



Quality Frameworks for NAP Services and Data

Mobility Data Days

November 6th, 2024



What is Quality?

~~Quality Frameworks for NAP Services and Data~~

Mobility Data Days

November 6th, 2024

Quality Frameworks: Agenda

- Welcome, background and goals *Peter Lubrich, BAST (DE)*
- A Quality Framework for UVAR Data *João Montenegro, ARMIS (PT)*
- A Quality Framework for Floating Car Data *Chrysostomos Mylonas/
Maria Stavara, CERTH (GR)*
- A Quality Framework for MMTIS Data *Anastasia Founta, ITxPT (BE)*
- A Methodology for assessing Cross-Domain Quality (CDQ) *Petr Bureš, TTR (CZ)*
- Future prospects *Peter Lubrich, BAST (DE)*
- Q&A with audience

How is Quality documented on mobility data portals?

Search for
„Qualität“
(Quality)

Search Results 569

Search: Topicality:

Map

Category

- ☐ Waterways and water bodies 440
- ☐ Roads 68
- ☐ Climate and weather 28
- ☐ Other 14
- ☐ Railway 4

Data Provider

- ☐ Bundesamt für Seeschifffahrt und Hydrographie (BSH) 432
- ☐ Umweltbundesamt 17
- ☐ Kraftfahrt-Bundesamt (KBA) 16
- ☐ RVR 11
- ☐ Genossenschaft 10

Type of Terms of Use

INFRASense: Mittels KI (Objekterkennung) detektierte Radweg- und Straßenschäden an Asphaltstrecken aus Oldenburg, Osnabrück, Köln und weiteren Städten

Offered by: Universität Oldenburg Visibility: Public

Der Datensatz umfasst detektierte Radweg- und Straßenschäden an asphaltierten Strecken, die im Rahmen des Forschungsprojektes INFRASense (FKZ: 19F2186E) erhoben wurden. Im Projekt wurde eine shapebasierte Bewertung für Radverkehrsanlagen entwickelt. Dabei wird das Streckennetz in Einzelschnitt...

CREATED	DATA MODEL	TYPE OF TERMS OF USE
30.10.2024	—	Licence, free use/open data

BROKERING TYPE	GEOGRAPHY	CATEGORY
Not brokered	See map	Cycle network

LSA--MAPEM Datensammlung-- Köln--C-ITS

Offered by: Stadt Köln Visibility: Public

Die MAPEM (MAP extended Message) ist ein ETSI ITS-Nachrichtentyp mit standardkonformen Header und MAP als relevantem Payload. Die MAP greift die Lücken der kommerziellen digitalen Karte im Fahrzeug auf und füllt sie mit Informationen der Lichtsignalanlage. Alle Details der Kreuzung werden gespe...

CREATED	DATA MODEL	TYPE OF TERMS OF USE
08.10.2024	Other	Licence, restricted use, free of charge

How is Quality documented on mobility data portals?

Search Results > Offer Details

Offered by: HMDKLV Visibility: Public

Verkehrsdaten Kfz (Infrarotdetektoren) Hamburg

CREATED	DATA MODEL	TYPE OF TERMS OF USE	BROKERING TYPE
08.08.2024	Other	Licence, free use/open data	Not brokered

GEOGRAPHY	CATEGORY
See map	Roads

Offer Details

mobilthek Data Offers Blog The Mobilthek Help Contact DE | EN BD

General Information

Allgemeine Informationen:
Der Datensatz umfasst Verkehrsdaten aller Standorte in Hamburg, an denen der Kraftfahrzeugverkehr (Kfz-Verkehr) mittels Infrarotdetektoren an 24h am Tag und allen Tagen des Jahres erfasst wird.
Die Daten enthalten Verkehrsstärken in Echtzeit und werden an für den Straßenquerschnitt zusammengefassten Zählstellen in 15-Minuten, 60-Minuten, Tages- und Wochen-Intervallen zur Verfügung gestellt.
Die Daten der Zählstellen werden außerdem in den entsprechenden Geoportalen der FHH, z.B. in Geo-Online und dem Verkehrsportal, visualisiert.

Neben den Echtzeitdaten sind auch historische Daten in folgendem Umfang verfügbar:
alle Daten für die letzten zwei Wochen in 15-Minuten-Intervallen, alle Daten für die letzten zwei Monate für die 60-Minutenintervalle, alle Daten für das aktuelle und das letzte Jahr in Tagesintervallen sowie alle Daten seit Beginn der Erfassung in Wochenintervallen.

Informationen zur Technik:
Die Infrarotdetektoren sind in der Regel an Lichtsignalanlagen, zu einem geringen Teil aber auch an anderen Masten, installiert. Die Detektoren erfassen und zählen den Verkehr über die Wärmeabstrahlung der einzelnen Verkehrsteilnehmenden. Da ausschließlich Infrarotbilder ausgewertet werden, ist der Datenschutz zu jeder Zeit gewährleistet.

Hinweise zur Datenqualität:
Die Daten werden in Echtzeit an die Urban Data Platform der FHH übertragen. So sind sie zeitnah für alle Nutzenden und Interessierten verfügbar. Durch die Echtzeitkomponente sind allerdings verschiedene Rahmenbedingungen zu beachten: Die Daten sind nicht umfassend qualitätsgesichert. Ungewöhnliche Abweichungen von den zu erwartenden Daten und Datenlücken werden zwar automatisch vom System erkannt, können aber nicht in Echtzeit korrigiert werden. Lücken, die z.B. durch einen Abriss der Datenübertragung auftreten, können im Nachhinein noch nachgeliefert werden. Unter Umständen und bei längeren Ausfällen können folglich noch nach ein paar Tagen Änderungen in den historischen Daten erfolgen.

„...Die Daten sind nicht umfassend qualitätsgesichert...”

“... The data are not comprehensively quality-assured...”

How is Quality documented on mobility data portals?

mobil|thek Data Offers Blog The Mobil|thek Help Contact DE | EN BD

Search Results > Offer Details

Offered by: Deutscher Wetterdienst Visibility: Public

Allgemeine Wettervorhersagen

DATA MODEL	TYPE OF TERMS OF USE	BROKERING TYPE	GEOGRAPHY
Other	Licence, free use/open data	Not brokered	See map

CATEGORY
Climate and weather

Offer Details

General Data Access Terms of Use

General Information

Description	Aussagen über die meteorologische Situation, insbesondere das zu erwartende Wetter als Ergebnis der Bearbeitung, Interpretation, Anpassung und Ausdeutung des meteorologischen Basismaterials (Wettermeldungen, NWV, Satelliten-, Radarinfos) unter Berücksichtigung der wichtigsten allgemein interessierenden Elemente (Bewölkung, Niederschlag, Temperatur, Wind) und ggf. besondere ... Show more
Category	Climate and weather
OpenData Category (GovData)	Environment
Keywords	Deutschland Qualitätsgeprüfte Angaben Vorhersage meteorology Wetter weatherForecasts Meteorologisch-geografische Kennwerte
Mode of Transport	Other
Network Coverage	Other

„Qualitätsgeprüfte
Angaben“

(“Quality-assured
information”)

How is Quality documented on mobility data portals?

The screenshot shows the 'mobil.thek' website interface. At the top, there's a navigation bar with links for 'Data Offers', 'Blog', 'The Mobilithek', 'Help', and 'Contact'. Below this, a breadcrumb trail reads 'Search Results > Offer Details'. The main content area displays the offer title 'VE 11 - Zeitreihe 2010-2021 - Verkehr europäischer Lastkraftfahrzeuge - Güter (7)' and its provider 'Kraftfahrt-Bundesamt (KBA)'. A table below the title lists metadata: 'CREATED' (01.03.2024), 'DATA MODEL' (Other), 'TYPE OF TERMS OF USE' (Licence, free use/open data), and 'BROKERING TYPE' (Not brokered). A 'Contact' button is visible. The 'Offer Details' section is active, showing 'General Information'. The 'Description' field contains a detailed paragraph about the data source and quality. The 'Category' is 'Roads' and the 'OpenData Category' is 'Transport'.

Offer Details

General Information

Description

Die Statistik "Verkehr europäischer Lastkraftfahrzeuge (VE)" stellt die Transporte des internationalen Güterkraftverkehrs über deutsches Bundesgebiet dar und vervollständigt die Transporte deutscher Lastkraftfahrzeuge (LkFz). Hierzu erhält das KBA, als national zuständige Behörde, aggregierte und hochgerechnete Jahresdaten gemäß Anhang D der Verordnung (EG) Nr. 6/2003 der Kommission aus dem Datenpool von Eurostat. Die übermittelten Daten liegen in unterschiedlicher regionaler Gliederungsstufe (NUTS) entsprechend der Verordnung (EG) 1059/2003 des Europäischen Parlaments und des Rates vor. Zu einigen Staaten des Geltungsbereichs der Verordnung gibt es zeitliche Datenlücken bzw. vorzeitige Datenmeldungen. Die Datengewinnung erfolgt zumeist (wie in Deutschland) als Stichprobenerhebung. Basis der Auswertung der Güterbeförderungen europäischer Fahrzeuge sind alle Fahrten aller gemeldeten Fahrzeuge, die deutsches Bundesgebiet befahren - INKLUSIVE der ausländischen Streckenanteile zur jeweiligen Fahrt. Ausgenommen sind Durchfahrten, da Informationen zu beförderten Gütern bei Durchfahrten für ausländische Fahrzeuge lediglich auf NUTS 0-Ebene zur Verfügung stehen. Bitte beachten Sie die Qualitätsberichte.

Category Roads

OpenData Category Transport

„Bitte beachten Sie die Qualitätsberichte.“

(„Please consider the quality reports.“)

How is Quality documented on mobility data portals?

mobil:thek Data Offers Blog The Mobil:thek Help Contact DE | EN BD

Search Results › Offer Details

Offered by: H2Think gGmbH Visibility: Public

AuDroK: Sounddaten von externen Partnern

CREATED	DATA MODEL	TYPE OF TERMS OF USE	BROKERING TYPE
01.08.2023	Other	royalty-free	Not brokered
GEOGRAPHY	CATEGORY		
See map	Other		

Contact

Offer Details

General Data Access Terms of Use **Quality Information**

Quality Description

In der bereitgestellten Datenbank werden sowohl hochqualitative eigene Messaufnahmen als auch recherchierte Geräuschdaten von Partnern oder externen Quellen referenziert. Weitere Hinweise sind dateibezogen in der audrok-Datenbank enthalten.

„In der Datenbank werden ..
hochqualitative Messaufnahmen ...
referenziert.“

““In the database ... high-quality
measurement recordings ... are
referenced.”

How is Quality documented on mobility data portals?

The screenshot shows the 'mobil:thek' website interface. At the top, there is a navigation bar with links for 'Data Offers', 'Blog', 'The Mobil:thek', 'Help', and 'Contact'. Below this, the page title is 'Search Results > Offer Details'. The offer is provided by 'heliopas.ai GmbH' and is 'Public'. The title of the offer is 'Hochaufgelöste Bodenfeuchtekarte - Frühjahr 2022'. Below the title, there is a table with the following information:

CREATED	DATA MODEL	TYPE OF TERMS OF USE	BROKERING TYPE
01.12.2022	—	Licence, free use/open data	Not brokered

Below the table, there is a 'Contact' button. The page also has a 'GEOGRAPHY' section with a 'See map' link and a 'CATEGORY' section with 'Climate and weather'.

Offer Details

General Data Access Terms of Use **Quality Information**

Quality Description

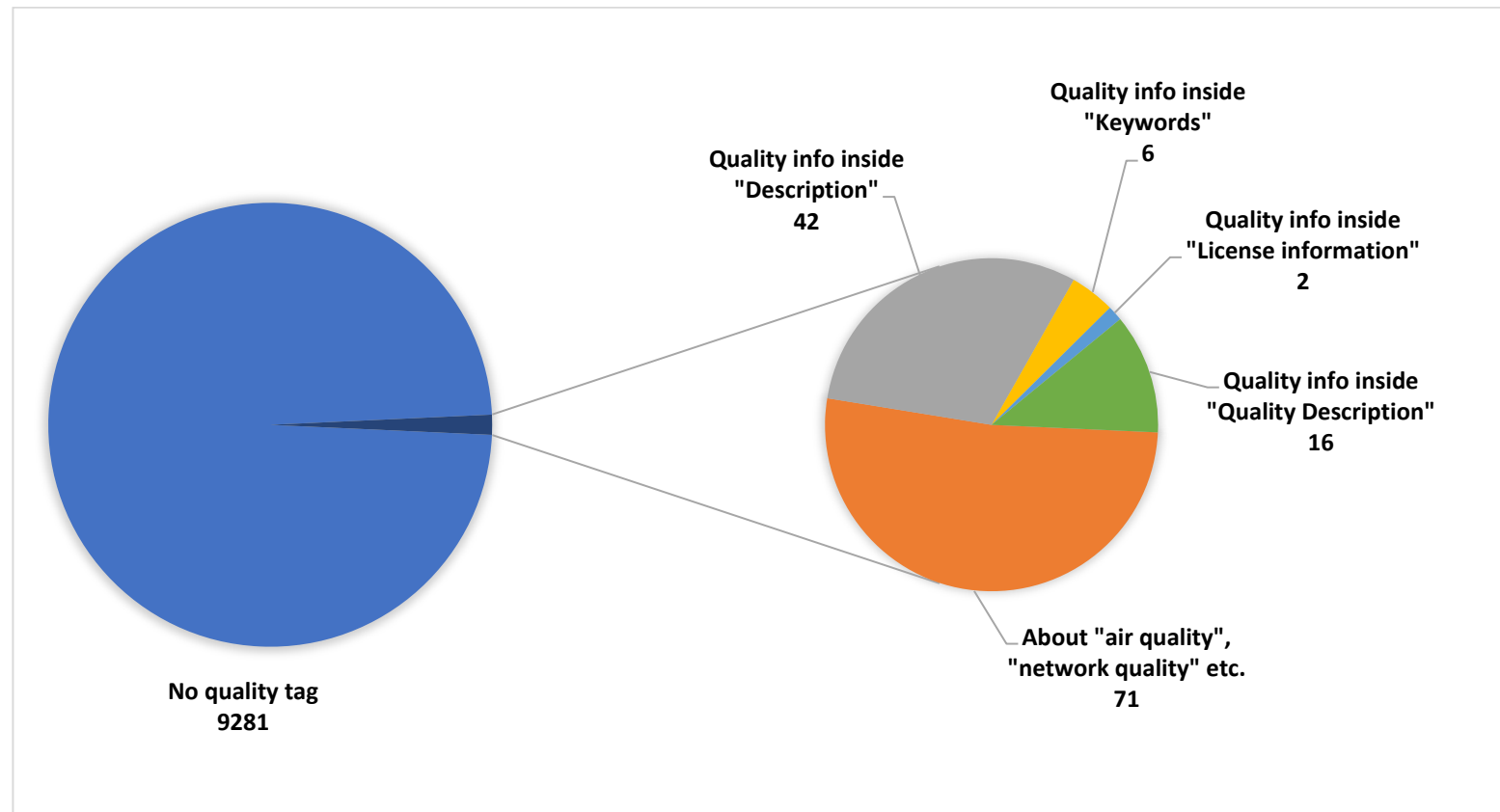
Im Rahmen Machbarkeitsstudie HABKIS wurde gezeigt, dass mit KI-basierten Methoden und geeigneten Daten, vor allem Satellitendaten, die räumliche Auflösung von Bodenfeuchtekarten von 50 km auf 10 m für beliebige Orte in Deutschland mit annähernd gleicher Qualität wie die des DWD AMBAV-Modells erhöht werden kann. Im Durchschnitt beträgt die Abweichung in der Studie 19,2 %nFK. Das Bestimmtheitsmaß liegt bei $R^2 = 0,736$.

„ Im Durchschnitt beträgt die Abweichung in der Studie 19,2 %nFK. Das Bestimmtheitsmaß liegt bei $R^2 = 0,736$.“

“On average, the deviation in the study is 19.2 %nFK. The coefficient of determination is $R^2 = 0.736$.“

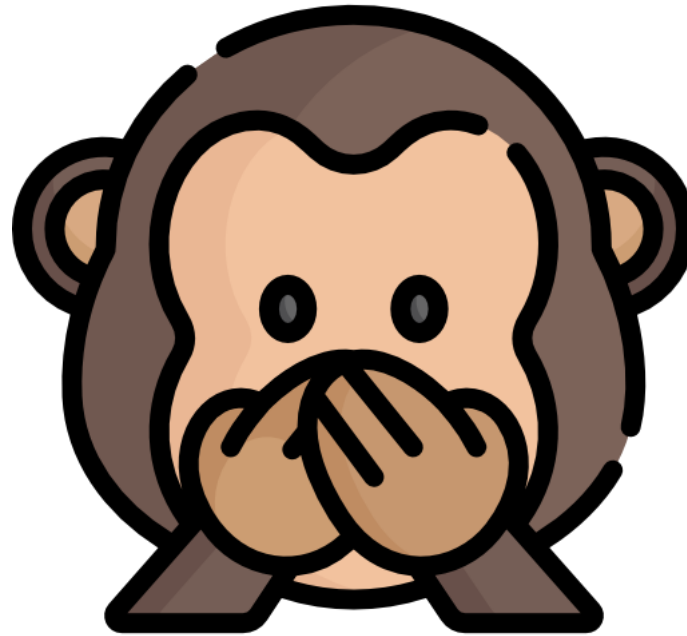
How is Quality documented on mobility data portals?

- Presence of „Quality“ information on German NAP (October 30th, 2024)



= amount of datasets

Defining data quality: Challenges



Defining data quality: Challenges



What is a Quality Framework?

