



Health Data Access Body-NL





Work Package 9

Data Quality

Public Summary Milestone 30 Data Quality Labelling and Maturity in the Netherlands
October 2025

CC-By © HDAB-NL

Samen gezond, fit en veerkrachtig

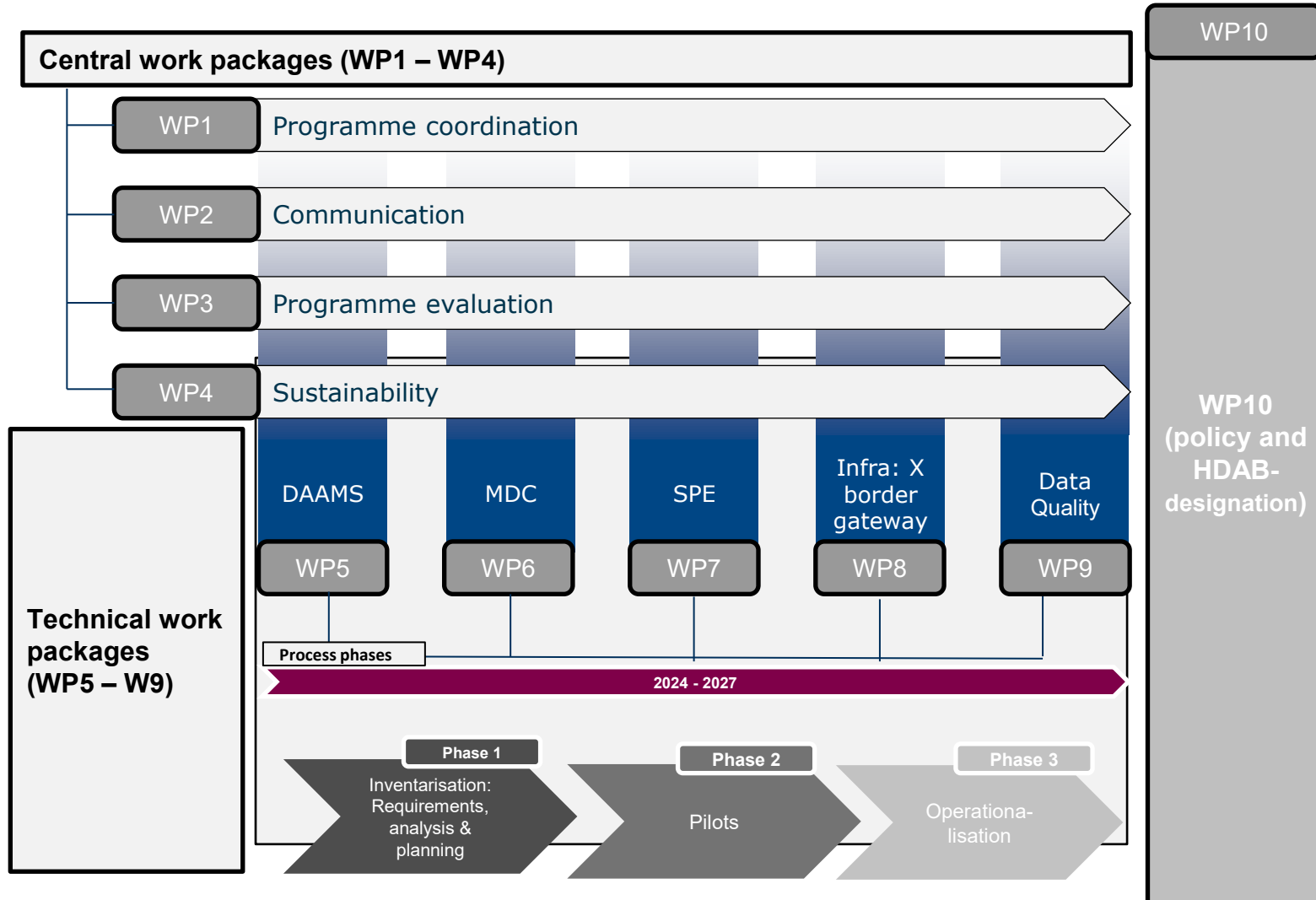
Programme HDAB-NL (2024 – 2027)

HDAB-NL programme:

- 4 central (c.q. horizontal) work packages (WP 1 t/m 4)
- 5 technical workpackages (WP 5 t/m 9)
- 1 policy work package (WP 10)

The technical work packages 3-phase approach:

1. Inventarisation (collect requirements and preferences, diverge)
[Jan. 2024 – March 2025]
2. Translate into concrete MVPs (converge, pilots)
[April 2025 – Aug. 2026]
3. Operationalisation
[Aug. 2026 – Nov. 2027]



HDAB-NL

Prepares important Building Blocks:

- Governance / processes
- Digital capabilities: DAAMS, SPE, MDC, X-gateway (NCP)



Ministry of Health, Welfare and Sport



National Institute for Public Health
and the Environment
Ministry of Health, Welfare and Sport



