

Hangzhou Forum on
United Nations Global Geospatial Information Management

Data Quality Control: Crowd-Sourcing Geospatial Information

Xizi Hotel, Hangzhou
24-26 May, 2012

Presented : Prof. Dr. Woosug Cho, Senior Advisor, NGII, Republic of Korea
Prepared : Prof. Dr. Ki-Joune Lee, Senior Advisor, NGII, Republic of Korea
Mr. Jun-Sung Kim, Deputy Director, NGII, Republic of Korea



Definition of Crowd-Sourcing

Crowd-sourcing is a process that involves outsourcing tasks to a distributed group of people. This process can occur both online and offline, and the difference between crowdsourcing and ordinary outsourcing is that a task or problem is outsourced to an undefined public rather than a specific body, such as paid employees.

Shortly, a distributed problem-solving and production model.

The term "crowdsourcing" is a portmanteau of "crowd" and "outsourcing," coined by Jeff Howe in 2006.

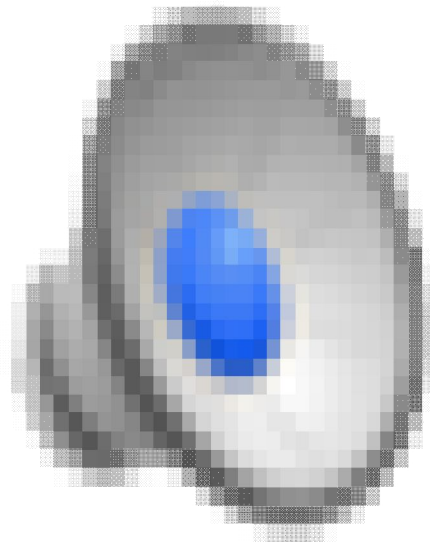
Crowd-Sourcing Geospatial Information

VGI (Volunteered Geographic Information) :
the harnessing of tools to create, assemble, and disseminate geographic data provided voluntarily by individuals (Goodchild, 2007)

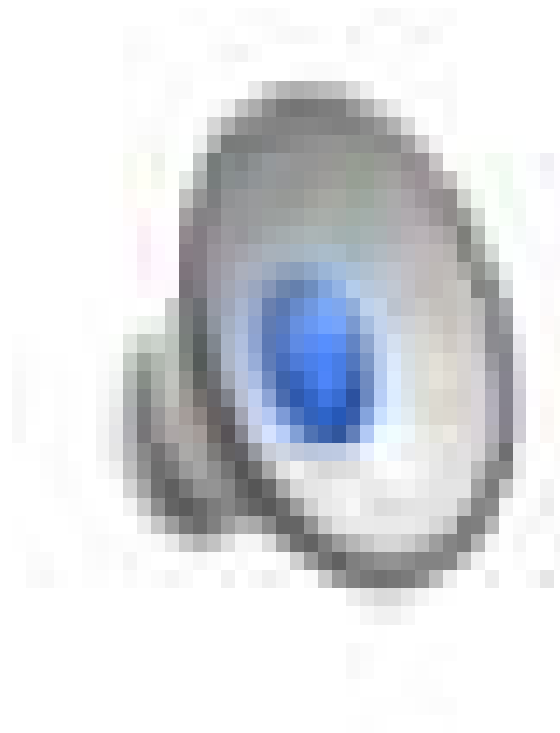
Examples on Crowd-Sourcing Geospatial Information:

1. OpenStreetMap (<http://www.openstreetmap.org/>)
<http://www.youtube.com/watch?v=lpXk5W6b298&feature=channel&list=UL>
2. Project Haiti: <http://vimeo.com/9182869>
3. Google Map Maker (<http://www.google.com/mapmaker>)
http://www.ted.com/talks/lalitesh_katragadda_making_maps_to_fight_disaster_build_economies.html

Project Haiti



Google Map Maker: Myanmar Case



Example in UK –Google Map & OSM

Map Compare

Choose map type: Google (Map) | Choose map type: OSM Mapnik

Data/Maps Copyright 2011 Geofabrik GmbH and OpenStreetMap Contributors | License: Creative Commons BY-SA 2.0 | Contact

-0.10357, 51.51887 zoom=17 Permalink

인터넷 | 북마크 | 지도 설정 | 100%

<http://tools.geofabrik.de/mc>

Example in Korea – Google Map & OSM

Map Compare

Choose map type: Google (Map) | Choose map type: OSM Mapnik

Data/Maps Copyright 2011 Geofabrik GmbH and OpenStreetMap Contributors | License: [Creative Commons BY-SA 2.0](https://creativecommons.org/licenses/by-sa/2.0/) | [Contact](https://www.geofabrik.de/)

