HES: Technical Specification No.1

Polygons:

Covering Polygon Classes:
- **Known Site Extent**
- **Discovery Area**

Mike Middleton (HES)

Introduction

This document is the technical specification to producing Known Site Extent and Discovery Area polygons in line with the methodology developed during the Defining Scotland's Places pilot.

This technical specification has been in use, by RCAHMS, since 2010. It is used as the underlying methodology for the Canmore Mapping programme.

Background

The Defining Scotland's Places pilot ran for one year from April 2010, funded by Historic Scotland and managed by the RCAHMS and developed in partnership with the SMR Forum and AlGAO Scotland. The project produced three interim reports detailing issued encountered during each trial:

- *Interim Report No.1: Renfrewshire Pilot*
- *Interim Report No.2: Midlothian Pilot*
- *Interim Report No.3: Northern Isles Pilot*

Detailed project background information can be found in *Interim Report No.1*.

All pilots were run with specific reference to:

- *Polygonisation* (RCAHMS *et al* 2009)

The main products of the pilot were:

- *Historic Environment Polygonisation Standards (Scotland)* (RCAHMS *et al* 2011)
  http://www.rcahms.gov.uk/historic-environment-polygonisation-standards-scotland.html
- *"Defining Scotland's Places Technical Guidance"* (RCAHMS et al 2011).

Definitions

1. A **Known Site Extent** polygon defines the limits of known upstanding and prospected remains. The polygon will be created based primarily on sources and professional judgement. Known site extent polygons should not be buffered and should be plotted against the largest map scale available. The shape of the polygon defines the limit of the known extent.

2. A **Discovery Area** polygon defines an area within which information has been recorded but where there is insufficient information to produce a Known Site Extent polygon. The shape of the polygon is arbitrary and may be anything from a grid square to a county boundary. A Discovery Area **does not** define a record's known extent.

3. **Attribution** refers to the attribution attached to a polygon class in a download format. A download format is an exchange format where the polygon data is independent of any associated master database. In a database a polygon is just one attribute among many associated with a record. The current national standard for heritage databases is Midas Heritage (http://www.english-heritage.org.uk/publications/midas-heritage/ [Accessed 20 Jan 2015]). In a download format a polygon must retain sufficient attribution to enable it to be used without ambiguity and it must include the unique identifier of a record in either the local or national record. This technical specification details the minimum attribution required to achieve this.

Known Site Extent   **Discovery Area**

A Known Site Extent (left). A monument's known extent is that shown by the polygon. A Discovery Area (right) where the polygon is an area whithin which a monument is thought to exist but it's exact location is not known.

General principles

1. Metadata for **Known Site Extent** and **Discovery Area** polygons can be found on the *Scottish Spatial Data Infrastructure Metadata Catalogue* at http://scotgovsdi.edina.ac.uk/geonetwork/apps/tabsearch/ [Accessed 18th December 2014] under the Defining Scotland's Places file identifier 4bb609a6-6560-47f6-99c1-280725c01761. It can also be found on the *European Public Data Catalogue* at http://publicdata.eu/dataset/defining-scotlands-places [Accessed 18th December 2014].

2. **Known Site Extent** and **Discovery Area** polygons are produced to the Open Government Licence (OG) (http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/) [Accessed 18th December 2014] and includes polygons also created using the Creative Commons Attribution (CC) - Share Alike 2.0 licence (https://creativecommons.org/licenses/by-sa/2.0/) [Accessed 18th December 2014]. The data can be copied and redistributed in any medium or format and the data can be remixed, transformed, and build upon for any purpose, even commercially as long as credit is attributed and the data is made available under OG or CC licence. Where polygons originate from a CC data source, the source of the data should be identified in the SOURCE field and (CC) or Creative Commons should be included in the NOTES field.

3. **Known Site Extent** and **Discovery Area** polygons must not not use any source that might limit of imposes restrictions on its dissemination. Ordnance Survey data under 50 years old must not be used as a source.

4. **Known Site Extent** and **Discovery Area** polygons have the same attribution structure.

