Multimodal Knowledge Graphs for Semantic Analysis of Cultural Heritage Data

Nitisha Jain

Hasso-Plattner-Institut, University of Potsdam, Germany
`nitisha.jain@hpi.de`

The availability of massive amounts of data has led to significant progress in the development of intelligent systems in different domains. Knowledge graphs (KGs) have become ubiquitous as structured repositories for several intelligent applications that are data-driven, such as search, recommendation [5] and question answering systems [2]. The data being generated today comes from diverse sources in various formats and modalities including texts, images, audios as well as videos. Therefore, multimodal knowledge graphs that can store and harvest the rich information from different modalities are gaining attention.

Cultural heritage is a prominent domain where a large variety and volume of data has been made available by the recent increase in digitization efforts for ancient historical and cultural materials [1,4]. The availability of digitized archives is geared towards facilitating faster and efficient analysis of the cultural heritage resources by intelligent systems. Knowledge graphs can play an important role in this setting by acting as semantically sound representations of cultural heritage entities. One of the desirable features of a KG for the cultural heritage domain is the representation of multimodal data, i.e. the visual and textual art-related entities, as well their relations. The meta-data associated with the art entities (such as the artistic style, technique and period for a painting) also play a definitive role in semantic analysis and therefore, should be included in the KG.

Multimodal representation and analysis of data is particularly appealing for the cultural heritage domain, where the textual and visual information often complements each other. In many cases, texts and images are correlated and essential to represent the information comprehensively. For instance, it is common for multiple distinct paintings to have the same title. In such cases, textual data i.e. titles are not sufficient for accurate identification of the paintings, and the image of the painting acts as a unique identifier. A KG with multimodal links and entities can help with this task and subsequently, enable multimodal named entity linking and named entity disambiguation for artworks. Several other semantic tasks can benefit substantially from multimodal analysis, e.g. visual question answering can retrieve images as answers, search and browsing can be based on both text and image, and semantic similarity and linking can be performed for a combination of text and image entities.

Constructing a multimodal knowledge graph based on data obtained from digitized art materials is a non-trivial task due to the missing association between the images of artworks and the textual data. A semantic alignment between the

images and the corresponding text with the help of multimodal embeddings can enable this matching [3]. This work explores the challenges and interesting research questions for the linking and representation of textual and visual cultural heritage entities so as to enable the construction of a multimodal knowledge graph for the cultural heritage domain.

References

1. Chen, C.c., Wactlar, H.D., Wang, J.Z., Kiernan, K.: Digital Imagery for Significant Cultural and Historical Materials. International Journal on Digital Libraries **5**(4), 275–286 (2005)
2. Cui, W., Xiao, Y., Wang, H., Song, Y., Hwang, S.w., Wang, W.: KBQA: Learning Question Answering over QA Corpora and Knowledge Bases. Proc. of the VLDB Endowment **10**(5) (2017)
3. Jain, N., Bartz, C., Krestel, R.: Automatic Matching of Paintings and Descriptions in Art-Historic Archives using Multimodal Analysis. In: 1st International Workshop on Artificial Intelligence for Historical Image Enrichment and Access (AI4HI-2020), co-located with LREC 2020 conference (2020)
4. Van Hooland, S., Verborgh, R.: Linked Data for Libraries, Archives and Museums: How to Clean, Link and Publish your Metadata. Facet publishing (2014)
5. Zhang, F., Yuan, N.J., Lian, D., Xie, X., Ma, W.Y.: Collaborative Knowledge Base Embedding for Recommender Systems. In: Proc. of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. pp. 353–362 (2016)