Knowledge Graph based Analysis and Exploration of Historical Theatre Photographs

Tabea Tietz[1,2], Jörg Waitelonis[3], Mehwish Alam[1,2], and Harald Sack[1,2]

[1] FIZ Karlsruhe – Leibniz Institute for Information Infrastructure, Germany
firstname.lastname@fiz-karlsruhe.de
[2] Karlsruhe Institute of Technology, Institute AIFB, Germany
[3] yovisto GmbH, Potsdam, Germany
joerg@yovisto.com

Abstract. Historical theatre collections are an important form of cultural heritage and need to be preserved and made accessible to users. Often however, the metadata available for a historical collection are too sparse to create meaningful exploration tools. On the use case of a historical theatre photograph collection, this position paper discusses means of automated recognition of historical images to enhance the variety and depth of the metadata associated to the collection. Moreover, it describes how the results obtained by image recognition can be integrated into an existing Knowledge Graph (KG) and how these generated structured image metadata can support data exploration and automated querying to support human users. The goal of the paper is to explore cultural heritage data curation techniques based on deep learning and KGs to make the data findable, accessible, interoperable and reusable in accordance with the F.A.I.R principles.

Keywords: Cultural Heritage · Linked Data · Knowledge Graphs · Exploratory Search · Image Recognition · Deep Learning.

1 Introduction

Theatre as a form of art and as a social institution has been firmly anchored in our societies for hundreds of years and has become indispensable. It is constantly evolving with artistic epochs, and plays which have existed for hundreds of years are just as relevant to today's theatre culture as new modern plays. Even though theatre has prevailed for such a long time, it does not exist on its own and is always embedded within the context of a time, culture, a society, political systems, and technological developments which grant possibilities but also set boundaries of creativity and artistic expression. Scientists from a variety of domains investigate how theatre develops in history and how it copes with these changing conditions [8,15,11]. Therefore, it has been recognized that preserving theatre data and making it openly accessible is vital.

The American Theatre Archive Project[4] emphasizes the meaning for theatre data preservation and published a manual and call to action to engage all theatres

[4] https://www.americantheatrearchiveproject.org/, last accessed: Oct. 29, 2019

in the U.S. in the endeavour [1]. In another effort, the Baden-Württemberg State Archives[5] collected and digitized data of Stuttgart State Theatres from the 1890s to the 1940s[6]. The collection consists of nearly 7.000 black and white photographs along with metadata. The photographs give insights into on-stage events like theater plays, operas and ballet performances as well as off-stage moments and theater buildings. The collection was made publicly available via the archive's Web platform[7]. However, solely publishing data on the Web does not automatically ensure that they are in practice found and used. Especially users who are unfamiliar with archival practises to structure information often find it challenging to access and explore the provided content [6,7].

To tackle these shortcomings, Linked Stage Graph was created [18]. The Knowledge Graph (KG) was built on the foundation of the aforementioned archival data about the Stuttgart State Theatres with the goal to enable researchers as well as the general public to access, analyze and explore the data in intuitive, interesting and useful ways[8]. Next to the KG, Linked Stage Graph so far contributes a publicly available SPARQL endpoint[9] to enable sophisticated queries for expert users, the extraction and linking of named entities mentioned in the metadata to the Wikidata KG and the German Integrated Authority File (GND) as well as timeline interfaces for data exploration[10][11].

The current visual means of explorations within Linked Stage Graph are time based. Timelines are a powerful tool for the exploration of historical content. However, a mere temporal data exploration is not always possible or useful for two reasons. Firstly, due to the complexity of cultural heritage data, [22] suggest that for a successful exploration by a broad range of users, multiple dimensions have to be taken into account when providing access to the data, because "every possible encoding method can capture only so much of a collection's composition". Another reason is data sparsity, because not every collection contains structured date information. While Linked Stage Graph contains 7000 images to explore in theory, there are only (structured) date information available for roughly 2600 of them, which means that less than half of the photographs can be explored (visually) through the provided temporal dimensions. However, data sparsity within the Linked Stage Graph metadata does not only regard temporal data. Besides the performances' titles and few additional information like the original author of the artwork there are at most sparse information on set or costume designers. These information were linked to the GND, Wikidata and DBpedia if possible. To provide further means of exploration, further metadata have to be generated.