zoom=15 Permalink

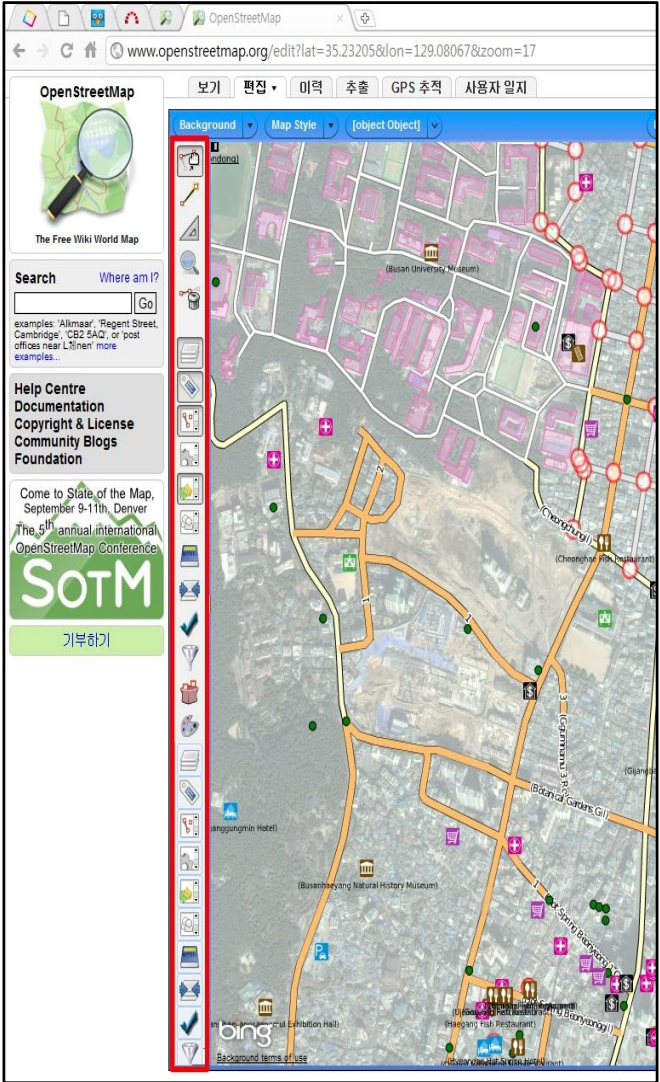
원본

인터넷 | 보호 모드: 설정

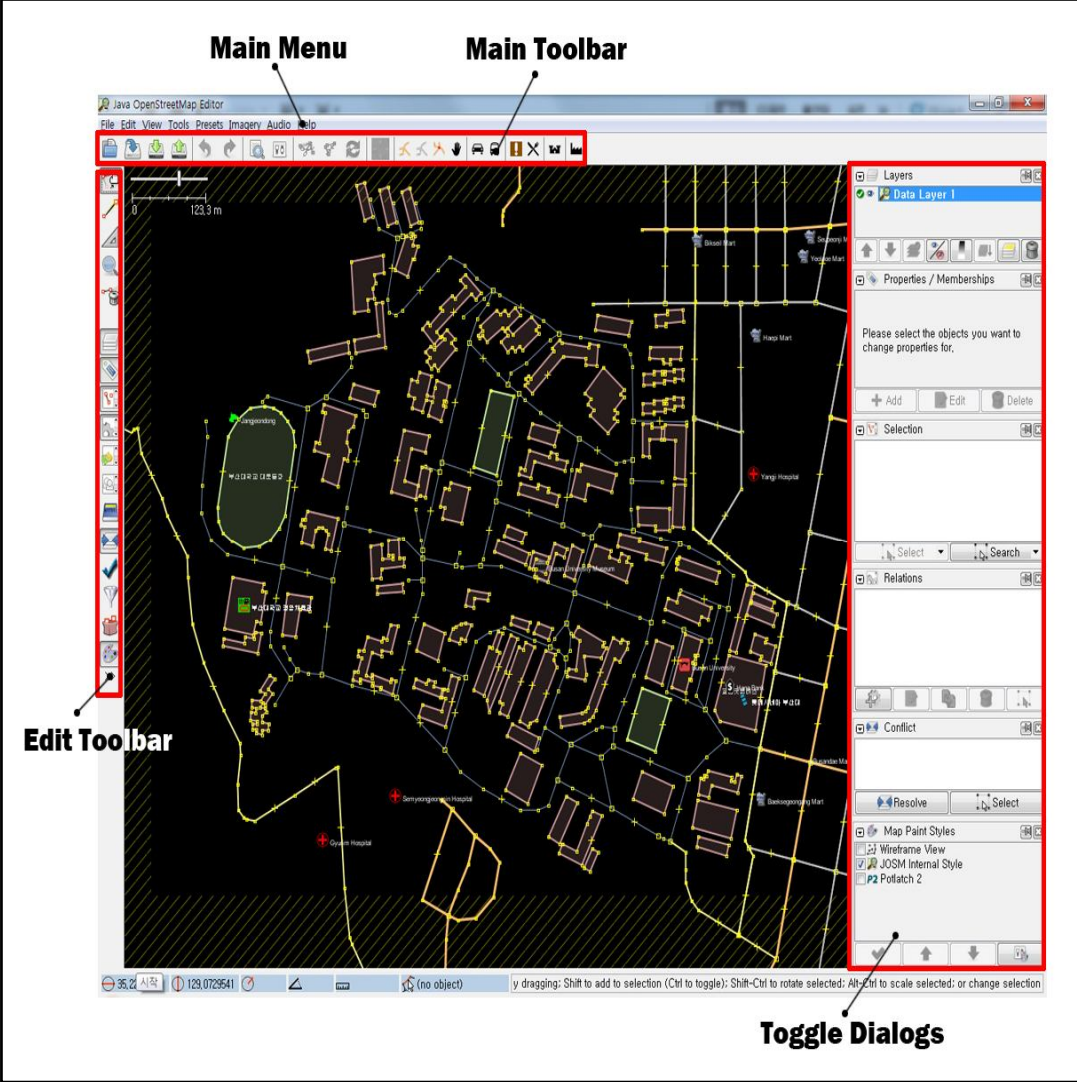
100%

<http://tools.geofabrik.de/mc>

PotLatch

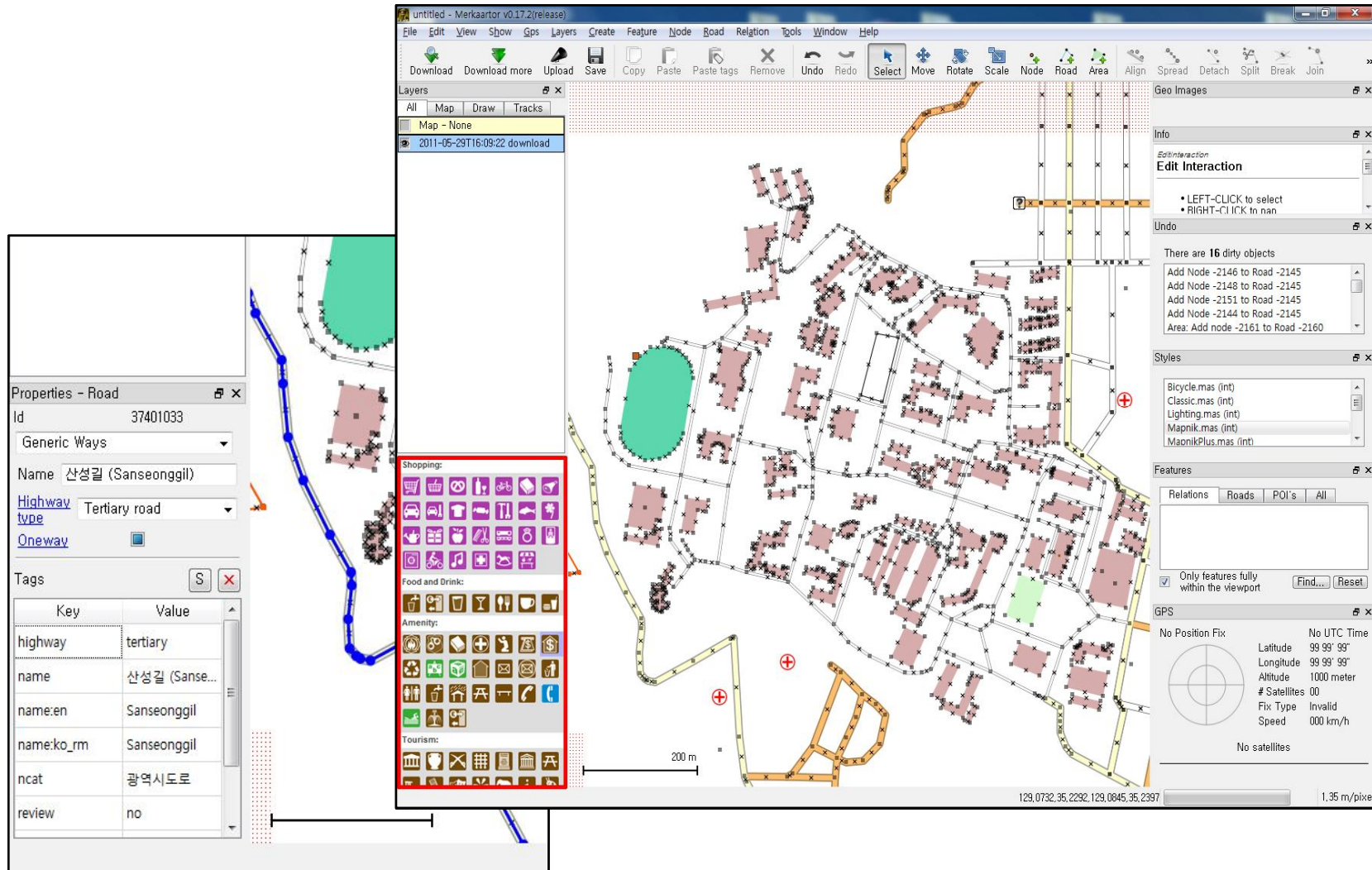


JOSM(Java OSM editor)

