A Toolkit for Archivists and Librarians Supporting Research and Teaching in Digital Humanities

Karen Sayers

Archivist, University of Leeds Special Collections
& RLUK/TNA Professional Fellow 2023-2024

April 2024

The Research Libraries (RLUK) and The National Archives (TNA) Professional Fellowship Scheme enables staff from both organisations to gain experience and insight from one another, strengthen and diversify the relationship between them, and to overcome some of the collective challenges facing research and cultural organisations.

This output was submitted as part of the successful completion of an RLUK/ TNA Professional Fellowship project conducted in 2023-2024 and was approved by the RLUK/TNA Professional Scheme Board.

More information about the scheme can be found on the RLUK website.

Contact the author:

Karen Sayers
Archivist
University of Leeds Special Collections

Email: k.sayers@leeds.ac.uk

Table of Contents

INTRODUCING DIGITAL HUMANITIES SUPPORT

The Aim of the Toolkit

This toolkit is a resource for archivists and librarians who support, or wish to support, digital humanities researchers and teachers and to become partners or collaborators in projects. In addition, it is of value to researchers, academic teachers and managers. The toolkit contains links to digital humanities projects, training resources and support networks.

During 2023-24, I undertook an RLUK/TNA Fellowship. My fellowship focused on digital humanities research using data from historical record series. To create this toolkit, I decided it was essential to understand the processes that a digital humanities researcher carries out during a project. I wanted to discover what skills were necessary, the training and support available, and how long a project might take.

I decided to carry out a mini research project 'Visualising the Medieval Manor'. I extracted data on legal cases involving women from the Wakefield Manor Court Rolls 1331-32 and used it to create visualisations. I also wrote a case study called 'My Experiences Carrying out a Mini Research Project to Learn the Skills Required to Support Researchers and Teachers in Digital Humanities'.

I have divided the toolkit into sections that outline the stages in a research project. It contains links to resources that provide support or training relating to the skills needed to carry out digital humanities research. The toolkit focuses on research skills that are especially relevant to digital humanities.

In the toolkit, I cite links to open access resources for supporting digital humanities researchers and teachers whenever possible. I have used or read most of them. It is important to remember that new resources are constantly becoming available. In the rest of the toolkit, I will use the term archivists to cover all information professionals.

Definitions

A historical record series is a sequence of records of the same type and origin. They have a relationship stemming from a shared history of creation, storage, or use. Records in the series will usually document the same activities or functions. People often created them as part of a shared filing system or process. Examples include court records, wage books and sales ledgers.

Data visualisation is the presentation of data in charts, graphs, and tables with the aim of making sense of the data and increasing our understanding and knowledge.

What is Digital Humanities Research?

Digital humanities research applies computer-based technology to the humanities. It is both methodological and interdisciplinary. Digital humanities researchers use digital tools to explore data derived from source material relating to subjects such as history, languages, literature, and music. Researchers mining historical record series employ computer-based technology to process and analyse large amounts of data. They often communicate some of their findings in graphs and charts generated by visualisation software.

Digital humanities research involves many different processes including identifying, gathering, and analysing data. These are some of the main ones:

- Exploring and describing data to understand its structure, patterns, and relationships.
- Organising, cleansing, and visualising data to create tables, charts, and graphs.
- Analysing data using statistical and/or machine learning techniques to understand any underlying patterns and relationships.
- Communicating the findings of the analysis clearly to an audience.

Resource

Berry, David. M. 2019. *What are the digital humanities?* 13 Feb. *British Academy.* [Online]. [Accessed 22 January 2024]. Available from: https://www.thebritishacademy.ac.uk/blog/what-are-digital-humanities/

DIGITAL HUMANITIES PROJECTS

The toolkit briefly examines some recent projects that have used digital humanities methods. These give an idea of some of the historical record series that researchers are employing in digital humanities and their outputs.

One of the most well-known projects based on a historical data set is the Proceedings of the Old Bailey. The main project work took place from 2000-05 with later enhancements. Researchers can search the records of 197,745 trials at the Old Bailey from 1674-1913. They can also access the API (Application Programming Interface) to download data sets to analyse. A webpage explains the methodology behind the project.

