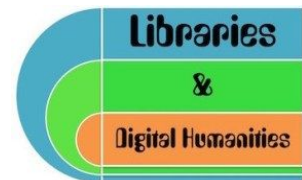


# Programme:

## Libraries as Research Partner in Digital Humanities

Preconference workshop, Digital Humanities 2019  
The Hague, July 8, 2019  
KB, National Library of the Netherlands

KB } national library  
of the netherlands



<b>Information and Acknowledgements.</b>	<b>4</b>
Local organizers	4
Programme committee	4
Volunteer reviewers	5
<b>Keynotes.</b>	<b>6</b>
Collections, Data, Research: perspectives at the national library of France	6
Community before Computing: Reflections on collections as data	6
<b>Lightning Talks.</b>	<b>7</b>
Digital scholarship and data provenance at The National Library of Scotland	7
Data stewardship on the map: A study of tasks and roles in Dutch research institutes	8
Creating a Collaboration Guidebook for Library Faculty Partnership Success	9
Project proposal for the training of Cuban librarians for Digital Humanities Research (Proyecto de formación de bibliotecarios cubanos para la investigación en Humanidades Digitales)	10
Building and Sustaining Fulcrum: A Digital Publishing Platform for Humanists and Social Scientists	11
Mnemosyne, Digital Library for Rare and Forgotten Texts (1868-1936): Collections and digital editions	11
Marketplace Lightning Talks	13
DARIAH: Core communities for collaboration: Libraries, Archives and Museums as essential partners in digital arts and humanities research	13
CERL: Data for DH at Consortium of European Research Libraries	13
IFLA Digital Humanities - Digital Scholarship Special Interest Group	13
CLARIN & research libraries: a very short introduction	14
<b>Short Papers.</b>	<b>15</b>
The case for a citation index by the humanities, for the humanities	15
Digitizing for impact: from pixels to scholarship	16
Next Generation of Scalable and Sustainable DH Services	17
Forschungspartnerschaften zu welchem Preis?	18
Collaboration with the Researchers – Content, Presentation systems and Techniques	19
Partnering up with researchers in a national library	20
Finding new environments for cooperation with digital research and new values for heritage collections: experiences at the National Library of Spain	21
Growing an international Cultural Heritage Labs community	23
Keyword searchable digital library of Serbian historic newspapers: Building the national infrastructure	24
The role of CLARIN and linguistic annotation in the digital transformation in research	25

Digital Initiatives at the Taylor Institution Library	27
Bringing coroner's records to life: a collaboration between the Glucksman Library and faculty to teach digital humanities to fourth-year history students at the University of Limerick	28
Las bibliotecas universitarias, detonante para las Humanidades Digitales dentro y fuera Universidad Autónoma de Querétaro	29
DH Clinics: Dutch librarians' tour of DH	30
<b>Workshop.</b>	<b>31</b>
Social Networks and Archival Context Cooperative (SNAC) and the Role of Cultural Heritage Members (90 minutes)	31

Last Edited July 5, 2019

# Information and Acknowledgements.

Libraries as Research Partners in Digital Humanities will be held on July 8, 2019, in Den Haag (The Hague, Netherlands), as a preconference to the 2019 Digital Humanities Conference in Utrecht. The event is hosted by the KB National Library of the Netherlands, with sponsorship from the KB, Consortium of European Research Libraries (CERL), and the Libraries and Digital Humanities Special Interest Group of the Alliance of Digital Humanities Organizations (ADHO).

This pre-conference aims to address the complexities of positioning libraries and other Cultural Heritage Organisations as digital humanities research partners. The question of CHOs and librarians as service providers or partners is a well-established, rich territory, with implications for labour and status. However, many are still finding their place in digital humanities due to questions of institutional commitments and external expectations. The potential for CHOs and their staff to serve as research partners is demonstrated here in keynotes, short papers, lightning talks, and interactive sessions on collections as data, library labs, technical infrastructure, organizational commitments and policies, staff-building, labour, and relationship-building.

Slides and other documentation of the event will be posted to and linked from our [Zenodo community site](#).

## Local organizers

Lotte Wilms, KB National Library of the Netherlands<sup>?</sup>

Martine Klaassen, KB National Library of the Netherlands

Steven Claeysens, KB National Library of the Netherlands<sup>?</sup>

Marian Lefferts, Consortium of European Research Libraries<sup>?</sup>

## Programme committee

Hege Hosoi, National Library of Norway

Sarah Potvin, Texas A&M University

Isabel Galina, Universidad Nacional Autónoma de México<sup>?</sup>

Silvia Gutiérrez, El Colegio de México

Andreas Degkwitz, Humboldt University of Berlin

Tamara Butigan, National Library of Serbia

Dawn Childress, University of California, Los Angeles

Steven Claeysens, KB National Library of the Netherlands<sup>?</sup>

Marian Lefferts, Consortium of European Research Libraries<sup>?</sup>

Lotte Wilms, KB National Library of the Netherlands<sup>?</sup>

## Volunteer reviewers

*These individuals generously joined programme committee members in contributing reviews of proposals*

Demmy Verbeke, KU Leuven

Liam O'Dwyer, Dublin City University

Caleb Derven, University of Limerick

# Keynotes.

## Collections, Data, Research: perspectives at the national library of France

Emmanuelle Bermès (Bibliothèque nationale de France)

From 2016 to 2019, the french national library (BnF) was engaged in a project called Corpus, aimed at exploring new services for researchers using digital collections as data. Starting from the idea of delivering big datasets of digitized material, web archives and metadata for text and data mining research projects, the Corpus initiative resulted in the definition of a wide array of services, including a physical space. This presentation will be the tale of this journey of librarians and academics through the vast land, still much to uncover, of digital humanities.

Bio:

Emmanuelle Bermès is deputy director for services and networks at BnF since 2014. From 2003 to 2011, she worked at the National library of France (BnF), first in digital libraries and digital preservation, then in metadata management. From 2011 to 2014, she was in charge of multimedia and digital services at the Centre Pompidou (Paris, France). In the course of her career, Emmanuelle has held a number of responsibilities at international level (within Europeana, the Library Linked Data W3C incubator group, the IFLA Semantic Web special interest group (SWSIG), the International internet preservation consortium (IIPC), and the International Image Interoperability Framework (IIIF) consortium). Within the BnF, Emmanuelle is leading the Corpus project, an initiative to create a digital scholarship service for researchers.

## Community before Computing: Reflections on collections as data

Sarah Potvin (Texas A&M University, United States)

How can collection stewards create and serve computationally-accessible collections as data? How can teachers, researchers, journalists, curators, and others who rely on collections as data find and access this data, and communicate their needs? How can collections as data move beyond large and well-resourced institutions and implementations? The Always Already Computational: Collections as Data project, which ran 2016-2018 with generous support from the Institute of Museum and Library Services' National Digital Platform grant program, sought to address these - and other - questions while documenting and sharing current and potential approaches to collections as data. Always Already Computational convened two national forums, organized multiple workshops, shared project outcomes in professional conferences, sought broad feedback,

and generated a framework comprising nearly a dozen publications meant to guide institutions. In this presentation, Always Already Computational project team member Sarah Potvin will situate and summarize the project's activities and framework, reflect on its critical emphasis on 'community' and 'computing,' and invite discussion of persistent questions for collections as data.

**Bio:**

Sarah Potvin is an Associate Professor in the Texas A&M University Libraries. As the Digital Scholarship Librarian, based in the Office of Scholarly Communication, she has experience coordinating a portfolio of digital humanities, digital collection, and library technology projects, with attention to platforms, standards, varieties of open access, and interoperability. At Texas A&M, she has contributed as a core project team member, with Amy Earhart, Rebecca Hankins, and Maura Ives, to the collaborative, interdisciplinary Digital Black Bibliographies (DiBB) project, which seeks to transform bibliographies into data and visualize the literary contributions of African Americans from the 1950s to the 1980s through the prism of two bibliographers' works, and to the Texas A&M node of the multi-institutional NSF-funded Open Explorations of Vertebrate Diversity in 3D project (OVert). Her current assignments emphasize a post-custodial approach to digital scholarship that extends the principles and practices of stewardship to collections outside of libraries' custody. Ms. Potvin recently concluded work as a project team member and co-Investigator on the IMLS-sponsored Always Already Computational: Collections as Data project (2016-2018). She is a founding editor of dh+lib and a founding convener of the Libraries and Digital Humanities Special Interest Group of the Alliance of Digital Humanities Organizations. Her research and service work is primarily focused on digital humanities, the politics and complexities of public and open access, and community engagement.