5. **Known Site Extents** define the extent of records. A single record may contain multiple sites.

6. **Known Site Extents** should not be buffered. On occasion, for particularly small sites, a small buffer is needed but buffering should be seen as the exception and any buffer must be identified in the attribution.

7. A **Discovery Area** can be any size. Discover areas may refer to any suitable area from grid squares up to county boundaries however, for clarity, programmes such as Canmore Mapping filter the display to include only discovery areas up to 1000m2.

Attribution

8. The following is a list of the data fields attached to both Known Site Extents and Discovery Area polygons. Field names are limited to eight characters. SHAPE, SHAPE_Length and SHAPE_Area are not listed as they are created automatically by the GIS software.

PID

Full Name	Format	Size	Multiples	Description	Recommendations
Project Identification Number	Auto number	auto	n/a	The project ID guarantees every record is unique. It is created automatically and allows data created in different locations to be combined without fear of data from one source overwriting another.	The PID is not the primary reference number. LAID or NUMLINK (see below) should always be used in preference to PID when referencing a record.

CLASS

Full Name	Format	Size	Multiples	Description	Recommendations
Monument Class	TEXT	254	Multiple values should be comma separated	The CLASS field is the top level thesaurus term. Ideally the data should be sourced from a Monuments Thesaurus.	Controlled data entry: The Scottish Thesaurus of Monument Types. (Available at: http://lmid1a.rcahms.gov.uk/rcahms/apps/f?p=210, accessed 24th March 2010)

TYPE

Full Name	Format	Size	Multiples	Description	Recommendations
Monument Type	TEXT	254	Multiple values should be comma separated	The TYPE field is the second tier thesaurus term.	Controlled data entry: The Scottish Thesaurus of Monument Types. (Available at: http://lmid1a.rcahms.gov.uk/rcahms/apps/f?p=210, accessed 24th March 2010)

POLYTYPE

Full Name	Format	Size	Multiples	Description	Recommendations
Polygon Type	TEXT	25	n/a	The field differentiates between data where the shape of the polygon is meaningful and data where the form of the polygon is an arbitrary shape to enable data discovery.	Controlled data entry: • FORM • DISCOVERY

POLYCLAS

Full Name	Format	Size	Multiples	Description	Recommendations
Polygon Class	TEXT	25	n/a	The field identifies what methodology was used when creating the polygon.	Controlled data entry: • KNOWN SITE EXTENT • DISCOVERY AREA

STATUS

Full Name	Format	Size	Multiples	Description	Recommendations
Monument Status	TEXT	150	n/a	This field provides the user with a guide to the status of the polygon they are consulting. For instance, a user is likely to be more cautious of a polygon with the status, 'Some or all of this area is a Scheduled Monument' compared to one with the status, 'Area of historic environment interest: This area has been identified as likely to contain evidence or objects relevant to the historic environment.'	Controlled data entry: • Some or all of this area is a Scheduled Monument • Some or all of this area is a Inventory Garden and Designed Landscape • Some or all of this area is a Conservation Area • Some or all of this area is a World Heritage Site • This area contains multiple designations • Area of historic environment interest: This area has been identified as likely to contain evidence or objects relevant to the historic environment

CONTACT

Full Name	Format	Size	Multiples	Description	Recommendations
Contact point for more information	TEXT	25	n/a	This field provides the user with guidance on who to contact for more information. The default value should be the local authority archaeological service unless the monument is designated in which case it should be Historic Scotland.	Controlled data entry: • See pick list below

Contact: Pick List
• Historic Scotland
• RCAHMS
• Aberdeen City Council - Sites and Monuments Record (SMR)
• Aberdeenshire Council - Sites and Monuments Record (SMR)
• Archaeology Service: Falkirk Museums
• Archaeology: Dumfries and Galloway Council
• Archaeology: Shetland Amenity Trust
• Comhairle nan Eilean Siar (Western Isles) Archaeology Service
• Dundee Historic Environment Record
• East Dunbartonshire Council SMR
• East Lothian Council Archaeology Service
• Fife Council Archaeological Unit
• Heritage Service: Scottish Borders Council
• Highland Council Historic Environment Team
• Archaeology: Inverclyde Council Planning Department
• Archaeology: Midlothian: Planning and Transport
• Archaeology: North Lanarkshire Council Planning Department
• Archaeology: Orkney Island Council
• Perth and Kinross Heritage Trust
• West of Scotland Archaeology Service (WoSAS)