[5] https://www.landesarchiv-bw.de/, last accessed: November 4, 2019
[6] https://www2.landesarchiv-bw.de/ofs21/olf/startbild.php?bestand=17460, last accessed: November 4, 2019
[7] https://bit.ly/2Ne7saD, last accessed: Oct. 29, 2019
[8] http://slod.fiz-karlsruhe.de/
[9] http://slod.fiz-karlsruhe.de/sparql/
[10] http://slod.fiz-karlsruhe.de/vikus/
[11] http://slod.fiz-karlsruhe.de/#Viewer

This position paper discusses the means to perform automated recognition on historical images to enhance a cultural heritage collection's metadata on the use case of historical theatre data. To support data exploration that is helpful for human users but also via automated querying, it is discussed how the image recognition results can be integrated into the existing KG. The goal of the paper is to explore cultural heritage data curation techniques at the intersection between machine learning and KGs to make the data findable, accessible, interoperable and reusable in accordance with the F.A.I.R principles [21]. While the use case described in this paper specifies on theatre data, the use case is generalizable to a broad range of collections containing historical images.

2 Related Work

There have been many recent initiatives which use methods from Computer Vision for analysing images from art history in digital humanities [20,10]. It is further used for two kinds of analysis, (i) qualitative data analysis and (ii) automated analysis. Atlas.ti[12] is a tool which allows qualitative data analysis over the images for digital humanities for facilitating the researchers/educators in creating, manipulating and curating the image data. In the current study we focus mainly on the automated analysis of the images.

Automated analysis includes the mechanism which automatically perform image captioning or object detection. Several studies have been proposed with the focus on object detection and visual searches of the images from art history. In [4,5], the authors propose a classification approach using Convolutional Neural Networks for detecting objects in paintings where the classifier is trained on natural images causing a problem of *domain shift*. It further discusses that the method is more effective when the model is learnt on the regions of images instead of the entire image. A combination of both levels of features increases the performance of the classifier.

In [16], the authors demonstrate how the links between the images can be established using similarity. Many other studies focus on providing an interface to the users to collect and explore art image collections. The work by Crissaff et al. enables art historians to explore and organize digital art, supporting image comparison [3]. Lang et al. allow the organisation of art work based on similarity [13]. WikiArt[13] is an encyclopaedia for visual arts which allows exploration by artists and the art work.

Given the methodologies described above, our work focuses on providing a tool which allows the exploration of images from Linked Stage Graph. One of the challenging aspects of Linked Stage Graph is that it contains images of one of the scenes in the play which carries limited information in comparison to paintings, meaning that, the paintings depict the whole scenario of what a painter wants to portray. From the algorithmic aspect, some of the many issues in the targeted domain include unknown objects at a certain point in time, fictitious objects,

[12] `https://atlasti.com/`

[13] `https://www.wikiart.org/`

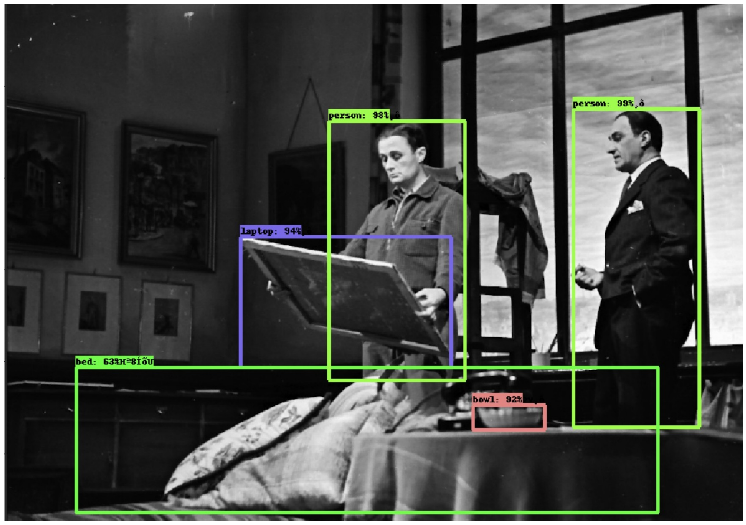

Fig. 1. Detection of a laptop (purple frame) as an example of modern objects falsely detected in historical images.

