



<https://publications.dainst.org>

---

# iDAI.publications

---

ELEKTRONISCHE PUBLIKATIONEN DES  
DEUTSCHEN ARCHÄOLOGISCHEN INSTITUTS

Dies ist ein digitaler Sonderdruck des Beitrags / This is a digital offprint of the article

Sabine Thänert – Marina Unger

## **Gelehrte, Ausgräber und Kunsthändler. Die Korrespondenz des Istituto di Corrispondenza Archeologica als Wissensquelle und Netzwerkindikator. Die Arbeiten der Jahre 2017 und 2018**

aus / from

### **e-Forschungsberichte**

Ausgabe / Issue

Seite / Page **1–7**

<https://publications.dainst.org/journals/efb/2175/6588> • urn:nbn:de:0048-journals.efb-2019-0-p1-7-v6588.4

Verantwortliche Redaktion / Publishing editor

**Redaktion e-Jahresberichte und e-Forschungsberichte | Deutsches Archäologisches Institut**

Weitere Informationen unter / For further information see <https://publications.dainst.org/journals/efb>

ISSN der Online-Ausgabe / ISSN of the online edition

ISSN der gedruckten Ausgabe / ISSN of the printed edition

Redaktion und Satz / **Annika Busching (jahresbericht@dainst.de)**

**Gestalterisches Konzept: Hawemann & Mosch**

**Länderkarten: © 2017 www.mapbox.com**

**©2019 Deutsches Archäologisches Institut**

Deutsches Archäologisches Institut, Zentrale, Podbielskiallee 69–71, 14195 Berlin, Tel: +49 30 187711-0

Email: [info@dainst.de](mailto:info@dainst.de) / Web: [dainst.org](http://dainst.org)

**Nutzungsbedingungen:** Die e-Forschungsberichte 2019-0 des Deutschen Archäologischen Instituts stehen unter der Creative-Commons-Lizenz Namensnennung – Nicht kommerziell – Keine Bearbeitungen 4.0 International. Um eine Kopie dieser Lizenz zu sehen, besuchen Sie bitte <http://creativecommons.org/licenses/by-nc-nd/4.0/>

**Terms of use:** The Research E-Papers 2019 of the Deutsches Archäologisches Institut is published under the Creative-Commons-Licence BY – NC – ND 4.0 International. To see a copy of this licence visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>



## GELEHRTE, AUSGRÄBER UND KUNSTHÄNDLER

Die Korrespondenz des *Instituto di Corrispondenza Archeologica* als Wissensquelle und Netzwerkindikator

Die Arbeiten der Jahre 2017 und 2018

**Zentrale Berlin und Abteilung Rom des DAI**  
von Sabine Thänert und Marina Unger



e-FORSCHUNGSBERICHTE DES DAI 2019 · Faszikel 1

*The aim of the international Instituto di Corrispondenza Archeologica, founded in 1829, was the collection and publication of archaeological discoveries. A network of correspondents situated all over Europe reported from their respective territories and collections. Their autographs are on one hand all-important sources of archaeological knowledge; on the other hand, they are evidence of the dense network formed around the Institute. In a first step the letters are indexed, digitized and made available online in an open access environment for the research community. The digital data is then analyzed using network and text-based approaches in order to visualize the network.*

„Gelehrte, Ausgräber und Kunsthändler: Die Korrespondenz des Instituto di Corrispondenza Archeologica als Wissensquelle und Netzwerkindikator“ heißt [das von der DFG geförderte Projekt](#) <sup>↗</sup>, welches im Frühjahr 2017 an der Abteilung Rom und an der Zentrale des DAI in Berlin angelaufen ist ([Projektseite auf der Homepage des DAI](#) <sup>↗</sup> und [iDAI.objects/Arachne](#) <sup>↗</sup>).

**Kooperationspartner:** CoDArchLab (Köln).

**Förderung:** Deutsche Forschungsgemeinschaft.

**Leitung des Projektes:** R. Förtsch (DAI Berlin), O. Dally (DAI Rom).

**Team:** F. Mambrini, Sh. Moshfegh Nia, M. Röhl, S. Schuth, S. Thänert, M. Unger, R. Wiedmann; Zuarbeit von H. Blattmann, S. Hohl, S. Wolter.