LAID

Full Name	Format	Size	Multiples	Description	Recommendations
Local Authority Unique ID	TEXT	50	n/a	This field is essential to enable the area data to be linked to more detailed data in local authority databases. However, this field is also problematic as the Scottish historic environment records have a mix of data structures. Some have a numeric LAID while others have a text LAID. There is no	

easy solution. It will be
difficult to link to all the DSP
data until data structures are
standardised across the
sector. As text cannot be
entered into a numeric field,
the project has been forced
to make LAID a TEXT field.

LAID2

Full Name	Format	Size	Multiples	Description	Recommendations
Local Authority Unique ID: Additional	TEXT	254	Multiple values should be comma separated	Occasionally there are multiple records for the same monument. It is outside the remit of DSP to resolve this issue by combining and cancelling records. One of the by-products of this project is that it highlights where data enhancement is needed in the underlying source data.	The record that best describes the monument should be listed in LAID. Additional ID numbers should be listed in LAID2

LAIDLINK

Full Name	Format	Size	Multiples	Description	Recommendations
Link to Local Authority website	HYPER LINK	254	n/a	Where the local authority record is available online, the link should be embedded.	

NUMLINK

Full Name	Format	Size	Multiples	Description	Recommendations
CANMORE unique ID	NUMBER	auto	n/a	This field is essential to enable the area data to be linked to more detailed data in the Canmore database.	

NUMLINK2

Full Name	Format	Size	Multiples	Description	Recommendations
CANMORE unique ID	TEXT	6	Multiple values should be comma separated	Occasionally there are multiple records for the same monument.	The record that best describes the monument should be listed in NUMLINK. Additional ID numbers should be listed in NUMLINK2

LINK

Full Name	Format	Size	Multiples	Description	Recommendations
Link to CANMORE website	HYPER LINK	254	n/a	Where the local authority record is available online, the link should be embedded.	

ACCURACY

Full Name	Format	Size	Multiples	Description	Recommendations
Polygon accuracy	TEXT	15	n/a	This filed gives the user information on how confident they can be in the accuracy of the polygon. The field is linked to the accuracy of the source data. For instance, an area polygon created using survey equipment in the field is more accurate than a polygon created from 19th century 1st edition Ordnance Survey mapping.	Controlled data entry: • Within 1m • Within 5m • Within 10m • Within 25m • Within 50m

SOURCE

Full Name	Format	Size	Multiples	Description	Recommendations
Polygon source information	TEXT	100	n/a	This field gives the user information on what source was used to create the polygon.	Controlled data entry: • See pick list below

Source: Pick list
- Metric survey tied to OD Datum
- HS Scheduled Monument (poly)
- HS IGDL (poly)
- HS Designated Wreck (poly)
- HS WHS (poly)

Accuracy note:
Within 1m, Within 5m or Within 10m
Within 10m
Within 10m
Within 50m
Within 10m

