsourced Data
 - With reference data
 - Intrinsic assessment
- Map Making with Crowdsourced Data
 - **Level of Detail heterogeneity**
 - Large scale maps
 - Small scale maps
- National Mapping Agencies (NMA) and Crowdsourcing

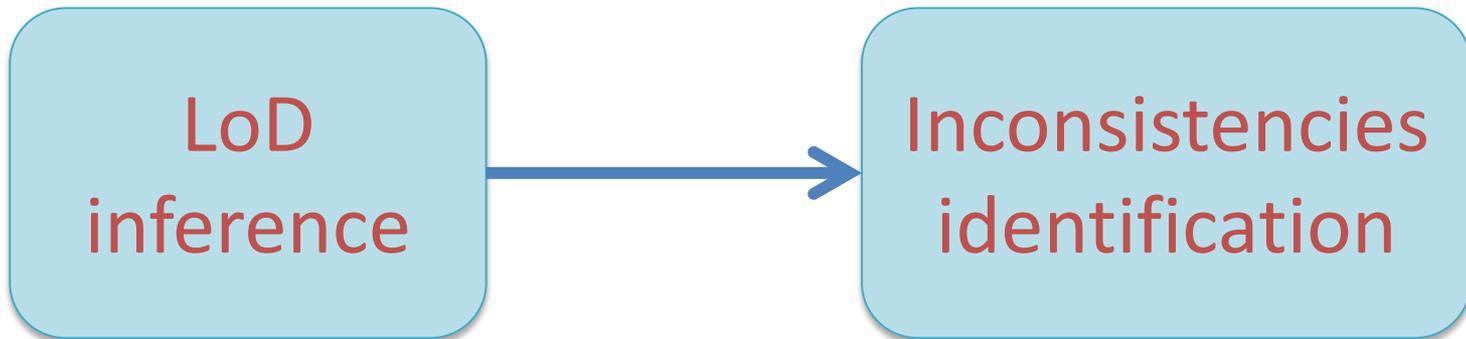
LoD heterogeneity



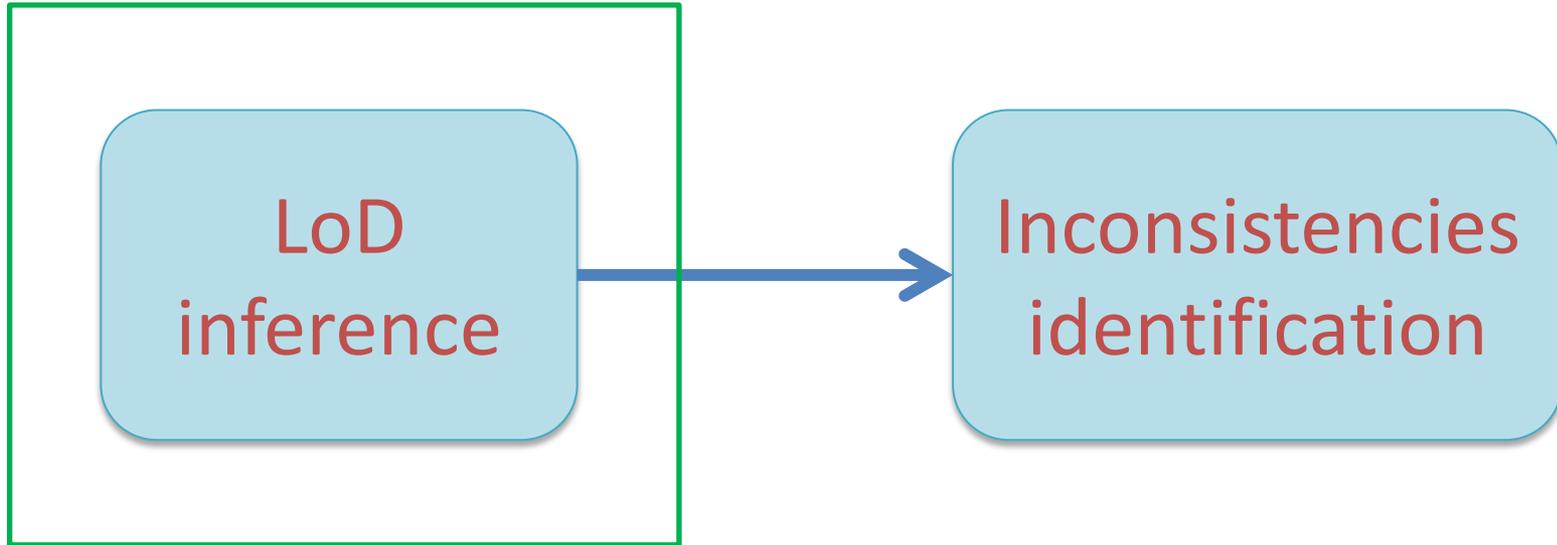
Causes:

- ⦿ Source heterogeneity (GPS, images Bing ...)
- ⦿ Contributor skills heterogeneity
- ⦿ No LoD specification in OSM

LoD heterogeneity



LoD heterogeneity



Automatic LoD inference

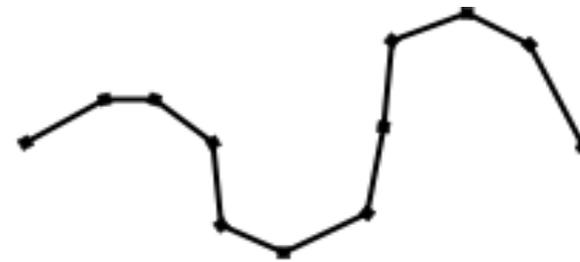
Multiple criteria
decision

- Feature type
- Vertex density
- Median length
- Capture technique
- Smallest segment
- Size
- Coalescence