etc. (see Section 3). To the best of our knowledge, this is the only study which considers a combination of KGs (from the theatrical domain) and methods from computer vision to provide exploration over these images.

3 Recognition of Historical Theatre Images

With the advent of Deep Neural Networks (DNN), a promising way to cope with the lack of available metadata for historical image archives is to exploit pre-trained models on images from a different domain and then fine-tune them on historical resources [2]. In the course of the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) sophisticated DNN architectures for visual recognition tasks have been developed and trained that are now publicly available as pre-trained models [12]. For visual content recognition, images are categorized according to one or several pre-specified or learned objects or object classes. ImageNet[14] is a large dataset of annotated images organized by the semantic hierarchy of WordNet[15]. ILSVRC uses a subset of ImageNet images for training and testing the models, of which some have been published as part of commonly used public machine learning programming libraries, as e.g. in Tensorflow[16].

Image Processing further comprises three major sub-tasks, i.e., (i) Image Classification, (ii) Object Detection and (iii) Image Captioning. Image classification refers to the task of categorizing entire image, whereas, object detection

[14] http://imagenet.stanford.edu/

[15] https://wordnet.princeton.edu/

[16] https://github.com/tensorflow/models/tree/master/research/slim

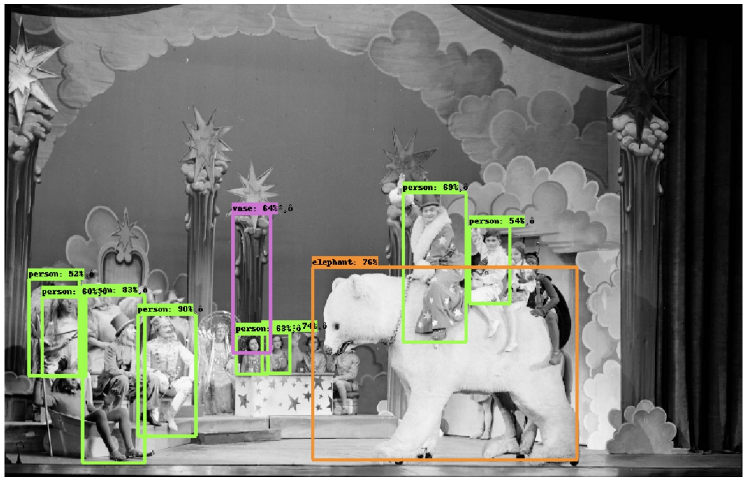

Fig. 2. A giant polar bear in a fictional scene

refers to locating and detecting specific objects in an image. Finally, image captioning processes the content of the whole image or uses it in combination with object detection algorithm to generate the textual description of the content of the entire image.

The above described algorithms can directly be applied to the historical content using pre-trained models. However, by directly applying these models trained with modern images, several problems can be identified concerning temporal inconsistencies and specific issues related to theatre and art domain. Temporal inconsistencies may occur when models are trained on objects present in the modern world such as modern telephones (e.g. smartphones) instead of historic cable phones, i.e., the tool will not recognize the telephones during the time period to which the image belongs. The same problem occurs if models are trained on modern objects which did not yet exist in history, e.g., laptops (as displayed in Figure 1). Issues which are specific to the theatre domain occur for example when objects and setup as displayed might not correspond to our reality. In many cases, the objects may be completely invented and do not exist in the real world such as *flux capacitor* in the movie *Back to the Future*. Moreover, many objects are stylized, e.g., the shadow of a fan projected on a wall might stand for a windmill. Furthermore, object scaling can be an issue. As an example, Figure 2 shows a theatre stage with a gigantic polar bear carrying several people.