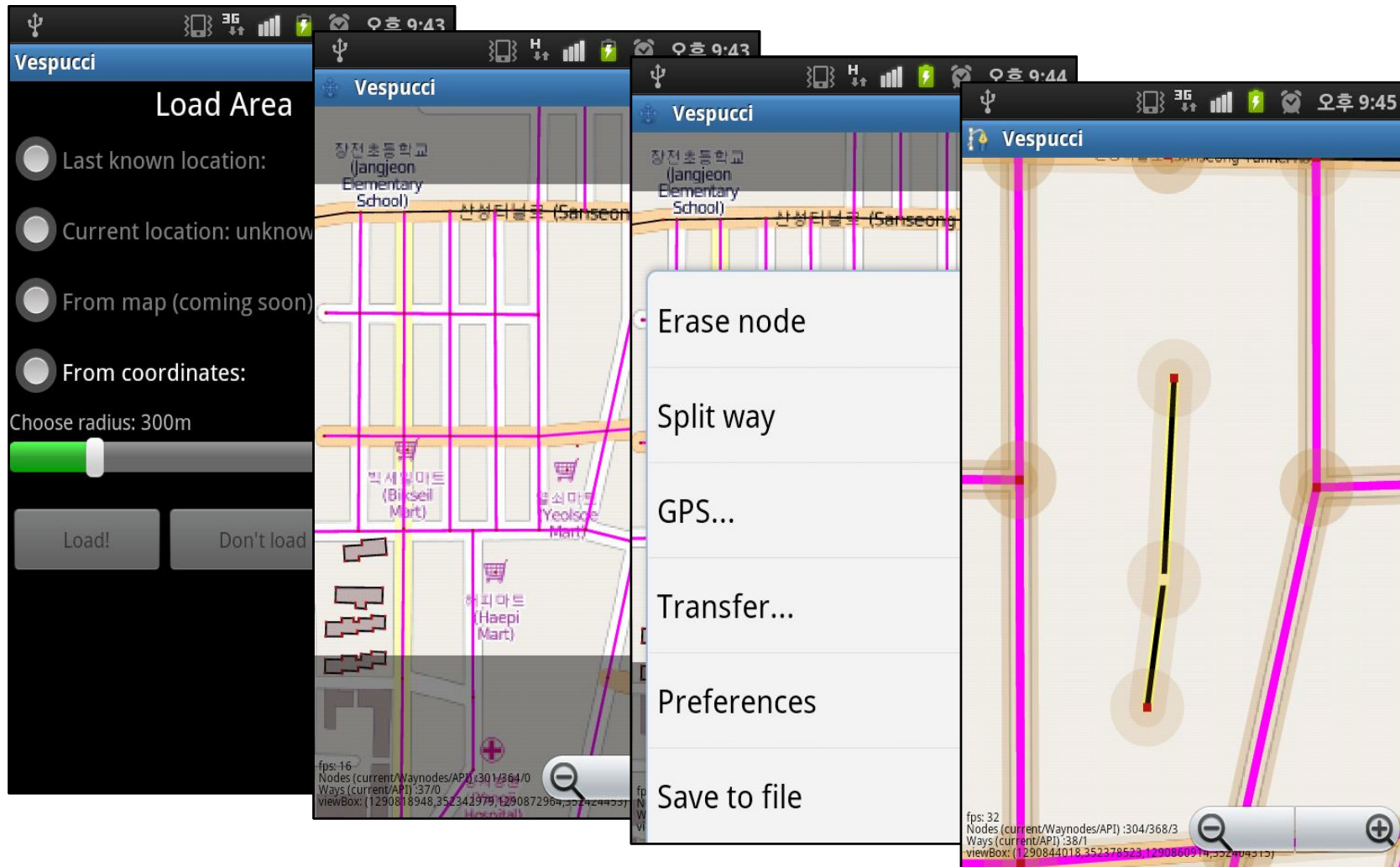


Data Quality Control of Crowd-Sourcing Geospatial Data

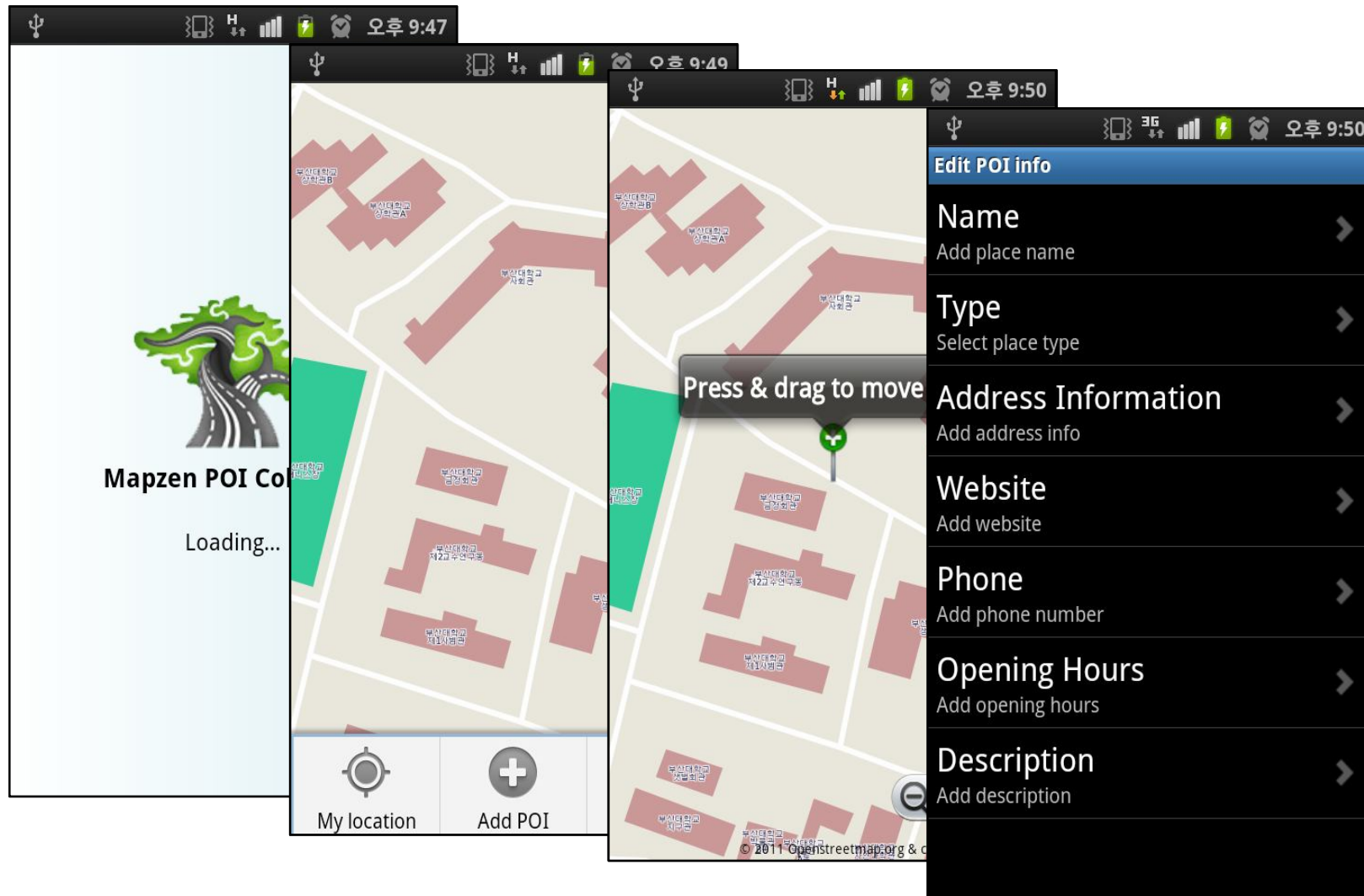
Merkaartor



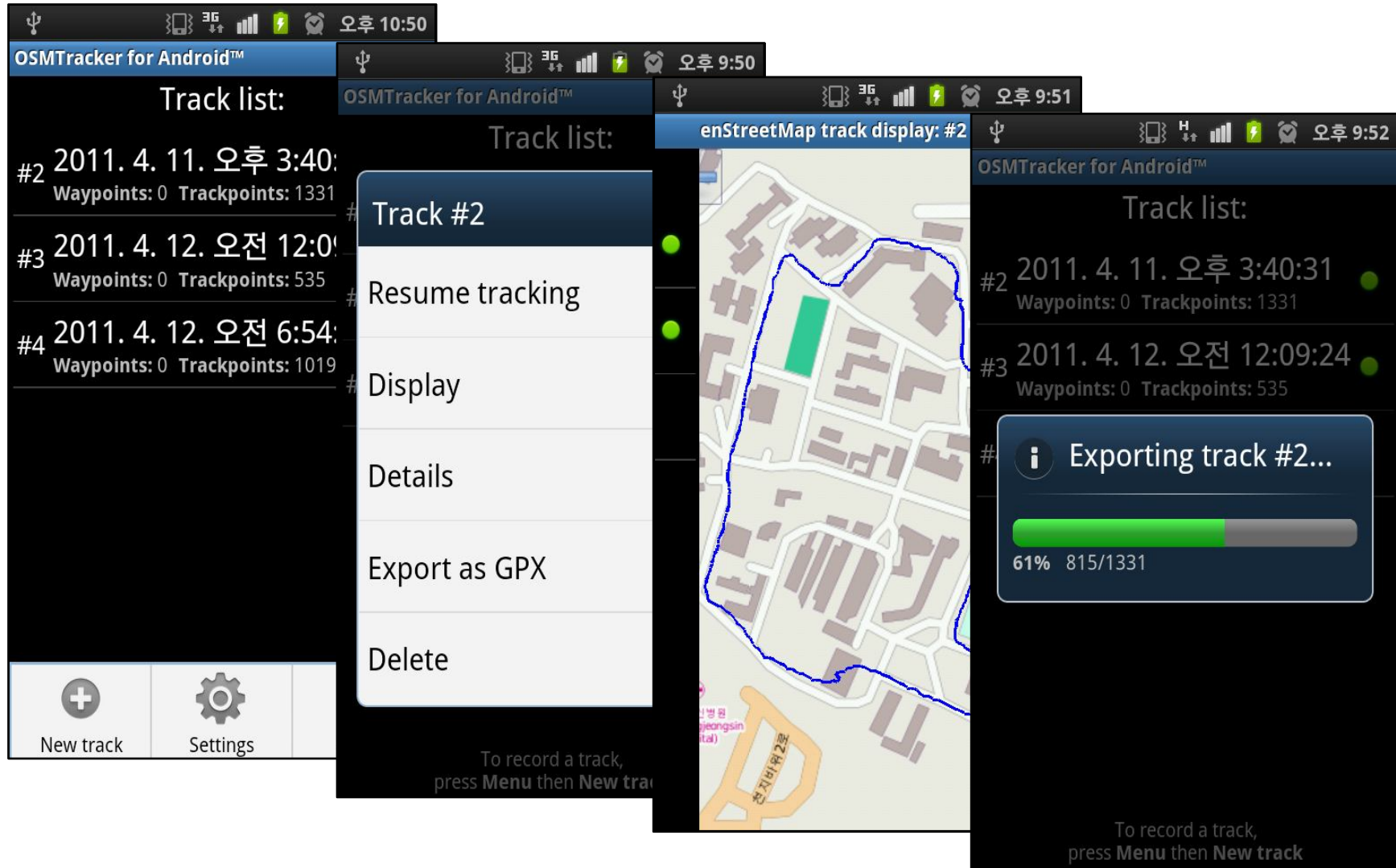
Vespucci



Mapen POI Collector



OSMTracker



A Comparative Study of Crowd-Sourcing GI

OSM vs. OS datasets

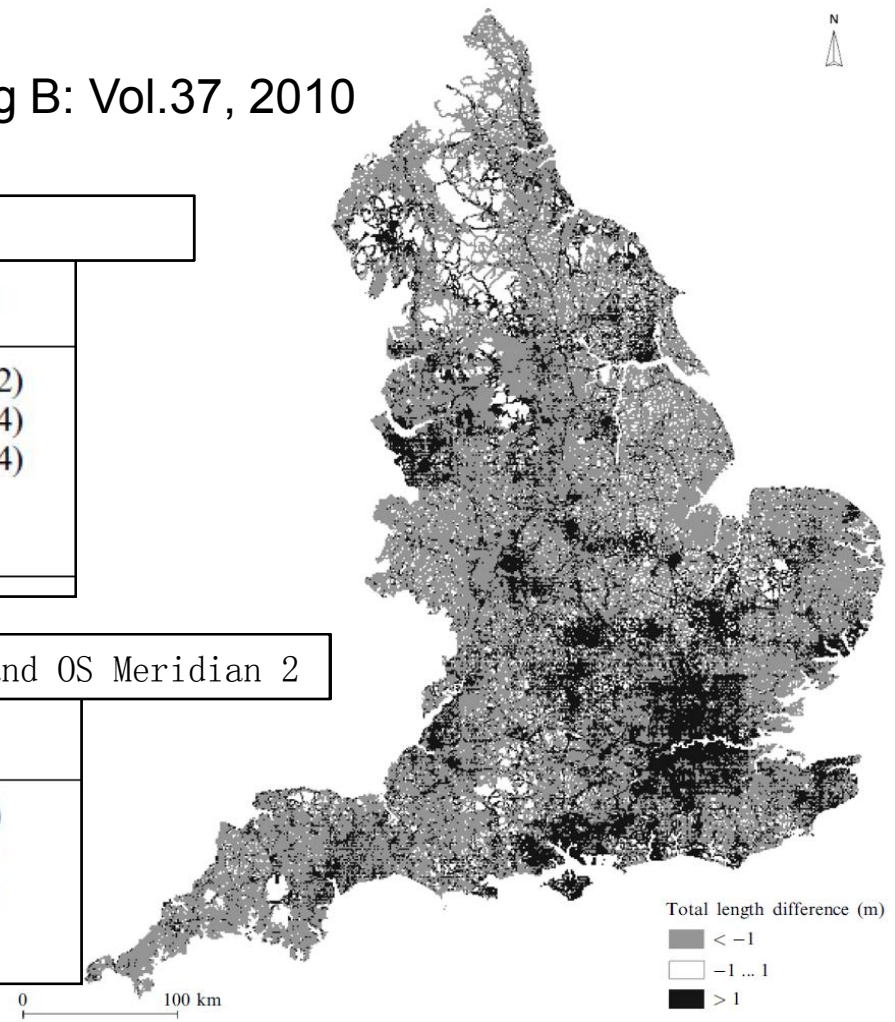
- M. Haklay, Environment and Planning B: Vol.37, 2010

Length comparison between OSM and OS Meridian 2

Cells	Area (km ²)
Empty cells	16 300 (13.2)
Meridian 2 more detailed than OSM	75 977 (61.4)
OSM more detailed than Meridian 2	31 437 (25.4)
Total	123 714

Length comparison with attributes between OSM and OS Meridian 2

Cells	Area (km ²)
Empty cells	17 632 ^a (14.3)
Meridian 2 more detailed than OSM	80 041 (64.7)
OSM more detailed than Meridian 2	26 041 (21.0)
Total	123 714