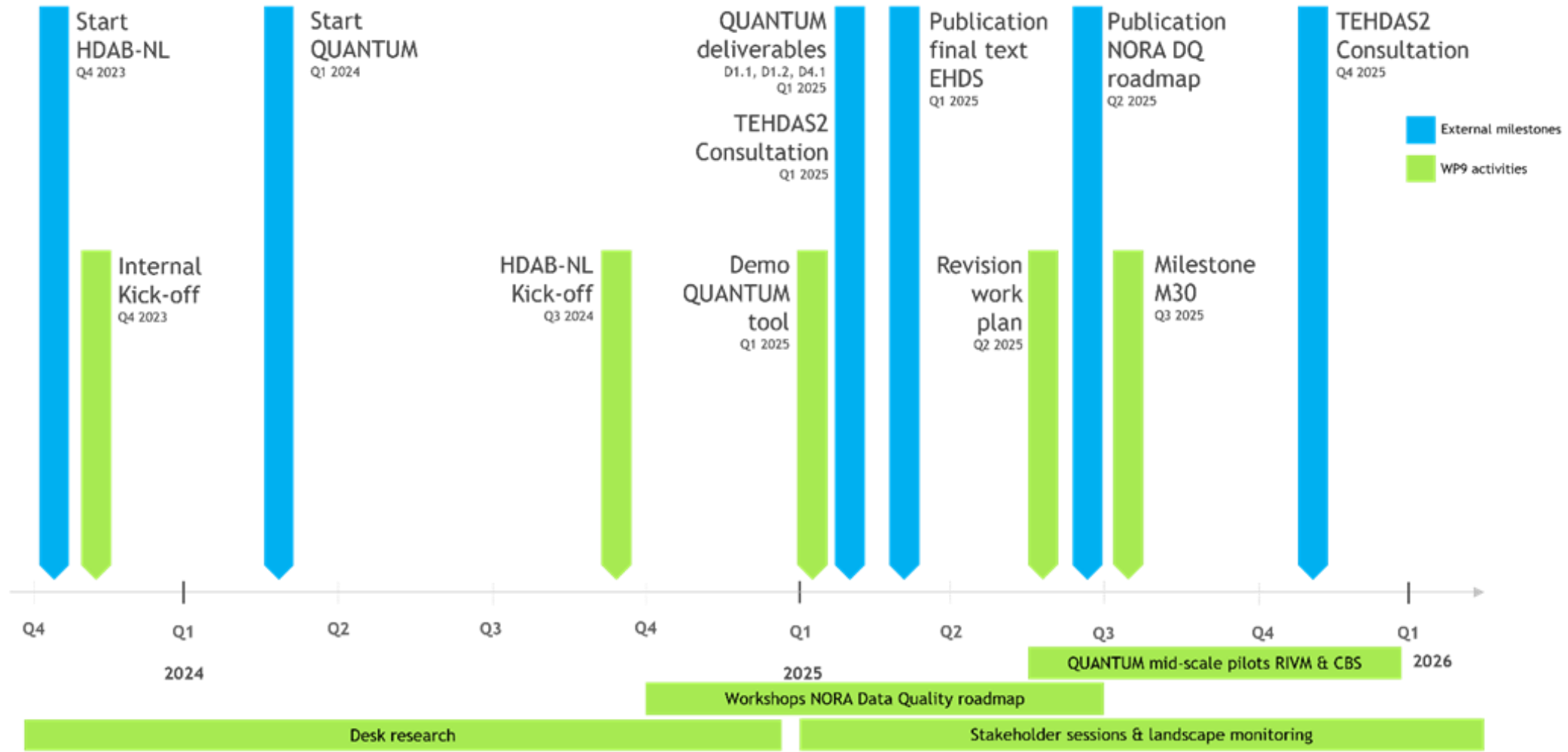
Milestone 30 Work Package 9: Data Quality labelling and maturity in the Netherlands

- First milestone
- Date of publication: 30 July 2025
- Publication: not public

- Summary of activities from Dec 2023 till June 2025

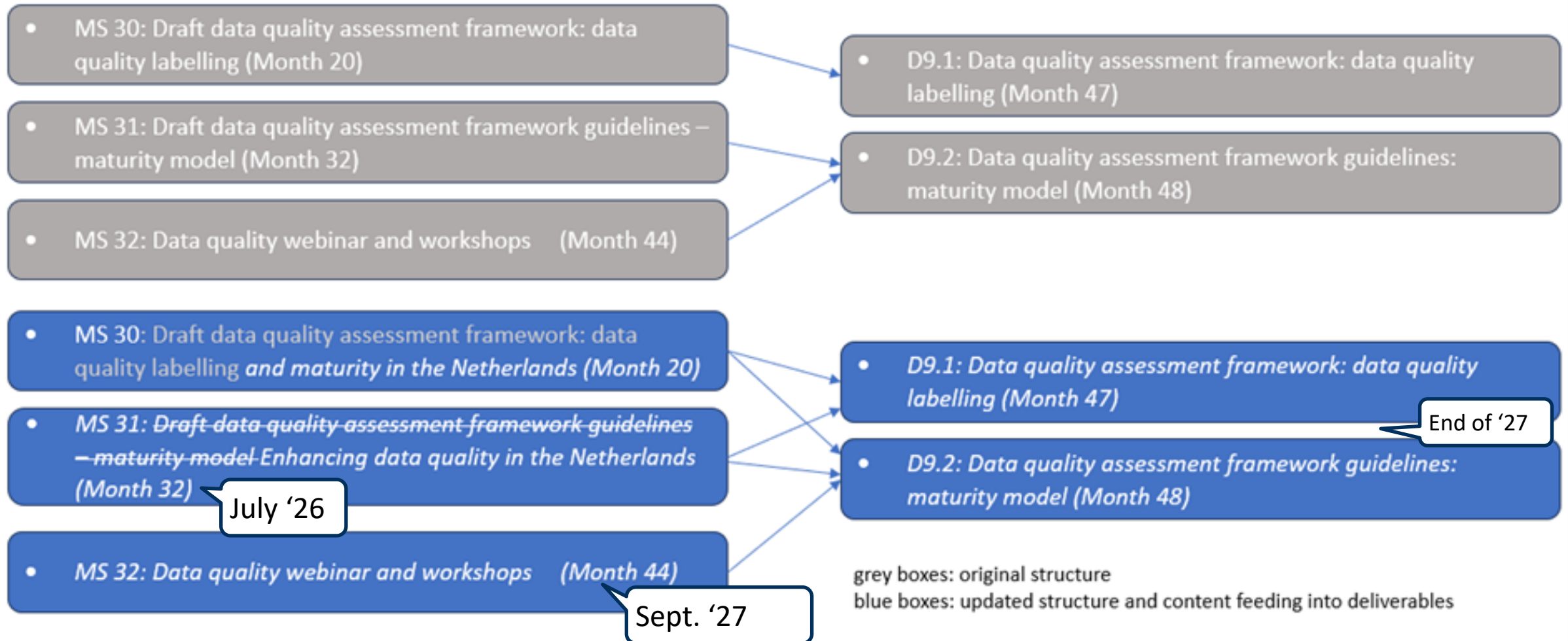
- **Original goal work package 9:** develop a data quality framework for a data quality label and maturity, support improvement of health data quality with guidelines, workshops and webinars.
- Focus changed because of the activities and products created by the European funded QUANTUM project.
- **New goal work package 9:** Usability evaluation of current data quality frameworks (a.o. from the [QUANTUM project](#)) in the context of the Dutch stakeholders (data holders, data users, HDAB) and the EHDS implementation.

