SPARQLing Ogham Stones: New Options for Analyzing Analog Editions by Digitization in Wikidata

Sophie C. Schmidt[1]

Florian Thiery[2]

[1] German Archaeological Institute (DAI)
Berlin, Germany
[2] Römisch-Germanisches Zentralmuseum,
Leibniz Research Institute for Archaeology (RGZM),
Mainz, Germany

Abstract

Linked Open Data (LOD) should be the standard for the modeling and publication of research data, as it greatly improves the research process by facilitating access and reducing the likelihood of mistakes. To encourage our readers to create LOD, we have developed a workflow for the enrichment of datasets taken from analog and digital sources by interlinking and publishing them in Wikidata. This paper will discuss the main principles of LOD, as well as current attitudes towards it in the digital humanities. Our vision for LOD has lead us to create SPARQL Unicorn, a series of tools designed to make Wikidata easier to use. As an example of SPARQL Unicorn in action, the second half of our paper will describe the Wikidata integration of the inscriptions found on the early medieval Ogham stones. Sources, workflow, tools, and data models will be explained in detail, as two examples of (geo-) statistical analysis showcase the benefits of enriched LOD to researchers.

In: Tara Andrews, Franziska Diehr, Thomas Efer, Andreas Kuczera and Joris van Zundert (eds.): Graph Technologies in the Humanities - Proceedings 2020, published at http://ceur-ws.org.

1 Introduction

This paper introduces a workflow for digitizing analog editions in the open access data storage facility Wikidata[1], and shows how the enriched data can be extracted and used for further research. For this purpose, we draw upon the Corpus Inscriptionum Insularum Celticarum (CIIC), the most complete analog catalog of Ogham stones assembled by R. A. S. Macalister (1945, 1949), the CISP project,[2] and the Ogham in 3D Project.[3] As we will show, the SPARQL Unicorn tool framework enables users unfamiliar with the SPARQL language to work with data stored in Wikidata. Our analysis of the created dataset by means of the SPARQLing Unicorn QGIS Plugin and the statistical scripting language R demonstrates the viability and usefulness of such an approach.

The SPARQL Unicorn tool framework and the ᚑᚌᚔ (Ogi) Ogham Project are developed and maintained by the Research Squirrel Engineers[4] research network and specific working groups. The Research Squirrel Engineers are an open and international association of researchers with a background in software engineering, geoinformatics, and cultural heritage, which was founded by the authors of this paper, Sophie C. Schmidt[5] and Florian Thiery[6]. Its members cooperate with other researchers and institutions to implement research projects and research software based on Linked Data and the Semantic Web, without institutional and funding dependencies.

This chapter begins with some basic information concerning the inscriptions on the Ogham stones and the Ogham alphabet itself. We will then introduce Linked Open Data (LOD) as a data modeling concept, and Wikidata and Wikimedia Commons as examples of this Open Science approach. Our vision of LOD lays the groundwork for the section on SPARQL Unicorn, a tool framework designed to facilitate the use of Wikidata. Last, but not least, we will describe the Ogi Ogham Project in detail: our sources, the workflow of the project, the import of data into Wikidata, and the modeling process.

[1] https://wikidata.org

[2] https://www.ucl.ac.uk/archaeology/cisp/database/manual/

[3] https://ogham.celt.dias.ie

[4] http://squirrel.link

[5] Research Associate at the Deutsches Archäologisches Institut (NFDI4Objects)

[6] Research Software Engineer at the Römisch-Germanisches Zentralmuseum, Department of Scientific Information Technology

2 What is Ogham?

Ogham (Old Irish: Ogam) stones (Figure 1) are monoliths inscribed with the Ogham script (Ferguson, 1864; Graves and Limerick, 1876), which is an early medieval alphabet used predominantly in Ireland and the western parts of Britain between the fourth and the ninth centuries CE. The so-called 'orthodox' inscriptions were employed from the fourth to the sixth centuries to record the Primitive Irish language, of which the stones are the earliest sources, whereas the 'scholastic' inscriptions of the sixth to the ninth centuries were used to represent Old Irish. The script was probably inspired by the Latin alphabet and contains 26 letters written from bottom to top, left to right, and top to bottom, in the form of a boustrophedon. Each letter is composed of strokes placed along a central line. Their relation to one another and the angle towards this central line, for which the ridge or natural arris of the stone is often used, encode their meaning (MacManus, 1997).

The texts found on Ogham stones mostly refer to persons, tribes, or family affiliations (Macalister, 1945). Since graph databases rely on linking statements between two entities to form a network, the formula used to record kinship relations on the stones makes this corpus especially well-suited to being represented in a graph form along the lines of X *MAQI* Y → X son of Y (e.g. Q69389090)[7] or X *MAQI MUCOI* Y → X son of the tribe Y (e.g. Q69388229)[8].

3 Introduction to Linked Open Data

The term 'Linked Open Data' describes the goal of interlinking data within the World Wide Web with other semantically related data. Tim Berners-Lee has proposed the Linked Data principles consisting of four simple rules (Figure 2):

1. Use URIs as names for things
2. Use HTTP URIs so that people can look up those names
3. When someone looks up a URI, provide useful information using the standards RDF* and SPARQL
4. Include links to other URIs, so that users can discover more things

However, the shortcomings that Berners-Lee pointed out in 2006 remain a challenge today: "a surprising amount of data isn't linked [...] because of problems with one or more of the steps." (Berners-Lee, 2006)

The concept of the Semantic Web as established by Berners-Lee (Berners-Lee et al., 2001) draws on Open Data to describe the web resources, links,

[7]https://www.wikidata.org/wiki/Q69389090

[8]https://www.wikidata.org/wiki/Q69388229

and machine-readable interfaces that are required for creating a Giant Global Graph (Berners-Lee, 2007). Applying LOD techniques helps in creating FAIR (findable, accessible, interoperable, reusable) (Wilkinson et al., 2016) data. Yet publishing data with unique identifiers on the internet (HTTP URIs) and providing links to other resources is of limited utility unless the data is both open and usable – only then does it qualify as LOD (Hausenblas and Boram Kim, 2015). In the best case scenario, the data in question adheres to Linked Open Usable Data (LOUD) principles (Sanderson, 2018), which are all combined in the seven-sphere data model proposed by Thiery (2019).

The model's seven spheres (Figure 3) are defined as follows:

1. OL (Data is available in the Web under an open license)
2. RE (Data is available as structured machine-readable data)
3. OF (Data is available in an open, non-proprietary, structured, and machine-readable data format)
4. RDF (Data is available as URIs and semantically modeled as RDF)
5. LOD (Data is available as a resource, semantically modeled as RDF, and semantically linked to other resources)
6. LOUD (LOD is available as usable data, according to the LOUD principles)
7. LOUD+FAIR (LOUD is available as findable, accessible, interoperable, and reusable data, according to the FAIR principles)

3.1 LOD in the Digital Humanities and Digital Archaeology

In 2011, Leif Isaksen described possible applications for Semantic Web (and Linked Data) technology in the discipline of archaeology (Isaksen, 2011). This was the starting point for Linked Data initiatives in the field of ancient studies in general, including seminal projects like Pelagios Commons[9] (Simon et al., 2016). In the wake of this development, a number of community projects and networks such as Linked Pasts[10] have been established, and increasing numbers of researchers are joining the scholarly LOD community. The work being carried out in this field includes the modeling and reasoning of vagueness and uncertainties in graph data with the help of the Academic Meta Tool (Unold et al., 2019).

In popular culture, the phrase *hic sunt dracones* ('here be dragons') is often used to designate areas on historical maps that are unknown to the authors. It has also been applied to the harvesting and storage of data (Wuttke, 2019): nowadays, researchers can share their research as LOD, enabling the com-

[9] https://pelagios.org

[10] http://linkedpasts.org/assets/lp_whitepaper.pdf

munity to participate in the scholarly discourse in order to create new knowledge; yet unfortunately, much of this data is neither findable nor accessible, and is therefore not FAIR, resulting in modern "unknown data dragons." Often, these 'data dragons' are not interoperable and thus unusable for further scholarly research. These shortcomings can be overcome with the standards and recommendations outlined above: Semantic Web and LOD, the FAIR principles, and LOUD (Thiery et al., 2019). To reduce the number of archaeological 'data dragons,' the Computer Applications and Quantitative Methods in Archaeology (CAA) Special Interest Group (SIG) on Semantics and LOUD in Archaeology (SIG-DataDragon)[11] was founded in 2019.

3.2 Wikidata and the Wikimedia Universe

Established in 2012, Wikidata[12] is a data hub for structured data that was not developed as a primary database for a specific project, but rather to gather existing entries from other databases (Vrandečić and Krötzsch, 2014). It is a free and open knowledge base where anybody can add and edit data, and serves as the central storage for structured data of Wikimedia projects such as Wikipedia and Wikimedia Commons. Data within Wikidata is available under a free licence (CC0, Public Domain), is multilingual, is accessible to both humans and machines (e.g. GUI, API, SPARQL), is exportable using standard formats (e.g. JSON, RDF, SPARQL), and is interlinked to other open datasets in the Linked Data Cloud. The data model used by Wikidata contains items (e.g. label, description, alias, identifier) and statements (e.g. property, value, qualifier, reference), and encourages attribution, links to other entities, and the indication of provenance and sources as references (Voß et al., 2014). It lends itself particularly well to the non-hierarchical representation of divergent interpretations in cultural heritage research, as it is capable of representing uncertainties and ongoing scholarly debates: doubts, ambiguities, and uncertainties can be made explicit by means of qualifiers such as P518 ('applies to part') and Q18122778 ('presumably'), which indicate that a part of a statement remains open to question.

