ScAPA

Scottish Archaeological Periods & Ages Implementation Report

03 July 2018

Ceri Binding, Hypermedia Research Group, University of South Wales

ceri.binding@southwales.ac.uk

University of South Wales
Prifysgol De Cymru

Table of Contents

1 Introduction

Archaeologists in Scotland have in the past experienced difficulty reaching a uniform consensus concerning historical periods and date ranges and there is no single agreed published list of Scottish periods for use in consistently indexing cultural heritage datasets of artefacts, monuments and sites. Such lists do exist in England and other countries and there is a need for initiatives to provide a possible way forward.

The process of *periodization* involves the subdivision of time into discrete named blocks representing periods, utilising knowledge of past events. Documenting named periods and associated date ranges within an openly available published periodization would be a useful initial step towards this goal. Location-specific periodization is often used in addition to absolute dating to contextualize cultural heritage records. Dates, periods and timespans need to be associated with geographical and cultural context to give them relevance, and periodization can provide such context by referencing material technologies, significant events, cultures, leaders/monarchs etc. One such example is the Historic England (HE) period list[1] - a particularly selective curation of terms from various categories including:

- Industrial stages / 3 age system (e.g. *Iron Age*)
- Regnal periods (e.g. *Victorian*)
- Ordinal centuries (e.g. *20th Century*)
- Wars (e.g. *First World War*)
- Cultures (e.g. *Roman*)

Described[2] as a list of "*…the main archaeological and cultural periods of Britain*", it represents a short and coarse grained periodization relating to certain aspects of British (or at least English) cultural history, and is often presented as a de-facto standard for the temporal orientation of UK cultural heritage data. However, whilst many of the concepts may be generally applicable to the UK the associated dates present may be debatable when applied specifically to Scottish chronology (E.g. Iron Age phasing). Gaining national consensus on a single standard periodization covering the whole of the UK is understandably difficult, however producing a region-specific periodization is possible. A number of regional research frameworks are being undertaken or considered under the Scottish Archaeological Research Framework (ScARF) [1] and a recent ScARF panel report described a summary chronology for the Scottish Bronze Age, together with a research recommendation that "*A specifically Scottish chronology that relates chronological data from many sources needs to be developed*" [2] ; other panel reports mention region-specific period terms such as "*Scottish Neolithic*" [3] and "*Scottish Iron Age*" [4] . These panel reports also contain useful initial pointers to bibliographic reference sources concerning Scottish chronology.

1.1 Project proposal

An application for funding for a short project was submitted to Historic Environment Scotland by the University Of South Wales (USW) Hypermedia Research Group, and the funding proposal was approved in December 2017. This short project would not go so far as to produce a detailed chronology of the whole of Scottish history, but rather a periodization as an extensible framework to provide relative context for dating of events and artefacts. The proposal was to compile and publish online a reference periodization incorporating period data specifically related to Scotland that could be useful for categorization and temporal orientation of Scottish cultural heritage records. Such a periodization could not in itself resolve any existing fundamental disagreements on phasing and

[1] HE period list http://heritage-standards.org.uk/wp-content/uploads/2015/08/Periods-List-HE-FISH-WP.pdf
[2] http://www.heritage-standards.org.uk/fish-vocabularies/

dating, but would at least provide a way forward in presenting the concepts and possibly accommodating multiple opinions via the scope notes for the dates identified. The outcome of the project would be an extensible body of interlinked historical terms from which subsets of data may be extracted in multiple serialization formats, and used as the basis for various visualizations.

1.2 Project scope

The concepts to be included would primarily be drawn from practical evidence of usage in indexing of cultural heritage resources. It was envisaged that cultural terms such as *Pictish* and *Norse* might be included in a Scottish periodization, having a spatial coverage property restricting the geographical context of the concepts to Scotland, together with finer grained scope notes to describe how these concepts might relate to particular Scottish regions e.g. the period of *Norse* control was generally relinquished in the three years following the *Battle of Largs* (1263) - but in Shetland and Orkney it ended by 1472, over 200 years later. Other significant events might also be included to give the major archaeological/cultural periods some additional context (e.g. the houses of *Alpin* and *Dunkeld*).

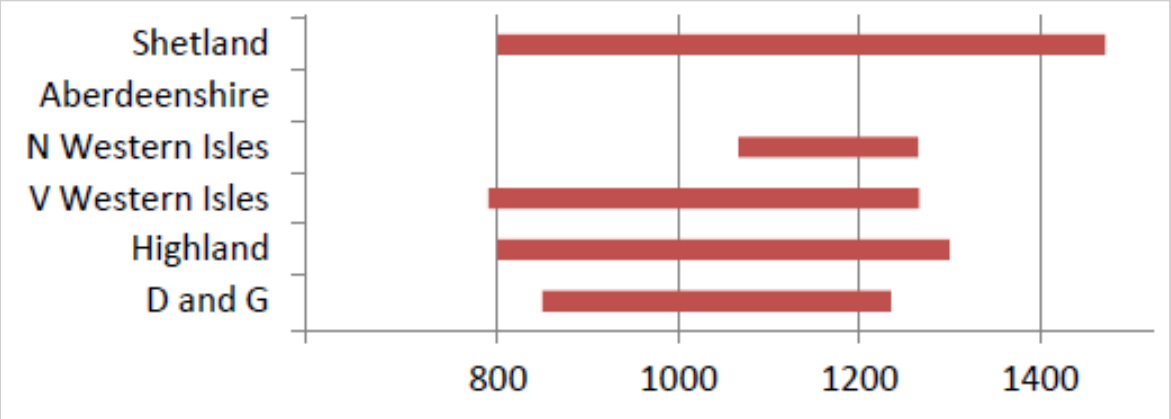

Figure 1: Norse / Viking periods occurring in different regions at different dates