- ✓ Define an agreed set of quality definitions and concepts
 - ✓ quality criteria
 - ✓ quality requirements
 - ✓ assessment methods
- ✓ Regarding data and information exchanged via NAPs and other mobility portals
- ✓ Some level of specialisation, but can be further concretised
- ✓ Current focus: data and services specified by the EC Delegated Regulations

Previous definitions of Quality Criteria

Ability to be Joined With Acceptability	Ability to Download	Ability to Identify Errors	Ability to Upload
Adaptability	Access by Competition	Accessibility	Accuracy
Age	Adequate Detail	Adequate Volume	Aestheticism
Auditable	Aggregatability	Alterability	Amount of Data
Breadth of Data	Authority	Availability	Believability
Clarity of Origin	Brevity	Certified Data	Clarity
	Clear Data	Compactness	Compatibility
	Responsibility		
Competitive Edge	Completeness	Comprehensiveness	Compressibility
Concise	Conciseness	Confidentiality	Conformity
Consistency	Content	Context	Continuity
Convenience	Correctness	Corruption	Cost
Cost of Accuracy	Cost of Collection	Creativity	Critical
Current	Customizability	Data Hierarchy	Data Improves Efficiency
Data Overload	Definability	Dependability	Depth of Data
Detail	Detailed Source	Dispersed	Distinguishable
			Updated Files
Dynamic	Ease of Access	Ease of Comparison	Ease of Correlation
			Ease of Understanding
Ease of Data Exchange	Ease of Maintenance	Ease of Retrieval	Ease to Question
Ease of Update	Ease of Use	Easy to Change	Ergonomic
Efficiency	Endurance	Enlightening	Extendability
Error-Free	Expandability	Expense	Flawlessness
Extensibility	Extent	Finalization	Integrity
Flexibility	Form of Presentation	Format	Historical Compatibility
Friendliness	Generality	Habit	Integrity
			Level of Standardization
Importance	Inconsistencies	Integration	Manipulable
Interactive	Interesting	Level of Abstraction	Minimality
			Normality
Localized	Logically Connected	Manageability	Orderliness
Measurable	Medium	Meets Requirements	Past Experience
Modularity	Narrowly Defined	No lost information	Portability
Novelty	Objectivity	Optimality	Purpose
Origin	Parsimony	Partitionability	Regularity of Format
Pedigree	Personalized	Pertinent	Reproducibility
Preciseness	Precision	Proprietary Nature	Retrievability
Quantity	Rationality	Redundancy	Robustness
Relevance	Reliability	Repetitive	Self-Correcting
Reputation	Resolution of Graphics	Responsibility	Source
Revealing	Reviewability	Rigidity	
Scope of Info	Secrecy	Security	
Semantic	Semantics	Size	
Interpretation			
Specificity	Speed	Stability	Storage
Synchronization	Time-independence	Timeliness	Traceable
Translatable	Transportability	Unambiguity	Unbiased
Understandable	Uniqueness	Unorganized	Up-to-Date
Usable	Usefulness	User Friendly	Valid
Value	Variability	Variety	Verifiable
Volatility	Well-Documented	Well-Presented	

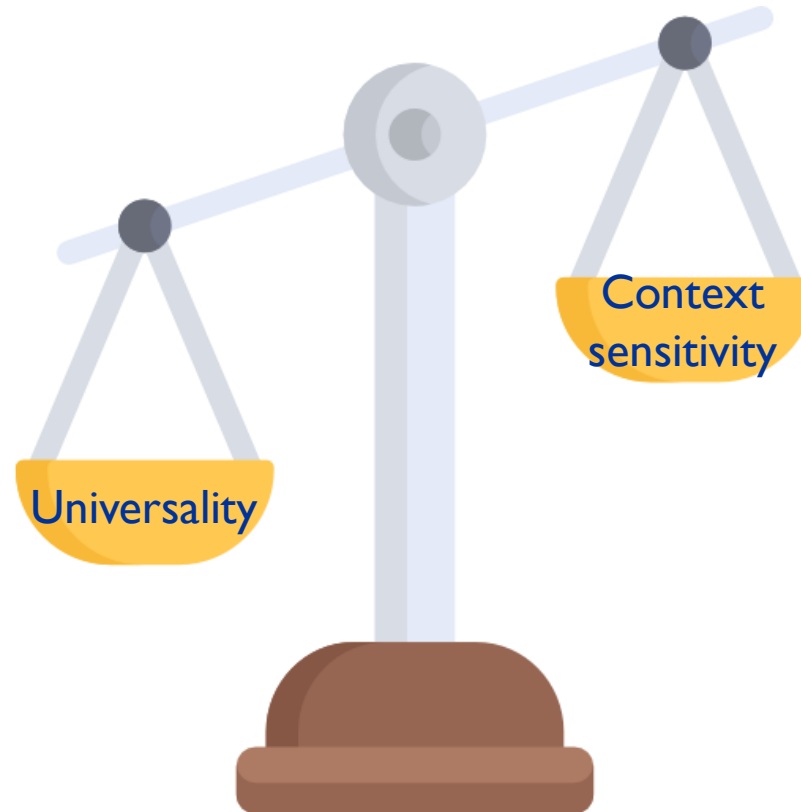
Figure 1. Data Quality Attributes Generated from the First Survey

Beyond Accuracy: What Data Quality Means to Data Consumers, Wang, R., Strong, D. M., Journal of Management Information Systems, Vol. 12, No. 4, 1996

Previous definitions of Quality Criteria

Traffic Data Quality Measurement: Final Report (FHWA, 2004)	Quality of Safety-Related and Real-Time Traffic Information Services (EU EIP, 2019)	ISO 19157-1:2023 Geographic information Data quality	ISO/TR 21707:2008 Data quality in ITS systems
Accuracy Completeness Validity Timeliness Coverage Accessibility	Geographical coverage Availability Timeliness Reporting period Latency Location accuracy Classification correctness Error Rate Event coverage Report coverage	Completeness Logical consistency Positional accuracy Temporal quality Thematic quality	Service completeness Service availability Service grade Veracity Precision Timeliness Location measurement Measurement source Ownership

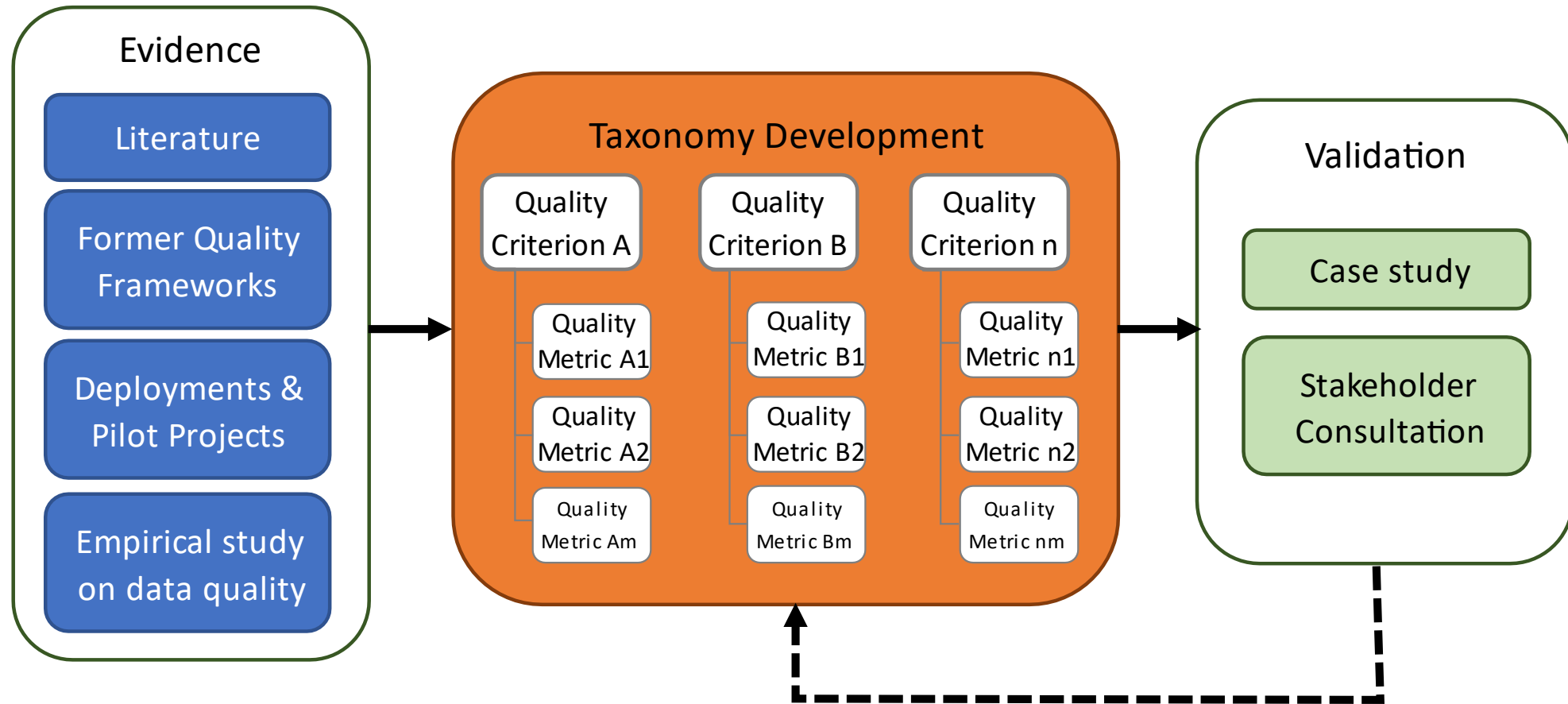
Defining data quality: Challenges



Overview on Quality Frameworks

Scope/domain	Baseline/approach	Project
Safety-Related and Real-Time Traffic Information Services (SRTI and RTTI)	Workshops, conceptualisation, validation with real-life data	 Link Published in 2019-2020
Multimodal Travel Information Services (MMTIS)	Stakeholder consultation	
Intelligent Truck Parking Services (ITPS)	Workshops, conceptualisation, validation with real-life data	
Cooperative ITS (C-ITS)	Pilot projects	
On-Street Parking Data	Literature, Real-life data assessment, Stakeholder consultation	 Link To be published in 2024
Urban Vehicle Access Regulations (UVAR)	Literature, pilot projects	
Alternative Fuels (AF)	Literature	
Multimodal Travel Information Services (MMTIS) - upgrade	Analysis of national approaches, Stakeholder consultation	
Floating Car Data (FCD)	Literature, Real-life data assessment	
Network Topology	Pilot projects	
A Cross-Domain Quality (CDQ) methodology for assessing data and services	InQMS project, workshops, validation with real-life data	

How to do a Quality Framework?



Quality Frameworks: Agenda

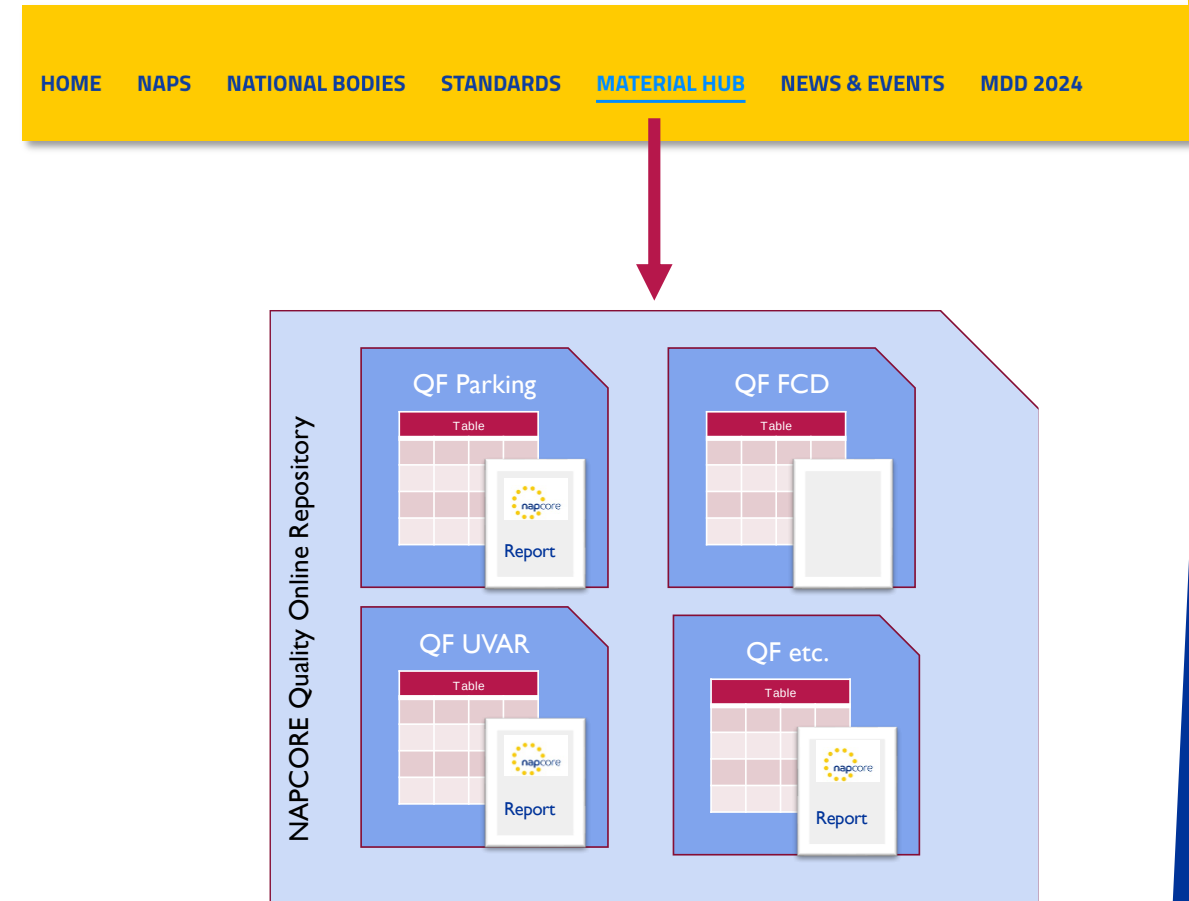
- Welcome, background and goals *Peter Lubrich, BAST (DE)*
- A Quality Framework for UVAR Data *João Montenegro, ARMIS (PT)*
- A Quality Framework for Floating Car Data *Chrysostomos Mylonas/
Maria Stavara, CERTH (GR)*
- A Quality Framework for MMTIS Data *Anastasia Founta, ITxPT (BE)*
- A Methodology for assessing Cross-Domain Quality (CDQ) *Petr Bureš, TTR (CZ)*
- Future prospects *Peter Lubrich, BAST (DE)*
- Q&A with audience

Quality Frameworks: Agenda

- Welcome, background and goals *Peter Lubrich, BAST (DE)*
- A Quality Framework for UVAR Data *João Montenegro, ARMIS (PT)*
- A Quality Framework for Floating Car Data *Chrysostomos Mylonas/
Maria Stavara, CERTH (GR)*
- A Quality Framework for MMTIS Data *Anastasia Founta, ITxPT (BE)*
- A Methodology for assessing Cross-Domain Quality (CDQ) *Petr Bureš, TTR (CZ)*
- **Future prospects** *Peter Lubrich, BAST (DE)*
- **Q&A with audience**

Publication of NAPCORE Quality Frameworks

- NAPCORE Quality Online Repository
 - Individual QF documents provided online
 - Can be uploaded / updated step-by-step
 - To be announced soon
 - Similar to EU EIP repository*



Align with other Quality Frameworks

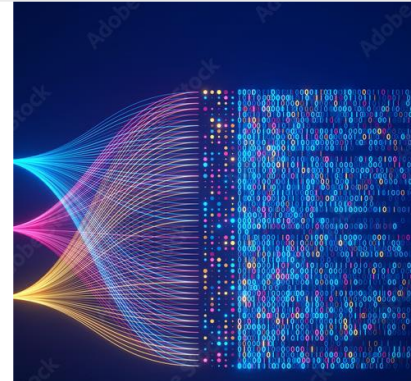


EU RTTI 5-Star Rating

Follow up Workshop
Hybrid Brussels/Online
12th March 2024



Task Force led by TISA



General Requirements for NAP
Functionality and Static + Dynamic Data

Requirements for
Road Works Data



Requirements for
Static Speed Limit Data

Requirements for
Road Closure Data



Enable feedback & interactivity

<http://dke.uqcloud.net/DataQualityPatterns/>

The image displays two overlapping screenshots of the Data Quality Patterns Portal. The background screenshot shows the homepage with a welcome message and a grid of data quality dimensions. The foreground screenshot shows a detailed view of the 'Precision' characteristic.

Homepage Content:

- Welcome to the Data Quality Patterns Portal**
- Reliability and Credibility**
 - > Source quality
 - > Objectivity
 - > Traceability
- Accuracy**
 - > Accuracy to reference source
 - > Accuracy to reality
 - > Precision
- Currency**
 - > Data timeliness
 - > Data freshness
- Availability**
 - > Data availability
 - > Continuity
 - > Data manageability
 - > Data accessibility
 - > Ease of use
 - > Data portability
- Consistency**
 - > Uniqueness
 - > Redundancy
 - > Semantic consistency
 - > Value consistency
 - > Format consistency
 - > Referential integrity
- Usability and Interpretability**
 - > Usefulness and relevance
 - > Understandability
 - > Appropriate presentation
 - > Interpretability
 - > Information value
- Completeness**
 - > Completeness of mandatory attributes
 - > Completeness of optional attributes
 - > Completeness of records
 - > Data volume
- Validity**
 - > Business rule compliance
 - > Metadata compliance
 - > Statistical compliance
 - > Standard compliance

Precision Detail View:

- Characteristic Name:** Precision
- Definition:** Attribute values should be accurate as per linguistics and granularity
- Dimension:** Accuracy
- Granularity:** Element
- Characteristic Type:** Declarative
- Implementation Form:** Rule-based approach
- Verification Metric:**
 - The number of tasks failed or under performed due to lack of data precision
 - The number of complaints received due to lack of data precision
- Validation Metric:**
 - To what extent required rules have been identified and implemented to maintain the declarative characteristic in concern.
- Background Guidelines:**
 - Definition:** Data values are correct to the right level of detail or granularity, such as price to the penny or weight to the nearest tenth of a gram.
 - Source:** ENGLISH, L. P. 2009. Information quality applied: Best practices for improving business information, processes and systems, Wiley Publishing.

Thank you for your attention!