Enslaved: Peoples of the Historical Slave Trade is a discovery hub where people can search and find information from data sets and digital projects. A partnership began the initial project in 2011

and it is ongoing. The website gives access to linked data from several projects with visualisations. It also offers a publishing platform the 'Journal of Slavery' for related humanities data sets.

A project at the University of Leeds Re-archiving the Individual: British Army Officers, 1790 –1820 is creating a 'life archive' as a tool to understand people in the past. The project runs from 2022-24. Researchers have extracted data from army lists and recorded it in a spreadsheet. The project aims to create an easily searchable online database with downloadable data for analysis.

Staff at The National Archives undertook the project Computing Cholera with Wellcome Trust funding in 2020. It is different from the above projects in using a data set derived from the catalogue entries for 272 volumes of correspondence between the General Board of Health, Local Boards, sanitary officials, and members of the public from 1846-71. Thus, the catalogue records served as the record series. Researchers employed computers to make statistical calculations that placed the data into subject groups. The groups were then analysed.

SKILLS REQUIRED BY ARCHIVISTS TO SUPPORT RESEARCHERS AND TEACHERS IN DIGITAL HUMANITIES

The abilities and knowledge archivists need to support digital humanities researchers, academic teachers and students vary depending on the depth of their involvement. They include:

- Professional archival skills including recognising data sets suitable for research and teaching
- Gathering, cleansing, and structuring data
- Processing data/ability to process data sets in spreadsheets
- Knowledge of visualisation software
- Programming skills
- Understanding of quantitative methods
- Networking, communication, and collaborative skills

Many archivists are in contact with researchers, academic teachers, and students and have a good knowledge of the previous usage of resources. This makes them valuable networkers who can provide contacts for new projects. They can also facilitate links to users.

This toolkit does not include a survey of the current level of skills of archivists relating to digital humanities, but it was evident from conversations with archivists that levels vary considerably. This can be because of individual role descriptions and responsibilities, the length of time in the profession and opportunity related to employers' priorities and budgets.

Professional Training

Archivists' collections-based knowledge means they are well able to identify records containing historical record series for digital humanities research projects and learning programmes.

Subject matter can indicate historical data sets; for example, people have often recorded legal, financial, and medical information in a standard way. The format of the data carrier can be a good indication of a historical record series; for example, ledgers, wage and sales books, membership lists, tables, and index cards.

Archivists' professional skills enable them to provide practitioners with essential information about the context of archival material including:

- provenance
- appraisal
- processing
- arrangement
- cataloguing

Archivists' knowledge of history can provide valuable background information and context to data.

If there is complementary research material in their repository in addition to that requested by researchers, academic teachers and students or, in other institutions, archivists often know. They may be able to facilitate contact with other people who have used the material.

Archivists can advise about GDPR (General Data Protection Regulation) restrictions that apply to data about living people in the records and any relevant terms agreed with a donor or depositor. In addition, archivists can give preliminary advice about copyright restrictions.

Identifying Data in the Historical Record Series

The data researchers or students on a learning programme may collect varies depending on their topic and sources. Four of the most common types of data required are:

- personal names
- dates
- geographic names
- quantities

Quantities or measures are categories such as wages, costs, prices, and fines. They have a numeric value.

Researchers and students will also gather data specific to the type of record series they are exploring. In legal papers, researchers may seek data on the type of court case, outcome, and the category of punishment. In business records, they may extract data about job roles, apprenticeships, products, and places of residence. Records need to contain several diverse types of data and examples over a sufficient timespan.

Excellent quality data is the basis for robust research results that can increase knowledge by providing new insights into topics. Archivists can often advise on the quality of data in historical record series. Essential data characteristics are:

- Completeness. Is the data comprehensive? Are there large gaps either in the individual data entries or in the chronological period over which people recorded it?
- Accuracy. Is the data error-free?
- Consistency. Did people record the data regularly and in a uniform manner? Avoid data sets where there are many missing or incorrectly recorded values.
- Reliability. Who recorded the data? Is it biased? This does not mean researchers cannot use it, but they must acknowledge its biases.
- Relevance. Is the data relevant to the research question?
- Sufficiency. There needs to be enough data to draw conclusions from a variety of categories of data and/or over a long enough timespan.

