

DIGHUMLAB – AN ECOSYSTEM SUPPORTING DIGITAL HUMANITY RESEARCH AND EDUCATION

By Birte Christensen-Dalsgaard

DIGHUMLAB

Public version 1.1

March 20, 2017

D	I	G
H	U	M
L	A	B

Version history

Version no.	Data	Author	Status	Changes
1.0	2016-12-20	BCD	Draft version for comment among theme leaders	
1.01	2017-03-02	BCD	Draft for steering group	Change of layout; sustainability added
1.1	2017-03-20	BCD	Public Version	

D	I	G
H	U	M
L	A	B

INDHOLD

Introduction	1
Mission, values and vision for DIGHUMLAB	1
DIGHUMLAB - A community of communities	2
DIGHUMLAB - A digital ecosystem	3
Existing Services in DIGHUMLAB.....	5
DIGHUMLAB tools.....	5

D	I	G
H	U	M
L	A	B

INTRODUCTION

EU defines research infrastructure¹ as: **facilities, resources and related services** used by the scientific community to conduct top-level research in their respective fields,.

DIGHUMLAB is a mix of cutting edge tools, examples and tutorials based on front line research and experts ready to share their experience with digital research activities within the humanities and social sciences.

The **facilities** in DIGHUMLAB are advanced equipment and top of the line software, the **resources** are both the digital objects offered by the cultural institutions and tutorials and advice from experts and **related services** are the support, training, outreach, and administrative structure built around the facilities.

MISSION, VALUES AND VISION FOR DIGHUMLAB

DIGHUMLAB is a **partnership** between institutions sharing the mission: *“to enhance and facilitate digital humanities in Danish research, thereby contributing to greater interdisciplinary cooperation, widespread knowledge transfer and global orientation and increased internationalisation of both research and education².”*

The partnership is based on **values** such as:

- **Research driven** - to ensure relevance of tools, data and related activities
- **Excellent** – research infrastructure is excellent on an international scale and allows researchers to perform cutting edge research
- **Inspirational** – to demonstrate the potential of digital methods in humanities
- **Shareable** – to allow others to do research based on tools supported by DIGHUMLAB
- **Easy Accessible and Usable** – to ensure uptake
- **Ethical** – to assist on ethical conduct of research
- **Sustainable** – to give stability

DIGHUMLAB supports the Danish research community in digital research activities within the humanities and the social sciences by:

- Providing access to cutting edge facilities for research and education in the disciplines covered
- Inspiring to new research questions by demonstrating the possibilities in digital research activities and point to relevant tools
- Supporting competence development by offering access to e-learning material

¹ (https://ec.europa.eu/research/infrastructures/index_en.cfm?pg=what)

² <http://dighumlab.com/about/vision-mission-goals/>

D	I	G
H	U	M
L	A	B

The activities in DIGHUMLAB are carried out by experts from research groups or the library community, who share tools and data from their own domain and offer inspiration and support to colleagues in the form of tutorials, workshops, personal advice and a helpdesk.

DIGHUMLAB identifies gaps between research activities and available tools and works toward ensuring that these are filled – either through international partnerships or by raising funding for the required development.

The **vision** for DIGHUMLAB is:

DIGHUMLAB is the *preferred Danish community* for inspiration to digital research activities within the humanities and the social sciences and for recommendation on and access to methodologies, tools and data.

The economy of DIGHUMLAB will enable experts to be compensated for their support activities. The economy will be based on payment from faculties for daily running costs and of grants and other funding for new activities.