## Lightning Talks.

### Digital scholarship and data provenance at The National Library of Scotland

Sarah Ames (National Library of Scotland)

**Short abstract:**

This talk will consider the importance of data provenance and discuss how The National Library of Scotland is seeking to contextualise its data collections as it establishes a new Digital Scholarship Service.

**Long abstract:**

As the collections as data movement increasingly gains momentum, and libraries seek to present collections in machine readable form, the gap between our print collections and how our digital collections represent these increases. What, then, is the role of data provenance, and how can this be conveyed?

At The National Library of Scotland, an in-house, mass-digitisation programme runs 14 hours and produces 10,000 images each day. As we produce data at scale, we need to ensure that our readers understand where the data they use has come from, what processes it has undergone, and why it exists in the first place. How to convey that information in an area that currently has few agreed standards, and how libraries can remain trusted sources of information when we are creating that information ourselves, is problematic.

This lightning talk will consider the importance of data provenance and discuss how The National Library of Scotland is seeking to contextualise its data collections as it establishes a new Digital Scholarship Service. Using METS to convey the structural, bibliographical and technical metadata, as well as rights information, for each digital object enables readers to piece together how data has been produced; in addition to this, by including further information about the reasoning behind why an item has been digitised, we are aiming for a level of transparency between library and reader which enables a clearer understanding of the library processes print items undergo when creating collections as data.

This ties in to the new responsibilities of libraries as we produce our own collections: with digitisation programmes creating not only OCR'd text, but also additional data from scanners, software and digital production tools, transparency about our data creation processes become important. Enabling readers to contextualise the data creation process aligns libraries with the need for reproducible and authentic data, and ensures that we do not preclude some forms of research.

## Data stewardship on the map: A study of tasks and roles in Dutch research institutes

Melanie Imming (Het Landelijk Coördinatiepunt Research Data Management / Dutch National Contact Point Research Data Management)

### **Short abstract:**

The Dutch LCRDM has delivered a report: 'Data stewardship on the map' which provides a basis for clearer job descriptions in the different domains.

### **Long abstract:**

The Datastewardship Task Force of the The National Coordination Point Research Data Management in the Netherlands has recently delivered a report: 'Data stewardship on the map: A study of tasks and roles in Dutch research institutes'. The outcomes of the report are based on different methods, such as a large survey amongst librarians and other RDM specialists. It provides a basis for clearer job descriptions for data steward roles, and some guidance on the implementation of datastewardship.



## KEY FINDINGS

Key findings of the report are:

### AMBITION

Datastewardship is high on the agenda. Many institutions are currently working on professionalizing data stewardship but it is under construction everywhere. Often the ambition is to increase the number of data stewards in the near future; in particular the number of data stewards with domain data specific knowledge.

### RECOGNITION

Everyone who takes on data stewardship tasks should get the correct recognition, including accompanying career perspectives and appropriate training and visibility. Furthermore, RDM should be embedded in the organization and it requires targeted coordination and institution-wide policy.

### TASK AREAS INSTEAD OF ROLES

Although the job titles are very different everywhere, there are clearly identifiable data stewardship tasks. The task areas identified in this report clearly show the main focus of roles. However, if new job profiles are drawn up, the recommendation is to keep them broad. The focus on task areas instead of well-defined job profiles can help in compiling new customized job descriptions.

### POOL OF EXPERTS

There is a lot of enthusiasm in the field for setting up a pool of data stewards, each with their own expertise, as a way to offer the correct (domain-specific) support.

### TRAINING

Investing in good basic education for data stewards and training programs for continuous training is necessary for a sustainable set-up of data stewardship.

## Creating a Collaboration Guidebook for Library Faculty Partnership Success

Jennifer O'Brien Roper (University of Virginia Library, United States)

Kara McClurken (University of Virginia Library, United States)

Carla Lee (University of Virginia Library, United States)

### Short abstract:

The University of Virginia Library has created a framework intended to guide creation of positive partnerships with faculty for creation of innovative and sustainable projects.

**Long abstract:**

In order to frame collaborations with faculty for success, the University of Virginia Library has created a Collaboration Guidebook. While intended for use with all types of collaborations and partners, a particular use case that was a driver in the Guidebooks's development is teaming with faculty on digital scholarship projects. Born of necessity based on previous projects wherein the goals and expectations of each party were not clarified, this document is framed as a forward looking guide to creating positive partnerships with faculty for creation of innovative and sustainable projects. Building on the work of the University of Texas at Arlington, the Guidebook provides in-depth guidance and advice on shaping an idea into a project proposal, identification and inclusion of stakeholders across the Library who will be integral to success, and understanding the importance of articulating achievable goals with timelines and end dates. Also included are a workflow that with a robust review cycle and templates and instructions for writing formal memoranda of understanding (MOU) to document agreements. This lightning talk will provide an overview of the document structure, highlights of the primary concepts provided as pillars for success, and a summary of the process by which it was created.

## Project proposal for the training of Cuban librarians for Digital Humanities Research (Proyecto de formación de bibliotecarios cubanos para la investigación en Humanidades Digitales)

Sulema Rodríguez Roche (Universidad de La Habana, Cuba)

**Short abstract:**

A project for the training of Cuban Librarians for Digital Humanities research is presented. It will be developed by the HDCICuba Research Group, from the University of Havana.

**Long abstract:**

Se presenta un proyecto de formación de habilidades para la investigación en Humanidades Digitales en bibliotecarios cubanos que tiene como objetivo contribuir a la transformación digital y la innovación social en Cuba. Se caracteriza el contexto profesional de las bibliotecas cubanas para la investigación en Humanidades Digitales. Se examinan las convergencias entre la Bibliotecología y las Humanidades Digitales para identificar los núcleos de contenido y definir las formas de organización de la enseñanza. Se utilizan el análisis documental y la entrevista como métodos de investigación. Se propone un plan de proyecto basado en la metodología de marco lógico. Este proyecto forma parte de una investigación doctoral en Ciencias de la Información, que propone transversalizar las Humanidades Digitales en el campo de las Ciencias de la Información en Cuba. Ambas se realizan en el marco del del Grupo de investigación académica Humanidades Digitales y Ciencias de la Información (HDCICuba), proyecto institucional de la Facultad de Comunicación de la Universidad de La Habana.

# Building and Sustaining Fulcrum: A Digital Publishing Platform for Humanists and Social Scientists

Emma DiPasquale (Michigan Publishing, University of Michigan Library, United States)

## **Short abstract:**

This presentation will discuss the creation and development of Michigan Publishing's Fulcrum, a community-based publishing platform focused on presenting the best possible representations of work in the humanities and social sciences.

## **Long abstract:**

Michigan Publishing--the hub of scholarly publishing at the University of Michigan situated in the University Library--saw a need for a platform that addressed the needs of scholarly authors who wished to link source materials to book-length interpretations of them in an integrated way. This is an emerging need in media-rich humanistic fields such as digital humanities; from accumulating collections of film and video clips for comparison to visualizing excavation records through three-dimensional interactive models, humanists today often end their research projects with large archives of digital data. With support from the Mellon Foundation, Michigan Publishing and the University of Michigan Library developed Fulcrum, a community-based, open source digital publishing platform to help these scholars present these materials. Rather than forcing authors to strip away the richness of these materials at the point of formal publication to meet the constraints of a book or journal article format, Fulcrum facilitates rich integration of such digital objects with the narratives that reference them. They can either be published as supplemental materials, with the narrative (a book, a journal article) elsewhere, or both narrative and data can be presented together as an embedded EPUB. We have been expanding our development to include more expansive digital humanities projects, though we recognize the difficulty these projects can oftentimes present, especially in terms of preservation and accessibility. This is why we approach such projects through the lens of our values: accessibility, flexibility, durability, and discoverability. Our work with various press and publishing partners are helping us to better understand the needs of these scholars and how to develop the platform to support these kinds of works. Several of our forthcoming projects have prompted research and development for interactive maps, geospatial mapping, interactive EPUBs, 3D modeling, and annotation tools.