Modeling decision-making processes is difficult in any kind of data base, and Wikidata is no exception. To resolve this and other potential drawbacks of Wikidata, it is possible to create a user-defined ontology based on, for example, the CIDOC CRM CRMsci extension (Doerr et al., 2020) or the PROV ontology (PROV-O).[13]

In Wikidata, versions of the entries are recorded, but while the entry his-

[11] http://datadragon.link

[12] https://wikidata.org

[13] https://www.w3.org/TR/prov-o/

tory may be viewed, links to a specific version of an entry are impossible, because the Semantic Web and URIs are not designed to allow this. As anybody can edit the entries, the original author does not have data sovereignty, and false information may be entered. What is more, information cannot be flagged as debatable – it can only be deleted and discussed in the history view, which, in the absence of structured moderation or editing, may lead to duplicates and inconsistencies. However, despite all of these shortcomings, Wikidata is nonetheless suitable for a number of Open Science applications, as will be discussed in the following sections.

3.3 Our Vision

LOD should be the standard in digital heritage data management, but unfortunately this is not yet the case. Currently, we see two main obstacles to achieving this goal: 1) data extraction requires SPARQL knowledge, and data hosting calls for extensive technical training, which only few researchers have the necessary time and resources to acquire; 2) researchers do not gain any professional recognition for opening up their archives to share and link their data, even though this would greatly enrich the research process in that it enables data to be found faster, reduces errors in transcription and citation, and leads to reproducible research practices. Wikidata allows for the direct creation of LOD as part of the Linked Open Data Cloud.[14] Yet despite a number of existing applications,[15] Wikidata lacks user-friendly tools for researchers and archaeologists, which is where the SPARQL Unicorn comes into play.

4 The SPARQL Unicorn

The SPARQL Unicorn was born at the Computer Applications and Quantitative Methods in Archaeology conference 2019 in Kraków, Poland.[16] Over the course of this conference, it became abundantly clear that data analysis and the creation and maintenance of databases play a central role in archaeological and digital humanities-related research. However, very few databases are freely and openly available and accessible, and even fewer are connected to the Linked Open Data Cloud, which greatly hampers the comparative analysis of records across multiple datasets. One data hub that has recently picked up momentum is Wikidata, a repository based on the SPARQL query language. Yet while supporters of free and open research data have a strong affinity towards community-driven and volunteer-

[14] https://lod-cloud.net

[15] https://hay.toolforge.org/directory/#/search/wikidata

[16] https://2019.caaconference.org

operated data collecting initiatives such as Wikidata, there is – in addition to the various issues outlined above – a shortage of user-friendly, free, and open tools for LOD technologies and repositories, including Wikidata itself. To help mitigate these problems, we developed the SPARQL Unicorn, which we would describe as a friendly tool series for researchers working with Wikidata and other related triple stores. More specifically, the SPARQL Unicorn aims to help researchers from the field of ancient studies use the community-driven data from Wikidata by making it more accessible to those without expertise in LOD or SPARQL (Thiery et al., 2020b).

4.1 The SPARQLing Unicorn QGIS Plugin

For geospatial-related research, community-based LOD repositories, their triple stores, as well as SPARQL endpoints (e.g. Wikidata,[17] LinkedGeoData,[18] and DBpedia[19]) have been established in the Linked Open Data Cloud. Additionally, gazetteer repositories such as GeoNames[20] or Pleiades,[21] as well as administrative providers such as Ordnance Survey UK[22] and OS Ireland,[23] also offer their geodata as LOD. Unfortunately, these resources have gained little traction in the geo community thus far. We believe that the reason for this is a lack of support for LOD-processing GIS applications. As of yet, triple stores, i.e. semantic databases and SPARQL, are not supported by GIS software, GeoServer implementations, or OGC services. Also, the handling of spatial information is quite a challenge for the Linked Data community: although there are vocabularies like GeoSPARQL,[24] which at least allow the modeling of spatial 2D information, even a simple conversion of a Linked Data geo dataset into another reference system can currently only be accomplished with significant scripting effort.

The SPARQLing Unicorn QGIS Plugin[25] (Thiery and Homburg, 2021) addresses the problem of a lack of tools for geospatial data in the Semantic Web. Our experimental plugin can be installed in QGIS[26] from the central QGIS repository[27] (latest release v0.12.2, 2021-06-11).[28] It allows for

[17] https://query.wikidata.org/
[18] http://linkedgeodata.org/sparql
[19] https://dbpedia.org/sparql
[20] http://factforge.net/repositories/ff-news
[21] http://sandbox.mainzed.org/pleiades/sparql
[22] http://data.ordnancesurvey.co.uk/datasets/os-linked-data/apis/sparql
[23] http://sandbox.mainzed.org/osi/sparql
[24] https://www.ogc.org/standards/geosparql
[25] https://github.com/sparqlunicorn/sparqlunicornGoesGIS
[26] https://www.qgis.org/en/site/
[27] https://plugins.qgis.org/plugins/sparqlunicorn
[28] https://plugins.qgis.org/plugins/sparqlunicorn/version/0.12.2/

the graphical user interface based execution of queries in (Geo)SPARQL to selected (archaeology and humanities-related) triple stores and geo-enabled SPARQL endpoints, and prepares the results of the queries as geospatial vector layers in QGIS (Thiery and Homburg, 2020b) for use by the geo community and other interested researchers.

As it stands, the plugin offers three main functions: a) the simplified querying of LOD sources, b) the enrichment of geodata, and c) the transformation of GeoJSON to RDF data.

Function A (Figure 4) permits assisted querying of predefined triple stores, e.g. Wikidata, Nomisma,[29] Linked Ogham Data,[30] and Roman Open Data.[31] It also makes it possible to add another SPARQL endpoint. The SPARQLing Unicorn QGIS Plugin will automatically search for semantically described geospatial data in the widely used vocabularies GeoSPARQL[32] and WGS84 Geo Positioning,[33] e.g. *geosparql:asWKT*, *geosparql:asGeoJSON*, *geo:lat* and *geo:long*.

To assist the user, the plugin provides query examples for the predefined triple stores which can be used as query templates. The plugin will automatically detect geospatial-related (inferred) classes/concepts (Figure 4, right). As shown in Figure 4, the user can select one of the geospatial concepts by clicking on it (e.g. *Stone*), and the plugin will automatically replace the class in the query template. Additionally, the user can restrict the spatial area by a bounding box (Figure 5) or a loaded geospatial layer. This results in a GeoSPARQL function[34] query using *geof:sfIntersects*, as shown in Figure 6 (if the triple store supports GeoSPARQL).

If a user employs the Linked Ogham Data SPARQL endpoint, the *Concept + Label* query template, the geospatial concept *Stone*, or the bounding box shown in Figure 5, this will result in a QGIS vector layer that can be saved in common geo data file formats, e.g. GeoJSON, or a shapefile. By way of example, Figure 7 shows the distribution of Ogham stones on the Dingle Peninsula.

Function B allows the enrichment of geo data using LOD resources from the Linked Open Data Cloud, especially Wikidata (e.g. Ogham site elevation levels). Geospatial data, in particular, can always be seen in an application context, which usually requires additional data from other knowledge domains. Semantically interpreted Linked Data may represent such a resource

[29] http://nomisma.org/sparql

[30] https://digits.mainzed.org/ogham/

[31] https://romanopendata.eu/sparql/epnet.html

[32] http://www.opengis.net/ont/geosparql#

[33] http://www.w3.org/2003/01/geo/wgs84_pos

[34] http://www.opengis.net/def/function/geosparql/

for data enrichment.

Function C converts geospatial information from formats such as GeoJSON into RDF formats like Turtle, which enables the publication of archaeology-related geo data as Linked Data.

5 Ogi Ogham Project

The ᚑᚌᚔ (Ogi) Ogham Project was initially created in 2019 by members of the Research Squirrel Engineers network in conjunction with the SPARQL Unicorn project,[35] and is now supported by the Wikimedia Foundation. The project is a work in progress, whose aims include the digitization and publication of the analog catalog compiled by Macalister (1945, 1949), as well as the linking of existing online resources into Wikidata.[36] The standards applied in the creation of this set of LOD mean that the data involved will become FAIR. The Wikidata repository can be used to support reproducible statistical and geostatistical analysis of the data. We will draw on Macalister's corpus to illustrate this point and to demonstrate the value of modeling analog catalogs as LOD.

5.1 Sources on Ogham Stones and Script

There are several sources, both analog and digital, for information on Ogham stones. As it is our goal to link existing datasets to each other, we will introduce them here in some detail.

