CULTURAL
ERITA

CULTURAL
HERITA

DIGITAL HUMANITIES -
AN INTRODUCTION

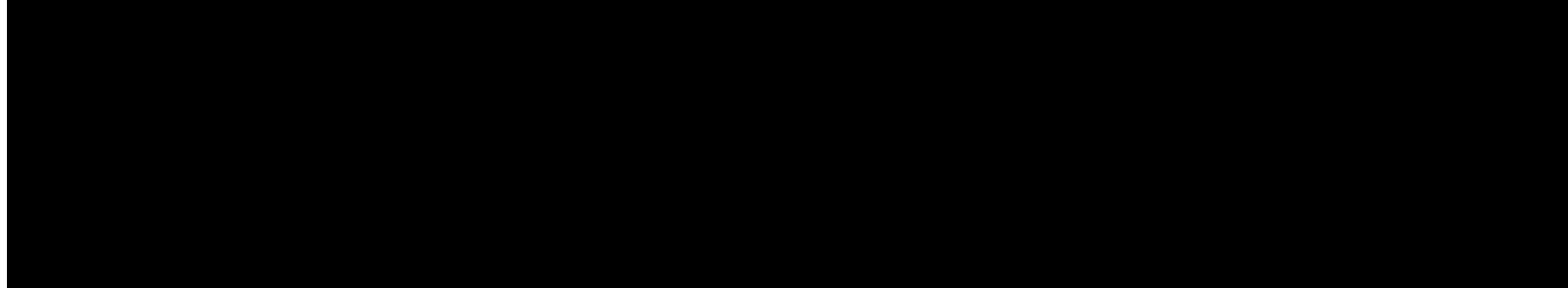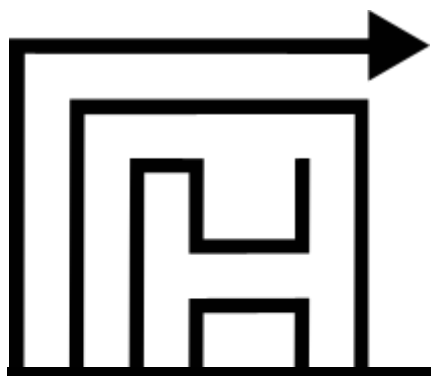## CULTURAL

HERITAGE

01 | INTRODUCTION

This OER aims to give a basic understanding of the scope and focus of the field "Digital Humanities" within research and cultural heritage. It will give a broad overview, how the field developed and how it relates to cultural heritage institutions. At the end the OER, it will give a starting point to further dive into Digital Humanities projects.

01

INTRODUCTION

| "DIGITAL HUMANITIES" - A FIRST DEFINITION

Digital Humanities (DH) is an interdisciplinary field on the interface of computer science and the humanities that is concerned with applying digital methods into humanities research. Moreover, it is concerned how digital practices – including programming, data processing, digitization and digital publication of physical objects or whole historical collections, networking data, data visualization or using digital tools in teaching – influence and change the way we do research in different disciplines of the humanities. These disciplines include for example art history, cultural studies, history, linguistics, literature, etc. (see Manovich, 2020; Schwandt, 2021). Thus, the main objective of this field is exploring how digital technology can be used to create new knowledge about human expressions and our culture at large. It is about new questions that can be asked about the objects of study that were not able to be asked without computational methods – thus the help of a computer in the exploration and generation of new insights about these expressions. A lot of other definitions of Digital Humanities can be found online at https://whatisdigitalhumanities.com.

Moreover DH is in line with the definition of **digital cultural heritage** that includes the digital storage, collection, and preservation of information about cultural objects: from digitization of archives and transforming them into databases, over the digital reproduction or reconstruction of cultural objects, the preservation of digital data to the networking of information or knowledge (see Zschocke et al., 2004).

DEFINITION

|AN (HISTORIC) OVERVIEW ON DIGITAL HUMANITIES

The history of Digital Humanities – which was referred to as "Humanities Computing" until the early 2000s –goes back to the 1940s with the work of Roberto Busa. He was working on a concordance of the works of St. Thomas Aquinas, which according to Tara Andrews from the University of Vienna is "an enormous list of words that we use in a work and the context in which they appear" (Dirtl, 2017). Roberto Busa aimed to research the concept of place in Thomas Aquinas' work, and thus was looking for all words that were relevant to a sense of place with the help of the concordance. He collaborated with IBM in a pilot project to use the new electronic computers for creating the concordance. This work served as inspiration for the digitization of archive or library-records in the 1960s / 1970s and beyond, with the goal to build research databases and develop machine-readable records. Machine readable records are digital versions of e.g. physical finding aids that a computer can process and therewith interpret. This enables, in a next step, the computer-aided analysis of data. One example for an early machine readable record-format in a library environment is "MARC – machine-readable catalog" (https://www.loc.gov/marc/umb/), which was developed for the Library of Congress in the USA.