# Mnemosyne, Digital Library for Rare and Forgotten Texts (1868-1936): Collections and digital editions

Dolores Romero-López (Universidad Complutense, Spain)

José Luis Bueren Gómez-Acebo (Biblioteca Nacional de España, Spain)

Lucía Cotarelo Esteban (Universidad Complutense de Madrid, Spain)

**Short abstract:**

*Mnemosyne* is a digital library for Spanish rare and forgotten literary texts (1868-1936). Its objective is to model data to create the digital history.

**Long abstract:**

The objective of *Mnemosyne*, Digital Library for Rare and Forgotten Literary Texts (1868-1936) is to select, categorize, and make visible in digital format literary texts that belong to a forgotten repertoire in order to allow the historiographical review of the period. *Mnemosyne* has a repertoire of texts and authors who have remained in the shadow of the great literary figures of the first third of the 20th century. This digital library pretends to be a field of international experimentation for the creation of interoperable semantic networks through which a large group of scholars could generate innovative research and theoretical reading models for literary texts. It can be accessed through <http://repositorios.fdi.ucm.es/mnemosyne/>.

*Mnemosyne* aims to be an open-access digital library allowing data modelling for **specific collections** (intellectual women, Madrid in the Silver Age literature, children literature, kiosk literature, science fiction literature, personal writers collections, etc.) in support of research and teaching on Silver Age Spanish literature. Through *Mnemosyne* it can be accessed digital edition of texts. The digitization of these works has been carried out by public and private institutions. The first version of the library is stored on the server of the Universidad Complutense de Madrid Library, itself linked to the collections of the digital library *HathiTrust*. In a search conducted in *HathiTrust* in 2012, the names, surnames, and pseudonyms of selected authors were used to locate a total of 2,873 digitized texts corresponding to “odd and forgotten” writers. The *Biblioteca Digital Hispánica*, which serves as the access portal for the digital collections of the *Biblioteca Nacional de España*, provided 2,448 works by male authors and 1,017 works by female writers. Now have finished a new collection of **interactive editions** “La Edad de Plata interactive” with the support of Biblioteca Nacional de España. It can be accessed through <http://www.bne.es/es/Colecciones/LibrosInteractivos/Subcolecciones/edad-plata-interactivo.html>.

Behind the scenes of *Mnemosyne*’s public presence online, the project is developing with the aid of the tool *Clavy*. *Clavy* is an RIA (Rich Internet Application) that is able to import, preserve, and edit information from collections of digital objects so as to build bridges between digital repositories and create collections of enriched digital content. *Clavy* also provides a basic system of data visualization, edition, and navigation. *Clavy* facilitates the import, export, and edition of records in multiple formats such as MARC21, as well as their integration into *Mnemosyne*’s predesigned model with a view to their export into other compatible formats like XLS (Excel Binary File Format) or XML (Extensible Markup Language). Using *Clavy*, metadata from *HathiTrust* and *Biblioteca Digital Hispánica* have already been imported into the *Mnemosyne* database. The data from these sources was described in MARC21, following the rules for library catalogs. An example could be accessed through <http://clavy.fdi.ucm.es/Clavy/?CollectionId=13>.

*Mnemosyne* is the digital resource made with the support of three research teams (LEETHI <https://www.ucm.es/leethi>, LOEP <https://www.ucm.es/loep>, ILSA <http://ilsa.fdi.ucm>) affiliated with the Universidad Complutense de Madrid and funded by the public research projects

(FFI2012-34666, RedR+Human: TIN2014-52010-R, y CetrO+Spec: TIN2017-88092-R ), and by the private financial aid (S2015/HUM-3426).

## Marketplace Lightning Talks

### DARIAH: Core communities for collaboration: Libraries, Archives and Museums as essential partners in digital arts and humanities research

Sally Chambers (Ghent Centre for Digital Humanities, Belgium)

The mission of DARIAH, the Digital Research Infrastructure for the Arts and Humanities, is to empower research communities with digital methods to create, connect and share knowledge about culture and society. Digital methods are a cornerstone of what we do, ensuring we focus on how technology is transforming not objects, but activities. For us, the digital is not a goal in itself, but a means to explore, discover and grow. Libraries, Archives and Museums are essential partners in this endeavour. DARIAH is committed to continuing to encourage, facilitate and celebrate this collaboration at the core of what we do. In this presentation a number of examples of existing partnerships with cultural heritage communities will be introduced, as well as proposing one or two suggestions for future collaborations.

### CERL: Data for DH at Consortium of European Research Libraries

Marian Lefferts (Consortium of European Research Libraries)

The Consortium of European Research Libraries (CERL – [www.cerl.org](http://www.cerl.org)) hosts a large number of databases related to early printed books (1450-1850) and manuscripts that are useful for those that study the history and economics of printing, the spread of ideas, or the knowledge networks of a by-gone age. In this presentation, Marian Lefferts, CERL Executive Manager, will give an overview of what is on offer and how you can access the data.

### IFLA Digital Humanities - Digital Scholarship Special Interest Group

Xuemao Wang (University of Cincinnati, United States)

A global discussion forum to share knowledge, and promote best practices of Digital Scholarship/Digital Humanities among international library communities. For additional information, please visit: <https://www.ifla.org/dhds>.

# CLARIN & research libraries: a very short introduction

Leon Wessels (CLARIN ERIC / Utrecht University, Netherlands)

## Short abstract:

This presentation introduces the European Research Infrastructure CLARIN, its partnerships with research libraries and invites you to discuss potential future collaboration.

## Long abstract:

The Common Language Resources and Technology Infrastructure or CLARIN is a research infrastructure initiated to make digital language resources and tools from all over Europe and beyond accessible through a single sign-on online environment. CLARIN supports academic researchers as well as students, journalists and citizen-scientists interested in digital language resources, such as parliamentary records, social media data, newspaper archives and spoken corpora. In 2012, CLARIN received the status of European Research Infrastructure Consortium (ERIC) and as of 2016 CLARIN is recognized as an ESFRI Landmark. [1]

Research libraries have always played a vital role in CLARIN. National libraries actively participate in several of the 25 national CLARIN consortia. Clearly, libraries are some of the largest providers of digital resources used by linguists, historians, literary scholars and other researchers engaged in the (digital) social sciences and humanities. In addition, many libraries encourage digital scholarship by offering analytical tools, training and data management support. [2]

On a supranational level, CLARIN has a number of partnerships with large cultural heritage infrastructures. In 2017, CLARIN signed a Memorandum of Understanding with LIBER, the Association of European Research Libraries. The goal of the Memorandum is to engage in and share information about policy and advocacy activities of mutual interests such as copyright regulations, identification of opportunities for joint capacity building and planning, and cooperation on user involvement in the design of services. [3]

CLARIN is also a partner in the Digital Service Infrastructure of Europeana, an EU platform for digital cultural heritage. As a result of this partnership, a selection of Europeana records can be found in CLARIN's main search engine, the Virtual Language Observatory. These records include newspapers from Austria, Finnish travel brochures and audio recordings of Romanian poetry. [4]

CLARIN ERIC continues to explore new ways to strengthen its relations with research libraries across the world and is very open to suggestions for further collaboration.

## References

[1.] Koenraad De Smedt, Franciska de Jong, Bente Maegaard, Darja Fišer, Dieter Van Uytvanck, 'Towards an Open Science Infrastructure for the Digital Humanities: The Case of CLARIN', Proceedings of the Digital Humanities in the Nordic Countries (DHN), 2018, pp. 139-151;

Franciska de Jong, 'CLARIN: Infrastructural Support for Impact through the Study of Language as Social and Cultural Data', in: Bente Maegaard and Riccardo Pozzo ed., *Stay Tuned to the Future. Impact of the Research Infrastructures for Social Sciences and Humanities*, Leo S. Olschki Editore, Florence 2019, pp. 121-129.

[2.] John W. White and Heather Gilbert ed., *Laying the Foundation: Digital Humanities in Academic Libraries*, vol. 7, Purdue University Press, West Lafayette, Indiana, 2016.