Timeline context WP 9



Data Quality HDAB-NL Context (external milestones in blue). Work Package 9 activities in green.

Milestones en deliverables WP9



Original and planned milestones and deliverables of Work Package 9. The milestones are not public, the deliverables will be public and expected end of 2027

What can we expect in the short term?

- **25 March 2025: [EHDS](#)**
 - Data Quality mentioned 27 times. Most relevant articles 51 (minimum categories) and 78 (data quality and utility)
 - A digital infrastructure for secondary use is being built: HealthData@EU. Next to the infrastructure for primary use Health@EU.
 - The HDABs are national contact points for secondary use.
 - The HDAB will provide the local infrastructure (metadata catalogue, SPE, DAAMs and X-border gateway support connection to HealthData@EU) and support to the local stakeholders.
 - **26 Maart 2027: implementing Act** EHDS for dataset descriptions.
- **August 2026:** end of the **QUANTUM** project – EU project. Development of the Data Quality and Utility framework and labelling tool, create insight in implementation challenges, data holders' data quality maturity and development of support materials. Currently 3 deliverables published. The second mid-scale test pilot will start end of October 2025.
- **[TEHDAS2 EU project](#) – ends December 2026** – TEHDAS2 is developing guidelines and technical specifications to enable seamless use of health data across Europe under the upcoming European Health Data Space (EHDS). [Public consultations](#) are used to ensure final output meets the needs of stakeholders. Data Quality is not a separate guideline as the Quantum project will prepare guidelines on this topic.
- **Obligations expected based on the EHDS:**
 - For **data holders:**
 - **2029:** Share dataset descriptions for most categories of data and be prepared to share the data.
 - **2031:** idem for all categories of data
 - For **data users:**
 - **2029:** Data requests are possible for most categories of data
 - **2031:** idem for all categories of data

Approach Work Package 9 (HDAB-NL) since Dec 2023

- **Study numerous resources on data quality:**
 - well-known data quality frameworks (more details follow)
 - EHDS and European sources
 - reports from projects TEHDAS1, QUANTUM
- **Participate in initiatives around data quality:**
 - Dutch [NORA working groep on data quality and maturity](#)
 - Project QUANTUM Delphi surveys on data quality dimensions
 - Mid-scale pilot tests QUANTUM data quality tool
 - Join Community of Practice (SG2) metadata catalogue and data quality
 - WP6 metadata catalogue in HDAB-NL
- **Knowledge exchange on data quality**
 - With stakeholders (data users, data holders, other HDABs) and data quality experts
 - By organising workshops, breakouts, specific stakeholder visits and meetings,
 - By joining congresses and seminars organised by HDAB-NL Work Package 6 (metadata catalogue) or stakeholders that include the topic of data quality. For example, by the QUANTUM project, NIVEL, and others

Data Quality in the EHDS (i)

- **Definitions** (art. 2(2))
 - *'data quality'* means the degree to which the elements of electronic health data are suitable for their intended primary use and secondary use;
 - *'data quality and utility label'* means a graphic diagram, including a scale, describing the data quality and conditions of use of a dataset;
- A data quality and utility **label** for datasets would inform health data users about the quality and utility characteristics of a dataset and enable them to choose the datasets that best fit their needs. (recital 85)
- The data quality and utility **label** should not prevent datasets from being made available through the EHDS (recital 85)
- Datasets with electronic health data collected and processed with the support of Union or national public funding shall have a data quality and utility label (art. 78(2))
- Specifications are still to be developed by the European Commission (art. 78 (5+6))

Data quality in de EHDS (2)

- Tasks data holder
 - Responsible for adding a data quality and utility label to a dataset (art. 78(1))
 - Responsible for providing documentation for verification of the label (art. 60(4))
- Tasks HDAB
 - Supervision of the label, has the right to withdraw a label (art. 57(1d)) & (art. 78(4))
 - Cooperating with data holders on consistent and accurate implementation of the label (art. 57(1d))
- Aspects - dimensions:
 - Documentation: metadata, formats, standards, primary source, data model
 - Technical quality: completeness, accuracy, validity, consistency
 - Quality management: degree of process maturity, audits, bias analysis
 - Coverage: period, population, representativity
 - Access & delivery: time to inclusion in dataset, time to provision after approval
 - Update / change information: merges, additions, links to other datasets

Analyse various frameworks, principles of data quality and maturity

Work Package 9 started with studying the definitions, scope and associated indicators of data quality dimensions in the various frameworks. Followed by studying approaches for improving data quality and data quality maturity. The team focused on similarities and differences between frameworks, data quality dimensions, and maturity approaches are analysed.

