

B.7 Data standards for the Historic Environment

- [B.7 Data standards for the Historic Environment](#)
- [B.7.1 What are data standards?](#)
- [B.7.2 Why are data standards needed?](#)
- [B.7.3 MIDAS Heritage and FISH Terminology](#)
- [MIDAS Heritage](#)
- [FISH Terminology](#)
- [B.7.4 The FISH Interoperability Toolkit](#)

B.7 Data standards for the Historic Environment#

B.7.1 What are data standards?#

Data standards can be simply defined as an agreed statement of "what information should be recorded, in what manner, to achieve a particular objective". The objectives for historic environment managers are set out elsewhere in this manual. This section describes;

- the value that shared data standards add to historic environment recording,
- data standards that are in current use,
- how HER officers can get involved in shaping the future direction of data standards to ensure that they meet the needs of the profession.

B.7.2 Why are data standards needed?#

Collection and recording of historic environment information is an expensive, and in some cases unrepeatable, process. It is essential that the information contained within HERs can be readily retrieved and understood by anyone and increasingly important that data from different HERs can be compared. Data standards ensure that information is recorded in a consistent and retrievable way so that the maximum benefit for the users of data is obtained from the investment.

The benefits offered include:

- **Reliability:** As standard procedures are tried and tested, if they are sensibly applied they will work well.
- **Consistency:** Data standards promote consistency in recording of information, which ensures that records can be retrieved easily, so that all known relevant information is available when needed.
- **Expertise:** Historic environment data standards incorporate the accumulated expertise and experience of several decades. Their adoption builds this experience into an information system, reducing the risk associated with developing a system from scratch.
- **Compatibility:** No one organisation holds all the data relating to a particular site or area. There are usually other organisations with an interest who may have useful data. When recorded to common standards, data held in different databases can be shared and exchanged much more easily. This is increasingly important as new ways of accessing data remotely are developed.
- **Communication:** Data standards assist in communication of concepts involved in the historic environment both between HER staff and between HER staff and their software consultants and organisations within and beyond the sector.
- **Benchmarking:** Data standards perform a crucial role in providing an independent benchmark against which HER data quality can be assessed.
- **Staff skills:** Adoption of common data standards enable staff to become familiar with systems more quickly when changing jobs and enable employers to recruit staff with relevant skills. They also facilitate organisation of training through events such as conferences and seminars.

- **Information system development:** No information system lasts indefinitely. At some point data needs to be migrated from one system to its successor. Standards simplify the process of specification of new information systems and migration is greatly simplified if standards are consistent within the data.

B.7.3 MIDAS Heritage and FISH Terminology#

This section describes the purpose of the central standard for historic environment recording (MIDAS Heritage) and the use of standardised terminology in its application.

MIDAS Heritage is a British cultural heritage standard for recording information on buildings, monuments, archaeological sites, shipwrecks and submerged landscapes, parks and gardens, battlefields, artefacts and ecofacts. The data standard suggests the minimum level of information needed for recording heritage assets and covers the procedures involved in understanding, protecting and managing these assets. It also provides guidelines on how to support effective sharing of knowledge, data retrieval and long-term preservation of data.

The use of standardised terminology alongside MIDAS Heritage allows consistent indexing practice which facilitates accurate and fast information retrieval. The main terminology used by HERs was originally held together as a national heritage reference dataset known as '[INSCRIPTION](#)' which comprised of a series of [wordlists](#) and [thesauri](#) developed by various heritage bodies, both national and regional. These word lists and thesauri are now managed and maintained by the national heritage organisations and the [FISH Terminology Working Group](#).

Further information about these are available online via the [Forum on Information Standards in Heritage \(FISH\) website](#).

Figure 7: The logo for the Forum on Information Standards in Heritage.

MIDAS Heritage#

Developed from data standardisation work in the 1980s and early 1990s, the first edition of the standard, [MIDAS ? A Manual and Data Standard for Monument Inventories](#), was published by the Royal Commission on the Historical Monuments of England (RCHME) in 1998. The organisation merged with English Heritage (now Historic England) in 1999 and development work in 2004 extended MIDAS to cover GIS type data as well as watercraft and aircraft recording. An updated version, MIDAS Heritage, was published in 2007 in collaboration with other UK heritage organisations including the [RCAHMS](#) and [RCAHMW](#). The standard was developed for and on behalf of the FISH and is maintained and updated by them in collaboration with Historic England, RCAHMS, RCAHMW and other contributing heritage organisations. MIDAS Heritage is available as a free downloadable electronic from the [Historic England website](#).

Figure 8: The logo of MIDAS Heritage, the national data standard for the content of historic environment records.

What is in MIDAS Heritage?

MIDAS Heritage is a content standard in that it defines the individual facts or ?units of information? that should be included in a standardised record of, for example, a monument or archaeological event. It is also an open data standard, in that although information schemes are described the exact structure in which data is to be recorded is not defined, so that MIDAS Heritage can be applied to a range of information systems. It is set out as a series of closely integrated data standards, rather than one single standard and is designed to be used in conjunction with separate standards covering specific types of applications or projects, which will give the necessary data elements. Examples

include [SPECTRUM](#) (for artefacts), [UK Gemini Discovery Metadata Standard](#) (for GIS), and [CIDOC Conceptual Reference Model](#) (for concepts and relationships). MIDAS Heritage complies with the [UK e-Government Metadata Standard](#) (e-GMS), which is based on [Dublin Core](#).

