Quality control of Large-scale Reference Data at the Flemish Geographical Information Agency

Lies De Meulenaer – Barbara Bulkens
Agenda

- FGIA (AGIV)
- LRD
- Quality control of LRD data
- MIRO
- Future
Flanders? Belgium?
Flanders? Belgium?
Who is FGIA?

- Flemish Geographical Information Agency
- Public organisation, founded in 1995
- Flemish Government
- Fast growing
  - 1998: 15 employees
  - 2014: 140 employees (geography, engineering, ICT, …)
- 2015-2016: Transition
  - Flemish Information Agency
Mission FGIA

- “Enabling an optimal application of geographical information in Flanders”
  - Leading provider of geodata
    - All data = open data (exception: Mobile mapping)
Mission FGIA

- IT service provider
- Geopunt, FLEPOS, KLIP, GIPOD, …

http://www.geopunt.be/
FGIA assignments

- Legal basis
  - Several decrees (Flemish Parliament)
  - INSPIRE (European Commission)

- Co-financed
  - Public private partnerships
Agenda

- FGIA
- LRD
- Quality control of LRD data
- MIRO
- Future
LRD concept

- **Large**
  - Scale 1/250 – 1/5000

- **Reference**
  - Uniform data for Flanders
  - Area-wide (since 2013)
  - Basis for link with external data

- **Data**
  - Central database
  - Object oriented
Mandatory use

- LRD is an official Flemish **authentic** data source
- LRD decree sets **mandatory use** and **notification obligation** for public bodies
- **Guarantee** for utility sector
Timeline – production phase

- 2013: LRD area-wide

International workshop on spatial data and map quality
Multi platform data collection
LRD objects

- 4,300,000 buildings
- 4,700,000 parcels
- 64,000 km transport network
- 1,000,000 sewer lids
- 1,300,000 poles
- 40,000 km curbstone
- 20,000 km canals
- 11,000 bus shelters
Updating LRD

- Since 2006
- Different update processes
  - Road + 20m: every 6-9 months update for each city/village
    - Terrain survey
  - Backsides: every 3 years
    - Photogrammetry
  - Parcels: yearly
    - CADMAP (Cadastre)
  - As built maps: continuous
- Based on abnormalities
  - Notifications from users
  - Notifications collected by FGIA
LRD updates - statistics

Amount of new buildings, demolitions and rebuildings

km renewed roads, new roads

Adaptation parcels
Agenda

- FGIA
- LRD
- Quality control of LRD data
- MIRO
- Future
Quality control? What?

- Production of LRD data is outsourced to private data providers
- Extensive documentation set
  - Specifications
  - Quality control specifications
  - Examples
  - …
- Deliveries are subject to quality control by FGIA
Quality control? What?

- Check if data delivered by contractors is conform data specifications
- Control time:
  - Receipt control: 24h
  - Full control: 4 weeks
- Result = Quality report with quality results and decision for contractor
  - Delivery approved
  - Delivery disapproved
Quality control - statistics

Quality controls LRD - production phase

- Total # Deliveries
- Approval
- Disapproval
- Inacceptable

deliveries

Quality control - statistics

Quality controls LRD - Update processes

- Total # deliveries
- Total approval
- Total disapproval
- Total # deliveries TU
- Approval TU
- Disapproval TU
- Total # deliveries MEL
- Approval MEL
- Disapproval MEL
- Total # deliveries BIN
- Approval BIN
- Disapproval BIN
- Total # deliveries KAD
- Approval KAD
- Disapproval KAD
- Total # deliveries ADP-update
- Approval ADP-update
- Disapproval ADP-update
Quality control - process
Quality control - process

- Receipt control
  - Checks the incoming data on the required file structure
  - If a delivery is not receptive, it will be rejected
    - Examples:
      - Format is not correct
      - File is not readable
      - Name of the file is not correct
      - ...
    - There will be no further quality controls
    - Contractor has to deliver a new file
Quality control - process

GIS control
- 100% digital controls
  - Unacceptable errors
    - Example: Double vertices, intersects, …
    - QC can be done
  - Integration
    - Control integration of the data with the surrounding data
  - Topology tests
- Sample controls (based on ISO 2859-1)
  - Visual screen controls
Quality control - process

- TOPO control
  - Terrain controls important for:
    - Positional accuracy
    - Completeness
    - Attribute correctness