[3.] Memorandum of Understanding between LIBER (Ligue des Bibliothèques Européennes de Recherche) and CLARIN ERIC, <https://office.clarin.eu/v/CE-2017-0997-MoU-with-LIBER.pdf>.

[4.] Twan Goosen and Nuno Freire, 'Bridging the Europeana and CLARIN infrastructures: Cultural heritage data for the social sciences and humanities', CLARIN.EU [weblog], 13 September 2017, <https://www.clarin.eu/blog/bridging-europeana-and-clarin-infrastructures>.

## Short Papers.

### The case for a citation index by the humanities, for the humanities

Giovanni Colavizza (University of Amsterdam, Netherlands)

Silvio Peroni (University of Bologna, Italy)

Matteo Romanello (École polytechnique fédérale de Lausanne (EPFL), Switzerland)

#### **Short abstract:**

We propose and exemplify motivations and design requirements for a humanities citation index. We suggest ways to make the index collaborative, distributed and open.

#### **Long abstract:**

Citation indexes are by now part of the research infrastructure in use by most scientists: a necessary tool in order to cope with increasing amounts of scientific literature being published. Commercial citation indexes are designed for the sciences and have uneven coverage and unsatisfactory characteristics for scholars in the humanities, while no broad-coverage citation index is proposed by a public organization. We argue that a citation index for the humanities is desirable, for four reasons: it would greatly improve the retrieval of sources, it would offer a way to interlink collections across repositories (such as archives and libraries), it would foster the adoption of metadata standards and best practices by all relevant actors (including libraries and publishers) and it would contribute data to fields such as bibliometrics and science studies. We also suggest that this citation index must be informed by a set of requirements relevant to the humanities, including comprehensive source coverage and substantial chronological depth. To substantiate our proposal, we showcase two infrastructure components of the envisaged humanities citation index: a prototype online application including a digital library and citation index for the historiography on Venice (named Venice Scholar), and two open linked data citation repositories made available by OpenCitations (i.e. the OpenCitations Corpus and the Crossref corpus COCI). The Venice Scholar digital library allows operators to curate machine-extracted

citation data, which will then be federated by OpenCitations into its persistent triplestores of citation data. These citation data can be pulled and served via the Venice Scholar citation index: an open online application. We conclude by proposing that the humanities citation index should be created taking a collaborative, distributed and open effort, as made possible by the proposed infrastructure: taking inspiration from library catalogs and in collaboration with GLAM institutions, primarily libraries.

## Digitizing for impact: from pixels to scholarship

Eric Shows (New York Public Library, United States)

Sara Rubinow (New York Public Library, United States)

Matt Knutzen (New York Public Library, United States)

### **Short abstract:**

We'll explore how transforming collections into a digital form can multiply research impact and how partnering with research communities can influence the library's Digital Collections strategy.

### **Long abstract:**

To further increase the discoverability, access, and use of the New York Public Library's collections, digitization offers our patrons manifold possibilities for research and a deep dive into the Library's encyclopedic holdings. Over the past decade, we have established clear rationales, processes, and workflows for digitizing research collections at the New York Public Library and are dedicated to building partnerships to explore how transforming collections into a digital form can multiply research impact.

In this presentation, we'll take a closer look at the complete digitization life cycle for a single collection, from direct patron request and conversations with curators to online publishing, museum exhibition, and structured data extraction. To surface this largely invisible work, we'll discuss the roles that the Library's Digital Imaging Unit, Metadata Services Unit, Registrar's Office, and Rights teams play in that life cycle. Along the way, we'll look at the opportunities for scholarship that arise when research materials have been digitized and become available online through our digital portals.

We'll discuss recent successes that aid genealogical research and our understanding of the urban landscape, experiments in computer vision to develop new ways of seeing and organizing items in a collection, and the latent potential in some recently digitized collections and corresponding research implications. We'll also discuss how feedback from researchers about the rising baseline expectations around collections as data has begun to influence how we prioritize digitization projects and how we think about partnering with the scholarly community.



# Next Generation of Scalable and Sustainable DH Services

Zach Coble (New York University Libraries, United States)

## **Short abstract:**

This talk will discuss the recent evolution and alignment of digital humanities support, services, and infrastructure at NYU, in particular infrastructure building collaborations between Libraries and IT.

## **Long abstract:**

This talk will discuss the recent evolution and alignment of digital humanities activities at NYU Libraries. While the Digital Scholarship Services (DSS) department is the public face of digital humanities at the library, our support, services, and infrastructure increasingly rely on collaborations between several departments across Libraries and IT.

For context, the scalable and sustainable approach to digital humanities in libraries, developed by DSS co-founders Jennifer Vinopal and Monica McCormick over seven years ago, has continued to guide our growth as we have adapted our approach to meet new needs. One of the core practices of this approach is to not embed ourselves in long-term projects that inevitably sink staff time and resources into maintaining projects that scholars gradually lose interest in. Instead, we provide training, services, and infrastructure that empower scholars to create, publish, and preserve their work. More recently, we have focused on empowering users by developing technical infrastructure that meets their unique needs, which is the result of taking a broad look and identifying similarities in the computational needs among DHers and scholars in other disciplines.

One aspect of this I will briefly touch on is our success in solving the “faculty website problem” using Reclaim Hosting’s Domain of One’s Own service. But perhaps more interesting is our work taking place under the umbrella of “Digital Repository Services for Researchers” (DRSR) This multi-phase project is a collaboration between several Libraries departments and the Research Technology group in Central IT to build and develop a robust suite of repository services that meet user’s complex needs at each stage of the research lifecycle, including discovery, access, publication, and preservation. For example, we have recently launched Research Workspace, a repository service providing fast, reliable, mountable storage for management of large-scale research data. Our Collections as Data work now lives on Research Workspace, and these collections can be easily mounted to a researcher’s laptop or NYU’s high performance computing cluster for analysis.

This talk will outline NYU Libraries’ current iteration of sustainable and scalable services, with a focus on the goals of our DRSR work. I will also provide use cases to describe in practical terms how NYU’s DH community has benefited from this work.

# Forschungspartnerschaften zu welchem Preis?

Jesko Reiling (Zentralbibliothek Zürich, Switzerland)

## Short abstract:

Der Beitrag erläutert die organisatorischen und finanziellen Aufwendungen, welche die Zentralbibliothek Zürich aufbringen muss(te), um sich (längerfristig) als Forschungsbibliothek etablieren zu können.

## Long abstract:

Die Zentralbibliothek Zürich (ZB) unterstützt seit einigen Jahren universitäre Forschungsprojekte und ist 2018 erstmals eine Forschungsk Kooperation mit einem Editionsprojekt eingegangen. Zusammen mit dem Editionsteam der Johann Caspar Lavater (1740-1801) Briefausgabe erarbeitet sie seither für die online-Briefedition ([lavater.com](http://lavater.com)) Metadaten und produziert zudem rund 110'000 Digitalisate der Korrespondenz. Dafür zahlt sie einen hohen Preis, der im Referat eingehender und aus verschiedenen Blickwinkeln zur Sprache kommen soll.

Der Beitrag illustriert anhand des Lavater-Pilotprojekts die Herausforderungen für eine Bibliothek, wenn sie sich als Forschungspartnerin aufstellen will. Vorgestellt werden die erforderlichen Ressourcen, die notwendigen organisatorischen Entwicklungsmassnahmen sowie betriebswirtschaftliche Mittel, die es braucht, um ein Einzelprojekt abzuwickeln und sich darüber hinaus längerfristig als Forschungspartnerin zu etablieren. Die Erfahrung zeigt, dass es eminent wichtig ist, die institutionellen Prozesse und Workflows ebenso umfassend wie detailliert zu identifizieren. So braucht es z.B. nicht nur Fachpersonal für die Erschliessung und das Scannen der Briefe, darüber hinaus mussten auch weitgehende restauratorische Massnahmen ergriffen werden, um die Digitalisierung des Materials überhaupt erst zu ermöglichen. Im Weiteren mussten auch technische Entwicklungen veranlasst werden, um der Edition den Datenimport auf die gewünschte Art und Weise zu ermöglichen.