Peter Lubrich
BAST (DE)
Email: lubrich@bast.de



Quality Framework for UVAR Data

Mobility Data Days
Turin, 6th November 2024

João Montenegro, ARMIS – PT

I. Introduction – Scoping

Urban Vehicle Access Regulations

- A set of traffic management policies that regulate access and movement of vehicles in specific urban locations

Most common UVAR types

- Low Emission Zones (LEZ)
- Pedestrian zones (PED)
- Limited Traffic Zones (LTZ)
- Congestion Charges (CS)
- Parking Regulations (PARK)



Pedestrian Zone in Amsterdam (NL)

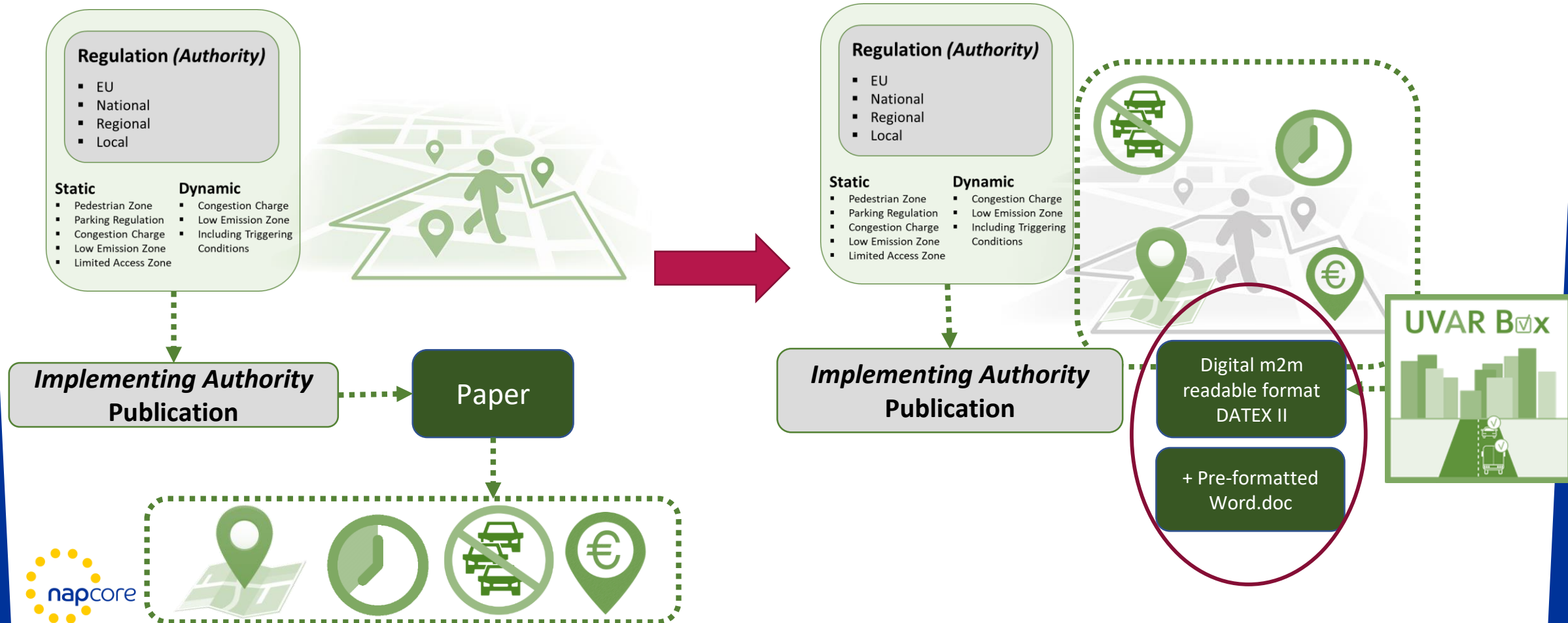


Low Emission Zone in Lisbon (PT)

I. Introduction – Background (UVAR Box)

Aims and scope

- Current UVAR schemes vs near future/proposed one



I. Introduction – Background (UVAR Box)

UVAR data attributes

General Information

- City/Town Name and Country
- UVAR Name
- Classification of UVAR type
- Status (active/not active)
- Issuing Authority or Regulation ID

Traffic Regulation

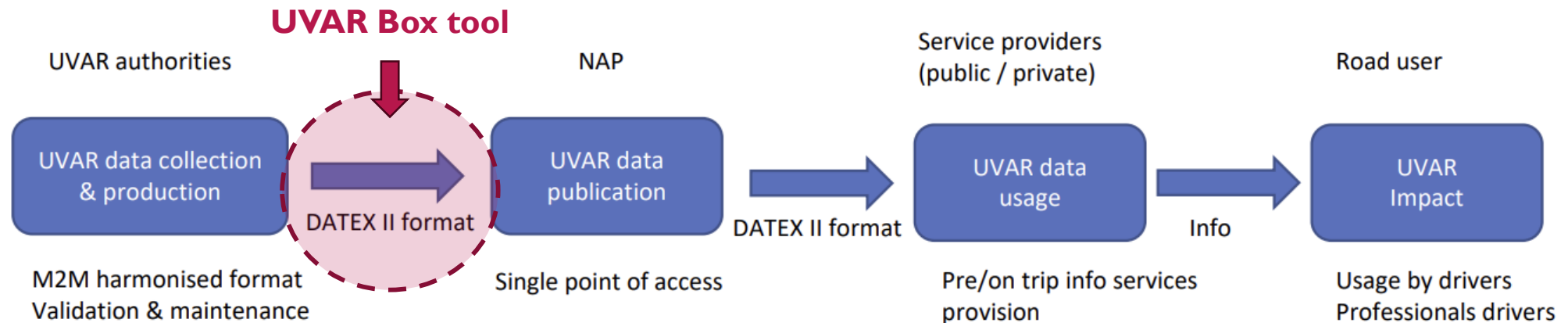
- Type of traffic regulation
- Location
- Hours of operation
- Who is affected/exempted
- Penalty/Tariffs

Additional Information

- URL for further information
- Alternative mobility options

I. Introduction – Motivation

UVAR data value chain



Adapted from: UVAR Box project

- Legislation (DRs) requires **traffic regulations' data** should be published at the NAP
- **Hesitation** by public authorities...
- ... concerns about UVAR data having enough **quality** for services like navigation
- **Common definitions** for quality in UVAR datasets are still lacking

**Motivation for
the UVAR QF**

2. Quality criteria – Methodology

Literature review on quality dimensions (for any use case!)

- Grouping quality dimensions frequently mentioned in scientific literature into 9 “categories”

#	Group	Dimensions
1	Openness	Accessibility, Availability
2	Exactness	Accuracy, Precision, Correctness
3	Comprehensiveness	Completeness, Amount of data, Granularity, Data Coverage, Variety of Data & Data Sources
4	Coherence	Consistency, Consistent Representation, Data specification, Reliability
5	Veracity	Believability, Traceability, Objectivity, Reputation
6	Promptness	Timeliness, Currency, Volatility
7	Usability	Understandability, Conciseness, Interpretability, Ease of Manipulation, Ease of Operation, Navigation
8	Safekeeping	Security, Safety, Integrity
9	Suitability	Relevance, Effectiveness, Efficiency, Value added, Transactability, Usefulness, Uniqueness

Adapted from: A. M. Carvalho, S. Soares, J. Montenegro, and L. Conceição, “Data Quality: revisiting dimensions towards new framework development,” in 6th International Conference on Industry 4.0 and Smart Manufacturing [in press], 2024.

Applying identified dimensions to UVARs

- Tailoring quality dimensions to the context of UVARs - Creating UVAR-specific definitions for each dimension
- Based on inputs from
 - Other projects (UVAR Box and UVAR Exchange)
 - DATEX II (UVAR profiles)
 - Legislative documents (ITS directive and delegated regulations)

2. Quality criteria – Results

Resulting quality criteria

- 7 dimensions with applicable criteria
- 12 criteria including 15 metrics
- Metrics with different formats – numeric, multi-class evaluation, binary... (next slide)

Quality dimension	Quality criterion	Quality metric
Accuracy	Location accuracy	Area deviation
		Hausdorff Distance
	Validity of data	TRO alignment
Completeness	Required data coverage	% general
		% traffic
	Provision of additional information	Binary classification
Accessibility	Level of Accessibility	Binary classification
	NAP Accessibility	Binary classification
Currency	Update frequency	Multiclass evaluation
	Timeliness	Multiclass evaluation
Consistency	Compliance with data standards	Binary classification
Believability	Traceability	System tracking
		Retrieval of versions
	Trustworthiness	Binary classification
Relevance	Data duplication	Binary classification

2. Quality criteria – Results

Resulting quality criteria

- Summary table – Description of criteria and metrics

Quality dimension	Quality criterion	Metrics	Mathematical/theoretical formulation	Possible values
Accuracy	Location accuracy	Area Deviation (AD)	$AD = \frac{ A_{dataset} - A_{Authority} }{A_{Authority}}$	0 to 100%
		Hausdorff Distance (d_H)	$d_H := \max\{sup_{x \in X}(x, Y), sup_{y \in Y}(X, y)\}$ Where: - X: The set of coordinates from the UVAR dataset under evaluation. - Y: The set of coordinates from the map provided by the issuing authority, considered as the ground truth.	0 to positive infinity
	Validity of data	Alignment with TRO (Multiclass evaluation)	Have any differences/misalignments between the information provided by the UVAR dataset (e.g., exemptions, hours of operation, UVAR type classification) and the Traffic Regulation Order (TRO) or UVAR traffic signs been detected?	0: Yes, there is significant misalignment - some elements clearly contrast with information from the TRO/traffic signs; 1: Slight differences identified between TRO/traffic signs and the dataset (not enough information to check if the dataset's information is correct) 2: No difference identified - dataset seems to match with either TRO or traffic signs
Accessibility	Level of Accessibility	Multi-class evaluation	Is the dataset publicly accessible? Are there access restrictions?	0: Dataset is not accessible 1: Dataset is accessible but content is not publicly accessible (meaning there are access restrictions) 2: Dataset is publicly accessible (no access restrictions)
Consistency	Compliance with data standards	Binary classification	Does the dataset comply with the DATEX II standard (UVAR profile)?	0: No; 1: Yes

Quantitative values

Binary-like and multiclass

3. Assessment methods

Generic approach

- Suggestions for which references to use (e.g.,TRO)
- Proposal for meta information for quality assessment
- Recommendations for documentation (reference to MobilityDCAT-AP)

Aggregated view

- Normalization + weighted mean of individual metrics
- Allows to obtain **quality indicators** for specific dimensions
- Values are computed automatically
- **Flexibility** – weights are assigned by the evaluator

$$Q_i = \frac{\sum_j \left(W_{i,j} \times A_{i,j} \times \frac{\text{Input value}_{i,j} - \text{Worst}_{i,j}}{\text{Best}_{i,j} - \text{Worst}_{i,j}} \right)}{\sum_j (W_{i,j} \times A_{i,j})}$$

Quality indicator ← Q_i

Weights ↑ $W_{i,j}$

Normalization ↑ $\frac{\text{Input value}_{i,j} - \text{Worst}_{i,j}}{\text{Best}_{i,j} - \text{Worst}_{i,j}}$

4. Case study

Dataset information

- Low Emission Zone in Lisbon
- Digitized into DATEXII during the UVAR Box project (using the UVAR Box tool)

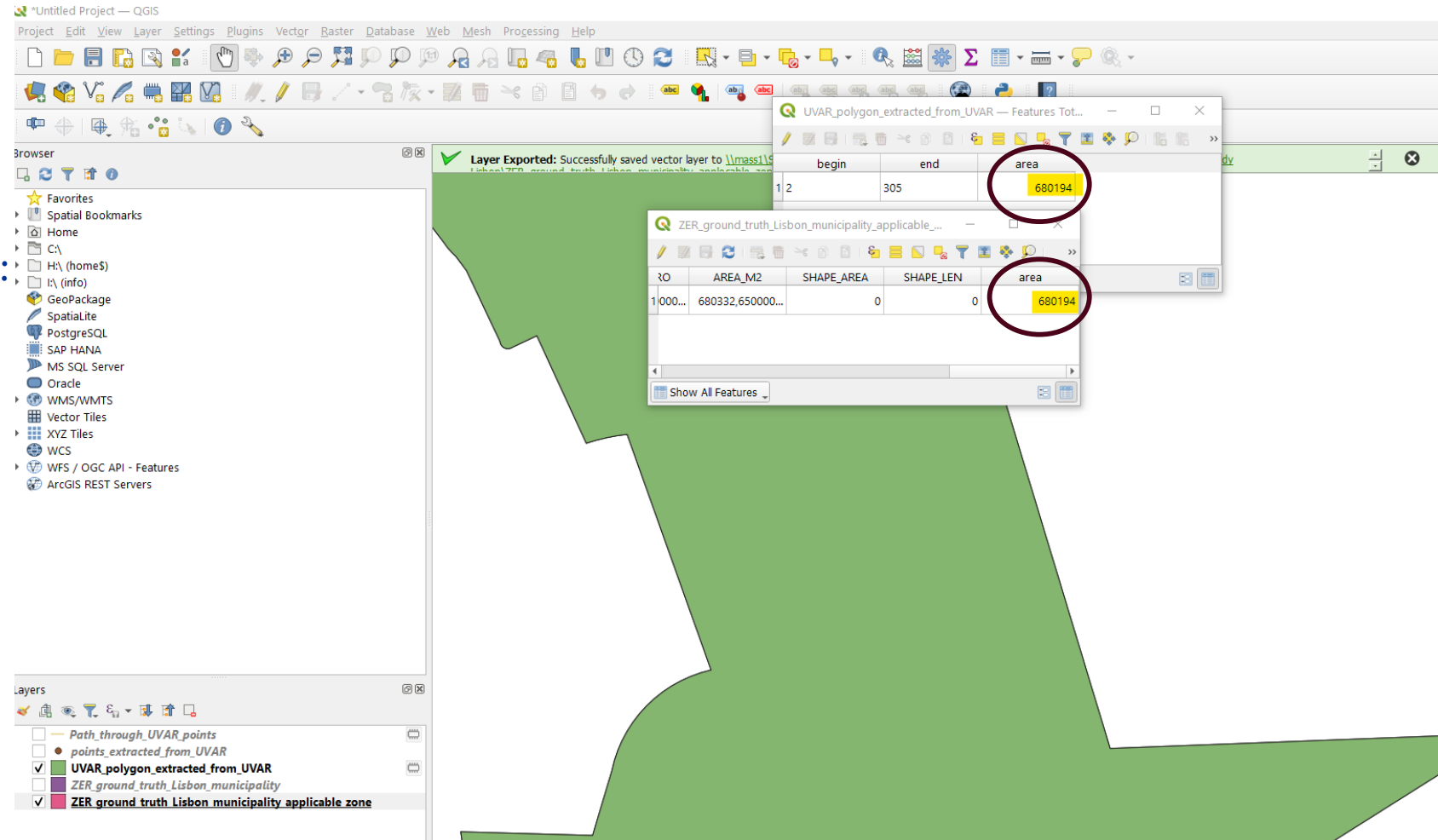


Geometry/location of UVAR dataset from the city of Lisbon

4. Case study

Application of quality criteria

- Dimension: Accuracy
- Criterion: Location Accuracy
- Coordinates match -> Metrics:
 - Area deviation = 0%
 - Hausdorff distance = 0



Areas of the Polygons formed by the ground truth's and dataset's coordinates (in QGIS)

4. Case study

Aggregated view

- Values assigned for each metric computed into the calculation sheet
- Weights selected by evaluators
- Quality indicators obtained

Quality dimension	Quality criterion	Quality metric	Weight	Input Value
Accuracy	Location accuracy	Area deviation	1	0%
		Hausdorff Distance	1	0
	Validity of data	TRO alignment	0.8	1
Completeness	Required data coverage	% general	0.7	100%
		% traffic	1	75%
Accessibility	Provision of additional information	Binary classification	0.2	1
	Level of Accessibility	Binary classification	0.6	0
Currency	NAP Accessibility	Binary classification	0.5	0
	Update frequency	Multiclass evaluation	0.5	0
Consistency	Compliance with data standards	Timeliness	0.8	N/A
		Binary classification	1	1
Believability	Traceability	System tracking	0.4	0
		Retrieval of versions	0.2	0
Relevance	Trustworthiness	Binary classification	0.7	1
		Binary classification	0.2	1



Quality dimension	Parameter weight	Quality indicator
Accuracy	0.93	0.88
Completeness	0.63	0.87
Accessibility	0.55	0.00
Currency	0.50	0.00
Consistency	1.00	1.00
Believability	0.43	0.54
Relevance	0.20	1.00

5. Conclusions

Main outcomes:

- Flexible approach accommodating the diverse requirements for UVAR data
- Objective quality criteria and feasible metrics
- Suggested method allows obtaining quality indicators
- Applied to a real-world dataset – attesting usability

Limitations:

- Requires GIS skills to calculate accuracy-related metrics
- Identify the “ground truth” might be challenging

Future work could:

- Explore a broader set of criteria and metrics
- Include multiple UVARs in a more comprehensive case study – may lead to re-scaling of metrics’ ranges
- Define minimum quality levels

Thanks

Do you have any questions?

Joao.montenegro@armis.pt





Quality Framework for Floating Car Data

Mobility Data Days

Turin, 6th November 2024

Chrysostomos Mylonas, CERTH-HIT

Maria Stavara, CERTH-HIT

I. Introduction – Scoping

Floating Car Data (FCD)

- Timestamped geolocations of connected (probe) vehicles collected by leveraging advanced onboard vehicular technologies, e.g., GPS/GSM/GPRS.
- Provision of insights into the operational state of a road network, including traffic flow, travel speeds, mobility patterns, level of congestion, etc.

Main use cases

- Digital route and map inference
- Traffic sensing and prediction
- Trajectory mining & O-D matrix estimation
- TM applications: info & warnings, advisory, traffic influence and control

Scope/objective

- Provide a framework to assess the quality of RTTI derived from FCD

I. Approach followed

- Analysis of the current SotA
- Creation of an inventory of best practices and challenges to be addressed
- Definition of the required quality dimensions and criteria
- Mathematical notation of FCD networks and metrics
- Association of quality criteria with specific metrics
- Testing on a large-scale road network (Thessaloniki)
- Analysis of results
- Ways forward

2. Summary of SotA

Empirical findings and challenges:

- The accuracy of FCD estimates depends on several parameters, including prevailing traffic conditions, **penetration rate**, and **sampling rate**.
- High sampling rate is generally beneficial for speed, and driving direction measurements; however, it may lead to the **overestimation** of **distance traveled** (propagation of measurement error).
- **Bias** in the **probe sample** may obscure certain **driving behaviors**, limiting the **representativeness** of the data.
- FCD **granularity** is crucial and useful for certain ITS services (e.g., end-of-queue collision prevention) but may create problems as well.
- Incomplete route coverage and adjacent network impact can bias FCD estimates, especially in travel time calculations.
- **Map-matching** accuracy is affected by both the applied **method** and **other factors**, such as sampling rate and **road density**.