5.1.1 Traditional Offline Sources

Our main source of information is the aforementioned Corpus Inscriptionum Insularum Celticarum (CIIC) compiled by Macalister (1945, 1949), in which a large number of Ogham stones in Ireland (CIIC 1-317), Wales (CIIC 317-456), England (CIIC 457-499), the Isle of Man (CIIC 500-505), and Scotland (CIIC 506-520) are recorded. In the first phase of the project, we concentrated on the Irish[37] Ogham stones (CIIC 1-317). The CIIC is the most important reference work for Ogham inscriptions, and its numbering system, CIIC, has also been used to reference stones in other sources.

Macalister describes two different word categories: formula words and nomenclature words. Examples of formula words are *MAQI* ᚋᚐᚊᚔ (son, e.g. CIIC 203, Q67978531) or *MUCOI* ᚋᚒᚉᚑᚔ (tribe/sept, e.g. CIIC 197, Q69388229). Examples of nomenclature words are *CUNA* ᚉᚒᚅᚐ (wolf/hound, e.g. CIIC 154, Q68002826) and *CATTU* ᚉᚐᚈᚈᚒ (battle, e.g.

[35] https://www.wikidata.org/wiki/Q73901970

[36] https://www.wikidata.org/wiki/Q70873595

[37] https://w.wiki/Wmz

CIIC 58, Q70892430). Other names refer to deities, such as the god Lugh (*LUC* ᚂᚒᚉ), who appears in names such as *LUGADDON* ᚂᚒᚌᚐᚇᚇᚑᚅ (CIIC 4, Q70899515).

As most of the stones are damaged, linguistic analysis often yields divergent readings and reconstructions (MacManus, 1997). A graph-based approach makes it possible to model these differing interpretations as equally significant readings of the same text. The encoded relations between inscriptions and the spatial topology of the sites of their discovery also lend themselves to visualization and analysis within a graph data system, for example in the form of historical network research, see e.g. Deicke (2017).

Of course, the archaeological and linguistic research conducted in the intervening years means that the CIIC is not the only relevant source. It is, however, the most complete, and since subsequent scholarship has continued to use Macalister's numbering system for the Ogham stones, the CIIC is an expedient starting point for our data acquisition. Later sources, such as MacManus (1997) and Forsyth (1996), can be added as a second step.

We used the spatial information given in Macalister's catalog – county, barony, and townland – as well as the inscriptions themselves to identify the stones in other resources. A townland (Irish: *baile fearainn*),[38] the most precise category describing the location of the stones, is a spatial unit of Gaelic origin, though it has been modified by later re-structuring processes such as the Ordnance Survey. This information was used for linking the data to OpenStreetMap with the help of Townlands,[39] and in some cases Logainm.[40] Most Ogham stones recovered so far have been re-used over the centuries and can now be found as lintels in huts, in the walls of churches etc. (MacManus, 1997). As a spatial category, the townland thus provides a sufficient level of detail, while avoiding the pitfall of implying a higher level of precision than is actually achievable.

5.1.2 Online Databases

Several online databases for Ogham-related research exist. The Thesaurus Indogermanischer Text- und Sprachmaterialien (TITUS) database[41] hosted by Jost Gippert dates from 1996. It is online, but it is neither open nor interoperable, and cannot be interlinked with other datasets, which prevented us from using it in our case study.

Another rich source of information on Ogham stones is the database of

[38] https://data-osi.opendata.arcgis.com/datasets/559dc2f725004823b51923e1d6ba5889_0/data

[39] https://www.townlands.ie/

[40] https://www.logainm.ie/en/

[41] http://titus.uni-frankfurt.de/ogam/frame.htm

the Celtic Inscribed Stones Project (CISP)[42] assembled by University College London, which gathered numerous inscriptions (not only in Ogham) and published them online in 2001. This database has no API, and one can only click through the various entries. It is online, but neither linked nor open, which is wholly understandable given the year of its publication. We would like to thank Dr. Kris Lockyear, who generously made the project's Access database available to us and gave us permission to use the data.

Some more recent open research projects offer online and open data, including the Ogham in 3D Project,[43] which creates 3D scans of known Ogham stones in Ireland (as of January 2021, there were c. 160 scanned stones[44]). An online platform[45] offers 3D data (3D PDF and OBJ files), as well as EPIDOC[46] files with annotated stone inscriptions for download under a Creative Commons License. An example is stone CIIC 180, Emlagh East (Imleach Dhún Séann), Co. Kerry,[47] which is shown in Figure 8.

Another important resource is the Historic Environment Viewer,[48] which contains a WebGIS that provides access to the records of the National Monuments Service's Sites and Monuments Record (SMR) and the National Inventory of Architectural Heritage (NIAH). While this repository is very helpful, there are no links to other datasets.

It is therefore safe to conclude that there is a lack of Ogham research data in the Linked Open Data Cloud. This results in various challenges, and makes comparative analysis of records across multiple datasets difficult. However, some useful information can be found in the Wikimedia universe with its variety of tools and projects under the umbrella of the Wikimedia Foundation, as will become clear in the following section.

5.1.3 Ogham in the Wikimedia Universe

The Wikimedia project universe[49] provides researchers with a large number of free and open access data hubs. This includes the free encyclopedia Wikipedia, the free media repository Wikimedia Commons, the free knowledge base Wikidata, as well as the bibliographic extension to Wikidata, Scholia. Wikipedia, for example, contains articles on the Ogham alphabet[50] and

[42] https://www.ucl.ac.uk/archaeology/cisp/database/

[43] https://ogham.celt.dias.ie/menu.php?lang=en

[44] https://ogham.celt.dias.ie/menu.php?lang=en&menuitem=81&overviewinfo=epidoc_stone_details

[45] https://ogham.celt.dias.ie/menu.php?lang=en&menuitem=81&overviewinfo=file

[46] https://epidoc.stoa.org/gl/latest/

[47] https://t1p.de/o3d180

[48] https://webgis.archaeology.ie/historicenvironment/

[49] https://meta.wikimedia.org/wiki/Wikimedia_projects

[50] https://en.wikipedia.org/wiki/Ogham

Ogham inscriptions.[51] Some bibliographic information concerning Ogham has already been assembled in Scholia, but the collection requires further attention[52] to expand the small number of existing items related to the topic.[53]

Wikidata Commons offers a set of free Ogham images,[54] including a chart of the Ogham alphabet,[55] Macalister's drawing of CIIC 81 at Garranes, Barony of Kinalmeaky (Figure 9), and the photograph of CIIC 81 in University College Cork's Stone Corridor (Figure 1) with the inscription *C[A]SSITT[A]S MAQI MUCOI CALLITI* (CASSITTAS son of the tribe CALLITI, the hard ones).

5.2 Workflow

The first step of our Ogi Ogham Project workflow (Figure 10) is the extraction of spatial information on the Ogham stones and their inscriptions from Macalister's offline catalog. This spatial information is then linked to Open Street Map data, other Geo LOD resources, as well as further Ogham data, which allows the mapping of databases according to location and information on the inscription (e.g. the CISP database). As a next step, this data can be imported into Wikidata, making it possible to extract the enriched information with SPARQL Unicorn and to analyze it further with software such as QGIS (using the SPARQLing Unicorn Plugin) or R (see Section 5.3) in a process that can be applied to data from other sources as well. In the following sections, we will first discuss Wikidata modeling and the import process for spatial information, before presenting two concrete examples of how our workflow can be applied.

5.2.1 Wikidata Import

The Association of Research Libraries (ARL, 2019) recommends using OpenRefine[56] for inserting, publishing, and maintaining information stored in Wikidata. Our basis for editing in OpenRefine are CSV files, which contain all information along with links to other Wikidata entities. We manually compiled these CSV files from sources such as Macalister, Logainm, and Open Street Map. Our first import[57] contained CSVs for

[51] https://en.wikipedia.org/wiki/Ogham_inscription
[52] https://ogham.celt.dias.ie/menu.php?lang=en&menuitem=80
[53] https://scholia.toolforge.org/topic/Q184661
[54] https://commons.wikimedia.org/wiki/Category:Ogham
[55] https://commons.wikimedia.org/wiki/Ogham_alphabet
[56] https://openrefine.org/
[57] https://github.com/ogi-ogham/ogham-wikidata

‘Ogham sites/townlands,’[58] ‘Ogham collections,‘[59] and ‘Ogham stones.’[60] OpenRefine offers the possibility of creating individual mapping schemes for Wikidata, which we applied to, among other things, Ogham sites[61] and Ogham stones[62]. The results of this mapping process can be transformed and exported by OpenRefine to QuickStatements.[63] The Quick Statements 2 Tool[64] is capable of creating new Wikidata entities and updating existing ones automatically based on the QuickStatements syntax.[65] In light of our own experience with this workflow, we can highly recommend using OpenRefine and QuickStatements as Wikidata import tools.[66]

5.2.2 Wikidata Modeling: Words

Wikidata contains a number of Ogham-related entities[67] which describe Ogham as an early medieval writing system (Q184661), Ogham stones as stones with Ogham inscriptions (Q2016147), Ogham sites as sites where Ogham stones have been found (Q72617071), and Ogham letters as characters in the Ogham alphabet (Q41812345).

As outlined above, a variety of formula and nomenclature words can be found in Ogham inscriptions. A list of these words according to MacManus (1997) accompanied by a translation and a list of references and variants has been published by the Ogi Ogham Project members on GitHub.[68] While some of the listed words exist as a Wikidata entity, additional entries will be published within the Ogi Ogham Project.