$$\text{vertex density} = \frac{\text{vertex number}}{\text{object length}}$$

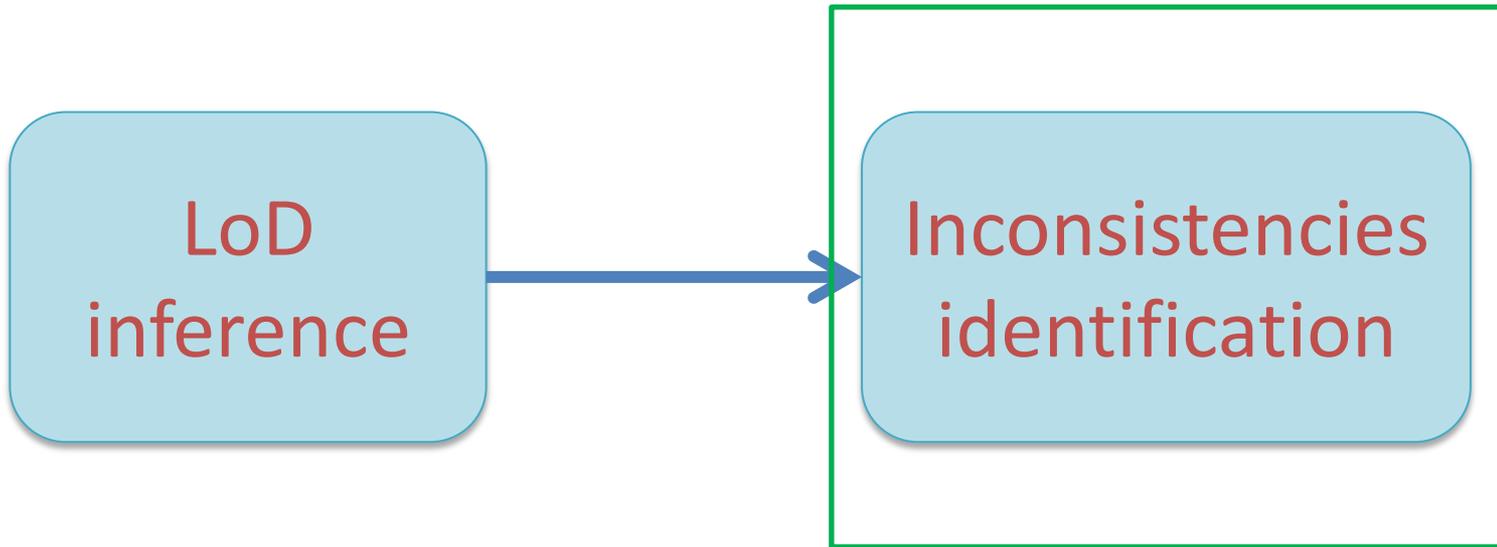


$$\text{vertex density} = \frac{23}{100} = 0.23$$

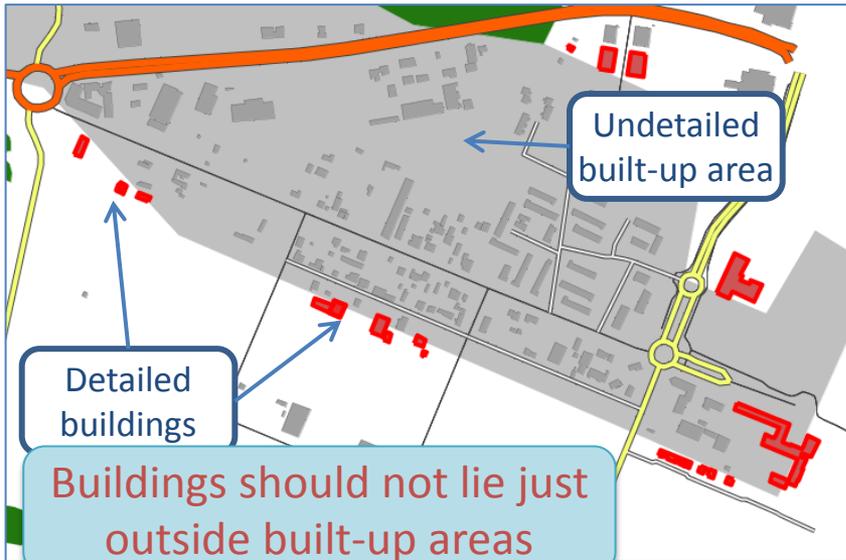


$$\text{vertex density} = \frac{12}{100} = 0.12$$

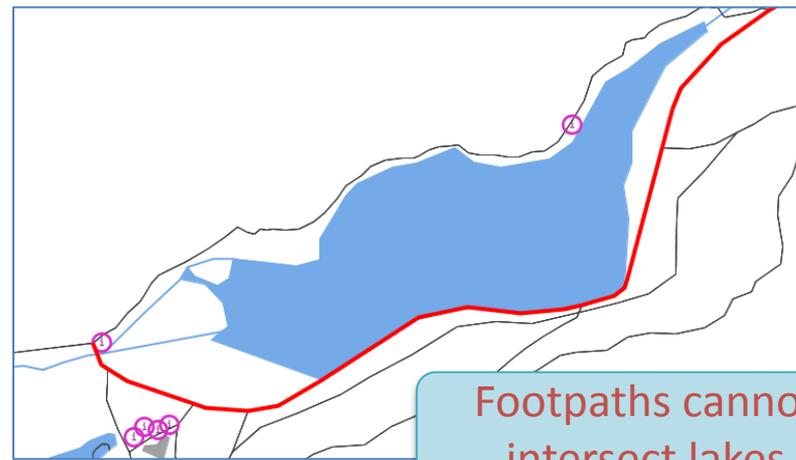
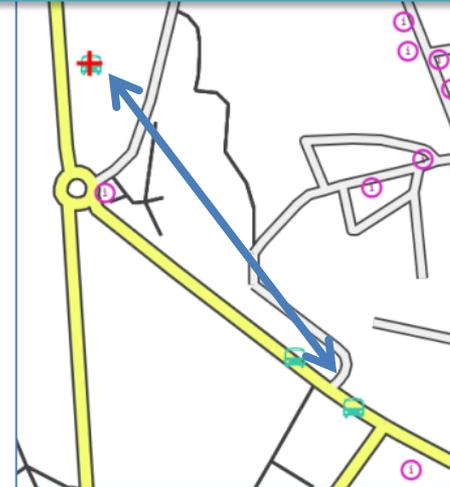
LoD heterogeneity



Identify LoD inconsistencies



Bus stops should be along roads



Outline of the presentation

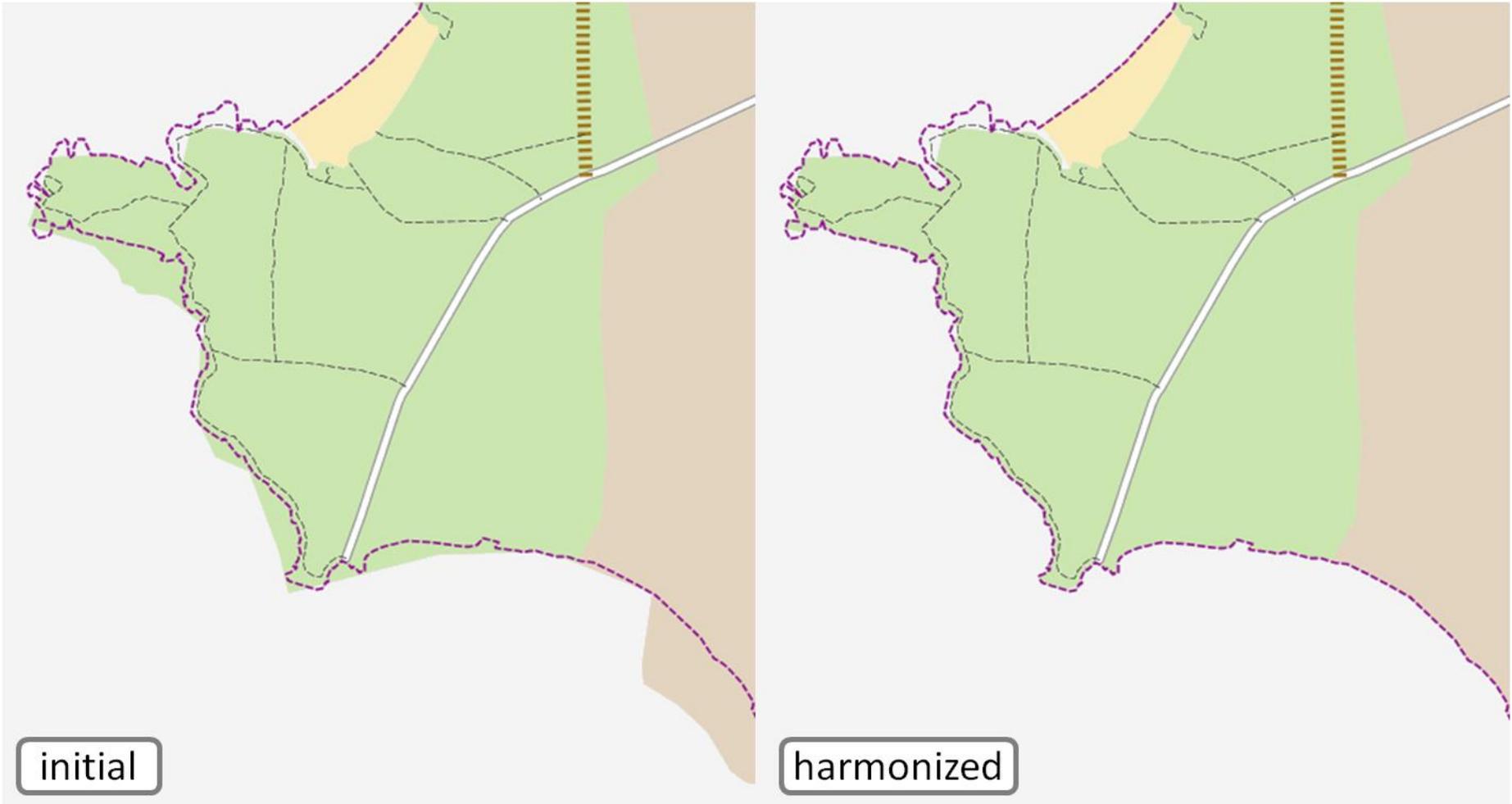
- Quality Assurance with Crowdsourced Data
 - With reference data
 - Intrinsic assessment
- Map Making with Crowdsourced Data
 - Level of Detail heterogeneity
 - **Large scale maps**
 - Small scale maps
- National Mapping Agencies (NMA) and Crowdsourcing