Clearly there would be some conceptual overlap with any existing HE period list concepts that truly have UK-wide scope (e.g. ordinal centuries), however due to the selective nature of the HE list there are only two ordinal centuries present (*20th Century, 21st Century*) - the Scottish list in comparison would contain a full set of ordinal centuries from *1st Century CE* through to *21st Century CE* inclusive, together with associated Early/Mid/Late subdivisions. All concepts produced by this project would specify Scotland as the spatial coverage attached to each concept and so do not purport to represent the whole of the UK.

1.3 Linked Open Data

In recent years Linked Open Data (LOD) has emerged as a prominent mechanism to make data available online as part of the semantic web for consumption by applications. A previous USW project (SENESCHAL, [5]) published a series of LOD reference vocabularies in SKOS RDF format. SKOS[3] (Simple Knowledge Organization System) is a standard RDF data model *"…to support the use of Knowledge Organization Systems such as thesauri, classification schemes, subject heading lists and taxonomies within the framework of the Semantic Web"*. The LOD vocabularies published during the SENESCHAL project originated from Historic England (HE), Historic Environment Scotland (HES) and the Royal Commission on Ancient and Historical Monuments of Wales (RCAHMW), including period vocabularies

[3] SKOS Simple Knowledge Organization System https://www.w3.org/2004/02/skos/

for both England[4] and Wales[5]. This project proposed to produce an equivalent period vocabulary for Scotland, and to publish the resultant data onto the same platform in the same format.

1.4 Perio.do

Perio.do [6] is a "*gazetteer of period definitions for linking and visualizing data*". The periods in this gazetteer have geographical context and globally unique identifiers to reference and distinguish them. Each period is effectively a spatio-temporal concept representing an intersection of subject, place & time. Archaeological periods do not typically have 'hard edges' with absolute dates, so each period may optionally include a start period and an end period, to represent transitions or uncertainty. Table 1 shows some examples of periods representing the *Bronze Age* in the context of the United Kingdom as listed in the Perio.do canonical dataset:

Identifier	Label(s)	Original Source	Spatial Coverage	Start	End
ark:/99152/p0gjgrs6qb2	Bronze Age	Portable Antiquities Scheme	United Kingdom	2350 BC	801 BC
ark:/99152/p0kh9ds7q8m	Bronze Age	HeritageData.org (Historic England Periods List)	United Kingdom	2600 BC	700 BC
ark:/99152/p0zj6g8rbvk	Bronze Age	ARENA Portal	United Kingdom	2500 BC	700 BC

Table 1: Examples of existing UK Bronze Age periods in Perio.do

Scotland currently has only limited independent representation in Perio.do, and although existing period concepts such as the examples shown in Table 1 purport to represent the whole of the United Kingdom, a lack of contextual detail (e.g. scope notes) coupled with contradictory start/end dates makes it difficult to align these periods to Scottish history. The Perio.do dataset contains many such terms related to the UK, but just 13 records are specific to Scotland (originating from LCSH - mainly house of Stewart regnal periods). The proposal included provision for data produced in this project to be submitted as a patch intended for inclusion in the canonical Perio.do dataset.

2 Methodology

2.1 Data collection

Listings of period terms in use for the description of archaeological artefacts, monuments and sites were solicited from a number of Scottish Historic Environment Record (HER) offices. These lists were received in Excel spreadsheet form, and aggregated into a composite spreadsheet from where decisions on term inclusion and final wording could be made. Some lists also contained dates and additional explanatory comments.

2.2 Data conversion

The composite spreadsheet produced was specifically formatted for use by the SKOS-Play service [7] allowing conversion of the data to SKOS RDF format. Multilingual listings of the concepts in alphabetical and hierarchical arrangements were also produced using the SKOS-Play service, together with various experimental graphical visualisations. These listings aided understanding of what was being produced during the course of the project as they were posted online for interested parties to examine, in addition to the raw underlying SKOS RDF data.

