Workshop
2-3 September 2021

Provincial Historical Archive of Alicante
Guillén de Castro 3, 03007 Alicante (Spain)

In person

Digital

INNOVATION ON NEW DIGITAL EXPONENTIAL TECHNOLOGIES IN THE ARCHIVES

2nd, morning.

Moderated by Zoltán Szatucsek

9:30 Welcome words. Spanish State Archives.

9:45 Kerstin Arnold. *Aggregation as a service - Automatic topic detection and collaborative topic tagging in Archives Portal Europe's multilingual environment.*
10:15 Artem Reshetnikov. *St. George on a bike: Enrichment of metadata of Cultural Heritage Objects using deep learning.*

10:45 BREAK

11:15 Panel discussion. Artem Reshetnikov, Joan Andreu Sánchez, Thomas Aigner and Kerstin Arnold.
12:30 LUNCH BREAK.

2nd, afternoon.

14:00 Hands On workshop. Joan Andreu Sánchez & Enrique Vidal. *Handwritten Text Recognition for the European Digital Treasures Collections.*

16:00 CLOSE OF FIRST DAY.

3th, morning.

Moderated by Thomas Aigner

9:30 Daniel Stromer. *Browsing through sealed historical documents: non-invasive imaging methods for document digitization.*
10:00 Torsten Hiltmann and Philipp Schneider. *Automated Processing and Exploitation of Heraldic Data from European Archives.*
10:30 Barbara Revuelta-Eugercios. *Link-Lives: building historical big data from archival records for use by researchers and the Danish public.*
11:00 BREAK.

11:30 Zoltán Szatucsek. *How did language technology affect the publication of the GULAG victims' records.*
12:00 Gregor Završnik. *Digital geospatial data, a tool for interpretation of our past.*
12:30 Lipót Répászky. *How many faces has a person? The AI/ML in action.*
13:00 Conclusions - Zoltán Szatucsek.
13:10 CLOSE OF SECOND DAY.

"The presentations will last 20 minutes followed by a 10-minute question and answer session"

In person

Digital

SPEAKERS

Kerstin Arnold
Archives Portal Europe Foundation, Netherlands.
Kerstin has been working in the archives domain for more than 15 years, after initially starting her career in public relations and editing. During this time, she has gained expertise in a variety of areas, from academic editions of archival collections via records management and transfer of born-digital materials to aggregation of cultural heritage metadata on national, international and cross-domain levels. Having been part of the projects creating and establishing Archives Portal Europe, Kerstin is now the initiative's acting COO in the role of the APEF Manager. She holds a Master degree each in Communication Science and in Library and Information Management and also is a member of the Technical Subcommittee on Encoded Archival Standards (TS-EAS) at the Society of American Archivists. She shares her knowledge and experience with colleagues and with the next generation of archivists at conferences, workshops and by teaching at the University of Applied Sciences in Potsdam, Germany.

Artem Reshetnikov
Barcelona Supercomputing Center, Spain.
Artem Reshetnikov is a deep learning researcher at Barcelona Supercomputing Center. Working in several companies and research centers, he received big experience in Computer Vision and Natural Language Processing and applying it to the tasks of different domains. All of his life he was curious about history and art. For a long time, he was thinking about how to join his two main passions: machine learning and art. The solution is the project where he works now. Saint George on a Bike is a project about the enrichment of metadata of paintings using Deep Learning and NLP approaches. His current area of interest is the intersection of NLP and Deep Learning and using these techniques for improving object detection and caption generation in the cultural heritage domain. He graduated from the Autonomous University of Barcelona in 2019 with a Master's Degree in Engineering. Before he worked for several commercial projects focused on Data Analysis, Computer Vision, and Anomaly detection in marketing and retails in such companies as Indra or Tecnocom in Spain. The projects mainly were based on the idea of applying deep learning in traffic counting using Computer Vision, analysis of time series data for detection of anomaly behavior of clients, and replanning of marketing strategies based on detected insights. His Bachelor Degree is in Computer Science and Applied math. He loves swimming and churros with chocolate.

