



# LOEWE

## ABSCHLUSSBERICHT



**LOEWE-Schwerpunkt  
Digital Humanities – Integrierte Aufbereitung  
und Auswertung textbasierter Corpora**

# Inhalt

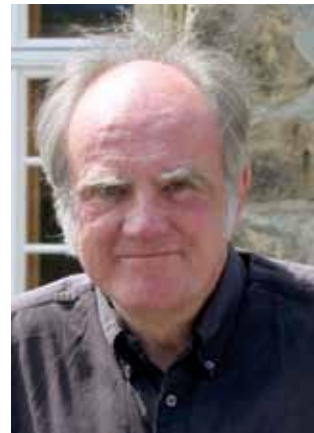
- 2 Statement des Koordinators
- 3 Projektinhalte
- 3 Wissenschaftlich-technische Ausgangslage
- 3 Im Rahmen des LOEWE-Projekts erreichte Erkenntnisse und getätigte Entwicklungen
- 5 Erreichte Strukturentwicklung
- 5 Erreichte Bedeutung/Stellung im Themen-/Forschungsfeld
- 6 Weitere Informationsmöglichkeiten
- 6 Zahlen und Fakten
- 7 Wichtigste Meilensteine des Projekts
- 8 Kurzvorstellung der beteiligten Hochschulen und Forschungsinstitute
- 9 Impressum

Unter der Bezeichnung *Digital Humanities* hat sich in jüngerer Zeit ein Methodenwandel in den Geisteswissenschaften etabliert, der davon ausgeht, dass die Digitalisierung verschiedenster forschungsrelevanter Inhalte eine bis dato unerreichbare empirische Basis für den wissenschaftlichen Fortschritt legen und die digitale Modellierung von Forschungsaufgaben zu einer neuer Qualität von Forschungsergebnissen führen kann.

Mit der Förderung des LOEWE-Schwerpunkts *Digital Humanities: Integrierte Aufbereitung und Auswertung textbasierter Corpora* von 2011 bis 2014 hat das Hessische Ministerium für Wissenschaft und Kunst einen bemerkenswerten Rahmen geschaffen, um die Tragfähigkeit dieses Anspruchs im Zusammenwirken einer ungewöhnlichen Breite von Disziplinen (Sprach-, Literatur-, Geschichtswissenschaften, Texttechnologie und Informatik) und anhand einer ebenso ungewöhnlichen Breite von Forschungsaufgaben auszuloten und mit eigenständigen Ansätzen zur weiteren Entwicklung der *Digital Humanities* beizutragen.

Die an den drei Standorten (Goethe-Universität Frankfurt, Technische Universität (TU) Darmstadt, Freies Deutsches Hochstift/Goethe-Museum Frankfurt) im Rahmen des LOEWE-Schwerpunkts durchgeführten Forschungsaufgaben haben in eindrucksvoller Weise bestätigt, dass der eingeschlagene „digitale“ Weg für die geisteswissenschaftlichen Fächer zukunftsweisend ist, sowohl im Hinblick auf „klassische“ Forschungsfragen wie diejenigen nach den wechselseitigen Abhängigkeiten von Bibelhandschriften als auch im Hinblick auf „neuartige“, durch die digitalen Medien selbst generierte, Fragestellungen wie diejenigen nach den Prozessen, die bei der Entstehung von Wikipedia-Artikeln ablaufen, oder den Möglichkeiten digital basierter Texteditionen.

Mit Hilfe der für die einzelnen Projekte gemeinsam auf- und ausgebauten Forschungsinfrastruktur hat der Schwerpunkt dem Arbeitsfeld der *Digital Humanities* neue Möglichkeiten erschlossen und das Rhein-Main-Gebiet als Zentrum für diesen Bereich etabliert. Dies wird deutlich in der Gründung des aus dem LOEWE-Schwerpunkt hervorgegangenen, vom BMBF geförderten „Centrums für Digitale Forschung in den Geistes-, Sozial- und Bildungswissenschaften“ (CEDIFOR), das das erworbene Know-how und die vorhandene Infrastruktur für weitere Projekte zur Verfügung stellt und dabei gezielt ausbauen wird.



A handwritten signature in black ink, appearing to read 'Jost Gippert'.