- **TEHDAS1:** [TEHDAS' proposals for data quality and utility in EHDS - deliverable](#)
- **QUANTUM:** [two deliverables DQ framework \(labelling dimensions & maturity\)](#)
- **NORA online:** [data quality framework](#) and [data quality maturity](#)
- **DAMA international en DAMA NL:** internationale datamanagement organisatie en [results Dutch working group](#)
- **EMA (European Medicines Agency):** [data quality framework EMA](#)
- **Health Data Research UK** [Data Utility Evaluation - HDR UK](#)
- **Data quality at a Ministry**
- **EOSC taskforce FAIR metrics and data quality** [Towards a Data Quality Framework for EOSC](#)
- **ALCOA++ principles** [guideline](#)
- **Metadata kwaliteit (EU)** [data.europa.eu methodology](#)
- **NICTIZ Nationale bibliotheek en terminologie server**
- **CoreTrustSeal** certification data repositories [CoreTrustSeal – Core Trustworthy Data Repositories](#)
- **HealthDCAT** metadata standard extension to DCAT-AP for health data [HealthDCAT AP | European Health Information Portal](#)

Existing frameworks data quality – dimensions (1)

TEHDAS1	NORA	DAMA	QUANTUM	
Dimension	Juistheid	Dimension of Quality	Category	Key dimension
Relevance	Compleetheid	Validity	Access & provision	Accessibility
Accuracy and Reliability	Validiteit	Completeness		
Coherence	Consistentie	Consistency	Coverage	Population coverage
Coverage	Actualiteit	Integrity	Data documentation	Population representativity
Completeness	Precisie	Timeliness		Compliance
Timeliness	Traceerbaarheid	Currency	Technical quality	Data Provenance
	Begrijpelijkheid	Reasonableness		Metadata scope
		Uniqueness/ Deduplication		Accuracy
		Accuracy		Coherence
				Completeness
				Consistency
				Precision
				Validity

EMA



Existing frameworks data quality – dimensions (1a)

TEHDAS1

Dimension	Definition
Relevance	How well data meets users' needs.
Accuracy and Reliability	How closely data reflects what it was designed to measure and whether this is consistent over time.
Coherence	How consistent is data across data sources and data holders and can be combined and compared.
Coverage	The degree of representativeness in the population to which the dataset refers to, its exposures and events.
Completeness	Level of missingness at variable level.
Timeliness	How up-to-date the information is collected and delivered.

Existing frameworks data quality – dimensions (1b)

NORA

Juistheid	Thematische juistheid	Classificatie juistheid	Positionele juistheid	Kwantitatieve juistheid		
Compleetheid	Dataset compleetheid	Object compleetheid	Over compleetheid	Historie compleetheid	Metadata compleetheid	Ruimtelijke dekking
Validiteit	Formaat validiteit	Domein validiteit	Geometrische validiteit			
Consistentie	Logische consistentie	Referentiële integriteit	Identificeerbaarheid	Homogeniteit	Topologische consistentie	
Actualiteit	Update frequentie	Versheid				
Precisie	Opslagprecisie	Geometrische precisie	Statistische precisie			
Plausibiliteit	Authenticiteit	Reputatie	Bewijsbaarheid	Waarschijnlijkheid	Representativiteit	
Traceerbaarheid	Herleidbaarheid	Reproduceerbaarheid	Gebruiks inzicht			
Begrijpelijkheid	Leesbaarheid					

Existing frameworks data quality – dimensions (1c)

DAMA

Dimension of Quality	Description
Validity	Validity refers to whether data values are consistent with a defined domain of values
Completeness	Completeness refers to whether all required data is present.
Consistency	Consistency is ensuring that data values are coded using the same approach, assessment and valuation criteria. Three levels on which consistency can be measured were identified: record-level consistency, cross-record consistency, temporal consistency.
Integrity	Integrity refers to the lack of incoherent values and broken relationships in data. This can be measured as coherence, or as parent-child i.e. referential integrity.
Timeliness	Data Timeliness refers to the time data needs to become accessible to a user after its capture or update.
Currency	Currency is the date that the data was last updated relative to now and the likelihood that it is still correct.
Reasonableness	Reasonableness asks whether a data pattern meets expectations.
Uniqueness/ Deduplication	Uniqueness states that no real-world entity exists more than once within the dataset.
Accuracy	Accuracy refers to the degree that data correctly represents 'real-life' entities. Accuracy is difficult to measure [...]. Most measures of accuracy use the other dimensions to imply accuracy.