5.2.3 Wikidata Modeling: Sites

As we have already pointed out, Macalister's CIIC established the Ogham stones' geospatial placement down to the level of individual townlands. Initial comparisons were made with Townlands (based on OSM) and Logainm, but several problems arose during this process: some locations were unknown to Macalister; he made mistakes (typographical errors, wrong place names), or gave imprecise information (occasionally, townlands in one barony have the same name). In some cases, a shift in the structure of baron-

[58] https://github.com/ogi-ogham/ogham-wikidata/tree/master/Townlands

[59] https://github.com/ogi-ogham/ogham-wikidata/tree/master/OgamCollections

[60] https://github.com/ogi-ogham/ogham-wikidata/tree/master/OgamStones

[61] https://t1p.de/gh-ogi-1

[62] https://t1p.de/gh-ogi-2

[63] https://www.wikidata.org/wiki/Help:QuickStatements

[64] https://quickstatements.toolforge.org/#/

[65] https://www.wikidata.org/wiki/Help:QuickStatements#Command_sequence_syntax

[66] https://github.com/ogi-ogham/ogham-wikidata/blob/master/HowTo/OpenRefine.md

[67] for a short list, see https://github.com/ogi-ogham/ogham-wikidata

[68] https://github.com/ogi-ogham/oghamextractor/blob/master/words/words.csv

ies, electoral divisions, and townlands had occurred between 1945 and 2020. This created a situation where some of the townlands referred to by Macalister could no longer be identified by us.

We used the database of archaeological finds uploaded by the Department of Culture, Heritage and the Gaeltacht (the above-mentioned Historic Environment Viewer), to check our information. Although the locations of some items in this database are listed as unknown, attempts at clarification of their precise whereabouts were undertaken by local experts, to whose findings we adhered (this applies to a total of 16 cases; concerning CIIC 204, for example, we followed the experts' evaluation that Macalister had erroneously given its location as Curraghmore West instead of East).

The website Logainm also proved helpful, as it linked a monument about which Macalister's catalogue yielded only imprecise information to a specific townland. In instances where Macalister provided further information on the stone's location by describing the area or certain landmarks, we were able to use these details to improve the precision of the spatial data. For example, Macalister's comment that the stone CIIC 48 was found south of the village allowed us to tentatively identify the townland in question. In the rare cases in which two townlands with the same name were located right next to each other, we opted for the larger of the two, crosschecking this educated guess with the Historic Environment Viewer.

In this way, 185 out of the 196 townlands mentioned by Macalister could be identified and the corresponding information entered into Wikidata: the townland's anglicized name; its Gaelic alias; the province, county, barony, civil parish, and electoral division it belongs to; and a point coordinate indicating the middle of the townland polygon. The dataset was linked by providing the OSM, Logainm, and OSi GeoHive IDs, as well as the link to the Townlands IDs from which we derived most of our data. In order to map the eleven Ogham stones whose location remained unknown, we resorted to the center of the barony as indicated by Macalister (Thiery et al., 2020a).

As stated above, decisions such as these are currently difficult to model using Wikidata. A bespoke ontology with self-defined semantic modeling rules could be a step towards more transparency in decision modeling.[69]

Table 1 shows the example of the Wikidata model for the Ogham sites in Garranes Townland, Co. Cork. Each site is described with a label, a description, an alias (Irish name), the country, the province, and other administrative divisions, a coordinate location, as well as external identifiers linked to OpenStreetMap and the Logainm website. Wikidata also offers the possibility of setting statements for reference information, thereby linking inform-

[69] more information at `https://doi.org/10.5281/zenodo.4091743`

ation to its source. This was done for several properties of an Ogham site including administrative divisions, location, and external identifiers (Table 2), so as to provide a reference for each piece of controversial information with regards to the stones' location.

5.2.4 Wikidata Modeling: Stones

The entries on Ogham stones in Wikidata follow a basic modeling structure: P31 (instance of) Ogham stone (Q2016147), P361 (part of) Ogham stones (Q67978809)/Ogi-Ogham Project (Q70873595), P17 (country) Ireland (Q27), P625 (coordinate location) derived from the Ogham site, P189 (location of discovery) county and Ogham site, P1684 (inscription) and P1545 (series ordinal). Some of the stones contain additional information as shown in Table 3 for CIIC 81.[70] Our goal is to model all known Ogham stones. We would like to add details on P186 (material used), P195 (collection), P2043/P2049/P2048 (length, width, height), P6568 (inscription mentions), P5816 (state of conservation) and P18 (image). The sources of this in-depth information are shown in Table 4. As discussed above, references such as these are especially relevant in cases of conflicting information regarding the texts and their interpretation.

5.3 Analysing Ogham

The information on the Ogham stone sites collected in Wikidata enables a multitude of analytical approaches. By way of example, we will concentrate on spatial analysis and the exploration of word counts.

To gather and import Wikidata data into a format that is usable in the statistical programming language R, a package called WikidataQueryServiceR (Popov, 2020) can be used. It is an API client for the Wikidata Query Service[71] and implements the function *query_wikidata*. *query_wikidata* transforms the data retrieved from Wikidata by means of an input SPARQL query into a dataframe, which is one of the most widely used classes for tabular data in R. The scripts for the analytical approaches described in the following sections can be found on GitHub.[72]

In our investigation of the spatial relations between the Ogham sites and word frequencies, several R packages were utilized.[73] In addition, we also

[70] https://www.wikidata.org/wiki/Q69385424

[71] https://query.wikidata.org/

[72] https://github.com/sparqlunicorn/oghamaps

[73] R Core (R Core Team, 2020); for data cleaning and visualization: tidyr, reshape, ggplot2, viridis (Wickham and Henry, 2020; Wickham, 2007, 2016; Garnier, 2018); for spatial analysis: ggspatial, rgdal, sp (Dunnington, 2018; Bivand et al., 2019, 2013).

employed QGIS and the SPARQLing Unicorn QGIS Plugin to conduct various types of spatial analysis on the modeled Ogham stones.

5.3.1 Spatial Analysis

It is evident from the density and distribution analysis carried out by means of the R project oghamaps[74] (Schmidt, 2021) and QGIS (Figures 11 and 12) that the majority of Ogham sites are located in the counties of Kerry, Cork, and Waterford (Table 5), and the Dingle Peninsula (western hotspot in Figure 12). The south of Ireland is the main distribution area with decreasing density of sites from west to east.

To further examine the various areas, percolation analysis was applied; an explorative clustering algorithm (Maddison, 2020) which lends itself well to the identification of clusters and the establishment of densities and distances between archaeological sites (Maddison and Schmidt, 2020). Euclidean distance is used to create clusters from points that are within a certain radius from each other. This radius is then increased in an iterative way, and maps as well as summaries of the development of certain parameters with regard to this radius are generated.[75]

In our example, the change in the mean cluster size is shown in relation to the clustering radius (Figure 14). At the points where the curve flattens, only a small number of sites appear to be integrated into the existing clusters. As a result, we can see how dense the points inside small clusters are, and how far away the next points lie. Many Ogham sites are about 10 km away from each other, with the next steps in the distribution occurring at a radius of 22, 34, and 44 km, after which almost all stones are incorporated into a single cluster. At 22 km, all Ogham stones in the south of Ireland belong to a single cluster, whereas several distinct clusters still exist in the north – a clear indicator of the lower density of Ogham stones in this area.

5.3.2 Analysis of Word Frequency

The foregoing discussion begs the question: which words are most frequently found on the 317 Ogham stones in Ireland? And do some of these words regularly appear in combination with others?

To answer these questions, we used the Ogham Extractor Tool, available on Github,[76] and the table ciictowords2 (Thiery and Homburg, 2020a), in which relevant words and the various readings of the Ogham inscriptions on

[74] source code at `https://github.com/ogi-ogham/oghamaps`

[75] see DOI 10.17605/OSF.IO/7EXTC for the R package, which has this algorithm implemented.

[76] `https://github.com/ogi-ogham/oghamextractor/`

the CIIC stones are recorded. As our focus was on their meaning and not on variances in spelling, we resorted to the most common forms of the words in question (for example, *MAQI* instead of *MAQ*).

The resulting word list was used to a) show how often the most common words occur (Figure 13) and b) how often words co-occur (Figure 15). As discussed above, *MAQI* (son) and *MUCOI* (tribe) are most common, followed by *LUG* (god), *AVI* (grandson, descendant), and *ANM* (name). This, combined with the preponderance of names in the inscriptions, underscores the importance of familial affiliation in the communities that used the Ogham alphabet. The religious dimension also appears to have played a significant role (MacManus, 1997).

The second word analysis looks at the co-occurrence of words on the Ogham stones (Figure 15). In this case, we wished to determine whether or not specific names or topics always appear in combination.

Our analysis shows a range of values between 0 and 21. The diagonal must be ignored, as it reflects how often a word occurs all in all (co-occurrence of words with themselves). The most frequent co-occurrence is of the words *MAQI* and *MUCOI* (21 times) – once again, son and tribe. As *MAQI* is the most common word overall, it also has the most co-occurrences with other words. Given that the dataset yielded by the Ogham inscriptions is rather small and consists of a limited number of words, the most common co-occurrence frequency is 0, as some words do not co-occur with others at all (e.g. *CARI*, loves).[77] By contrast, two words that are seldom used – *VIR* (man, 7x) and *CATTU* (battle, 3x) – do co-occur, even though the compound word *CATTUVIR* (man of battle) also exists. This particular case of co-occurrence gives us another brief glimpse at the workings of the society behind these monuments and the importance it attached to warfare.