The MIDAS Heritage data standard has a three-level structure with six main themes, each containing several Information Groups, which in turn contain a number of Information Units.

- **Level 1: Themes** - Describe the main subject areas at the broadest level
- **Level 2: Information Groups** - Outline the specific standard which should be included in any entry covering a particular subject
- **Level 3: Units of Information** - Provide the basic 'facts' or items that make up an entry

The six main themes and the information groups within each are set out in the diagram below:

Figure 9: An overview of the structure of MIDAS Heritage themes and information groups.

Who should use MIDAS Heritage?

MIDAS Heritage is intended for a wider audience than just the HER community and is recommended reading for anyone involved in inventory projects. Worked examples in the manual cover, for example, its application to academic research databases and voluntary sector thematic surveys. The same requirement for standards exists for these other sectors if the maximum value is to be obtained from their efforts. This is of course particularly important if the intention exists to deposit the results of these surveys with the local HERs.

In addition to the data standard itself, MIDAS includes a manual to assist those managing and maintaining an inventory of the historic environment. HER staff should not only be familiar with MIDAS themselves but should also recommend its use to local societies working in their area.

FISH Terminology#

Effective searching of HERs relies on the quality of indexing. Good indexing ensures effective retrieval of records relevant to a search; poor indexing means that records will be overlooked.

Most HER information systems support indexing using standardised terminologies built into the actual software. These include simple lists, such as types of legal protection for a monument, more complex hierarchical lists such as at the [Thesaurus of Event Types](#) or full indexing thesauri such as the English Heritage [Thesaurus of Monument Types](#) (see [Panel 4](#) and [Panel 5](#) in [Section B.8](#) for examples of wordlists and heirarchical thesauri).

The main terminologies ([wordlists](#) and [thesauri](#)) used by HERs were originally held together as a national heritage reference dataset known as '[INSCRIPTION](#)'. These are now managed and maintained by the national heritage organisations and the [FISH Terminology Working Group](#).

These worlists and thesauri are recommended for use in conjunction with MIDAS Heritage units of information. For each case where MIDAS Heritage recommends the use of a controlled terminology, a suitable indexing terminology has been developed by one or more of the partners in the Forum on Information Standards in Heritage. Details about each of the available wordlists and thesauri, along with downloadable versions (in either PDF and CSV format) can be found on the [FISH website](#). They are also available to browse online via the [Historic England website](#).

B.7.4 The FISH Interoperability Toolkit#

Interoperability is the ability to reuse data created in one software application in another. It is essential if efficient sharing of data between HERs and their users is to become a reality. Interoperability relies on data standards that are understood and shared within the HER community such as MIDAS Heritage and FISH Terminologies, plus technical standards to support the creation of data files that different software packages will be able to process.

To provide HER managers and their IT advisors with the technical tools needed to make interoperability possible the Forum on Information Standards in Heritage have developed a ?toolkit? of protocols and formats aimed at HERs and their technical advisors. It is based upon MIDAS Heritage and incorporates the HER benchmark Level 1 standard. Online validation of data against controlled terminologies is supported.

Whereas MIDAS Heritage and FISH Terminologies are aimed at HER staff, the Toolkit is aimed at IT specialists and software developers. HER staff should familiarise themselves with the functions that the Toolkit supports, but are advised to contact their IT support staff for discussion of how the Toolkit might be implemented in their systems.

The toolkit consists of four main components which check, map and verify content before exchange of textual and geographical information between systems.

- *The XML Mapping Tool*

This tool allows the user to upload their own non-MIDAS XML schema (or sample file) use a visual mapping tool to map between this file and MIDAS XML and then carry out a transformation. The mapping itself is stored in an XSL file under a user specific login for later reuse in this system or it can be exported to allow the transformation to be carried out elsewhere. This tool currently requires an additional Login process.

- *The Toolkit Data Validator*

This online application validates the content of MIDAS XML files. The presence or absence of data required by standards such as the English HER Level 1 Benchmark is checked. Reports are issued on compliance and a watermark is embedded in a new XML file for download.

- *The Concordance Tool*

This online application parses two MIDAS XML files to find updated records. Reference and New XML files are checked for the updated or new records. Reporting consists of description and two automatically generated XML files. The first one consists of only the differences between the two files (i.e. new or updated records only) and the second one consists of Reference XML file plus new records and updated records.

- *The Geospatial Tool*

This tool reads a valid XML file, finds coordinates in OSGB36 format and adds additional spatial elements for the same coordinates in WGS84 format. If the original coordinate format is WGS84 decimal degrees then additional elements in OSGB36 are added

Further information on the FISH Interoperability Toolkit can be accessed online at <http://archaeologydataservice.ac.uk/fishtoolkit/> or <http://fishforum.weebly.com/fish-toolkit.html>.