2. Summary of SotA

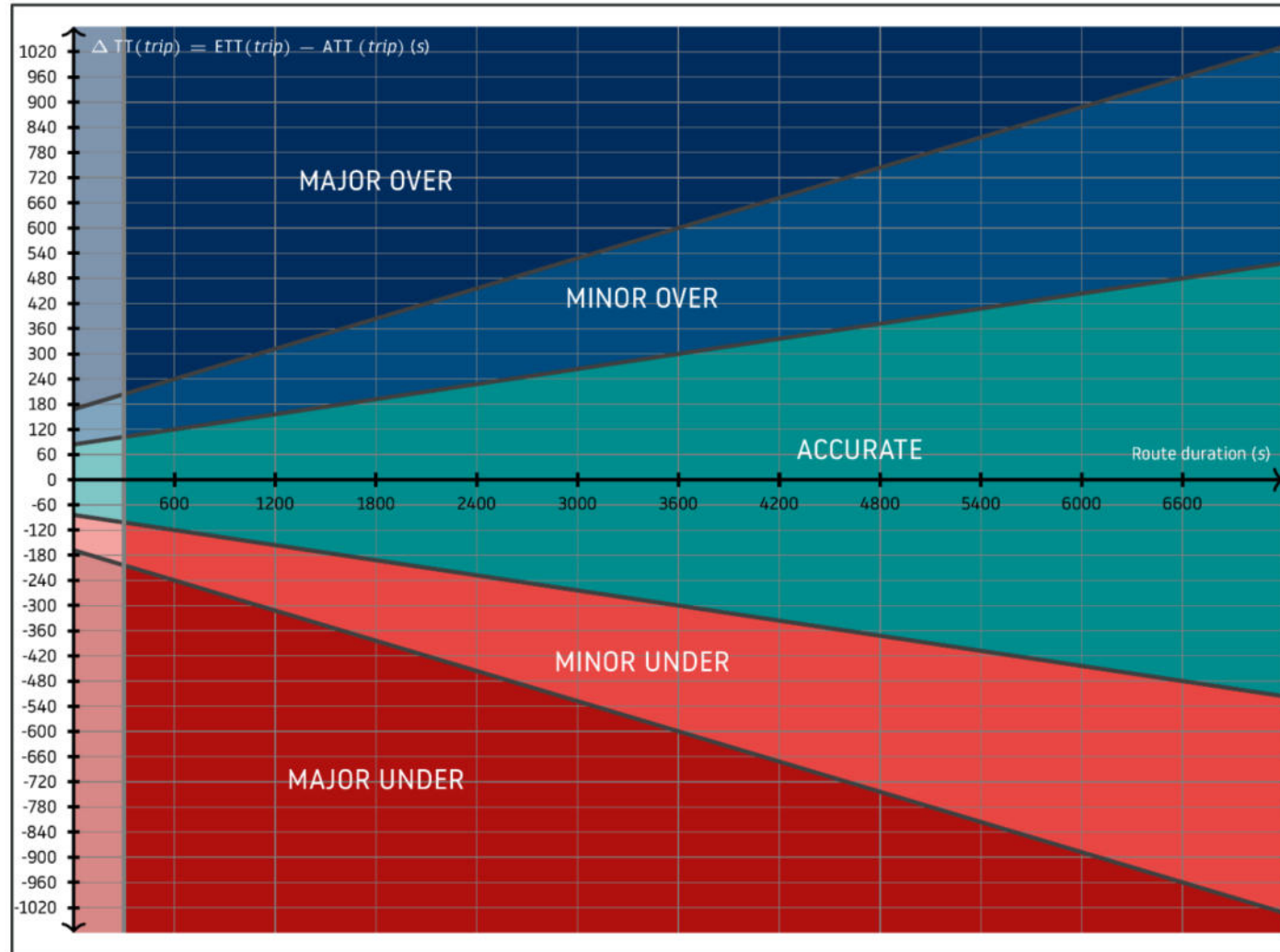
Quality Frameworks considered:

- EU-EIP framework for RTTI/SRTI
- TISA Guidelines (ETT)
- Data quality white paper FHWA
- Austrian framework for the “National FCD platform”
- Other frameworks available in the scientific literature

3. Quality dimensions, criteria, metrics

Quality criterion	Type	Quality metric	Formulation	Value range
Dimension: Correctness				
Accuracy – error rate	Direct	Pearson Correlation Index (CI)	$CI(k) = \frac{\sum_{i \in T} (x_{FCD}(k,i) - \bar{x}_{FCD})(x_{GT}(k,i) - \bar{x}_{GT})}{\sqrt{\sum_{i \in T} (x_{FCD}(k,i) - \bar{x}_{FCD})^2 \sum_{i \in T} (x_{GT}(k,i) - \bar{x}_{GT})^2}}$, where $\bar{x}_{FCD} = \sum_{i \in T} x_{FCD}(k,i) / \sum_{i \in T} 1$	-1 to 1
		Mean Absolute Percentage Error (MAPE)	$MAPE(k) = \sum_{i \in T} \left \frac{x_{FCD}(k,i) - x_{GT}(k,i)}{x_{GT}(k,i)} \right / \sum_{i \in T} 1$	0 to 1
		Root Mean Squared Percentage Error (RMSPE)	$RMSPE(k) = \sqrt{\sum_{i \in T} \left(\frac{x_{FCD}(k,i) - x_{GT}(k,i)}{x_{GT}(k,i)} \right)^2 / \sum_{i \in T} 1}$	0 to 1
		Per Mile Absolute Error (PMAE)	$PMAE(k) = \sum_{i \in T} x_{FCD}(k,i) - x_{GT}(k,i) / \sum_{i \in T} 1(k)$	≥ 0
Bias rate	Direct	Mean Bias Percentage Error (MBPE)	$MBPE(k) = \sum_{i \in T} \left(\frac{x_{FCD}(k,i) - x_{GT}(k,i)}{x_{GT}(k,i)} \right) / \sum_{i \in T} 1$	-1 to 1
Classification correctness	Direct	Congestion Classification Error in Congestion (CCEC)	$CCEC(k) = \frac{1}{\sum_{i \in T} \delta_{x_{GT}(k,i) > a \times tt_f(k)}} \sum_{i \in T} \delta_{x_{FCD}(k,i) > a \times tt_f(k)} - \delta_{x_{GT}(k,i) > a \times tt_f(k)} $, where $\delta_{x_{GT}(k,i) > a \times tt_f(k)}$ and $\delta_{x_{FCD}(k,i) > a \times tt_f(k)}$ equal to 1 if the encompassed conditions are satisfied and 0 otherwise.	0 to 1
Reliability	Direct	Minor Threshold Reliability	$R_{minor}(k) = 1 - (P_{minor\ under}(k) + P_{minor\ over}(k))$, where $P_{minor\ under}(k) = P_{i \in T} (x_{FCD}(k,i) - x_{GT}(k,i) < QT_{under}^{minor}(k,i))$ and $P_{minor\ over}(k) = P_{i \in T} (x_{FCD}(k,i) - x_{GT}(k,i) < QT_{over}^{minor}(k,i))$	0 to 1
		Major Threshold Reliability	$R_{major}(k) = 1 - (P_{major\ under}(k) + P_{major\ over}(k))$, where $P_{major\ under}(k) = P_{i \in T} (x_{FCD}(k,i) - x_{GT}(k,i) > QT_{under}^{major}(k,i))$ and $P_{major\ over}(k) = P_{i \in T} (x_{FCD}(k,i) - x_{GT}(k,i) > QT_{over}^{major}(k,i))$	0 to 1
		Overall Reliability	$R(k) = (R_{minor}(k) + R_{major}(k)) / 2$	0 to 1

3. Quality dimensions, criteria, metrics



3. Quality dimensions, criteria, metrics

Quality criterion	Type	Quality metric	Formulation	Value range
Dimension: Completeness				
Network coverage	Direct	Link Coverage (LC)	$LC = \frac{\sum_{j \in L^*} 1}{\sum_{j \in L} 1}$	0 to 1
		Weighted Link Coverage (WLC)	$WLC = \frac{\sum_{j \in L^*} w_{(2,j)}}{\sum_{j \in L} w_{(2,j)}}$	0 to 1
Spatial coverage	Direct	Grid Coverage (GC)	$GC = \frac{\sum_{j \in A^*} 1}{\sum_{j \in A} 1}$	0 to 1
		Weighted Grid Coverage (WGC)	$WGC = \frac{\sum_{j \in A^*} w_{(1,j)}}{\sum_{j \in A} w_{(1,j)}}$	0 to 1
Missing values	Direct	Percentage of Missing Values (PMV)	$PMV = \frac{\sum_{i \in T} \sum_{j \in S} \delta_{i,j}}{\sum_{i \in T} \sum_{j \in S} 1}$, where $\delta_{i,j}$ is an indicator variable with $\delta_{i,j} = 1$ if $x_{FCD}(i,j) \neq \emptyset$ and $\delta_{i,j} = 0$ otherwise	0 to 1
		Weighted Percentage of Missing Values (WPMV)	$WPMV = \frac{\sum_{i \in T} \sum_{j \in S} \delta_{i,j} w_{(3,i)}}{\sum_{i \in T} \sum_{j \in S} w_{(3,i)}}$, where $\delta_{i,j}$ is an indicator variable with $\delta_{i,j} = 1$ if $x_{FCD}(i,j) \neq \emptyset$ and $\delta_{i,j} = 0$ otherwise	0 to 1

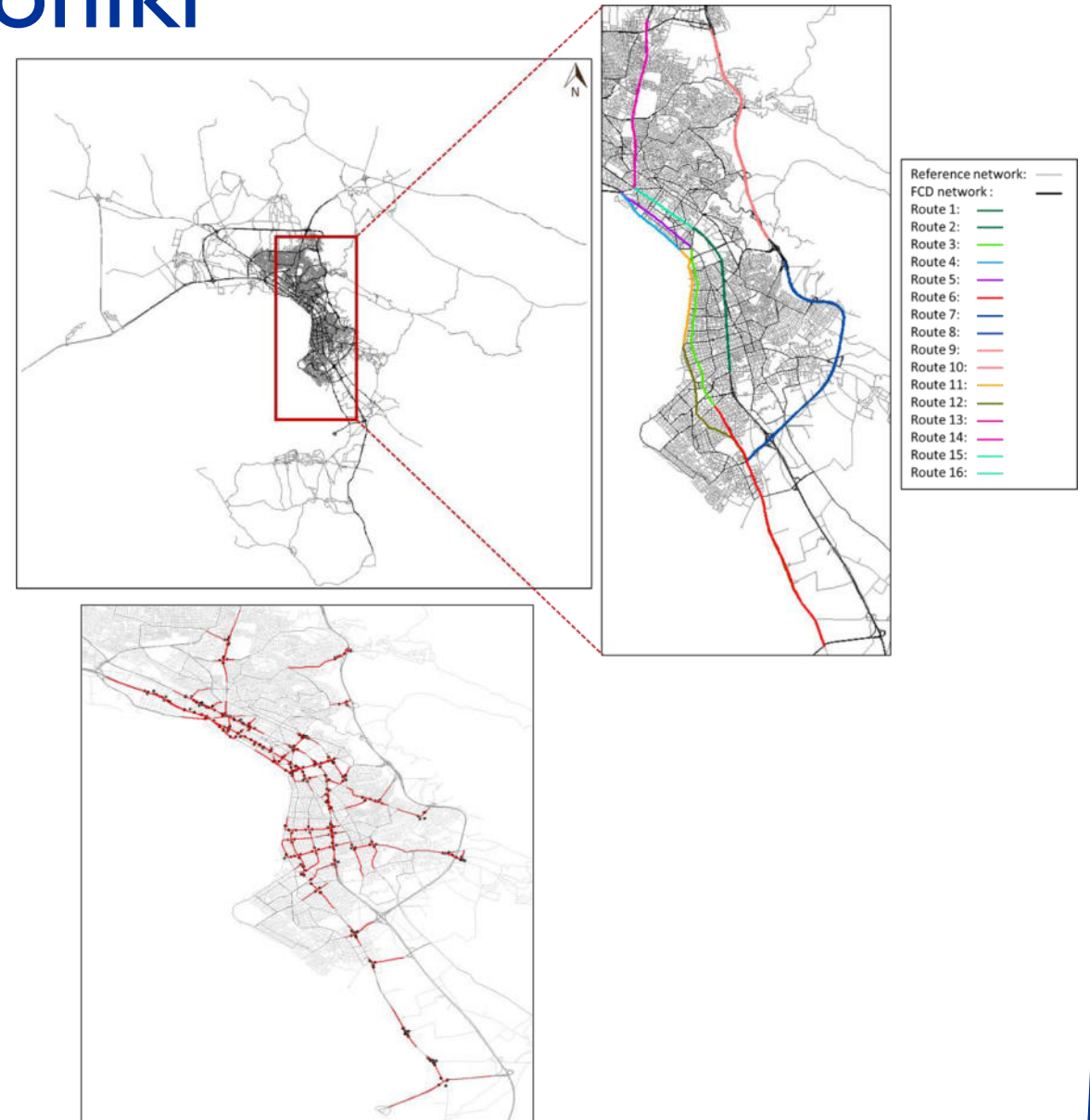
Quality criterion	Type	Quality metric	Formulation	Value range
Dimension: Timeliness				
Freshness	Direct	Average Update Frequency (AUF)	$AUF = \sum_{i \in T[0: T -1]} t(i+1) - t(i) / \sum_{i \in T[0: T -1]} 1$	≥ 0
Latency	Direct	Average Time Delay (ATD)	$ATD = \sum_{i \in T} t'(i) - t(i) / \sum_{i \in T} 1$	≥ 0

3. Quality dimensions, criteria, metrics

Quality criterion	Type	Quality metric	Formulation	Value range
Dimension: Hybrid				
Granularity	Informative	Segmentation Density (SD)	$SD = \sum_{k \in P} \frac{\sum_{k \in S(k)} 1}{l(k)} / \sum_{k \in P} 1$	≥ 0
		Topological Discrepancy Index (TDI)	$TDI = \sum_{i \in S} \sum_{j \in L^*(i)} \frac{l(i) - l(j)}{l(i)} / \sum_{i \in S} \sum_{j \in L^*(i)} 1$	$(-\square, +\square)$
Penetration	Contextual	Average Penetration Rate (APR)	$APR = \sum_{i \in T} \sum_{j \in S_c} \frac{N_{UProbes(i,j)}}{V(i,j)} / \sum_{i \in T} \sum_{j \in S_c} 1$	0 to 1
Sampling	Contextual	Average Sampling Rate (ASR)	$ASR = \sum_{i \in T} \sum_{j \in S} \frac{N_{UProbes(i,j)}}{N_{ProbeT(i,j)}} / \sum_{i \in T} \sum_{j \in S} 1$	≥ 0
Map matching	Contextual	Map-Matching Method (MMM)	Possible options: (1) Geometrical, (2) Topological, (3) Statistically enhanced	-
Fleet composition	Contextual	Fleet Composition Score (FCS)	$FCS = \sum_{i \in T} \sum_{j \in S} VF(i,j) / \sum_{i \in T} \sum_{j \in S} 1$	≥ 0

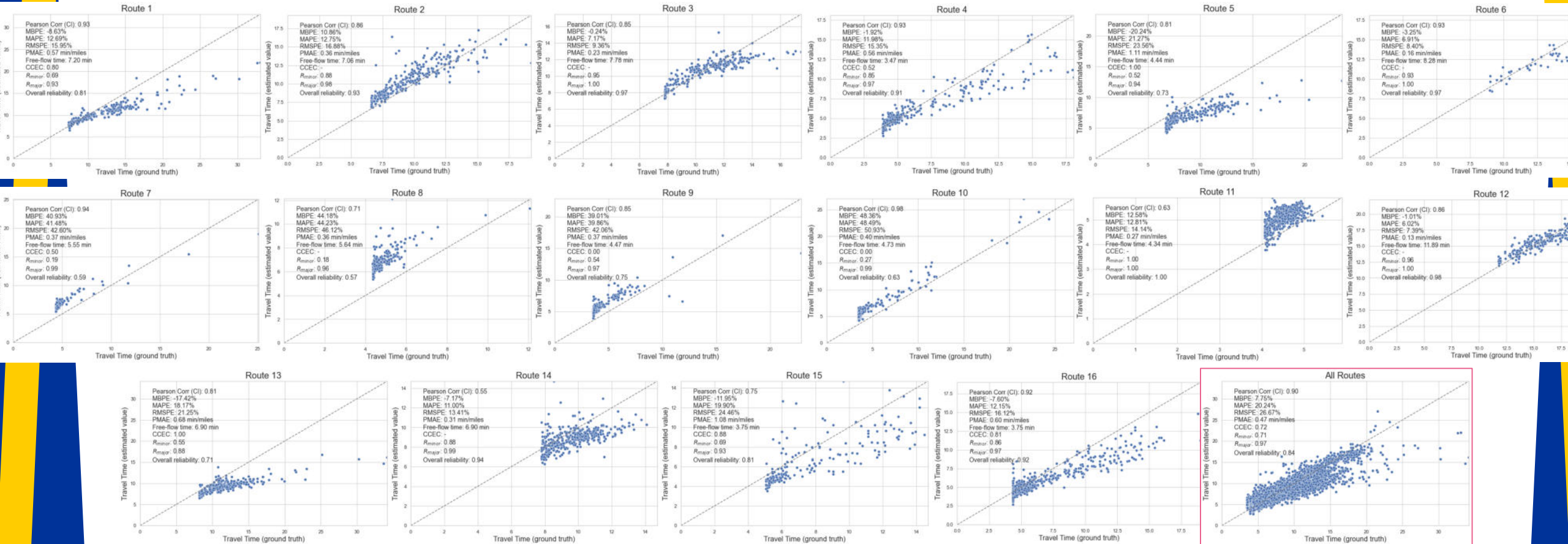
4. Case study in Thessaloniki

- Second biggest city in Greece (~1 million inhabitants) with limited alternative transport modes and frequent traffic bottlenecks
- Evaluation of an open FCD service derived from a large taxi fleet (~1200 vehicles)
- Comparison with a commercial FCD service (“ground truth”)
- Evaluation period: 10 days
- Time aggregation intervals: 15 minutes
- Use of 16 traffic routes (accuracy dimension)
- Use of data derived from inductive loop detectors (penetration rate)
- Use of Overpass API for POIs extraction
- Use of QGIS and Python libraries for data analysis and road network classification