6 Conclusion and Outlook

In this paper, we discussed LOD and its role as the emerging standard for research data management and semantic data modeling in the digital humanities, digital archaeology, and archaeoinformatics, as well as its practical implementation in Wikidata. Acknowledging that the SPARQL query language is not yet widely known, we introduced the SPARQL Unicorn tool framework developed by the community of Research Squirrel Engineers with the aim of helping researchers use SPARQL. One of the cornerstones of this framework is the SPARQLing Unicorn QGIS Plugin, which is already published and enables researchers to query Wikidata for geospatial and other LOD within the

[77] It should be noted, however, that the dataset has been reduced to the words that occur at least three times in total.

QGIS environment.

A short overview of the Ogi Ogham Project demonstrated the practical application of our digitization and linking workflow for the creation of LOD, and showcased how OpenRefine and QuickStatements can be employed to import that data into Wikidata. Two analytical approaches that have already been implemented illustrated the potential of our toolkit to foster open and reproducible research.

In the future, we will focus on the steps necessary for modeling our own bespoke Ogham ontology and hosting it in a triple store, which will increase our flexibility with regard to semantic modeling, and will allow us to use the Wikidata entries as links to the Linked Open Data Cloud. This planned development work is part of the project Irische ᚑᚌᚆᚐᚋSteine im Wikimedia Universum and is funded by the Wikimedia Deutschland Fellow-Programm Freies Wissen.[78]

As has become clear, the Ogham data is perfectly suitable for text analysis (e.g. text as a graph using a graph database like Neo4J[79]), and will lend itself equally well to network analysis of the relationships between persons and tribes – all that is needed is the further enrichment of Ogham data through the addition of sources and information (including provenance information) by the community.[80]

Acknowledgements

Timo Homburg and Martina Trognitz, fellow members of the Research Squirrel Engineers network, have been of great support, both in the development of our project and its presentation in this paper. We would also like to thank Dr. Kris Lockyear for his encouragement and for kindly making the CISP database accessible to us, and Dr. David Wigg-Wolf for language corrections.

[78]http://ogham.squirrel.link

[79]https://neo4j.com

[80]The Research Squirrel Engineers welcome new squirrels interested in feeding the Linked Data Cloud with Ogham nuts!

Figures and Tables

Figure 1: UCC Stone Corridor, Stone 4, CIIC 81 [Florian Thiery, CC BY 4.0 via Wikimedia Commons]

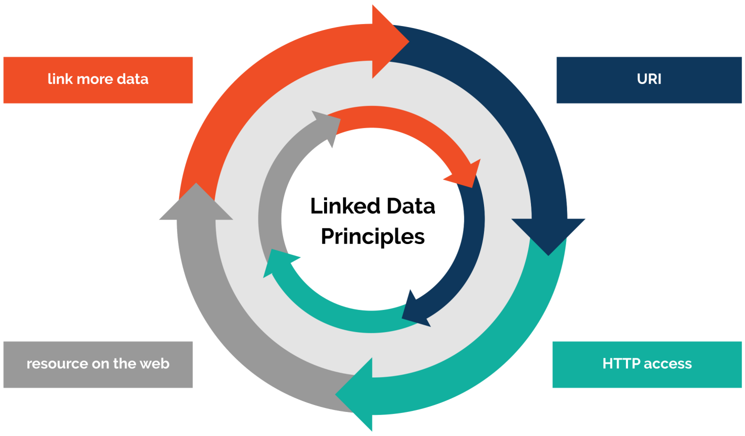

Figure 2: Linked Data principles [Florian Thiery, CC BY 4.0 via Wikimedia Commons]

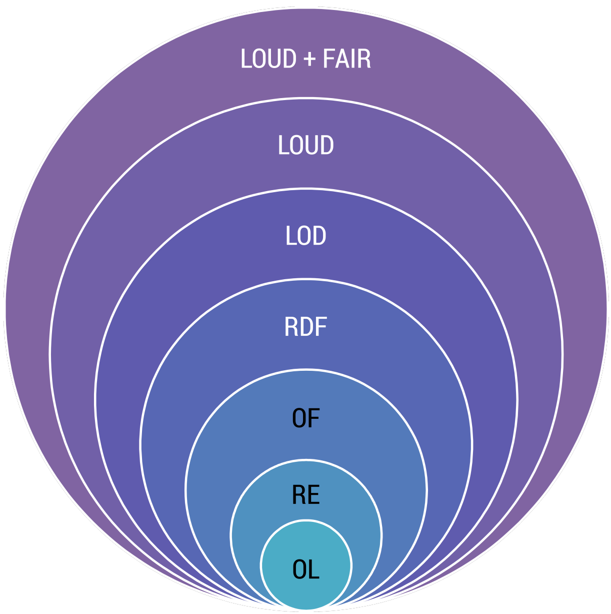

Figure 3: 7-sphere data model [Florian Thiery, CC BY 4.0 via Wikimedia Commons]

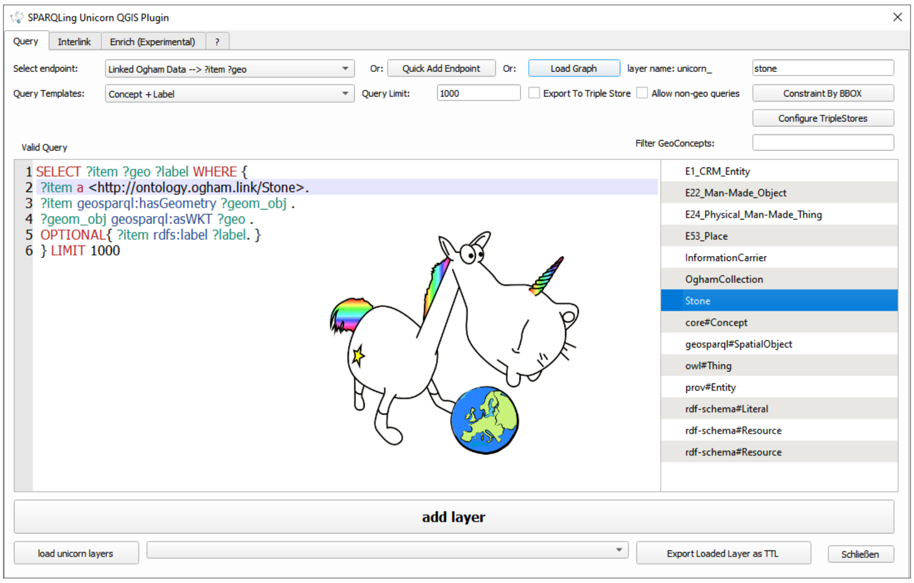

Figure 4: Query tab to search Ogham stones using the SPARQLing Unicorn QGIS Plugin

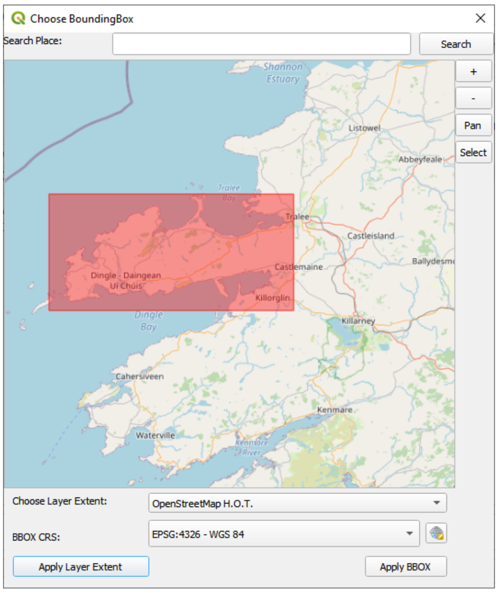

Figure 5: Bounding box pop-up in the SPARQLing Unicorn QGIS Plugin

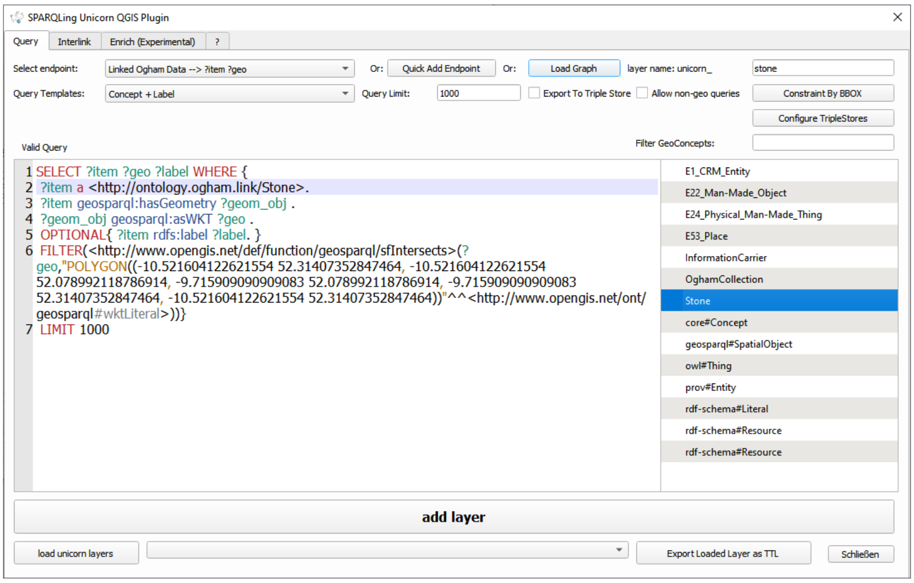

Figure 6: Query tab to search Ogham stones with the bounding box in Figure 5 using the SPARQLing Unicorn QGIS Plugin