Large scale maps

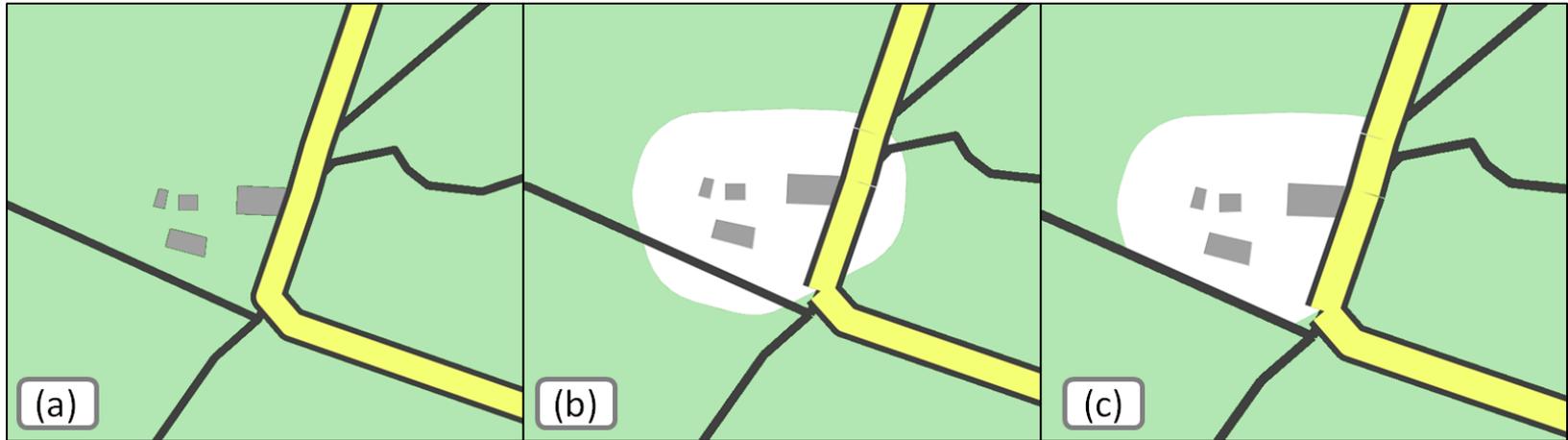


How can we represent undetailed features?

LoD harmonization



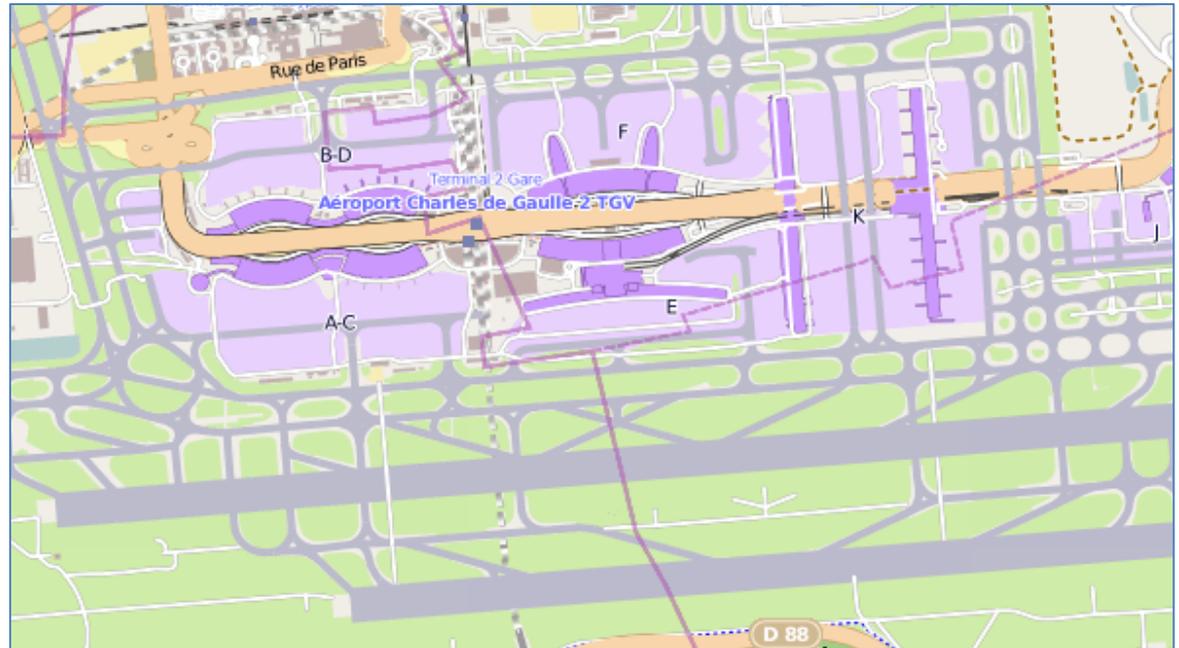
LoD harmonization



Outline of the presentation

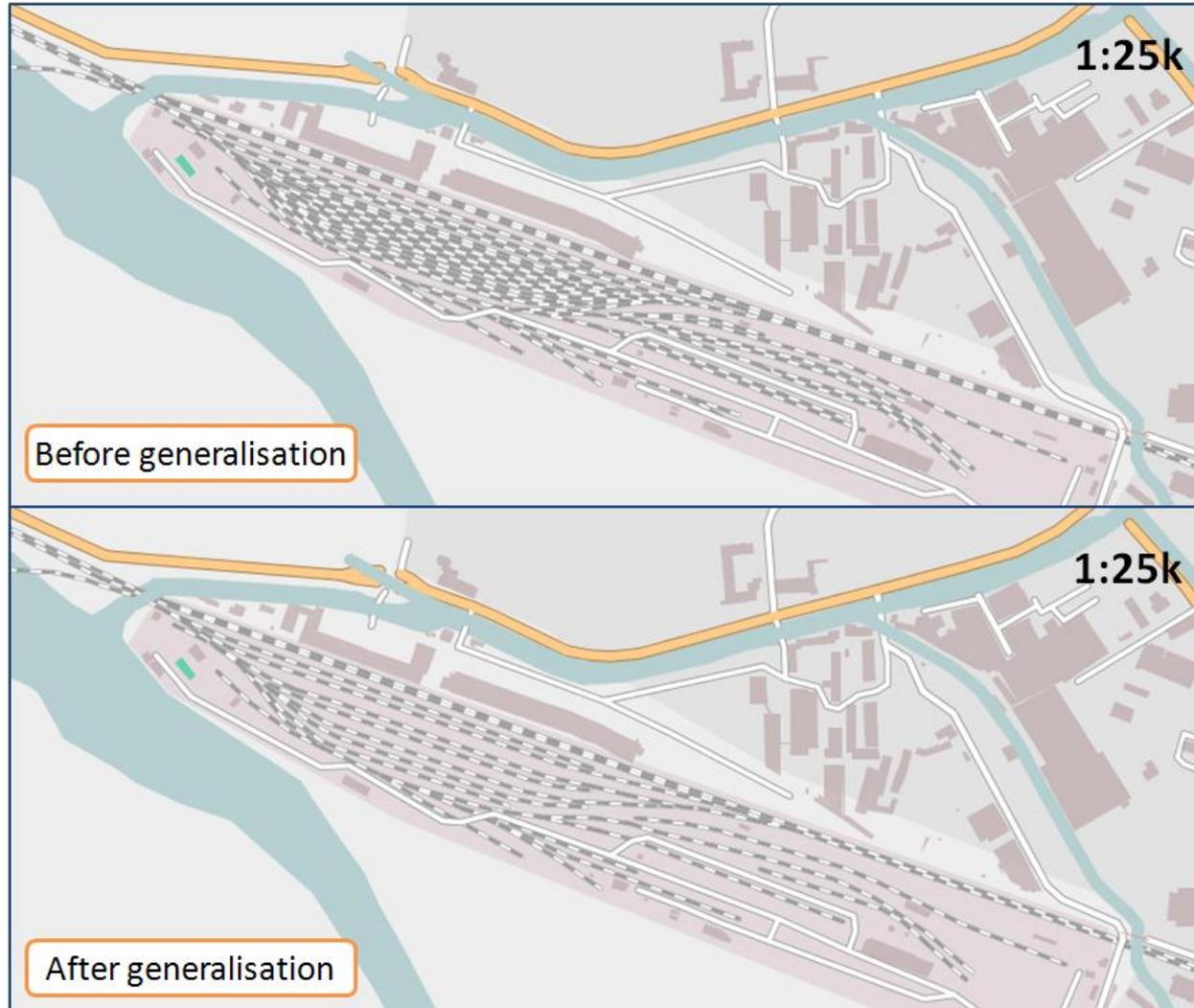
- Quality Assurance with Crowdsourced Data
 - With reference data
 - Intrinsic assessment
- Map Making with Crowdsourced Data
 - Level of Detail heterogeneity
 - Large scale maps
 - **Small scale maps**
- National Mapping Agencies (NMA) and Crowdsourcing