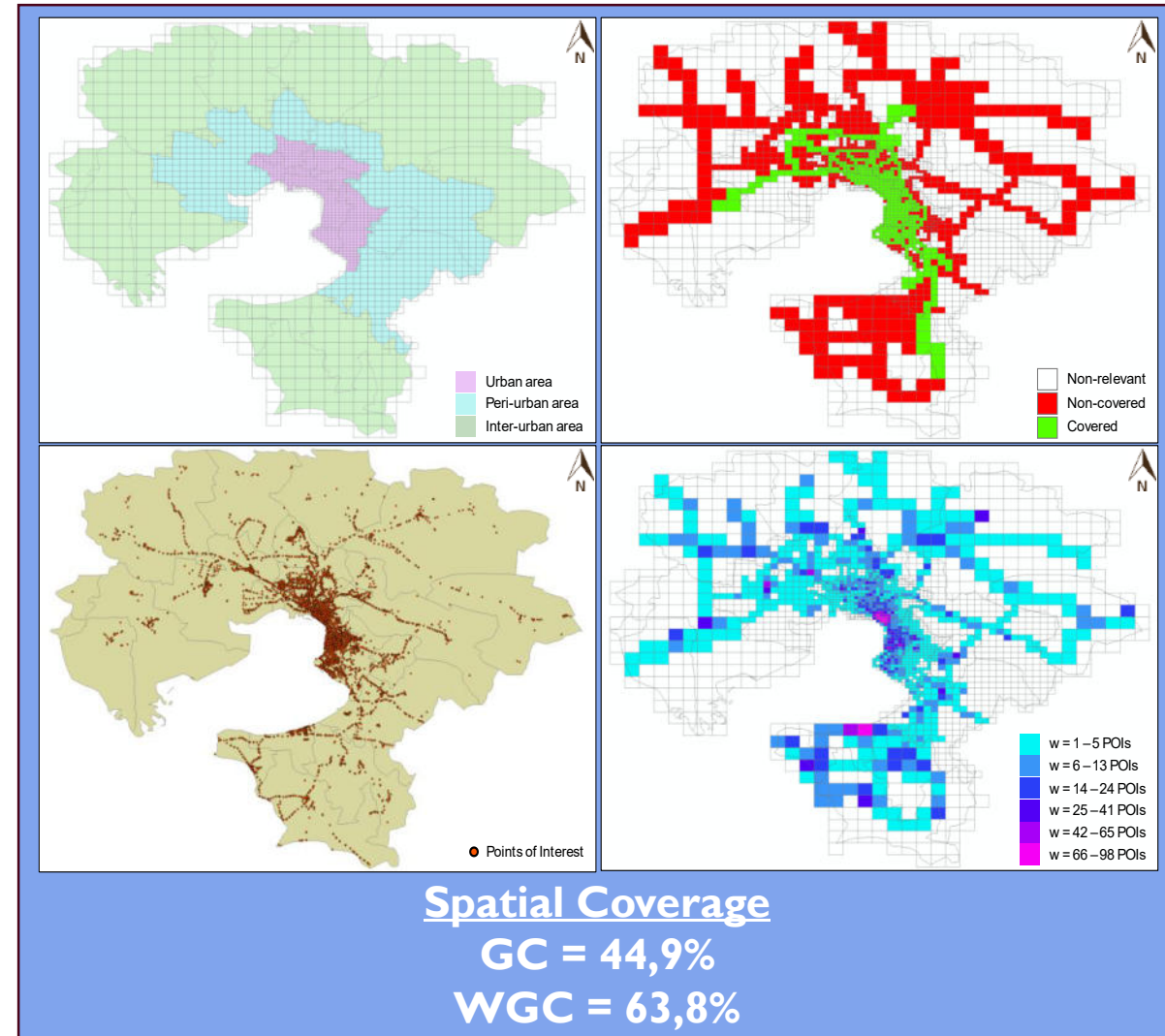
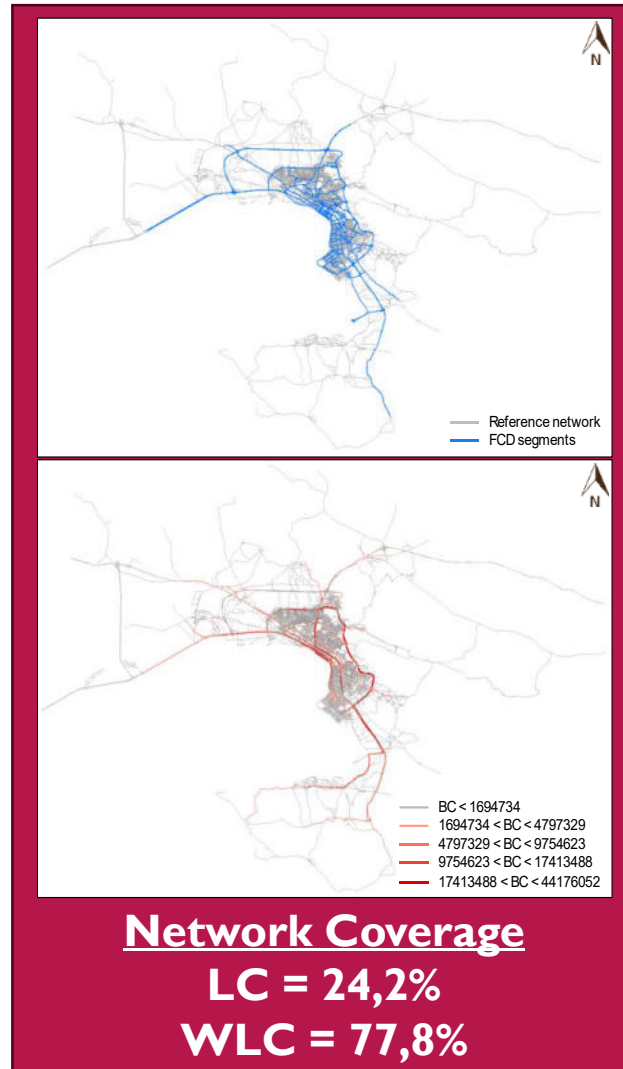
4. Case study in Thessaloniki

Correctness dimension



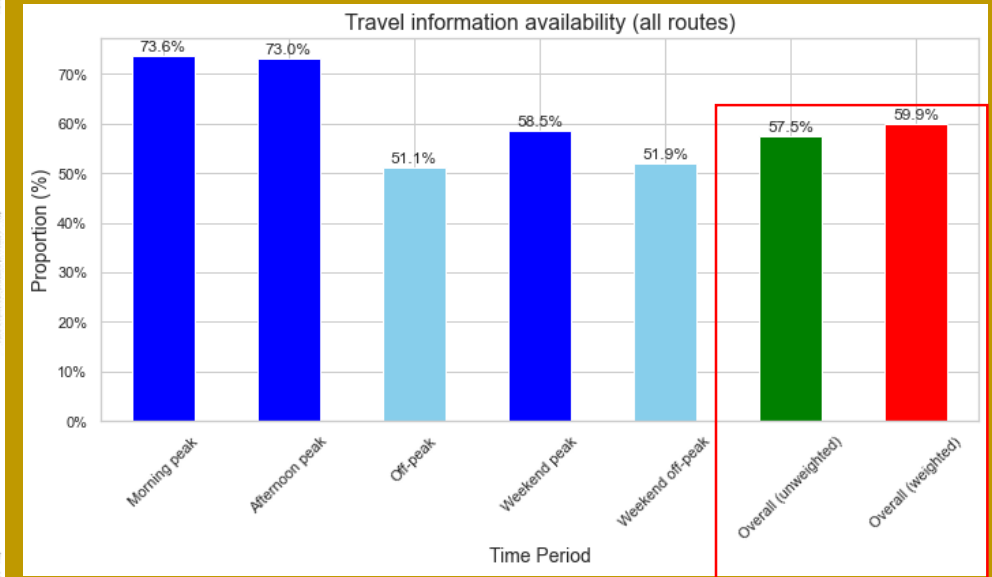
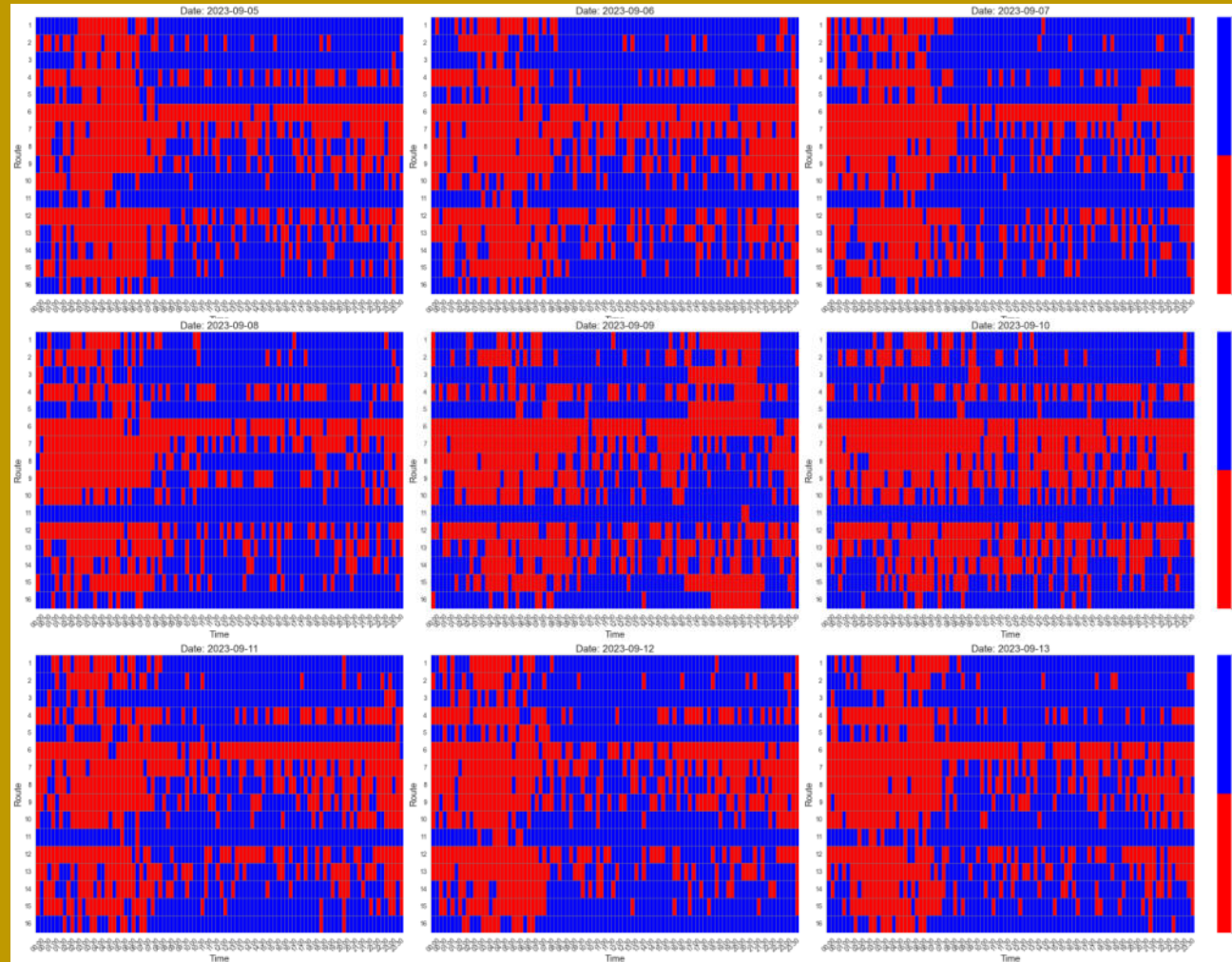
4. Case study in Thessaloniki

Completeness dimension



4. Case study in Thessaloniki

Completeness dimension



Missing values

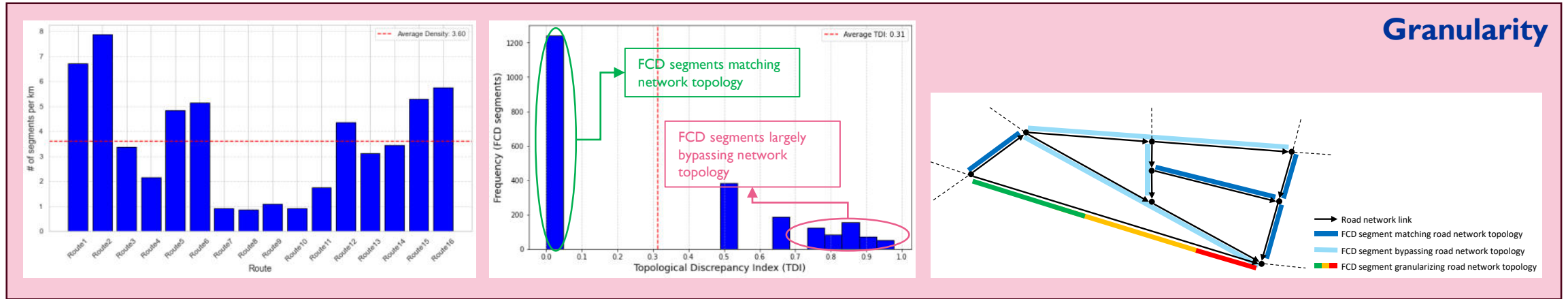
4. Case study in Thessaloniki

Timeliness dimension

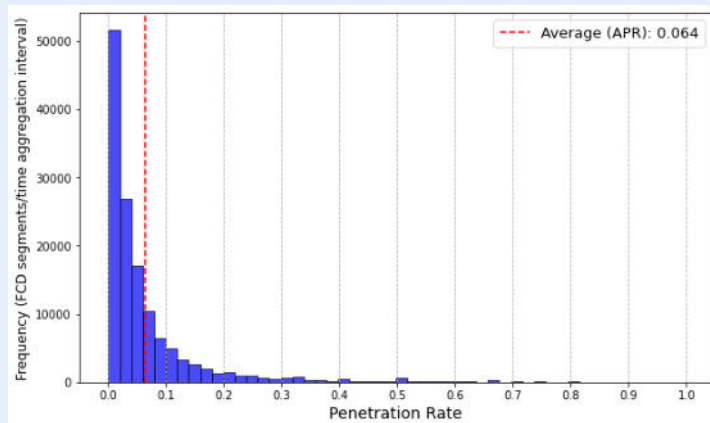
- Average Update Frequency (AUF) = 15 minutes (pre-fixed)
- Average Time Delay (ATD) = 10.8 seconds (30 measurements)

4. Case study in Thessaloniki

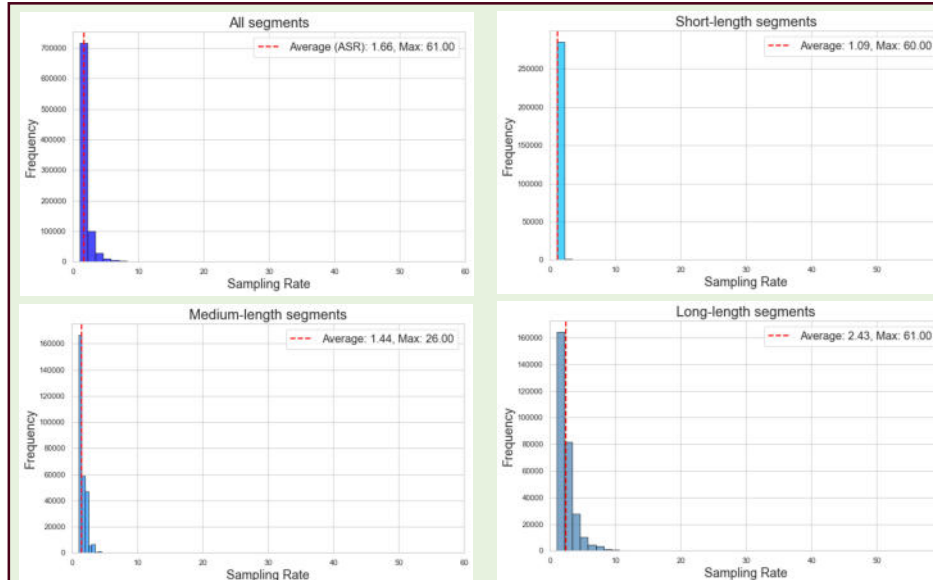
Hybrid dimension



Penetration rate

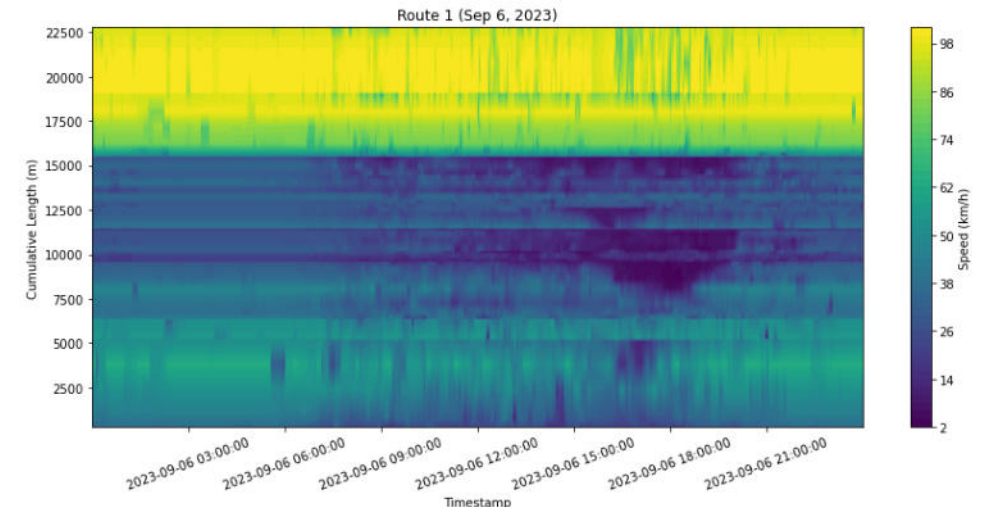
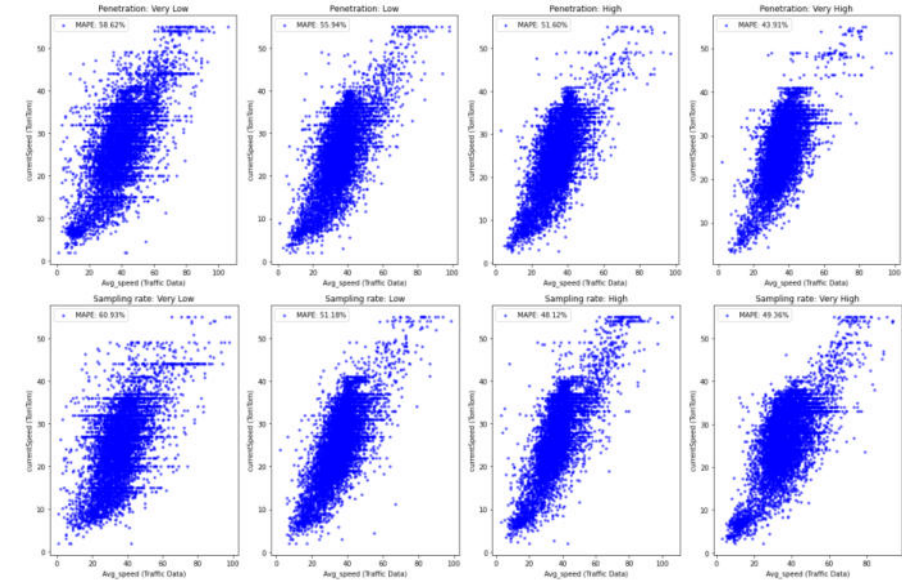


Sampling rate



5. Ways forward

- Experimental analysis of derived results
 - Effect of penetration rate on accuracy
 - Effect of sampling rate on accuracy
 - Effect of update frequency on ...
 - ...
- When do we need to reconstruct travel times following a back-office approach?
- Discussion of specific quality criteria
 - Is bias rate always bad?
 - What is an appropriate FCD segmentation?
- Publication of a paper
- Set of criteria to be used by NAPs and short guidelines



Thanks

Do you have any questions?

chmylonas@certh.gr

mstavara@certh.gr





Quality Framework for Multimodal Travel Information Services (MMTIS) Data

**Mobility Data Days
Turin, 6th November 2024**

Anastasia Founta, Information Technology for Public Transport

Agenda

1 Introduction – Why a Quality Framework on MMTIS data?

2 NAPCORE contribution - Methodology

3 Quality dimensions and criteria

4 Next steps

Why a MMTIS Quality Framework

Regulatory requirement

Delegated
Regulation
2024/490

“Member States and ITS stakeholders should cooperate **to reach an agreement on common definitions of data quality** with a view to use common data quality indicators ..., such as the **completeness, accuracy and up-to-dateness of the data** as well as the quality checks applied, in particular for **crossborder services**.