Trotz Forschungsgelder durch den Schweizerischen Nationalfonds (SNF), dem nationalen Förderer universitärer Forschung, muss die ZB den Grossteil der Kosten selber bestreiten. Das Referat zeigt die enorme Kluft zwischen den gesprochenen und den benötigten Geldern detailliert auf. Hier stellt sich diejenige Frage, die wohl alle Bibliotheken beschäftigt: Bis zu welchem Preis kann man als Bibliothek (längerfristig) Forschungspartnerschaften eingehen?

Dabei gilt es freilich auch die Vorteile von Forschungsk Kooperationen zu bedenken: durch Forschungsgelder eröffnen sich GLAM-Institutionen zusätzliche Möglichkeiten der Erschliessung: Es können weitere, bislang unerschlossene Bestände aufgearbeitet werden oder man kann einen Bestand noch genauer und vertiefter erschliessen. Als Forschungspartnerin kommt man als Bibliothek dem eigenen institutionellen Sammlungs-, Erschliessungs- und Vermittlungsauftrag nach. Zudem trägt jedes Forschungsprojekt dazu bei, den eigenen Bestand sichtbarer zu machen. Das kann schlussendlich zu weiteren Folgeprojekten führen.

Die ZB, davon soll abschliessend die Rede sein, entwickelt(e) deshalb (erstmal in ihrer Geschichte) eine Forschungsstrategie. Ein Teil davon ist es, zusätzliches (geisteswissenschaftlich und bibliothekarisch geschultes, aber auch technisch ausgebildetes) Fachpersonal einzustellen. Diese Mitarbeitende sollen nicht mehr, wie bis anhin, lediglich in einer Fachabteilung (Handschriften, Alte Drucke, Graphik oder Karten etc.) tätig sein, sondern arbeiten als 'Universalisten' fachbereichsübergreifend in zukünftigen Forschungsprojekten und -kooperationen mit. Mit diesen neuen Strukturen will sich die ZB als zuverlässige und effiziente Forschungspartnerin profilieren.

## Collaboration with the Researchers – Content, Presentation systems and Techniques

Tuula Pääkkönen (National Library of Finland)

Jussi-Pekka Hakkarainen (National Library of Finland)

### **Short abstract:**

This talk tells about the researcher services evolution in National Library of Finland (data clinics, data catalogue and presentation system features), now and in the future.

### **Long abstract:**

In the National Library of Finland, there is constant attempt to facilitate discussion with researchers in many ways. In the annual summit of the Helsinki Centre of Digital Humanities the opening words of the new approach towards the researcher user segment (Lilja and Hakkarainen 2018), was a welcomed step by the researcher community. The communication approach of the National Library vary in size. There are 1-on-1 meetings, small group events in so called "morning coffee" sessions where researchers can present their work and how they utilize the collections. Then recently there has been data-clinics, where in an on-site workshops, where the programmatic ways how to utilize the data from the collections are explained. In this talk, we approach the research questions, what kind of documentation different researchers would need from the content, presentation systems and techniques? Are the end-user documentation (screen casts, basic user guides) enough also for the researchers, or is there need for more?

Recently we created an initial version of the instructions on how to utilize the collections on their own, i.e. self-help instructions on how to download desired parts of the collections or metadata. This is a local experiment bit similar to the Programming Historian (PH 2019). Our focus has been to answer to the most common questions of the researchers and enable them to help them to help themselves. While doing this toolset we realized, that the contextual bubble within the library differs from that of the researchers of themselves. We need to benchmark good practices to overcome the possible technological barriers researchers might face (Zhang, Liu, and Mathews 2015) and evaluate the strategy of the future.

Discussion with the researchers in the data-clinics has turned out to be very effective. The data-clinics, as a face-to-face meeting, lowers the communication threshold. We need to learn new ways for the cross-institutional collaboration (Nyhan and Fay 2015) and embrace the collaborative nature of the digital humanities (Wong 2016) in different lifecycle phases of the library service creation. We need to investigate further the expectations for the National Library and towards our contact points in other organizations – together we could weave a network where each organization can provide some aspect to bring a more coherent whole. In the long run, this helps the researches, which helps in the increasing the societal impact of the collections.

## References

- Lilja, Johanna, and Jussi-Pekka Hakkarainen. 2018. 'Re-Defining Our Services - National Library's New Initiatives to Support Open Science'. In Helsinki.
- Nyhan, Julianne, and Ed Fay. 2015. 'Webbs on the Web: Libraries, Digital Humanities and Collaboration'. *Library Review* 64(1/2): 118–34.
- PH. 2019. 'The Programming Historian'. *Programming Historian*.  
<https://programminghistorian.org/> (April 20, 2019).
- Wong, Shun Han Rebekah. 2016. 'Digital Humanities: What Can Libraries Offer?' *portal: Libraries and the Academy* 16(4): 669–90.
- Zhang, Ying, Shu Liu, and Emilee Mathews. 2015. 'Convergence of Digital Humanities and Digital Libraries'. *Library Management; Bradford* 36(4/5): 362–77.

## Partnering up with researchers in a national library

Martijn Kleppe (KB National Library of the Netherlands)  
Steven Claeysens (KB National Library of the Netherlands)  
Sara Veldhoen (KB National Library of the Netherlands)  
Lotte Wilms (KB National Library of the Netherlands)

### Short abstract:

This paper presents how the KB collaborates with researchers in various forms to better serve their needs but also to improve our own services.

### Long abstract:

The KB National Library of the Netherlands has been actively engaged with digital humanities through collaborations with academia. This paper present how the KB is partnering with the research community on digital humanities projects using the more than 100 million digitised pages of printed text. Since our digitisation started in the 1990s our focus shifted from the physical to the digital collection. We actively search for partners in the research community to use and study our digitised materials, with support from research software engineers in the digital humanities team of the KB.

Of course, collaboration is easier when access to the data is easy. This is why in 2013 a data services coordinator was installed in the library. Providing researchers with a single contact person to get access to the APIs and how to work with them facilitates use of the collections.

We aim to collaborate with researchers in different stages in their career. Since 2004, we have had the KB/NIAS Fellowship ([kb.nl/en/organisation/kb-fellowship](http://kb.nl/en/organisation/kb-fellowship)). This fellowship programme went from a programme for foreign researchers focused on the physical collections (2004-2013) to a national fellowship for professors working in the digital humanities (2014-2018) who were invited to spend a research sabbatical in the library for 5 months.

A similar programme was launched in 2014, aimed at early-career researchers: the funded researcher-in-residence programme ([kb.nl/en/organisation/research-expertise/researcher-in-residence](http://kb.nl/en/organisation/research-expertise/researcher-in-residence)), which works with a yearly call for proposals. Two researchers are selected by a committee of scholars from various disciplines and two library representatives. They spend 6 months working part-time in the Research department, together with the DH team. All output is published on the KB Lab ([lab.kb.nl](http://lab.kb.nl)).

Through the years we have built up a network of researchers through these forms of collaboration. We found that they are often looking for real world cases for their students to work on, as theses or research projects. We regularly have interns in the Research department or provide data and cases to students working on a project assignment.

Another type of collaboration that we have pursued recently aims to obtain input from several researchers at a time for a specific case that we formulate. For instance in the ICT with Industry workshop ([ict-research.nl/ict-with-industry/ictwi2019/#case1](http://ict-research.nl/ict-with-industry/ictwi2019/#case1)), organised by the ICT-research Platform Netherlands, where a team of researchers spend a full week working on a problem put before them by industry. We are also trying out this approach by providing data to benchmarking contests, such as DMAS ([primaresearch.org/DMAS2019/](http://primaresearch.org/DMAS2019/)). Asking for input from researchers to help solve issues we have in the library means we get a broad view of possible solutions while also contributing to the scholarly community with research topics with practical applications.

This paper will present the various forms of collaboration described above, and will go into more detail about lessons learned and recommendations for collaboration with researchers as a library.

## Finding new environments for cooperation with digital research and new values for heritage collections: experiences at the National Library of Spain

Elena Sánchez Nogales (National Library of Spain)

**Short abstract:**

Reusable open data, crowdsourcing... and more. Sharing the strategies developed at the National Library of Spain to strengthen, support and multiply connections with digital research.