Small scale maps: generalization



*How can we generalize **very** detailed features?*

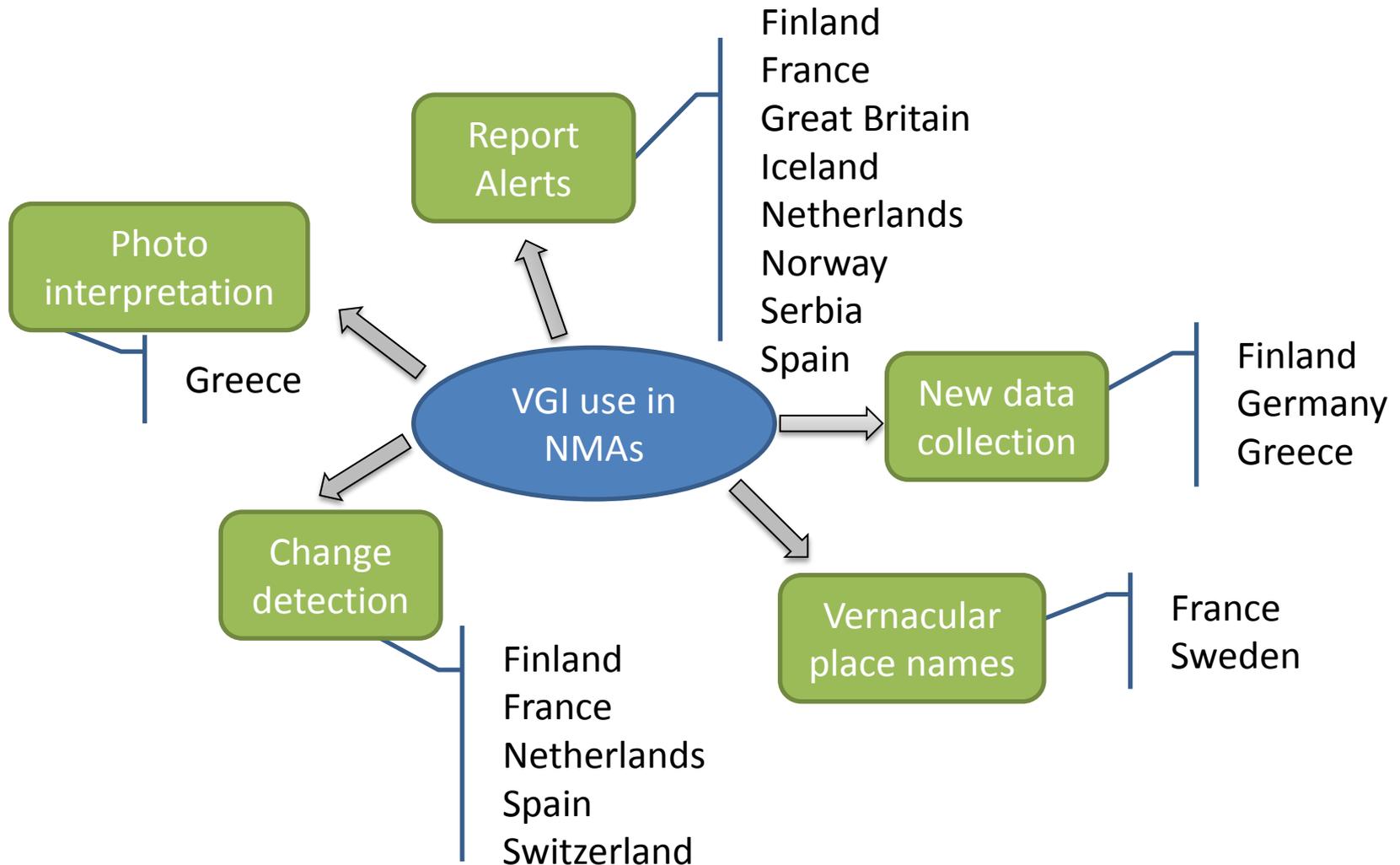
Small scale maps: generalization



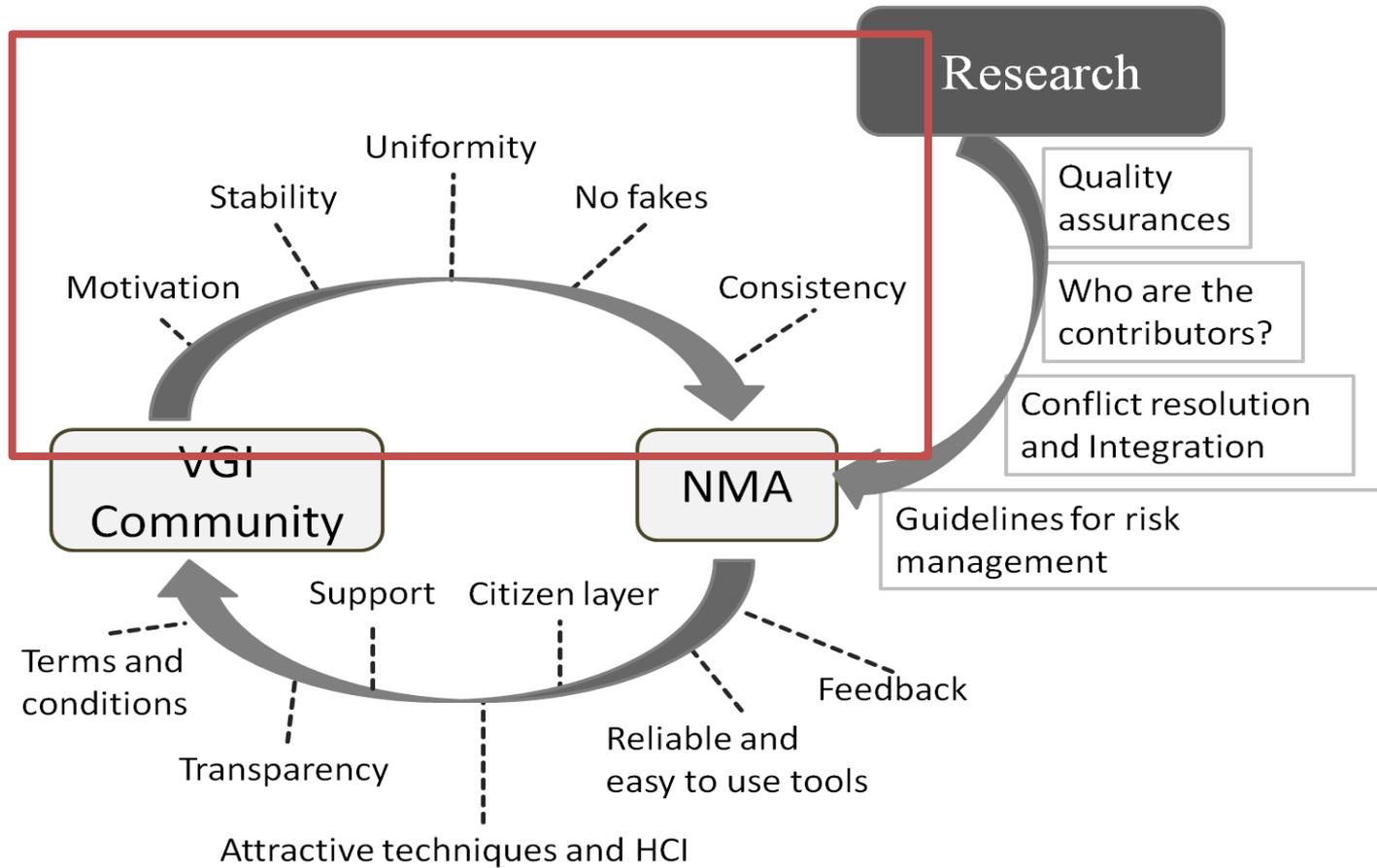
Outline of the presentation

- Quality Assurance with Crowdsourced Data
 - With reference data
 - Intrinsic assessment
- Map Making with Crowdsourced Data
 - Level of Detail heterogeneity
 - Large scale maps
 - Small scale maps
- **National Mapping Agencies (NMA) and Crowdsourcing**