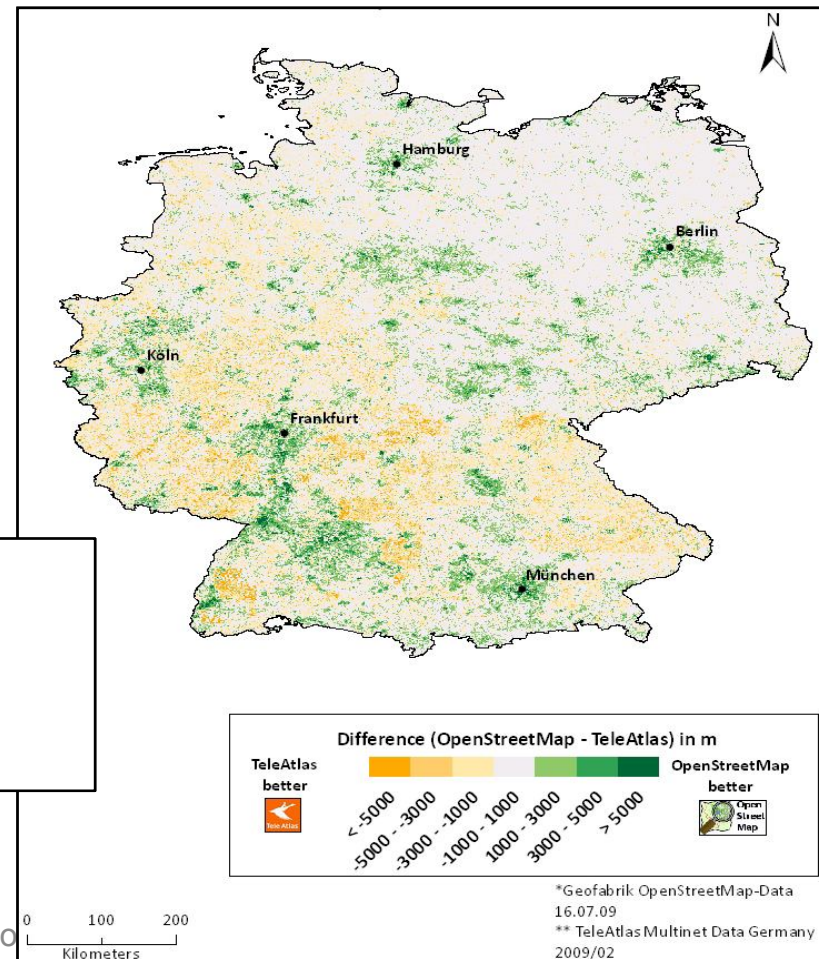


A Comparative Study of Crowd-Sourcing GI

OSM vs. Tele Atlas in Germany

- D. Zielstra and A. Zipf, Agile Conf., 2010

Total Length Differences of all Street Network Data between OpenStreetMap and Tele Atlas per Km²



Weakness of Crowd-Sourcing GI

NO Quality Control on GI and NO Metadata

- Data quality depends on each individual volunteer
- No metadata
- Only some lineage data (change-sets) is provided

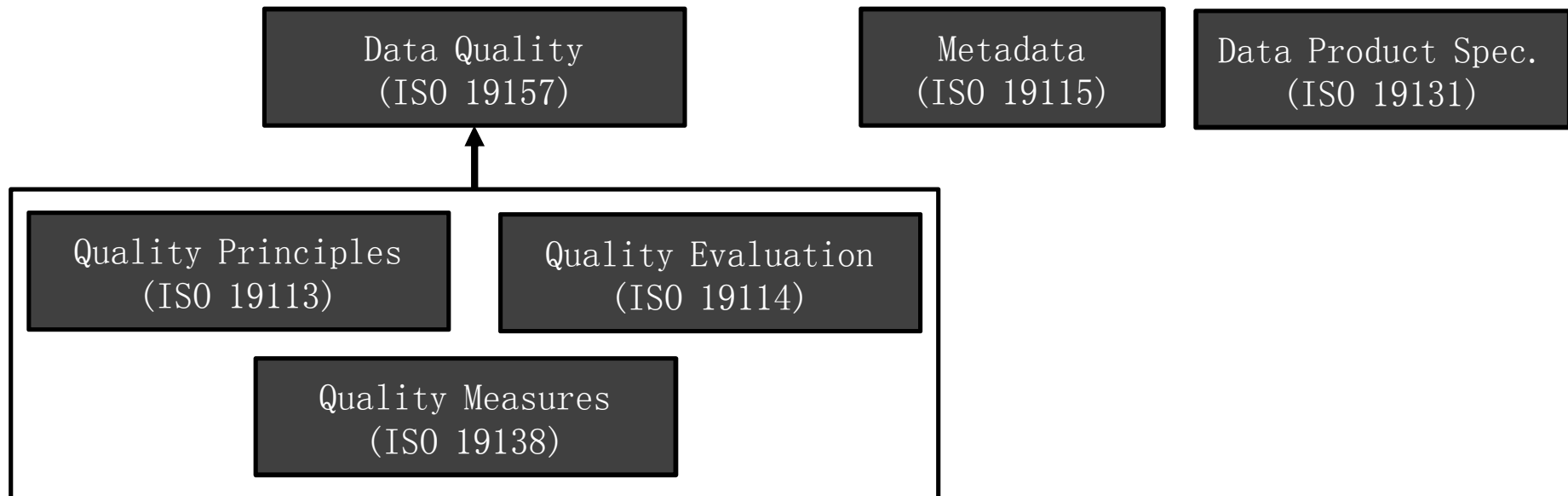
Changesets

Changesets within -0.14419,51.48869,-0.10729,51.52836

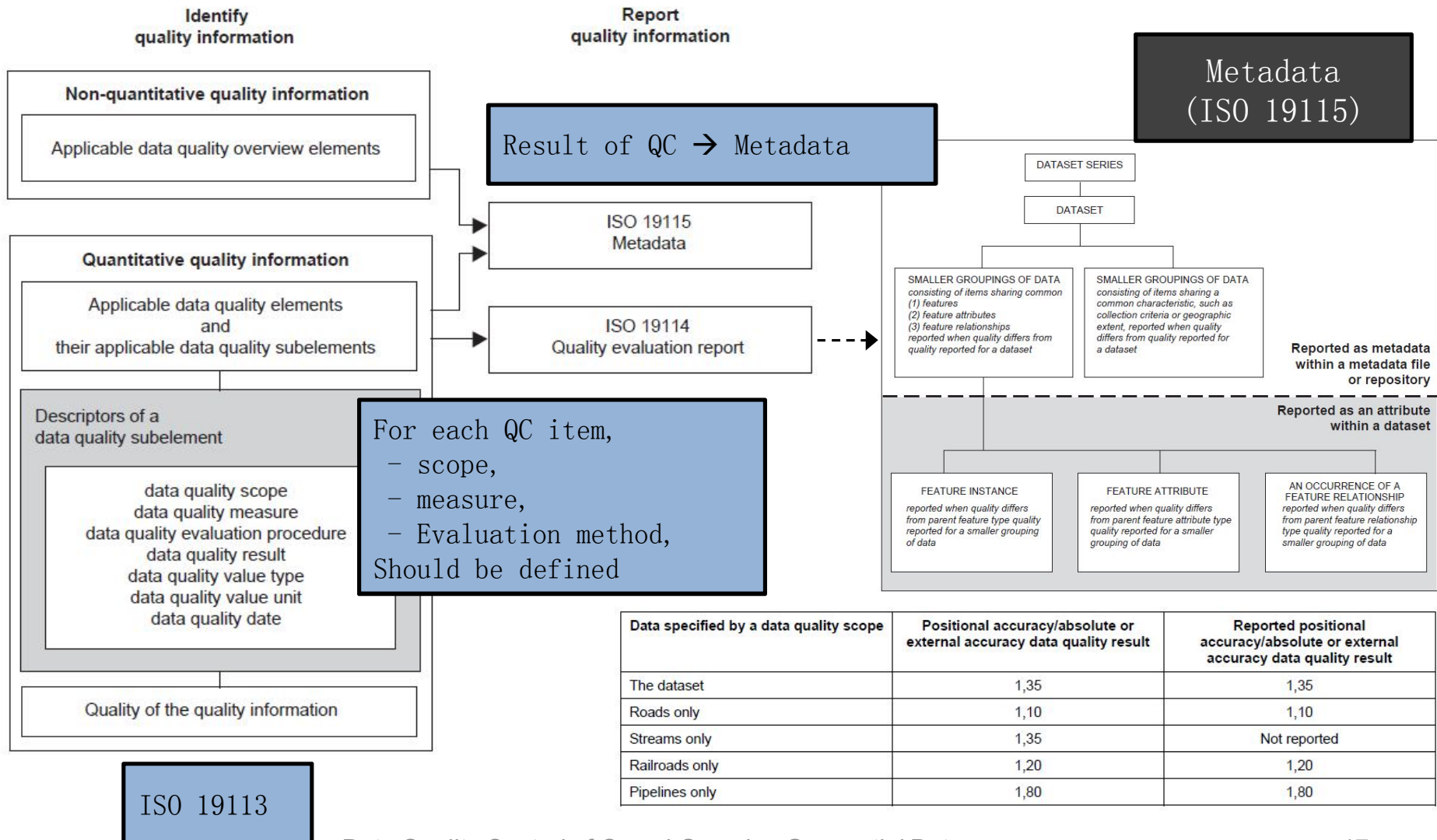
« Previous | Showing page 1 | Next »