**Long abstract:**

The National Library of Spain (BNE, Biblioteca Nacional de España), preserving the most important bibliographic and documentary collection in the country, has historically had a close and deep relationship with the research community.

Digitisation of collections, initiated in a massive way in 2008, changed the nature of this relationship, and while expanding and improving access to resources, it also implied the new risk of losing a direct ('physical') connection with researchers and their needs, and losing visibility of how the Library resources were used.

Over the last years, the BNE has worked to strengthen, support and multiply connections with digital research. Calling for proposals of new digitisations and actively cooperating in digital projects was a first step, but the Library continued beyond by setting a specific strategy to promote the use and reuse of data and digital resources within the framework of their BNElab, (<https://bnelab.bne.es/en/>). It comprises different lines of action 2018-2022, in cooperation with a community of experts in digital transformation and of course with researchers.

One of these lines looks at collections as data, and here catalogues (bibliographic and authorities) and other data produced by the Library, have been converted into open reusable formats and published in a dedicated space <https://bnelab.bne.es/en/data/> - as well as in official open data catalogues.

Helped by experts in data and digital humanities, the Library is promoting the use of these new resources. And the process is also having an interesting internal dimension: 'questioning' these data files from a researcher's view has enabled cataloguing teams and digital collections managers to learn about new uses for our catalogues, and also be aware about digital researchers needs, problems posed by our data and enrichment possibilities. By using tools such as Open Refine, considering new external sources for integration or learning about visualization tools, this work has indeed initiated a change in the way we look at our data and cataloguing practices, and at our catalogue as an information tool itself (beyond description and discovery).

But the Library is also offering new tools and digital platforms such as the 'Interactive books' project or ComunidadBNE (<https://comunidad.bne.es/>), a crowdsourcing initiative recently launched and very well received by the general public and also by academic and research groups. ComunidadBNE is conceived as a digital environment of collaboration between general public, specialists and librarians, for cooperative transcription, georeferencing, identification or tagging of images from the Library's digital resources, and also for the enrichment of the bibliographic catalogue and authorities. The platform was built on an open source base code, Pybossa – adapted, developed, and also made open and licenced for free reuse in GitHub.

Since the initiative was presented, different research groups have already confirmed interest in participating and using the platform for their academic projects.

This paper proposal aims at sharing the strategy, experience and initial results, as well as future projects within these new lines of action at the BNE.

## Growing an international Cultural Heritage Labs community

Sally Chambers (Ghent Centre for Digital Humanities, Belgium)

Mahendra Mahey (British Library Labs, British Library, London, United Kingdom)

Katrine Gasser (Royal Danish Library, Denmark)

Milena Dobрева-McPherson (UCL Qatar, Doha, Qatar)

Kristy Kokegei (History Trust of South Australia, Adelaide, Australia)

Abigail Potter (Library of Congress, Washington D.C., United States)

Meghan Ferriter (Library of Congress, Washington D.C., United States)

Rania Osman (Bibliotheca Alexandrina, Alexandria, Egypt)

### **Short abstract:**

Cultural Heritage Labs help people to innovate and experiment with digital collections. We will present our growing 'Labs' international community and outline our future activities.

### **Long abstract:**

'Cultural Heritage Labs' in galleries, libraries, archives and museums around the world help researchers, artists, entrepreneurs, educators and innovators to work on, experiment, incubate and develop their ideas of working with digital content through competitions, awards, projects, exhibitions and other engagement activities. They do this by providing services and infrastructure to enable, facilitate and give access to their data both openly online and onsite for research, inspiration and enjoyment.

In September 2018, the British Library Labs team organised a 'Building Library Labs' international workshop. The event provided the opportunity for colleagues that are planning or already have digital experimental 'Labs' to share knowledge, experiences and lessons learned. The workshop, which attracted over 40 institutions from North America, Europe, Middle East, Asia and Africa, demonstrated a clear need and enthusiasm for establishing an international support network. Within 6 months, a second international workshop was organised at the Royal Danish Library in Copenhagen in March 2019. In total we have brought together some 120 participants and an even wider community of around 250 people online. Some have been sharing their experiences in setting, using and running innovation labs, but there was a sizeable group of attendees who are planning to set up such labs and need advice and support in how to do this.

The aim of this short paper is to present the journey and development of the International Labs community and outline our future activities.

The principle of the network is that by fair sharing and 'paying forward' our expertise, knowledge and experiences, the group hopes to ensure that organisations don't have to 're-invent the wheel'. Organisations can learn from each other and enable collaboration across borders through their digital collections, data, services, infrastructure and practice. This we hope this will result in building better digital 'Labs' for their organisations and their users and help to further open up data and services for everyone.

People are the essence of the international Labs network. From the results of an initial global Building Library Labs survey, including 40 responses from 23 countries, there was significant interest from the wider cultural heritage sector, beyond libraries. With currently 250 people, from over 60 institutions, based in over 30 countries affiliated with the network, a solid set of communication tools were needed. The network has a shared Google drive, a mailing list and a Wiki, as well as an active WhatsApp group, a Slack channel and meets regularly via Zoom.

With two successful events behind us, and plenty of enthusiasm and willingness to continue activities further, we now looking to the future. Planned activities include: a booksprint to capture significant knowledge and expertise within the Labs network serving as a reference guide for people wanting to build their own lab, populating our wiki and creating a global directory of Cultural Heritage labs. Further regional and international events have and are also being organised. In less than a year, the Labs network has come a long way, and this is only the beginning!

## Keyword searchable digital library of Serbian historic newspapers: Building the national infrastructure

Adam Sofronijevic (University of Belgrade library)

Tamara Šević (Ministry of Culture and Media of Republic of Serbia, University of Belgrade)

### **Short abstract:**

The paper presents the process of building the national infrastructure that allows for keyword searchable digital library of Serbian historic newspapers.

### **Long abstract:**

The paper presents the process of building the national infrastructure that allows for keyword searchable digital library of Serbian historic newspapers. The infrastructure building encompass implementation of technical infrastructure, defining and setting up the working processes and drafting and making legal documents official. The process started in 2013 when University library Belgrade and National library of Serbia took part in Europeana newspaper project that allowed for first METS-ALTO files of Serbian historic newspapers to be created at project partner institution - University of Innsbruck. The first step in building technical infrastructure was creation of keyword searchable digital library at University library Belgrade that provides search and display of digitized historical newspapers and other documents. The website



<http://www.unilib.rs/istorijske-novine/browse> offers simple and advanced search interfaces, and an interface for browsing through the digitized works. The website relies on the BnLViewer software by the National Library of Luxembourg to display digitized METS-ALTO documents, while the search backend uses Lucene/Solr search engine with enhancements made by the University library Belgrade. The website infrastructure is presented in the paper in some details. Collection descriptions and other texts of the website itself are published under the Creative Commons Attribution (CC-BY) license, to facilitate their reuse in other projects. In 2015 a proprietary software has been acquired for creation of METS-ALTO files - docWorks by CCS Hamburg. University library Belgrade and National library of Serbia got one Basic license each and few years the second biggest library in Serbia Matica sprska Library acquired one license as well. The paper presents in some detail the process of creation of compatible work processes in these three libraries accomplished through project and other special activities. Activities that provided for training of librarians and volunteers to allow them to work with this software is also depicted. Collaboration with different smaller libraries from Serbia and the region that provided jpg images to be refined with docWorks is described along with defining the role of libraries that are content providers and promoters of keyword searchable digital library infrastructure. Most recent steps in building technical infrastructure such as additional design and branding of the interface so that it is more functional and branded as central point of access to historic newspapers in Serbia supported by six big libraries and hosted in government data center are explained. The paper specially present the process of drafting and making legal documents regarding keyword searchable digital library of Serbian historic newspapers official. In November of 2018 document entitled National framework for digitization of periodicals has been signed by four big libraries and ministry of culture and information of Republic of Serbia. This document establish METS-ALTO as recommended format for digitizing newspapers and other periodicals in Republic of Serbia. The legal documents drafted in 2019 that allow for implementation of the National framework are also presented. Paper also presents aspects of the process of building this national infrastructure that encompass collaboration with private companies, media, publishers and other stakeholders.