[4] LOD publication of HE periods list http://purl.org/heritagedata/schemes/eh_period
[5] LOD publication of RCAHMW periods list http://purl.org/heritagedata/schemes/11

Scope notes tailored to Scotland were produced and included within the composite spreadsheet. Once the wording of these scope notes was reviewed and became more stable they could be translated to include Scottish Gaelic versions of the notes. It should be noted here that whilst some scope notes were included in the initial ScAPA vocabulary, due to the very short timescale available many were not completed before the formal end of the project. Since that time draft Scope notes have been published on the ScARF website for open peer review and editing. Once the scope notes have been finalised, it is hoped to translate into Gaelic then refresh the Scapa resource on HeritageData.org

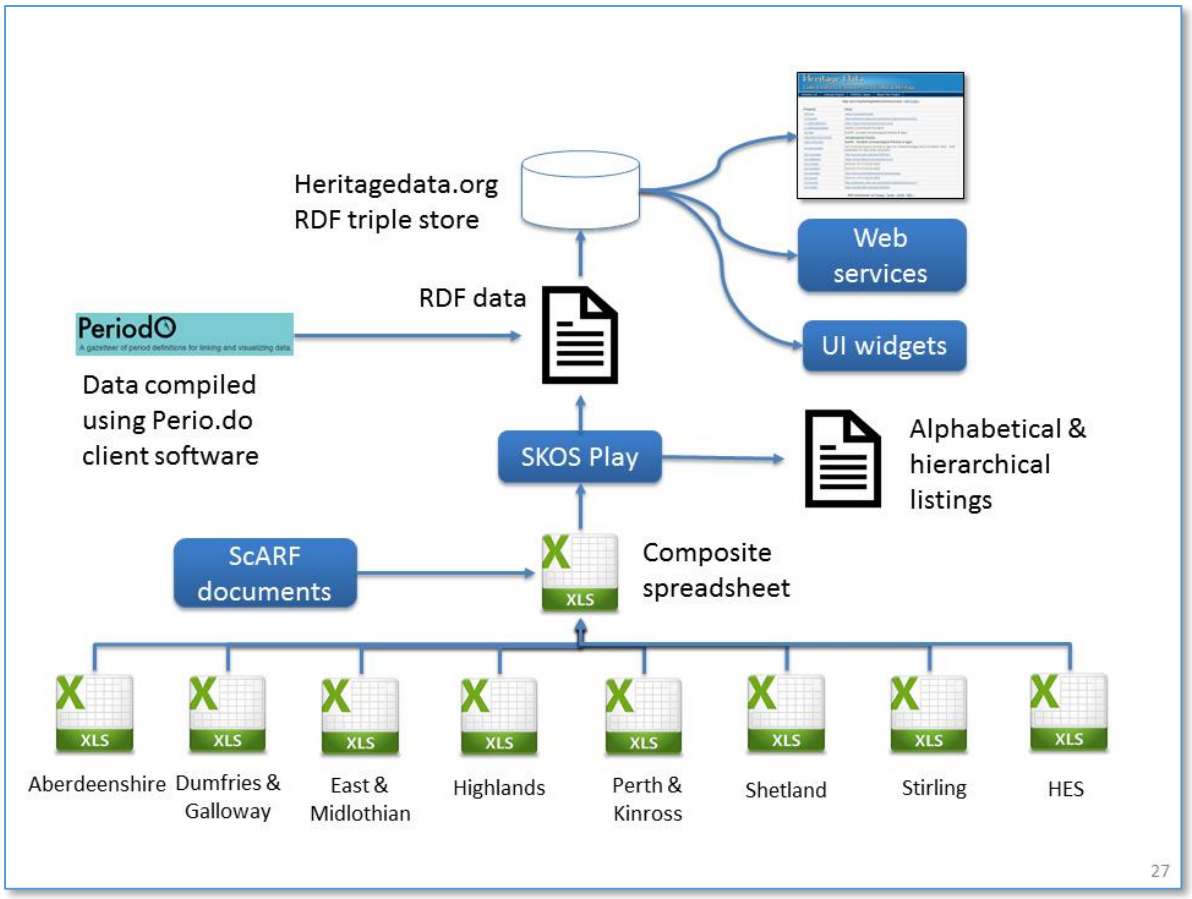

Figure 2 : ScAPA project – overview of general workflow

Although not fully distinguished within the data structure, a number of distinct period classes were observed in the received lists, exhibiting similarities to the period classes previously identified in the HE period list:

- Three age system e.g. *Bronze Age, Long Iron Age*
- Cultural terms e.g. *Pictish, Viking, Saxon, Anglian*
- Regnal periods e.g. *Georgian, Antonine, Robert I*
- Ordinal centuries e.g. *5th Century*
- Wars e.g. *First World War*
- Other terms specific to Scottish history e.g. *Industrial, Improvement, Crofting*
- Non-periods e.g. *Period unknown, Period unassigned*

In some instances the HER period lists received contained terms that were not actually periods at all e.g. *Nil-Antiquity, Period unknown, period unassigned*. Clearly though these are still important terms

used in practice in the context of temporal fields to indicate an absence of evidence or other information. These terms were therefore included in the ScAPA vocabulary to reflect actual usage.

Some dual periods were identified in the originally received term lists e.g. *Neolithic / Bronze Age*. There is ambiguity as to whether this refers to a larger period encompassing the extents of both - or whether it means a smaller transitional period between the two. Lacking information on the precise original usage of this and other similar terms it was decided not to include dual period terms in the initial publication of the ScAPA list. This could be revisited if it is felt that such terms are of some value for inclusion.