ID	User	Date	Description
#11602259	wheelmap_visitor	(현재 수정 중)	Modified via wheelmap.org
#11598988	Robert Whittaker	2012년 5월 14일, 20시 18분 50초 UTC	Fix tagging
#11598378	mapper999	2012년 5월 14일, 22시 33분 03초 UTC	railways
#11594934	fitzsimons	2012년 5월 14일, 12시 46분 48초 UTC	Added missing Costa Coffee branch after a visit last week.
#11592660	DavidThackray	2012년 5월 14일, 07시 13분 40초 UTC	(없음)
#11592640	wheelmap_visitor	2012년 5월 14일, 22시 57분 14초 UTC	Modified via wheelmap.org
#11591455	timhwhitehouse	2012년 5월 14일, 00시 47분 08초 UTC	footpath addition near Derwent and Bredsall north of Derby
#11588276	Robert Whittaker	2012년 5월 13일, 17시 18분 27초 UTC	Change royal_cypher=none to royal_cypher=no as per wiki
#11588070	Maarten Deen	2012년 5월 13일, 17시 01분 42초 UTC	e-road e40
#11586576	Maarten Deen	2012년 5월 13일, 14시 54분 25초 UTC	reverted empty relation to v2878
#11583438	wheelmap_visitor	2012년 5월 13일, 17시 24분 53초 UTC	Modified via wheelmap.org
#11581375	mercurium	2012년 5월 13일, 00시 33분 49초 UTC	Added road

Data Quality Control in ISO 191xx



Data Quality Evaluation – ISO

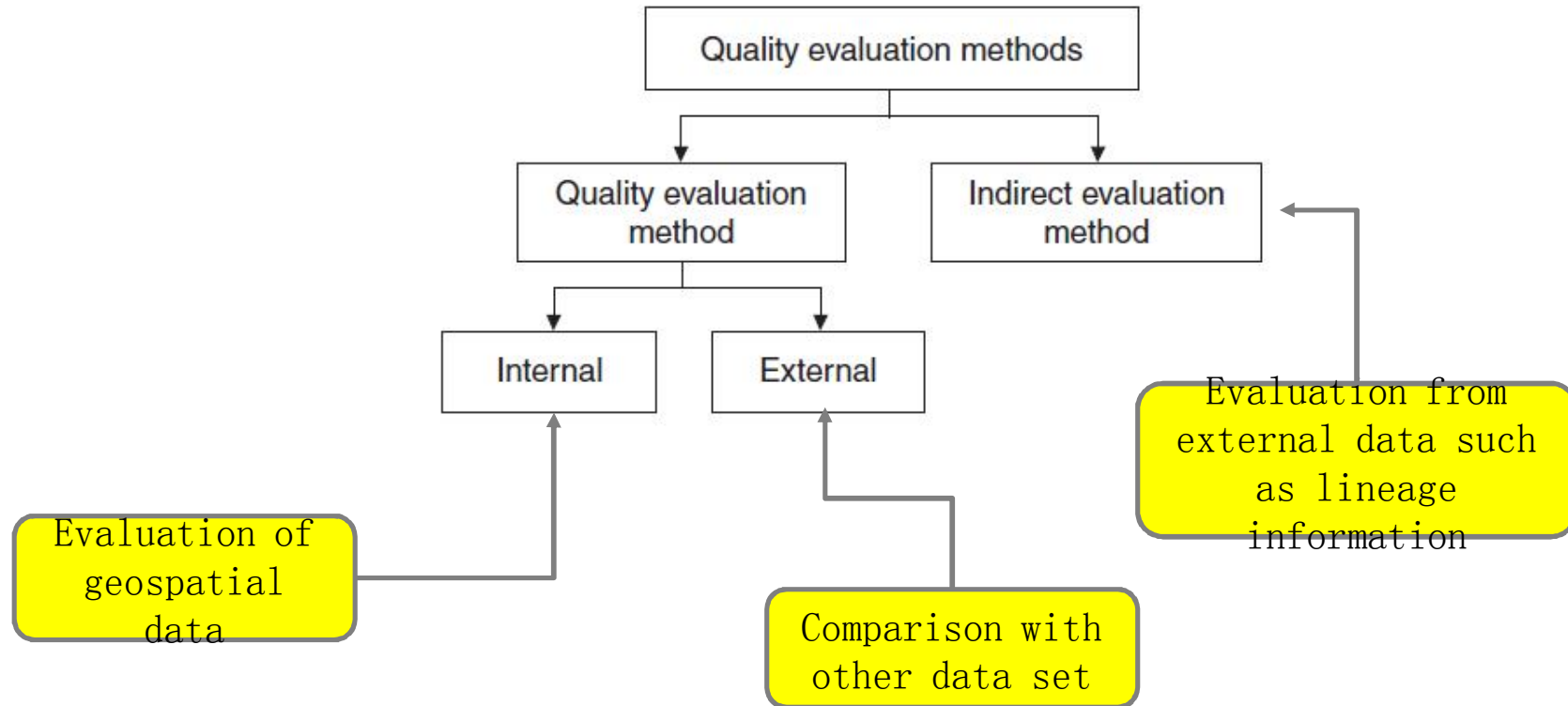


For each QC item,
 - scope,
 - measure,
 - Evaluation method,
 Should be defined

ISO 19113

Data specified by a data quality scope	Positional accuracy/absolute or external accuracy data quality result	Reported positional accuracy/absolute or external accuracy data quality result
The dataset	1,35	1,35
Roads only	1,10	1,10
Streams only	1,35	Not reported
Railroads only	1,20	1,20
Pipelines only	1,80	1,80

Quality Evaluation Method – ISO 19114



Data Quality Control for Crowd-Sourcing GI

Impossible to enforce data quality control process to volunteer map makers

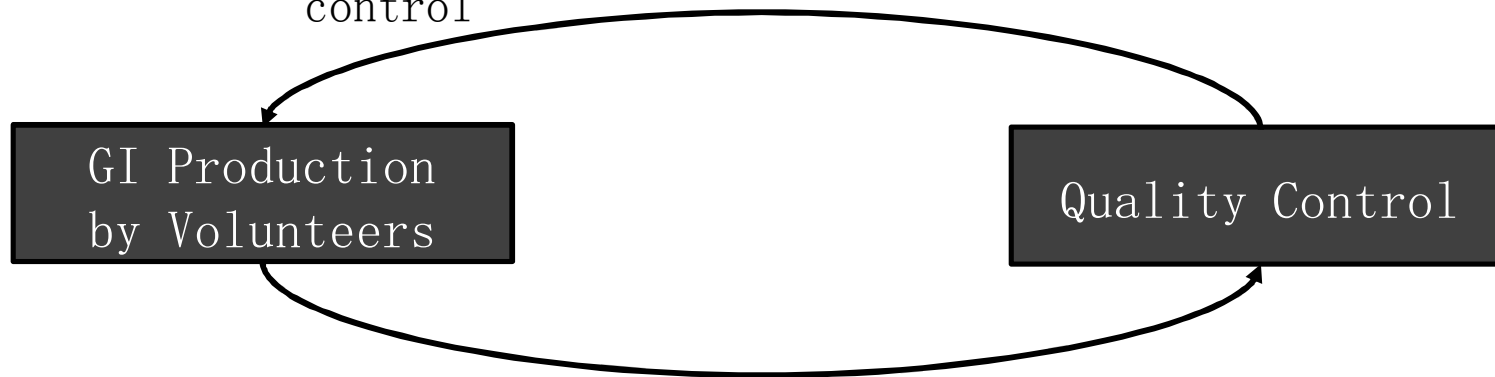
- Possibly by governmental or semi-governmental institutes