These projects opened up archives and databases for wider audiences and aided the finding of documents and books. Besides digitizing library records, these early activities also included building corpora – a larger collection of related cultural data such as texts by one author – as well as creating standards for text-encoding (e.g. TEI – Text encoding initiative, developed in 1987) to aid structured transcription and therewith an enriched digitization of texts (see Burdick et al., 2012, p. 8 ff).

Digital Collection Management

While this development started mainly in libraries and text-based research, there were parallel developments in the use of computers to record collection information for museum collections or visual archives in the 1960. However, museum collections had different needs from libraries, as Elana Carpione describes in her thesis (Carpione, 2010). While bibliographic information of a library usually do not change, records in the museum are always changing with new information about museum objects that needed to be added to the records – such as new research findings from curators, changes within exhibitions or the like. The daily doing of a museum needed to be reflected in the museum data. Subsequently, specialized collection management systems for museums were developed starting in the 1960s. Over time they grew from record-keeping systems for accountability over databases modelling workflows of specific museums into commercial collection management software like "The Museum System (TMS)" (initially developed for the Metropolitan Museum of Art in New York). These databases for collection management became crucial tools by increasing public access to collections information, including all information about a museum, interpretive material and even the digitization of museum objects as soon as computers could deal with multimedia material in the 1990s. Thus, successively the computer became a tool for publication and dissemination, for which new, natively digital publication formats were developed: such as database narratives, digital editions, hypertext / hypermedia storytelling or digital exhibitions.

Networking of Information

In the 1990s, the advent of the Internet and the Worldwide Web shifted the focus from data processing to networking of information and data. According to media theorist Wolfgang Ernst, the task of an archive changed "to meaningfully link up different information nodes" (Ernst, 2010, p. 84) and building relations between information e.g. through traversable hyperlinks. Hence also the focus of museums shifted to networking information not only inside their own collection, but also with sources of other institutions. Museums started to publish collection information as well as image-reproductions

of their collection on the Worldwide Web, following the growing demand to make information about their collections easily publicly accessible and findable. This lead to a potential democratization of information, but also increased the need for standards and conventions within the data to allow an interplay of different sources. Standardized vocabularies such as the Getty Arts and Architecture Thesaurus (AAT) (https://www.getty.edu/research/tools/vocabularies/aat/) or ontologies like CIDOC CRM (https://cidoc-crm.org/) gained importance. Where controlled vocabularies enable to build relations between cultural objects through coherent categorization, an ontology formalizes and represents a specific subject area or snippet of the world and describes how it functions as relation between objects and subjects. This enables researchers and cultural institutions to use this abstract model of the world to fill it with concrete information in order to document a specific subject area.

In the 2000s, collection portals such as "Europeana" (https://www.europeana.eu/en) enabled search and access to collection data across multiple European collections. Such a repository development on a massive scale was only possible by collaborating with multiple partners and stakeholders and fosters the big goal "to make cultural legacy available to broad publics for a wide range of puposes" (Burdick et al., 2012, p. 9). This vision is also inherent in the OpenGLAM movement (www.openglam.at). The use of semantic web technologies lead to Linked (Open) Data (https://lod-cloud.net) that enabled to computationally infer meaning and new knowledge that was not explicitly expressed within the data. This opened up the cultural data for new research questions.

Cultural Analytics

Besides the gathering of large scale datasets, that are usually too large for manual analysis by a single human being, the computer is used for generating insights based on these massive amounts of data algorithmically by using e.g. statistical analysis or machine learning. Methods from information visualization or geospatial representations are used for knowledge generation but also for the visual representation and mediation of the data (see Burdick et al., 2012, p. 9). Graphical user interfaces enable us to get an overview on massive amounts of data (generous interfaces), or allow a deep dive and close reading of specific digital cultural objects (ambiguous interfaces) (see Schofield, 2014; Whitelaw, 2009). These visualizations and interfaces are at the heart of the research paradigm Cultural Analytics by Lev Manovich. He refers to Cultural Analytics as "the use of computational and design methods – including data visualization, media and interaction design, statistics, and machine learning – for exploration and analysis of contemporary culture at scale." (Manovich, 2020, p. 9) With the term "scale" Manovich refers to making sense of a vast amount (millions) of images that are produced today. This facilitated the vision to get more encompassing insights into "culture" at large, if one does not have to be as exclusive with the size of the sample than in the era of manual analysis. Because Manovich is focusing predominantly on visual media, methods of visualization play an important role in his research practice. They "give us a language to represent gradual and continuous temporal changes. We can now describe the characteristics of cultural processes that are hard to capture linguistically – for example, gradual historical changes in visual culture over long periods or temporal changes in visual form in the career of an artist" (Manovich, 2020, p. 10). These visualizations themselves are not only research or mediation tools but can themselves become cultural expressions and aesthetic objects that can be exhibited in a museum.