Das 1829 in Rom als privater, internationaler Verein offiziell ins Leben gerufene *Instituto di Corrispondenza Archeologica* oder Institut Archäologischer Korrespondenz setzte – wie der Titel impliziert – auf den kooperativen Gedanken des Informationsaustauschs durch Korrespondenz mit dem Institut als Mittelpunkt: Durch ein weitgespanntes Netzwerk von Archäologen, Künstlern und Archäologieinteressierten aller Art sollten aktuelle Entdeckungen, Ausgrabungen und Forschungen zentral gesammelt und publiziert werden. Jeder, der archäologische Nachrichten liefern konnte, eignete sich prinzipiell als korrespondierendes Mitglied. Die bewusst gewählte und explizite Internationalität des Vereins sowie die liberale Mitgliedschaftspolitik sind für die frühe Phase des Instituts kennzeichnend. Die spätere Umwandlung des Instituts in eine Preußische Anstalt und die sukzessive Gründung anderer Nationaler Schulen in Rom hat die Wahrnehmung jedoch entscheidend geformt. Internationale Diplomaten-, Gelehrten- und Künstlerkreise sowie Bildungs- und Grand-Tour-Reisende prägten dabei aber nicht nur das päpstliche Rom des frühen 19. Jahrhunderts, sondern auch das Institut, das dort seinen Ursprung nahm.

Dennoch bestand von Anfang an eine Verbindung nach Berlin. Zunächst eher lose durch die Schirmherrschaft des Kronprinzen, doch spätestens mit dem Wechsel Eduard Gerhards (1795–1867) an das Königliche Museum auch eine personelle: Gerhard, als einer der Gründer, behielt weiterhin seine zentrale Rolle am Institut. Mit der Zeit verfestigte sich die Bindung an Preußen, das sukzessive immer mehr Anteile der Finanzierung übernahm. Zugleich etablierte sich in Berlin das Gremium der *Direzione Centrale* als feste Einrichtung, das bis heute seine Funktion als Zentralkommission ausübt.

### Bestände

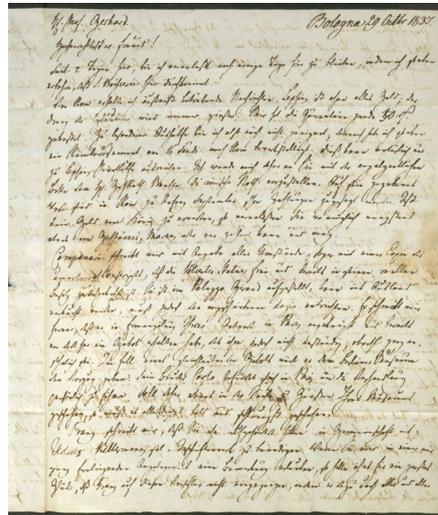
Die namensgebende Korrespondenz des *Instituto di Corrispondenza* wurde in Rom von Beginn an in einem gesonderten Archiv systematisch abgelegt. Dieses Charakteristikum behielt das Institut bis Anfang des 20. Jahrhunderts bei, entsprechend erstreckt sich die Laufzeit dieses Archivbestandes – der sog. Gelehrtenbriefe – von ca. 1829 bis ca. 1915. Es handelt sich heute um einen geschlossenen Bestand im [Archiv des DAI Rom](#) <sup>↗</sup>.

In Berlin wurde auf die Initiative des damaligen Institutspräsidenten Gerhard Rodenwaldt (1886–1945) zum Hundertjahrjubiläum des Instituts das „Archiv zur Geschichte des Instituts und der Archäologie“ (heute [Archiv der Zentrale](#) <sup>↗</sup>) gegründet, in dem die Nachlässe namhafter Archäologen aufbewahrt werden. Die älteren Nachlässe gelangten durch Erwerb oder Schenkung der Erben ins neu gegründete Archiv. Inzwischen umfasst das laufende Archiv um 250 kleinere und größere Nachlässe insbesondere von Archäologen und Bauforschern. Für das Projekt wurden der Briefnachlass von Eduard Gerhard sowie die Nachlässe von Otto Jahn (1813–1869), Friedrich Gottlieb Welcker (1784–1868) und Karl Witte (1800–1883) ausgewählt. Außerdem wurden einzelne Briefe aus der sog. Autographensammlung, einem weiteren Bestand des Berliner Archivs, in das Projekt integriert, beispielsweise Briefe an Christian Karl Josias von Bunsen (1791–1860) und Theodor Panofka (1800–1858).