To overcome this problem, training data on historical images and images fitting to the theatre domain have to be generated. This can be achieved for instance by crowd-sourcing approaches and by expert-user annotations. This way of obtaining training data is expensive and time consuming. This is one of the reasons why existing approaches use the same methodology, i.e., using the pre-trained model over contemporary images (see Section 2). The current study

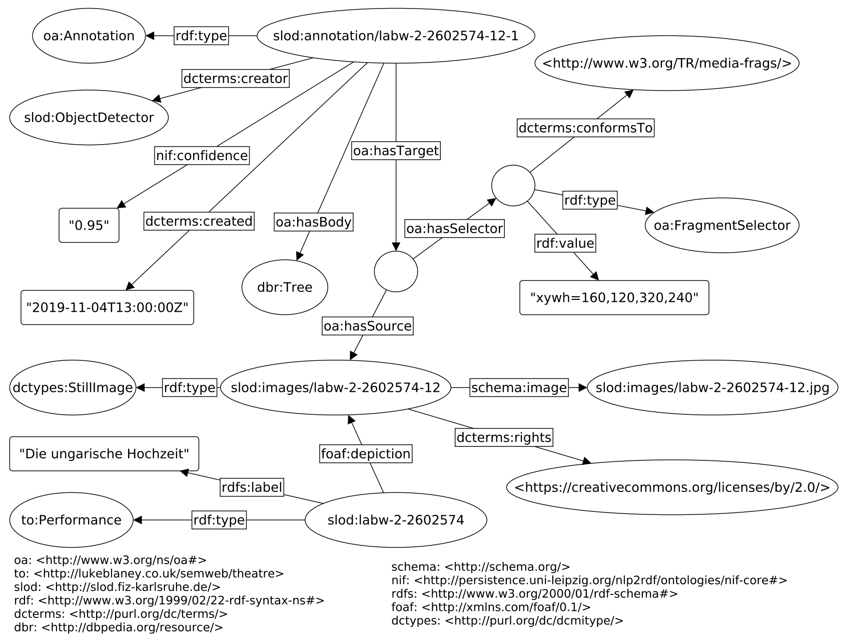

Fig. 3. Example representation of an object detection result integration into the KG

uses the pre-trained model as a starting point. Further inference or rule based approaches will be employed over the contextual information obtained from the plot of the play to verify the correctness of the objects detected by the algorithm.

4 Knowledge Graph Integration

Once the results of the analysis as described in section 3 are computed, they have to be integrated into the existing KG. The goal is to create machine understandable and thus meaningful metadata about each image in the dataset to enable further dimensions of exploration and to allow for all analysis results to be reproducible and transparent. For these purposes, the Web Annotation Vocabulary[17] provides a useful framework. Figure 3 shows an example of one image annotation. A resource, e.g. the play `slod:labw-2-2602574` is depicted via an image which is analyzed using object detection. The resulting image annotation is modeled via the annotation target, consisting of the source and the spatial fragment selector and the annotation body, in this case the DBpedia resource `dbr:Tree`. In order to make the origin of the annotation transparent, the annotation system as well as the date and time of the annotation are also modelled. The confidence of the automated annotation is integrated using the NIF2 ontology [9]. Once the results are integrated into the KG as described, the

[17] `https://www.w3.org/TR/annotation-vocab/`, last accessed: Oct. 31, 2019

analysis results are transparent and reproducible. Furthermore, when the KG is enhanced with further metadata, these structured information can be utilized for content exploration tools.

5 Exploration and Exploratory Search

The presented image analysis methods allow to implement means of exploration beyond time based exploration tools already implemented and the traditional archival practices. This is especially useful for users non-familiar with the typical structure of archives such as retroconverted archival finding aids. With the newly generated content-related descriptions of the images other perspectives on the image data might be adopted. The image descriptions thereby include *categories* from image classifications, *objects* detected within the images or textual *captions* describing an image motif.

