WARREN — Language

<0 images>

Language, Cultural Influences, and Intelligence in Historical Gazetteers of the Great War

Warren, R.

This talk concerns itself with the binarization of British and German Western Front maps during the Great War and their modelling using Linked Open Data (LOD) ontologies. Through the use of image processing algorithms, the scanned images of British and German trench maps were binarized into vector representations that allow us to compare each actor’s perception of the other as well as infer military intelligence activities by tracking the provenance of toponyms.

## \* \* \*

As with many armed conflicts, the initial intentions and plans of the belligerents were quickly changed when the outcomes of their actions were unsatisfactory. All sides expected a war of maneuver using the classical arms of previous conflicts: infantry, cavalry, and artillery. Commanders from the previous centuries would have been comfortable with the pre-war planning of both sides. However, terrain and innovations in artillery and small arms rate of accurate fire constrained the parties to a war of attrition where movement was no longer possible. Military thinking moved from planning large strategic maneuvers to limited tactical engagements that required local planning.

Having to deal for the first time with large numbers of tactical situations requiring good maps, the different groups of belligerents began looking for mapping solutions. The British Expeditionary Forces fighting on Belgian and French soil benefited from Belgian cartographic knowledge from the onset that was then augmented from other French sources (Jack, 1920). The Germans were faced with the problem of invading a hostile country with uncooperative officials and having to build a mapping system from the ground up using opportunistically captured material (Boelcke, 1920; Hinks, 1919; Great Britain War Office, 1917) as well as previous large-scale maps.

This talk will cover the geometry and cartography of the Battle of Vimy Ridge in 1917 as a case study using both British and German cartographic material.

## Merging Geometries

The military maps were created from an amalgam of Belgian and French sources augmented by field surveys and aerial photography. Overlaying these maps allows us to determine what each belligerent believed about his opponent at that point in time. In some cases, friendly trench details were omitted from maps to avoid leaking information in case of capture. However, what trenches are shown represent an armies’ ability to keep track of its own operational activities.

Several editions of each map exist, and some instances of these dated editions survive and have been scanned in digital form. We binarize trench representations by separating the colour of the trench from the image while using Thiessen polygons. Registration of the digitized map is done using OpenCV’s implementation of Hough Line Transform driving a web service implementation previously described in Warren and Evans (2014).

Multiple Linked Open Data vocabularies were used to record this information, including GeoSparql (OGC, 2012), NeoGeo, and W3C geo:Points. By recording the individual points within the polygon in a manner similar to that of Linked Geo Data (Stadler et al., 2012), we are able to annotate individual locations within the polygon. This effectively creates polygons formed by a series of geo:Points that can each record not only the shape of the trenches but also calculated positional errors, linkages to other trenches, and trench width.

Depending on the specific type of trench, its width could be anywhere from two feet, six inches to one foot, six inches, depending on the application, and the depth about five feet, six inches (Great Britain War Office, 1914). The depth information is not recorded in the maps, but the width of the trench is sometimes traced from aerial photography, which allows for the measurement of high-traffic trenches. When the scaled trench trace exceeds the minimum one-foot, six-inch width, the width calculated from the pen trace on the map is used instead.

A recent improvement brought about by GeoSPARQL is the differentiation between Feature and Geometry. The Geometry is the materialized physical location of the Feature (thing), and the separation between thing and location allows for greater flexibility, for example, in stating the existence of a place without necessarily defining where it is. It also allows us to link different representations of the same trench across maps by assigning the same trench (Feature) multiple Geometries as represented in multiple source documents at different times.

This also allows us to reduce redundant information by merging similar geometries across maps where no change has occurred. Small variations in locations (due to wear and tear as well as minor survey errors) occur, and these are taken into account by the merging process by calculating the theoretical accuracy of both map scales against each other. For large variations of friendly trenches, the change can be assumed to be a change in the layout of the trench, while for large variation in the observed enemy trench, the change is assumed to be done through a better intelligence estimate. Comparing both these sets of trench geometries over time is a means of obtaining a quantitative measure of the quality of both military operation and intelligence activities.

## Toponyms and Intelligence

The toponyms used on the maps varied greatly, and translated Belgian and French town names were generally used when known. Other visually prominent features such as farms, woods, or isolated trees (e.g., The Lone Tree) were named on an ad-hoc basis based on local events or the cultural baggage of the unit involved in the survey work (e.g., Regina Trench).

A common practice for all sides was to re-use their opponent’s toponyms when these were known for printing on maps issued to frontline officers. For example, Prinz Arnolf Graben, a trench near modern-day Beaumont-Hamel, France, is shown on British maps using the German Army name and spelling. This was done for both frontline tactical purposes and making the navigation of enemy trenches easier during an attack or raid. The re-use of names also marks a turning point in British and German attitudes to military intelligence. Intelligence is no longer seen as a cloak-and-dagger craft involving shady individuals but as a core component of military operations. At the onset of the war, cavalrymen still carried a sketch board onto which they were supposed to draw enemy dispositions for a commander’s review. By 1916, the information flowed from frontline operations back to the printing sections that created maps on an as-needed basis for local tactical use.

Through the use of algorithmic methods the language of trench names is isolated and their source identified across different maps. Within this specific application, this allows us to cross-reference the same trench across the maps of multiple belligerents while establishing the source, or earliest known use, of the toponyms. Previous recording of the name of things, or gazetteers, tended to record objects as a single name, occasionally with a localization (Chasseaud, 2006). Through the use of the SKOS eXtension for Labels (SKOS-XL) vocabulary, we improved on this by adding label-specific provenance, which allows the tracking of the original owner of the name. This recording of the nomenclature provenance can be used as a proxy that identifies the intelligence processes of the belligerents.

## Conclusion

These military maps are an important source of historical locations that have been long forgotten or whose importance has changed over time. Through the binarization of scanned trench maps, the locations of smaller engagements can be extracted and the location of named features can be placed in a modern context. Some of these features, such as the location of former ammunition dumps, remain important to this day.

In closing this presentation, the use of the semantic web will be reviewed in managing a historical gazetteer of geometries, features, and names during the Great War on the Western Front. The careful tracking of provenance information will be explained, and the novel use of existing semantic web standards allows for the discovery of both the quality of the cartographic work done by both sides and the cultural influences between belligerents.

References

**Boelcke, S.** (1920). Die engländer und das deutsche kriegsvermessungswesen. *Petermanns Geographische Mitteilungen,* **66**(6–9).

**Chasseaud, P.** (2006). *Rats Alley: Trench Names of the Western Front, 1914–1918*. Spellmount.

**Great Britain War Office.** General Staff. (1914). *Notes on Field Defenses*. His Majesty’s Stationery Office.

**Great Britain War Office.** General Staff. (1917). *Handbook of the German Army in War, January, 1917*. n.p.

**Hinks, A. R.** (1919). German War Maps and Survey. *Geographical Journal,* **53**(1): 30–40.

**Jack, E. M.** (1920). *Report on Survey on the Western Front, 1914–1918*. His Majesty’s Stationery Office, London.

**OGC.** (2012). OGC GeoSPARQL—A Geographic Query Language for rdf Data. *Open Geospatial Consortium*, Technical Report OGC 11-052r4, September.

**Stadler, C., Lehmann, J., Höffner, K. and Auer, S.** (2012). Linkedgeodata: A Core for a Web of Spatial Open Data. *Semantic Web Journal,* **3**(4): 333–54.

**Warren, R. and Evans, D.** (2014). Translating Maps and Coordinates from the Great War. In *Proceedings of the Terra Cognita Workshop at ISWC 2014*, Riva Del Garda, Italy, October 2014.