## The role of CLARIN and linguistic annotation in the digital transformation in research

Martin Wynne (Bodleian Libraries, University of Oxford, United Kingdom)

### **Short abstract:**

The CLARIN research infrastructure supports the linguistic annotation of digital texts, which allows smarter searching, and can make new forms of research possible.

### **Long abstract:**

Large and growing quantities of textual materials are being made available by libraries in digital form. Many library users will find that this is a convenient way to discover and read texts, but digital texts can offer much more than a shortcut to the delivery of texts to readers. Digital texts

can be searched, clusters and patterns of words can be identified, and changes in language usage over time can be traced. Corpus linguistics has developed advanced techniques for the storing, annotating, searching, exploring, and visualizing results. Resources and techniques which have been developed for the purposes of addressing linguistic research questions can also be used to explore and analyse the content of texts, by researchers in disciplines across the human and social sciences. While these opportunities are currently being exploited in many projects, there is still not yet a concerted effort to make all digital texts available in suitable formats to facilitate this sort of digital research. What more do digital libraries need to do in order to make texts available in ways that can really support advanced digital scholarship?

This question has been addressed via surveys of existing practice via the networks of the CLARIN research infrastructure, and via experimental services offered by projects within the Bodleian Libraries at the University of Oxford. The conclusion has been drawn that minimum levels of annotation are necessary for text collections in digital libraries.

Annotation of texts should include structural markup, metadata, and linguistic annotation, including:

- Standardized metadata for basic categories such as language, relevant dates, author, title and text type;
- Part-of-speech tagging;
- Lemmatization; and
- Modernized (or otherwise normalized) forms, where historical or other forms of variation are likely to be a barrier to effective search.

When these annotations are made searchable, it can enable the following:

- Enhanced search, for inflected and variant forms of words;
- Cross-searching the content of numerous texts, with results refined on the basis of content and metadata values;
- Visualization of results;
- 'Distant reading', data mining and extraction of results across large numbers of texts;
- Limited access via search and exploration interfaces to restricted texts, which would otherwise not be available to readers;
- Application of further levels of annotation and analysis by users based on the basic levels of annotation;
- Linking to resources held in other repositories, allowing cross-repository resource discovery and content search.

How can it be achieved? The preliminary results will be presented of a CLARIN initiative to gather information on NLP for the annotation of historical documents, and efforts to make it easier for resource curators to use the annotation tools.

Some challenges remain. Digital libraries are not necessarily centres of expertise in NLP, and some new initiatives and collaborations are often necessary. The repository will need to develop and

maintain interfaces, which need development and maintenance, and to offer user support. The suggested annotations are likely to be considered provisional, and subject to ongoing updates, improvement and enhancement.

## Digital Initiatives at the Taylor Institution Library

Emma Huber (Bodleian Libraries, University of Oxford, United Kingdom)

Frank Egerton (Bodleian Libraries, University of Oxford, United Kingdom)

### **Short abstract:**

The Taylorian uses Digital Humanities teaching initiatives in order to lower barriers to engaging in digital methods and raise the profile of librarians as skilled DH advisors.

### **Long abstract:**

The Taylor Institution has created a number of innovative Digital Humanities teaching initiatives over the past year, aimed primarily at graduate students and early career researchers, in order to lower barriers to engaging in digital methods. These initiatives include a popular Digital Editions course (running for the last three terms at full capacity), a Humanities Coding Club, and an Introduction to Digital Humanities course (see

<https://www.bodleian.ox.ac.uk/taylor/about/courses-and-training> - the last two are being run for the first time from April 2019). These courses give students the skills they need to use Digital Humanities tools for their own research without requiring a research grant.

The courses are written, planned and run by Subject Librarians at the Taylor Institution Library, one of the Bodleian Libraries at the University of Oxford, and the largest library dedicated to Modern Foreign Languages in the UK. Librarians are ideally placed to teach Digital Humanities, as they combine subject knowledge, expert information management skills and a wide knowledge of other support services. There is currently a gap in the market for introductory level courses for students with an interest in DH, but who have not yet decided to pursue a DH-specific course. Many existing training opportunities are for people who have already secured funding for DH projects, or are showcasing cutting-edge projects, which can be intimidating for beginners. Our interest is in providing students without research grants the opportunity to use digital methods, for free, to enhance their existing studies, and to inspire them to pursue further study in Digital Humanities.

In the Digital Editions course, librarians offer introductory training on text selection, imaging, transcription, xml encoding using TEI (two sessions), quality assurance, preservation and delivery. The aim is to give participants an overview of the main elements of a digitisation project, raising awareness of the issues to be considered (and budgeted for) in funded digitisation projects, and where to seek advice. "Library" skills such as research data management, open access, use of information databases and preservation are taught alongside Digital Humanities skills. The Humanities Coding Club is a weekly peer-support group for anyone teaching themselves to code, as well as anyone with expertise to share. Library staff are offering their support and guidance, and may go on to provide further structured courses in specific areas if there is a clear demand. The Introduction to Digital Humanities course teaches basic principles of data modelling, with

practical hands-on activities leading to the creation of databases, visualisations based on statistical analysis, network graphs and geographic visualisations, as well as an introduction to linked data, using SPARQL and wikidata.

The students participating in these courses are the funded researchers of the future, and we feel that by building strong relationships at this early stage, and by teaching library priorities such as the use of standards and preservation, we are creating a strong foundation for future partnerships.

## Bringing coroner's records to life: a collaboration between the Glucksman Library and faculty to teach digital humanities to fourth-year history students at the University of Limerick

Kirsten Mulrennan (Glucksman Library, University of Limerick, Ireland)

Sinéad Keogh (Glucksman Library, University of Limerick, Ireland)

### **Short abstract:**

This paper examines how library staff used historical Coroner's Inquest records to develop a new model for teaching library and archive skills to undergraduate students.

### **Long abstract:**

The Glucksman Library runs a series of programmes to support research, teaching and learning at the University of Limerick (UL). The Special Collections and Archives department within the library utilises its unique and distinctive collections to enhance student experience, encouraging 'hands-on' learning for students of all types. As a pilot project in the Autumn Semester of 2018, archivist Dr Kirsten Mulrennan and librarian Sinéad Keogh were invited to co-teach a portion of a fourth-year medical history module run by the Department of History. This module was designed to do a number of things at once: educate the students about the importance of archival material; encourage critical thinking about the use of medical records for historical research; demonstrate the potential of digital humanities tools for the exploitation of linked data; and introduce the students to the basic concepts of metadata, text mark-up and the Text-Encoding Initiative (TEI). Working in groups with digitised historic records from the Irish Coroner's Court held in the National Archives of Ireland, the students produced verbatim archival transcriptions of five records, highlighting linked data elements they felt would be most beneficial for future research – elements such as names, places, organisations, dates, and causes of death. The output from the module demonstrated a high level of student engagement with the process, and feedback from both the faculty and students at the end of the module was incredibly favourable. As a result, this module design will be further developed for the coming academic year. Mulrennan and Keogh have used this experience to build their teaching profile within the university, as well as establish lasting links with the history department. Overall, this project illustrates the diverse ways in which academic libraries can work with faculty as partners in digital humanities to grow graduate information literacies and skills, creatively use technological tools to get both staff and students engaged with historic sources, and ultimately, that the increasing integration of rare books and

archival material into the curriculum allows for the continued development of unique academic programmes at UL.

## Las bibliotecas universitarias, detonante para las Humanidades Digitales dentro y fuera Universidad Autónoma de Querétaro

Carlos Alberto Martínez Hernández (Universidad Autónoma de Querétaro, Mexico)

Rubén Cantor Pérez (Universidad Autónoma de Querétaro, Mexico)

### **Short abstract:**

The Digital Humanities are the ideal field for librarians to participate, collaborate and cooperate in the generation of new knowledge, as well as being the link to create bridges for students, teachers, and librarians.

### **Long abstract:**

En este artículo tenemos la intención de mostrar nuestras experiencias dentro de la Universidad Autónoma de Querétaro, para proponer nuestra metodológica, que hemos desarrollado a lo largo de cuatro años, en donde nos hemos preocupado por planear e impartir a la comunidad universitaria talleres, cursos y conferencias enfocados en el desarrollo de habilidades informativas, Humanidades Digitales y competencias lectoras, cuyo objetivo de éstos es, principalmente, incidir en la formación universal de la comunidad universitaria. En otras palabras: la biblioteca universitaria ha sido el detonador y la precursora de las Humanidades Digitales en la Universidad Autónoma de Querétaro.