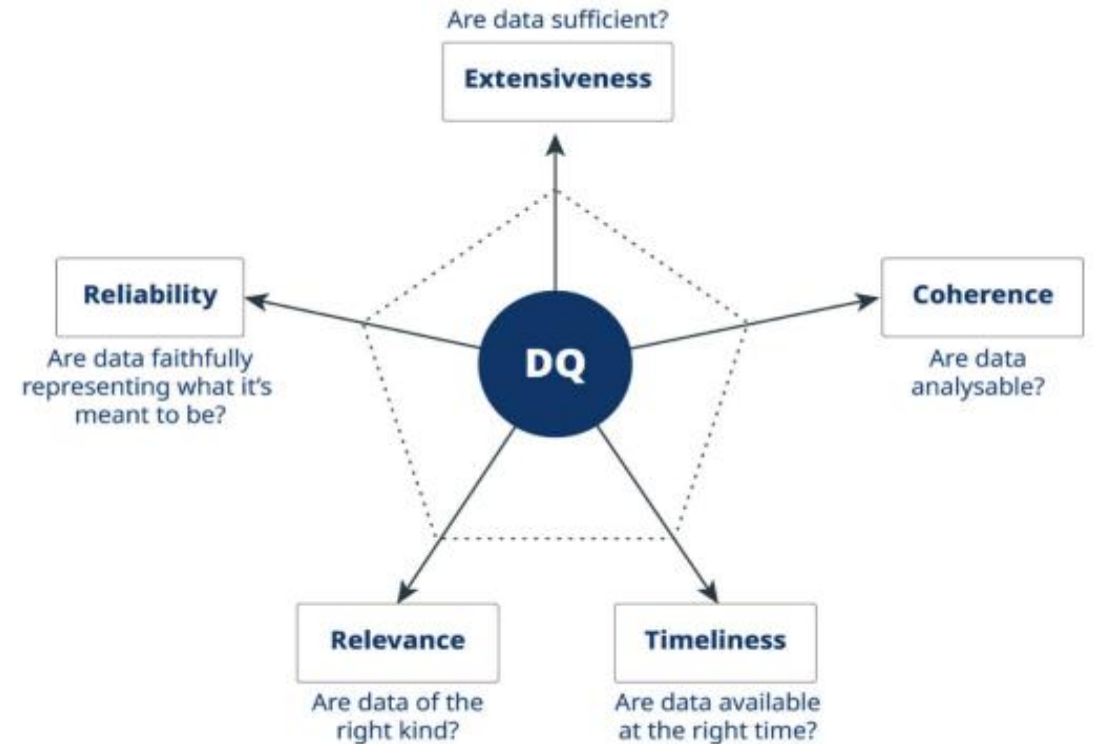
Existing frameworks data quality – dimensions (1d)

QUANTUM

Category	Key dimension	Measure labels
Access & provision	Accessibility	Availability of a data access & usage policy at the time of release of the dataset
		Average time from data access application to data release for a specific dataset
Coverage	Population coverage	Coverage Rate (percentage of the eligible population represented in the dataset)
	Population representativity	How closely does the observed population represent the expected population?
Data documentation	Compliance	Is there documentation of compliance with ethical standards, conventions, protocols or regulations?
	Data Provenance	Is the source of the dataset documented? Are the processes and operations on the data documented?
	Metadata scope	Existence of comprehensive standardised metadata Existence of an exhaustive data dictionary at variable level
Technical quality	Accuracy	Is the accuracy of the dataset documented?
	Coherence	Is coherence of the dataset documented?
	Completeness	Is completeness of the dataset documented?
	Consistency	Is the consistency of the dataset documented?
	Precision	Is the precision of the dataset documented?
	Validity	Availability of a conformance report for the data model

Existing frameworks data quality – dimensions (1e)

EMA



Preliminary findings comparing DQ frameworks (dimensions) (1)

- The frameworks each have **different dimensions of consistency and coherence**. In addition, the same dimensions with different definitions are mentioned in different frameworks.
- The exact interpretation and measurement methods for dimensions differ per framework, making **mapping** these dimensions **complex** and requires in-depth knowledge.
- The comparison of these frameworks at the level of the dimensions is not yet complete and will be continued.

QUANTUM labelling tool

100%

1. Access and provision

1.1. Accessibility

Definition: Accessibility refers to the dataset being accompanied by clear and transparent access and usage conditions.
Dimension Weight: 9.95 out of 100.

Metric #1

Definition
Availability of a data access & usage policy at the time of release of the dataset

Recommended Measurement Approach
https://www.w3.org/TR/vocab-dcat/#Propertydistribution_access_rights

Weight
50%

Comprehensive policy available

Metric #2



Quality & Utility Assessment score



Total Q&U Score: 65.67 / 100

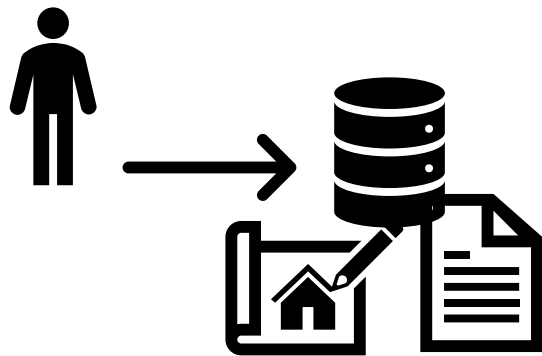
Maturity score: 42.0% (21/50)

Other findings (2)

