



Dealing with Doubts: Modelling Approaches in site georeferencing

Florian Thiery¹

¹Research Squirrel Engineers Network, Mainz, Germany

Abstract. This paper discusses three data-driven modelling approaches handling vague and uncertain geo-references (findspots) based on literature as Linked Open Data.

Keywords: uncertainty, RDF, Linked Data, Linked Open Data, FAIR

DOI: <https://doi.org/10.5281/zenodo.10403509>

1 Introduction

Humanities and natural sciences research must handle doubts such as vagueness, uncertainty and ambiguities in data modelling. This occurs especially in georeferencing. However, creating reproducible and comprehensible data for re-use while guaranteeing data quality in archaeological data involves disclosing doubts and ambiguities [4]. Especially vagueness and uncertainty must be modelled to work with geodata. However, for linking data and FAIRification graph-based modelling as Linked Open Data (LOD) proposed by Berners-Lee (2006) [1] (cf. Schmidt et al. (2022) [3]) is the method and technique of choice. Due to the huge variety of research domains, an interdisciplinary, commonly understandable modelling of uncertainties and vagueness in research data is extremely challenging. We will present three data-driven modelling approaches handling vague and uncertain geo-references (findspots) based on literature as Linked Open Data.

2 Modelling Approaches

The main idea and goal is to publish and model the following information:

- describe where the geoinformation comes from
- describe the method of how the coordinate was created
- describe the uncertainty issue(s)
- use references into the Semantic Web

2.1 Wikidata

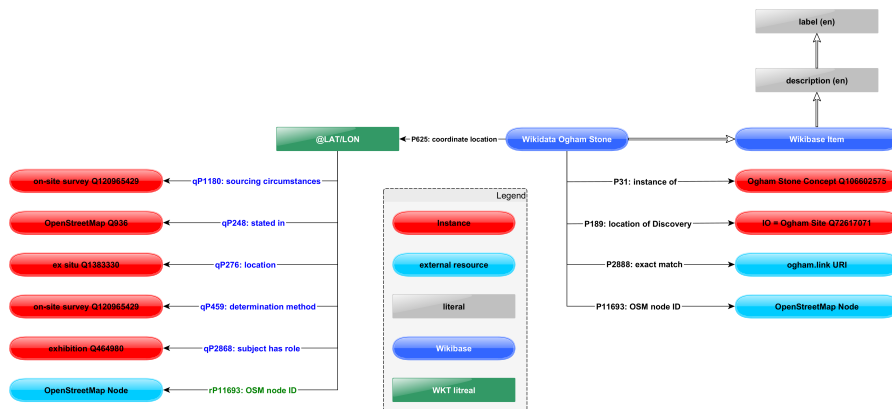


Figure 1: Ex-situ Ogham Stone with on-site survey. Florian Thierry CC BY 4.0.

Entities in Wikidata (here, an Ogham Stone) may have coordinates (P625). These coordinates contain uncertainty and reference information, which can be modelled using Wikidata qualifiers and references. In the case of Ogham Stones, which have been visited in on-site surveys, this can be done as follows (Figure 1):

- qualifier: P1180 sourcing circumstances
- qualifier: P248 stated in
- qualifier: P276 location
- qualifier: P459 determination method
- qualifier: P2868 subject has role
- reference: P11693 OpenStreetMap node ID

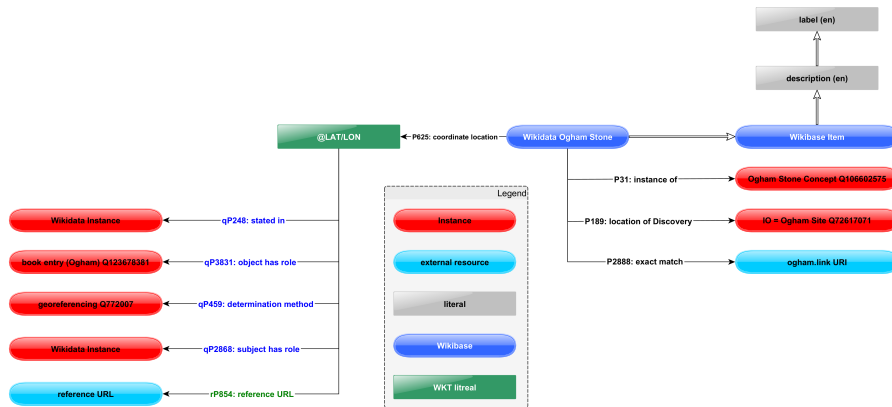


Figure 2: Ogham Stone: only Literature as Source. Florian Thiery CC BY 4.0.