Prof. Dr. Jost Gippert  
Koordinator des LOEWE-Schwerpunkts *Digital Humanities*  
Goethe-Universität Frankfurt am Main

## Projekthalte

### Wissenschaftlich-technische Ausgangslage

Verschiedene geisteswissenschaftliche Fächer haben wissenschaftsgeschichtlich jeweils disziplinspezifische Verfahren entwickelt, die der Erschließung des kulturellen Erbes und damit der materialen Grundlegung der wissenschaftlichen Arbeit dienen. Wenige Jahre vor Beginn des LOEWE-Schwerpunkts haben jedoch interdisziplinäre Kooperationen im geisteswissenschaftlichen Umfeld unter der Bezeichnung *Digital Humanities* erfolgreich die Digitalisierung und Verfügbarmachung verschiedenster forschungsrelevanter Inhalte begonnen und damit eine bis dahin unzugängliche empirische Basis für den wissenschaftlich-methodologischen Fortschritt gelegt. Förderpolitisch wurden diese Aktivitäten nur partiell flankiert; vielfach war eine Vereinzelung entsprechender Vorhaben zu beklagen.

Auf der technischen Seite hatte auch die Entwicklung geeigneter Tools und Infrastrukturumgebungen begonnen. Viele Aspekte der für die Erforschung von Sprach- und Bildstrukturen in historischen und gegenwartsprachlichen Corpora erforderlichen computertechnischen Umgebung wie etwa die Multimodalität (Beschreibung von Bild-Bild- und Bild-Text-Bezügen), die Möglichkeit der Mehrebenenannotation und der Einsatz eines flexiblen Annotationsmodells waren in der benötigten Form allerdings noch nicht verfügbar.

### Im Rahmen des LOEWE-Projekts erreichte Erkenntnisse und getätigte Entwicklungen

Im Bereich der *Digital Humanities* tätige Forscherinnen und Forscher gehen davon aus, dass die Digitalisierung verschiedenster forschungsrelevanter Inhalte eine zuvor unerreichbare empirische Basis für den wissenschaftlichen Fortschritt legen und die digitale Modellierung von Forschungsaufgaben zu einer neuen Qualität von Forschungsergebnissen führen kann.

Mit der Einrichtung des LOEWE-Schwerpunkts *Digital Humanities*: Integrierte Aufbereitung und Auswertung textbasierter Corpora zum 01.01.2011 hat das



Stemmagerierung am Bildschirm. Die Abhängigkeiten verschiedener Manuskripte desselben Textes werden durch elektronische Analysen ermittelt und als Stammbaum sichtbar gemacht.



Elektronische Aufbereitung durch Multispektralaufnahmen. Die obere Abbildung zeigt das Farbfoto der Seite eines Palimpsests (wiederverwendetes Manuskript nach Löschung des ursprünglichen Textes), die untere dieselbe Seite nach der elektronischen Aufbereitung von Aufnahmen bei unterschiedlichen IR-, Licht- oder UV-Frequenzen mit Hilfe der Multispektralkamera. Die obere Schrift (in altgeorgischer Sprache) wird dabei weitgehend optisch entfernt und die untere (in kaukasisch-albanischer Sprache) wird verstärkt und tritt nun deutlich hervor.



Einsatz der Multispektralkamera zur Sichtbarmachung getilgter Textstellen

Hessische Ministerium für Wissenschaft und Kunst eine bemerkenswerte Chance geschaffen, um die Tragfähigkeit dieses Anspruchs im Zusammenwirken einer ungewöhnlichen Breite von Disziplinen (Sprach-, Literatur-, Geschichtswissenschaften, Texttechnologie und Informatik) und anhand einer ebenso ungewöhnlichen Breite von Forschungsaufgaben auszuloten und mit eigenständigen Ansätzen zur weiteren Entwicklung der *Digital Humanities* beizutragen.