To conform with guidelines for open research, researchers and students need to record their methodology for exploring data and make it available. Academic institutions often require researchers to deposit the methodology in a central database with their project outputs. The methodology will include why they decided to extract certain types of data, how they did it and the software tools and services used.

Resources

UK Research and Innovation. *Open Research.* [Online]. [Accessed 14 January 2023]. Available from: https://www.ukri.org/manage-your-award/good-research-resource-hub/open-research/#:~:text=sharing%20of%20research%20outputs%20must,value%20for%20money%20and%20transparency.

University of Sydney. *Finding Datasets*. [Online]. [Accessed 22 January 2024]. Available from: https://www.library.sydney.edu.au/research/finding-data.html

Extracting Data

One of the first issues researchers and students encounter when extracting data from records is how to do it. Typed data can often be scanned using OCR (Optical Character Recognition) techniques and transferred to a spreadsheet. Increasingly, researchers can scan handwritten data using HTR (Handwritten Text Recognition) which they train to recognise particular handwriting.

However, if the data contains abbreviations and non-standard punctuation it may be easier to input it to a spreadsheet manually as machine applications may not be able to separate the data into meaningful columns. This is particularly the case with medieval and early modern written documents because of the use of abbreviations and inconsistent spelling of names.

One option is outsourcing the gathering of data. It is important to consider who will extract the data so the costing can be included in a funding bid, if necessary, when planning a project. The Libraries, Reading Communities and Cultural Formation in the 18th Century Atlantic project employed volunteers to extract data. Commercial transcription services are also available.

Resources

Adobe. *What is OCR and why is OCR software important?* [Online]. [Accessed 22 January 2024]. Available from: https://www.adobe.com/acrobat/guides/what-is-ocr.html

Burnard, Lou. 2014. *What is the Text Encoding Initiative?* [Online]. [Accessed 22 January 2024]. Available from: https://books.openedition.org/oep/679?lang=en

Dunley, Richard. 2018. *Machines reading the archive: handwritten text recognition software*. 19 March. *The National Archives.* [Online]. [Accessed 22 January 2024]. https://blog.nationalarchives.gov.uk/machines-reading-the-archive-handwritten-text-recognition-software/

Library Carpentry. *Our Lessons, Introduction to Data for Archivists and Introduction to Working with Data (Regular Expressions)*. [Online]. [Accessed 22 January 2024]. Available from: https://librarycarpentry.org/lessons/ (need to be booked as part of a workshop)

Text Encoding Initiative. *Text Encoding Initiative*. [Online]. [Accessed 22 January 2024]. Available from: https://tei-c.org/

TownsWebArchiving. *Data Capture, Handwritten Text Recognition.* [Online]. [Accessed 22 January 2024]. Available from: https://www.townswebarchiving.com/datacapture/handwritten-text-recognition-htr/

READ-COOP. *Transkribus. Unlock Historical Documents with AI.* [Online]. [Accessed 22 January 2024]. Available from: https://readcoop.eu/transkribus/

Internet Archive. *Wakefield Manor Court Rolls Vol. 03.* [Online]. [Accessed 22 January 2024]. Available from: https://archive.org/details/YASWCR003/page/n1/mode/2up

Gathering, Structuring and Cleansing Data in Spreadsheets

It is essential that researchers and students ensure that the data they extract from historical record series is accurate and standardised. Archivists have a good understanding of the need for uniform, structured data from their cataloguing work. Many archivists have knowledge of the capabilities of spreadsheets from inputting cataloguing data to them for import into collections management systems.

The cleansing of the data is often the most time-consuming task in a digital humanities project. OpenRefine is a useful tool for cleansing it, particularly the facet option. Library Carpentry offers workshops to help library and information professionals automate tasks and teaches OpenRefine. Online tutorials are available on YouTube. Researchers and students can also use Excel to cleanse data.