- The QUANTUM tool is tested, and feedback is sent and will be send after the second test that has just been started to the QUANTUM project. Detailed report in Deliverable 6.2 (work package 6 HDAB-NL) next Summer.
- It is expected that health data catalogues will work with a metadata schema as DCAT-AP extended with HealthDCAT-AP. Whereas the data quality elements used in the QUANTUM tool to create the label are not yet included in HealthDCAT-AP; It is strongly recommended to **add and integrate these DQ elements into the metadata scheme** and develop controlled vocabularies for these.
- Insights from the EOSC Task Force underline that it is **difficult to define absolute data quality**, because the necessary quality dimensions and levels depend on the specific (secondary) use, stakeholders' context and the lifecycle of a dataset, it means that one dataset can never meet all possible requirements.
- Although active maintenance provides significant advantages to users, not all **frameworks** evaluated are currently under active **maintenance**. It is recommended for the QUANTUM framework to designate an owner and governance process after the project so that the framework can be updated and remains relevant.
- A recurring recommendation is to **add context typical usage information** to the metadata or available documentation, such as usage instructions, limitations, anonymization techniques applied, and relationships between attributes, so that users can make better informed choices. (see also next slide)

Fit-for-use (utility), fit-for-purpose (utility for specific purpose) and perceived data quality

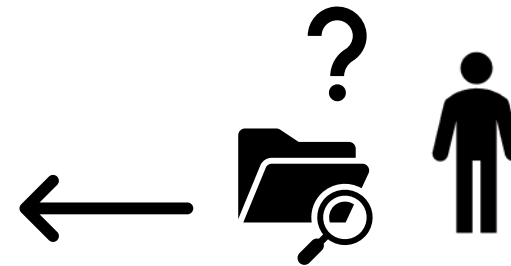
data holder



HDAB



data user



The data is fit-for-use and fit-for-purpose and is described in metadata (HealthDCAT)
The Quantum tool for creating the data quality label – focuses on recording technical quality (dimensions) and less on other documentation, maturity

Advantage of improving metadata and documentation on the part of the data holder: fewer mismatches, fewer questions from a hdab and potential data users.

User can find data in the HDAB metadata catalogue. The question is, will the data that has been found be usable: actual utility = fit-for-purpose and how is the data quality assessed? The **perceived data quality** of the data is likely to depend on: completeness of metadata, available additional documentation (e.g. data dictionary, use of standards), additional information / communication with data holder and the data itself. What does the user need to find a good match?

Data Quality maturity: Dutch NORA

Six themes in data quality;
1. context, 2. description dataset, 3. rules, 4. measuring data quality, 5. response, 6. implementing measures.

Three levels of maturity
(basis, advanced, expert)

	Basis – eerste serieuze stappen	Gevorderd - professioneel	Expert - geavanceerd
1. Bepalen context	<ul style="list-style-type: none"> Start met concreet probleem, behoefte, nulmeting en/of wetgeving Zicht op de bedrijfscontext Uitvoerende verantwoordelijkheid is duidelijk Betrokkenheid van eindverantwoordelijke 	<ul style="list-style-type: none"> Jaarlijkse verbetercyclus Focus op dataset zelf Uitvoerende en eindverantwoordelijke rollen gedefinieerd en geformaliseerd Competente medewerkers 	<ul style="list-style-type: none"> Focus op relatie met andere datasets (m.n. consistentie) Processen en verantwoordelijkheden geformaliseerd Borging competenties in functiehuis en in de werving
2. Beschrijven gegevens	<ul style="list-style-type: none"> Eerste vastlegging van metagegevens, minimaal op het niveau van datasets Eerste aanzet tot kritieke gegevenselementen 	<ul style="list-style-type: none"> Centrale metagegevens repository, tot op het niveau van gegevenselementen Begrippen, informatie- en/of gegevensmodel gedefinieerd Verbinding met organisatiebrede master- en referentiegegevens gedefinieerd 	<ul style="list-style-type: none"> Data lineage volledig inzichtelijk Koppeling van metagegevens en modellen Metagegevens actief en continu verzamelen, inzichtelijk maken en gebruiken om bij te sturen Harmonisatie met organisatiebrede master- en referentiegegevens
3. Beschrijven regels en eisen	<ul style="list-style-type: none"> Eerste inventarisatie van regels en eisen Informeel afspraken over kwaliteit 	<ul style="list-style-type: none"> Centrale vastlegging van regels en eisen Gestandaardiseerd formaat Data profiling met professionele tool Gegevensleveringsovereenkomsten 	<ul style="list-style-type: none"> Formeel sturen op gegevensleveringsovereenkomsten Geautomatiseerde controle, bewaking en bijsturing van regels en eisen
4. Meten gegevenskwaliteit	<ul style="list-style-type: none"> Eerste controles, met focus op validiteit en consistentie binnen dataset 	<ul style="list-style-type: none"> Geautomatiseerde metingen en duplicaatdetectie Organisatiebreed kwaliteitsdashboard 	<ul style="list-style-type: none"> Pro-actief signaleren en oplossen problemen met gegevenskwaliteit (data observability)
5. Afhandelen signalen	<ul style="list-style-type: none"> Informeel proces voor melden en herstellen Meldingen via e-mail Bepalen oorzaken van problemen 	<ul style="list-style-type: none"> Formeel proces voor melden en herstellen Standaard terugmeldmechanisme Bewaken van signaalafhandeling 	<ul style="list-style-type: none"> Geautomatiseerde workflows Sturen op kwaliteitseisen Zicht en sturing op effect van verbetering
6. Implementeren maatregelen	<ul style="list-style-type: none"> Informeel proces 	<ul style="list-style-type: none"> Formeel proces, projectmatig Implementeren maatregelen als formele beheersmaatregelen 	<ul style="list-style-type: none"> Kwaliteitscyclus op beheersmaatregelen Gegevenskwaliteit by design