***Recommendation**: ambiguous terminologies should not be used. Where there is an uncertainty, for instance 'the building is 18th or 19th century in date' or 'this late Neolithic to Bronze Age Stone Circle', database users should index each term separately.*

2.3 Concept temporal boundaries

Century boundaries have been the subject of repeated debate over many years, the issue concerns whether centuries start at year zero or year 1. Bibliographic evidence [8] demonstrates similar arguments being rekindled on the cusp of each of the previous three centuries. The approach taken in the ScAPA project was to align the century bounds with the few existing Historic England century concepts present – so AD/CE centuries starting at year 1, ending at year 100. E.g. *5th Century CE* = 401 CE → 500 CE inclusive.

Centuries are then subdivided as *Early*, *Mid* and *Late*. These subdivisions were interpreted in terms of years as follows:

- *Early*: 1 → 33 E.g. *Early 5th Century CE* = 401 CE → 433 CE
- *Mid*: 34 → 66 E.g. *Mid-5th Century CE* = 434 CE → 466 CE
- *Late*: 67 → 100 E.g. *Late 5th Century CE* = 467 CE → 500 CE

Alternative dates for these subdivisions could be equally valid here due to the fuzzy nature of exactly what the terms *Early/Mid/Late* actually mean. Overlapping subdivisions of 1→ 40, 30 → 70 and 60 → 100 respectively might also be considered reasonable - so e.g. *Mid-5th Century CE* would then become 430 CE → 470 CE.

Temporal bounds for other concepts were obtained from the received spreadsheets, explaining regional differences within the scope notes. This information was input to the Perio.do client software as a local resource.

Figure 3: Extract of ScAPA data entered in the Perio.do client during development

The Perio.do client produces data conforming to a model that is a blend of SKOS RDF and the W3C Time ontology[6]. This data was then exported to be merged with the existing RDF produced by the SKOS-Play service.

Recommendation: *for consistency century boundaries should be defined as starting at the year 1 and ending at the year 100.*

2.4 Gaelic translations

As one strand of previous project work on the SENESCHAL project [5] Gaelic translation of terms and scope notes for the Scottish Monuments Thesaurus had been undertaken on behalf of Historic Scotland with funding provided by Bòrd na Gàidhlig. This data was converted to SKOS format and included in the Linked Open Data published on HeritageData.org. The inclusion of Gaelic terms at the time provided an illustration of the usefulness of the SKOS concept-based approach. An example

[6] W3C Time Ontology http://www.w3.org/2006/time#

Scottish Monument Types thesaurus concept http://purl.org/heritagedata/schemes/1/concepts/447 is expressed in an English preferred term as *SCHOOL HOUSE* and in Scottish Gaelic as *TAIGH-SGOILE*, and a Scottish Gaelic scope note is also present for this concept.

A similar data structure and approach was taken for the ScAPA dataset, and a few preliminary Scottish Gaelic term translations obtained via Google translate were included in the draft dataset merely as temporary placeholders for illustration and discussion during the workshop. Google translate is obviously not a viable mechanism for the production of quality translated resources, and it was acknowledged that whilst there were not necessarily the resources available immediately to provide valid Gaelic translations for all ScAPA terms and scope notes, it remained a firm aspiration to include these in the dataset at some point in the near future.

Recommendation: *Historic Environment Scotland provide authoritative translations of terms and scope notes used in the ScAPA project.*

3 Outcomes

3.1 Publication as Linked Open Data

The initial composite spreadsheet data (160 concepts) was converted to SKOS RDF via the SKOS-Play service [7] and published on HeritageData.org as a new Scottish period list - ScAPA [9] . Linked Open Data concepts have unique and persistent web-accessible identifiers (URI) e.g.

- http://purl.org/heritagedata/schemes/scapa/concepts/PA (*Palaeolithic*)
- http://purl.org/heritagedata/schemes/scapa/concepts/NO (*Norse*)
- http://purl.org/heritagedata/schemes/scapa/concepts/PI (*Pictish*)
- http://purl.org/heritagedata/schemes/scapa/concepts/IMP (*Improvement*)

Each concept has a series of descriptive properties and links to further linked data resources as shown in Figure 4. This data is downloadable in a number of different RDF serialisation formats for direct programmatic integration and usage (NTriples, Turtle, JSON, XML).