Diese beiden Bestände stehen im Mittelpunkt des laufenden DFG-Projekts, wobei für die erste Finanzierungsphase die Korrespondenz von der Gründungszeit des Instituts bis ca. 1860 ausgewählt wurde. Die frühesten für das Projekt relevanten Briefe im Archiv der Zentrale datieren in die ersten Jahrzehnte des 19. Jahrhunderts. Die Korrespondenzen einzelner Autoren werden vollständig bearbeitet, auch wenn die Laufzeit die gesetzte Grenze überschreitet, sodass sowohl Briefe vor 1829 als auch bereits einige Briefe aus den 1870er- und 1880er-Jahren erschlossen und digitalisiert werden.

### Ziele und Stand der Arbeiten

Ziele des Projekts liegen einerseits in der digitalen Erschließung, Metadaten-generierung und Digitalisierung der Bestände (Abb. 1), andererseits geht es um die Integration dieser Daten in die digitalen Systeme des DAI (Abb. 2) sowie um die Vernetzung mit anderen Systemen und Normdaten und um die Entwicklung neuer Visualisierungsanwendungen und Präsentationsformen. Mit der Erschließung und Digitalisierung beider Bestände werden zudem einige Briefwechsel virtuell zusammengeführt sowie der Forschung vereint und komplettiert online zur Verfügung gestellt.



1 Digitalisat des Briefes von Emil Braun an Eduard Gerhard A-II-BraE-GerE-131. (Abb.: DAI)

### Erschließung und Digitalisierung

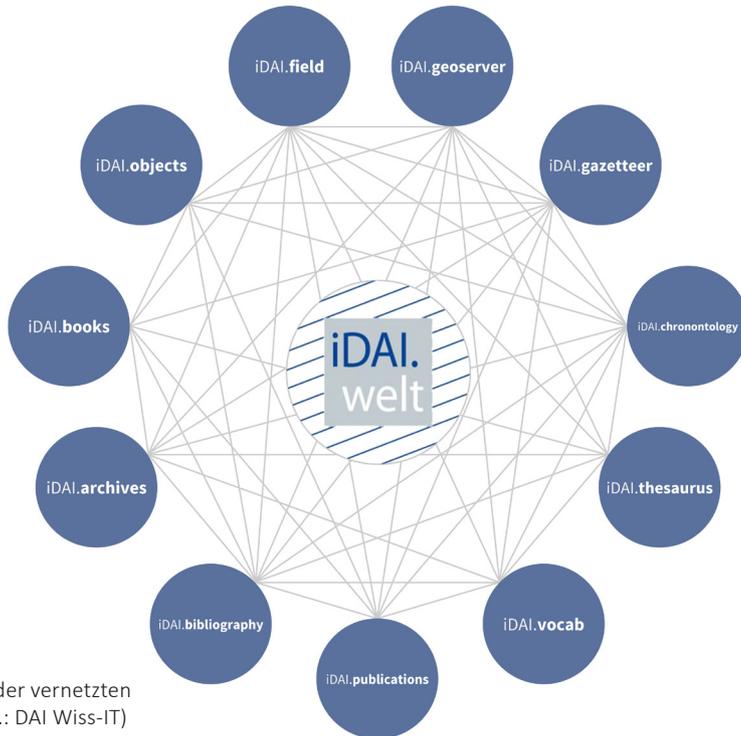
Die Digitalisate der einzelnen Briefe werden in die Objektdatenbank [iDAI.objects/Arachne](#) <sup>↗</sup> eingespielt (Abb. 3). Dort werden sie mit den erhobenen Metadaten verbunden, die in [Kalliope](#) <sup>↗</sup>, dem überregionalen Verbund und zugleich nationalen Nachweisinstrument für Nachlässe, Autographen und Verlagsarchive generiert werden. Zusätzlich erfolgt der Nachweis aller Briefmetadaten auch in [iDAI.bibliography/ZENON](#) <sup>↗</sup>.

