



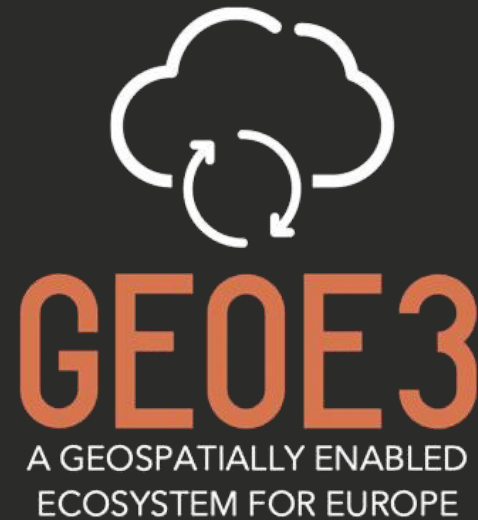
Co-financed by the Connecting Europe  
Facility of the European Union

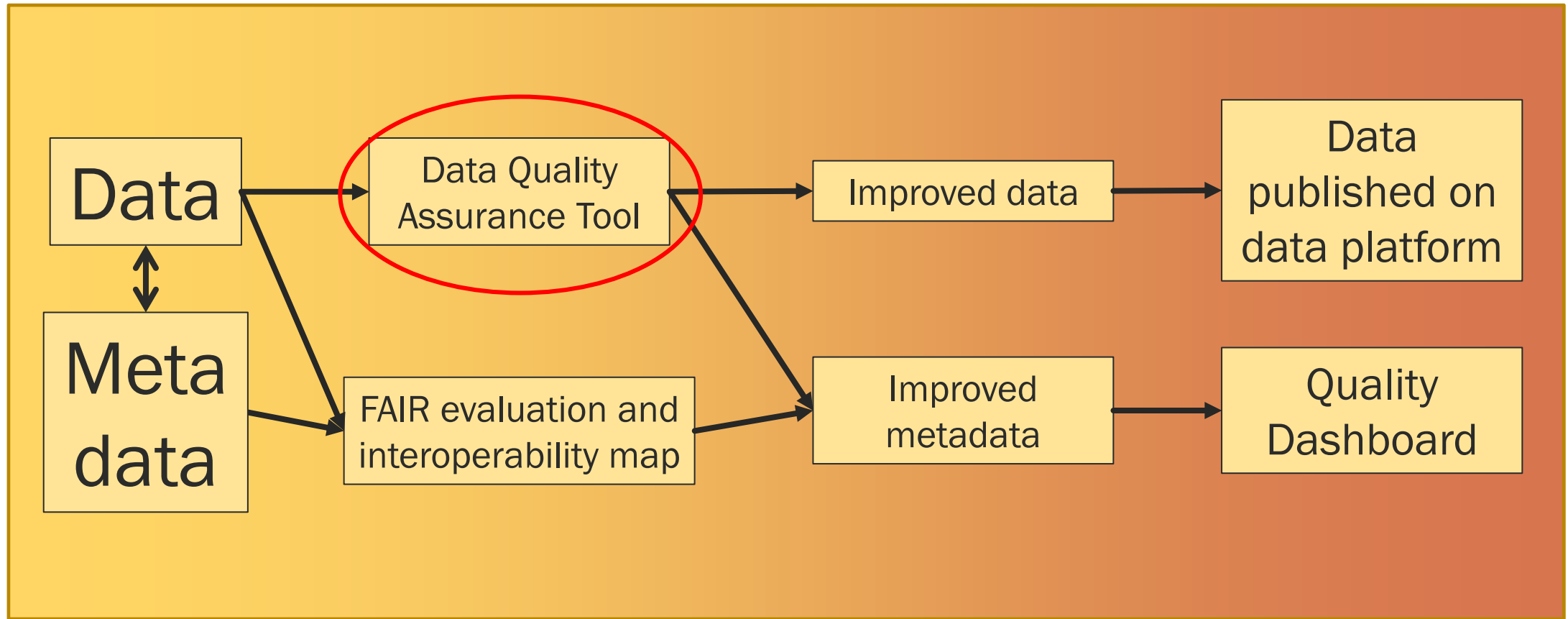
# ROLE OF QUALITY ASSURANCE AND DASHBOARDS IN DATA ECOSYSTEMS

Alpo Turunen

Spatial data specialist

National Land Survey of Finland





Data producer

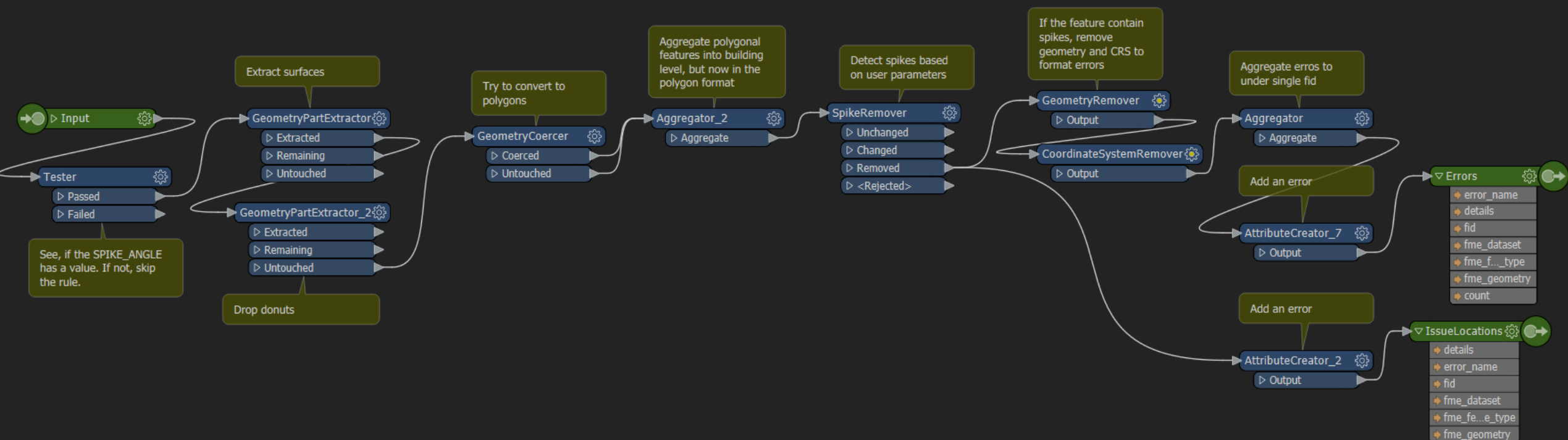


Data user

# DATA QUALITY ASSURANCE TOOL

FME-based quality assurance software:

- Takes 3D building data set (CityGML / CityJSON) as an input
- As an output, produces a quality report and fixes some errors automatically
  - Results can be written either to CSV, CityGML or CityJSON file.



