Data narratives with Linked Open Data, the case of mythLOD storytelling

Pasqual, Valentina
valentina.pasqual2@unibo.it
Digital Humanities Advanced Research Centre (/DH.arc), Department of Classical Philology and Italian Studies, University of Bologna, Bologna, Italyorcid

Tomasi, Francesca
francesca.tomasi@unibo.it
Digital Humanities Advanced Research Centre (/DH.arc), Department of Classical Philology and Italian Studies, University of Bologna, Bologna, Italyorcid

Introduction

Cultural heritage (CH) data are constantly growing from digitisation projects carried out by many GLAM institutions. The Digital Humanities community widely adopted Linked Open Data (LOD) as a standard for knowledge organisation in CH domain promoting interoperability and exchange (Hyvönen 2012, Alexiev 2018). Several cultural heritage platforms disseminate their data through browsable online catalogues with faceted search interfaces, e.g. Europeana[1]. Graphic User Interfaces (GUI) make data accessible also to non-IT experts such as museum curators, humanities scholars, students, or the general public, by hiding - sometimes complex - SPARQL queries. More recently, dissemination practices rely on *narration* as a new expedient to interactively involve users in a more efficient and effective way (Eide et al. 2019, Windhager et al. 2019). Storytelling and data storytelling practices can visualise latent relations and new perspectives over data, augmenting users' knowledge. While faceted search allows users to customise data exploration depending on individual interests, storytelling drives the users through a defined narrative path. In this context, mythLOD collection (Pasqual and Tomasi 2021) is the revalorisation of Mythologiae data collection[2] through the conversion of its data from tabular data to LOD. The project workflow followed the curation of digital objects life-cycle described by Tomasi (2022), addressing the data dissemination task with two visualisations, an online catalogue browsable through facets and a data storytelling experiment applied to Aeneid by Virgil case study.

Modelling Data and visualisation

mythLOD[3] is a Knowledge Base (KB) storing more than 4'000 heterogeneous museum artwork from different artists, locations and epoques. In particular, besides descriptive metadata (e.g. title, date of creation, current location), domain experts annotated each artwork with a list of literary sources (from classical to contemporary production) sharing the same mythological theme or scene. For example, the artwork "Landscape with Aeneas at Delos"[4] depicts "Aeneas towards the Italic Peninsula"[5], which can be found in *Aeneid, III, 69-123* by Virgil and *Divina Commedia* by Dante Alighieri. Additionally, mythLOD KB preserves interpretations' contextual information (i.e. the above mentioned example is an hermeneutic analysis with iconographic criterion performed by person P). Moreover, literary works are recorded in mythLOD KB with related descriptive metadata (e.g. author, period of creation). In particular, canonical citations, modelled with Humanities Citation Ontology (HuCit) (Romanello and Pasin 2013), are recorded in the form work, book, chapter, verse.

mythLOD data are accessible via two main visualisations, namely the catalogue and the storytelling sections in the website (cf. the following section). mythLOD catalogue[6]presents a complete overview of the artworks data which can be browsed through facets filters as their metadata (collocation, periods, keywords, type, author, mythological category), related references (reference type, reference author) and experts' interpretations (author, interpretation type and criterion).

While offering a complete overview of the dataset, the two visualisations have been implied as an actual testing activity to the whole dataset with the domain experts. In fact, the visualisations graphically represent the project competency questions hiding their technicalities. Both visualisations have been exposed to domain experts involved in Mythologiae to review both the representativeness of mythLOD data model and the correctness of data in the Knowledge Base.

Storytelling paths

Telling Aeneis through museal art in mythLOD is the data-storytelling section about Aeneid contents in mythLOD collection. The storytelling consists of a restricted set of visualisations that investigate the relation between mythLOD artworks and Aeneid literary reference by re-proposing the project competency questions in four categories.

- Artworks temporality (*when*): a timeline illustrates when the museum works concerning the themes of the Aeneid were produced
- Artworks spatiality (*where*): a map shows where the museum artworks concerning the themes of the Aeneid are located
- Artworks contents (*what*): as exemplified in figure 1, a heatmap represents the most frequently cited Aeneid passages in the artworks depicted scenes; a word cloud represents the most frequent terms implied in the description of the artworks related to Aeneid; a network represents the literary references which co-occur with Aeneid in the dataset.
- Agents involved with artworks (*who*): a network shows the relations between authors who created an artwork whose depicted scene has also been recognised to be in common with Aeneid.