→ Sampling based on ISO 2859-1
Quality control - process

- QC Validation
  - Bringing together errors from GIS and TOPO control
  - Evaluate results
    - Approval: All the quality results are conform the predefined AQL values
    - Disapproval:
      - To many “unacceptable errors”
      - Systematic errors
      - At least one test exceeds the acceptance level
  - Making a quality report + error shape
Rapport GRB22_24_01_01

1 Ontvangstnaarzicht
   OK

2 Niet toegelaten fouten
   OK

3 Kwaliteitstypes

3.1 POS

3.1.1 Positionele nauwkeurigheid
   Totaal aantal gecontroleerde items: 16
   Aantal foutieve items: 0
   Aanvaardingsgetal: 0

3.2 VT

3.2.1 Volledigheid/Typologie
   Totaal aantal gecontroleerde items: 32
   Aantal foutieve items: 1
   Aanvaardingsgetal: 2

Fouten

<table>
<thead>
<tr>
<th>ID</th>
<th>Entiteit</th>
<th>IDNR 1</th>
<th>TPC</th>
<th>Entiteit2</th>
<th>IDNR 2</th>
<th>Omschrijving test</th>
<th>Verduidelijking</th>
</tr>
</thead>
<tbody>
<tr>
<td>166175</td>
<td>wez</td>
<td>30029052</td>
<td></td>
<td></td>
<td></td>
<td>Foutieve geometrie, geometrie van het exemplaar is niet correct</td>
<td></td>
</tr>
</tbody>
</table>

Bepalingen

<table>
<thead>
<tr>
<th>ID</th>
<th>Entiteit</th>
<th>IDNR 1</th>
<th>TPC</th>
<th>Entiteit2</th>
<th>IDNR 2</th>
<th>Omschrijving test</th>
<th>Verduidelijking</th>
</tr>
</thead>
<tbody>
<tr>
<td>166175</td>
<td>gra</td>
<td>30029016</td>
<td></td>
<td></td>
<td></td>
<td>Foutieve geometrie, geometrie van het exemplaar is niet correct</td>
<td></td>
</tr>
</tbody>
</table>

Illustatie voor #166175

3.4 GRB

3.4.1 Ligging
   Totaal aantal gecontroleerde items: 1136
   Aantal foutieve items: 0
   Aanvaardingsgetal: 114

3.4.2 Codering
   Totaal aantal gecontroleerde items: 1130
   Aantal foutieve items: 0
   Aanvaardingsgetal: 57

3.4.3 Juistheid
   Totaal aantal gecontroleerde items: 1130
   Aantal foutieve items: 2
   Aanvaardingsgetal: 57

Fouten

<table>
<thead>
<tr>
<th>ID</th>
<th>Entiteit</th>
<th>IDNR 1</th>
<th>TPC</th>
<th>Entiteit2</th>
<th>IDNR 2</th>
<th>Omschrijving test</th>
<th>Verduidelijking</th>
</tr>
</thead>
<tbody>
<tr>
<td>166100</td>
<td>wze</td>
<td>40029051</td>
<td>103</td>
<td></td>
<td></td>
<td>attributewaarde is niet correct: tpo (mee)</td>
<td></td>
</tr>
<tr>
<td>165101</td>
<td>alin</td>
<td>30025023</td>
<td>2</td>
<td></td>
<td></td>
<td>attributewaarde is niet correct: tpo (win)</td>
<td>ontsierende knikpuntzone</td>
</tr>
</tbody>
</table>

3.4.4 Grafische consistentie 100%
   Totaal aantal gecontroleerde items: 1130
   Aantal foutieve items: 0
   Aanvaardingsgetal: 1

3.4.5 Grafische consistentie steekproef
   Totaal aantal gecontroleerde items: 13
   Aantal foutieve items: 0
   Aanvaardingsgetal: 0

3.4.6 Integratiecontrole
   Totaal aantal gecontroleerde items: 6
   Aantal foutieve items: 0
   Aanvaardingsgetal: 0

4 Bepalingen
   geen

5 Beoordeling
   Goedgekeurd
Agenda

- FGIA
- LRD
- Quality control of LRD data
- MIRO
- Future
QC system: MIRO

Design principles:
• Reduce manual work
• Support controllers during the complete quality process
• Central database where all specifications and tests are defined
QC system: MIRO

Additional process

• QC configuration
  • Definition of tests
  • Definition of quality criteria
MIRO architecture and technology

- Technology
  - .NET
  - Net Topology Suite
  - SharpMap
MIRO database: simplified data model