They should also be encouraged to work further to **establish associated methods of quality measurement and monitoring of the different data types**. Member States should be encouraged to share with **each other their knowledge, experience and best practices** in the data quality field in on-going and future coordination projects.”

Why a MMTIS Quality Framework

Practical need

Current status

- Few countries have a national framework or implement a process for quality checks
 - Different levels of NAP implementation lead to different priorities regarding data quality
 - Different level of experience from each MS regarding data quality methods
- Creates barriers for the usage of the NAP data

NAPCORE contribution

Support the effort of MS to adopt a harmonised approach

- Encourage knowledge, experience and best practices exchange between MS
- Establish common quality dimensions and criteria
- Propose assessment methods and tools



A Quality Framework for Multimodal Travel Information Services (MMTIS)

Report by Task 3.2: European NAPs data quality

Status: draft
Version: 0.x
Date: xx

NAPCORE Methodology



Multimodal Travel Information Services (MMTIS) Quality Package

EU EIP 4.1 Task 2: Determining Quality of
Multimodal Travel Information Services (MMTIS)

Version 2.0

31 October 2019

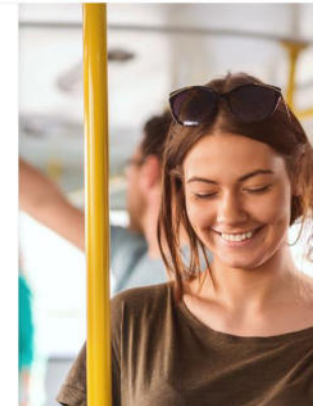


DATA VALIDATION TOOL
VERIFY YOUR NETEX DATA

Data4PT

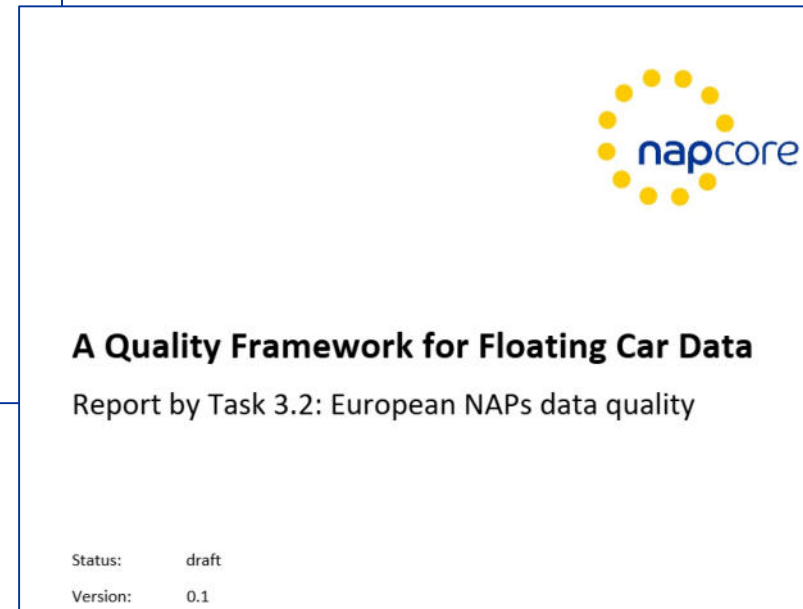
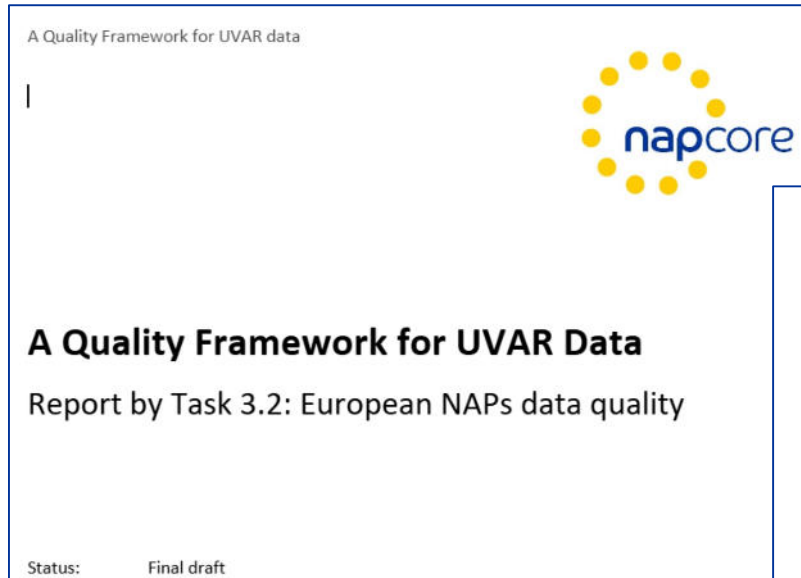
The DATA4PT project aims to advance data-sharing practices in the public transport sector by supporting the development of data exchange standards and models, to fulfil the needs of multimodal travel information service providers.

By supporting EU Member States in deploying a set of harmonised European public data standards (Transmodel, NeTEx and SIRI), DATA4PT wants to



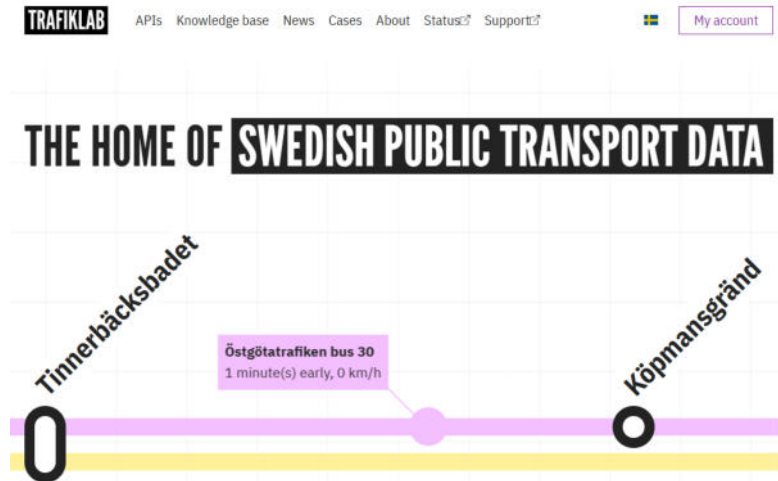
Review of existing documentation
(EIP, DATA4PT, projects, peer-reviewed papers)

NAPCORE Methodology



Alignment across *data categories* Quality Dimensions and Criteria

NAPCORE Methodology

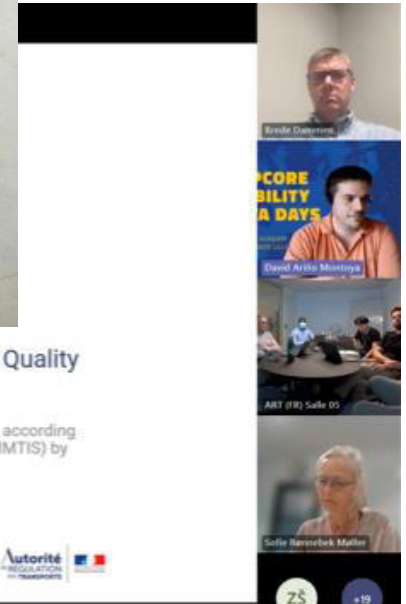


Interviews,
questionnaire
desktop research



3rd Workshop on MMTIS Quality Framework

Control of mobility open data and use according to delegated regulation 2017/1926 (MMTIS) by French transport regulatory body



Online workshops with MS exchanging knowledge about currently adopted practices (Norway, Switzerland, France, Finland, Austria)

Considerations derived from workshops

- Different requirements and different approaches for static (planned) data and dynamic (real-time) data
- Quick reaction to mistakes/errors is important for the reliability to the validation process
- Monitoring data quality at different stages of data pipeline
- Feedbacks loop for data providers and consumers
- Cost of validation for data need to be taken into consideration in the process of data publication

Quality dimensions and criteria

Correctness

- Accuracy, Error rate etc.

Completeness

- Spatial (geographical) coverage, Network coverage, Completeness of dataset etc.

Timeliness

- Freshness, Latency etc.

Reliability and usability

- Logical precision, Uniqueness etc.

Next steps

Map MMTIS data types with specific:

→ quality criteria , metrics (and possible values), possible tools and methods

Quality dimension	Quality criterion	Criterion description	Metrics	Metric description	Mathematical / Theoretical formulation	Possible values		Notes/Comments	DATA TYPES TARGETED
Reliability (technical and conceptual) & usability	Logical precision	Extent to multimodal dataset contains logical errors (e.g., vehicles arrive to final stop before leaving first stop).	Frequency of logical errors (FLE)	Evaluated the frequency with which records present logical errors.	Does the dataset present logical errors?	0: Yes; 1: No			ALL DATA TYPES
	Uniqueness	Indicates whether there are data elements in a dataset which need to be unique, and they are not	Binary (yes/no)	Check unique IDs in a dataset / datafile	Are the datasets / datafiles IDs unique?	0: No; 1: Yes			SPECIFIC DATA TYPES

Quality dimension	Quality criterion	Criterion description	Metrics	Metric description	Mathematical / Theoretical formulation	Possible values		Notes/Comments	DATA TYPES TARGETED
Correctness	Accuracy	Extent to which data is represented in an accurate & precise manner. The real-world situation or conditions it describes without inaccuracies, discrepancies, or mistakes.	Location accuracy – horizontal		This is taken from EU EIP MMTIS QF (page 41): The errors are calculated as: $e_i = \sqrt{(x_{mi} - x_{ti})^2 + (y_{mi} - y_{ti})^2}$ The mean positional uncertainties of the positions are calculated as: $\bar{e} = \frac{1}{N} \sum_{i=1}^N e_i$				LOCATION-related
			Location accuracy – vertical		This is inspired from EU EIP MMTIS QF (page 41): The errors are calculated as: $e_i = z_{mi} - z_{ti}$ The mean positional uncertainties of the positions are calculated as: $\bar{e} = \frac{1}{N} \sum_{i=1}^N e_i$			Jasper Beernaerts: When combining different metrics into one value, cfr the UVAR metrics, this one should have less weight than the horizontal accuracy *not applicable in some use cases	

Thanks

Do you have any questions?

anastasia.founta@itxpt.org

david.arinomontoya@itxpt.org



Cross Domain Quality Evaluation Framework

Mobility Data Days
Turin, 6th November 2024

Petr Bureš, FTS CTU in Prague, TamTam Research



Agenda

1 Introduction – scoping & approach followed

2 Summary of State of the Art (SotA)

3 Quality parameters, grades and scoring

4 Case study DATEX@NAPs (by CERTH-HIT and TamTam Research)

5 Vision 2025 (and beyond)

Horizontal formal and technical quality framework														
Horizontal common features framework														
STTP	SRTI	RTTI	MMTIS	emerging services										
Data category quality package	Data category quality package	Data category quality package	Data category quality package	Data category quality package	Data category quality package	Data category quality package	Data category quality package	Data category quality package	Data category quality package	Data category quality package	Data category quality package	Data category quality package	Data category quality package	Data category quality package

I. Introduction – Scoping

- **Cross-Domain**
 - Every dataset or service **has to be described** (documentation)
 - This is common for any data set or service. More complete = better
 - Provide guidance and best practices for data description
- **Main use cases**
 - Description of data set and service by provider at NAPs
 - Evaluation of data set and service description by provider or by national body
 - Adopting proposed best practices when describing and providing data sets
- **Scope**
 - Framework to assess the quality of information provided **about ANY** dataset

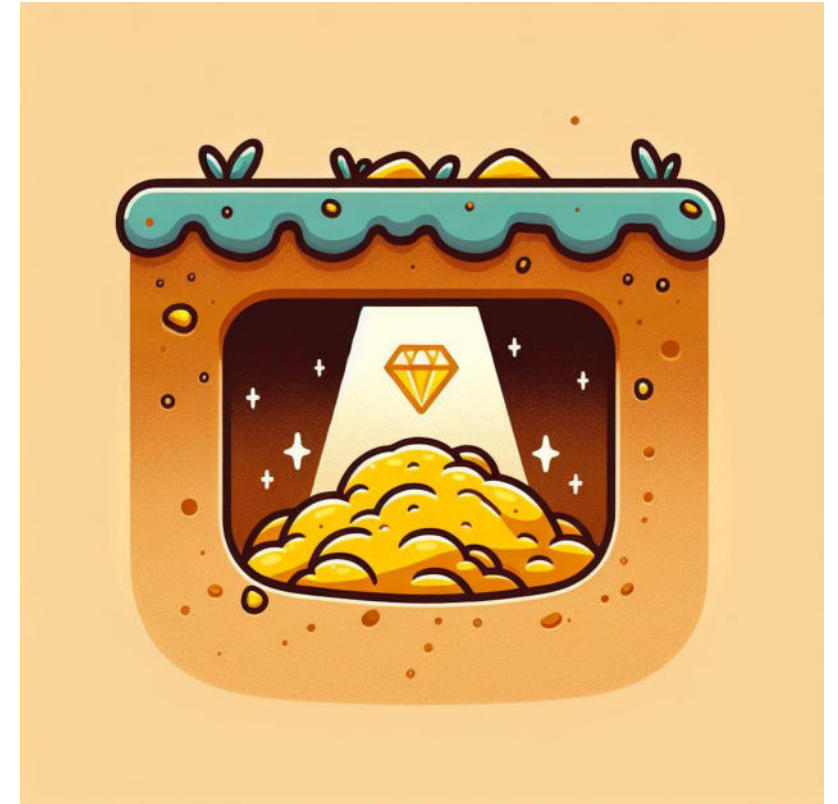
I. Approach followed

- **WHY:** Analysis of the rationale and benefits of the methodology
- **HOW:** Analysis of the current State of the Art
 - Identification of best practices and approaches
 - Choosing a candidate for adoption in NAPCORE
 - Testing and adopting methodology in a closed environment
 - Testing on many live dataset descriptions
 - Analysis of results and methodology improvement (grading)
 - Future challenges

Rationale = why is the methodology relevant?

Metadata (Documentation) quality

- Only some information filed in, only partially
- Only in the local language
- Filed-in fields do not describe the dataset
- Missing samples, schemas, documentation
- Missing classification
- Missing conditions to use and how to get it
- Missing contact information



2. Summary of the SotA

- Framework with automated checks = MQA
 - Used by the EU and by the [Norway data portal](#)
 - Validates 5 aspects of quality
- Framework utilized by NBs / NAPs
 - structure, delivery, availability, metadata
- Framework from [InQMS project](#)
 - evaluation of aspects of data set quality in different phases
 - Validates metadata, documentation as well as the structure of the data set

Metadata quality for the dataset description of
National dataset for public transport information in Norway 46% BETA

[Go to dataset](#)

Criterion	Metadata quality
Accessibility	50%

Wheel monitors

GovData

Publisher: Hanse- und Universitätsstadt Rostock

Updated: 30.10.2022 06:31

Metadata Quality

The Metadata Quality Assurance is intended to help data providers and data portals to check their metadata against various indicators. For information on which metrics we use for indicator measurements, please have a look at our methodology page.

Accessibility		
① Download URL		80 %
① Most frequent accessURL status codes	200	100 %
① Most frequent downloadURL status codes	200	100 %

Reusability		
① Access restrictions		yes
① License information		100 %
① Access restrictions vocabulary		yes
① Contact point		yes
① Publisher		yes

Contextuality		
① File size		80 %
① Rights		100 %
① Datasets: Modification date		yes
① Datasets: Date of issue		yes
① Distributions: Modification date		100 %
① Distributions: Date of issue		100 %

Findability		
① Keyword usage		yes
① Categories		yes
① Geo search		yes
① Time based search		no

Interoperability		
① DCAT-AP compliance		no
① Format		80 %
① Media type		80 %
① Format / Media type from vocabulary		80 %

2. Summary of the SotA

- NAPs are mostly not ready for automatic frameworks (based on a catalog)
- We utilized the InQMS **methodology** as a baseline for development.