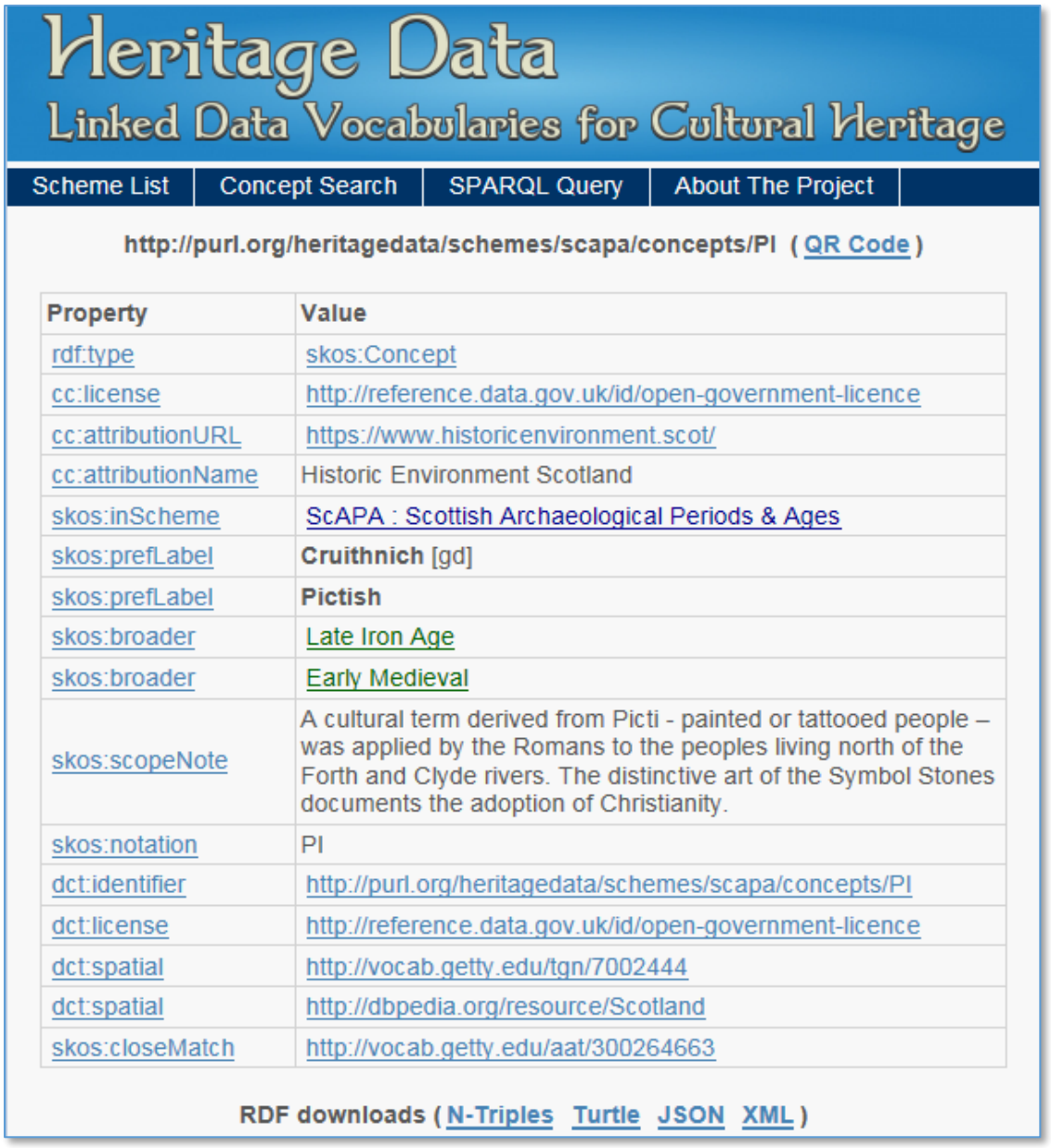

Figure 4: Example ScAPA concept data displayed via the HeritageData.org site

The data compiled would also be used to supplement the canonical Perio.do dataset. Alphabetical and hierarchical listings[7] of the ScAPA concepts, relationships and scope notes together with a listing of Gaelic-English equivalent terms were generated from the SKOS RDF output using the SKOS-Play service. Note the Gaelic terms appearing here (and in Figure 5) are only intended as placeholders showing how Gaelic would be accommodated in the dataset – they have not been provided or even verified by a native Gaelic speaker and will be updated accordingly as and when such resources are available.

[7] ScAPA vocabulary - alphabetical and hierarchical listings
http://heritagedata.org/data/hes/hes_scapa_20180329_alpha.pdf (PDF, 732 KB)