The nodegoat webpages can be useful for individual research or teaching projects. The site allows people to build, filter and analyse their own data sets.

Resources

Alex the Analyst. 2023. *Cleaning Data in Excel. Excel Tutorial for Beginners.* [Online]. [Accessed 22 January 2024]. Available from: https://www.youtube.com/watch?v=_jmiEGZ6PIY

Alex the Analyst. 2023. *Full Project in Excel, Excel Tutorials for Beginners.* [Online]. [Accessed 22 January 2024]. Available from: https://www.youtube.com/watch?v=opJgMj1IUrc

Carroll, Emma. 2019. *OpenRefine Beginners Tutorial - how to process and reconcile datasets.* [Online]. [Accessed 22 January 2024]. Available from: https://media.ed.ac.uk/media/OpenRefine+Beginners+Tutorial+-+how+to+process+and+reconcile+datasets+by+Emma+Carroll/1_ujc73yrb

GitHub. *Structuring tabular data in spreadsheets.* [Online]. [Accessed 22 January 2024]. Available from: https://nbisweden.github.io/module-organising-data-dm-practices/105-tabular-data/index.html

Kocsis, Andrea. 2022. *The Challenges of Working on the Operation War Diary Records.* 6 April. *National Archives.* [Online]. [Accessed 22 January 2024]. Available from:

https://blog.nationalarchives.gov.uk/the-challenges-of-working-on-the-operation-war-diary-records/

Library Carpentry. *OpenRefine.* [Online]. [Accessed 22 January 2024]. Available from: https://librarycarpentry.org/lc-open-refine/

nodegoat. *Formulating Ambiguity in a Database*. [Online]. [Accessed 22 January 2024]. Available from: https://nodegoat.net/blog.s/21/formulating-ambiguity-in-a-database

OpenRefine. *OpenRefine.* [Online]. [Accessed 22 January 2024]. Available from: https://openrefine.org/

If you have access to in Learning on LinkedIn, there are courses such as:

Du Soleil, Oz. 2022. *Managing and Analyzing Data in Excel.* [Online]. [Accessed 22 January 2024]. Available from: https://www.linkedin.com/learning/excel-managing-and-analyzing-data/managing-and-analyzing-data-in-excel?u=57895809

Using Visualisation Software

Researchers and students may use visualisation software to communicate their findings to their audience. Archivists already have technical skills from inputting data to collection management systems. These skills help when learning data visualisation software.

My project used Power BI and Tableau Public to create visualisations. Some institutions will provide access to Power BI, as part of a subscription to Microsoft products. PowerBI Desktop is a free cut down version. Tableau Public is accessible online if people create an account. The full functionality of Tableau is only available through subscription. Excel can also create effective visualisations.

Resources

in Learning. 2022-23. *Power BI.* [Online]. [Accessed 22 January 2024]. Available from: https://www.linkedin.com/learning/topics/power-bi

in Learning. 2021-23. *Tableau.* [Online]. [Accessed 22 January 2024]. Available from: https://www.linkedin.com/learning/topics/tableau

Microsoft. *Learn Power BI.* [Online]. [Accessed 22 January 2024]. Available from: https://powerbi.microsoft.com/en-gb/learning/

Microsoft. *Support. Charts and other visualizations in Power View.* [Online]. [Accessed 22 January 2024]. Available from: https://support.microsoft.com/en-us/office/charts-and-other-visualizations-in-power-view-141bd462-9853-4973-ac37-842e8345f51e

Microsoft. *Power BI. Uncover powerful insights and turn them into impact.* [Online]. [Accessed 21 January 2024]. Available: https://www.microsoft.com/en-gb/power-platform/products/power-bi/