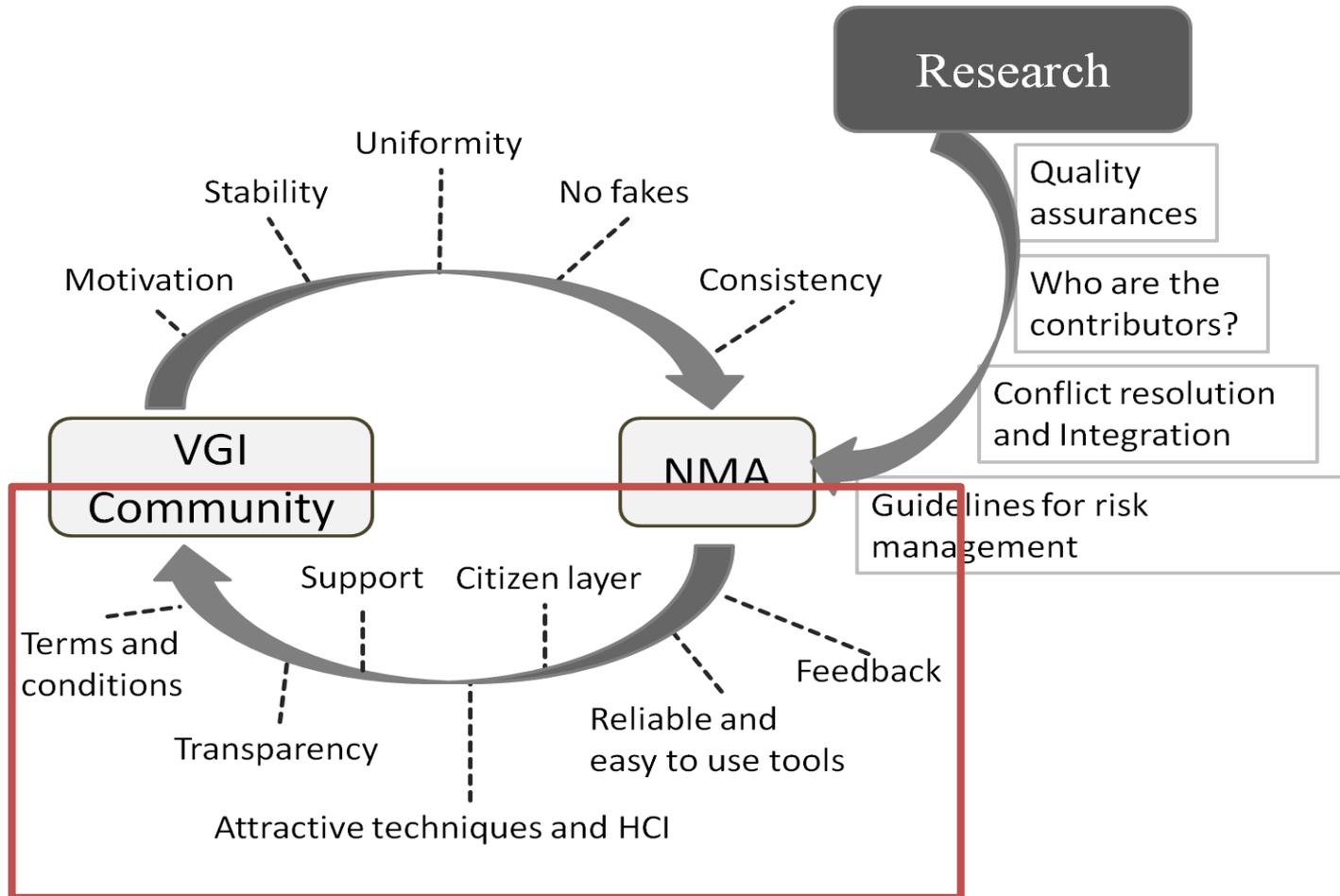
A survey within European NMAs



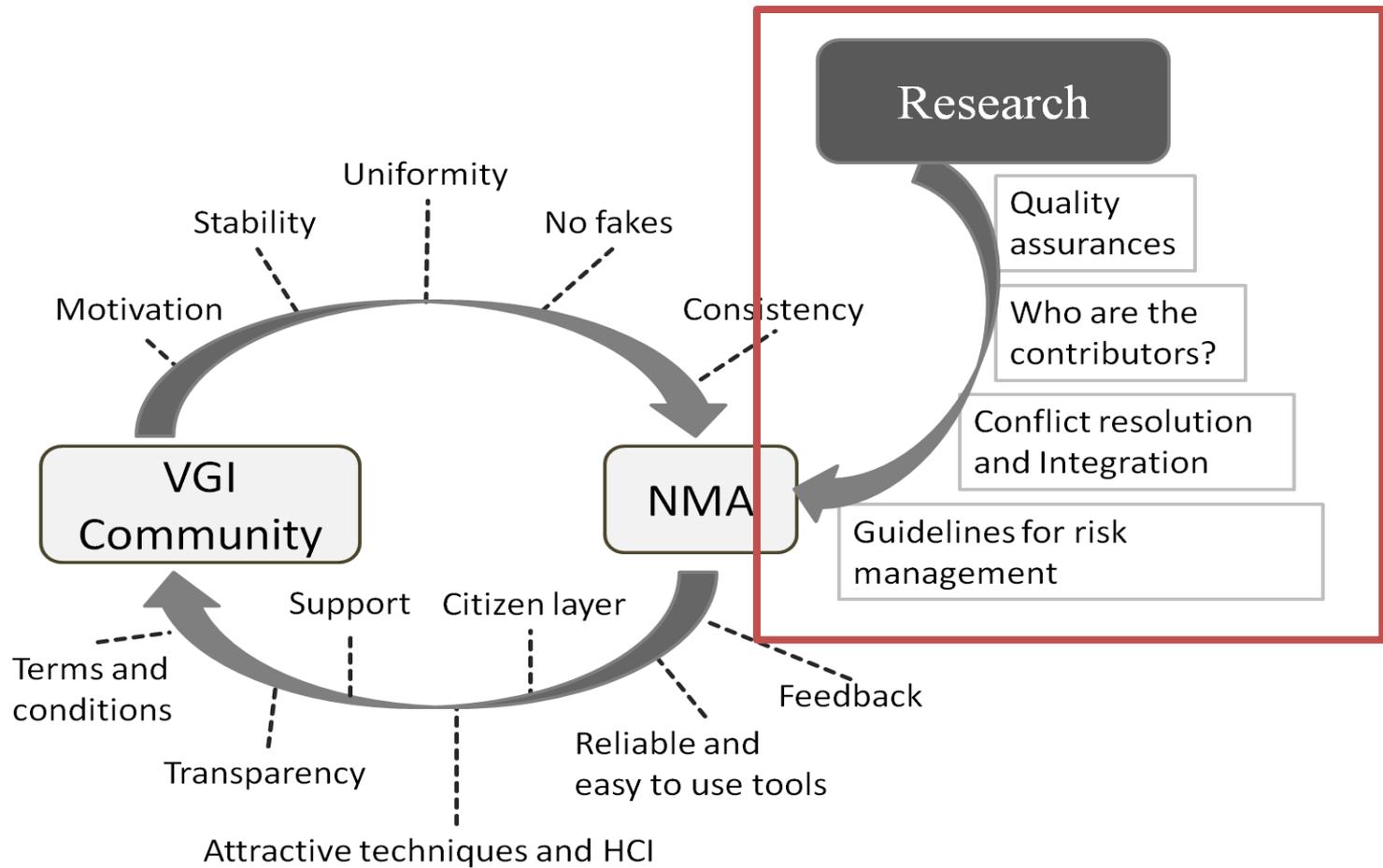
Mutual expectations



Mutual expectations



Mutual expectations





INSTITUT NATIONAL
DE L'INFORMATION
GÉOGRAPHIQUE
ET FORESTIÈRE



Collaborative Data Production: Research Perspectives

Thanks for your attention
Any question?

Guillaume Touya

EuroSDR Meeting – May 25th 2016

Bibliography

- Touya, G. and Savino, S. (2015). Enrichissement automatique et généralisation de réseaux ferrés. *Cartes & Géomatique - Revue du Comité Français de Cartographie*, 226:71-80.
- Touya, G. and Reimer, A. (2015). Inferring the scale of OpenStreetMap features. In Jokar Arsanjani, J., Zipf, A., Mooney, P., and Helbich, M., editors, *OpenStreetMap in GIScience, Lecture Notes in Geoinformation and Cartography*, pages 81-99. Springer International Publishing.
- Touya, G. and Girres, J.-F. (2014). Generalising unusual map themes from OpenStreetMap. In *Proceedings of 17th ICA Workshop on Generalisation and Multiple Representation*, Vienna, Austria. ICA.