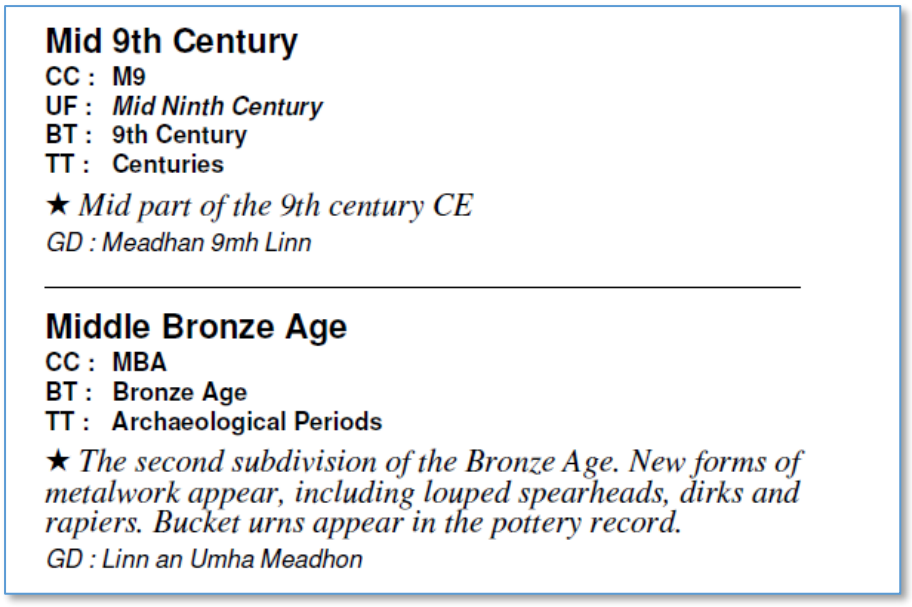

Figure 5: Excerpt from the alphabetical listing of the ScAPA vocabulary

Mappings to external vocabulary resources were manually made wherever possible linking the newly published ScAPA LOD concepts to concepts within both the HE and RCAHMW period lists, and also to Getty AAT LOD concepts. Each concept also has a spatial context specified, with links out to both the DBPedia and Thesaurus of Geographic Names (TGN) concepts for Scotland.

Source Concept URI	Match relationship URI	Target concept URI
scapa:PA	skos:closeMatch	http://vocab.getty.edu/aat/300019258 http://purl.org/heritagedata/schemes/eh_period/concepts/PA http://purl.org/heritagedata/schemes/11/concepts/505122

Table 2: Example inter-thesaurus concept mappings for the ScAPA concept of "Palaeolithic"

3.1.1 Update of existing Scottish LOD vocabularies

A general data refresh of all HES data on HeritageData.org was included in the original project proposal, and this involved obtaining a new download from the vocabulary component of the live HES database to update the HeritageData.org LOD representations. The received data was converted using the USW STELLAR application [10] as this was the approach used in the initial creation of these LOD resources. New SKOS RDF representations of the following existing LOD thesauri were created and uploaded:

- **Scottish Monument Types Thesaurus** (http://purl.org/heritagedata/schemes/1) – 2,723 concepts; 40,705 RDF triples. The update work also involved retaining the supplementary Gaelic translations that had already been previously produced (see section 2.4), as these translations are not included within the thesaurus data in the HES database.
- **Scottish Archaeological Objects Thesaurus** (http://purl.org/heritagedata/schemes/2) – 469 concepts; 6,152 RDF triples.
- **Scottish Maritime Craft Thesaurus** (http://purl.org/heritagedata/schemes/3) – 202 concepts; 2,555 RDF triples.

The opportunity was also taken during this exercise to revise the existing HeritageData.org web pages referencing the older 'RCAHMS' to refer to 'Historic Environment Scotland' instead.

3.1.2 Perio.do submission

The temporal elements of the ScAPA concepts had been compiled using the Perio.do client application as a local database (see section 2.3). This separate data was then exported in RDF format from the Perio.do client application for merging with the HeritageData.org LOD, and is intended to also be submitted as a "patch" for inclusion in the main canonical Perio.do dataset.

3.2 Dissemination activities

A presentation describing and promoting ScAPA (and other work) was made to the FISH-HEIRNET Terminology Working Group at the National Library of Wales in Aberystwyth during their biannual meeting on Tuesday 6th February 2018. Feedback was generally encouraging, one interesting audience observation was that HER's often have access to lists of period terms that people have actually searched for, which may be quite different to the terms used in controlled term indexing, but empirically curated query terms could be used to supplement the controlled entry vocabulary in a period thesaurus in order to facilitate more flexible searching on thesaurus indexed resources.

A full workshop was then held on Thursday 15th March 2018 at the Historic Environment Scotland offices in Edinburgh. The workshop was well attended with presentations from HES, HE and ScARF, ADS, and USW. The opening presentations provided a context for the current research, namely to provide authoritative reference vocabularies for the redevelopment of OASIS and to encourage interoperability between OASIS and the emerging Regional Research Frameworks. The USW ScAPA presentation generated a good level of interest and debate around the subject of Scottish periods, there was particular enthusiasm shown for the demonstrated principle of including Gaelic translations of terms and scope notes. The audience also felt that the definition and date outputs should be shared with the education and interpretation wings of HES and promoted through ScARF. Chronologies play a central part in people's understanding of sites and material – it's therefore important the dates and periods being supplied as part of this work are accurate, conform and are consistent.

3.3 ScAPA usage with existing tools and services

The existing SENESCHAL project had previously produced a number of useful web services and embeddable user interface widget controls, and at the time of writing these still continue to function. Once the ScAPA SKOS RDF dataset had been uploaded to the RDF triple store on the HeritageData.org server, the new data became immediately available and usable within the context of these existing services and widgets.