Der ausgewählte Berliner Bestand von etwa 1400 Briefen ist bereits vollständig erschlossen und digitalisiert, nur die Onlinestellung der Digitalisate befindet sich im Bearbeitungsprozess. Die Erschließung und Digitalisierung des für die frühe Zeitperiode deutlich größeren römischen Bestandes dauert noch an. Von den in dieser Projektphase geplanten etwa 12.500 Briefen wurden bereits 6825 Einzelbriefe von über 100 Autoren an insgesamt etwa 175 Adressaten erfasst (Stand August 2018). In beiden Kategorien (Empfänger/Absender) liegen Wilhelm Henzen (1816–1887), Eduard Gerhard sowie Emil Braun (1809–1856) zahlenmäßig mit deutlichem Abstand vorn.

Gerhard gehörte zu den Initiatoren und war der erste Sekretär des Instituts, also geschäftsführender Direktor im modernen Sprachgebrauch. 1833 verließ er Rom, um Archäologe des Königlichen Museums zu Berlin zu werden. Braun wurde in dieser Zeit zu seinem Assistenten und stieg in wenigen Jahren vom Unterarchivar und Bibliothekar zum Sekretär auf. Er leitete die Geschicke des Instituts in Rom bis zu seinem plötzlichen Tod 1856. Seit 1843 stand ihm Henzen als zweiter Sekretär zur Seite, der 1856 die Institutsleitung für die folgenden 30 Jahre übernahm. Damit prägten diese drei Gelehrten das Institut im 19. Jahrhundert wie kein anderer – eine Tatsache, die sich auch in dem Briefbestand deutlich widerspiegelt (Abb. 4).

### Text- und Data-Mining

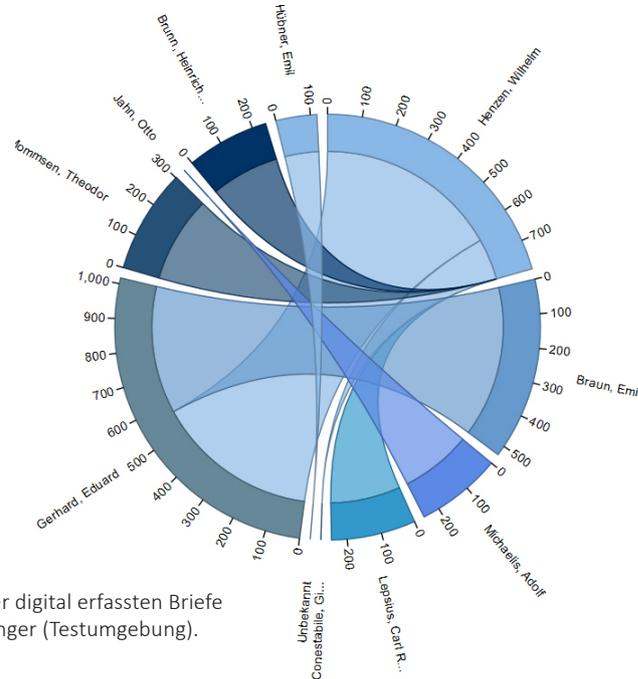
Zu einem Teil der im DAI Rom befindlichen Briefe wurden in den 1970er- und 1980er-Jahren maschinenschriftliche Abschriften angefertigt, **die bereits digitalisiert und mit OCR vorliegen** <sup>↗</sup>. Es handelt sich um etwa 70 Konvolute mit insgesamt ca. 20.000 Seiten, wobei als Konvolut ein DIN-A4-Ordner zu verstehen ist. Da die Texterkennung an Vorlagen unterschiedlicher Qualität



2 Visualisierung der vernetzten iDAI.welt. (Abb.: DAI Wiss-IT)

The screenshot shows the iDAI.objects.arachne interface. The main content area displays the record for 'Brief von Emil Braun an Eduard Gerhard'. It includes a map of the location (Bologna, Italy), a thumbnail of the handwritten letter, and various metadata fields such as 'Autor', 'Titel', 'Publikationsort', 'Physische Beschreibung', 'BibId', and 'Download'. The interface is in German and features a search bar at the top and navigation links.

- 3 Datensatz des Briefes von Emil Braun an Eduard Gerhard [A-II-BraE-GerE-131 in iDAI.objects](#)<sup>7</sup>. (Abb.: DAI).

