Non-representational approaches to visualise complex information in the Cultural Heritage domain

Pasqual, Valentina
valentina.pasqual2@unibo.it
Università di Bologna, Italy

Pedretti, Carlo Teo
carloteo.pedretti@uniroma1.it
Università La Sapienza di Roma, Italy

Schimmenti, Andrea
andrea.schimmenti2@unibo.it
Università di Bologna, Italy

Tomasi, Francesca
francesca.tomasi@unibo.it
Università di Bologna, Italy

Vitali, Fabio
fabio.vitali@unibo.it
Università di Bologna, Italy

Introduction

Data visualisation tools help users explore connections in a dataset, discover new meanings and foster an understanding of a domain. They are considered crucial epistemic aids (Windhager, Salisu, and Mayr 2019), especially for Linked Data (LD) collections on the Semantic Web. In recent years, semantic models have been subject to new formalisations that can effectively convey complex information, including meta-information (provenance, reliability) and contexts (concerning a graph or subgraph), which are especially relevant in the cultural heritage (CH) domain. Traditionally, some of the most common RDF data visualisation metaphors employ graphs or tables. Most online tools support graphs, while most knowledge bases (e.g. DBpedia, Wikidata, etc.) employ tables (Desimoni and Po 2019). A pretty limited panorama in respect of the heterogeneity of CH data.

Nonetheless, in the humanities, uncertainty is an integral part of the domain, and it is a scholarly appropriate habit to include both preferred as well as contested assertions (Daquino et al. 2021; Barabucci, Tomasi, and Vitali 2021).

By introducing uncertainty and contextuality in the data modelling pipeline, it becomes more evident that they are not data, they are not natural representations of facts: the non-representational paradigm in Drucker (2011) calls them *capta*, stressing that data is always interpreted, situated and constructed from a perspective. This paradigm shift must also affect the traditional visualisation approaches for LD collections.

Furthermore, most tools available for visualising RDF data fail to consider the differences between end-users and employ identical visualisation metaphors for their datasets regardless of their audiences' goals, interests and proficiency. In particular, while those metaphors may be evident to experienced RDF users, they may fall flat and sometimes be incomprehensible to other users.

Henceforth, we need new approaches for visualising complex semantic data while considering the different competencies of intended users: we must identify a broader set of evaluation metrics for this purpose.

Some examples of novel visualisation models

The following are some examples of visualisation models which have yet to be explored by the semantic community to tackle the challenge of effectively communicating complex information to users. While we have not implemented these visualisation models, they represent potential approaches that could be applied to different domains and datasets.

Higraphs: Graphs are suited for power users: they offer an overview of the graph content and are essential for discerning clusters of linked data. Yet they offer little structures for complex concepts such as provenance or uncertainty. Hypergraphs (or higraphs) use closed curves around graphs, creating sets that can have, in turn, edges to each other or other nodes. A graph can become a structure, delimited by its area, appropriate for naming and labelling purposes (Harel 1988). This graph formalism is particularly apt to describe Named Graphs and several approaches to meta-information about entities and relationships (including RDF Star and Nanopublications), efficiently conveying fairly complex information in a clean visual format.

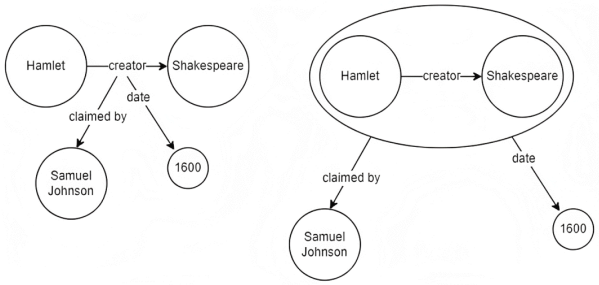

Figure 1: Graph as displayed by isSemantic (left) and Higraph (right).

Maps: Domain experts and non-experienced users can benefit from visual metaphors that stretch across different semantic fields while supporting qualitative information that enables distortions, fuzziness or blurriness of an otherwise plain surface. For example, *emotional maps* allow places to be shown as a space indicated by geographic coordinates and as symbols where pivotal historical or fictional episodes took place, or authors lived. At the same time, topographical representations can be distorted by emotional factors concerning distance travelling (Drucker 2011). This creates a dialogue between the end-user and the system via a non-representational approach.

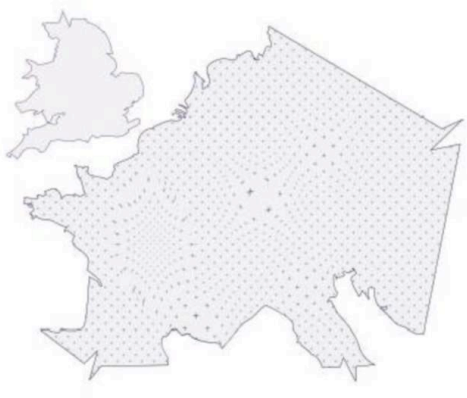

Figure 2: Map with space warped by the experience of travel so that the standard distances are distorted by effect of emotion and infrastructure (Drucker, 2011).

In particular, complex information such as uncertainty can be conveyed in maps using techniques that alter existing symbology through manipulating visual variables (e.g. colour) or adding new objects (e.g. glyphs). Moreover, adjacent views between certain and uncertain data and dynamic interfaces (e.g. popup) could be employed for the same purposes (Kinkeldey, MacEachren, and Schiewe 2014).