# QUALITY RULES

<https://github.com/opengeospatial/GEOE3>

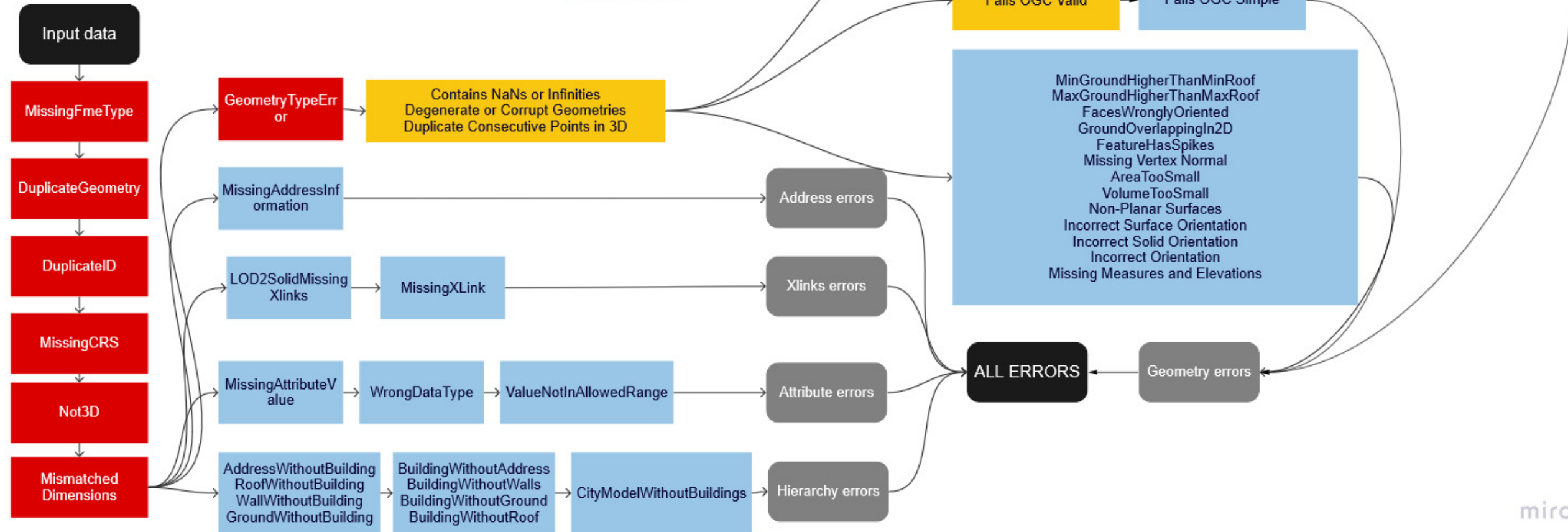


- Over 60 quality rules were created for 3D buildings
- Mostly geometrical checks
- Based on the standards (e.g. CityGML standard), recommendations (SIG3D Modelling Guide), FME transformers, or our previous experience
- All of them can find errors automatically, and some of them can fix errors

B	C	D	E	F	G	H
RULE NAME	RULE ID	QUALITY ELEMENT	FME IMPLEMENTATION	IMPLEMENTATION STATUS	SEVERITY LEVEL	DESCRIPTIVE QUESTION
MissingParentValue	ValidateAttributes6	Omission	ValidateAttributes -> Tester	Implemented	NOTE	Does every 2nd-level cityobject has parent value?
Contains NaNs or Infinities	ValidateGeometry2	Omission	ValidateGeometry -> RemoveInvalidGeometries	Implemented	FIXED	Does the geometry contain NaNs or infinities?
Degenerate or Corrupt Geometries	ValidateGeometry3	Conceptual Consistency	ValidateGeometry -> RemoveInvalidGeometries	Implemented	FIXED	Does the geometry contain degenerated or corrupted geometries?
Non-Planar Surfaces	ValidateGeometry5	Conceptual Consistency	ValidateGeometry -> CheckAllOtherProblems	Implemented	NOTE	Is the Face or BoundarySurface planar based on thickness or normal deviation?
Self-Intersections in 2D	ValidateGeometry6	Conceptual Consistency	ValidateGeometry -> CheckSelfIntersectionsIn2D	Implemented	FAIL	Does the feature intersect itself in 2D?
Duplicate Consecutive Points in 3D	ValidateGeometry7	Commission	GeometryValidator -> RemoveInvalidGeometries	Implemented	FIXED	Does the feature contain duplicate consecutive points?
Missing Vertex Normals	ValidateGeometry8	Conceptual Consistency	GeometryValidator -> CheckNormals	Implemented	FIXED	Does the feature has vertex normals?
Incorrect Surface Orientation	ValidateGeometry9	Conceptual Consistency	GeometryValidator -> CheckAllOtherProblems	Implemented	NOTE	Does the surface has an correct orientation?
Incorrect Solid Orientation	ValidateGeometry10	Conceptual Consistency	GeometryValidator -> CheckAllOtherProblems	Implemented	NOTE	Does the solid has an correct orientation?
Incorrect Orientation	ValidateGeometry11	Conceptual Consistency	GeometryValidator -> CheckAllOtherProblems	Implemented	NOTE	Does Areas, such as polygons, ellipses, and donuts, have an correct orientation?

## Rule hierarchies of the CityGML workspace

Severity levels:  
**Red:** FAIL  
**Orange:** FAIL&FIXED  
**Blue:** NOTE



Visual Preview x

Table

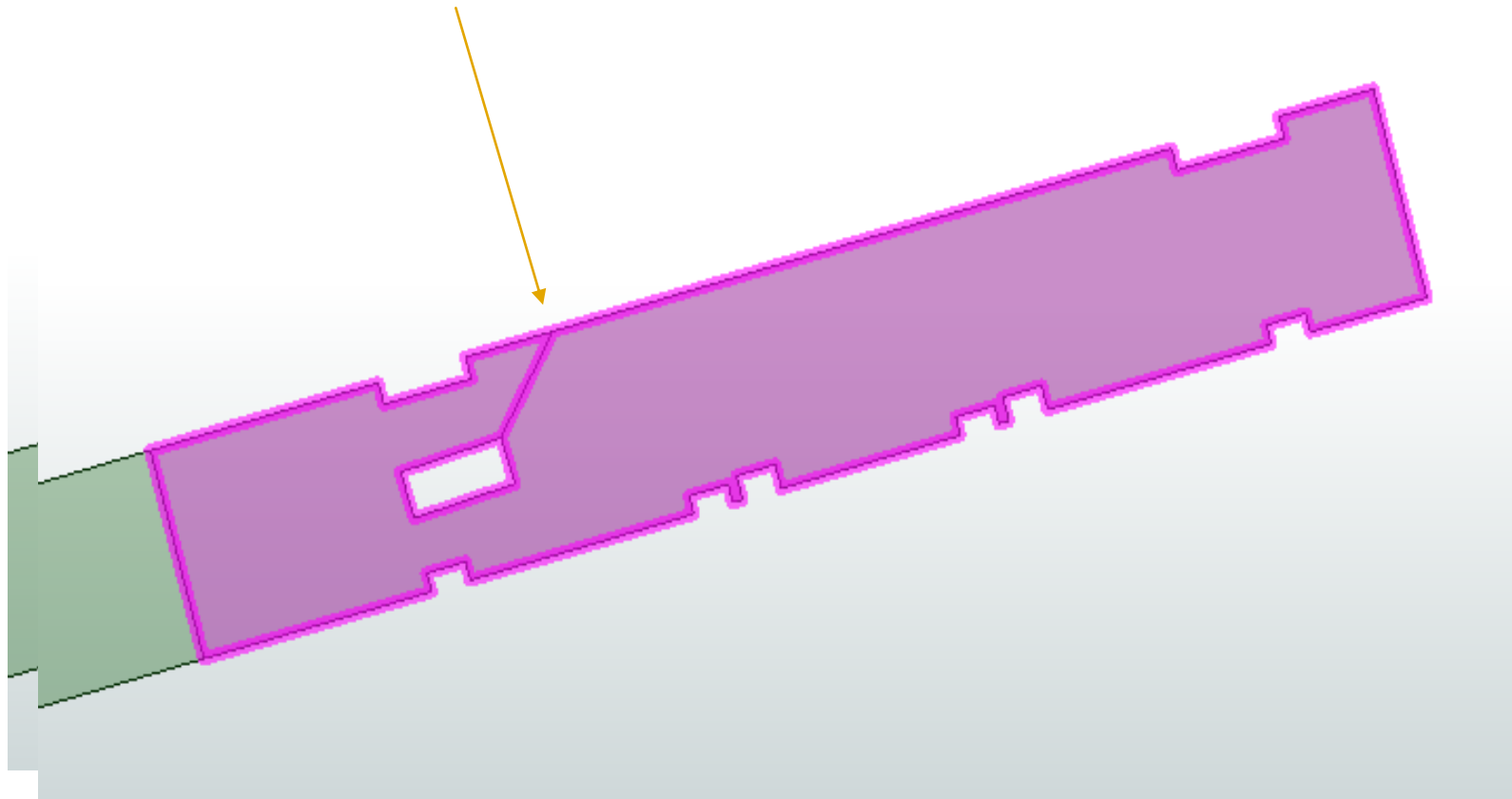
ErrorsPerDataset\_Summary Columns...