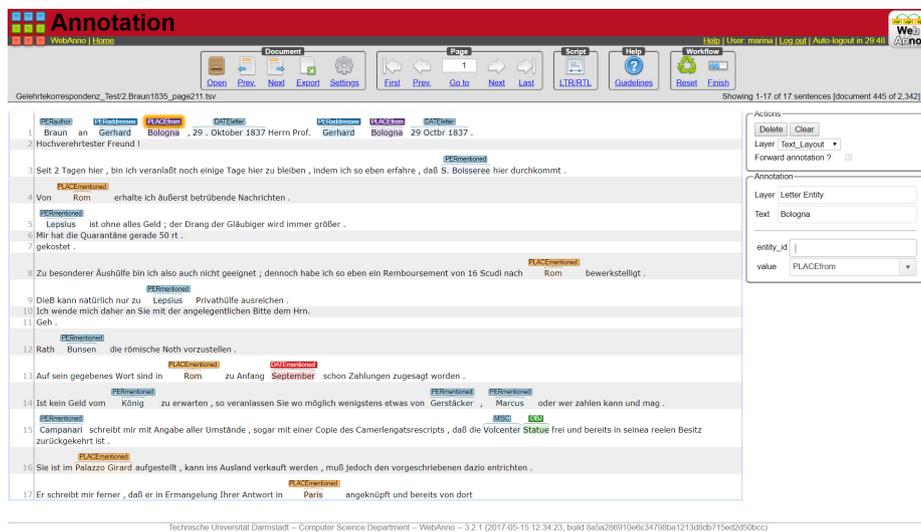


- 4 Chord-Diagramm der bisher digital erfassten Briefe nach Absender und Empfänger (Testumgebung). (Abb.: DAI)

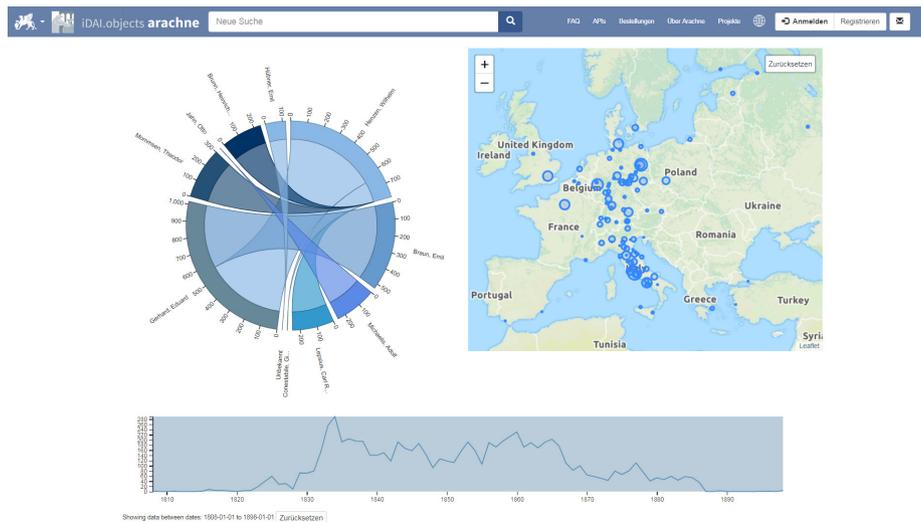
durchgeführt wurde, variiert auch die Qualität der OCR-Texte. Einige enthalten z. B. handschriftliche Korrekturen, die die maschinelle Lesbarkeit erschweren.

Mit Methoden des *Machine Learning* und der *Named Entity Recognition* (NER) werden die Texte weiterverarbeitet. Die maschinenschriftlichen Texte werden bis auf Wortebene in ihre Bestandteile zerlegt. Anhand verschiedener Kriterien können Wortarten und Syntax des Textes analysiert werden. Diese dienen wiederum als Kriterien zur Identifizierung von Eigennamen, also genannter Personen, Orte und Organisationen, aber auch chronologischer Daten und archäologischer Objekte. Dabei kommt ein Verfahren des Maschinellen Lernens, sog. *Conditional Random Fields*, zum Einsatz, das auf Basis manuell annotierter Trainingsdaten ein automatisiertes Tagging weiterer Texte ermöglicht.

