# CEDIFOR
## Centre for the Digital Foundation of Research in the Humanities, Social, and Educational Sciences

Approved under the name of **Frankfurter eHumanities-Zentrum (FeHZ)**

Goethe-Universität Frankfurt
Technische Universität Darmstadt
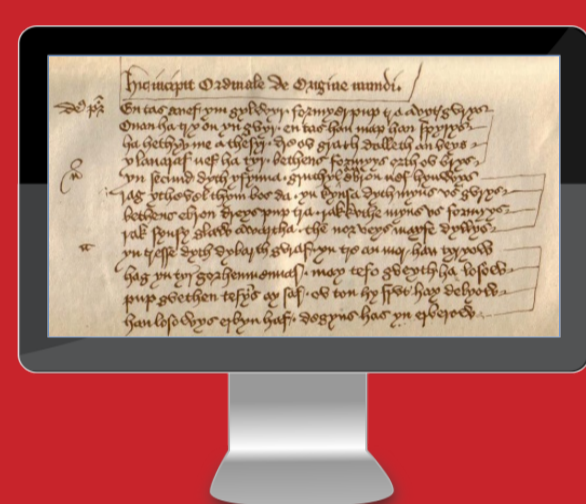Deutsches Institut für Internationale Pädagogische Forschung

GOETHE UNIVERSITÄT FRANKFURT AM MAIN · TECHNISCHE UNIVERSITÄT DARMSTADT · DIPF Educational Research and Educational Information

---

## CEDIFOR: Vision

- We serve as a **primary contact point for eHumanities projects** in universities, research institutes, archives, libraries, and museums in the Rhein-Main area and beyond
- We cover research projects in the **humanities, social sciences, and educational science**
- Our results are **sustainable** and our services are **long-term**
- We work with **text, image, video, and audio**
- We work and publish **nationally and internationally**
- We base our work on the results of the **LOEWE Research Cluster "Digital Humanities"** (2011-2014)

---

## Innovative Research Conceptions

### Epistemological Innovation

**Consulting and Supervising Novel Research Concepts**

- Integration of humanities with computer science for mutual benefit
- Discovery, elaboration, and realization of research projects

**Developing Tools for Multimedia Research Subjects**

- Integration of multi-code, multimodal, and multimedia research subjects
- Creation of highly structured corpora: internally linked and stratified across time

### Technological Innovation

### Infrastructural Innovation

**Creating a Research Infrastructure for the Common Usage of Tools and Data**

- Enable shared usage of research subjects, data, and tools
- Create sustainable results, workflows, and best practices

---

## eHumanities: The Current State

**Research Environments**
- **Annotation:** CSniper, WebAnno, OWLnotator
- **Processing:** NLTK, GATE, UIMA, DKPro, eHumanities Desktop,
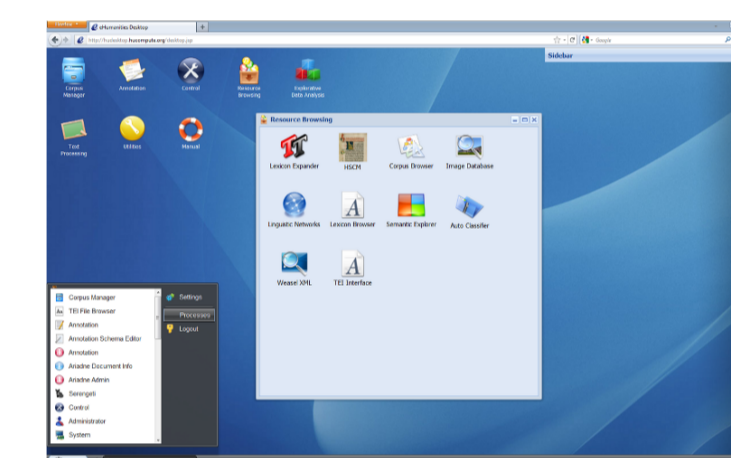- **Discovery:** Semantic CorA

**Assignment and Linking of (Meta-)data**
- Based on the **Lexical Markup Framework:** UBY
- Based on the **Semantic Web** standard: OWL, RDF

**Development and Analysis of Multimodal/Multimedia Data**
- **Image:** ImageDB
- **Video/Audio:** ELAN

**Infrastructure**
- DARIAH-DE
- CLARIN-D

CLARIN-D
DARIAH-DE

*Tools developed and/or maintained by ourselves are highlighted.*

### Challenges

- Researchers in humanities have difficulties to manage the coordination, interpretation, and linking of growing amounts of metadata
- Lack of accessible tools for the analysis of multimedia data (image, video and audio)
- Error rates in applied Natural Language Processing methods are too high in practice

