

National Underground Asset Register (NUAR)

NUAR Data Model Publication

Principles underlying the NUAR Harmonised Data Model

Publication History

Date	Version	Description
12/08/2024	0.1	First publication of the NUAR Data Model

Purpose of this document

This document is intended to outline the high-level principles which underpin the development and ongoing maintenance of the NUAR Harmonised Data Model and guide the direction and evaluation future developments.

The principles are presented in descending order of importance, such that if a requirement results in a conflict between principles, consideration of the higher priority principle should prevail.

Revisions to the Data Model will be considered in the light of these underlying principles by the prevailing governance processes, which will also be responsible for reviewing and maintaining these principles.



Data Model Principles

In descending order of importance, the NUAR Harmonised Data Model should:

1. **Provide a consistent framework** for communicating information to data consumers in a common language.
2. Allow for the representation of data that is comprehensively **fit for purpose** for data consumers and enables the data requirements for the supported use cases to be implemented.
3. Support the broad **participation of data providers** in the NUAR ecosystem.
4. Enable scaling to **comprehensive geographical coverage** of data.
5. Draw on **community consensus** to enable a standardised representation of common concepts across geographies and sectors.
6. Facilitate the **onboarding of new data providers** to the NUAR ecosystem in new and existing sectors.
7. Facilitate the clear identification of **data ownership and custodianship**.
8. Support and enable, but not require, **data quality improvement** for datasets owned and managed by data providers.
9. Not inhibit **outbound interoperability** with other systems, services and initiatives.
10. Be based on a **broad and extensible framework** to support the principle of extensibility in the absence of specific extensibility requirements.

Function of a Data Specification for the NUAR Harmonised Data Model

A Data Specification for the NUAR Harmonised Data Model should:

1. Give data owners (or their agents) **definitive guidance on data expected** (types, content and structures) from a data provider for ingestion into the NUAR ecosystem.
2. Give data consumers **detailed information about the data that is being ingested or made available** through the NUAR ecosystem.
3. State the **minimum compliance conditions**, both generally and for specific use cases (which may result in a subset or profile of the Data Model).
4. State **conformance requirements** and define criteria for dealing with **non-conformant** data (e.g. accepting data but feeding back to data providers to enable data quality improvement). These requirements and criteria may vary depending on the Data Maturity Level of the supplier and the type of data (as suitably defined and agreed).
5. Provide definitions and explain the **purpose** of the various elements of the Data Model in terms of supporting the requirements of the NUAR ecosystem.



Appendix – Some questions underlying the formation of these principles

Why do we need a data model?

There is a requirement for a common language to communicate information to data consumers in the NUAR ecosystem, and a data model provides a framework for this common language.

How does the NUAR Data Model relate to the MUDDI Model?

The MUDDI Conceptual Model provides broad and high-level framework which defines the key concepts and entities required to represent elements of a utility network (and the broader surface and subsurface context within which it is installed and operated). The NUAR Harmonised Data Model may be considered as a “UK Excavation” profile of the MUDDI Conceptual Model, implementing as it does the scope, terminology and level of detail appropriate for that jurisdiction and use case.

What are the data model requirements to enable scalability to comprehensive geographical and sector coverage?

The data model design and terminology should aspire to be sufficiently generic and have sufficient flexibility to accommodate data from new data providers with a minimal requirement for redesign.

Should data ever be excluded from the NUAR ecosystem on the grounds of quality?

The principle of a minimum data quality threshold is valid, and this is represented in a Data Specification for a given use case and in an emerging Data Maturity Model which takes into account valid differences across organisations and sectors.

Based on the principle of broad participation in the NUAR ecosystem, and the principle of asset owners retaining responsibility for data management and data quality, the lowest threshold should generally represent a low bar for participation in the current use case: broadly data should be of sufficient detail and quality to allow it to be incorporated in a manner that at a minimum identifies the source/owner, the type of feature, and allows spatial geometry to be faithfully presented.

How important is standardisation of data values across geographies and sectors?

There is value in presenting common concepts in a standardised way across geographies and sectors, as this improves clarity and reduces ambiguity, and provides a framework which can be used to assess data conformance in a standard way across multiple organisations, sectors and geographies. This common terminology should be based on a community consensus within the industry or sector.



In keeping with the foregoing principles, data should not be excluded from the NUAR ecosystem on the basis of non-conformance with standardised terminology, but feedback may be provided on such non-conformance to data owners in order to support and enable data quality improvement and standardisation in source systems.

To what extent should attribute values in the NUAR ecosystem be faithful to source attribute values?

Where one-to-one mappings to standardised values are unambiguously defined by community consensus for the use case, source values may be mapped. The Data Model also supports representation of “original values” in domains where standardised values are implemented, in order to maintain a clear audit trail to original source data.

Where mapping to a standardised value is not possible, the original value may be retained, and the failure to map provided as feedback to the data owner, thus potentially enabling data quality improvement. The level of conformance of input data to standardised values may be reflected in metadata on data quality/maturity level. Repeated or widespread non-conformance with standardised values may also be grounds for consideration of a Data Model change, as per the prevailing governance processes.

With the approach outlined above, the NUAR Data Model may support and enable data quality improvement and support ongoing validation of the model itself.

To what extent should NUAR aim to improve data quality?

NUAR should aim to **promote**, **support** and **enable** data quality improvement. Responsibility for data quality improvement remains with the asset owner, and NUAR should not change incoming data, unless as per explicitly agreed and documented standardisation rules (see above).

There is an opportunity to drive data quality improvement via multiple feedback channels - from the end user, but also from the act of transforming data to a standardised target model.

How important is interoperability with other systems?

Interoperability may be considered to be the ability of other systems or processes which support other use cases or added-value services to interact with and extend or enrich the data held in the NUAR ecosystem.

While the focus of the NUAR programme is the excavation use case, the resulting platform does not exist within a vacuum, but rather sits within an ecosystem of systems and processes addressing different, but related, use cases. It is important that the design of the NUAR Data Model should not inhibit interoperability in principle.



How important is extensibility to other use cases?

Extensibility may be considered to be the ability to serve additional or extended requirements within the NUAR Harmonised Data Model.

The focus of the NUAR programme has been firmly on the excavation use case, and other potential use cases are being explored and evaluated by the NUAR Discovery project. While there are therefore no well-defined extensibility requirements at the time of publication, there is a desire to support the principle of extensibility by building the data model on an extensible framework.

Extensibility may be supported by implementing a broad but comprehensive data model, with specific use cases expressed in a narrower profile with tailored scope and level or detail, or we may implement a more restricted core to answer a particular use case or set of use cases and implement extensions to the core for further use cases in future.

The approach taken to the development of the NUAR Harmonised Data Model keeps both options largely open depending on the outcome of the NUAR Discovery Project. The MUDDI conceptual model provides a broad (but shallow) umbrella from which further profiles could be easily developed based on the experience of developing the “UK Excavation” profile for NUAR. Equally, the NUAR Harmonised Data Model itself has been built with flexibility and extensibility in mind (to mitigate the risk of encountering new and unexpected data scenarios as the coverage and onboarding scales), so use cases which extend or enrich the excavation use case could potentially be accommodated quite easily.