Ogham Stones, which are only available in literature and/or online databases such as “Ogham in 3D” or CISP (Figure 2):

- qualifier: P248 stated in
- qualifier: P3831 object has role
- qualifier: P459 determination method
- qualifier: P2868 subject has role
- reference: P854 reference URL

2.2 Fuzzy Spatial Locations Ontology (FSLO)

The Fuzzy Spatial Locations Ontology is based on PROV-O, SKOS and GeoSPARQL. It follows the PROV-O concept of Entity, Activity and Agent (Figure 3). In the case of this ontology, Sites (entity) have a Geometry and were created using a Method (activity) by a Person (agent), c.f. Figure 4. The site and geometry include two properties to describe certainty: `fsl:certaintyDesc` and `fsl:certaintyLevel`; and sites on top properties to describe references to, e.g., books, `fsl:hasReference` and to online resources, e.g. via `exactMatch` properties from the SKOS vocabulary. The method can be described via sources (`fsl:hasSource`, `fsl:hasSourceType`), references (`fsl:hasReference`), method descriptions (`fsl:activityDesc`), and certainty information (`fsl:certaintyLevel`, `fsl:certaintyDesc`). The resulting Linked Open Data as RDF can be transformed into human-readable HTML files using the SPARQL Unicorn Ontology Documentation research tool [2].

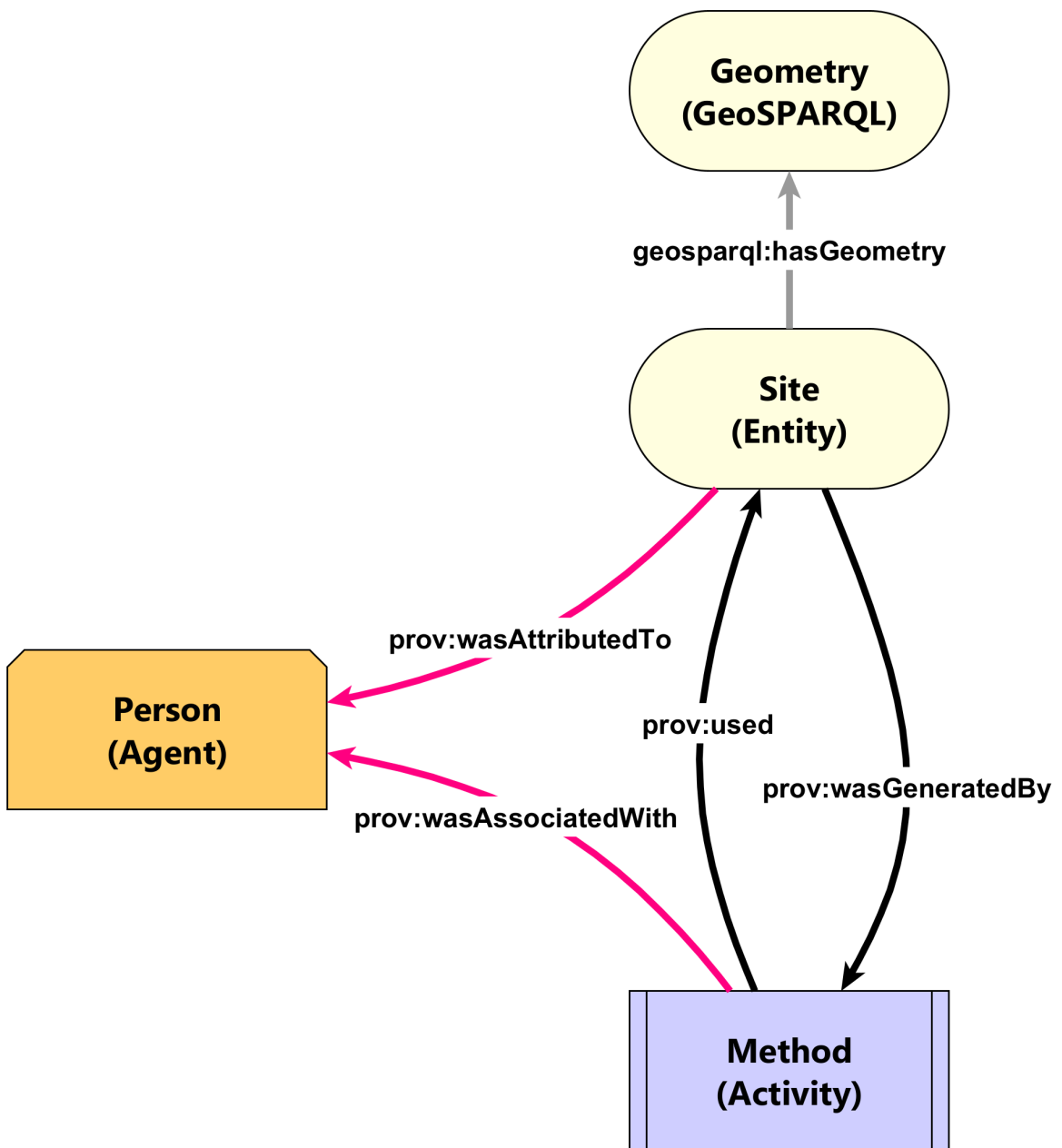


Figure 3: Schematic view of the Fuzzy Spatial Locations Ontology based on the PROV-O ontology. Florian Thiery CC BY 4.0.