- **It evaluates:**

- Documentation
- Structure
- Distribution
- Content
- of the data, in different phases of subscription scenarios
- Manual and automatic

Docs	Structure	Distribution	Content	Evaluation procedure	Phase
✓				M: Completeness of service documentation	Planning
✓				M: Usability of service documentation	Planning
		✓		M: Distribution quality	Staging
	✓			M: Structural quality	Staging
		✓		A: Distribution quality	Production
	✓			A: Structural quality	Production
			✓	A: Content quality	Production
	✓	✓	✓	M: Overall quality of event based service	Ad hoc
	✓	✓	✓	M: Overall quality of status based service	Ad hoc
	✓	✓	✓	M: Content quality	Ad hoc

3. Quality parameters, grading and scoring

- The InQMS provides a **set of methodologies** for complete dataset evaluation
- In **Cross-domain** quality framework we further developed FIRST methodology.
- **Completeness and correctness of data description**
 - Manual inspection and evaluation dataset characteristics as described at NAP
 - The characteristics to be evaluated are mainly, but not entirely **based on CMC** (Coordinated Metadata Catalogue).
 - Each characteristic is extracted from the provided information, ease of extraction is graded (A – F) and commented by the evaluator in the form.
 - Individual grades are combined into the overall score.

3. Quality parameters, grading and scoring

- Individual characteristics grading A to F, are combined based on:
- obligation** and **importance**

<i>obligation</i>	<i>w</i>	<i>importance</i>	<i>w</i>
<i>optional</i>	0,25	<i>normal</i>	1
<i>partial</i>	0,75	<i>important</i>	2
<i>mandatory</i>	1	<i>very important</i>	3
<i>conditional</i>	1		
<i>recommended</i>	0,5		

- Into a **score Q**

overall quality	B	2,16
-----------------	---	------

- $$Q = \frac{\sum_i^n obligation * importance * grade}{\sum_i^n obligation * importance}, \text{ where } n \text{ is a number of characteristics}$$



grade	Value	Description
A	1	Metadata is provided in the expected format with prescribed values as defined by the relevant standard (i.e., CMC). = Parameter value(s) in the metadata exactly as expected.
B	2	Metadata is provided as part of easily identifiable substructures. It does not have to be directly in the expected format (e.g., contact data in different forms, different wordings of the keywords), but it is semantically compatible. = Parameter value(s) in the metadata, in a different form than expected but semantically compatible.
C	3	Metadata is not provided directly but is inferable from the collected evidence. It does not “fit” in terms of format and structure. = Parameter value(s) are not in the metadata, but findable elsewhere and semantically compatible.
F	4	Metadata is not provided and cannot be “easily” derived from the collected evidence (including samples and schemas provided). = Parameter value(s) are not found/filled in. Assumptions and guessing are needed to fill in the item, and correctness is not assured.
-	0	(Not relevant) The item is not evaluated because it makes no sense in the context.

3. Quality parameters, grading and scoring

- Groups of evaluated parameters
 - Metadata information (4 parameters, [CMC:4])
 - Content information (10 parameters, [CMC:7])
 - Temporal information (2 parameters, [CMC:2])
 - Geographical coverage (3 parameters, [CMC:3])
 - Transportation system (1 parameter, [CMC:1])
 - Responsibilities (2 parameters, [CMC:2])
 - Conditions for use (2 parameters, [CMC:2])
 - Access information (12 parameters, [CMC:8])
 - Quality information (9 parameters, [CMC:3])
- Total 45 evaluated parameters in 9 groups

Category	Parameter Name	id	oblig	evaluat	data provided or implied	yes	comment
Overall quality				B	samples, schema and detailed documentation of format and protocol provided. However many of the parameters required by CMC were either completely missing or had to be derived from textual descriptions.		
Evaluation metadata	evaluated dataset url			-	https://register.dapreventinfo.cz/en/so-urces/cz-ndic_d2-common/		THIS is a sample evaluation done for updating and evaluation of the assessment. It does not have to be complete nor TRUE!
	evaluation date			-	01.12.2019		
	evaluator			-	Petr Remeš		
Metadata information	Metadata date	2.2.1.1	mandatory	F	01.12.2019	yes	data formally not provided, not part of the metadata set. Date was derived from date accessed (web page)
	Metadata language	2.2.1.2	mandatory	B	cz-cyber	yes	not part of the metadata set. Data was derived from context (web page)
	contact point	2.2.1.3	optional*	B	title: Manager of traffic information distribution department phone: +420-556663578		data derived from the web, provider information section, it is however unclear if this contact is also metadata provider (the person who put metadata together)
	dataset identifier		recommended	A	x-source-cz-ndic_d2-common		dataset url
	name of the dataset	2.2.2.1	mandatory	A	Common		within a dataset specification (cz-cyber)
	description of dataset	2.2.2.2	mandatory	A	Common traffic information (accidents etc.), provided by NDIČ for the whole Czechia, pushed on occurrence, in a common traffic information profile of DATEX II format.		within a dataset specification (cz-cyber)
	resource type	2.2.2.3	mandatory	A	Data set		Directly in metadata as literals (listed in a table)
	dataset type category	2.2.2.4	mandatory	A	road work information, Unexpected road events and conditions, Real-time traffic data		Directly in metadata as literals (listed in a table)
	dataset detailed type	2.2.2.5	optional	A	Long-term road works, Short-term road works, Accidents and incidents, Poor road conditions, Weather conditions affecting road surface and visibility, Location and length of queues		Directly in metadata as literals (listed in a table)
	service type category	2.2.2.6	conditionally*	-			not applicable for dataset only services. The value is however not part of the metadata since services are not described within metadata repository.
	dataset language	2.2.2.7	mandatory	C	cz-cyber	yes	not part of the metadata set. Derived from context, samples in format documentation.
	georeferencing method	2.2.2.8	recommended	B	coordinates, linearReferenceElement, AlertCLineae, globalNetworkLineae	yes	not part of the metadata set. Derived from format documentation and samples in documentation. Only linear locations are used, extension for local road network and local coordinates
	where documentation		recommended	A	https://register.dapreventinfo.cz/en/so-urces/cz-ndic_d2-common/		available at web as html, consisting of format and protocol documentation
	related linked data sources		recommended	A	x-source-cz-ndic_d2-geodetected-location-set		available as parameter "Required Data Sources"
Temporal information	start date of publication	2.2.3.1	mandatory	B	01.01.2018		available as a note
	end date of publication	2.2.3.2	optional	-	-		not provided for dataset
Geographical coverage	area covered by publication	2.2.4.1	mandatory	B	C2D	yes	derived from region parameter, however not in required format (NUTS)
	network coverage	2.2.4.2	mandatory	F	Motorways, Arterial road network, Regional roads, Urban and local roads	yes	not part of the metadata set. Derived from context, description of the datasets.
	network coverage description	2.2.4.3	optional	F			not part of the metadata set.
Transportation system	transportation modes covered	2.2.5.1	mandatory	F	Personal / car, truck	yes	not part of the metadata set. Derived from context, description of the dataset.
	publisher	2.2.6.1	optional*	A	title: Manager of traffic information name: Česká Wescoteln title: Manager of traffic information distribution department phone: +420-556663578	yes	provider information section
Responsibilities	data owner	2.2.6.2	optional*	B			implied from provider section
Conditions for use	contract or license	2.2.7.1	mandatory	B	License and free of charge		linked from each data source, separate, part of the provider description, not in the required format however.
	conditions for use	2.2.7.2	conditionally*	A	http://register.dapreventinfo.cz/en/so-urces/cz-ndic_d2-common/		url and the process description provided
	data format - encoding	2.2.8.1	optional	C	UTF-8	yes	implied from formal documentation section, validation schema

3. Quality parameters, grading and scoring

- Overall Quality = grade and comment
- information about the evaluation (not evaluated :-)
- **category, name, chapter ref., obligation** Exactly as specified by CMC

overall quality				B	samples, schemas and detailed documentation of format and protocol provided. However many of the parameters required by CMC were either completely missing or had to be derived from textual descriptions.		
Category	Parameter Name	ref.	obligation	evaluation	data provided or implied	implied ?	comment
Evaluation metadata	evaluated dataset url			-	https://registr.dopravniinfo.cz/en/sources/cz-ndic_d2-common/		THIS is a sample evaluation done for updating and evaluation of the assessment. It does not have to be complete nor TRUE!
	evaluation date			-	01.12.2019		
	evaluator			-	Petr Bureš		
Metadata information	Metadata date	2.2.1.1	mandatory	F	01.12.2019	yes	data formally not provided, not part of the metadata set. Date was derived from date accessed (web page)
	Metadata language	2.2.1.2	mandatory	B	cze;eng	yes	not part of the metadata set. Date was derived from context (web page)
	Contact point	2.2.1.3	partially*	B	title: Manager of traffic information distribution department phone: +420-596663578		data derived from the web, provider information section, it is however unclear if this contact is also metadata provider (the person who put metadata together).
	Dataset identifier		recommended	A	x-source:cz-ndic_d2-common		dataset uri

3. Quality parameters, grading and scoring

- Content information (10 parameters, [CMC:7])

Category	Parameter Name	ref.	obligation	evaluation	data provided or implied	implied ?	comment
Content information	Name of the dataset	2.2.2.1	mandatory	A	DATEX II Situation Publication - Common		within a dataset specification [cze;eng]
	Description of dataset	2.2.2.2	mandatory	A	Common traffic information (accidents etc.), provided by NDIC for the whole Czechia, pushed on occurrence, in a common traffic information profile of DATEX II format.		within a dataset specification [cze;eng]
	Resource type	2.2.2.3	mandatory	A	Data set		Directly in metadata as literals (listed in a table)
	Dataset type category	2.2.2.4	mandatory	A	Road work information, Unexpected road events and conditions, Real-time traffic data		Directly in metadata as literals (listed in a table)
	Dataset detailed type	2.2.2.5	optional	A	Long-term road works , Short-term road works, Accidents and incidents , Poor road conditions, Weather conditions affecting road surface and visibility, Location and length of queues		Directly in metadata as literals (listed in a table)
	Service type category	2.2.2.6	conditionally*	-			not applicable for dataset only services. The value is however not part of the metadata since services are not described within metadata repository.
	Dataset language	2.2.2.7	mandatory	C	cze	yes	not part of the metadata set. Derived from context, samples in format documentation.
	Georeferencing method	2.2.2.8	recommended	B	coordinates, linearWithinLinearElement, alertCLinear, globalNetworkLinear	yes	not part of the metadata set. Derived from format documentation and samples in documentation. Only linear locations are used, extension for local road network and local coordinates
	Dataset documentation		recommended	A	https://registr.dopravniinfo.cz/en/sources/cz-ndic_d2-common/		available at web as html, consisting of format and protocol documentation
	Related linked data sources		recommended	A	x-source:cz-ndic_d2-predefined-location-set		available as parameter "Required Data Sources"

3. Quality parameters, grading and scoring

- Temporal information, Geographical coverage, Transportation system, Responsibilities and Conditions for use

Category	Parameter Name	ref.	obligation	evaluation	data provided or implied	implied ?	comment
Temporal information	Start date of publication	2.2.3.1	mandatory	B	01.01.2018		available as a note
	End date of publication	2.2.3.2	optional	-	-		not provided for dataset
Geographical coverage	Area covered by publication	2.2.4.1	mandatory	B	CZ0	yes	derived from region parameter, however not in required format (NUTS)
	Network coverage	2.2.4.2	mandatory	C	Motorways, Arterial road network, Regional roads, Urban and local roads	yes	not part of the metadata set. Derived from context, description of the dataset.
	Network coverage description	2.2.4.3	optional	F			not part of the metadata set.
Transportation system	Transportation modes covered	2.2.5.1	mandatory	C	Personal / car, truck	yes	not part of the metadata set. Derived from context, description of the dataset.
Responsibilities	Publisher	2.2.6.1	partially*	A	title: Manager of traffic information name: Ctirad Weissmann		provider information section
	Data owner	2.2.6.2	partially*	B	title: Manager of traffic information distribution department phone: +420-596663578	yes	implied from provider section
Conditions for use	Contract or license	2.2.7.1	mandatory	B	Licence and Free of charge		linked from each data source, separate, part of the provider description, not in the required format however.
	Conditions for use	2.2.7.2	conditionally*	A	http://registr.dopravniinfo.cz/en/providers/cz-ndic/#process-ddrcontract		url and the process description provided

3. Quality parameters, grading and scoring

- Access information (12 parameters, [CMC:8]), with data sample and schema

Category	Parameter Name	ref.	obligation	evaluation	data provided or implied	implied ?	comment
Access information	Data format - Encoding	2.2.8.1	optional	C	UTF-8	yes	implied from format documentation section, validation schema
	Data format - Syntax	2.2.8.2	mandatory	C	XML	yes	implied from format documentation section, validation schema
	Data format - Grammar	2.2.8.3	optional	C	XSD	yes	implied from format documentation section, validation schema
	Data format - Data Model	2.2.8.4	mandatory	C	DATEx II profile	yes	implied from format documentation section, validation schema
	Data sample		recommended	A	https://registr.dopravniinfo.cz/docs/x-format/cz-ndic_d2-common-v1.1-cs-html/examples.html		provided also as separate files
	Data schema / validation		recommended	A	https://registr.dopravniinfo.cz/docs/x-format/cz-ndic_d2-common-v1.1-cs-html/format.html		provided also as separate file
	Data format description	2.2.8.5	conditionally*	B	https://registr.dopravniinfo.cz/en/formats/cz-ndic_d2-common-v1.1/		provided also as separate file, comprehensive description including concepts, samples and schema.
	Access interface	2.2.8.6	mandatory	B	HTTP/HTTPS	yes	derived from protocol documentation section
	Security mechanisms		recommended	C	IP filter	yes	derived from protocol documentation section
	Communication method	2.2.8.7	conditionally*	A	Push on occurrence		provided as name and type of the protocol
	Protocol documentation		recommended	A	https://registr.dopravniinfo.cz/en/protocols/cz-ndic_push-v1.0/		provided also as separate file
	Access URL	2.2.8.8	conditionally*	B	https://datex.rsd.cz		provided as part of the providers information, a separate distribution server
	Update frequency	2.2.9.1	mandatory	C	On occurrence		derived from free text data description

3. Quality parameters, grading and scoring

- Quality information (9 parameters, [CMC:3]), with characteristics from Quality parameters (EU-EIP, ISO), problem reporting, reaction time, etc.

Category	Parameter Name	ref.	obligation	evaluation	data provided or implied	implied ?	comment
Quality information	Update frequency	2.2.9.1	mandatory	C	On occurrence		derived from free text data description
	Availability		recommended	F	-	yes	not provided within the documentation
	Data collection method		recommended	C	agenda systems, sensoric systems, call centrum, operators	yes	data collection mehod described partially in the text
	Cross verification		recommended	C	partially, some of the information are manuall checked by operators	yes	cross verification described partially in the text
	Problem reporting process		recommended	F	-		not provided within the documentation
	Reaction time to respond		recommended	F	-		not provided within the documentation
	Quality description	2.2.9.2	mandatory	F	quality information is unknown		quality information is unknown
	National body assessment status	2.2.9.3	optional	F	-		"No assessment done", this information was not however provided
	Applicable DRs		recommended	A	SRTI, RTTI		references to fulfilled delegatred action provided, it is not clear to what degree though.

- Filling in the EVAL FORM serves as a gathering of information, provides guidance WHAT is expected from providers (NAPs), and as a reference.

4. Case study DATEX@NAPs

- **DATEX@NAPs** study, by WG3.3, real status of
 - existence of **NAPs** and availability of **catalog records**,
 - Metadata completeness and availability and clarity of documentation provided for both formats and protocols,
 - ability to subscribe to data distributions offered by various **publishers**,
 - technical feasibility for consumers to access and use the data,
 - technical quality of the content, including whether it conforms to the provided **schema**.
- Methodology: **Cross-domain Quality Evaluation Framework**

DATEX@NAPs

An In-Depth Look at DATEX II on National Access Points

Introduction

We are writing a travel journal from our journey across European countries and their National Access Points (NAPs), where we search for something delicious — DATEX II Safety...



Key Concepts

This page provides an overview of the most relevant concepts and terms in our study.



Quick Start

This chapter shall help the inpatient readers willing to use this website for something practical as soon as possible.