ISO Quality Control

- As defined by ISO 19113, 19114, and 19138
- Direct Evaluation
 - Comparison with other map data such as national maps
- Indirect Evaluation, such as
 - Number of participants at given area,
 - Profile of each participant, and
 - Frequency of updates and last update date

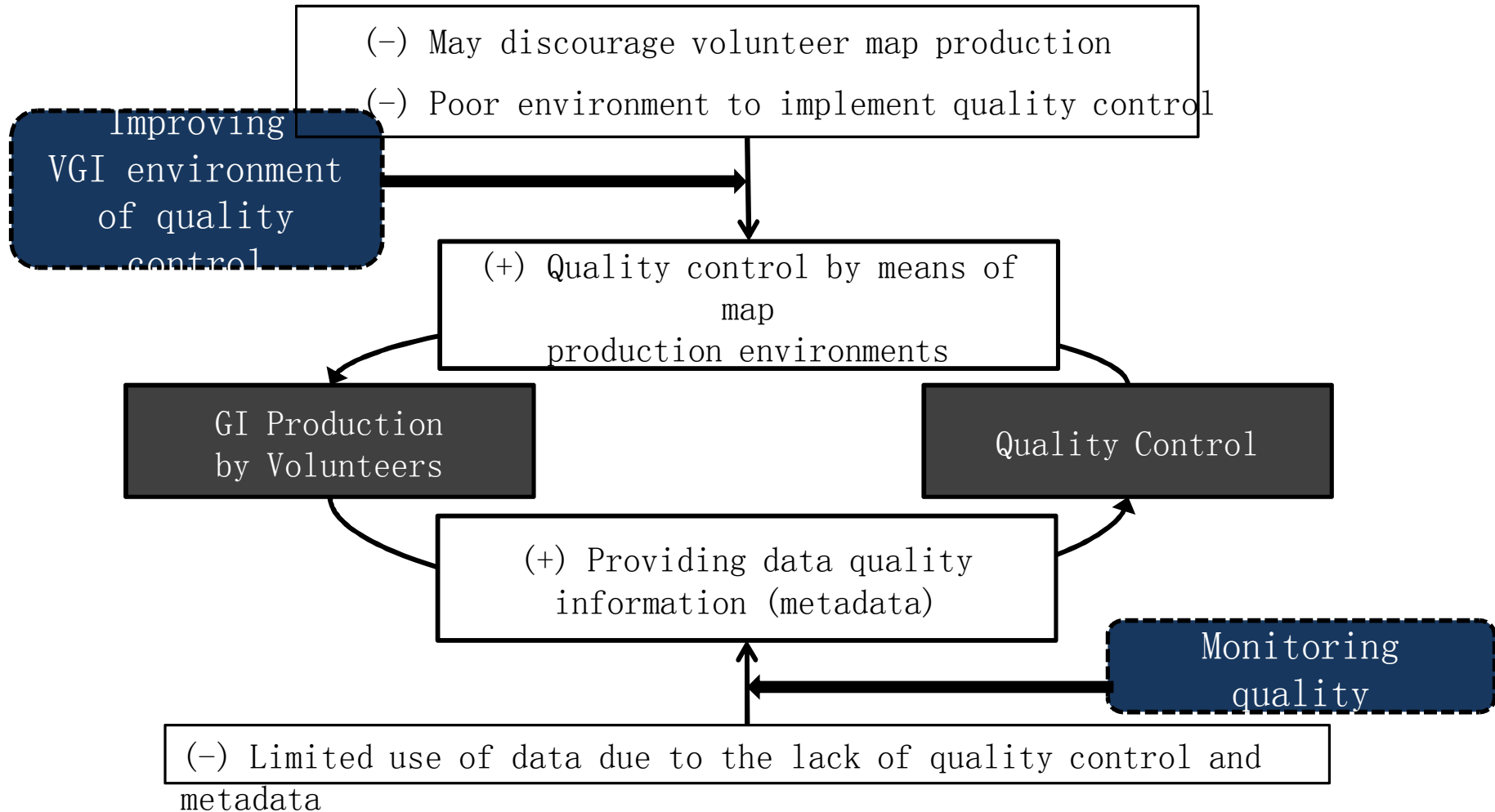
Data Quality Control for Crowd-Sourcing GI

- (-) May discourage volunteer map production
- (-) Poor environment to implement quality control