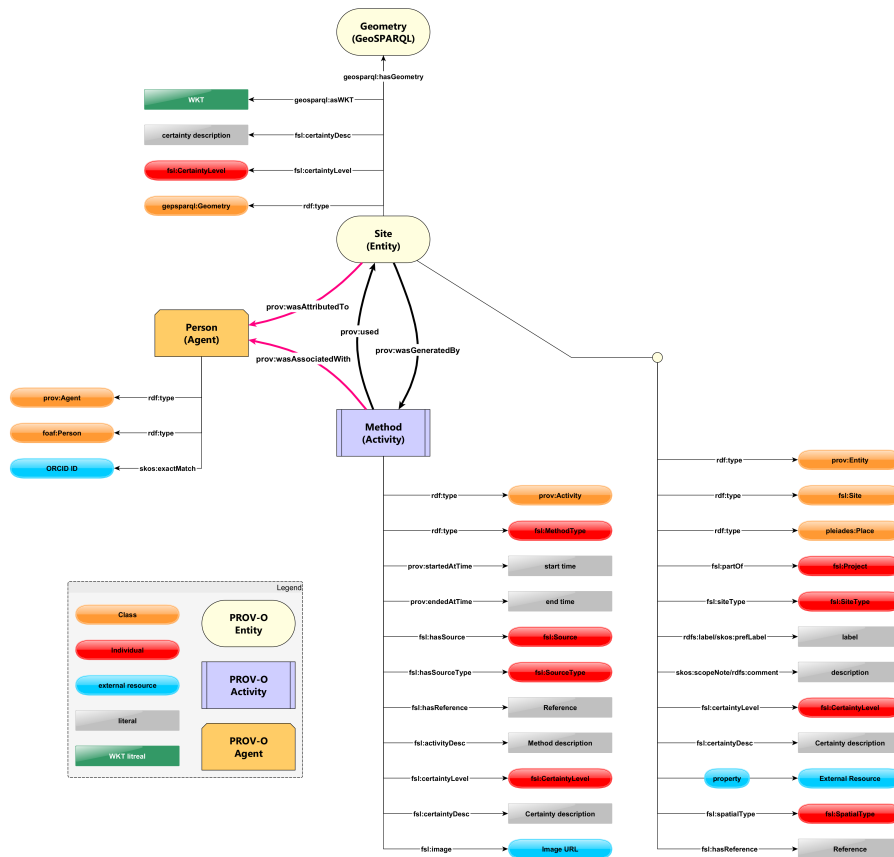


Figure 4: Schematic view of the Fuzzy Spatial Locations Ontology and its relations, based on the PROV-O ontology. Florian Thiery CC BY 4.0.

2.3 Wikibase

The Wikibase modelling (Figure 5) is related to the Wikidata modelling approach. Here, a site also has a lat/lon coordinate, which has qualifier to describe it further:

- has certainty level
- certainty description
- method used
- acting person
- has source
- has source subtype
- method description

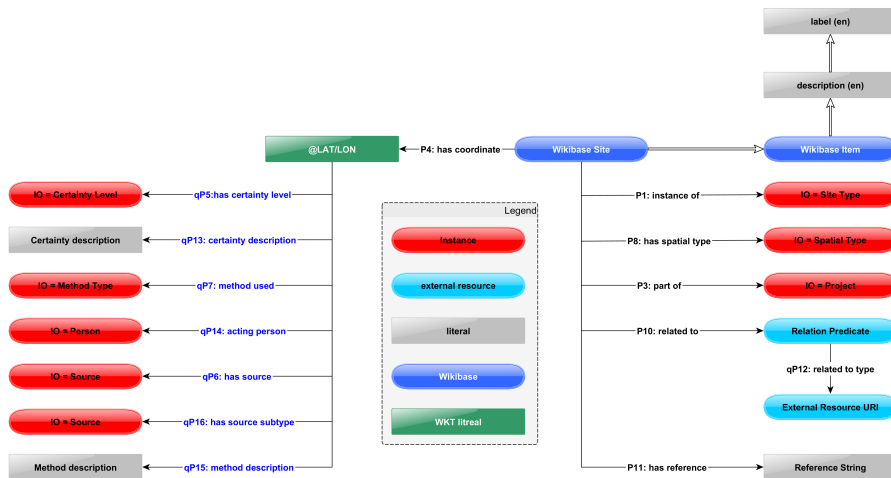


Figure 5: Wikibase Data Model of a site. Florian Thiery CC BY 4.0.

REFERENCES

- [1] Berners-Lee, T.: Linked Data - Design Issues, 2006. <https://www.w3.org/DesignIssues/LinkedData.html>.
- [2] Homburg, T.; Thiery, F.: Sparql unicorn ontology documentation, 2023. <http://doi.org/10.5281/zenodo.8190763>.
- [3] Schmidt, S.C.; Thiery, F.; Trognitz, M.: Practices of Linked Open Data in Archaeology and Their Realisation in Wikidata. Digital, 2(3), 333–364, 2022. ISSN 2673-6470. <http://doi.org/10.3390/digital2030019>.
- [4] Thiery, F.; Mees, A.: Taming Ambiguity - Dealing with doubts in archaeological datasets using LOD, 2023. <http://doi.org/10.15496/PUBLIKATION-87762>. Publisher: Tübingen University Press.