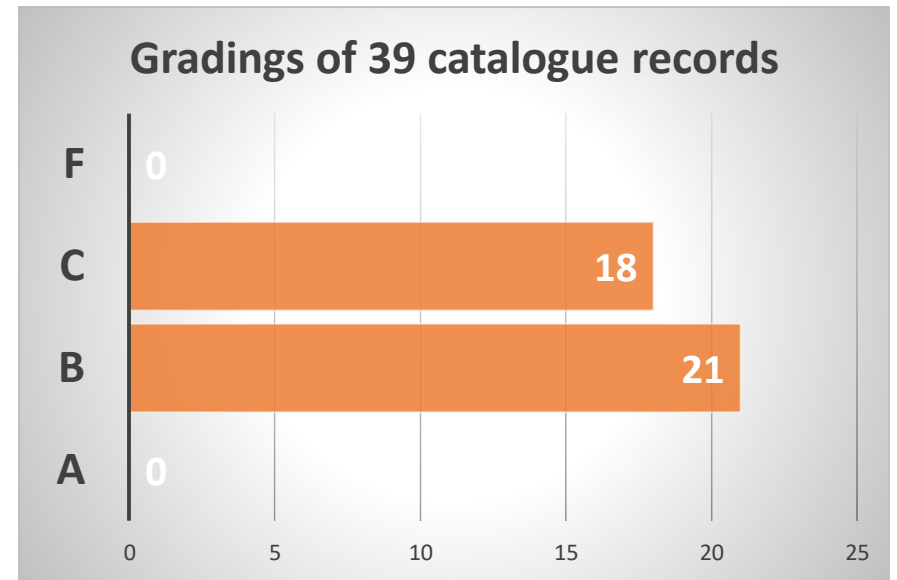
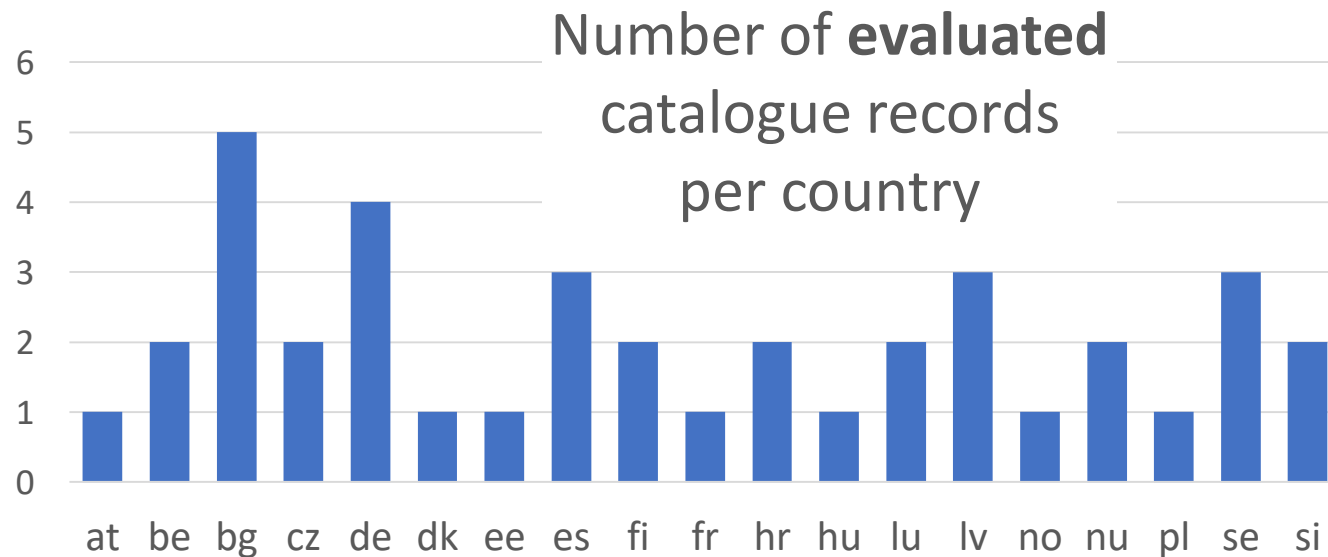
NAPs Browser

Concepts are structured as tree of resources (parents owning subresources):



4. Case study DATEX@NAPs

- Scope of the Study: **SRTI**
- Evaluation of **39 catalog records**, from **19 countries**
- In some countries data were not available
- Some characteristics, were not met in many cases



4. Case study DATEX@NAPs

- The following tables show **grades** of evaluated catalog records per **country** and their **score**

B	B	B	C	C
B	B	B	C	B
B	B	B	B	B
C	B	B	C	B
C	C	B	C	C
C	C	B	C	C
C	C	B	B	C
C	C	B	B	

at	hr	de	lu	ee
be	hr	de	lu	es
be	cz	de	nu	es
bg	cz	de	nu	es
bg	dk	hu	no	se
bg	fi	lv	pl	se
bg	fi	lv	si	se
bg	fr	lv	si	

2,1	1,8	1,9	3,1	2,9
2,1	1,8	2,1	3,1	1,9
2	1,6	2,1	2,4	2
2,9	1,6	1,9	2,7	2,1
2,9	3	2,3	2,7	2,5
2,9	3,1	2,2	2,8	2,5
2,9	3,1	2,3	1,8	2,6
2,9	3,1	2,3	1,8	



1
2
3
4

- Details of evaluations are available at <https://datex2.naps.inqms.cz/>

5. Future work

- Improve methodology by the findings of the DATEX@NAPs study
- Adapt the methodology to mobilityDCAT-AP

Further work (also scope of the DATEX@NAPs study)

- Adopt other parts of cross-domain methodology
 - Fitness of data considering the provided description
 - Validation of the data/service against provided information
 - Validation against the provided schema, availability of the data
 - Validation of the technical content of the distribution

Thanks

Do you have any questions?

Want to join the work?

petr.bures@tamtamresearch.com



<https://datex2.naps.inqms.cz/>

<https://inqms.cz/>

Information on parts of the InQMS framework

- Providing best practices in data description, and data content:
- Converted and adapted methodology
 - Completeness and correctness of data description
- Methodologies in the queue
 - Completeness and correctness of data description
 - Fitness of data considering provided description
 - Validation of the data/service against provided information,
 - Validation against provided schema,
 - Validation of the technical content of the distribution
- Next slides provides examples of those further methodologies

2. Fitness of data considering provided description

- Covering relevancy of provided information and best practices for referred data / service,
- Considering availability of samples and schema and that it validates
- Considering data / service content: used location referencing, event description + event impact, format and provision mechanisms

Evaluation of objective documentation quality

v0.2

overall quality	B	samples validates against provided schema, profile exists, documentation is relevant, however some datasets used in source are not available to public and publication by HTTP has low security mechanism and does not follow standard.
-----------------	----------	---

Category	Element Name	evaluation	description
Evaluation metadata	provider contact	-	https://registr.dopravniinfo.cz/en/providers/cz-ndic/
	Provider dataset name	-	DATEX II Situation Publication - Common
	evaluation date	-	01.12.2019
	evaluator	-	Petr Bureš
	comment	-	evaluation done for research purposes not provided, downloaded from www.registr.dopravniinfo.cz

Category	Element Name	evaluation	description
validation objective	Metadata quality	B	samples, schemas and detailed documentation of format and protocol provided. However many of the parameters required by CMC were either completely missing or had to be derived from textual descriptions.
	Is the source uniquely identified?	A	unique identifier of the source exists
	Is sample available?	A	there are 3 samples present in documentation as separate files, covering different aspects of data.
	Is scheme available?	A	there is schema present in documentation as separate file, the schema is a profile, not just whole model.
	Samples validate against the schema	A	All provided samples validate against provided schema
	are all required datasets identified and available?	C	documentation and samples identify datasets used for location description, it is: - TMC location table (description how to get it is included) - Global Network road data (private dataset available only to public authorities)
	Relevance of the referenced documentation	B	Documentation describes in detail formats with samples and schemas and protocols and procedures how to get data. Description of a source lack specific information that must be inferred from samples.
	usable description of data subscription	A	Exact procedure how to get the data is described, licence conditions are available to download.
best practice objective	What combination of location descriptions is used?	B	ALERT-C is used wherever possible for location description, coordinates are however described in local S-JTSK reference standard, OpenLR is not available and location description extension uses a private set of reference data.
	How does the event textual description look like?	B	Text is only in CZ, sometimes hard to comprehend, only in one comment without the specification of content.
	What data format is used?	A	DATEX II v2.3 is used, profile with extension for local road network localization.
	What is the method of data provision?	C	Described in separate document, uses HTTP, FTP, email delivery. Security is provided by optional IP filtering no authentication used for HTTP.
	Is impact described within event phrases?	A	Event is described by a single phrase and impact where possible. Is used in sample, described in concepts and allowed in schema.

3. Validation of the data/service against provided information

- Covering ease of the service consumption, update time
- Considering technical aspects of data delivery (safety and security, content type, compression, conditional requests etc.)

Evaluation of web service distribution quality

v0.2

overall quality	C	administrative proces even thpough sufficiently described requires paper trail and takes long time, subscription process requires email
-----------------	---	---

Category ▾	Element Name ▾	evaluation ▾	description ▾
Evaluation metadata	provider contact	-	https://registr.dopravniinfo.cz/en/providers/cz-ndic/
	Provider dataset name	-	DATEX II Situation Publication - Common
	evaluation date	-	01.12.2019
	evaluator		Petr Bureš
	comment	-	-

Category ▾	Element Name ▾	evaluation ▾	description ▾
primary evaluation	type of administrative process	Licence and Free of	non standard free licence available at NAP pages
	administrative process evaluation	C	licence is available at the NAP, must be downloaded, printed, signed and sent via post to provider, after countersigning the subscription could be set up. Data subscription is set up partly through a web interface and partly via email communication with subscription manager. During the subscription process user is not informed about any development. Communication only in czech language. External support may be required.
	administrative process duration	less than a month	the paper trail takes most of the time to process. Data subscription done via email + web is reasonably quick (a week).
	service delivery mechanism	push	HTTP push mechanism is used
	PULL subscription address	-	not relevant, PUSH mechanism is in place
	delivery protocol evaluation	C	(-) uses obsolete delivery mechanism with no security and compression, (+) frequent pushes with small average size
	safety and security level	none	no inherent security is provided, data is pushed to consumers server in plain http, optionally consumer could set up IP filtering.
	declared content type	other	wroong content type is provided by the push mechanism
	content compression (gzip)	no	data are provided in a uncompressed way, push served does not provide for gzip functionality
	conditional request	not relevant	HTTP push mechanism is used
	update interval [min]	2,5	data is pushed on occurence, during 1h observation period 2,2 +- 1,2 minutes average interval was measured
	message envelope	soap	standard soap envelope is used in provided XML
	average message size [kB]	26	during 1h observation period 26 +- 18 kB, max 98 kB and min 3 kB was measured

3. Validation of the data/service against provided information

- Covering ease of the service consumption, richness of information
- Considering all aspects evaluated on sample but now in reality

Evaluation of technical information quality in web service

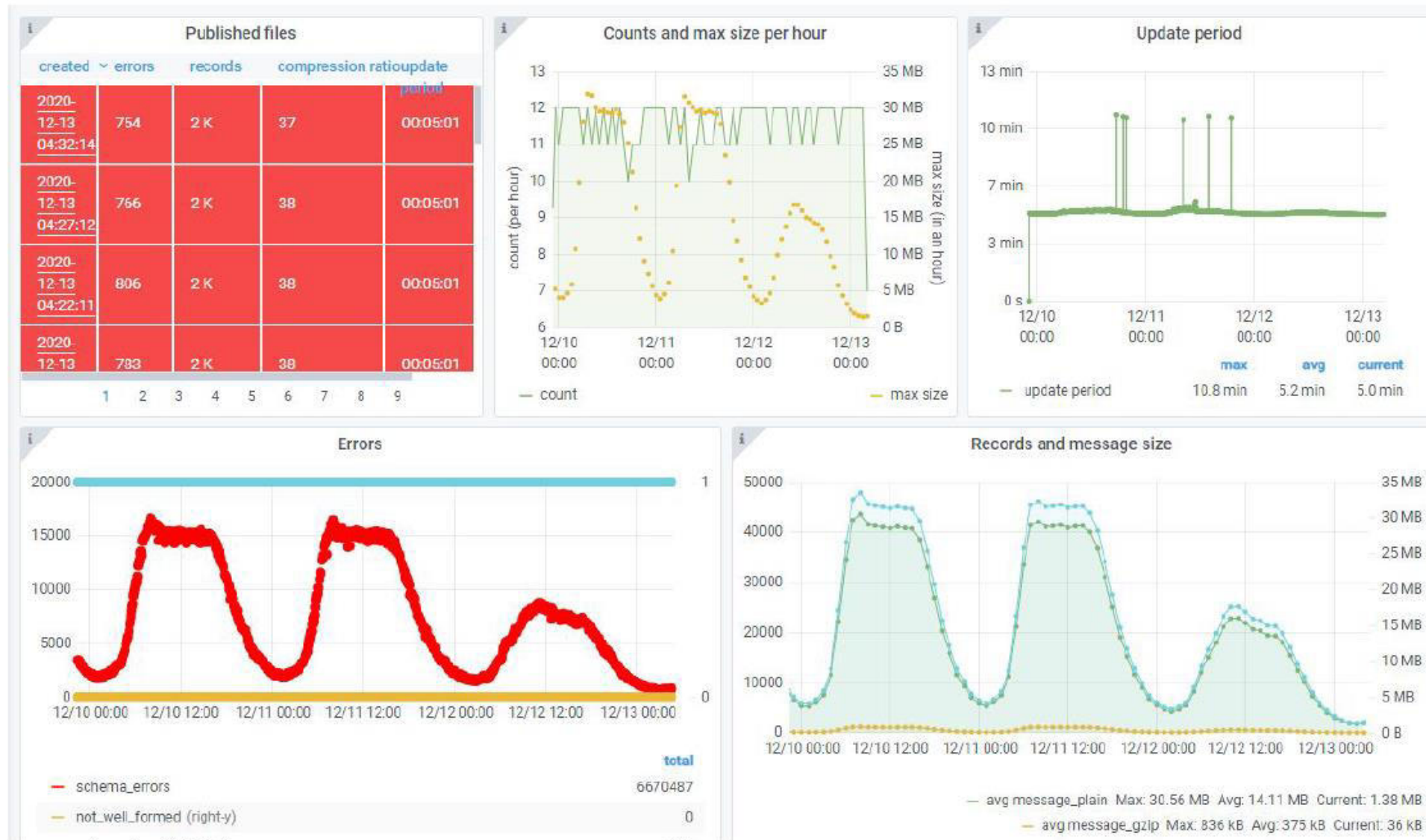
v0.2

overall quality	C	samples does not validate against schema, one location dataset used extensively for all messages (sometimes solely) is not available, description in czech only and impact sometimes misleading.
-----------------	---	--

Category ▾	Element Name ▾	evaluation ▾	description ▾
Evaluation metadata	provider contact	-	https://registr.dopravniinfo.cz/en/providers/cz-ndic/
	Provider dataset name	-	DATEX II Situation Publication - Common
	evaluation date	-	01.12.2019
	evaluator	-	Petr Bureš
	comment	-	

Category ▼	Element Name ▼	evaluation ▼	description ▼
primary content evaluation	Samples validate against the schema	F	errors present. Soap envelope must be removed before running the test
	Are all required datasets identified and available?	C	received samples uses extended location referencing globalnetwork dataset without possibility to get this dataset. (not problem when this is not the only method)
	What combination of location descriptions is used?	C	samples uses following methods as expected from documentation: supplementaryPositionalDescription (when makes sense), alertCLinear (where possible), linearWithinLinearElement (on highways) and globalNetworkLinear extension (in all messages). There are no WGS84 coordinates in the dataset, only SJTSK in globalNetworkLinear extension. Also globalNetworkLinear is used alone without any additional methods, making location referencing impossible.
	How does the event textual description look like?	B	quite descriptive, but only in czech and without discrimination by event and location.
	Is impact described within event phrases?	B	impact is used wherever possible, sometimes however does not make sense (1 of 2 lanes closed, 2 lanes remaining?). (+) Closed is modelled as capacity remaining 0,

4. Validation against provided schema, availability of the data



5. Validation of the technical content of the distribution

Following parameters can be **checked in live feed** to assess its fitness:

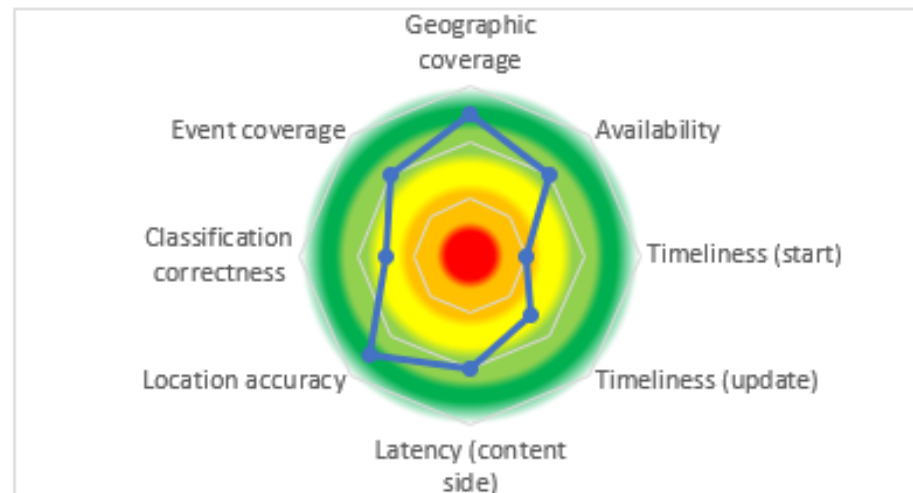
- How much data is published regularly,
- Distribution interval of new files.
- Total number of events in effect at the end of the interval.
- Total number of events received per unit of time.
- Average "message lifetime" for messages in the monitored interval.
- The number of new, existing, updated and cancelled messages, or messages that expired during the interval.
- Number of valid events by event type and location type.
- Average time between the provision and the start of the event.
- The ratio of events that are terminated explicitly / expire automatically

Extra – overall service quality

- Based on quality objects and EU-EIP frameworks

Evaluation of overall service quality

v0.2



overall assesment

B

The service has very good geographical and event coverage, availability and accuracy. Room for improvement is in timeliness updates and technical processing of supplied data that lead Isometimes to missclassification

Category	Element Name	evaluation	description
Evaluation metadata	provider contact	-	https://registr.dopravniinfo.cz/en/providers/cz-ndic/
	Provider dataset name	-	DATEX II Situation Publication - Common
	evaluation date	-	01.12.2019
	evaluator	-	Petr Bureš
	commnet	-	

Category ▼	Element Name ▼	evaluation ▼	description ▼
overall content evaluation	Service type	event oriented - only ghost driver	the service evaluated is basic and most comprehensive service of NTIC covering event oriented and a selected few of status oriented data.
	Geographic coverage	**** (expert)	the service covers 100 % of TERN and above it all national 1st-3rd class roads and selected urban streets based on where the event is reported it can be manually added on all possible road locations.
	Availability	*** (advanced)	the service is available 99,9 % of the year, nonstop operator operation and inclusion to technical backbone of state relevant security services.
	Timeliness (start)	* (basic)	best effort, expert assessment: the accidents and incidents are reported after police is at the scene, until then a probability measure is given, start of road closures is by planned time not by actual time the workers started to close the road.
	Timeliness (update)	** (enhanced)	best effort; expert assessment: the accidents are reported by police as cleared but this could happen when police starts to sort administrative task (while saving lives is the top priority) also firebrigade cleans the spot after the accident, so accidents sometimes tend to have fixed duration time in 1h increments. The road closures are reported closed by planned time, not by actual work, updates of the plan are resolved in office, so delay is expected. Call center could alleviate this to a certain degree.

Category ▼	Element Name ▼	evaluation ▼	description ▼
Evaluation	Latency (content side)	*** (advanced)	advanced level; information are processed under 3 minutes time, information from authorised sources come through immediately after its acceptance and cross validation in the system.
	Location accuracy	**** (expert)	expert level; expert assesment: Location of start and end of the event has accuracy in meters, this is however reported in SJTSK coordinate system in location reference method that relies on private data set (GlobalNetwork), location accuracy for services using ALTER-C is between two location points, (intersections, selected milleage posts on highways, etc). Location on a milleage is provided on highways.
	Classification correctness	** (enhanced)	enhanced level; while classification of event type and location accuracy levels are very advanced, the technical systems sometimes assign road segments in wrong way making location referencing impossible, classification of time aslo suffers from unreported end times.
	Event coverage	*** (advanced)	> 80 % of all relevant events, expert assesment: not all events are reported to police, call center is not the main channel of communication with road users, some of the events can be misclassified and therefore not count into coverage criteria.