An important idea of this paradigm is what Manovich calls "thinking without categories". The main idea is to "explore and study collections of cultural media and records of cultural behaviors without systems of categories that languages impose on reality" (Manovich, 2020, p. 11), since "[a]ny numerical measurements, linguistic categories, sets of tags, networks, or other forms of representation – regardless of their descriptive power and capacity to make visible similarities and other forms of relations between objects – are also forms of omission" (Manovich, 2020, p. 11), leaving out potential insights. To bypass this bias, the algorithmic analysis of the cultural data provides a "computational point of view" (Wiencek, 2019) that can in a second step serve as a basis and starting point for further human analysis and therewith an novel, non-human entry point to the observation of massive cultural datasets.

| DIGITAL HUMANITIES PROJECTS

How does Digital Humanities work look like in practice? DH-projects can take many different forms. To get an overview over Digital Humanities projects in Austria please have a look at https://digital-humanities.at/en/dha/projects.
The projects are well documented and will give you a broad picture of what falls into Digital Humanities practice: from

- digital editions, such as the correspondence of Arthur Schnitzler (https://digital-humanities.at/en/dha/s-project/arthur-schnitzler-correspondence-writers);
- digitization-projects (such as https://digital-humanities.at/en/dha/s-project/3d-digitizing-glaser-squeezes);
- building up and maintaining digital repositories such as the ARCHE (https://arche.acdh.oeaw.ac.at/browser/);
- the collaborative archive ADA, which involves a community of artists and researchers to document digital art (https://digitalartarchive.at/nc/home.html;
- gamification and citizen science projects like ThingTAG (https://thingtag.at/home);
- to large scale research infrastructures such as InTaVia (In/Tangible European Heritage - Visual Analysis, Curation & Communication, https://intavia.eu/).

| BIBLIOGRAPHY

Burdick, A., Drucker, J., Lunenfeld, P., Presner, T., & Schnapp, J. (2012). *Digital_Humanities*. MIT Press. http://mitpress.mit.edu/sites/default/files/titles/content/9780262018470_Open_Access_Edition.pdf

Carpione, E. C. (2010). *Museum Collections Management Systems: One Size Does Not Fit All* [MA Museum Professions, Seton Hall University]. https://scholarship.shu.edu/dissertations/2366/

Dirtl, T. (2017). *Digital digging into the past*. https://medienportal.univie.ac.at/uniview/semesterfrage/digitalisierung/detailansicht/artikel/digital-digging-into-the-past/?no_cache=1

Ernst, W. (2010). Underway to the Dual System – Classical Archives and / or Digital Memory. In D. Daniels & G. Reisinger (Eds.), *netpioneers 1.0 - contextualising early net-based art* (pp. 81–99). Sternberg Press. http://143.50.30.21/static/publication09/np_ernst_09.pdf

Manovich, L. (2020). *Cultural Analytics*. MIT Press.

Schofield, T. (2014). Ambiguous Interfaces. In *Vimeo Channel of information + visualization lecture series, FH Potsdam*. https://vimeo.com/96062027

Schwandt, S. (2021). Introduction - Digital Humanities in Practice. In S. Schwandt (Ed.), *Digital Methods in the Humanities - Challenges, Ideas, Perspectives* (pp. 7–22). Transcript.

Whitelaw, M. (2009). Visualising the Digital Archive: the Visible Archive project. *Archives & Manuscripts*, *37*, 22–40. http://mtchl.net/papers/VisualisingArchivalCollections.pdf

Wiencek, F. (2019). *Digital Mediation of Art and Culture. A Database Approach.* [PhD Thesis, Jacobs University Bremen]. http://nbn-resolving.org/urn:nbn:de:gbv:579-opus-1008454

Zschocke, N., Blome, G., & Fleischmann, M. (2004). *Cultural Heritage. Kunstvermittlung mit Digitalen Medien*. Netzspannung.Org. http://netzspannung.org/media-art/topics/cultural-heritage/?lang=en