Salesforce. *Tableau helps people see and understand data.* [Online]. [Accessed 21 January 2024]. Available from: https://www.tableau.com/en-gb/trial/tableau-software?d=7013y000002aJFmAAM&utm_campaign=Prospecting-CORE-ALL-ALL-ALL-ALL&utm_medium=Paid+Search&utm_source=Google+Search&utm_campaign_id=2017049&utm_language=EN&utm_country=UKI&adgroup=&adused=STAT&creative=&gclid=EAIaIQobChMI7OSdr8mHgwMVKj4GAB0yEwRwEAAYASAAEgIff_D_BwE&gclsrc=aw.ds

Salesforce. *Tableau Public. Learn.* [Online]. [Accessed 21 January 2024]. Available from: https://public.tableau.com/app/learn/how-to-videos

Salesforce. *Welcome to Tableau Public.* [Online]. [Accessed 21 January 2024]. Available from: https://public.tableau.com/app/discover

There are many YouTube video tutorials on Power BI, Tableau and Tableau Public.

The following free data visualisation tools are also available:

Datawrapper. *Enrich your stories with charts, maps, and tables.* [Online]. [Accessed 11 January 2024]. Available from: https://www.datawrapper.de/

Gephi. *The Open Viz Graph Platform.* [Online]. [Accessed 11 January 2024]. Available from: *https://gephi.org/*

Stanford. Humanities + Design. *Palladio.* [Online]. [Accessed 11 January 2024]. Available from: https://hdlab.stanford.edu/palladio/

RAWGraphs. *A free and open source tool for data visualisation.* [Online]. [Accessed 11 January 2024]. Available from: https://www.rawgraphs.io/

Understanding the Effectiveness of Visualisations

Visualisations are a communication tool aimed at releasing the value from abstract data by analysing and interpreting it and presenting research findings in an intelligible format. Researchers and students can convert data to knowledge to inform actions and/or decision making. Visualisations often have more impact on an audience than presenting outputs in text or as raw data.

Researchers should be clear about the purpose of their visualisations and understand the needs and graph literacy of their intended audience. The choice of chart or graph and its design all influence the audience's perception and understanding. PowerBI and Tableau training includes courses on choosing the most appropriate visualisations.

Visualisations can, by their nature, be challenging to make accessible. Researchers should aim to meet current Web Content Accessibility Guidelines (WCAG) wherever possible to support digital assistive technologies, keyboard use and voice over. It is important to provide the original data with a visualisation so users with specific accessibility needs can download the data in a common format such as CSV or JSON and access it in their choice of application. The WCAG provide more information.

Resources

Few, Stephen. 2009. *Show me the Numbers: Designing Tables and Graphs to Enlighten*. Burlingame, California: Analytics Press.

Healy, Kieran. 2018. *Data Visualization: A Practical Introduction.* Oxford, Oxfordshire: Princeton University Press.

Kirk, Andy. 2016. *Data Visualisation: A Handbook for Data Driven Design*. Los Angeles: SAGE.

Microsoft, Power BI. *Learn Power BI.* [Online]. [Accessed 22 January 2024]. Available from: https://powerbi.microsoft.com/en-gb/learning/

Office for National Statistics. *Data Visualisation*. *Guidance for Creating Charts and Tables and Best Practice for Using Colour in Your Work.* [Online]. [Accessed 13 March 2024]. Available from: https://service-manual.ons.gov.uk/data-visualisation

Salesforce. *Tableau Public. Learn.* [Online]. [Accessed 22 January 2024]. Available from: https://public.tableau.com/app/learn/how-to-videos

W3C. 2023. *Web Content Accessibility Guidelines (WCAG) 2.2.* [Online]. [Accessed 16 January 2024]. Available from: https://www.w3.org/TR/WCAG22/

Understanding Quantitative Methods

An understanding of quantitative methods helps researchers and students with various stages of a project. These include the design of the study and the collection and analysis of data. Researchers with a knowledge of quantitative methods know which statistical techniques to select to analyse and interpret their data and the most effective visualisations to convey their findings. The

Financial Times GitHub pages host a resource which gives people guidance on choosing the right visualisation for the data relationship they wish to highlight. Future researchers can build more easily on research outputs, which have a solid statistical basis.