Schwierigkeiten bereiten veraltete Schreibweisen des 19. Jahrhunderts sowie die retrodigitalisierten PDFs, in denen die Text- und Seitenstruktur nicht abgebildet ist und somit Textbestandteile ebenfalls mithilfe der Analysetools erkannt werden mussten. Dies führt zur Aufsplitterung der Kategorien wie Person, Ort und Organisation in Unterkategorien, die die Rolle des Wortes im gegebenen Text widerspiegeln – beispielsweise den Namen oder Ort des Absenders bzw. Empfängers oder aber im Text genannte dritte Personen bzw. Orte. Auch sind auf den Abschriften Daten mit unterschiedlichen Funktionen angegeben, sodass auch hier zwischen Unterkategorien des Feldes „Datum“ unterschieden werden musste: Briefdatum, Eingangsdatum, Poststempel etc. Diese Prämissen haben das Trainieren des Modells für die NER erschwert und mehrere Trainingsrunden erfordert, bis Ergebnisse mit einer akzeptablen Fehlerquote erzielt werden konnten. Für das Training wurden an einigen Konvoluten die Ergebnisse mit [Webanno](#)<sup>7</sup>, einem universellen webbasierten Annotationstool, manuell überprüft und ggf. korrigiert (Abb. 5). Im nächsten Schritt werden die Nennungen von Personen und Orten mit Identifiern aus DAI-eigenen, wie dem [iDAI.gazetteer](#), und externen Systemen, wie der Gemeinsamen Normdatei der Bibliotheken (GND), angereichert. In diesem Zusammenhang mussten und müssen eine Vielzahl von Personen, aber auch Ortsnamen zunächst identifiziert werden, um ggf. in die Normdatenbanken aufgenommen zu werden.



5 Annotationen in der Abschrift des Briefes von Emil Braun an Eduard Gerhard A-II-BraE-GeE-131 (WebAnno). (Abb.: DAI)



6 Visualisierungen der bisher digital erfassten Briefe (Testumgebung). (Abb.: DAI).

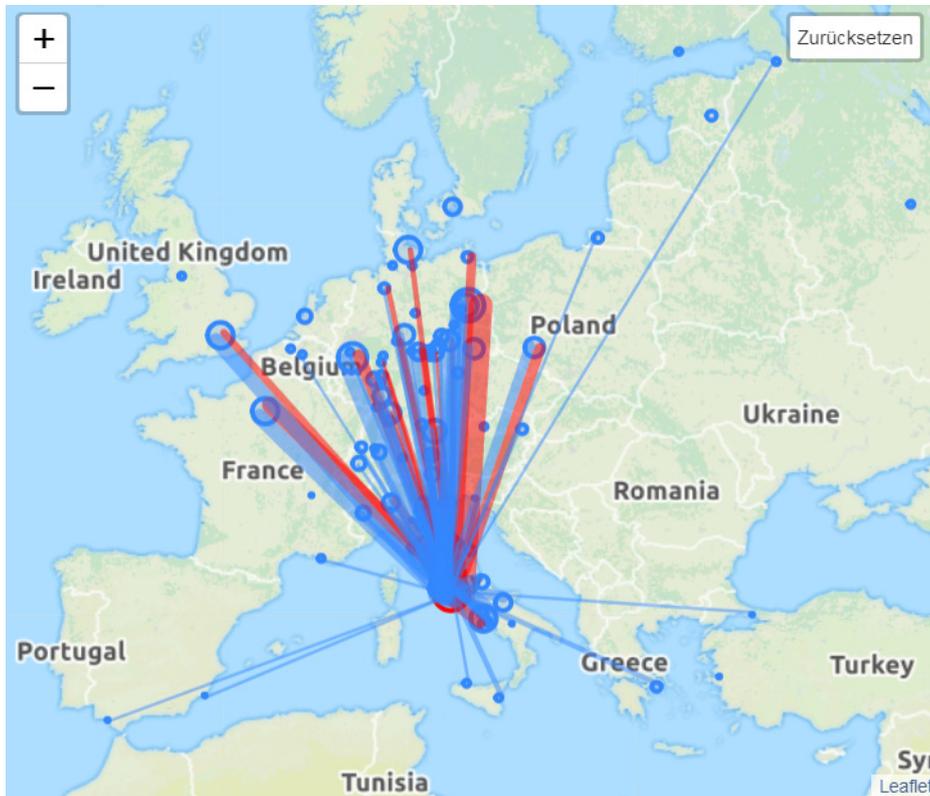
Eine besondere Herausforderung stellen Nennungen archäologischer Objekte und Monumente dar, da Objektbezeichnungen oft keinen klaren Kriterien unterliegen wie Personen- oder Ortsnamen und daher im Text schwieriger maschinell identifiziert werden können. Eine Möglichkeit für Texte in deutscher Sprache ist die Erkennung bestimmter Suffixe in Komposita, die mit archäologischen Gattungsbezeichnungen übereinstimmen, wie beispielsweise „-vase“, „-relief“ oder „-sarkophag“. Vielfach werden figürlich verzierte Objekte jedoch schlicht mit ikonographischen Begriffen bezeichnet, sie sind also für eine maschinelle Erkennung nicht als archäologische Objekte identifizierbar. Oft ist es schwierig, ein in den Briefen genanntes Artefakt eindeutig mit einem real existierenden archäologischen Objekt zu verbinden und mit einem Identifier zu versehen, wie es durch die Normdatenbanken bei Personen und Orten möglich ist.

