Kartography CIC

A Critical and Practical Approach to Data and Knowledge

Date: 25th / 26th March 2024 **Supported by:** The National Archives UK **Venue:** The National Archives, Richmond, London
http://kartography.org

Introduction

Data is a broad term that refers to facts and figures without context. In its simplest form, data can be numbers, text, images, or other symbols that, in themselves, do not carry purposeful and consistent meaning. For example, the words ‘Harold’, ‘Hastings’, the number 2000, and the year ‘1066’, are all pieces of data, but without additional context and relationships it is difficult to convey the details of an event, like a battle, and the many influences that led to it. We rely on disciplinary conventions, software and other sources to help fill in the gaps. This explains why databases are often limited to references and finding aids, and why data from different sources is difficult to integrate. Creating data involves significant resources and time. This course is about significantly improving the benefits and value of data - maintaining and improving its relevance (context) into the future.

While a database uses data types, semantic data also uses semantic typing. The CIDOC CRM defines universal concepts, like a person, a place, an event, or a concept, to clarify language and provides meaningful relationships to generate a human and computer ‘data narrative’. Using this framework we can continually expand and integrate knowledge across many different subjects, for example, art and architecture, literature, music, archaeology, anthropology, and history generally. It allows interpretations and arguments, which are usually represented in separate textual narrative, to be part of the semantic data network. In adopting semantic data, and using it to collaborate, we can reveal the bigger picture.

Kartography CIC is a non-profit social enterprise with a charter directed at diversity and inclusion in data.

The two day course has five main themes;

1. To introduce practical data skills that enable a transition from traditional data to richer and meaningful representations and help you understand the long term benefits of semantic data.
2. To provide practical skills in Semantic Linked Data - and specifically through the use of CIDOC CRM - for creating and using knowledge graphs.
3. To emphasise how the creation of semantic data leads to effective visualisations and communication - furthermore it enables publication and sharing of data with context, provenance and history.

4. To explain the design principles workings of the ResearchSpace platform, a native Linked Data system that uses CIDOC CRM ontologies at its core.
5. To provide a critical theory perspective on using technology for historical representation using current tools.

Additional Aspects

- How to prepare a traditional spreadsheet to produce data that can be transformed into Semantic Linked Data.
- How to convert from two common data formats (CSV / XML) into Semantic Linked Data.
- How to use public resources (in this case Wikidata) to prime richer data environments using the Linked Data query language - SPARQL.
- How to manipulate data and adjust it using regular expressions.
- Understand Semantic Linked Data and take ownership of its use.

Pre-requisites

- It makes no assumptions about current technical skills and is designed for people working in knowledge professions and researchers using examples that are humanities related.
- The course is relevant across many disciplines and organisational knowledge processes.
- It provides the knowledge to help plan new digital projects and apply for funding.

Course Details

Day 1 - Practical Data Skills

The course starts with an introduction to the wider issues with current published data and the need for new approaches. It provides simple guidelines for using spreadsheets to record data that can be easily converted to semantic Linked Data, and how to perform this transformation. There is an examination of two different data resources from The Old Bailey Proceedings (TEI) and Wikidata and how to 'clean' and convert the data using various freely available tools.

Day 2 - ResearchSpace Semantic Knowledge System?

After consolidating the work from day 1, data will be imported to the ResearchSpace system (http://researchspace.org) and various aspects of the system will be presented to explain the underlying design principles and the issues of working in a native semantic environment. This includes a practical overview of the system to inform decision making and the design of a semantic knowledge base project.

Practical

- Attendance is £450 - which includes post course support and manual.[1]

Participants will receive a course manual to reference in the course. A laptop is needed with Visual Code installed (Instructions for installation will be provided on registration). Online tools will be used in most exercises.

To request registration please send an email with your name, organisation, and position to,

knowledge_course@kartography.org

[1] Ability to ask specific technical questions related to the technical aspects course content for 6 months.