- Touya, G. and Baley, M. (2014). Harmonizing level of detail in OpenStreetMap based maps. In Duckham, M., Stewart, K., and Pebesma, E., editors, *Proceedings of GIScience 2014 - Poster session*, Vienna, Austria.
- Touya, G. and Brando-Escobar, C. (2013). Detecting Level-of-Detail inconsistencies in volunteered geographic information data sets. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 48(2):134-143.
- Girres, J.-F. and Touya, G. (2010). Quality assessment of the french OpenStreetMap dataset. *Transactions in GIS*, 14(4):435-459.

Results for ATMs



- 347 ATMs in Paris
- 6,6% outside buildings: bad quality
- 65% on the wall of the buildings: good quality
- 5,9% >15m off centroid and 2m off walls: bad quality

| | Mean | Median | Standard deviation |
|----------------------|------|--------|--------------------|
| Distance to centroid | 17,6 | 13 | 12,9 |
| Distance to wall | 1,1 | 0 | 4,6 |

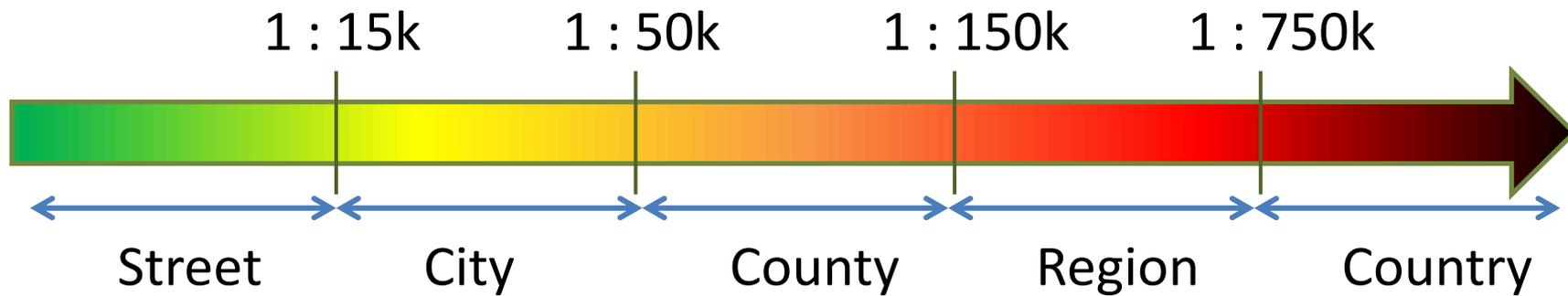
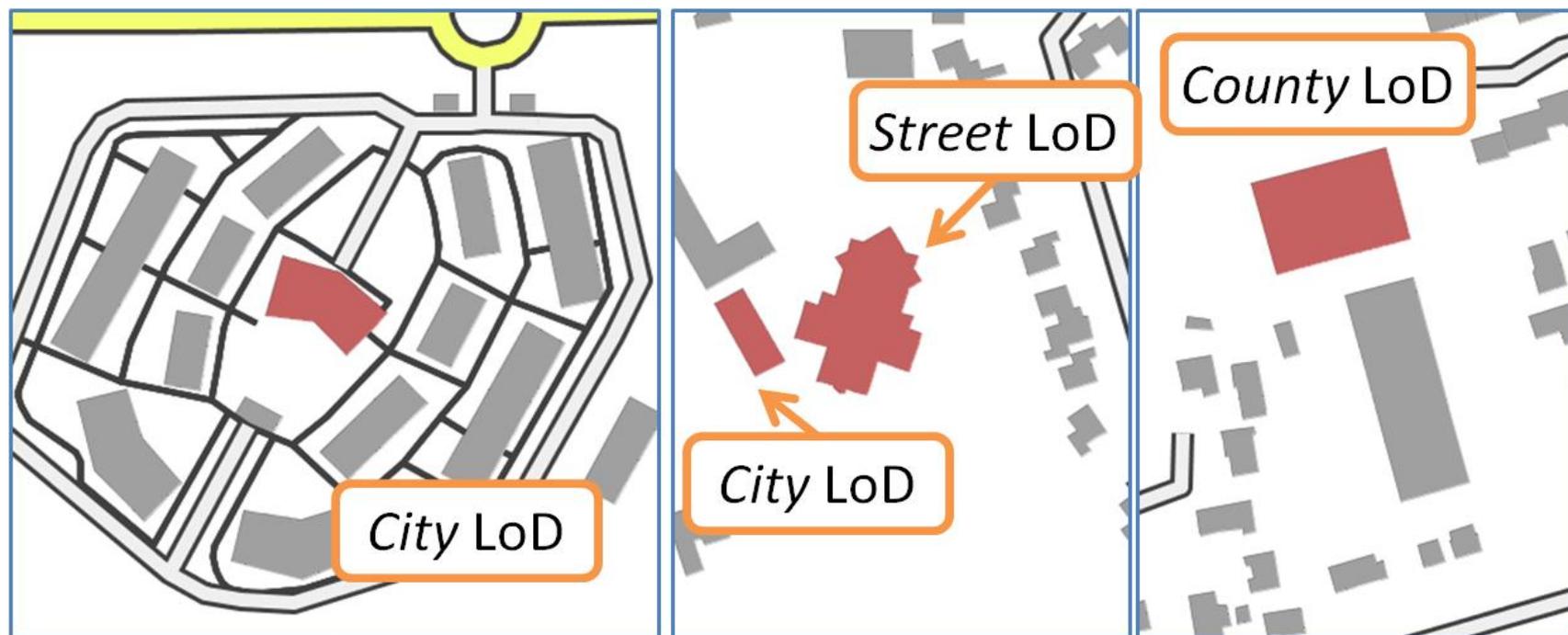
Results for bus stops