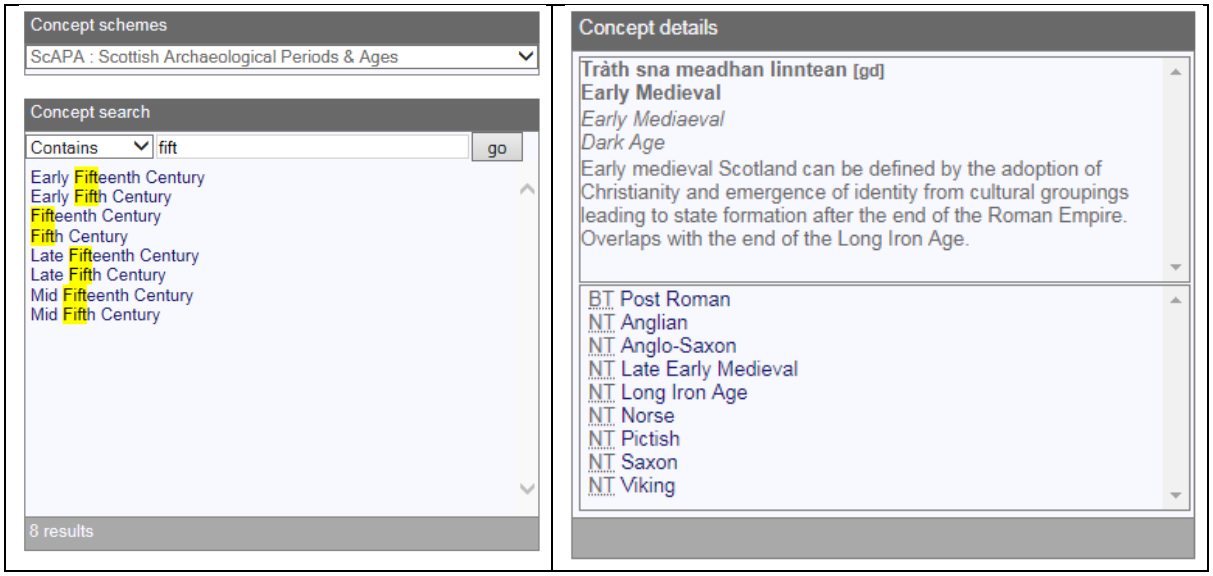

Figure 6 : ScAPA data working within existing SENESCHAL widget controls

Figure 6 shows a search being conducted on ScAPA terms using the original SENESCHAL 'Concept Search' user interface widget control (left hand pane), and the details of an individual ScAPA concept being displayed in the SENESCHAL 'Concept Details' user interface widget control (right hand pane).

Web service call

http://www.heritagedata.org/live/services/getConceptLabelMatch?schemeURI=http://purl.org/heritagedata/schemes/scapa&startsWith=EARLY&limit=3&pretty

Web service response content (JSON)


```
[
  {
    "uri":"http://purl.org/heritagedata/schemes/scapa/concepts/EM",
    "label":"Early Medieval",
    "label lang":"en"
  },
  {
    "uri":"http://purl.org/heritagedata/schemes/scapa/concepts/EBA",
    "label":"Early Bronze Age",
    "label lang":"en"
  },
  {
    "uri":"http://purl.org/heritagedata/schemes/scapa/concepts/E01",
    "label":"Early 1st Century",
    "label lang":"en"
  }
]
```


Figure 7: ScAPA data retrieved using an existing SENESCHAL web service call

Figure 7 demonstrates use of one of the existing SENESCHAL web services to obtain ScAPA data programmatically – searching within only the ScAPA concept scheme to retrieve the first three concepts found having terms starting with the word 'EARLY'.

As part of USW contribution to the ARIADNE FP7 project [11] a Vocabulary Matching Tool[8] had been produced mainly for the purpose of mapping LOD vocabularies to Getty AAT concepts, but it also facilitated dynamic search, visualisation and navigation of the LOD vocabularies. This utility tool (by default) points to the HeritageData.org SPARQL endpoint, so again the new ScAPA data became immediately available and usable within the context of this existing tool.