From these information different exploration means might be provided such as semantic *facet-based* browsing [14]. Each facet might represent a category or object. Activating a facet selects all images relevant to that item. The user now can use this selection as an entry point to further browse the archive content. Based on the data in the current use case, some useful facets are:

- **persons:** male, female, seated, standing, laying
- **interior/exterior stage design:** dining table, chair, mirror, tree, grass
- **requisites/costumes:** military uniform, spiked helmet, sword, bowtie, dress
- **animals:** pigs, sheep, birds, horses
- **means of transportation:** boats, cars, bicycles
- **buildings:** palace, church, restaurant/pub

The usefulness of these facets depends of course on the intersection of what the AI models provide and what the user's information need is. To bring these two together it is sometimes necessary to act on a higher abstraction level. If categories, objects, and also the captions are linked to formal KG, a taxonomy might be used to further aggregate the items according to super classes or categories (bold faced in the item list).

However, when linked to a KG, any kind of relation might be utilized to enable *relation-based browsing*. The user then is able to navigate along semantic relations and therewith move through the archive content. An interactive visualization as demonstrated by [17] could enable that and might also help to discover previously unknown relations between archival units.

In general, the entire portfolio of *semantic search and recommendation* technologies can be applied to provide added value to the user. In a semantic search application on top of the Linked Stage Graph, with concept search more precise search results are achievable. Implementing a ranking function incorporating the semantic relatedness between concepts would help to increase the hit ratio (recall) of the search result [19]. Likewise, semantic graph recommendations are desirable especially when working with relatively small data collections as the Linked Stage Graph is. A user not familiar with the content domain will get insights about the content more quickly when supported by meaningful recommendations. For example, if a theatre researcher is interested in the 'presentation of military' in different plays, it would be beneficial if the system would propose the items of the collection relevant to related concepts such as 'military uniform', 'spiked helmet', 'rifles', etc.

6 Conclusion

Theatre has been anchored in our societies for hundreds of years and even plays which were performed during the early 1600s like Midsummernight's Dream are of high relevance in modern theatre. However, theatre has always also been shaped by the context of its social, political and technological developments. Therefore, initiatives to digitize and curate historical theatre data have become relevant to researchers. Linked Stage Graph is a KG based on the historical theater data of Stuttgart theatres. Next to a SPARQL endpoint, the already implemented means of explorations are time based. However, to enable a more complete exploration of the dataset, further metadata is needed. This paper discussed the potential to analyze the images in the dataset by various means, integrate the results into the KG and develop exploration environments. Even though these means for analysis and curation of historical data is discussed on the use case of theatre data, the approach is generalizable to other domains in which historical photographs are provided.

References

1. Brady, S., et al.: Preserving theatrical legacy: An archiving manual for theatre companies. American Theatre Archive Project (2015)
2. Chatfield, K., Simonyan, K., Vedaldi, A., Zisserman, A.: Return of the devil in the details: Delving deep into convolutional nets. CoRR **abs/1405.3531** (2014)
3. Crissaff, L., Ruby, L.W., Deutch, S., DuBois, R.L., Fekete, J., Freire, J., Silva, C.T.: ARIES: enabling visual exploration and organization of art image collections. IEEE Computer Graphics and Applications **38**(1), 91–108 (2018)
4. Crowley, E.J., Zisserman, A.: In search of art. In: Computer Vision - ECCV 2014 Workshops - Zurich, Switzerland, September 6-7 and 12, 2014, Proceedings, Part I. pp. 54–70 (2014)
5. Crowley, E.J., Zisserman, A.: The Art of Detection. In: Computer Vision - ECCV 2016 Workshops - Amsterdam, The Netherlands, October 8-10 and 15-16, 2016, Proceedings, Part I. pp. 721–737 (2016)
6. Ferro, N., Silvello, G.: From users to systems: Identifying and overcoming barriers to efficiently access archival data. In: ACHS@ JCDL (2016)
7. Freund, L., Toms, E.G.: Interacting with archival finding aids. Journal of the Association for Information Science and Technology **67**(4), 994–1008 (2016)
8. Halbach, F.: Judenrollen: Darstellungsformen im europäischen Theater von der Restauration bis zur Zwischenkriegszeit, vol. 70. Walter de Gruyter (2008)
9. Hellmann, S.: NIF 2.0 Core Ontology. Ontology Description, AKSW, University Leipzig, `http://persistence.uni-leipzig.org/nlp2rdf/ontologies/nif-core/nif-core.html` (2013)