Delivery

Control

Error

Specification

Quality Class

Specific test

Data model

Basic test
MIRO suite

- Interface for MIRO database
- Quality control configuration
  - Configuration of specification
    - Data model
    - List tests
    - Definition of quality criteria
  - Registration of delivery
Specification

- Type
- Code
- Description
Quality Class

- 100% or sampling
- AQL
- Composition of lot
Basic test – specific test
Basic test – specific test

- 70 basic tests → define specific tests with parameters
- Parameters
  - Features
  - Attribute restrictions
  - Topological restrictions
  - Intersection matrix (DE-9IM)
Basic test – specific test: example

- Example “Each road segment must coincide the road area”

  - Basic test = GRBtestAll
    - Calculate intersection matrix for each pair of features
    - If Calculated IM <> Defined IM
      - Add error

- Parameters
  - Entity 1 = road segment
  - Entity 2 = road area
  - Intersection matrix= **F** **F***

<table>
<thead>
<tr>
<th>Road segment</th>
<th>Road area</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Interior</td>
</tr>
<tr>
<td>Interior</td>
<td>*</td>
</tr>
<tr>
<td>Boundary</td>
<td>*</td>
</tr>
<tr>
<td>Exterior</td>
<td>*</td>
</tr>
</tbody>
</table>
Error

- Geometry
- Status
  - Evaluated
  - No error
  - ...
- Description
- Explanation
- Feature ID

January 21st 2015
International workshop on spatial data and map quality
Error: processes

- Evaluate error only once
  - Inheritance of status
- Error shape in case of redelivery
  - Contractor reports which errors were solved and why errors were not solved
  - Automatic join with remaining errors
- Communication between GIS and TOPO
Receipt control

- Preparation
  - Unzip delivery
  - Create folders
  - Load tests from MIRO database
- Format consistency evaluation
  - Run tests
  - Evaluate tests
- Load results to MIRO database
GIS control

Functionality
- Map interface
- Calculate statistics (lot size, sample size, acceptance number)
- File compare
- Evaluation of logical consistency
  - Automatic topological tests + evaluation output
    - Aim: evaluate error only once (in case of redelivery, update,…)
  - Manual tests
    - Facilitates sampling of the data (ISO 2859-1)
    - Enables the registration of errors

Process
- Load control information from MIRO database
- Load delivery from file server
- Perform control
- Load results to MIRO database
GIS control: interface
GIS control: sampling

- **Statistics**
  - [Image: Table displaying quality classes, AQL, Goodness criteria, Inspection level, Lot size, Sample size, Acceptance number]

- **Composition of sample**
  - Drawing area
  - Random selection

International workshop on spatial data and map quality

January 21st 2015
TOPO control

Quality control configuration

Delivery
- Receipt control
- GIS control
- TOPO control
- QC validation

Reporting

Preparation
- Calculating sample size
- Defining measurement areas

Field measurements

Processing
- Compare datasets
- Locate errors
- Load results
TOPO control: interface
TOPO control: preparation
TOPO control: processing
MIRO suite

- Management environment
- QC validation
  - Bringing together errors GIS and TOPO
  - Decision approval or not
- Reporting
  - Quality report
  - In case of disapproval: error shape
  - In case of approval: shape with remaining errors
    - Solve at FGIA
    - Upload to database and solve in next update
- QC configuration
Agenda

- FGIA
- LRD
- Quality control of LRD data
- MIRO
- Future
AS IS <> TO BE

- FGIA has different quality control systems for different vector datasets
  - MIRO → LRD
  - Other datasets have their own systems
- No longer feasible for FGIA to maintain all these different environments
  - High cost for maintenance
  - Centralize knowledge quality control
  - More efficient controls

→ 2014: start project “Quality control system for all FGIA vector data”
Main goal project

Guarantee the quality of FGIA vector data in a cost efficient way

LOWER COSTS FOR QUALITY CONTROL

+ 

GUARANTEE QUALITY DATA
Roadmap

- Project with AGILE approach
- 2014: Inventory AS IS situation
  - Interviews with product owners
  - Collecting documentation
  → Result = Report
- December 2014
  - Workshop with product owners
  - Collecting needs
- Begin 2015
  - Making product backlog
  - Prioritise
- 2015 - …
  - Starting development
Lies De Meulenaer
lies.demeulenaer@agiv.be

Barbara Bulkens
barbara.bulkens@agiv.be