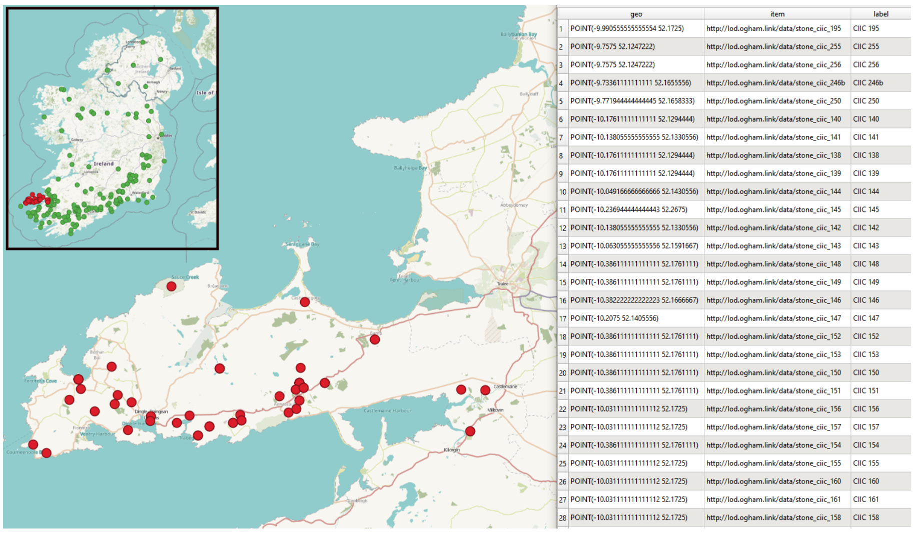

Figure 7: top-left: Irish Ogham stones overview, red: stones filtered by the bounding box; bottom-left: Ogham stones on the Dingle Peninsula filtered by the bounding box represented in QGIS (over an OpenStreetMap H.O.T. map) queried by the SPARQLing Unicorn QGIS Plugin; right: excerpt of the resulting attribute table

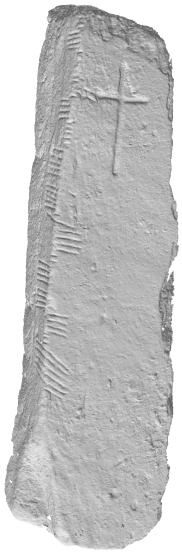

Figure 8: CIIC 180, Emlagh East (Imleach Dhún Séann), Co. Kerry based on a 3D scan by the Ogham in 3D Project [Florian Thiery, CC BY 4.0 via Wikimedia Commons]

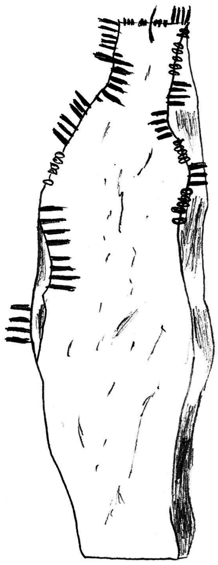

Figure 9: Example of an Ogham-related Wikimedia Commons entry: drawing of CIIC 81 by Macalister (1945) [Florian Thiery, CC BY 4.0 via Wikimedia Commons]

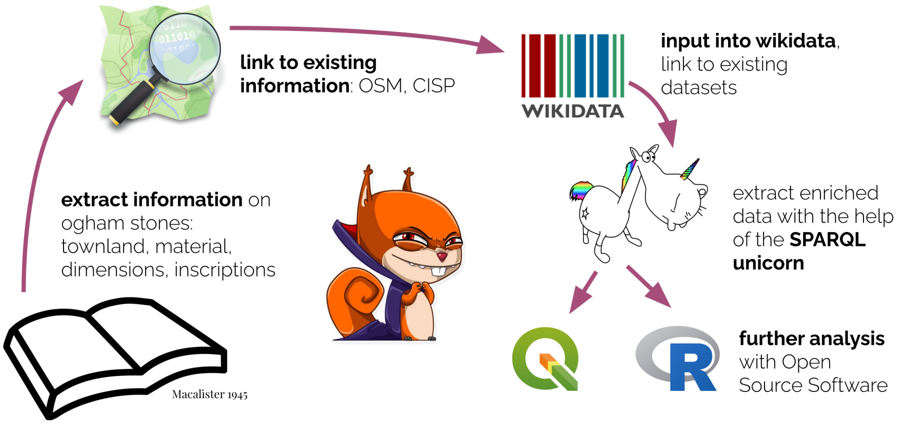

Figure 10: Workflow in the Ogi Ogham Project

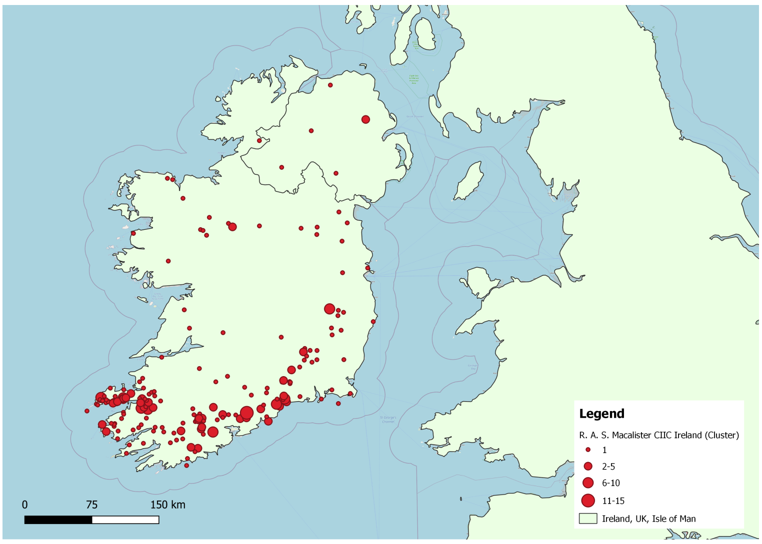

Figure 11: CIIC Ogham Ireland cluster created with QGIS [Florian Thiery, CC BY 4.0]

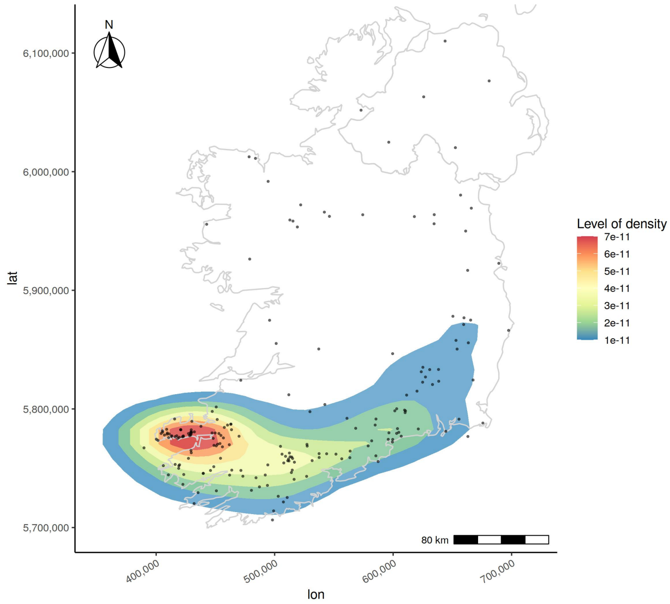

Figure 12: CIIC Ogham Ireland density map created with R (oghamaps)

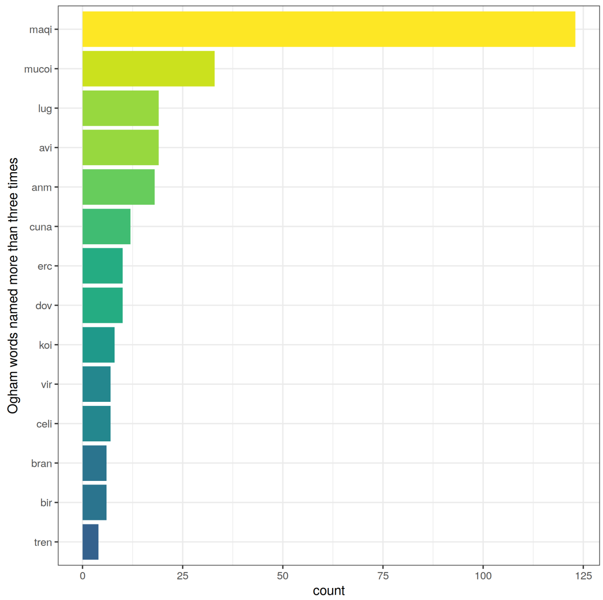

Figure 13: Count of CIIC Ogham words occurring more than three times created with R (oghamaps) [Code in the oghamaps - project on github, Sophie C. Schmidt, CC BY 4.0]