- RCAHMS data (point)   Within 10m, Within 25m or Within 10m
- RCAHMS data (poly)   Within 10m
- RCAHMS field survey (point/line/poly)   Within 1m, Within 5m or Within 10m
- RCAHMS linear (poly)   Within 10m
- RCAHMS AP transcriptions (line)   Within 10m
- RCAHMS areas for OS (poly)   Within 10m
- RCAHMS FESP (point/line/poly)   Within 25m
- RCAHMS HLA (poly)   Within 25m
- RCAHMS 10K record sheets (raster)   Within 10m
- LA data (point)   Within 10m, Within 25m or Within 10m
- LA event (poly)   Within 10m, Within 25m or Within 10m
- LA linear (line)   Within 10m
- LA site area (poly)   Within 10m
- LA Archaeological Area (poly)   Within 10m, Within 25m or Within 10m
- LA Conservation Area (poly)   Within 10m
- OS Mastermap (line/poly)   DO NOT USE
- OS 10K (raster)   DO NOT USE
- OS 6" 1st Ed: 1843-1882 (raster)   Within 25m
- OS County Series, 1st Ed: 1847-1884 (raster)   Within 25m
- OS County Series, 2nd Ed: 1893-1912 (raster)   Within 25m
- OS County Series, 3rd Ed: 1900-1949 (raster)   Within 25m
- OS County Series, Provisional: 1931-1969 (raster)   Within 25m
- 25cm Orthorectified AP   Within 10m (Note caution in areas of extreme height change)
- Lidar and remote sensed
- OTHER

BUFFER

Full Name	Format	Size	Multiples	Description	Recommendations
Buffer size in metres	NUMBER	auto	n/a	This field details the size of any buffer included in a polygon. The field should contain a number in metres or the word BESPOKE capitalised where a buffer of varying width has been used. If no buffer has been used, '0' should be entered as the default entry.	- Do not add buffers unless absolutely necessary. - Buffers should only be added to Known Site Extents - Buffers should never be added to Discovery Areas.

X

Full Name	Format	Size	Multiples	Description	Recommendations
Easting	NUMBER	auto	n/a	This filed should contain a six figure grid reference based on the Ordnance Survey	

OSGB36 datum. The
coordinate entered should
be a single location that best
locates the record. The point
location should sit within the
area polygon.

Y

Full Name	Format	Size	Multiples	Description	Recommendations
Nothing	NUMBER	auto	n/a	This filed should contain a six figure grid reference based on the Ordnance Survey OSGB36 datum. The coordinate entered should be a single location that best locates the record. The point location should sit within the area polygon. Note: Due to the length of Britain some six figure grid references are seven figures long in the Y axis.	

POOLED

Full Name	Format	Size	Multiples	Description	Recommendations
The organisation that has pooled the polygon and populated the attribution.	TEXT	20	n/a	This filed identifies which organisation carried out the initial baseline data pooling process. This field is needed because the DSP set up may be carried out by a different organisation than that which goes on to host and maintain the data. Also, different areas may be pooled by different organisations and any organisational variation in the process can be tracked and corrected using this field.	

NOTES

Full Name	Format	Size	Multiples	Description	Recommendations

Notes	TEXT	254	n/a	Free text notes. (Note: Limited to 254 in shape file download format.) (Note2: May be unlimited MEMO field if data is stored in a database. This will be truncated to 254 if the data is exported in a download format. If this is the case it is important to make sure key notes are stored in the first 254 characters of this field or the notes state: "Long note: Please consult if truncated"	

COMPDATE

Full Name	Format	Size	Multiples	Description	Recommendations
Compilation date	DATE	auto	n/a	The date the polygon is processed. This may also be the date the polygon was created but if this is not known, it is not essential. The compilation date becomes significant once the data becomes actively managed at which point it can be used for statistical studies like tracking the number of new sites added in a given period.	

UPDATE

Full Name	Format	Size	Multiples	Description	Recommendations
Date of last update	DATE	auto	n/a	The date of the most recent update. Knowing the date of update enables different parties to be sure they are using the same information. Different situations will require clarity based on the date of update. For instance, during legal public enquiry the data can be 'frozen' at a certain date to make sure all parties are using the same data	

Discussion

New data

1. New polygons can be created from a range of sources, some of which have not yet been identified. It is particularly important to record the SOURCE and the ACCURACY of the source against all polygons.

Metric Survey

2. Data captured using modern survey equipment and tied to the Ordnance survey datum is likely to be highly accurate. Survey captured in this way should record an accuracy of within 1m. For data tied more loosely to the OS datum use within 5m or within 10m as appropriate.

3. New polygons created as part of a metric survey should use:
    a. "Metric survey tied to OD Datum" in SOURCE and
   b. Name of the organisation creating the polygon in NOTES.