10. Hentschel, C., Wiradarma, T.P., Sack, H.: Fine tuning CNNS with scarce training data - adapting imagenet to art epoch classification. In: 2016 IEEE International Conference on Image Processing, ICIP 2016, Phoenix, AZ, USA, September 25-28, 2016. pp. 3693–3697 (2016)
11. Kornetis, K.: Stage of emergency: Theater and public performance under the greek military dictatorship of 1967-1974. Journal of Greek Media and Culture **3**(1), 125–131 (2017)
12. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. Commun. ACM **60**(6), 84–90 (May 2017)
13. Lang, S., Ommer, B.: Attesting similarity: Supporting the organization and study of art image collections with computer vision. DSH **33**(4), 845–856 (2018). https://doi.org/10.1093/llc/fqy006
14. Osterhoff, J., Waitelonis, J., Jäger, J., Sack, H.: Sneak Preview? Instantly Know What To Expect In Faceted Video Searching. In: Proceedings of 41. Jahrestagung der Gesellschaft für Informatik (INFORMATIK 2011). Lecture Notes in Informatics, vol. P192. Gesellschaft für Informatik (GI) (2011)
15. Räthel, C.: Wieviel Bart darf sein? Jüdische Figuren im skandinavischen Theater. Narr Francke Attempto Verlag (2016)
16. Seguin, B., Striolo, C., diLenardo, I., Kaplan, F.: Visual Link Retrieval in a Database of Paintings. In: Computer Vision - ECCV 2016 Workshops - Amsterdam, The Netherlands, October 8-10 and 15-16, 2016, Proceedings, Part I. pp. 753–767 (2016)
17. Tietz, T., Jäger, J., Waitelonis, J., Sack, H.: Semantic Annotation and Information Visualization for Blogposts with refer . In: V. Ivanova, P. Lambrix, S.L.C.P. (ed.) Proceedings of the Second International Workshop on Visualization and Interaction for Ontologies and Linked Data (VOILA '16). vol. 1704, pp. 28 – 40. CEUR-WS (2016), `http://ceur-ws.org/Vol-1704/paper3.pdf`
18. Tietz, T., Waitelonis, J., Zhou, K., Felgentreff, P., Meyer, N., Weber, A., Sack, H.: Linked Stage Graph. In: Alam, M., Usbeck, R., Pellegrini, T., Sack, H., Sure-Vetter, Y. (eds.) Proceedings of the Posters and Demo Track of the 15th International Conference on Semantic Systems co-located with 15th International Conference on Semantic Systems (SEMANTiCS 2019). vol. 2451, pp. 131–135. CEUR-WS (2019), `http://ceur-ws.org/Vol-2451/#paper-27`
19. Waitelonis, J., Exeler, C., Sack, H.: Linked Data enabled Generalized Vector Space Model to improve document retrieval. In: Proceedings of the Third NLP & DBpedia Workshop co-located with the 14th International Semantic Web Conference 2015 (ISWC 2015). vol. 1581 (2015)
20. Wevers, M., Smits, T.: The visual digital turn: Using neural networks to study historical images. Digital Scholarship in the Humanities (01 2019). https://doi.org/10.1093/llc/fqy085
21. Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.W., da Silva Santos, L.B., Bourne, P.E., et al.: The fair guiding principles for scientific data management and stewardship. Scientific data **3** (2016)
22. Windhager, F., Federico, P., Schreder, G., Glinka, K., Dörk, M., Miksch, S., Mayr, E.: Visualization of cultural heritage collection data: State of the art and future challenges. IEEE transactions on visualization and computer graphics **25**(6), 2311–2330 (2018)