	error_name	details	severity_level	fme_dataset	error_name.total_count
1	ValueNotInAllo...	The following v...	NOTE	C:\Users\ATURU...	9
2	WrongDataType	The following a...	NOTE	C:\Users\ATURU...	9
3	MissingAttribut...	Following attrib...	NOTE	C:\Users\ATURU...	9
4	NonPlanarByNo...	Triangulated no...	FIXED	C:\Users\ATURU...	5
5	MissingAttribut...	Following attrib...	NOTE	C:\Users\ATURU...	4
6	ValueNotInAllo...	The following v...	NOTE	C:\Users\ATURU...	3
7	WrongDataType	The following a...	NOTE	C:\Users\ATURU...	3
8	WrongDataTyne	The following a	NOTE	C:\Users\ATIIRII	1

Search:  in any column 9 row(s)

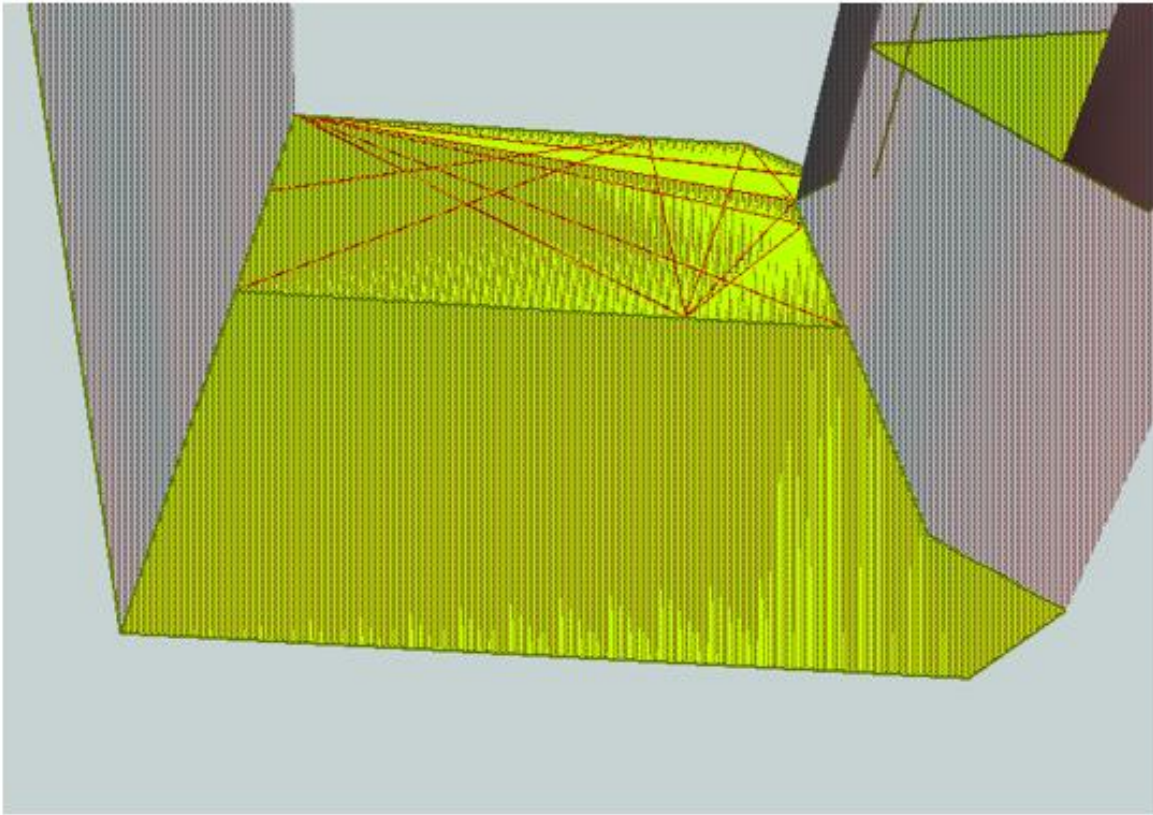
Visual Preview Translation Log

# SELF-INTERSECTION IN 2D

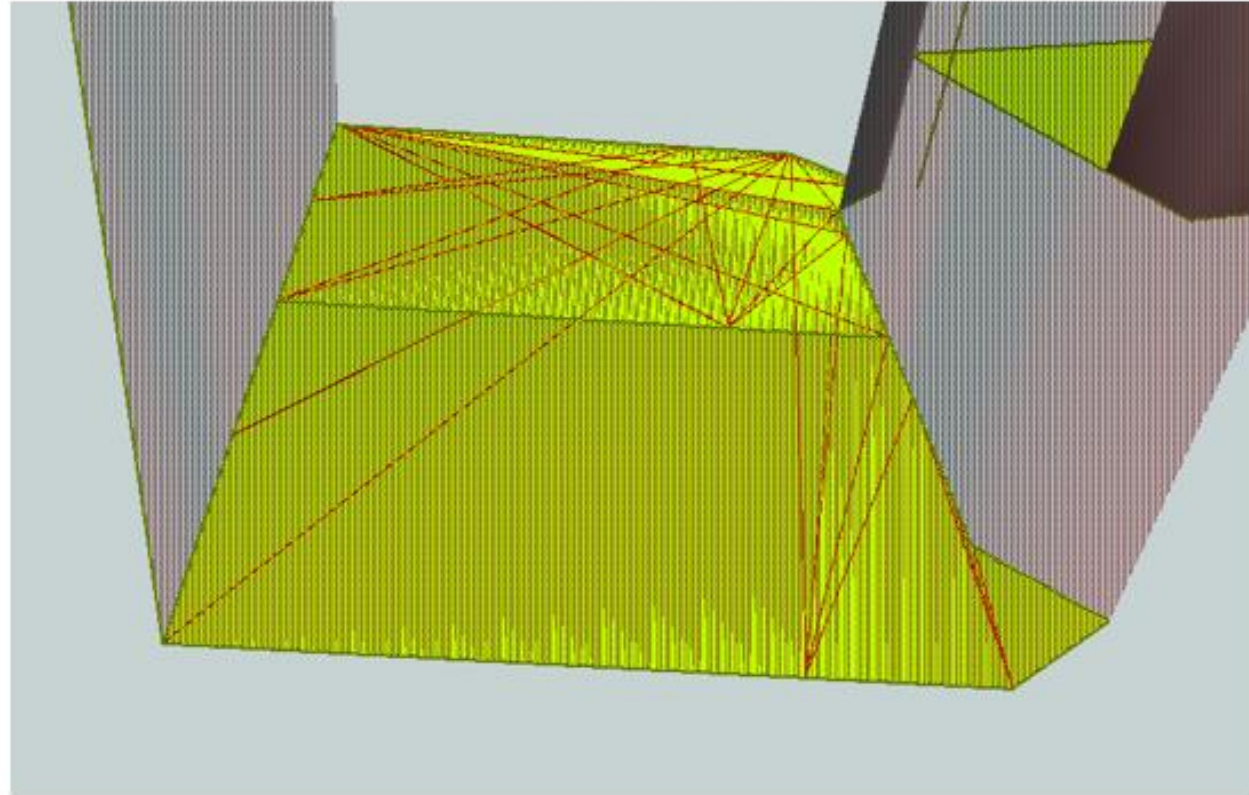


# GROUND OVERLAPPING IN 2D

Area 1:



Area 2



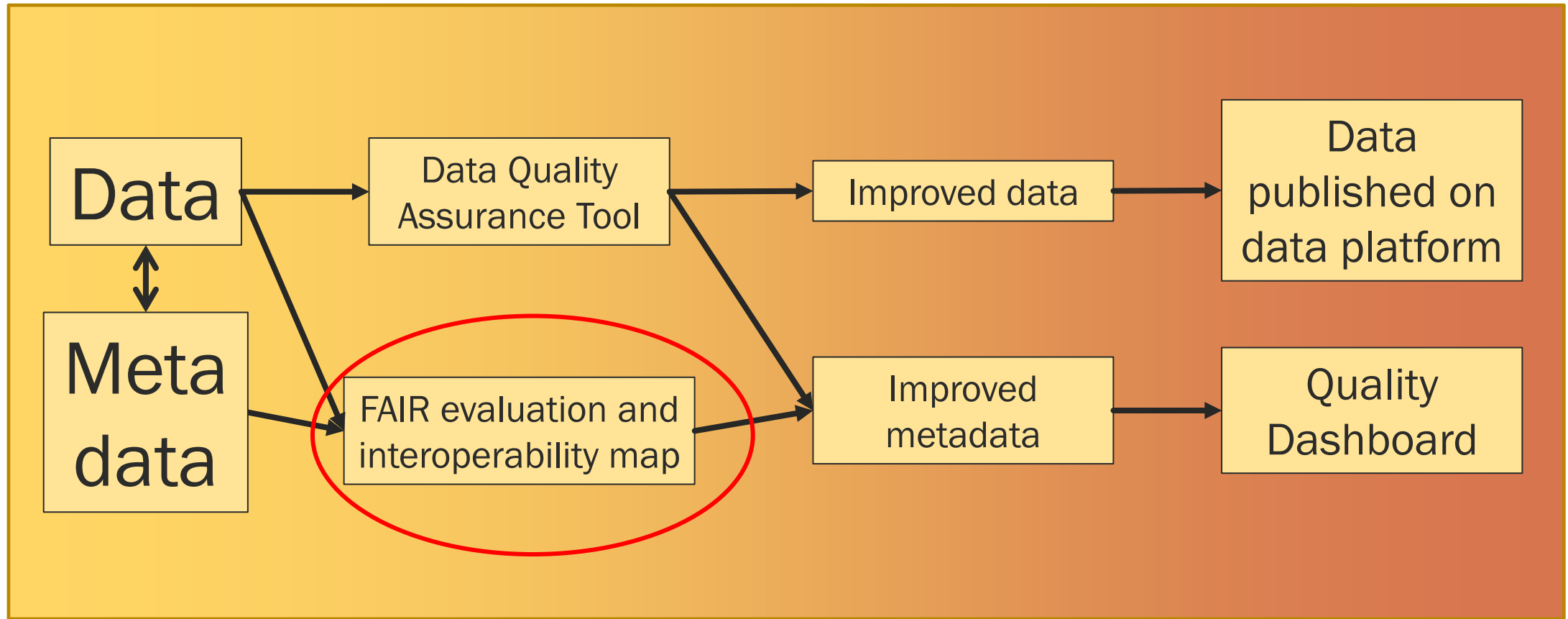


## SEE OUR GITHUB REPO

<https://github.com/opengeospatial/GEOE3>



Added three images	Added three images
Detailed_Error_Descriptions.docx	Updated grammar
GeoE3-Service-Architecture-official...	Service Architecture
QualityRules_CityGML.xlsx	New functionalities
QualityRules_CityJSON.xlsx	New functionalities
QualitySoftwareCityGML.fmw	Bug fixed
QualitySoftwareCityJSON.fmw	Bug fixed
README.md	Update README.md



Data producer



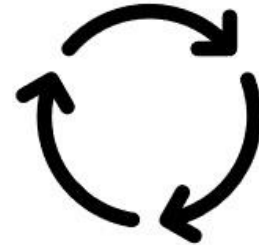
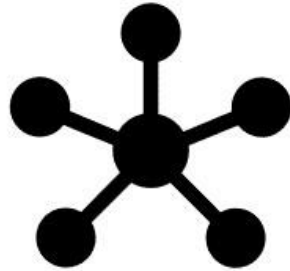
Data user