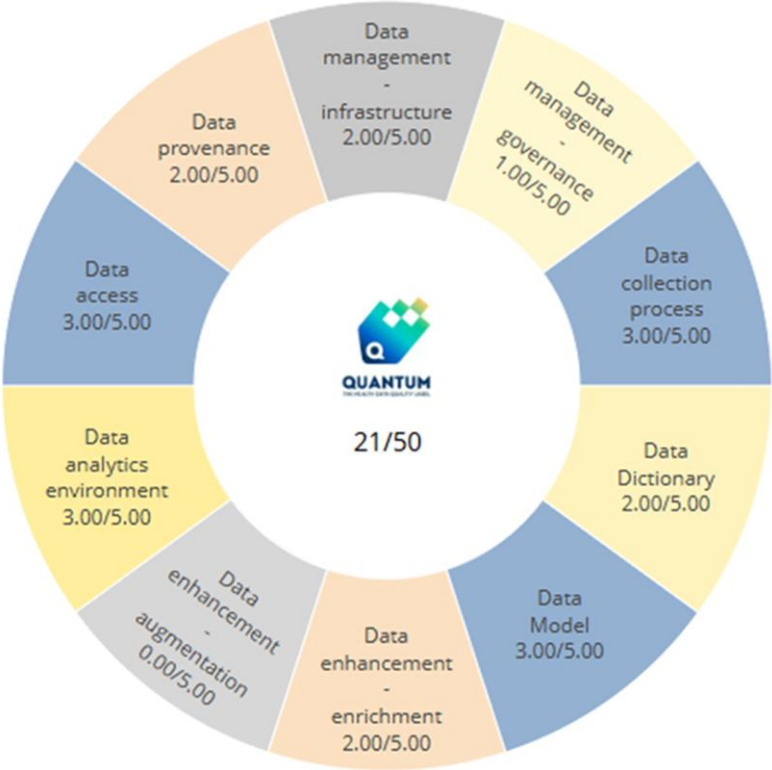
Quantum maturity model (1)

DIMENSION	DEFINITION	Level 1	Level 2	Level 3	Level 4	Level 5
		INITIAL: A NEW OR UNDOCUMENTED PROCESS, UNPREDICTABLE, POORLY CONTROLLED AND REACTIVE	EMERGING (OR DOCUMENTED, STRUCTURED): THE PROCESS IS DOCUMENTED SUFFICIENTLY SO REPEATING THE SAME STEPS MAY BE ATTEMPTED	DEFINED: THE PROCESS IS DEFINED AND CONFIRMED AS A STANDARD PROCESS	MANAGED: THE PROCESS IS QUANTITATIVELY AND PROACTIVELY MANAGED IN ACCORDANCE WITH AGREED-UPON METRICS	OPTIMISED: PROCESS MANAGEMENT INCLUDES DELIBERATE PROCESS OPTIMISATION
DATA COLLECTION PROCESS	The processes relating to the capture of data from multiple sources for secondary purposes	No data collection or capture process for secondary purposes	Ad hoc processes exist for data collection	Defined processes exist for some data collection	Defined processes exist for all data collection and are used for secondary purposes	Processes are automated for all data collection, focus on continuous process improvement
DATA MANAGEMENT – GOVERNANCE	The level of maturity of the data management process	No documented data management governance	A documented data management plan covering collection, auditing, and management is available for the dataset	Evidence that the data management plan has been implemented is available	Demonstrated compliance with the data management plan	Externally verified compliance with the data management plan
DATA MANAGEMENT - INFRASTRUCTURE	The level of implementation and development of the data holder’s data management infrastructure	No data management infrastructure	An emerging data management infrastructure, some validation and verification	Data management infrastructure defined and confirmed as a standard process	Data management infrastructure, with partially automated, verified and validated [real time] data management	A robust and comprehensive data management infrastructure, with fully automated, verified and validated real time data management
DATA PROVENANCE	Clear description of the lineage of the dataset providing transparent and comprehensive documentation of the data pipeline to get the dataset	No documented provenance	Source of the dataset is documented	Source of the dataset and any transformations, rules and exclusions	All original data items listed, all transformations, rules and exclusion listed and impact of these	Ability to view earlier versions, including “raw” or “source” dataset and review the impact of each stage/step
DATA ACCESS	How well defined and implemented are data access processes, from a legal, ethical and technical perspective	No data access processes or procedures	Have the processes and procedures but don’t respond in timely and consistent manner	Have the processes and procedures, and respond in timely and consistent manner	Data access system that covers both technical and policy areas, in accordance with agreed metrics	A comprehensive data access system that covers technical, ethical and policy areas (allowable uses, API documentation, access and approvals) compliant with EU policy

Quantum maturity model (2)