- ~2000 bus stops in Paris
- 13,5% too close (<2m for roads and <1,5m for buildings)
- Only 0,4 too far
- Quite homogeneous quality

| | Mean | Standard deviation | Maximum | % too close | % too far |
|-----------------------|------|--------------------|---------|-------------|-----------|
| Distance to roads | 6,2 | 2,9 | 18,5 | 6,8 | 0,4 |
| Distance to buildings | 7,4 | 5,8 | - | 6,7 | - |

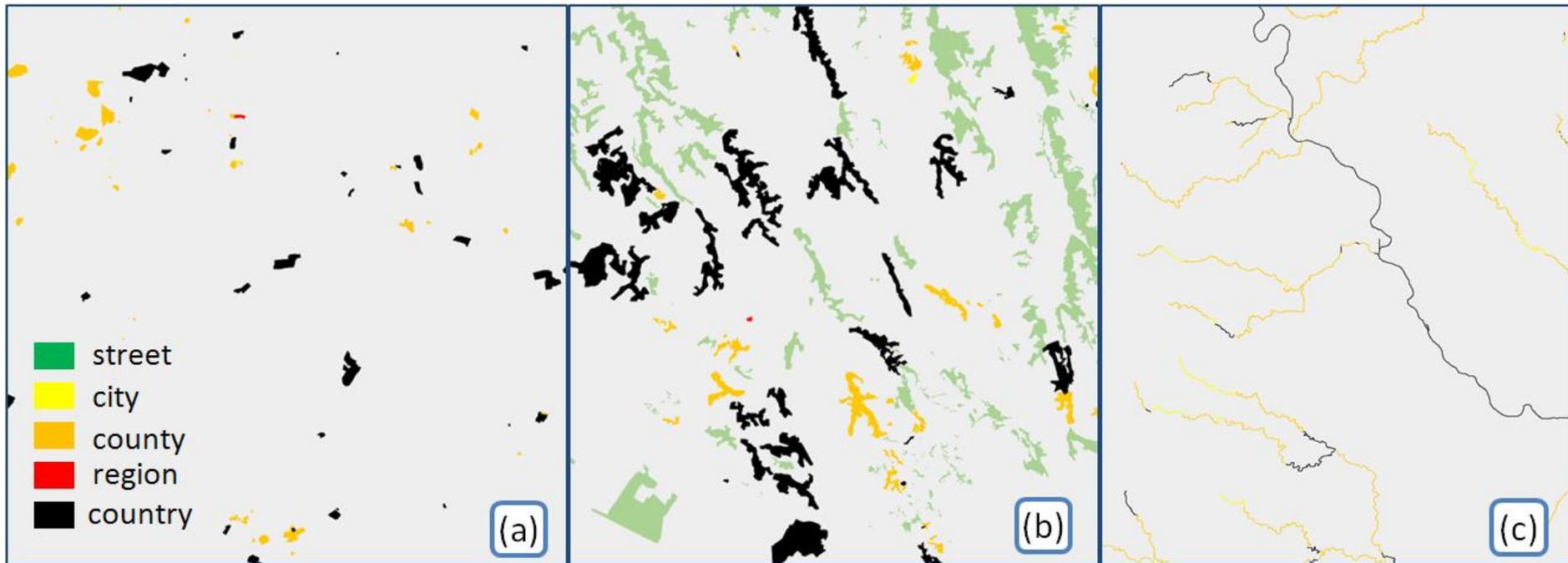