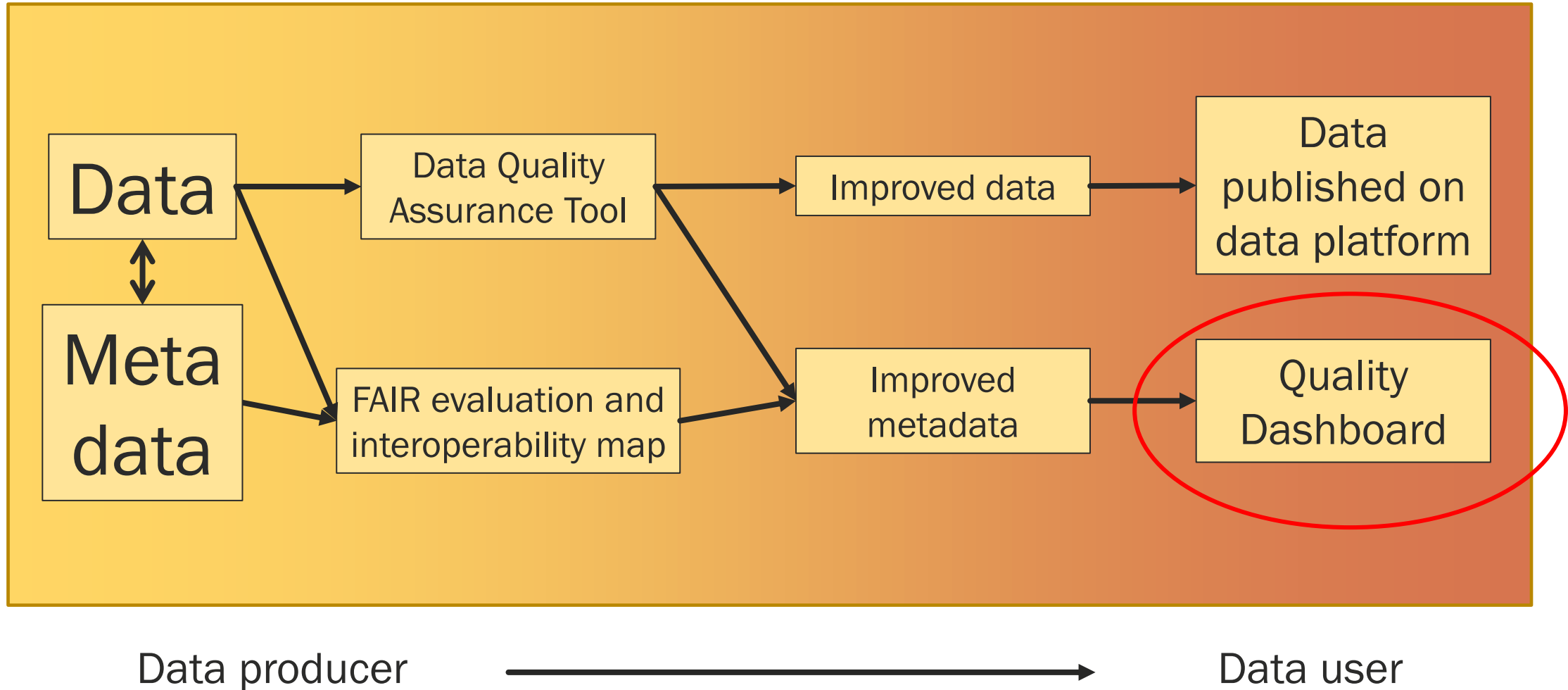
# INTEROPERABILITY MAP

- In practice, it is a maturity matrix (for high-value data sets)
- So it describes interoperability of a certain data set
- We evaluated 2D/3B buildings, DSM, DTM, and road, weather, and climate data sets
- Not automatised

Categories	Level 0: Not interoperable and cannot be integrated	Level 1: minimal interoperability and can be integrated with extra effort	Level 2: Intermediate interoperability and can be integrated mostly automatically	Level 3: Advanced /Optimal interoperability and can be integrated automatically
National data accessibility and integration arrangements	Data cannot be provided due to legal requirements or is not considered as open data	Data accessible through different agencies, no national integration arrangements (data available without restrictions or minimum restrictions as defined in LIFO)	Data available mostly through national platform but some data missing. This could be for example attribute data.	Data available through national platform and data integration arrangements in place
metadata discoverability	No metadata available	Metadata available nationally	Metadata provided through APIs.	Metadata provided through DCAT AP 2.0 or OGC API records .
data accessibility	No data available	Data available with legacy APIs	Data available with OGC APIs.	Data available with OGC APIs.
Vocabulary and data specifications	Vocabulary/ data descriptions not available and cannot be integrated	Vocabulary and data specifications including data content and data quality are described, but not according to any standards. Minimal definitions available and can be integrated with extra effort	Vocabulary and data specifications including data content and data quality are described, but not according to any standards. Intermediate interoperability Partly or full machine readable (MR) but automatic utilization not fully possible	Vocabulary and data specifications are fully machine readable in RDF/OWL. Advanced/Optimal vocabulary/definitions in machine readable format (MR) and can be utilized automatically
Data content and data quality	Data content and data quality are not described and cannot be integrated	Data content and data quality are described, but not according to any standards or in machine readable form.	Data content is sufficient for the expected usage in machine readable form.	Data content and quality are well described in machine readable form (e.g. UML).
Quality assessment	No quality assessment information available	Quality assessment done but not available through metadata	Quality goals defined and available through metadata	Quality assessment available through Data Quality Vocabulary (DQV)

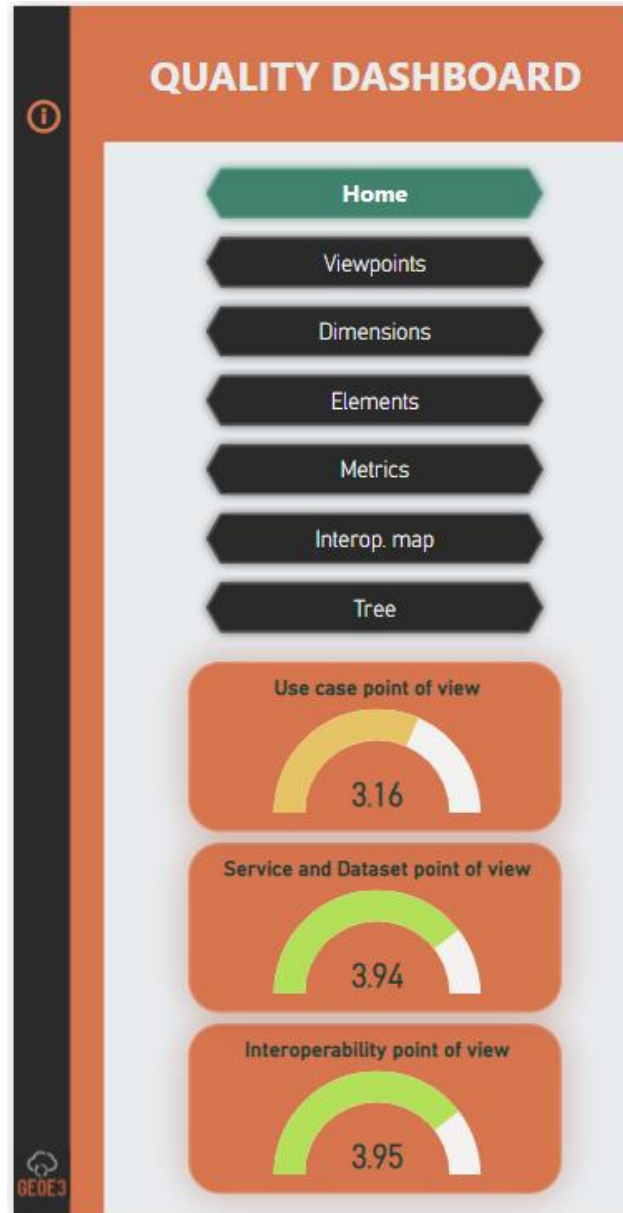
**F**indable **A**ccessible **I**nteroperable **R**eusable






# QUALITY DASHBOARD

- GeoE3 quality dashboard - a method for scoring services and data using metadata and monitoring information.



# DATA SOURCES:

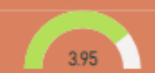
- Quality evaluation results of the FME tool
- Dataset metadata 
- Service metadata (or description capabilities document)
- Interoperability map
- Service availability information by Spatineo

```

1377     <!-- Added : data quality -->
1378     <gmd:dataQuality>
1379         <gmd:Completeness>
1380             <gmd:Commission>
1381                 <gmd:itemDuplicate>0</gmd:itemDuplicate>
1382                 <gmd:itemExtra>45</gmd:itemExtra>
1383                 <gmd:count>1</gmd:count>
1384                 <gmd:rate>0.6</gmd:rate>
1385             </gmd:Commission>
1386             <gmd:Omission>
1387                 <gmd:itemMissing>yes</gmd:itemMissing>
1388                 <gmd:count>2</gmd:count>
1389                 <gmd:rate>1.3</gmd:rate>
1390             </gmd:Omission>
1391         </gmd:Completeness>
1392         <gmd:ThematicQuality>
1393             <gmd:ClassificationCorrectness>
1394                 <gmd:date_evaluation>|
1395                     <gco:date>2021-10-01</gco:date>
1396                 </gmd:date_evaluation>
1397                 <gmd:missclassification>
1398                     <gmd:rate>2.2</gmd:rate>
1399                     <gmd:count>52</gmd:count>
1400                 </gmd:missclassification>
1401             </gmd:ClassificationCorrectness>