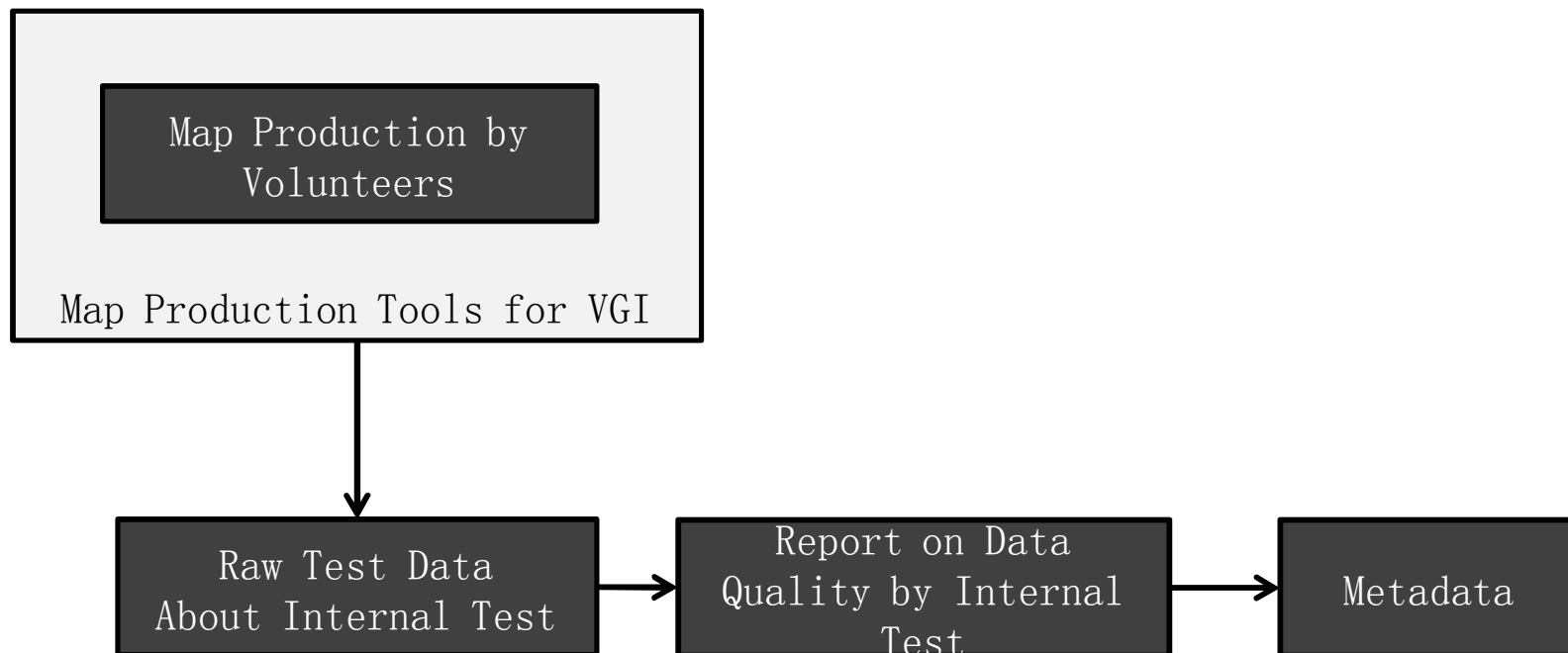


- (-) Limited use of data due to the lack of quality control and metadata

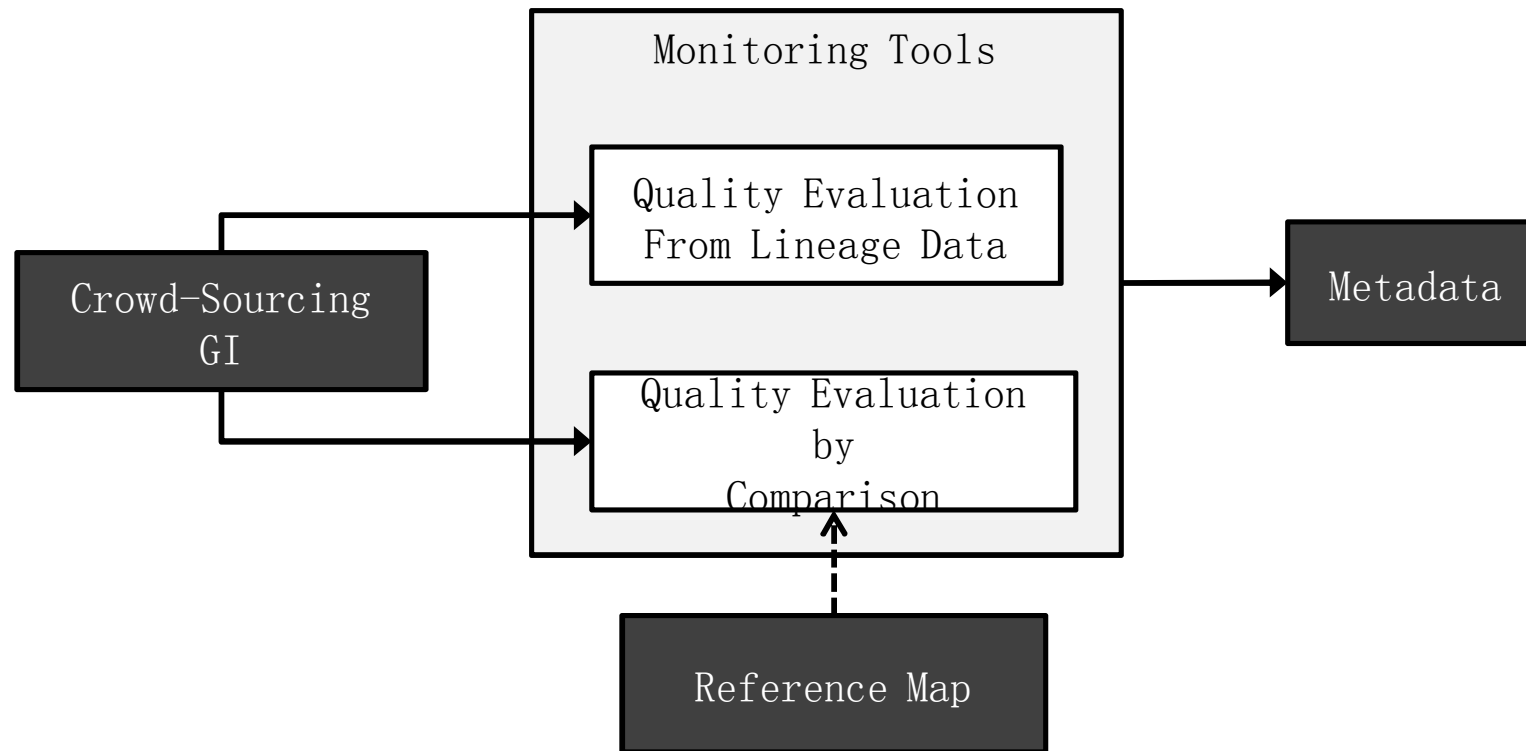
Data Quality Control for Crowd-Sourcing GI : How to break the vicious circle



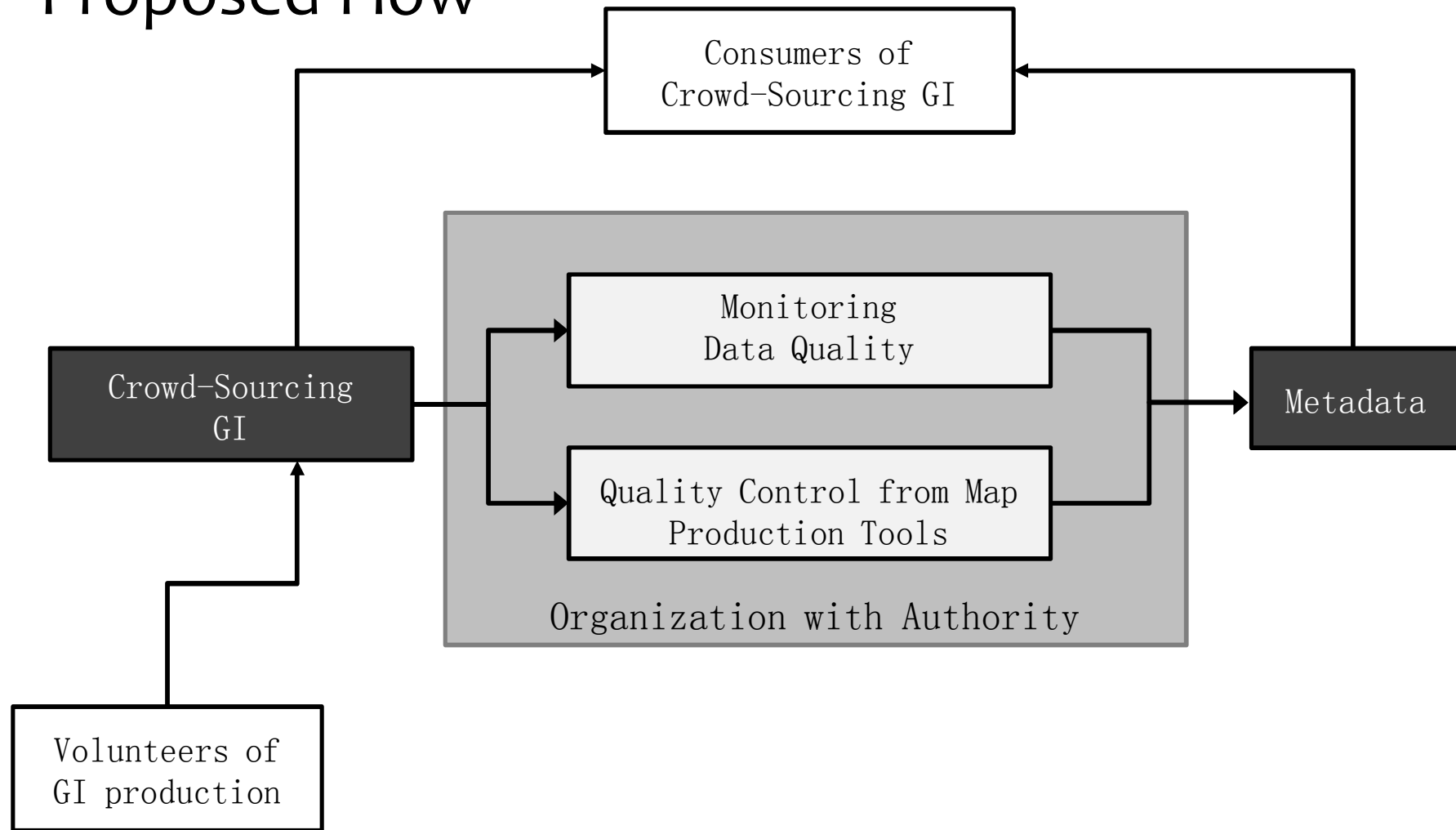
Data Quality Control for Crowd-Sourcing GI : Improving Map Production Environment



Data Quality Control for Crowd-Sourcing GI : Monitoring Data Quality



Data Quality Control for Crowd-Sourcing GI : Proposed Flow



Summary

Crowd-Sourcing GI

- An Important Source of GI
- Very helpful for developing countries

Difficult to enforce rigorous quality control

Quality Control of Crowd-Sourcing GI

- Map Production Tool with QC functions
- Monitoring QC of Crowd-Sourcing GI