Derzeit wird an einem Präsentationstool gearbeitet, in dem die mit Metadaten angereicherten Texte mit den entsprechenden Informationsressourcen inner- und außerhalb der iDAI.welt verbunden werden können. Bei den Nennungen archäologischer Objekte wird auf die Objektdatenbank des DAI iDAI.objects/Arachne verwiesen oder es werden Begriffe – auch als weiterführender Hinweis – aus dem iDAI.thesaurus herangezogen (vgl. [Thänert – Colombi – Vitt, eDAI-F 2017-1](#) <sup>7</sup>), sofern die Objekte nicht eindeutig identifizierbar sind. Auf diese Weise werden die Korrespondenzen auch in diese Ressource der iDAI.welt eingebunden.

### Netzwerk-Visualisierung

Die bei der Einzelbrieferschließung erhobenen Metadaten sollen in Form eines Netzwerks visualisiert werden. Zwar lassen sich in Kalliope bereits Netzwerke anzeigen, diese geben aber nur Auskunft über Verbindungen zwischen Personen, während beispielsweise Intensität der Briefwechsel oder die topographischen und chronologischen Informationen der Korrespondenzen darin nicht abgebildet werden können.

Um diese Desiderate zu erfüllen, wird derzeit an neuen Formen der Visualisierung für die Anwendungen der iDAI.welt gearbeitet, in denen u. a. die in den Meta- und Normdaten enthaltenen personellen, geographischen und