```

## Interoperability

Interoperability  


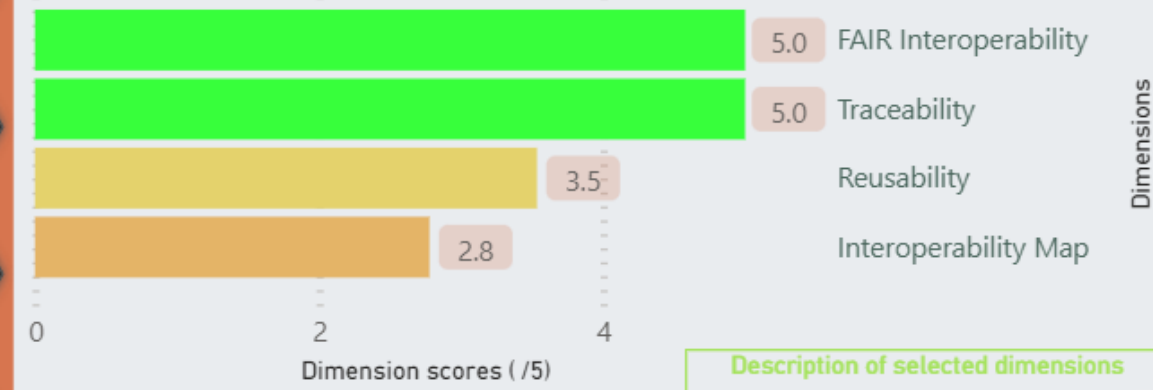
### Viewpoint

Select the viewpoint to examine ↩

- Use Case
- Service And Dataset
- Interoperability

Viewpoint	Viewpoint_Description
Interoperability	Viewpoint focused on the interoperability capacity of the dataset and service. Includes evaluations from the GeoE3 Data Maturity Model and Interoperability Map.

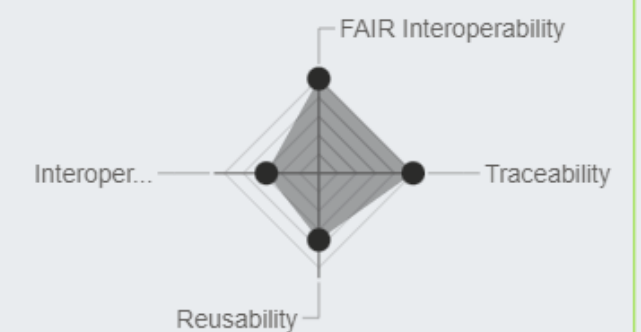
### Dimension Scores



Dimension	Score
FAIR Interoperability	5.0
Traceability	5.0
Reusability	3.5
Interoperability Map	2.8

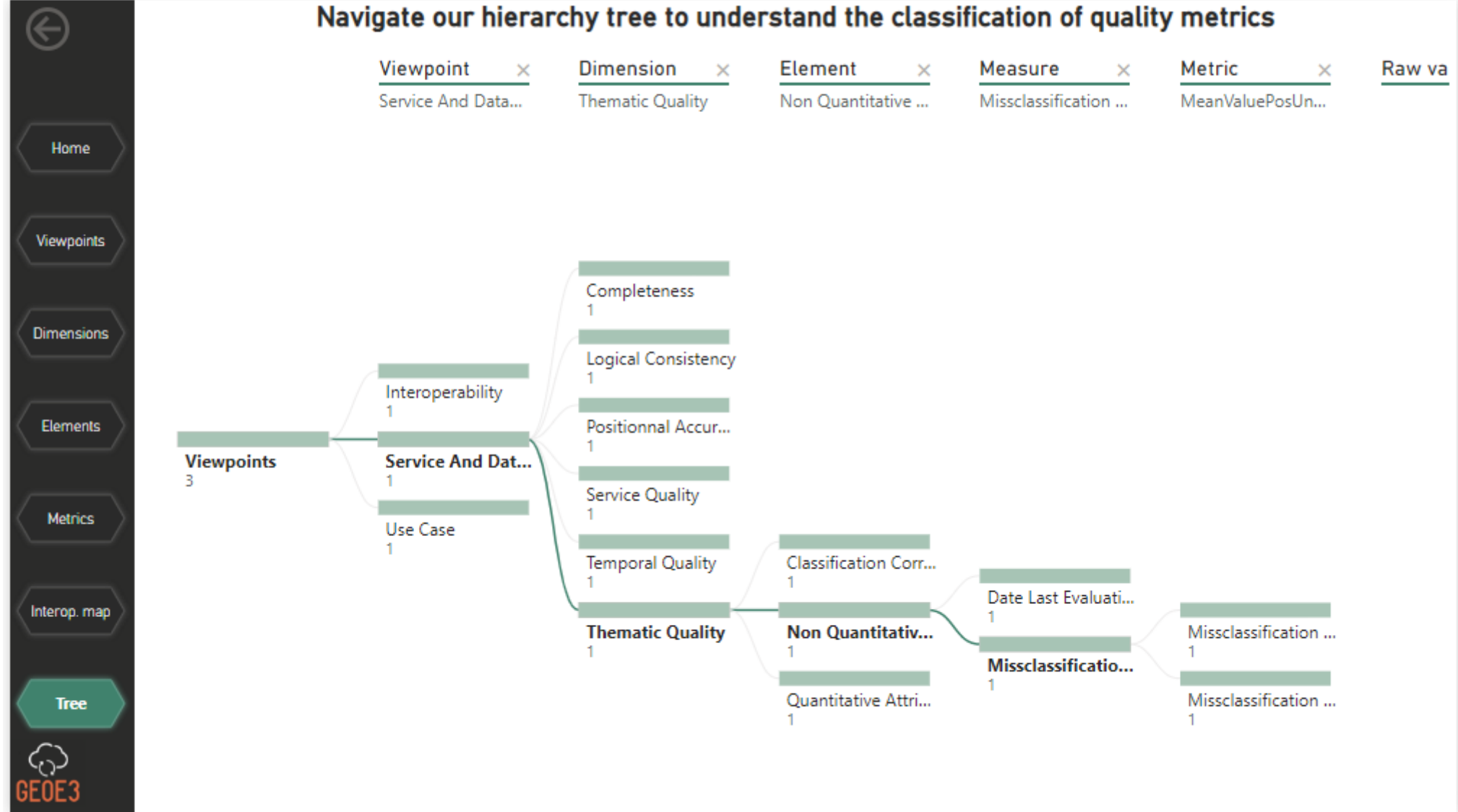
#### Description of selected dimensions

Dimension	Dimension_Description
Interoperability Map	GeoE3 Interoperability map
Traceability	Describes whether the data process is described adequately.
Reusability	Data and metadata are sufficiently annotated so machine and human users can determine