Die an den drei Standorten (Goethe-Universität Frankfurt, TU Darmstadt, Freies Deutsches Hochstift/Goethe-Museum Frankfurt) im Rahmen des LOEWE-Schwerpunkts in Angriff genommenen Forschungsaufgaben haben in eindrucksvoller Weise bestätigt, dass der eingeschlagene „digitale“ Weg für die geisteswissenschaftlichen Fächer zukunftsweisend ist, und dies sowohl im Hinblick auf „klassische“ Forschungsfragen wie diejenigen nach den wechselseitigen Abhängigkeiten von Bibelhandschriften wie auch im Hinblick auf „neuartige“, durch die digitalen Medien selbst generierte, Fragestellungen wie diejenigen nach den Prozessen, die bei der Entstehung von Wikipedia-Artikeln ablaufen, oder den Möglichkeiten digital basierter Texteditionen. Der beschränkte Raum verbietet eine Auflistung der in den einzelnen Projekten erzielten Ergebnisse und Erkenntnisse. Es kann hier nur ein Überblick über die Bandbreite der Themenstellungen gegeben werden.

Die im LOEWE-Schwerpunkt realisierten Vorhaben sind verschiedenen Projektbereichen zugeordnet worden, die den unterschiedlichen inhaltlichen bzw. methodischen Aufgabenstellungen entsprechen.

Gegenstand des Projektbereichs „Historische Corpora“ waren Forschungsaufgaben, die sich auf die Textüberlieferung älterer Sprachstufen des Deutschen, aber auch anderer – damit verwandter oder nicht verwandter – Sprachen wie des Lateinischen, des iranischen Avestischen oder kaukasischer Sprachen beziehen. Die behandelten Fragestellungen erstrecken sich von den Bezügen zwischen Paralleltexten wie z. B. den verschiedenen altgermanischen Bibelübersetzungen und -adaptationen und ihren mutmaßlichen lateinischen Vorlagen über die Wechselbeziehungen zwischen einzelnen sprachlichen Elementen in derartigen Texten bis hin zur hermeneutischen Auswertung von Textcorpora für geschichtswissenschaftliche Fragestellungen (im Sinne der „Historischen Semantik“).

Der Projektbereich „Gegenwartssprachliche Corpora“ konzentrierte sich, seinem Titel gemäß, auf aktuelles Sprachmaterial, das er im Hinblick auf drei einschlägige Forschungsfragen behandelte: 1) die Entwicklung von Verfahren der automatischen Klassifikation für größere Textmengen unter Einschluss von Methoden des Maschinellen Lernens, 2) das Erken-

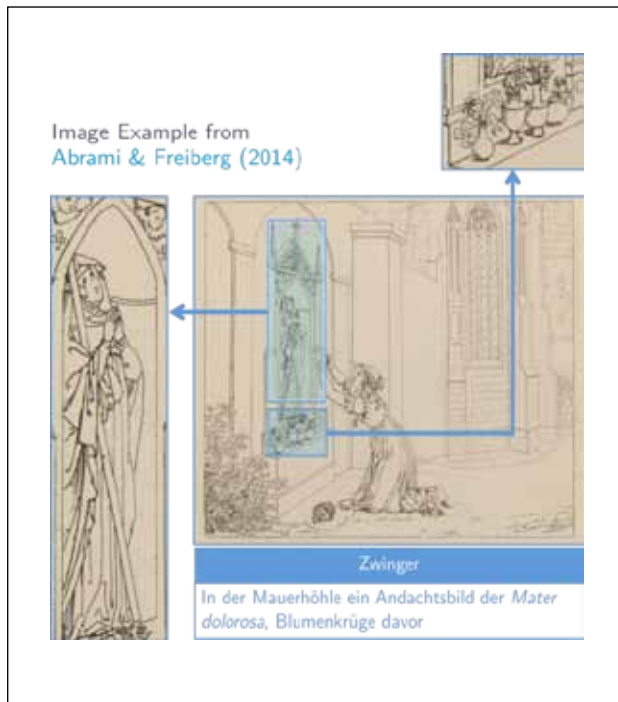
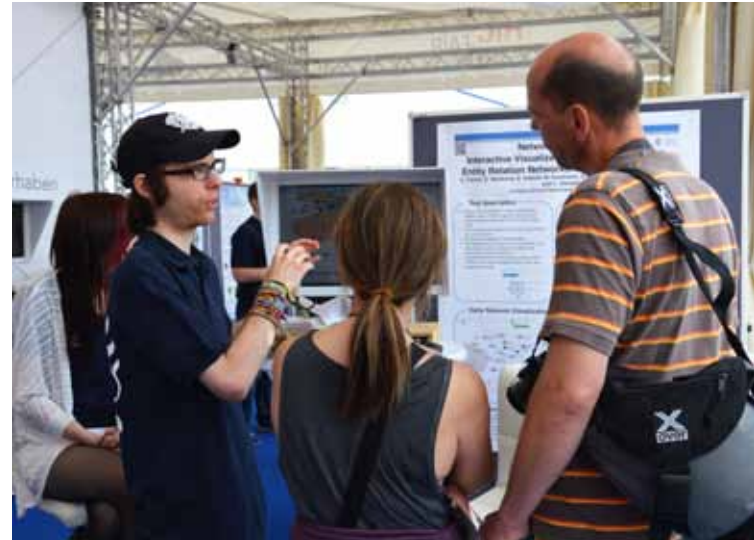


