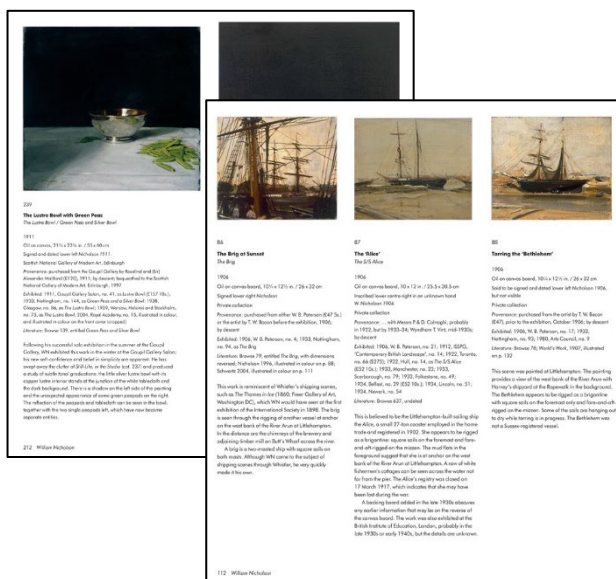# Multimodal Analysis for Cultural Data

Cultural heritage data assumes a pivotal role in the understanding of human history and culture. A wealth of information is buried in historical art archives that can be extracted via the digitization and analysis of these resources. This can facilitate search and browsing, help art historians to track the provenance of artworks and enable wider semantic text exploration for digital cultural resources.

The information in cultural data is contained in **images** of artworks as well as **textual descriptions** or annotations accompanied with the images. For example, in exhibition or auction catalogues, one or more images are printed along with their descriptions. However, with the digitization of such resources for the extraction of the data, the valuable associations between the images and texts are frequently lost. As such, the identification of artworks from the text or from the images alone, and the subsequent linking of these artworks to existing knowledge bases is a non-trivial task. We would like to retrieve these associations between the images and the texts for artworks to enrich the data associated with the artworks and enable their structured and semantic representation.

Project Outline

This project lies at the intersection of two prominent research areas – **computer vision** and **natural language processing**. The aim of this project is to address the issue of identification and linking of artwork images to their corresponding text annotations or descriptions. We explore a large collection consisting of auction catalogues, exhibition catalogues as well as art books and handwritten letters etc., from which text was extracted with the help OCR and the images were obtained via scanning.

Project Approach

This project involves the analysis of unstructured texts as well as images of artworks to explore the techniques of linking the images to the correct texts. We want to use deep learning methods in order to link images of artworks with their textual description. For instance, automatic image captioning can be used to ascertain that a particular image depicts a house with mountains in the background, while the description of text contains similar words and provides clues about the image. Several images might be depicting common themes such as portraits of kings and queens, or landscapes with similar features. In such cases, linking of the images with the text would require further inspection of the particular traits that distinguish similar images, perhaps the color tone of a painting or the style of a particular artist.

If you want to work with us in this project, you should be interested in doing research in one or more of the following research areas:

- Machine learning (especially deep learning)
- Computer vision
- Text mining
- Digital cultural heritage

http://dissertationreviews.org/archives/8672
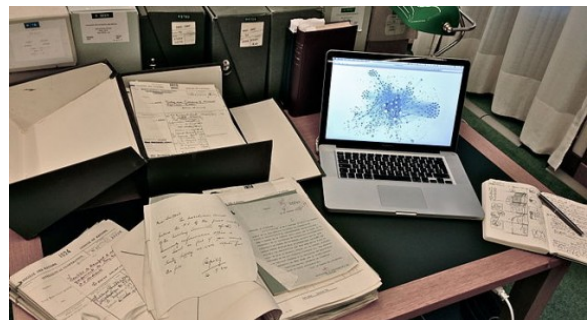
Contact

This project will be jointly supervised by the Information Systems chair and the Internet technologies and Systems chair. If you have any questions, please do not hesitate to contact Christian Bartz (christian.bartz@hpi.de) or Nitisha Jain (nitisha.jain@hpi.de).

Related Work

[1] Jain, Nitisha, and Ralf Krestel. "Who is Mona L.? Identifying Mentions of Artworks in Historical Archives." In International Conference on Theory and Practice of Digital Libraries, pp. 115-122. Springer, Cham, 2019.

[2] Bartz, Christian, et al. "LoANs: Weakly Supervised Object Detection with Localizer Assessor Networks." Asian Conference on Computer Vision. Springer, Cham, 2018.

[3] Huang, Xingsheng, Sheng-hua Zhong, and Zhijiao Xiao. "Fine-art painting classification via two-channel deep residual network." Pacific Rim Conference on Multimedia. Springer, Cham, 2017.

[4] Elgammal, Ahmed, et al. "The shape of art history in the eyes of the machine." Thirty-Second AAAI Conference on Artificial Intelligence. 2018.