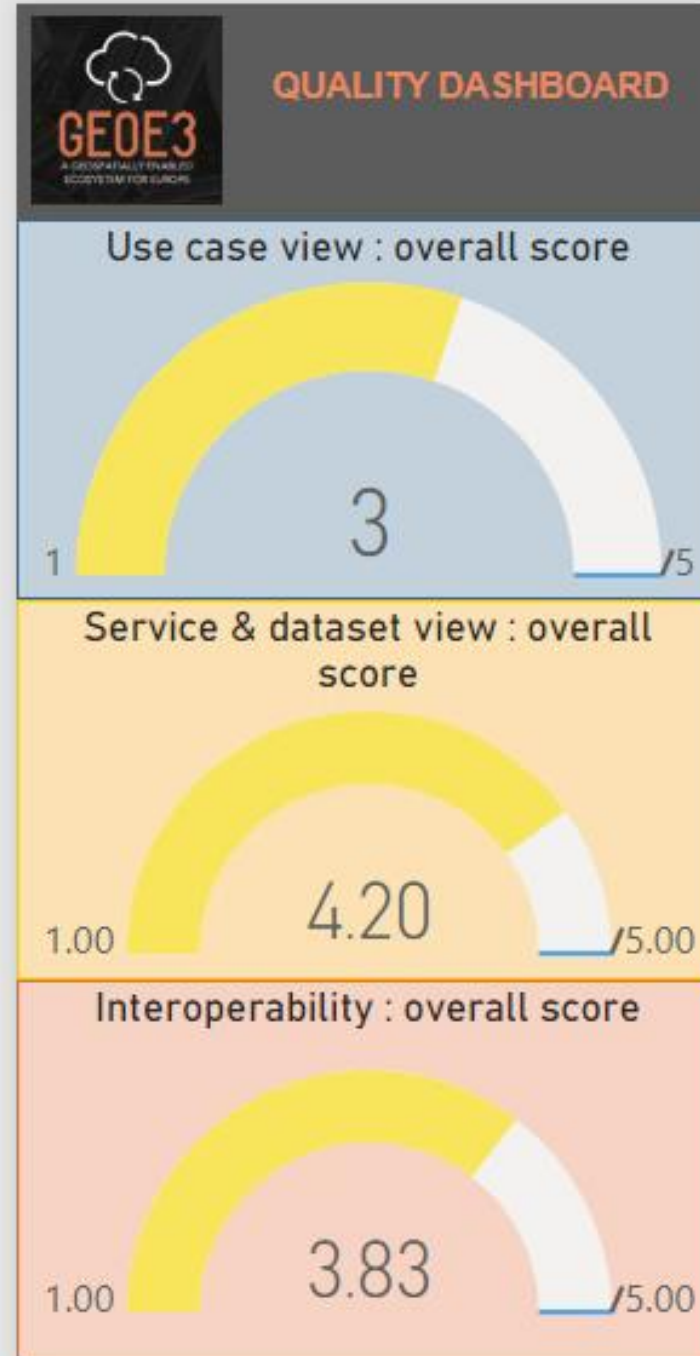


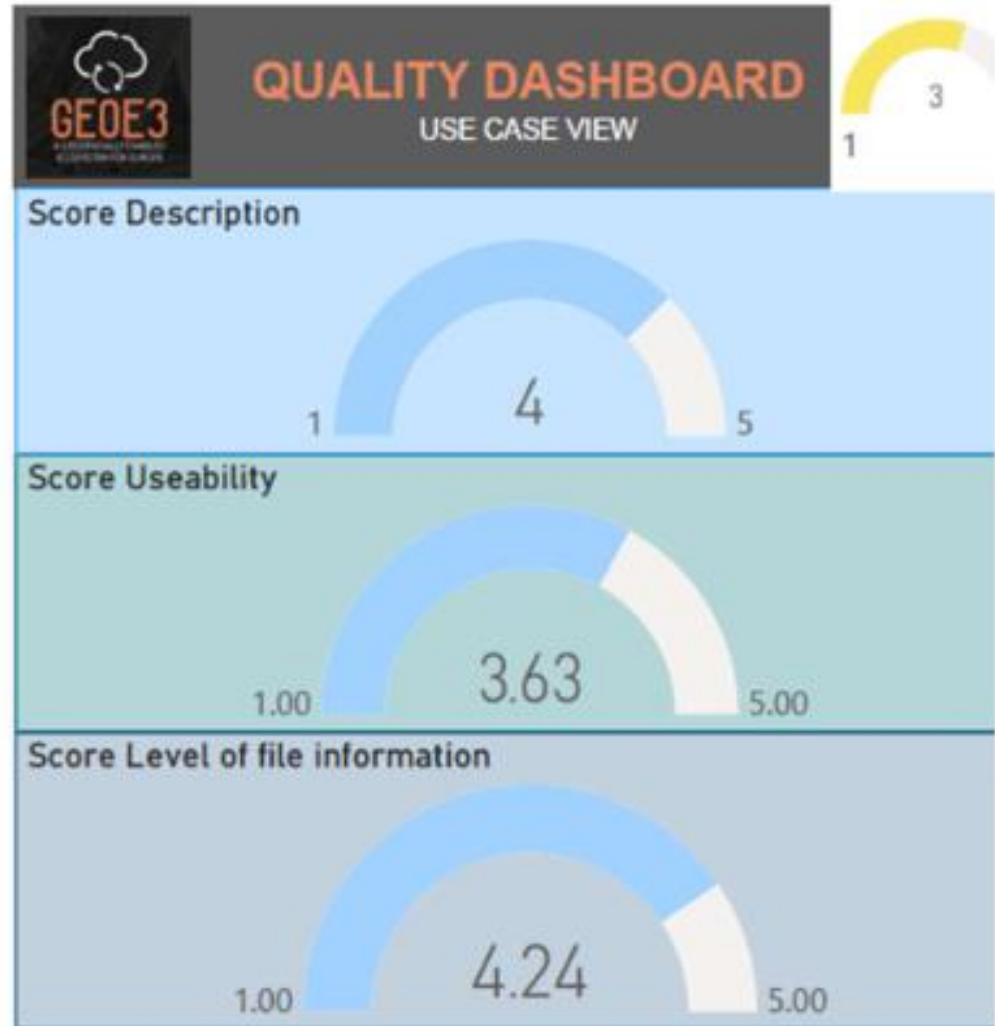


## Navigate our hierarchy tree to understand the classification of quality metrics



- 1. Use Case View:** Designed for novice users, it assesses dataset suitability based on reliability, description clarity, and relevance.
- 2. Data & Service View:** For data providers and integrators, it evaluates data service and quality using ISO-defined metrics like completeness and thematic quality.
- 3. Interoperability View:** Focuses on technical and semantic interoperability, legal and organizational aspects, portability, and information security, providing insights into dataset and service compatibility.





*Figure 4-2: Current version of the quality dashboard: user point of view*



Figure 4-3: Current version of the quality dashboard: service & dataset point of view