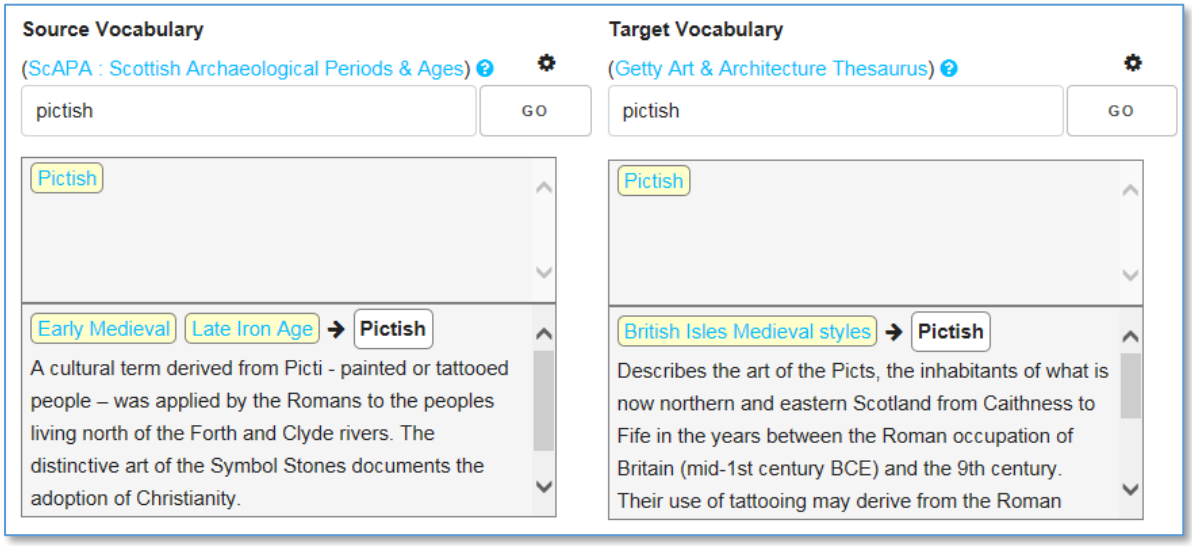

Figure 8: ScAPA data displayed within an extract of the USW Vocabulary Matching Tool

Figure 8 shows ScAPA data displayed in the context of the Vocabulary Matching Tool, comparing the ScAPA concept of *Pictish* to a (potentially) similar Getty AAT concept. This side by side comparison allows a more comprehensive review for concept mapping purposes - by studying the associated scope notes and conceptual relationships in addition to the terms used.

4 Conclusions

Although restricted in scope and timescale the ScAPA project nevertheless represents a useful initial step in the production of bilingual Linked Open Data reference periodization resources specifically related to Scotland. Scope notes and translations do require expert input however these can be added into the existing online resources as and when they can be made available.

Once published as Linked Open Data the new vocabulary resources were immediately applicable in a multiplicity of contexts including the existing SENESCHAL web services & embeddable user interface widgets (deployed 5 years prior to ScAPA), and the Vocabulary Matching Tool (deployed 3 years prior to ScAPA). This was fully anticipated and illustrates some of the many advantages of adopting the use of standard knowledge organization formats, web technologies and Linked Open Data approaches.

5 Further Work

Future work will involve supplementing, improving, extending and translating the initially produced vocabulary with scope notes, and the inclusion of Gaelic equivalents for all terms. Should there be particular demand for it the vocabulary could potentially be extended for additional context to include the names of significant battles, or names of Scottish rulers/leaders for dates preceding the *House of Alpin* (if the information was available) - however there was in practice no particular evidence from

[8] Vocabulary Matching Tool https://github.com/cbinding/VocabularyMatchingTool

the received spreadsheet lists that such terms are currently used to represent periods in the temporal indexing of Scottish cultural heritage resources.

New demonstration applications could be developed; an interactive graphic representation of the timeline would be a useful entry point to the data. Section 3.3 also describes a few examples of using the ScAPA data in the context of user interface components and services that could be further improved and extended, e.g.

- Query expansion services for use in improving online search. E.g. a search on "*Late Prehistoric*" should also retrieve resources indexed as "*Chalcolithic*", "*Neolithic*" or "*Pictish*".
- Indexing tools to 'nudge' cataloguers towards the use of appropriate SCAPA concept identifiers, rather than relying on (potentially ambiguous) free text.
- Auto-indexing tools to suggest SCAPA concepts annotations for free text description fields. This could be combined with systematic extraction of numeric timespans from the many variant expressions commonly in use (with any output suggestions subject to expert intellectual review).

Such applications and services could serve to improve visibility and encourage take-up and usage of the ScAPA reference vocabulary resources.

6 Links

[1] ScARF: Scottish Archaeological Research Framework http://www.scottishheritagehub.com/
[2] Downes, J (ed.) (2012). *Bronze Age Panel Report*. Scottish Archaeological Research Framework: Society of Antiquaries of Scotland, p.15. http://tinyurl.com/clxgf5s
[3] Downes, J (ed.) (2012). *Neolithic Panel Report*. Scottish Archaeological Research Framework: Society of Antiquaries of Scotland, p. 117. http://tinyurl.com/d73xkvn
[4] Fraser, H & Carruthers, M (eds.) (2012). *Iron Age Panel Report*. Scottish Archaeological Research Framework: Society of Antiquaries of Scotland, p.107. http://tinyurl.com/cx4nlt8
[5] USW SENESCHAL project (2013). *Semantic ENrichment Enabling Sustainability of arCHAeological Links.* http://hypermedia.research.southwales.ac.uk/kos/SENESCHAL/
[6] Perio.do: *A gazetteer of period definitions for linking and visualizing data* http://perio.do/
[7] Francart, Thomas (2017*). SKOS-Play data conversion service.* http://labs.sparna.fr/skos-play/
[8] Freitag, Ruth S (1995). *The Battle of the Centuries – A List of References.* ISBN 978-9998204751 https://www.loc.gov/rr/scitech/battle.html
[9] USW ScAPA project (2018). *Scottish Archaeological Periods & Ages*. http://purl.org/heritagedata/schemes/scapa
[10] USW STELLAR project (2011). *Semantic Technologies Enhancing Links and Linked data for Archaeological Resources.* http://hypermedia.research.southwales.ac.uk/kos/stellar/
[11] ARIADNE FP7 project (2017). http://www.ariadne-infrastructure.eu/