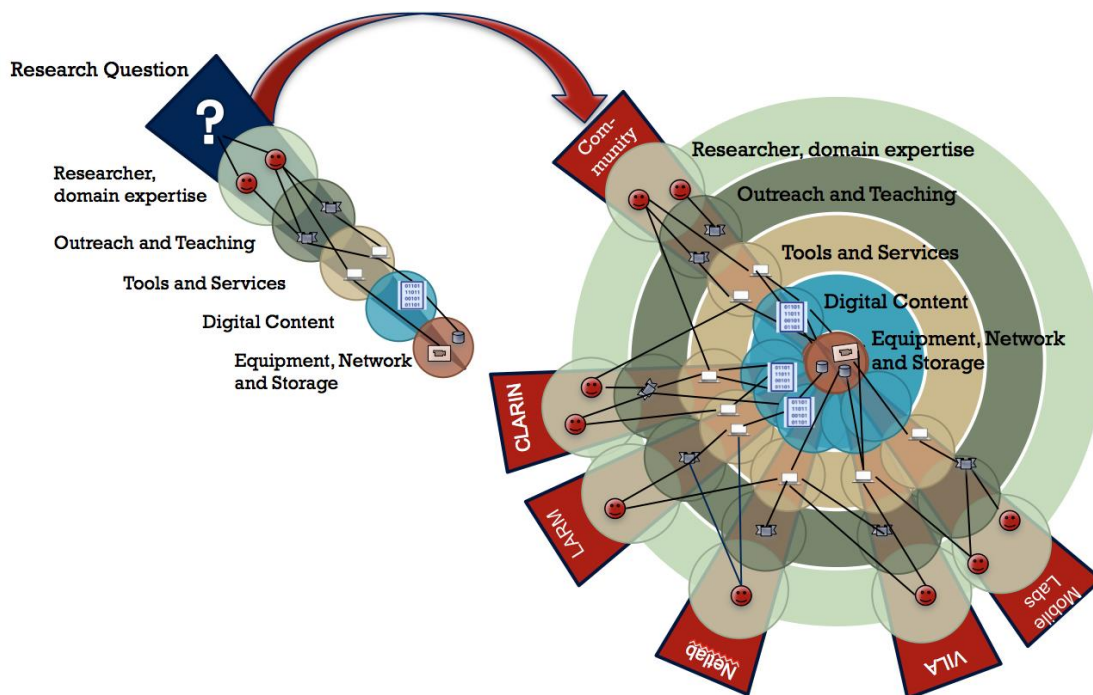
The non-commercial tools and online tutorials are available openly. The technical and expertise support and training are offered to the institutional partners behind DIGHUMLAB. In return, the research group will have easy access to a pool of tools and data and will be able to draw on the experts in other fields and in general be part of the community setting the agenda for the development of the DIGHUMLAB infrastructure.

DIGHUMLAB - A COMMUNITY OF COMMUNITIES

DIGHUMLAB is based on research centers and research groups who share tools and data from their own domain and offer inspiration and support to colleagues in the form of tutorials, workshops and a helpdesk.

As DIGHUMLAB is an infrastructure it is based on communities formed for specific purposes. It can be centered on research related to specific collections (e.g. the radio collection and the net archive) or research using specific tools (e.g. for natural language processing or video processing).

D	I	G
H	U	M
L	A	B



By pooling the knowledge components from the different communities, the combination can be useful both for the researcher herself as she gets access to the tools, methodologies etc. from other researchers, but also as a general base of knowledge, which can be accessed by others. New research questions can be inspired by research in related fields, tools and programming expertise are available to support new research activities, competence development can be based on the online tutorials, etc.

DIGHUMLAB will rely on availability of super computer power, storage and network; however these are not part of DIGHUMLAB and are expected to be provided by DeIC and/or university computing departments. DIGHUMLAB provides access to equipment in labs such as video cameras and computers with the relevant software installed.

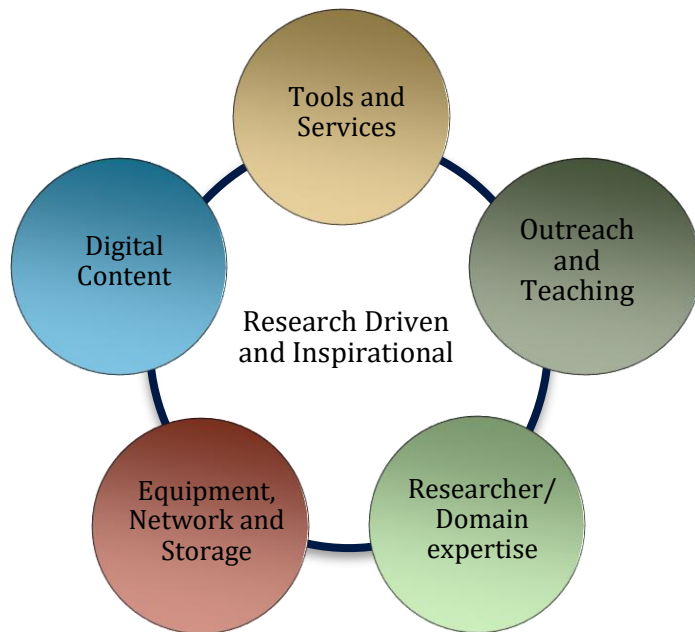
Tools, data and learning objects will adhere to standards when possible thereby supporting the cross-community use. Also open source/data are preferred to closed data; This said we notice that most of the digital content is restricted by copyright and/or due to privacy.