Abb. oben: Die beiden LOEWE-Mitarbeiter G. Abrami (rechts) und M. Freiberg (Mitte) präsentieren das Teilprojekt *Faust-Illustrationen* auf dem Hessentag 2014.

Abb. unten: Bei den Illustrationen zu Goethes *Faust* werden Bildsegmente extrahiert und mit den entsprechenden Textstellen verknüpft (Bild-Bild- und Bild-Text-Verknüpfungen).

nen und Identifizieren von untypischen grammatischen Konstruktionen des Modernen Englischen mit automatischen Verfahren und 3) das Entstehen von Textcorpora unter den Bedingungen kollaborativer Autorschaft am Beispiel der Wikipedia.

Im Projektbereich „Archiv, Corpus, Edition“ ging es um die Auslotung von Möglichkeiten zur wissenschaftlichen Absicherung und zur Ausgestaltung digitaler Textausgaben: 1) die Einbindung von Bildmaterialien in online bereitzustellende Neuauflagen, 2) die Untersuchung von Verfahren zur elektronischen Ermittlung und Abbildung von Handschriftenstambäumen und 3) die digitale photographische Analyse schwer lesbarer Handschriften (z. B. Palimpseste) und die Einbindung entsprechender Digitalisate in wissenschaftliche Editionen.



Auch nach dem Projektabschluss: U. Fahrer vom LOEWE-Schwerpunkt Digital Humanities zeigt auf dem Hessentag 2015 Wegweiser in der heutigen Datenflut.

Für die daten- und programmtechnische Unterstützung der geisteswissenschaftlichen Forschungsvorhaben stand der Projektbereich „Multimodales Corpusmanagement“, der sich zugleich in erheblichem Maße der Entwicklung neuartiger digitaler Werkzeuge, Methoden und Strukturen widmete.

## Erreichte Strukturentwicklung

Eine Nachhaltigkeit der geschaffenen Strukturen ist durch die Schaffung und Besetzung einschlägiger Professuren mit zugehöriger Ausstattung an den beteiligten Universitäten gewährleistet. Für eine nachhaltige Sicherung der im LOEWE-Schwerpunkt geschaffenen Forschungsansätze, Infrastruktur und Synergien stehen darüber hinaus vor allem zwei Verbundprojekte, die teils infrastrukturell (das seit 2014 vom BMBF geförderte eHumanities-Zentrum CEDIFOR), teils didaktisch (das ab 2015 von der DFG geförderte Graduiertenkolleg GRK 1994 AIPHES) ausgerichtet sind.

## Erreichte Bedeutung/Stellung im Themen-/Forschungsfeld

Der LOEWE-Schwerpunkt *Digital Humanities* hat den Raum Frankfurt/Darmstadt als Forschungszentrum in diesem Bereich national etabliert. Durch die Durchführung der internationalen Konferenz *Historical Corpora 2012*, deren Tagungsband im März 2015 erschienen ist, sowie durch die Vorstellung der Projekte auf vielen Tagungen und in zahlreichen Fachveröffentlichungen ist der Schwerpunkt auch international in Erscheinung getreten. Hieran knüpft insbesondere