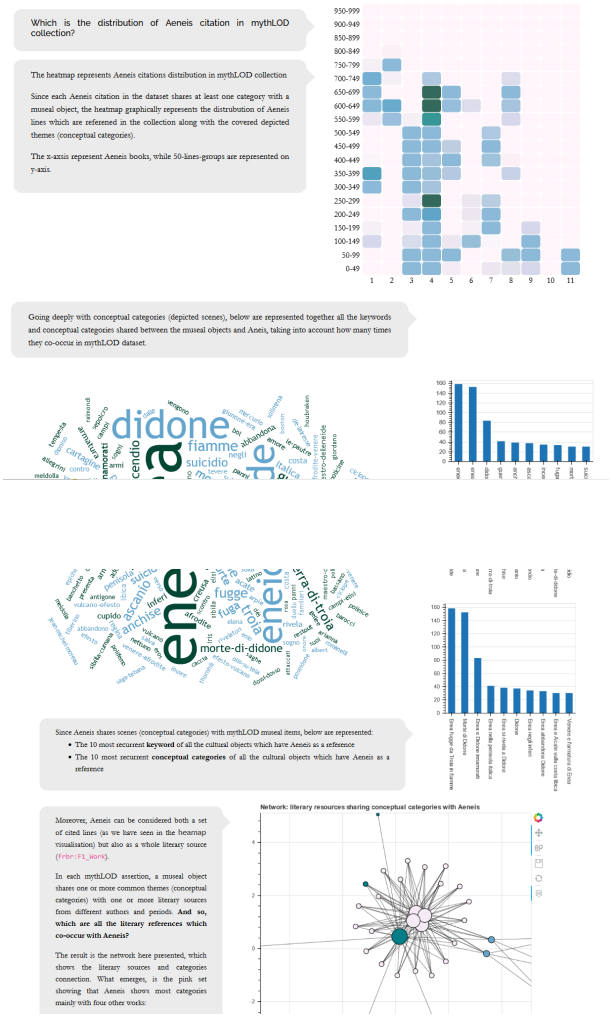

Figure 1. Snapshots of contents visualisation (what) in mythLOD storytelling

Besides the four categories of interest, a textual introduction and tips for each visualisation have been added to guide the user through the four sections.

Visualisations are strictly dependent from the whole data curation process, as knowledge organisation (e.g adopted data models and ontologies), data cleaning and alignment with external resources. For example, the granular representation of canonical citations is then presented in the *heatmap* (figure 1), highlighting Aeneid most cited lines and motives in the KB. Additionally, dates have been converted in a machine-readable format used for the creation of the timeline and fostering the valorisation of temporal data. Moreover, geospatial coordinates of the locations have been retrieved from Wikidata to visualise the instances on the map.

Conclusions

In conclusion, the visualisations fully exploit together museal and bibliographic sources. In particular, mythLOD visualisations show how the format migration from table to LOD naturally allows shifting the point of view from the collection (the artworks as subject of the catalogue) to the literary sources (the Aeneid as subject of the storytelling) and testimonies that data-visualisation can reveal the latent knowledge that was only intuitable in Mythologiae tabular collection. Additionally, visualisation, as demonstrated, is thus an example of how to employ procedures aimed at enhancing the role of the end user in accessing resources (for example as an actual part of the testing activity). It is one of those research actions that want to see LODs as a new way not only of publishing, but also of semanticising the cultural heritage, i.e. of enriching the access experience by developing new forms of organising, and thus of using, knowledge. In a phrase, the visualisation of narratives is not just a matter of interface, but a new service to acquire new knowledge from data.

Notes

1. https://www.europeana.eu/it, accessed 24/10/2022
2. https://mythologiae.unibo.it/, accessed 24/10/2022
3. https://dharc-org.github.io/mythlod/, accessed 24/10/2022. The ID of this artwork in mythLOD collection is item/534.
4. https://www.nationalgallery.org.uk/paintings/claude-landscape-with-aeneas-at-delos, accessed 24/10/2022
5. Cf. translation from italian “Enea verso la penisola italica”
6. https://dharc-org.github.io/mythlod/catalogue/, accessed 24/10/2022

Bibliography

Alexiev, Vladimir (2018): “Museum Linked Open Data: Ontologies, Datasets, Projects”, in: Digital Presentation and Preservation of Cultural and Scientific Heritage 8 (September): 19–50. https://doi.org/10.55630/dipp.2018.8.1.

Daquino, Marilena / **Pasqual, Valentina** / **Tomasi, Francesca** (2020): “Knowledge Representation of Digital Hermeneutics of Archival and Literary Sources”, in: JLIS.It 11 (3): 59–76. https://doi.org/10.4403/jlis.it-12642.

Eide, Øyvind / **Wieners, Jan** / **Schubert, Zoe** / **Türkoglu, Enes** (2019): “The intangibility of tangible objects: re-telling artefact stories through spatial multimedia annotations and 3D objects”, https://doi.org/10.5281/zenodo.3878966.

Hyvönen, Eero (2012): Publishing and Using Cultural Heritage Linked Data on the Semantic Web. Synthesis Lectures on Data, Semantics, and Knowledge. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-031-79438-4.

Pasqual, Valentina / **Tomasi, Francesca** (2022): “Linked open data per la valorizzazione di collezioni culturali: il dataset mythLOD”, in: AIB studi 62 (1): 149–68. https://doi.org/10.2426/aib-studi-13301.

Romanello, Matteo / **Pasin, Michele** (2013): “Citations and Annotations in Classics: Old Problems and New Perspectives”, in: Proceedings of the 1st International Workshop on Collaborative Annotations in Shared Environment: Metadata, Vocabularies and Techniques in the Digital Humanities, 1–8. DH-CASE ’13. New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/2517978.2517981.

Tomasi, Francesca (2022): Organizzare la conoscenza: Digital Humanities e Web semantico. Milano: Editrice Bibliografica.

Windhager, Florian / **Federico, Paolo** / **Schreder, Günther** / **Glinka, Katrin** / **Dörk, Marian** / **Miksch, Silvia** / **Mayr, Eva** (2019): “Visualization of Cultural Heritage Collection Data: State of the Art and Future Challenges”, in: IEEE Transactions on Visualization and Computer Graphics 25 (2019): 2311-2330