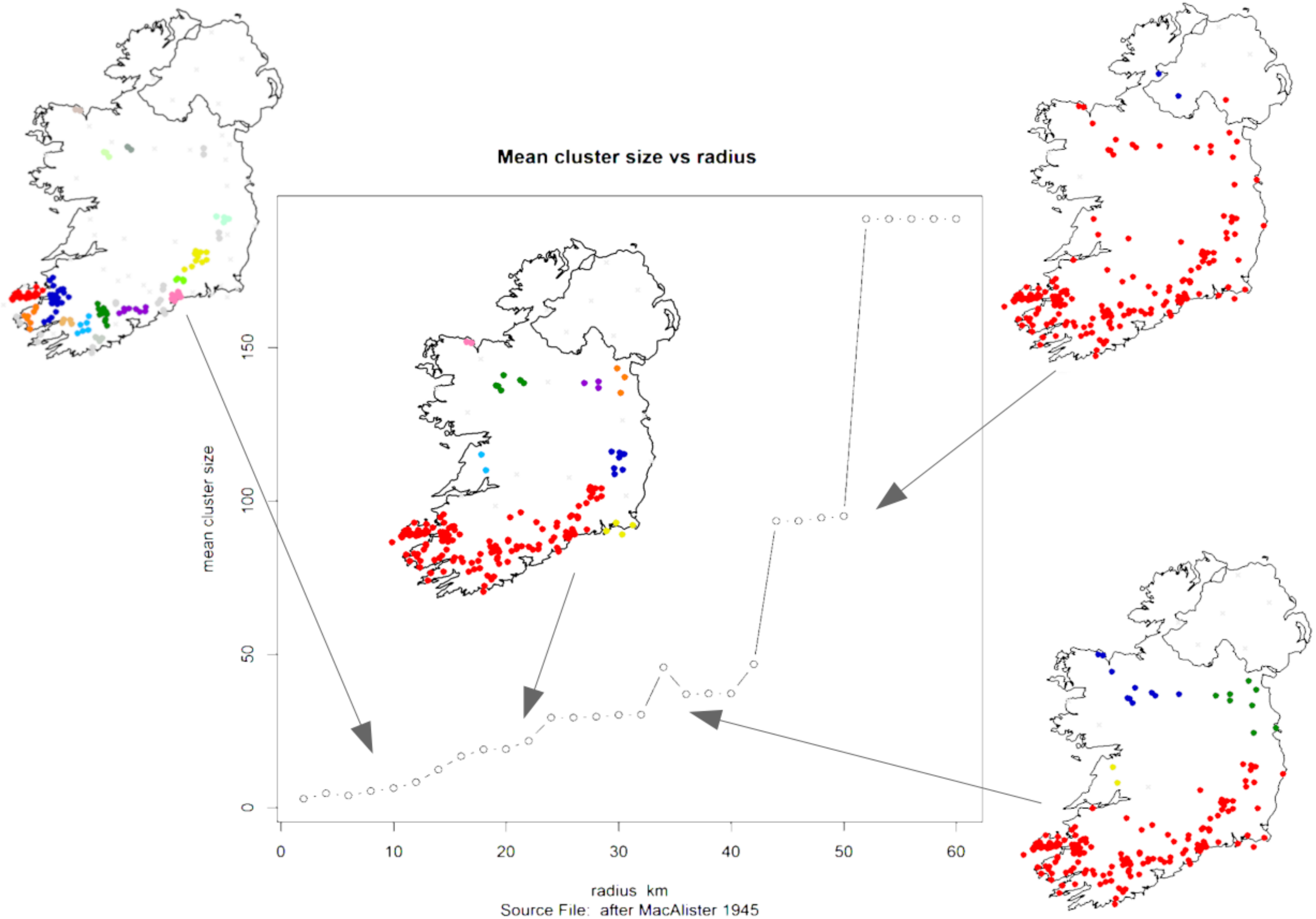

Figure 14: Percolation analysis. Graph showing the mean cluster size at differing radii. The maps represent particularly significant points in the distribution [Sophie Schmidt, CC BY 4.0]

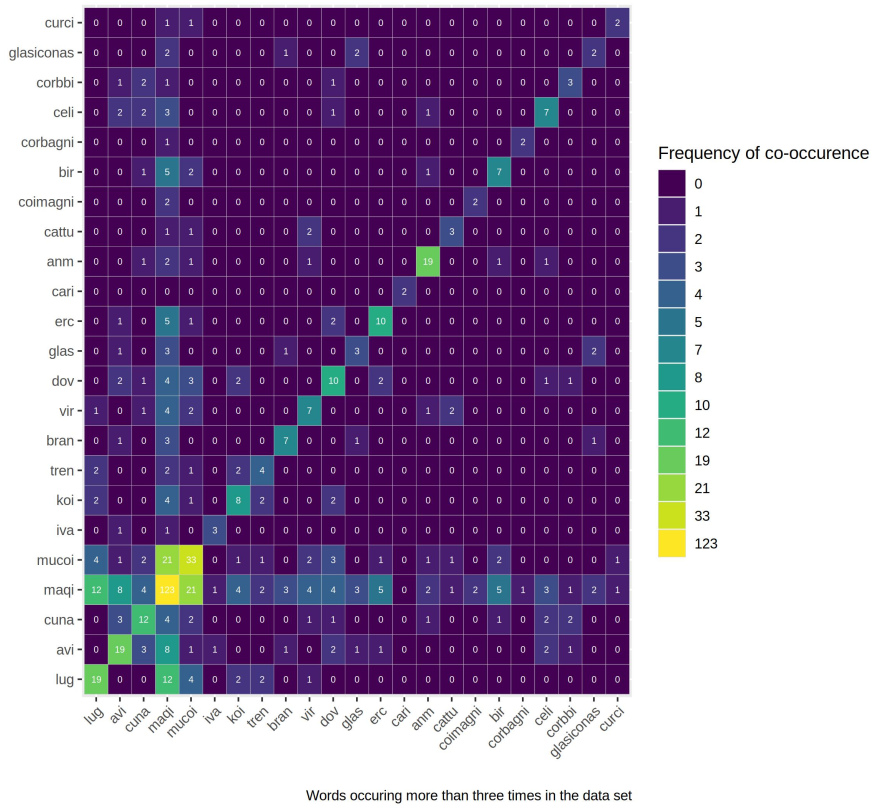

Figure 15: CIIC Ogham Ireland co-occurrence statistics created with R (oghamaps) for words that occur more than three times [Code in the oghamaps - project on github, Sophie C. Schmidt, CC BY 4.0]

property	**description**	**example**
label	desc	Garranes Townland
description	description	Garranes Townland, Co. Cork
alias	add. information	An Garrán
P31 (instance of)	townland geographic location county town	townland (Q2151232)
P361 (part of)	project	Ogi-Ogam Project (Q70873595)
P17 (country)	Ireland Northern Ireland	Ireland (Q27)
P131 territorial entity	province county barony civilparish electoraldevision	Province Munster (Q131438) County Cork (Q162475) Barony Kinalmeaky (Q20616069) Templemartin (Q60557037) Templemartin (Q59724324)
P625 (coordinate location)	lat lon	51.816389 -8.761667
P402 (OSM)	OpenStreetMapID	6168494
P5097 (Logainm)	LogainmID	8299

Table 1: Wikidata townland modeling example (Q69385525)

Qualifier	**description**	**example**
Ireland	P248 (stated in)	Q70256237
County Cork	P248 (stated in)	Q85384744
Kinalmeaky	P248 (stated in)	Q85384744
Templemartin civil parish	P248 (stated in)	Q85384744
Templemartin electoral district	P248 (stated in)	Q85384744
Munster	P248 (stated in)	Q85384744
51.816389;-8.761667	P248 (stated in)	Q85384744
LogainmID	P248 (stated in)	Q85384744
OpenStreetMapID	P248 (stated in)	Q85384744

Table 2: Wikidata townland modeling example (Q69385525 references)

property	description	example
label	desc	CIIC 81
description	description	Ogam Stone CIIC 81
alias	add. information	Ogam Stone 4 (Stone Corridor at UCC)
P31 (instance of)	ogham stone	Q2016147
P361 (part of) P361 (part of)	Ogi Ogam Project Ogam Stones	Q70873595 Q67978809
P18 (image)	photo or drawing	figures 1 and 9
P17 (country)	modern country	Ireland (Q27)
P625 (coordinate location) P2868 (subject has role)	lat lon qualifier	51.816389 -8.761667 archaeological site (Q839954)
P186 (material used)	stone material	stone (Q22731)
P195 (collection)	collection of ogham stones	Stone Corridor University College Cork (Q69379477)
P189 (location of discovery)	findspot	County Cork (Q162475) Garranes Townland (Q69385525)
P2043 (length)	stone length	17.78 cm
P2049 (width)	stone width	48.26 cm
P2048 (height)	stone height	175.26 cm
P1684 (inscription)	ogham inscription	C[A]SSITT[A]S MAQI MUCOI CALLITI
P518 (applies to part)	qualifier	presumably (Q18122778)
P1545 (series ordinal)	number	81,4
P6568 (insc. mentions)	ogham words	MAQI (Q67381254) MUCOI (Q67999759)
P5816 (state of conservation)	stone status	preserved (Q56557591)

