

Data Dictionary Maturity Assessment

Please select exactly one answer per question. Select the option that best describes your current situation today, not the situation you are working toward.

Dimension 10 — Data Dictionary

Definition*: A documented data dictionary and terminologies. The data dictionary explains, field by field, what each item means, what values it may take, and how it should be interpreted, so that somebody outside your organization can use the data correctly.

1. Do you have a data dictionary?*

- | | |
|--|--------|
| We have nothing that documents what our fields mean. | 1 pt. |
| ✓ Definitions exist for some fields, scattered across documents. | 2 pts. |
| All definitions are collected in one dictionary of our own, published online. | 3 pts. |
| The dictionary uses national definitions. | 4 pts. |
| Our definitions follow international standards, and the mapping to them is documented. | 5 pts. |

You can support this answer with: the data dictionary file or page itself.

2. How complete are the definitions?*

- | | |
|---|--------|
| There are no definitions. | 1 pt. |
| Definitions for some fields. | 2 pts. |
| Definitions for all fields used for secondary purposes. | 3 pts. |
| As above, including value sets, units, and permitted ranges. | 4 pts. |
| As above, including missing-value conventions and known limitations for each field. | 5 pts. |

You can support this answer with: a sample of dictionary entries for your most-used dataset.

3. Who can read the data dictionary?*

- | | |
|--|--------|
| Only our own team. | 1 pt. |
| Anyone, on request. | 2 pts. |
| It is published online, so users can read it before requesting data. | 3 pts. |
| It is published online and supplied with every dataset release, versioned in step with it. | 4 pts. |
| It is published online in machine-readable form, for example, through an API or catalog entry. | 5 pts. |

You can support this answer with: the public link to your dictionary, or the API endpoint.

4. What is the dictionary standardized against?*

- | | |
|---|--------|
| Nothing. Each field is described in whatever wording the author chose. | 1 pt. |
| An in-house term list is used for part of the dictionary. | 2 pts. |
| Standardized against a local or organizational standard. | 3 pts. |
| Standardized against national definitions. | 4 pts. |
| Based on international standards, with the mapping documented, for example, SNOMED CT, LOINC, or ICD. | 5 pts. |

You can support this answer with: your mapping table from local fields to the standard used.

5. How is the dictionary kept up to date?*

- | | |
|--|--------|
| It is not maintained. | 1 pt. |
| Updated when somebody notices an error. | 2 pts. |
| Updated whenever the data model changes. | 3 pts. |
| Reviewed on a fixed schedule, with a version number and date. | 4 pts. |
| Reviewed on a fixed schedule, versioned, with a change log published to users. | 5 pts. |

You can support this answer with: the version and date on your current dictionary, its change log.

YOUR NOTES FOR THIS DIMENSION

What sits behind these answers: the practices and tools you rely on, recent improvements, and the gaps that explain the lower scores. This is the material for the supporting narrative on the Overall Score sheet.

DIMENSION SCORE

Add up the points of all 5 answers and divide by 5. Keep one decimal.

YOUR MATURITY LEVEL

Round the dimension score down to the nearest whole number.

Level 1 Initial Starting out · 1 Point	Level 2 Emerging Written down · 2 Points	Level 3 Defined In practice · 3 Points	Level 4 Managed Measured · 4 Points	Level 5 Optimized Improving · 5 Points
--	--	--	---	--

The five maturity levels in this dimension

Your five answers average to one level. Each row describes what that level typically looks like across all five questions, so you may be ahead of it on some points and behind on others.

Level 1 Initial	No dictionary, no definitions, visible only internally, in your own wording, not maintained.
Level 2 Emerging	Scattered definitions for some fields, available on request, in-house terms, fixed when noticed.
Level 3 Defined	One dictionary covering all fields for secondary use, published online, updated with the model.
Level 4 Managed	Value sets, units, and ranges included; national definitions; versioned and reviewed on schedule.
Level 5 Optimized	International standards with documented mapping, machine-readable, with a published change log.

Your next step: find the row for the level you reached, then read the row directly below it. That row is what you need to achieve to move up one level in this dimension.

Note: Every question in this dimension uses the same 1–5 scale, and the same real-world situation is worth the same number of points in every question, so both the lowest and the highest level are reachable. Transfer your maturity level for this dimension to the *Maturity Overall Score Calculation* sheet, which brings all ten dimensions together.

References*

Question 1 follows the maturity levels of this dimension in the QUANTUM data holders' maturity model (Deliverable 1.2 v3.5); the wording is EHDS4ALL's own. Underlying source given there for this dimension: Adapted HDRUK / EHDS. Questions 2 to 5 are EHDS4ALL, developed from the WP3 workshop of August 2024 and the expert interviews of October 2024. The five level statements below combine all five questions, because the dimension score is the average of all five answers.

Jones, M., & Bernal-Delgado, E. (2025). **Deliverable 1.2 — Specification for the assessment of data holders' maturity** (Version 3.5.0). QUANTUM project. Zenodo. doi: 10.5281/zenodo.14944767

Gordon, B., Barrett, J., Fennessy, C., et al. (2021). Development of a data utility framework to support effective health data curation. *BMJ Health & Care Informatics*, 28(1), e100303.

European Commission. (2022). Proposal for a Regulation of the European Parliament and of the Council on the European Health Data Space. COM(2022) 197 final. Adopted as Regulation (EU) 2025/327.

EHDS4ALL (2024). Protocol, Workshop Kick-off WP3, Cologne, 19 August 2024. Work Package 3, dimension validation and maturity levels.

EHDS4ALL (2024). Expert interviews with health data holders and maturity model developers, October 2024. Conducted by inav and the University of Cologne; participants anonymized.