There are an increasing number of seminars and webinars offering training on using Chat GPT to learn quantitative methods. This is helpful to researchers, academic teachers, students and archivists who may need to use various forms of statistics occasionally, but whose research and training does not emphasise this as a core skillset.

Resources

ChatGPT. *Get started.* [Online]. [Accessed 16 January 2024]. Available from: https://chat.openai.com/auth/login

Diamantopoulos, Adamantios and Schlegelmilch, Bodo B. 2000. *Taking the Fear Out of Data Analysis: A Step-by-Step Approach*.

GitHub Financial Times. *Visual Vocabulary Designing with data.* [Online]. [Accessed 30 January 2024]. Available from: https://ft-interactive.github.io/visual-vocabulary/

Sage Research Methods. *Datasets.* [Online]. [Accessed 22 January 2024]. Available from: https://methods.sagepub.com/datasets/discipline

The University of Bristol. *Glossary Of Statistical Terms.* [Online]. [Accessed 27 January 2024]. Available from: https://www.bristol.ac.uk/Depts/History/Safe/ITMA/p006cb.htm

Programming Skills

Researchers and students may want to analyse their findings and generate visualisations using a programming language. Archivists often find it challenging to allocate time to learning and implementing programming skills in their day-to-day work unless this is a recognised part of their role. However, archivists processing born-digital materials may already have some knowledge of programming.

Some information professionals find the languages Python and R useful. Library Carpentry teaches a suite of lessons at workshops for librarians and archivists who want to learn more. Another resource is Programming Historian, which provides online tutorials for beginners.

There are videos on YouTube and in Learning.

Resources

Library Carpentry. *Our Lessons, Introduction to Python and Introduction to R*. [Online]. [Accessed 14 January 2024]. Available from: https://librarycarpentry.org/lessons/ (need to be booked as part of a workshop)

Programming Historian. *The Programming Historian.* [Online]. [Accessed 14 January 2024]. Available from: https://programminghistorian.org/en/

Python. *Python.* [Online]. [Accessed 22 January 2024]. Available from: https://www.python.org/about/

r-project organisation. *What is R?* [Online]. [Accessed 22 January 2024]. Available from: https://www.r-project.org/about.html

Open Research Data

This section outlines some of the main factors in making research data accessible and ensuring the sustainability of outputs. Researchers and academic teachers need to consider copyright when formulating digital humanities projects and learning programmes. Copyright applies to research data and its sharing and reuse and most research outputs such as spreadsheets, publications, reports, and data sets fall under literary works. However, people can reproduce the outputs for non-commercial teaching or research purposes without infringing copyright under fair dealing.

The creator of a data set may choose to make it freely available under a Creative Commons licence. Outputs of the Proceedings of the Old Bailey are available under a Creative Commons Attribution-NonCommercial 4.0 International licence.

UK Research and Innovation's Open Access Policy includes a data access statement asking researchers to describe how they will make any data sets underlying their research available. Universities require their researchers to deposit their project outputs, including data. The outputs need to be accessible to users without charge and in a format that does not need proprietary, that is, commercial software.

Some organisations and projects host data sets. The DHI Data Service at the Digital Humanities Institute at the University of Sheffield hosts data sets from projects based at various organisations. Large projects such as Enslaved have downloadable data sets on their webpages.

Copyright

Creative Commons. *Better Sharing, Brighter Future.* [Online]. [Accessed 14 January 2024]. Available from: https://creativecommons.org

The National Archives. 2022. *Copyright and Related Rights*. [Online]. [Accessed 14 January 2024]. Available from: https://cdn.nationalarchives.gov.uk/documents/information-management/copyright-related-rights.pdf

UK Data Service. *Copyright.* [Online]. [Accessed 14 January 2024]. Available from: https://ukdataservice.ac.uk/learning-hub/research-data-management/rights-in-data/copyright/

UK Research and Innovation. *UK Research and Innovation open access policy.* [Online]. [Accessed 14 January 2024]. Available from: https://www.ukri.org/publications/ukri-open-access-policy/uk-research-and-innovation-open-access-policy/

Linked Open Data

Researchers and academic teachers should think about data integration and linking when formulating digital humanities projects and learning programmes. Data integration is the process of combining data sets from different software platforms to give users a unified view of them. Making data easy to import and export increases its accessibility and reuse, which in turn will encourage collaboration among researchers across organisations and promote new discoveries.