DIGHUMLAB - A DIGITAL ECOSYSTEM

An often-used metaphor for digital infrastructures is “digital ecosystem”, indicating that the environment and usage change so fast, that building monolithic systems is outdated and doomed to fail. This also applies to DIGHUMLAB, which in its construction based on the actual used methodologies and their associated tools, consists of an array of systems and services.

From an end-users perspective, DIGHUMLAB is a digital ecosystem of tools, digital content, services and experts. To make it function, it is based on availability of computer, networks and storage.

D	I	G
H	U	M
L	A	B



The components of DIGHUMLAB are digital objects, services and support – developed in response to real research questions. The **interoperability** is ensured through adherence to exchange standards for metadata, for the objects, for the software and for communication. The **usability** is ensured through user involvement in future developments, through manuals and/or training material and through an expert support system.

The technical infrastructure consists of network and the hardware provided and maintained by the IT-departments at the partner institutions, by DeiC or by a cloud provider. This mix of responsibilities is reflected in the sustainability model for DIGHUMLAB.

The profile of DIGHUMLAB as a tool and content provider/promoter is to ensure a relevant, adequate selection set in a context. In particular:

- Participating in and supplementing the national and international development in selected areas developing missing components of special interest for the Danish community.
- Promoting and mobilising the work from the international communities into a Danish context – e.g. via the case stories.
- Creating guides and tutorials to tools and content, which are used by students and researchers at the Danish universities.
- Working towards creating a richer content base for the digital humanistic scholars through exposing and/or removing barriers for using existing digital material residing in archives, museums and other content collecting institutions.

The entry to the tools is using a common taxonomy for digital activities within the humanities and the social sciences developed by the European project DARIAH. The tools promoted for DIGHUMLAB mostly are for 'Capture of material' or for 'Enrichment', as can be seen on the figure below, where the DIGHUMLAB tools are shown as the outer circle – on top of the two inner circles containing the elements of the TaDiRah taxonomi.

D	I	G
H	U	M
L	A	B

DIGHUMLAB works towards supplementing its portfolio to contain tools and services supporting the analysis part. Here, however, we are fortunate to have a lot of existing initiatives to use, such as the open source Gephir for link analysis and the commercial Tableau as an integrated analysis tool.

DIGHUMLAB will use the community to evaluate existing activities related to tools, content and services to ensure DIGHUMLAB represents the needs of the user community. Furthermore, DIGHUMLAB needs to stimulate the development of new ideas for the infrastructure. The decision on activities to be terminated or initiated lies with the steering group.

EXISTING SERVICES IN DIGHUMLAB

The students and researchers of the participating institutions of DIGHUMLAB can freely use and seek advice on tools and services as described below.

DIGHUMLAB TOOLS

The following tools have been developed as part of or has been embedded in DIGHUMLAB:

Tools, which has been developed by members of DIGHUMLAB and to which DIGHUMLAB has a responsibility of maintaining the tools:

CLARIN infrastructure tools for:

- Linguistic annotations (POS-Tagging openNLP, Brill's PosTagger, Name Recogniser, CST lemmatiser, Bohnets parser) and for
- Conversation (TEIP5 tokenizer, TEIP5 segmenter, CoNLL converter)

LARM infrastructure giving access to more than one mio. radio- and television programs and supporting the following features:

- Collaboration
- Shared annotations and enrichment of the material
- Access to all material in mediastream (newspapers, broadcasts and manuscripts for news-broadcasts)

Lab-facilities to video-capture events and to annotate these:

- Mobile Video-capture facility – based on GoPro camera
- Video Research Lab in Aalborg
- Video Editing Lab in Kolding

The following open source or/and commercial tools are used by members of the DIGHUMLAB community:

A set of tools for gathering of webbased material explained and used in a workshop on creating web archives:

- Single pages: (Web Snapper, Paparazzi)
- Sound and video: (Video Download Helper, Musicbox, WireTap Studio, Videobox)
- Screen capture: (Snagit)

D	I	G
H	U	M
L	A	B

- Downloading of entire websites: (HTTrack)

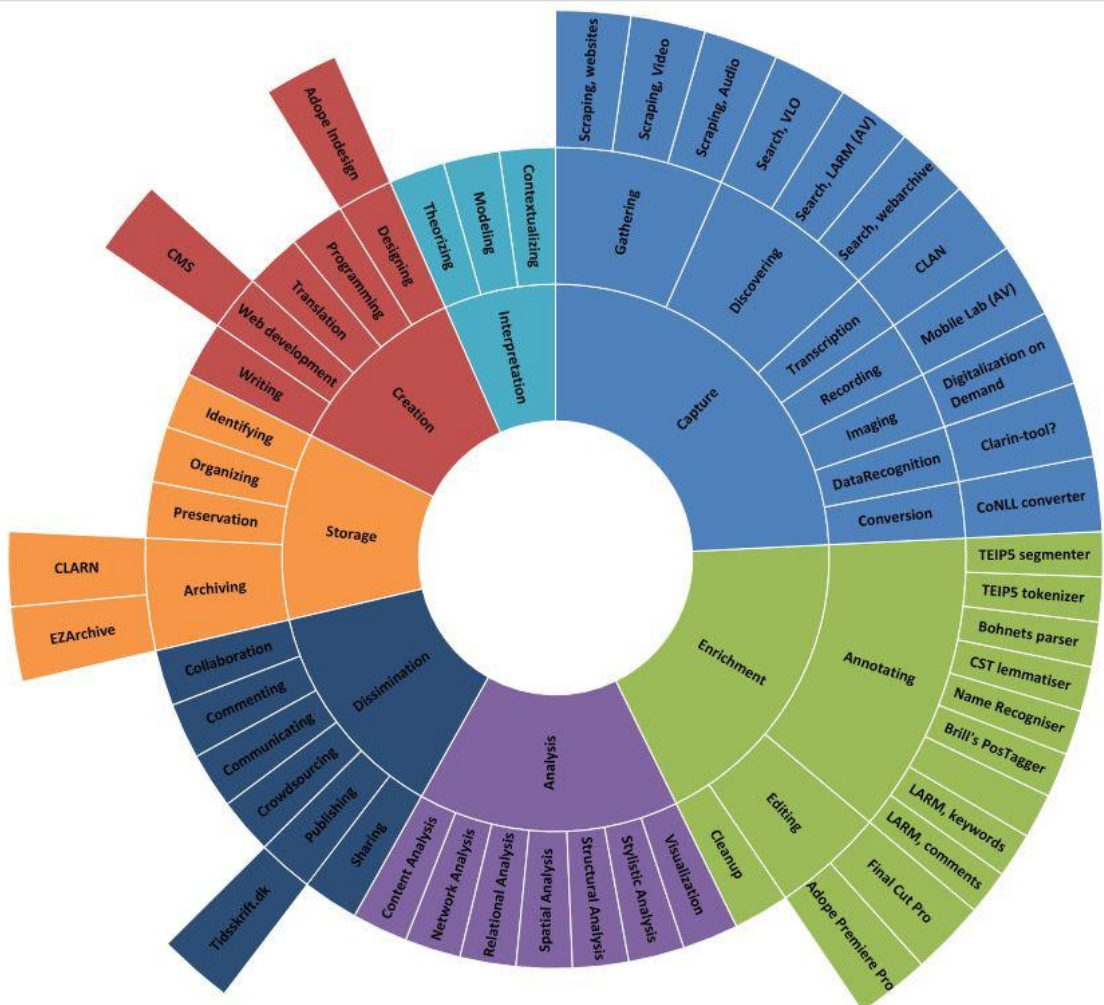
A set of tools for working with video including tutorials explaining how to use these:

- Transcription tools: (CLAN, ELAN, Praat)
- Video-editing tools: (FinalCut Pro, Premiere Pro)

A tool for document design

- InDesign

The tools of DIGHUMLAB are visualized in the figure below, where the taxonomy of TaDiRah is used as a basis for characterizing the tools:



Most of the DIGHUMLAB tools have been developed as open source and have been submitted to GitHub.