Unusual situations and exceptions

Linear features

4. Linear features can be very long. This makes digitising them difficult and the data awkward to deliver when sections of the data are downloaded or used. To resolve these issues, linear monuments should be polygonised separately in each 5 km gradicule (grid square). The polygons should contain exactly the same information except for the grid reference which should be specific to that gradicule. The resulting linear monuments composed of multiple separate polygons should not be converted into a multipart polygon. Selecting the monument by name or unique number will return the entire monument. Selecting the record by grid square will return only records for the grid squares selected.

Grouped structures

5. Multipart polygons should be used where distinct monuments are grouped and more than one polygon is needed to represent a single record.
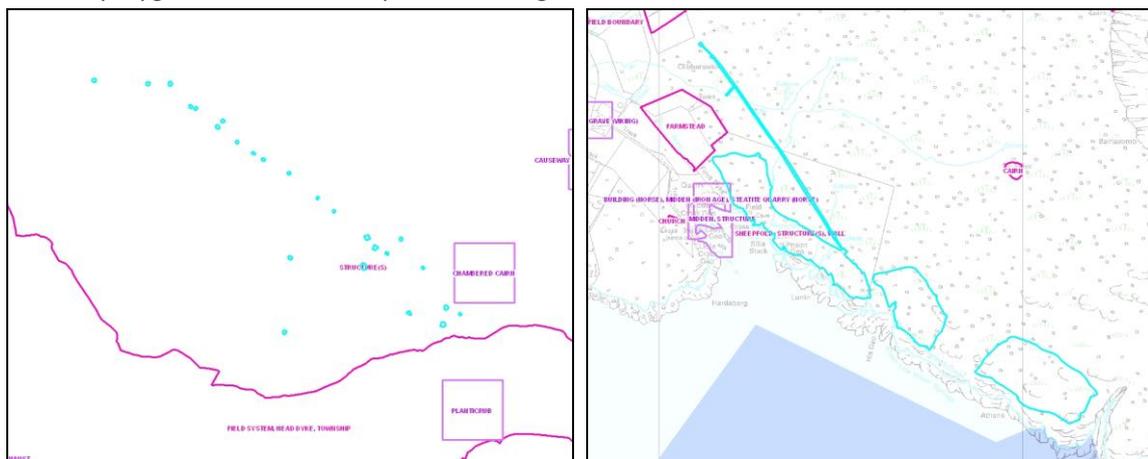
Two examples (light blue) of where multipart polygons have been used to define the extent of a single record.

Landscape features

6. On occasion, a single record has been used to describe both very large defined landscapes and individual monuments within those landscapes. Examples of this encountered during the pilot included:
   a. Townships defined by a hill dyke and the farmsteads and structures in the township like mills, nausts and field systems.
   b. Country estates and the buildings, lodges and gatehouses once on the estate.
   c. Defence sites and the buildings, fuel dumps and gun emplacements on the site.
   d. Industrial sites and the individual factories, railways, canals and quays that serviced the site.
7. The individual monuments in the landscape should be polygonised as a multipart polygon grouped as described in *Grouped Structures* above.
8. The landscape feature should be polygonised as a single entity.
9. Both the landscape feature and the grouped structures polygon should be given the same attribute information.

Maritime

10. Where the Known Extent of a maritime record can be defined this should be created.
11. Discovery Areas should not be produced for maritime records.

Display recommendations

12. Point data should be displayed down to 1:50,001:
13. Polygons should display at 1:50,000
14. Captions should default to TYPE and should display at 1,10,000
15. One caption should be shown per polygon (note: for mulitpint polygons this will be one polygon for the entire group).
16. Symbology and labelling:

	Label	Colour	Font	Expression	Scale Range
				(see attribution)	Do not show out beyond
KNOWN SITE EXTENT	Y	Ginger Pink	Arial Bold 9 pt	CLASS	1:10,000
DISCOVERY AREA	Y	Heliotrope	Arial Bold 9 pt	CLASS	1:10,000