Joan Andreu Sánchez
Universitat Politècnica de València, Spain.
Joan Andreu Sánchez is Assistant Professor at Universitat Politècnica de València. He is Director of the Pattern Recognition and Human Language Technologies Research Center. His main research area is Machine Learning and Formal Languages applied to Text Recognition and Math Recognition. He has participated in several European projects and he leaded the tranScriptorium project. He has published more than one hundred articles related to Machine Learning and Text Recognition in international conferences and journals. He is one of the founders of the tranSkriptorium AI spin-off.

Enrique Vidal
Universitat Politècnica de València, Spain.
Enrique Vidal is a Emeritus professor of the Universitat Politècnica de València (Spain) and former co-leader of PRHLT research center in this University. He has published more than two hundred and fifty research papers in the fields of Pattern Recognition, Multimodal Interaction and applications to Language, Speech and Image Processing and has led many important projects in these fields. In the last 15 years Dr. Vidal has focussed his research in the area of Handwritten Dcument Analysis and Recognition and has lead the development of the successful Probabilisti Indexing technology. He is one of the founders of the tranSkriptorium AI spin-off company. Dr. Vidal is a fellow of the International Association for Pattern Recognition (IAPR) and his current h-index is 53 (according to Google Scholar).

Workshop

2-3 September **2021**

Provincial Historical Archive of Alicante
Guillén de Castro 3, 03007 Alicante (Spain)

In person

Digital

Daniel Stromer

Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany.

Daniel started his career at Siemens Healthineers back in 2003 before starting to study medical engineering at Friedrich-Alexander-University Erlangen-Nürnberg (FAU) in 2010. In 2019, he finished his PhD (Dr.-Ing.) in Computer Science with the title "Non-invasive imaging for Digital Humanities, Medicine, and Quality Assessment". His focus was to develop AI algorithms for image processing of non-invasive imaging systems where he mainly worked on digitization of historical documents by means of X-ray imaging. Besides, he had a research stay for seven months at Massachusetts Institute of Technology where he developed machine learning algorithms for ophthalmology applications. Daniel is now heading the Learning Approaches for Medical Big Data (LAMBDA) group of the Pattern Recognition Lab (PRL) at FAU. In parallel, he joined Siemens Healthineers as collaboration manager for Digital Health. He is currently project lead of a multimodal document digitization project as well as interested in research where pure image data is being enriched and correlated with new data sources.

Torsten Hiltmann

Humboldt-Universität zu Berlin, Germany.

Torsten Hiltmann is a professor for Digital History at Humboldt-Universität zu Berlin since 2020. His research focuses on the integration of Machine Learning and Semantic Web Technologies into historical studies and on the epistemological change of historical research through the application of digital methods, as well on visual communication in the late medieval and early modern period. As a trained medievalist, his dissertation dealt with heraldic compendia from the late Middle Ages in France and Burgundy. He heads the research project Coats of Arms in practice", where currently a heraldic database is being built as and with linked data on coats of arms and the contexts of their use.

Philipp Schneider

Humboldt-Universität zu Berlin, Germany.

Philipp Schneider is a research assistant at the chair of Digital History at Humboldt-Universität zu Berlin since 2020. He works in a project called "Coats of Arms in practice", where he is among other things responsible for modeling and contextualizing heraldry as a historical source with the help of Semantic Web Technologies. His dissertation project – also situated in this project – deals with the study of coats of arms on wall and ceiling paintings.

Barbara Revuelta-Eugercios

Rigsarkivet, Denmark.

Barbara Revuelta-Eugercios (PhD History) is associate research professor with special responsabilities at the National Archives of Denmark and associate research professor at the SAXO-Institute, University of Copenhagen. Her focus is on historical demography, mortality inequality and digital methods in history. After her doctoral studies in Spain, she has worked in research institutions in Sweden, France and Denmark in the fiels of economic history, demography and history. She co-directs the Link-Lives project since 2019, that aims at developing methods to reconstructing lifecourses for individuals in 19th century Denmark from information from censuses, parish records and other archival sources.
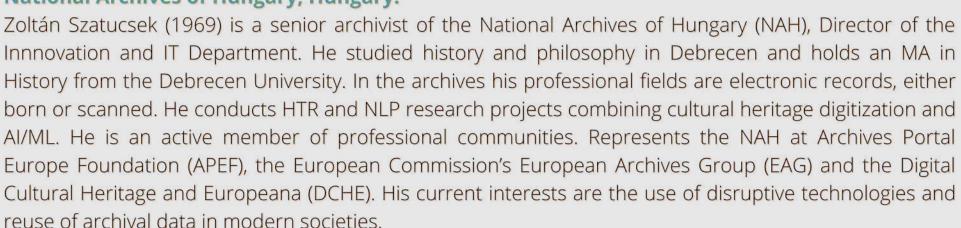
Workshop