LoD inference results



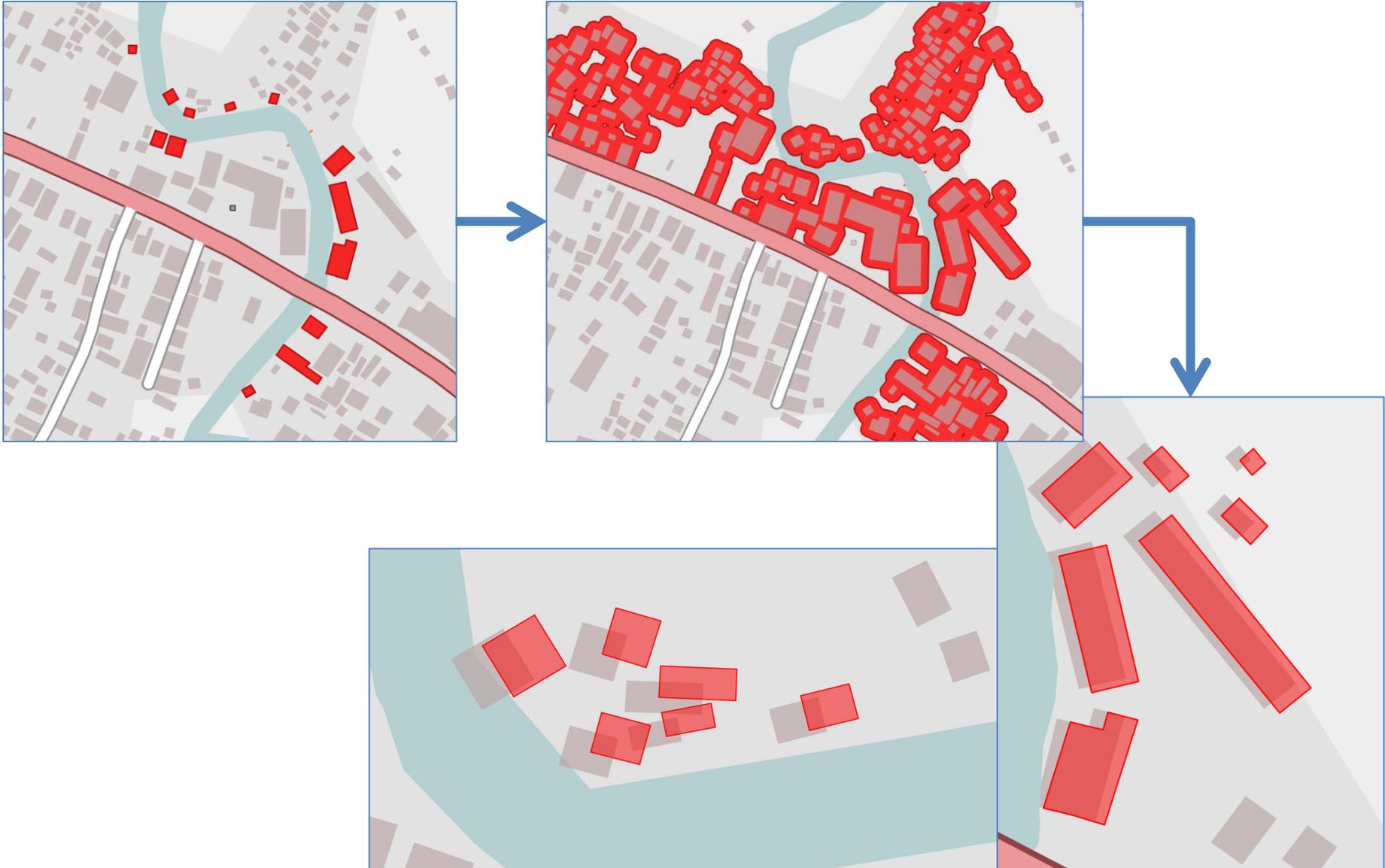
LoD inference results



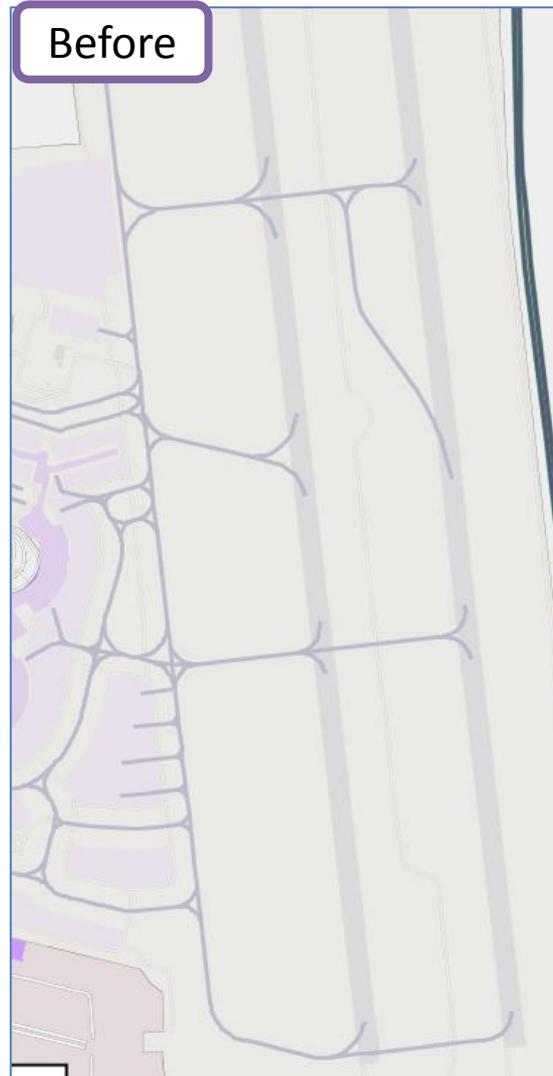
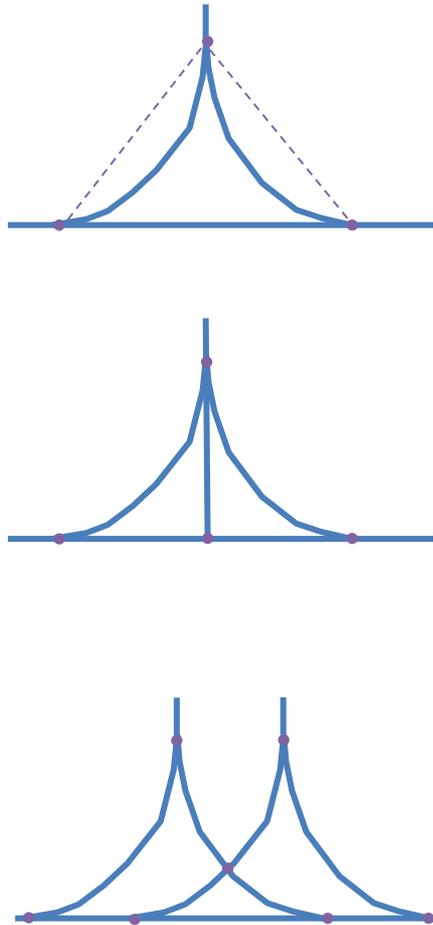
LoD inference results



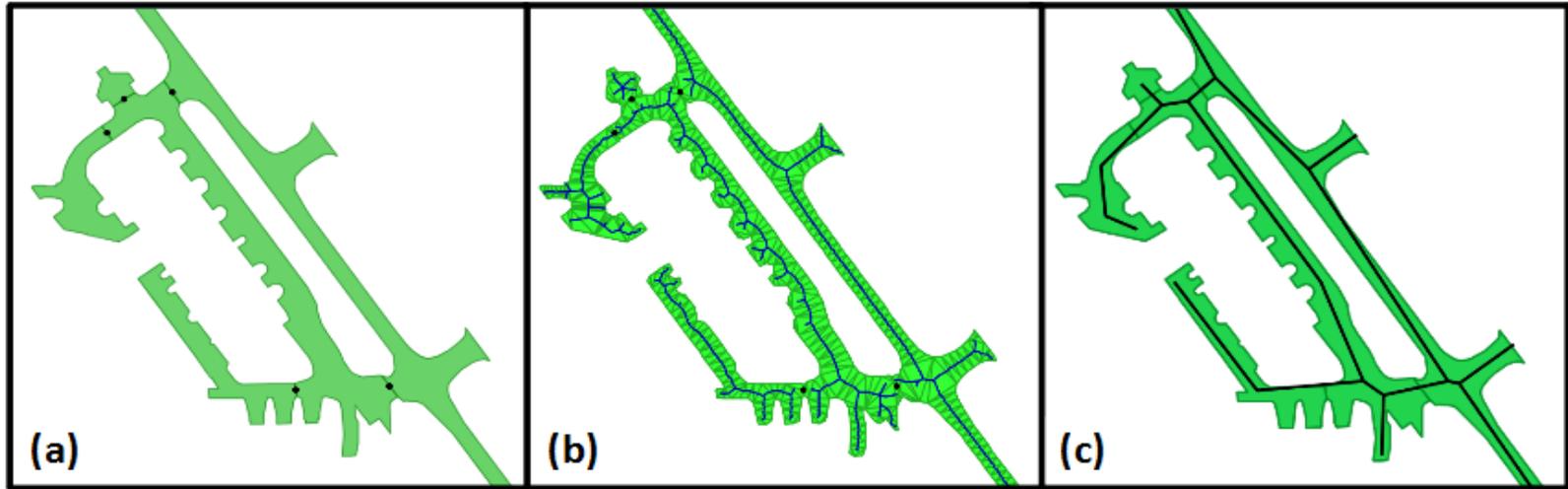
LoD harmonization results



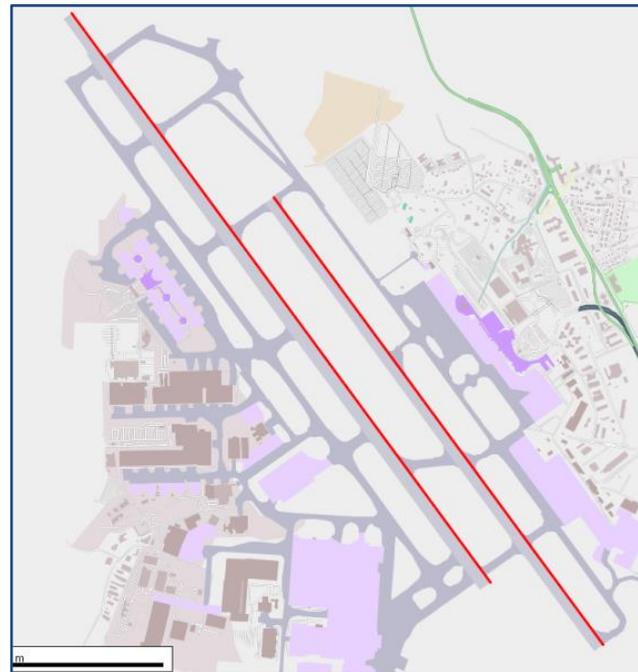
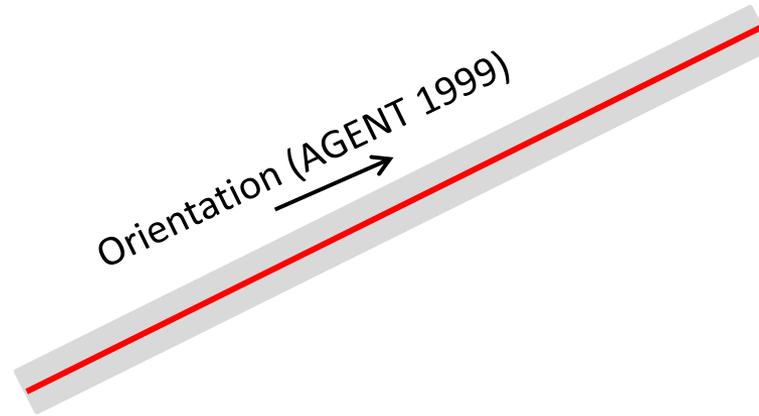
Taxiway generalization



Taxiway area generalization



Runway generalization



Main railways generalization