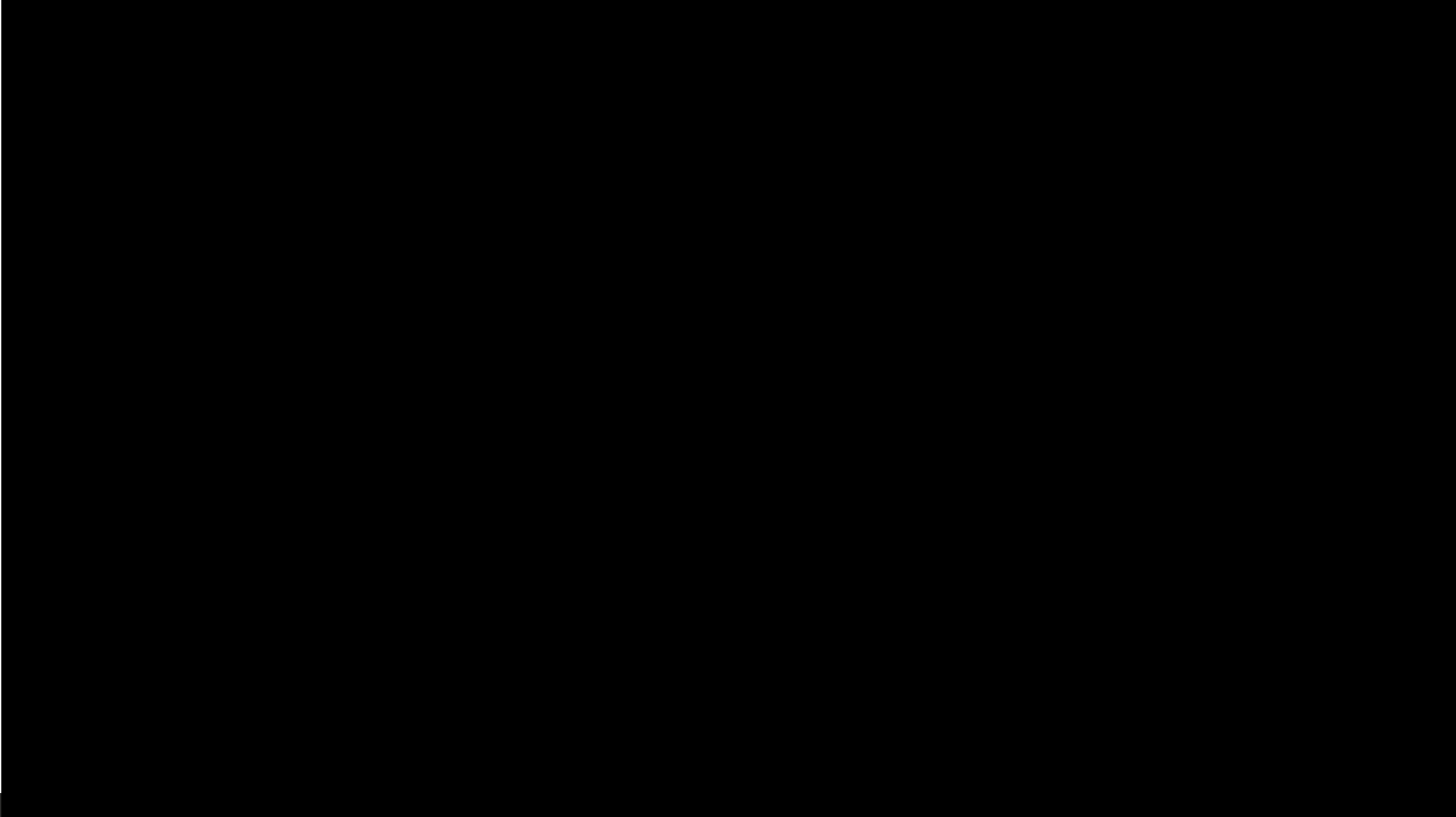


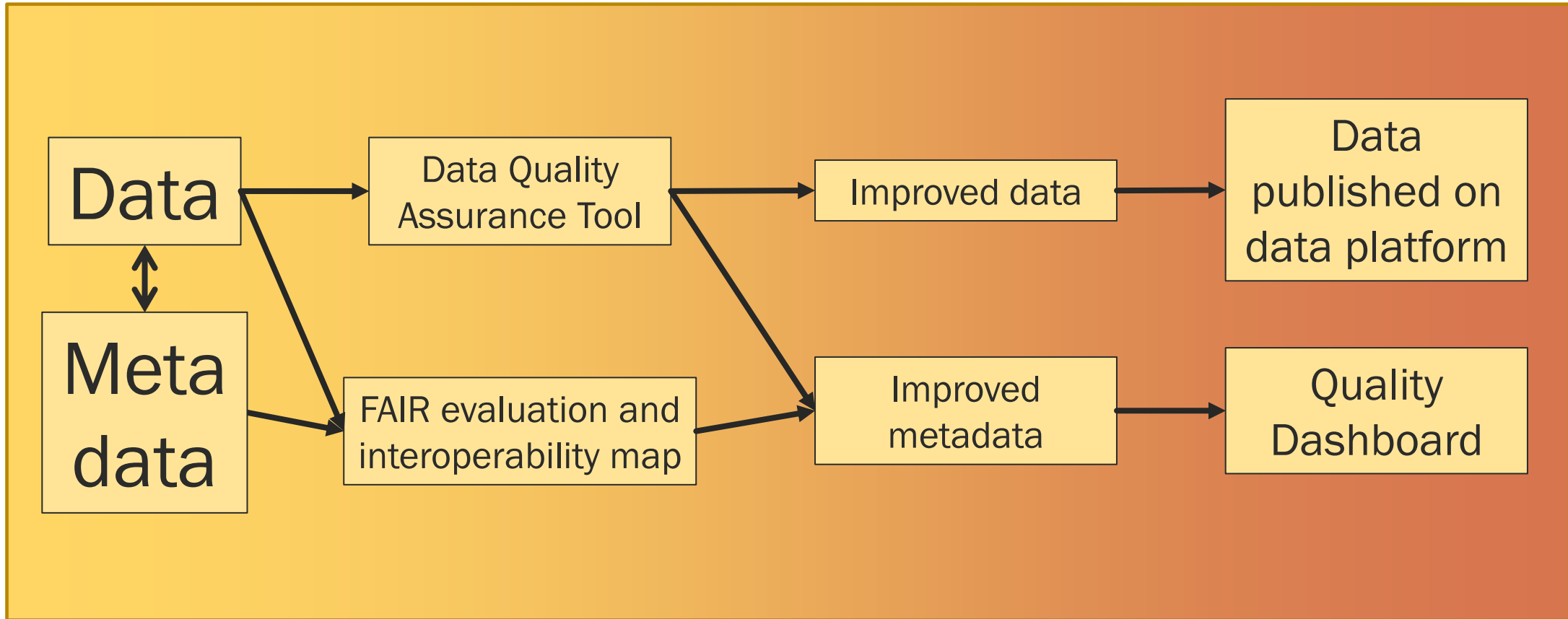
Figure 4-4: Current version of the quality dashboard: Interoperability point of view (interoperability map is to be integrated in the future)

1	Type of	Dimension	Element	Measure	Metric	Additions	User input	Extraction rule	
39		Thematic quality	Classification		Rate	A	Yes	/gmd:MD_M	
40				Date of last evaluation	Date of last evaluation	A	Yes	/gmd:MD_M	
41				Correctness	Missclassification rate	Missclassification rate	A	Yes	/gmd:MD_M
42			Non Quantitative Attribute		Missclassification	Number of incorrectly classified items	A	Yes	/gmd:MD_M
43				Date of last evaluation	Date of last evaluation	A	Yes	/gmd:MD_M	
44				Missclassification rate	Rate of incorrect attribute values	A	Yes	/gmd:MD_M	
45		Correctness	Missclassification	Number of incorrect attribute values	A	Yes	/gmd:MD_M		
46		Quantitative	Date of last evaluation	Matches last update or Pass / fail	A	Yes	/gmd:MD_M		

	Source of value for	Evaluation rule	Weight	Wei	Wei	Wei	Notes	Description/	Standards / Sources of	Id for ISO 19157
1	Dataset metadata	comparison <	7						ISO 19157-3	3
1	Dataset metadata	date	5	8					ISO 19157	
1	Dataset metadata	comparison <	7	5	7				ISO 19157-3 table D. 71	61
1	Dataset metadata	comparison <	1							ISO 19157-3 table D. 70
1	Dataset metadata	date	5	8					ISO 19157	
1	Dataset metadata	comparison <	7	5	7	5			ISO 19157-3 table D. 77	67
1	Dataset metadata	comparison <	1							
1	Dataset metadata	date	5	8					ISO 19157	

# DASHBOARD VIDEO





Data producer



Data user







## THANK YOU! QUESTIONS, COMMENTS?

<https://geoe3.eu/>

<https://www.linkedin.com/in/alpoturunen>

<https://www.github.com/opengeospatial/geoe3>

[Alpo.turunen@nls.fi](mailto:Alpo.turunen@nls.fi)

Cossec, Camille. "Creation of quality dashboard for geospatial data and services." (2022).

Turunen, Alpo. "Data Quality Assurance of 3D Building Features in Data Integration Processes." (2022).