Table 3: Wikidata Ogham Stone modeling example (Q69385424)

qualifier	description	example
ogham stone	P248 (stated in)	Q70256237
Ogham Stones	P248 (stated in)	Q70873595
Ireland	P248 (stated in)	Q70256237
51.816389;-8.761667	P248 (stated in)	Q70256237
stone	P248 (stated in)	Q70256237
Stone Corridor UCC	P248 (stated in)	Q70383873
County Cork	P248 (stated in)	Q70256237
Garranes Townland	P248 (stated in)	Q70256237
length/ width/ height	P248 (stated in)	Q70256237
inscription	P248 (stated in)	Q70256237
series ordinal (81)	P248 (stated in)	Q70256237
series ordinal (4)	P248 (stated in)	Q70256237
MAQI MUCOI	P248 (stated in) P248 (stated in)	Q70256237 Q70256237
preserved	P248 (stated in)	Q70256237

Table 4: Wikidata Ogham Stone modeling example (references of Ogam Stone CIIC 81 = Q69385424)

County	Number of Ogham stones
Kerry	127
Cork	81
Waterford	47
Kilkenny	12
Kildare	8
Mayo	8

Table 5: The counties with the largest number of Ogham stones in Ireland

References

ARL (2019). ARL White Paper on Wikidata Opportunities and Recommendations. https://scholarworks.iupui.edu/bitstream/handle/1805/18902/ARL-white-paper-on-Wikidata.pdf.

Berners-Lee, T. (2006). Linked Data - Design Issues. https://www.w3.org/DesignIssues/LinkedData.html.

Berners-Lee, T. (2007). Giant Global Graph. https://web.archive.org/web/20160713021037/http://dig.csail.mit.edu/breadcrumbs/node/215.

Berners-Lee, T., Hendler, J., and Lassila, O. (2001). The Semantic Web. *Scientific American*, 284(5):34–43, DOI: 10.1038/scientificamerican0501-34.

Bivand, R., Keitt, T., and Rowlingson, B. (2019). *rgdal: Bindings for the 'Geospatial' Data Abstraction Library*, https://CRAN.R-project.org/package=rgdal. R package version 1.4-8.

Bivand, R. S., Pebesma, E., and Gomez-Rubio, V. (2013). *Applied spatial data analysis with R, Second edition*. Springer, New York, NY, https://asdar-book.org/.

Deicke, A. (2017). Networks of Conflict: Analyzing the "Culture of Controversy" of Polemical Pamphlets of Intra-Protestant Disputes (1548-1580). *Journal of Historical Network Research*, page 71–105, DOI: 10.25517/jhnr.v1i1.8.

Doerr, M., Kritsotaki, A., Rousakis, Y., Hiebel, G., et al. (2020). Definition of the CRMsci - An Extension of CIDOC-CRM to support scientific observation. http://www.cidoc-crm.org/crmsci/sites/default/files/CRMsci%20v.1.2.8.pdf.

Dunnington, D. (2018). *ggspatial: Spatial Data Framework for ggplot2*, https://CRAN.R-project.org/package=ggspatial. R package version 1.0.3.

Ferguson, S. (1864). Account of Ogham Inscriptionsin the Cave at Rathcroghan, County of Roscommon. In *Proceedings of the Royal Irish Academy*, volume 9, page 160–170.

Forsyth, K. S. (1996). *The Ogham Inscriptions of Scotland: An Edited Corpus*. PhD thesis, Harvard University.

Garnier, S. (2018). *viridis: Default Color Maps from 'matplotlib'*, https://CRAN.R-project.org/package=viridis. R package version 0.5.1.

Graves, C. and Limerick, C. (1876). The ogham alphabet. *Hermathena*, 2(4):443–472.

Hausenblas, M. and Boram Kim, J. G. (2015). 5-star Open Data. http://5stardata.info/en/.

Isaksen, L. (2011). *Archaeology and the Semantic Web*. Thesis (Doctoral), University of Southampton, School of Electronics and Computer Science, https://eprints.soton.ac.uk/id/eprint/206421.

Macalister, R. A. S. (1945). *Corpus inscriptionum insularum Celticarum*. Number Vol. I in Corpus inscriptionum insularum Celticarum. Stationery Office, Dublin.

Macalister, R. A. S. (1949). *Corpus inscriptionum insularum Celticarum*. Number Vol II. in Corpus inscriptionum insularum Celticarum. Stationery Office, Dublin.

MacManus, D. (1997). *A Guide to Ogam*. Number 4 in Maynooth monographs. An Sagart, Maynooth, Ireland.

Maddison, M. S. (2020). Percolation analysis. In Gillings, M., Hacıgüzeller, P., and Lock, G., editors, *Archaeological Spatial Analysis: A Methodological Guide*, page 77–92. Routledge, DOI: 10.4324/9781351243858.

Maddison, M. S. and Schmidt, S. C. (2020). Percolation Analysis – Archaeological Applications at Widely Different Spatial Scales. *Journal of Computer Applications in Archaeology*, 3(1):269–287, DOI: 10.5334/jcaa.54.

Popov, M. (2020). WikidataQueryServiceR: API Client Library for 'Wikidata Query Service'. https://CRAN.R-project.org/package=WikidataQueryServiceR.

R Core Team (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, https://www.R-project.org/.

Sanderson, R. (2018). LOUD: Linked Open Usable Data. https://linked.art/loud/.

Schmidt, S. C. (2021). oghamaps. DOI: 10.5281/zenodo.3969519.

Simon, R., Isaksen, L., Barker, E., and de Soto Cañamares, P. (2016). The Pleiades Gazetteer and the Pelagios Project. In Berman, M. L., Mostern, R., and Southall, H., editors, *Placing names: enriching and integrating gazetteers*, The spatial humanities, pages 97--109. Indiana University Press, Bloomington, IN.

Thiery, F. (2019). Sphere 7 Data: LOUD and FAIR Data for the Research Community. DOI: 10.5281/zenodo.2643469.

Thiery, F. and Homburg, T. (2020a). oghamextractor. DOI: 10.5281/zenodo.3969514.

Thiery, F. and Homburg, T. (2020b). QGIS - A SPARQLing Unicorn? Eine Einführung in Linked Open Geodata zur Integration von RDF in QGIS Plugins. DOI: 10.5281/ZENODO.3719127.

Thiery, F. and Homburg, T. (2021). SPARQLing Unicorn QGIS Plugin. DOI: 10.5281/zenodo.3786814.

Thiery, F., Homburg, T., Schmidt, S. C., Trognitz, M., et al. (2020a). SPARQLing Geodesy for Cultural Heritage – New Opportunities for Publishing and Analysing Volunteered Linked (Geo-) Data. DOI: 10.5281/ZENODO.3751769.

Thiery, F., Schmidt, S. C., Homburg, T., and Trognitz, M. (2020b). The SPARQL Unicorn: An introduction. DOI: 10.5281/ZENODO.3742185.

Thiery, F., Trognitz, M., Gruber, E., and Wigg-Wolf, D. (2019). Hic Sunt Dracones! The Modern Unknown Data Dragons. DOI: 10.5281/zenodo.3345711.

Unold, M., Thiery, F., and Mees, A. (2019). Academic Meta Tool. Ein Web-Tool zur Modellierung von Vagheit. In Kuczera, A., Wübbena, T., and Kollatz, T., editors, *Die Modellierung des Zweifels – Schlüsselideen und -konzepte zur graphbasierten Modellierung von Unsicherheiten*, number 4 Special issue in Zeitschrift für digitale Geisteswissenschaften. DOI: 10.17175/SB004_004.

Voß, J., Bausch, S., Schmitt, J., Bogner, J., et al. (2014). *Normdaten in Wikidata Handbuch*. Hochschule Hannover Fakultät III Abteilung Information und Kommunikation, http://nbn-resolving.de/urn/resolver.pl?urn:nbn:de:bsz:960-opus4-4381.

Vrandečić, D. and Krötzsch, M. (2014). Wikidata: A Free Collaborative Knowledgebase. *Communications of the ACM*, 57(10):78–85, DOI: 10.1145/2629489.

Wickham, H. (2007). Reshaping Data with the Reshape Package. *Journal of Statistical Software*, 21(12), http://www.jstatsoft.org/v21/i12/paper.

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer, New York, NY, https://ggplot2.tidyverse.org.

Wickham, H. and Henry, L. (2020). *tidyr: Tidy Messy Data*, https://CRAN.R-project.org/package=tidyr. R package version 1.0.2.

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., et al. (2016). The FAIR Guiding Principles for Scientific Data Management and Stewardship. *Scientific Data*, 3:160018, DOI: 10.1038/sdata.2016.18.

Wuttke, U. (2019). "Here be dragons": Open Access to Research Data in the Humanities. https://ulrikewuttke.wordpress.com/2019/04/09/open-data-humanities/.