---

## First Year Pilot Projects

### Historical Semantics, Cooccurrence Analysis, and Collation

- extension of the Historical Semantics Corpus Management tool
- development of a digital, diachronic dictionary for classical and later forms of Latin
- online editions in cooperation with the Monumenta Germaniae Historica
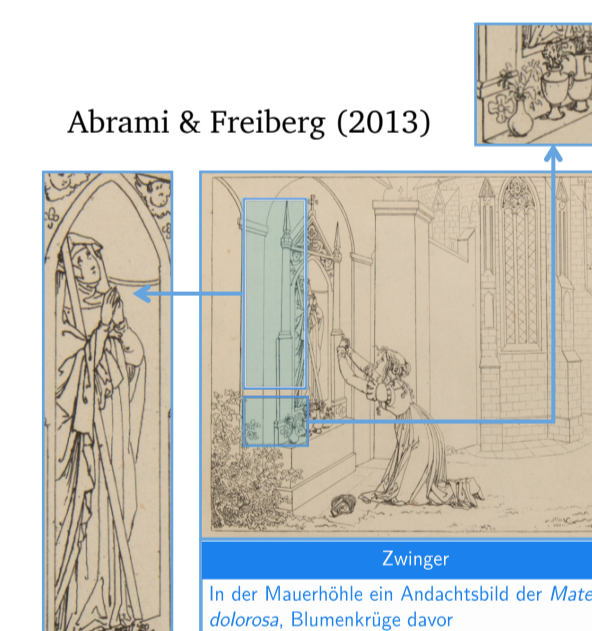
### Academic Trails in Multimedia Spaces

- intelligent system for processing image-text spaces in art
- semi-automatic extraction of highly structured corpora from large databases of image-text media
- recommendation-driven access to archive data for the investigation of content

### Academic Trails in Multimedial Spaces

**Goals:**
- Development of a system for computer-supported traversing of multimedia spaces in the context of fine arts
- Development of an ontology for the structural annotation of image-text spaces ("Guided Trails")
- Development of an explorative system to be integrated with OWLnotator, which enables a semi-automated intermedia and crossmedia annotation
- Elaborating and addressing research questions dealing with intermediality by the help of highly structured corpora

**Users:**
- Art historians, historians, philologists, media scientists
- Museums, archives

Abrami & Freiberg (2013)

### Computer-Supported Discovery of Metaphors

- context-based discovery of metaphors in large corpora
- computer-supported visualization and processing of the data
- theory and content-based classification of metaphors

### Inferences Between Linguistic Attributes and Higher Level Features

- correlations between linguistic attributes and personality traits
- example traits: performance motivation, emotional stability, locus of control
- use case: analysis of student texts

### Personality Analysis Based on Linguistic Attributes

**Research Questions:**
- How well can we predict various traits from text automatically? (e.g. age, extraversion vs. introversion)
- Which aspects of language usage can be indicators for certain personality types, moods, attitudes? (e.g., do depressed people talk more about themselves? do extroverts talk more in general?)
- How robust are the results across different data sets? (e.g., do gender differences in style hold for both school essays and social media?)

**Results from Preliminary Studies on Personality Traits:**
- Accuracy automated classification of "Big 5" personality traits on student essays: 0.56-0.60
- Linguistic indicators for *Extraversion:* "love", "boyfriend", "so much" etc.

---

## eHumanities: Our Approach

### Epistemological Innovation

- **Supervision** of pilot projects ("translation" between humanities and computer science)
- Creation of **reference workflows**
- Increasing national and international **visibility** (common publications, open access)
- **long-term** support for end users

### Technological Innovation

- Technical support for **digitizing**/converting data
- **Extending existing tools** according to project-specific needs
- „On-demand" support
- Technical **documentation**
- Open Source Software **Releases**

### Infrastructural Innovation

- Creation of a **research agenda to ensure reusability**
- **Organization** of workshops, providing a technical research infrastructure
- **Documentation** and Publication (Workflows, Best Practices)

---

## Partner Institutions

**Rhein-Main**
- Leibniz-Institut für Europäische Geschichte Mainz
- Universität Gießen
- Hochschule für Musik und Darstellende Kunst Frankfurt
- Universitätsbibliothek Frankfurt
- Deutsche Nationalbibliothek Frankfurt
- Städel Museum Frankfurt
- Freies Deutsches Hochstift / Frankfurter Goethe-Museum
- MPI für Europäische Rechtsgeschichte Frankfurt
- ULB Darmstadt

**National and International**
- Monumenta Germaniae Historica München
- Bibliothek für Bildungsgeschichtliche Forschung Berlin
- Universität Hildesheim
- Universität Bamberg
- Max Planck Institut Nijmegen
- Harvard University
- Centre for the Study of Manuscript Cultures Hamburg
- Europeana Den Haag

---

**Executive Board**
Prof. Dr. Jost Gippert (Goethe-Universität Frankfurt)
Prof. Dr. Iryna Gurevych (Technische Universität Darmstadt)
Prof. Dr. Alexander Mehler (Goethe-Universität Frankfurt)
Prof. Dr. Marc Rittberger (Deutsches Institut für Internationale Pädagogische Forschung)

**References**
- Abrami, G. & Freiberg, M. (2013). Final Presentation for the Evaluation of the LOEWE Research Cluster "Digital Humanities".

**Image Sources**
http://commons.wikimedia.org/wiki/File:Origo_Mundi_kynsa_gwersow.jpg

Contact: Dr. Ralf Gehrke, gehrke@rz.uni-frankfurt.de