Los profesionales de la información deben ser capaces de derrumbar mitos y lugares comunes para poder seleccionar, evaluar y acercar a la sociedad conocimientos con certeza. Sin embargo, esto implica que los bibliotecarios tienen que trabajar en consolidar sus habilidades informacionales, sus saberes, competencias lectoras, entre otros conocimientos.

Nuestra propuesta se enfoca en el fortalecimiento de nuestro personal bibliotecario, sobre todo en lo relacionado a tres grandes rubros: Humanidades Digitales, desarrollo de habilidades informacionales, así como de competencias lectoras.

Desde nuestro espacio hemos creado y participado en tres cursos y un taller de humanidades digitales y nuestro principal objetivo es el siguiente: que nosotros, los bibliotecarios podamos pensar de manera distinta, es decir, permitiendo experimentar la forma en que investigamos y enseñamos, y laborar de manera conjunta con otras disciplinas.

Las humanidades digitales son un espacio natural para que los bibliotecarios participen, colaboren y cooperen en la generación de conocimiento. El rol del bibliotecario se está convirtiendo en una profesión con un carácter tanto docente como analista de datos, en la que los formatos digitales tienen como soporte más conocimiento e información progresivamente, que exigen la comprensión de herramientas y técnicas que utilizan, por ejemplo, las humanidades digitales. Los profesionales de la información tienen que convertirse en el vínculo colaborativo para conectar a los alumnos y docentes con el conocimiento en un entorno de investigación y enseñanza.

El papel de los profesionales de la información en el desarrollo y la promoción del contenido de las Humanidades Digitales para la investigación, la enseñanza y el aprendizaje en las bibliotecas académicas actuales es un área que aún no se ha explorado por completo. En este mismo orden de ideas, y para explorar este camino se necesita que los bibliotecarios sean personas alfabetizadas mediáticamente, informacionalmente y tecnológicamente. Estas iniciativas, junto con las ya mencionadas, se han creado e implementado en diversas áreas de la UAQ y se busca incidir en futuras reformas curriculares y en las habilidades y metodologías de los alumnos, profesores y trabajadores de nuestra Universidad.

## DH Clinics: Dutch librarians' tour of DH

Michiel Cock (Vrije Universiteit Amsterdam, Netherlands)

Ben Companjen (Leiden University Libraries, Netherlands)

Lotte Wilms (KB National Library of the Netherlands)

### **Short abstract:**

In this paper we will present the lessons learned from organizing a set of five full-day training sessions for librarians on Digital Humanities.

### **Long abstract:**

In the fall of 2017 we organized a set of five full-day training sessions on Digital Humanities (DH) for librarians and other cultural heritage organisation in the Netherlands, called the DH Clinics. After our initial poster presentation at DH2017 describing our design process, in this short paper we will address the final programme, its reach and reception, the lessons learned and the impact on DH at research libraries in the Netherlands.

### **The programme**

We had noticed that librarians at our institutions - and at research libraries in the Netherlands in general - who wanted to get involved in DH needed more training to adequately partner with researchers and students in this field. The DH Clinics were intended to facilitate this need for education in DH by organising training events, with the following goals:

teach academic librarians the basics of DH

strengthen relationships between libraries and academic community

build a knowledge network of Dutch librarians working in DH

potentially lightening the work of librarians by introducing them to new techniques

We aimed to create a programme that gave librarians a basis in DH from where they can identify remaining gaps in knowledge or skills that could addressed by self-directed learning, to support researchers and (possibly) to (semi-)automate daily library work.

Each DH Clinic consisted of two lectures in the morning ( by local researchers) and a hands-on workshop in the afternoon. Each day had its own theme, based on the Taxonomy of Digital Research Activities (TaDiRah), of which we selected the following themes; Capture, Creation,

Enrichment, Analysis (text and non-text), hereby covering subjects ranging from digitization to Natural Language Processing and computer vision.

Reach, reception, lessons learned and impact

In this paper we will discuss the following lessons learned;:

- In total, the DH Clinics attracted between 75 and 80 one-time and returning participants, reaching a large enough number of our intended audience.
- The audience consisted mainly of research librarians, but also drew participants from archives and museums.
- Respondents of the evaluation questionnaire were positive about the clinics matching their expectations.
- Participants indicated they had a need for more examples of what libraries actually do or can do to support DH; we organized a follow-up session in November 2018 that addressed this theme.
- The programme has broadened the interest in and knowledge of DH in research libraries in The Netherlands, resulting e.g. in further training sessions at individual libraries.
- The clinics have laid a good foundation for DH at the research libraries, but more active work is needed to further develop the expertise of librarians and the network of librarians involved in DH in The Netherlands.

With this paper we intend to share our experiences with organising the DH Clinics and present an example of how libraries can introduce training to their staff in a non-formal manner.

## Workshop.

### Social Networks and Archival Context Cooperative (SNAC) and the Role of Cultural Heritage Members (90 minutes)

Daniel Pitti (University of Virginia, United States)

Susan Perdue (University of Virginia, United States)

#### **Short abstract:**

The workshop will demonstrate SNAC (<https://snaccooperative.org/>) as a public research tool that facilitates access to descriptions of CPF entities and dispersed, related historical documents; and a demonstration of SNAC as an editing and maintenance platform. This should be of interest to curators and research users of cultural heritage (CH) resources.

#### **Long abstract:**

SNAC (<https://snaccooperative.org/>) is a cooperatively maintained resource for the description of corporate bodies, persons and families (CPF) interrelated to one another and to the historical

resources that provide evidence of their lives and work. As a cooperatively maintained resource, SNAC serves interrelated functions that benefit both curators and research users of cultural heritage (CH) resources. For curators, SNAC serves as a platform for sharing labor-intensive identity resolution and description of shared CPF entities. It also serves as a platform for explicitly connecting dispersed historical resources held in repositories around the world through the CPF entities. SNAC thus aspires to be a vast global social-document network research tool. SNAC provides research users with two major benefits: integrated access to dispersed CH resources related to CPF entities and access to the social-professional-intellectual networks of the CPF entities documented in the resources. As a cooperative description platform, SNAC has focused on serving CH professionals, but it also aspires to expand the professional editorial community to include digital humanists interested in being partners in its development and maintenance: a CH-DH cooperative.

The intended audience for this workshop are persons involved in processing and describing CH holdings to facilitate access and understanding; persons providing reference services to researchers; research users of CH historical resources or with an interest in historical social networks; and digital humanists with an interest in both researching historical social networks and collaborating with CH professionals in documenting the global social-document network.

The workshop will demonstrate SNAC as a public research tool that facilitates access to descriptions of CPF entities and dispersed, related historical documents; and a demonstration of SNAC as an editing and maintenance platform.

The public research tool demonstration will cover the following:

- Basic and advanced searching
- Exploring biographies and histories of CPF entities
- Exploring related historical resources through interactive lists of resource descriptions and interactive geographic display of holding institutions
- Exploring social networks of CPF entities through simple lists of related entities and network graphs
- Exploring links to descriptions of the same entity in other services such as VIAF, Wikidata/Wikipedia, various national and institutional authority files, FamilySearch, and others
- Searching other CH resource description and access systems: WorldCat, ArchivesGrid, DPLA, Europeana, Archives Portal Europe

The editing platform demonstration will briefly cover the following:

- Editor authentication and permissions overview
- Editing of name entries: interface, rules, and structure
- Biographical and historical description, dates of existence, related places
- Demographic description of persons



- Summary descriptions of historical resources linked to holding repository and link to local description (a detailed finding aid, for example)
- Linking CPF entities to digital historical resources, for example, an image of a letter, or a scholarly edited letter (encoded in TEI)
- Linking CPF entities to one another (social network)
- Citing sources down to the assertion level
- Exploring the detailed history of any CPF description: including who deleted, added, or revised an existing description, and when

Twenty minutes will be allowed for questions and discussion at the end.