das vom BMBF geförderte, im Dezember etablierte eHumanities-Zentrum *CEDIFOR* an, an dem neben der Goethe-Universität und der TU Darmstadt auch das Deutsche Institut für Internationale Pädagogische Forschung (DIPF) in Frankfurt beteiligt ist. *CEDIFOR* nutzt das erworbene Know-how und die aufgebaute Infrastruktur zur Beratung, Betreuung und gemeinsamen Durchführung geistes-, sozial- und bildungswissenschaftlicher Digital-Humanities-Projekte. Es will dabei auch Fachrichtungen erschließen, die diesen Methoden noch fernstehen, und soll den Standort Rhein-Main dauerhaft als Digital-Humanities-Zentrum etablieren. Auch das DFG-geförderte Graduiertenkolleg GRK 1994: *Adaptive Preparation of Information from Heterogeneous Sources (AIPHES)*, das von der TU Darmstadt in Zusammenarbeit mit der Universität Heidelberg und dem Heidelberger Institut für Theoretische Studien getragen wird, baut auf der Erfahrung des Schwerpunkts auf und wird den Bereich der *Digital Humanities* durch die Ausbildung junger, qualifizierter Wissenschaftler weiterentwickeln.

## Weitere Informationsmöglichkeiten

- <http://www.digital-humanities-hessen.de/>  
Kurzform: <http://www.dhhe.de/>  
Homepage des LOEWE-Schwerpunkts mit Informationen zu den Arbeitsbereichen und Einzelprojekten
- <https://www.cedifor.de/de/cedifor/>  
Homepage des auf dem LOEWE-Schwerpunkt aufbauenden Zentrums CEDIFOR
- <https://www.aiphes.tu-darmstadt.de/de/aiphes/>  
Homepage des auf dem LOEWE-Schwerpunkt aufbauenden Graduiertenkollegs AIPHES
- <http://www.comphistsem.org/home.html>  
Homepage des Projektteils „Historische Semantik“

## Zahlen und Fakten

Förderzeitraum	01.01.2011 – 31.12.2014	Bemerkungen
bis Ende des Förderzeitraums verausgabte LOEWE-Mittel	4.528.388 Euro	
bis Ende des Förderzeitraums verausgabte Drittmittel	3.224.082 Euro	
eingeworbene Drittmittel	14.793.407 Euro	längste Laufzeiten bis 2020
Anzahl der beteiligten Personen	12 ProfessorInnen 34 wissenschaftliche Mitarbeiterinnen 3 technisch-administrative MitarbeiterInnen	
Anzahl an innerhalb des Förderzeitraums abgeschlossenen Promotionen	5	
Anzahl an Veröffentlichungen in Fachzeitschriften innerhalb des Förderzeitraums	121	4 Buchveröffentlichungen als Autor/ Herausgeber
Anzahl an Konferenzbeiträgen innerhalb des Förderzeitraums	85	
Anzahl an innerhalb des Förderzeitraums zugeteilten Patenten	0	

## Wichtigste Meilensteine des Projekts



Im Rahmen des Festivals der Wissenschaften präsentierte sich der Schwerpunkt am 29. Mai 2012 der Öffentlichkeit mit einem Tag der offenen Tür. Hier eine Aufnahme der Posterpräsentation.



Links das Poster zur internationalen Konferenz „Historical Corpora 2012“, die der LOEWE-Schwerpunkt im Dezember 2012 ausgerichtet hat, rechts daneben der 2015 erschienene Tagungsband.



Wenn ich sage, dass ich zur „Bank“ gehe, will ich mich dann hinsetzen oder Geld abheben? Der Sinn einer Aussage erschließt sich oft erst im Kontext. C. Bieermann, Juniorprofessor im LOEWE-Schwerpunkt Digital Humanities, entwickelt zusammen mit seinem Team eine Software, mit der Computer selbstständig lernen, Sprache zu „verstehen“: Indem sie aus großen Textmengen Sprachstrukturen erfassen, „lernen“ die Maschinen, Wörter aus ihrem Zusammenhang heraus richtig zu deuten. Im Februar 2014 erhielt Bieermann dafür einen mit 50.000 US-\$ dotierten IBM Shared University Research Award. Bereits im November 2013 war ihm der Adolf-Messer-Preis mit 50.000 € verliehen worden. Die Preisgelder fließen in die Forschung, unter anderem in einen neuen Rechner, der große Datenmengen verarbeiten kann. ([http://www.proloewe.de/files/ausgabe\\_01.2014\\_-\\_zur\\_ansicht.pdf](http://www.proloewe.de/files/ausgabe_01.2014_-_zur_ansicht.pdf))