2-3 September **2021**

Provincial Historical Archive of Alicante
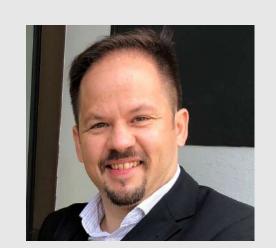Guillén de Castro 3, 03007 Alicante (Spain)

In person

Digital

Zoltán Szatucsek
National Archives of Hungary, Hungary.

Zoltán Szatucsek (1969) is a senior archivist of the National Archives of Hungary (NAH), Director of the Innnovation and IT Department. He studied history and philosophy in Debrecen and holds an MA in History from the Debrecen University. In the archives his professional fields are electronic records, either born or scanned. He conducts HTR and NLP research projects combining cultural heritage digitization and AI/ML. He is an active member of professional communities. Represents the NAH at Archives Portal Europe Foundation (APEF), the European Commission's European Archives Group (EAG) and the Digital Cultural Heritage and Europeana (DCHE). His current interests are the use of disruptive technologies and reuse of archival data in modern societies.

Gregor Završnik
Geoarh, Slovenia.

Gregor Završnik is an international consultant with more than 20 years of experience, working in the Geospatial field. He holds a Master's degree in Geodesy from the University of Ljubljana. His geospatial digital records preservation engagement started on the E-ARK project in 2015 and continues until today with his work on the CEF eArchiving Building Block. As a member of the E-ARK3 team, he is leading the development of specifications and training for the eArchiving Building block in the field of geospatial information. His clients range from national Archives to mapping agencies and digital preservation solution providers. He is also a member of the Digital Information LifeCycle Interoperability Standards Board (DILCIS) and a short term consultant for the World Bank.

Lipót Répászky
National Audiovisual Archive (NAVA), Hungary.

Lipót Répászky, graduated as a communication engineer in 2007 at the Faculty of Mechanical Engineering of Szent István University in Gödöllő, where he graduated in audiovisual engineering with a degree in digitisation solutions of 8/16/35 mm films and analogue tapes (beta, coll, vhs). After two years in the professorship, he joined NAVA's IT team as a broadcast systems engineer, where he was responsible for developing long-term storage and video encoding solutions. From 2014 he works as a technical manager and from 2016 as NAVA's managing director, his main goal is to create a long-term, open, community archive by combining machine (cognitive) knowledge with a modern, humanities-based expertise. "Changes during the operation of a steadily increasing archive pose challenges to the employees. Responses to such alterations do not come immediately but later, therefore the use of a poorly chosen coding technique or of an incorrect format in metadata space would require years of sorting out. These are the challenges in which I feel we can make decisions together as a good team, so that in 40-60 years, users coming to nava.hu (or at that point who knows what kind of virtual reality) website will find a working and usable archive."

Thomas Aigner
ICARUS, Austria.

Dr. Thomas Aigner MAS is a studied historian and archivist. He has assumed senior management positions related to cultural heritage: Director of the Archives of the Diocese of St. Pölten/AT since 1995; President of ICARUS and a major driving force behind its international activities, including cross-border and EU-funded projects supporting digitisation activities and open-access to digital content; and member of various international and national expert committees. Vice-president of Time Machine Organisation since 2018. He holds various national and international awards related to these activities (AT, CZ, DE, HR).

Links to join the workshop remotely:

September 2nd: https://youtu.be/VOsFvtNFyTU

September 3rd: https://youtu.be/qOjaNXIagxY