Microsoft Power Query, PowerBI and Tableau are among the increasing number of tools that can combine data sets from different platforms. Researchers and students can also use programming languages including R and Python.

Data integration relies on practitioners structuring information in the same formats to create interoperable data sets. This structured information is Linked Open Data (LOD), which machines can easily read. LOD relies on:

- Using a recognised LOD standard format
- Referring to an entity consistently and as other people do
- Publishing data openly

Recognised LOD formats that are familiar to many archivists include the Virtual International Authority File (VIAF) and Wikidata. Both these tools assign unique references to individuals, for example, which help with disambiguation. The Programming Historian has an introduction to LOD.

There are resources such as the UNESCO thesaurus and the Getty Art and Architecture thesaurus that provide a standardised approach to describing subjects. The FAST (Faceted Application of Subject Terminology) has been developed from the Library of Congress Subject Headings to offer a general subject terminology that is easy to apply.

Resources including the thesauri Geonames and Getty TGN help researchers, students and archivists to assign standard names and geo-coordinates to geographic data.

Resources

Microsoft. *PowerBI. Tutorial: Shape and combine data in Power BI Desktop.* [Online]. [Accessed 14 January 2024]. Available from: https://learn.microsoft.com/en-us/power-bi/connect-data/desktop-shape-and-combine-data

Microsoft. *Power Query.* [Online]. [Accessed 14 January 2024]. Available from: https://powerquery.microsoft.com/en-us/

Programming Historian. *Introduction to the Principles of Linked Open Data.* [Online]. [Accessed 14 January 2024]. Available from: https://programminghistorian.org/en/lessons/intro-to-linked-data

Salesforce. Tableau. *Tableau Desktop and Web Authoring Help. Join Your Data.* [Online.] [Accessed 14 January 2024]. Available from: https://help.tableau.com/current/pro/desktop/en-us/joining_tables.htm

Resources on Structured Data

Geonames. *About GeoNames.* [Online]. [Accessed 14 January 2024]. Available from: https://www.geonames.org/about.html

Getty. *Getty Thesaurus of Geographic Names Online.* [Online]. [Accessed 14 January 2024]. Available from: https://www.getty.edu/research/tools/vocabularies/tgn/about.html

OCLC. *FAST. Faceted Application of Subject Terminology.* [Online]. [Accessed 14 January 2024]. Available from: https://www.oclc.org/research/areas/data-science/fast.html

OCLC. *VIAF. Connect authority data across cultures and languages to facilitate research.* [Online]. [Accessed 14 January 2024]. Available from: https://www.oclc.org/en/viaf.html

UNESCO Thesaurus. *About UNESCO Thesaurus.* [Online]. [Accessed 14 January 2024]. Available from: https://vocabularies.unesco.org/browser/en/about

Wikidata. *Wikidata: Introduction.* [Online]. [Accessed 14 January 2024]. Available from: https://www.wikidata.org/wiki/Wikidata:Introduction

HELPFUL ORGANISATIONS AND RESOURCES

This list is not exhaustive but contains the names of organisations and resources that can support research and teaching. Your organisation may have a digital humanities hub can help.

Alliance of Digital Humanities Organizations (ADHO) promotes and supports digital research and teaching. It has guidance on research, publication, collaboration, and training.

Digital Humanities Quarterly (DHQ) is an open access, peer-reviewed, digital journal that covers aspects of digital media in the humanities.

The Digital Research Infrastructure for the Arts and Humanities (DARIAH) supports digitally enabled research and teaching in arts and humanities. It provides a European network of people, expertise, knowledge, content, methods, tools, and technologies. DARIAH offers free online tutorials in topics such as data analysis and knowledge organisation systems.