DIMENSION	DEFINITION	Level 1	Level 2	Level 3	Level 4	Level 5
		INITIAL: A NEW OR UNDOCUMENTED PROCESS, UNPREDICTABLE, POORLY CONTROLLED AND REACTIVE	EMERGING (OR DOCUMENTED, STRUCTURED): THE PROCESS IS DOCUMENTED SUFFICIENTLY SO REPEATING THE SAME STEPS MAY BE ATTEMPTED	DEFINED: THE PROCESS IS DEFINED AND CONFIRMED AS A STANDARD PROCESS	MANAGED: THE PROCESS IS QUANTITATIVELY AND PROACTIVELY MANAGED IN ACCORDANCE WITH AGREED-UPON METRICS	OPTIMISED: PROCESS MANAGEMENT INCLUDES DELIBERATE PROCESS OPTIMISATION
DATA ANALYTICS ENVIRONMENT	Analytical services, tooling and access to [secure] data environments	No data environment available	Requested analysis can be undertaken by internal teams and provided back in anonymized format to data requestors	The dataset can be used in a secure data environment (SDE)	The dataset can be used in and SDE and other data and tools can be brought in as required	The dataset can be used in federated organised environment
DATA ENHANCEMENT - AUGMENTATION	The application of various techniques to make data more useable for specific purposes	No data augmentation	Some techniques to make data more useable for specific purposes	Defined techniques to make data more useable for specific purposes	Managed techniques to make data more useable for specific purposes	Comprehensive application of various techniques and mapping to data model i.e. OMOP to make more useable for specific purposes
DATA ENHANCEMENT – ENRICHMENT	Data sources enriched for example with annotations, image labels, phenomes, derivations, NLP-derived data labels	The data has no additional derived fields, or enriched data	The data include additional derived fields, or enriched data	The data include additional derived fields, or enriched data used by other available data sources	The derived fields or enriched data were generated from, or used by, a peer reviewed algorithm	The data includes derived fields or enriched data from an [inter]national report
DATA MODEL	Availability of clear, documented data model that provides structure and standardization	There is no data model	Known and accepted data model but some key fields uncoded or free text	Key fields codified using a local standard and updated over time	Key fields codified using a national or international standard and updated	Data model conforms to [inter] national standard and key fields codified using a national / international standard
DATA DICTIONARY	Provided documented data dictionary and terminologies	No data dictionary	Data definitions available	Definitions compiled into local data dictionary which is available online	Dictionary relates to national definitions	Dictionary is based on international standards and includes mapping

Example QUANTUM tool maturity score for a dataset



Other findings data quality maturity

- It is preferable to **use mature, community-supported and maintained data quality frameworks**, but not all frameworks discussed are equally well-known, widely adopted, or fully developed.
- The higher the data quality maturity, the better the quality of individual datasets; however, **different frameworks assess maturity at different levels**, such as the level of the **dataset** (e.g. QUANTUM and EMA), the **organization** (e.g. DAMA and NORA) or the **community of discipline** (such as the EOSC Task Force).
- The improvement of data quality is also related to:
 - **quality of metadata,**
 - **quality of data repositories,**
 - **quality of federated data systems,**
 - **effective collaboration environments** and
 - the use and development of **standards for interoperability of infrastructures**
 - **semantic standards and harmonization standards** for interoperability of **datasets**.Together, these ensure consistency, harmonization and better usability of data.

Summary

The milestone presents the progress of work package 9 (Improving data quality) which focuses on two main tasks: developing and promoting a data quality assessment framework through both a data quality labelling system (T9.1), and through guidelines and a maturity model (T9.2). Considering that the QUANTUM project has already developed both a labelling system and a maturity model within the EHDS context, it was decided to focus on gaining insight on:

- current data quality practice and needs among Dutch data holders and users of health data,
 - current gap between the requirements of EHDS and current practice
 - the applicability of the QUANTUM label, tool, and training materials
- and to promote applicable data quality frameworks and develop guidelines for data quality improvement in the Dutch context.

Between December 2023 and July 2025, the main activities have been studying resources on data quality, knowledge exchange with Dutch stakeholders and participating in data quality initiatives.

- Special attention was given to data quality resources aligned with health data management, such as
 - project results from TEHDAS and QUANTUM, as well as frameworks used in commercial (e.g., DAMA DMBOK) and public organisations (e.g., NORA), and other widely referenced models (e.g., EMA, Health Data Research UK, EOSC FAIR taskforce).
 - development of the metadata standard HealthDCAT

Currently, there is limited insight into which data quality frameworks and principles are most widely adopted by organisations handling health data in the Netherlands. To address this, stakeholder consultations have begun and will continue in the coming year.

MS30 Highlights

1. The enhancement of data quality depends on several interconnected factors that are not on the level of the dataset nor the data quality maturity of the organization but can be seen as **external factors** such as the quality and interoperability of standards and infrastructures.
2. **High quality metadata** provides clear descriptions about the data's content, origin, format, and how it can be used. This makes it much easier for users to assess whether a dataset meets their specific needs (in other words, whether it is **fit-for-purpose**).
3. The Dutch NORA roadmap for **data quality improvement** is community-based and focuses on very practical steps and organisational change, making it an excellent resource for **actionable guidance on an organisational level**.
4. The **QUANTUM framework** for data quality provides a **valuable foundation**, though it does have some limitations. It would benefit from ongoing maintenance, evaluation and updates based on feedback from data holders, data users and HDABs.
5. The **importance** of certain data quality dimensions seems to **depend on the purpose** for which the data is used. However, the QUANTUM data quality and utility label uses fixed weights of certain dimensions.
6. **Data quality** can be improved **both at dataset and organisational level**. The QUANTUM framework focuses mainly on data quality maturity on a dataset level and not on organisational level.
7. Fit-for-purpose and fit-for-use requirements depend on the perspective (data holder, data user). **The QUANTUM framework focuses mainly on fit-for-use** and enhanced quality of data supplied by the data holder.



Thank you