## Kurzvorstellung der beteiligten Hochschulen und Forschungsinstitute

### Goethe-Universität Frankfurt am Main

<http://www.uni-frankfurt.de/>

Die Goethe-Universität ist eine forschungsstarke Hochschule in der europäischen Finanzmetropole Frankfurt. 1914 gegründet mit rein privaten Mitteln von freiheitlich orientierten Frankfurter Bürgerinnen und Bürgern fühlt sie sich als Bürgeruniversität bis heute dem Motto „Wissenschaft für die Gesellschaft“ in Forschung und Lehre verpflichtet. Viele der Frauen und Männer der ersten Stunde waren jüdische Stifter. Am 1. Januar 2008 gewann die Goethe-Universität mit der Rückkehr zu ihren historischen Wurzeln als Stiftungsuniversität ein einzigartiges Maß an Eigenständigkeit. Heute ist sie eine der zehn drittmittelstärksten und drei größten Universitäten Deutschlands mit drei Exzellenzclustern in Medizin, Lebenswissenschaften sowie Geisteswissenschaften.



### Technische Universität Darmstadt

<http://www.tu-darmstadt.de/>

Die TU Darmstadt zählt zu den führenden Technischen Universitäten in Deutschland. Ihre rund 300 Professorinnen und Professoren, 4.230 wissenschaftlichen und administrativ-technischen Mitarbeiterinnen und Mitarbeiter sowie rund 26.000 Studierenden widmen sich entscheidenden Zukunftsfeldern wie Thermofluidodynamik und Verbrennungstechnologie, moderne Materialien, Teilchenstrahlen und Materie, integrierte Produktionstechnologie sowie Future Internet.



### Freies Deutsches Hochstift/Frankfurter Goethe-Museum

<http://www.goethehaus-frankfurt.de/freies-deutsches-hochstift>

Das Freie Deutsche Hochstift ist eine gemeinnützige Forschungsinstitution und eines der ältesten Kulturinstitute Deutschlands. Zu ihm gehören Goethes Elternhaus sowie das Goethe-Museum und eine Gemäldegalerie der Goethezeit und der Romantik. Das Hochstift verfügt über bedeutende Kunstsammlungen, eine umfangreiche Handschriftensammlung und eine Forschungsbibliothek. Den Schwerpunkt der Forschungsarbeit bilden derzeit drei historisch-kritische Editionen: zu Clemens Brentano, dessen Nachlass hier aufbewahrt und erschlossen wird, zu Hugo von Hofmannsthal auf der Grundlage der umfangreichen Sammlung von Werkhandschriften und Briefen sowie zu Goethes Faust, eine Ausgabe, die den reichen Bestand der Goetheschen Arbeitshandschriften zu Faust umfassend erschließen wird.



HESSEN



Das Forschungsförderungsprogramm LOEWE ist eine Förderinitiative des Hessischen Ministeriums für Wissenschaft und Kunst.

## Impressum

### Herausgeber:

Hessisches Ministerium für Wissenschaft und Kunst  
Rheinstraße 23 – 25  
65185 Wiesbaden

### Inhalt:

LOEWE-Schwerpunkt *Digital Humanities* –  
Integrierte Aufbereitung und Auswertung textbasierter Corpora

### Redaktion:

LOEWE-Geschäftsstelle im  
Hessischen Ministerium für Wissenschaft und Kunst

### Layout:

Christiane Freitag, Idstein

### Fotos:

LOEWE-Schwerpunkt *Digital Humanities* –  
Integrierte Aufbereitung und Auswertung textbasierter Corpora  
Titel: © Jost Gippert; S. 3 Abb. 1: © Beate Meichsner;  
S. 4: © Ralf Gehrke; S. 5 Abb. 1 und 3: © ProLOEWE;  
S. 5 Illustration: © G. Abrami, M. Freiberg, A. Mehler;  
S. 7 Abb. 1: © Ralf Gehrke; S. 7 Abb. 4: © André Kind/TU Darmstadt