European Association for Digital Humanities (EADH) represents the Digital Humanities in Europe. It supports the creation of digital humanities interest groups defined by region, language, methodological focus, and other criteria.

GLAM Workbench. This is a collection of tools, tutorials, examples, and hacks for working with data from galleries, libraries, archives, and museums with a focus on Australia and New Zealand.

The N8 Centre of Excellence in Computationally Intensive Research (N8 CIR) is a partnership of the following universities: Durham, Lancaster, Leeds, Liverpool, Manchester, Newcastle, Sheffield, and York. It provides training in CIR methods, skills, and facilities.

JISC is a UK digital, data and technology agency focused on tertiary education, research, and innovation. It hosts a digital research community group to facilitate sharing, discussion, and collaboration among researchers. There is a digital research community JiscMail list.

Research Libraries UK (RLUK) provides a Digital Scholarship Network, a professional peer network for members developing and delivering digital scholarship services in libraries. RLUK published a report by C. Kamposiori in 2017 *The Role of Research Libraries in the creation, archiving, curation, and preservation of tools for the Digital Humanities. RLUK Report/* The report discusses the role of librarians as active partners and collaborators in digital humanities research projects.

The Pattern Recognition & Image Analysis Research Lab (PRImA) at the University of Salford provides expertise, carries out research and projects in document digitisation, OCR, and other image analysis applications. It also develops document analysis tools and hosts several downloadable data sets.

School of Advanced Study at the University of London hosts many digital humanities research projects and has a Digital Humanities Research Hub. The Hub offers various courses, some open to external participants.

Sheffield Digital Humanities Institute is a centre for research and teaching in digital humanities, digital culture, and cultural data. It hosts digital humanities projects from various institutions. The website links to digital humanities project websites and has a DHI data service that hosts data sets that people can explore, download, and reuse.

UK eInformation Group (UKeiG) is a forum for information professionals, users and developers of electronic information resources that offers a wide range of resources.

UK-Ireland Digital Humanities Association brings researchers, practitioners, and organisations together to foster collaborations and capacity for research and teaching in digital humanities.

INSTITUTIONAL SUPPORT AND BENEFITS

The attitude of their parent organisation is of paramount importance to archivists wishing to learn digital humanities skills to assist researchers and academic teachers. Expanding expertise needs resourcing so archivists may have to advocate to their managers for the benefits of providing support and becoming actively involved in partnerships and collaborations. Benefits include enhancing the quality of research and teaching outputs and raising the profile of the parent organisation. Speaking to archivists and researchers who have contributed to collaborative digital humanities projects helps to gather evidence for advocacy.

Employers would benefit from giving their staff the time, resources, and financial support to learn digital humanities related skills. Archivists will need ongoing opportunities to use their newly acquired expertise to ensure they sustain it. However, it would be difficult for one person to learn all the skills outlined in this toolkit to a high degree of competency and maintain them.

Researchers may not be aware of archivists' valuable professional and technical skills when formulating digital humanities projects. In addition, academic teachers can overlook archivists' valuable professional and technical skills when designing learning programmes for their students. Archivists should discuss with researchers and academic teachers how they can contribute to digital humanities research and learning.

The level of support that archivists can provide to research and teaching can be low key, mid-range or full involvement. Organisations need to decide whether it is possible to integrate support with archivists' existing work and if it is sustainable in the long term.

CONCLUSION

Information professionals clearly have professional and technical skills that they can re-purpose to support researchers and academic teachers. Often information professionals' digital abilities are already expanding to meet the challenges of supporting digital humanities, but they may need to advocate in their institutions for the resources and time to learn additional skills. Networking with digital humanities researchers and teachers ensures that they are aware of the benefits of involving information professionals in research projects and learning programmes.

Research and teaching in the humanities increasingly involves computer-based technology. Researchers are also building digital humanities research projects based on partnerships and collaborations in which all the participants contribute their specific skills. Information professionals have valuable knowledge and experience to offer and can contribute to high quality digital humanities research and teaching which enhances the reputation of their institutions.