Text: The text translation of RDF can be considered a type of extractive visualisation, i.e. the transformation of content into another form, which is more in the register of non-experienced users than traditional RDF linearisations. Natural Language Generation techniques to transform RDF into text have been explored since 2006 (Sun and Mellish 2006). The WebNLG+ challenge explored translation possibilities from RDF datasets to text and vice-versa. However, so far, they have yet to be used to represent RDF (Castro Ferreira et al. 2020). Rendering RDF as text can be a new medium for non-experienced users to perform CRUD operations on RDF databases and provide a widely understandable visualisation.

```
wd:Q254032 wdt:P21 wd:Q6581072 ;
wdt:P106 wd:Q36180,
         wd:Q49757,
         wd:Q212980 ;
wdt:P569 "1962-01-29T00:00:00Z"^^xsd:dateTime ;
wdt:P19 wd:Q457865 ;
wdt:P800 wd:Q21076759,
         wd:Q9171173,
         wd:Q11827997 .
```

"Olga Tokarczuk is a female writer, poet and psychologist. She was born on 29 January, 1962 in Sulechów. Their notable works are The Books of Jacob, Drive Your Plow Over the Bones of the Dead and Flights".

Figure 3: A subgraph regarding writer Olga Tokarczuk (Q254032) from Wikidata, with a possible corresponding "visualisation" in natural language.

Text may be paired with other visualisations (e.g. maps, timelines etc.) and question-answering interfaces. For example, any uncertainty modelled in RDF can be shown in natural language texts through adverbs, footnotes or multiple competing statements and represented according to well-known linguistic patterns.

User-oriented metrics for visualisation approaches

Complex information may increase the noise of a visualisation in a given interface and affect the cognitive load of users based on their competencies and tasks. Evaluation for information visualisation tools is only partially suited to capture the conceptual impact of uncertainty-based systems (Hullman 2020). We propose an evaluation matrix based on four metrics, considering users' needs and motivations to measure the effectiveness of varying visualisation approaches. These parameters reflect a more holistic and user-centred approach to evaluating the effectiveness of information visualisation tools.

- Comprehension metric: verify if the tester can infer simple information from the visualisation; targets any user.
- Retention metric: verify if the tester can remember simple information from the visualisation after a given time; targets any user.
- Weirdness detection metric: verify if the tester can notice apparent inconsistencies in the visualisation; targets data curators and owners.
- Drafting metric: verify if the tester can (rapidly) produce a record about a new entity according to existing models and syntaxes of the current visualisation; targets data curators and owners.

Conclusions

While new paradigms for RDF data modelling have spread, there still needs to be more effective approaches and tools for visualising complex, contextual and uncertain information. In this paper, we proposed novel non-representational approaches to support complex information visualisations while also trying to meet the needs of diverse user types. Higraphs, emotional maps and textual representations expand existing visual metaphors and standards while enabling novel ways of exploring datasets and complex information. Non-representational approaches can be adopted without disrupting standard visualisation by building onto them, especially to represent uncertainty, provenance and contextual information. By incorporating a more user-centred set of evaluation metrics, our proposed evaluation matrix provides a more thorough quality assessment of visualisation interfaces and their representativeness of complex, contextual, and uncertain information. Although the proposed models have yet to be fully implemented, it is essential to move in this direction, as a humanistic approach to data visualisation can provide greater complexity and depth to users.

Bibliography

Barabucci, Gioele, Francesca Tomasi, and Fabio Vitali. 2021. "Supporting Complexity and Conjectures in Cultural Heritage Descriptions." 104-115. https://ntnuopen.ntnu.no/ntnu-xmlui/handle/11250/2736994.

Castro Ferreira, Thiago, Claire Gardent, Nikolai Ilinykh, Chris van der Lee, Simon Mille, Diego Moussallem, and Anastasia Shimorina, eds. 2020. Proceedings of the 3rd International Workshop on Natural Language Generation from the Semantic Web (WebNLG+). Dublin, Ireland (Virtual): Association for Computational Linguistics. https://aclanthology.org/2020.webnlg-1.0.

Daquino, Marilena, Valentina Pasqual, Francesca Tomasi, and Fabio Vitali. 2021. "Expressing Without Asserting in the Arts."

Desimoni, Federico, and Laura Po. 2019. Empirical Evaluation of Linked Data Visualization Tools.

Drucker, Johanna. 2011. "Humanities Approaches to Graphical Display." Digital Humanities Quarterly 005 (1).

Harel, David. 1988. "On Visual Formalisms." Communications of the ACM 31 (5): 514–30. https://doi.org/10.1145/42411.42414.

Hullman, Jessica. 2020. “Why Authors Don’t Visualize Uncertainty.” IEEE Transactions on Visualization and Computer Graphics 26 (1): 130–39. https://doi.org/10.1109/TVCG.2019.2934287.

Kinkeldey, Christoph, Alan MacEachren, and Jochen Schiewe. 2014. “How to Assess Visual Communication of Uncertainty? A Systematic Review of Geospatial Uncertainty Visualisation User Studies.” The Cartographic Journal 51 (September): 1743277414Y.0000000099.

Sun, Xiantang, and C. Mellish. 2006. “Domain Independent Sentence Generation from RDF Representations for the Semantic Web.”

Windhager, Florian, Saminu Salisu, and Eva Mayr. 2019. “Exhibiting Uncertainty: Visualizing Data Quality Indicators for Cultural Collections.” Informatics 6 (3): 29.