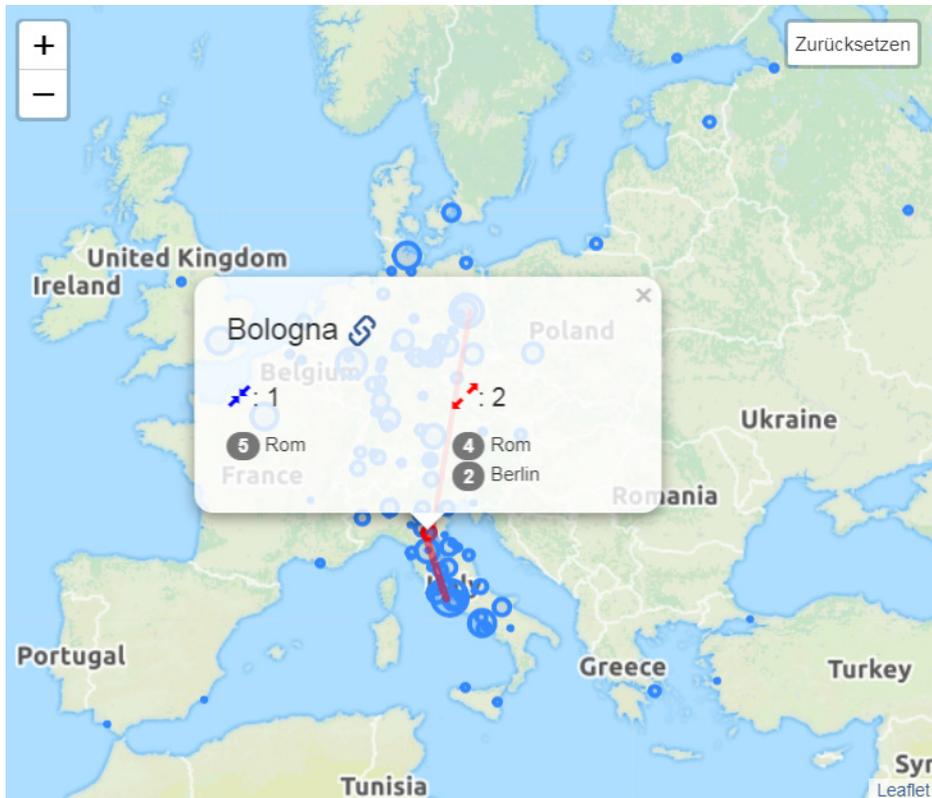


chronologischen Komponenten ausgewertet und dargestellt werden können. Das Präsentationstool befindet sich bereits im Teststadium und kann daher an dieser Stelle als „Werkstattbericht“ gezeigt werden (Abb. 6). Mittels der Koordinaten aus den Normdaten werden die Absender- und Empfängerorte kartiert, verschiedene Farben visualisieren die Richtung der Korrespondenz und die Stärke der Linien zeigt die Intensität an (Abb. 7). Die Anzahl der Briefe, die es zu einem Ort gibt (als Absende-/Empfängerort), bestimmt die Größe eines Kreises, der um den jeweiligen Ort dargestellt wird. Über ein Pop-Up zu dem ausgewählten Ort ist einerseits der Sprung in den iDAI.gazetteer möglich, andererseits sind die Anzahl der ein- und ausgehenden Briefe sowie ihre Absende- bzw. Destinationsorte angegeben, zu denen man auf der Karte weiter navigieren kann (Abb. 8). Weitere iDAI.welt-Systeme, insbesondere die Datensätze einzelner Briefdigitalisate oder ganzer Suchmengen in iDAI.objects/Arachne, sollen ebenfalls direkt angesteuert werden können.

Über eine interaktive Zeitleiste kann der in der Karte visualisierte Briefbestand chronologisch eingegrenzt werden. In der Gegenrichtung zeigt die Zeitleiste die Daten der auf der Karte ausgewählten Briefe/Städte an. Zur Eingrenzung der Suchergebnisse werden noch eine A-Z-Liste der Briefempfänger und -absender sowie Facettensuchen implementiert, die dynamisch aus den Daten erzeugt werden. Durch die Auswertung der verarbeiteten Briefmetadaten können verschiedene Fragestellungen mit den Komponenten Person, Ort und Zeit sichtbar gemacht werden.

Eine weitere Visualisierung in Form eines Chord-Diagramms, mit der die „Vielschreiber“ des Netzwerks im Bezug zueinander gestellt werden, kann hier ebenfalls im Entwicklungsstadium präsentiert werden (Abb. 4. 6). In dem kreisförmigen Graphen werden Absender und Empfänger der meisten Briefe miteinander verbunden, wobei jede Person eine eigene Farbe erhält. Außen am Kreis kann die Anzahl der jeweils geschriebenen/erhaltenen Briefe abgelesen werden. Die Verbindung zwischen den Korrespondenten erhält die Farbe desjenigen, dessen Briefe zahlenmäßig überwiegen. Mit einem Mouse-over kann eine Person oder ein einzelner Briefwechsel ausgewählt und die entsprechenden Informationen angezeigt werden. Auch diese

7 Kartenbasierte Visualisierung der bisher digital erfassten Briefe aus und nach Rom (Testumgebung). (Abb.: DAI)



Visualisierung kann mittels einer Auswahl auf der Karte oder Zeitleiste eingegrenzt werden, sodass beispielsweise die Korrespondenten einer bestimmten Stadt oder eines bestimmten Zeitraums angezeigt werden.

Mit diesem DFG-Projekt stellt das DAI nicht nur die einzigartigen Briefbestände und -nachlässe online weltweit zur Verfügung, die zu den besonders häufig angefragten Archivbeständen gehören, sondern ermöglicht mit der Analyse und visuellen Aufbereitung der Metadaten auch neue Zugangswege für die Forschung. Vorbehaltlich einer weiteren Förderung sollen im Rahmen des Projektes die genannten Bestände bis ins frühe 20. Jahrhundert weiter erschlossen und digitalisiert sowie die Visualisierungsmöglichkeiten verfeinert bzw. neue entwickelt werden, aus denen sich neue, spannende Forschungsfragen ergeben können.