

# Memorandum zum Stand des Projekts "Historische Räume in Texten und Karten" (Biondo-Projekt)

---

Günther Görz und Chiara Seidl, FAU und BHMPI

Stand: 31.12.2019

## Inhalt

1. Einleitung und Projekt-Überblick
2. Organisation und Datenerzeugung: Text
3. Organisation und Datenerzeugung: Bild
4. Semantische Modellierung
5. Implementation der WissKI-Plattform und Publikation von Linked Open Data
  - A.1 Ausgewählte Italien-Karten des 14. und 15.Jh. für das BIONDO-Projekt
  - A.2 Hinweise zur Benutzung von Recogito 2

## 1. Einleitung und Projekt-Überblick

---

In der [Projekt-Webseite](#) "Historische Räume in Texten und Karten" des Instituts wird das Projekt eingeführt und seine Ziele in allgemeiner Form beschrieben. Dem vorliegenden Bericht liegt die Annahme zugrunde, dass aus der Einführung in das Projekt präzise Forschungsfragen abgeleitet werden können, welche zu operationalisierten Hypothesen führen, die ihrerseits an Daten überprüft werden können. Im Folgenden geht es um den Daten-Aspekt des Projekts, d.h. die Generierung, Organisation, Erschließung, Speicherung und Publikation von Forschungsdaten.

Zugleich war durch den zugrunde gelegten theoretischen Rahmen (kunst-/historisch, kognitiv) festgelegt worden, welche Daten aus den verfügbaren Quellen in welcher Weise gewonnen werden sollten. Die Quellenauswahl wurde durch verschiedene Parameter geleitet: Verfügbarkeit und Aufwand zur Bereitstellung sowie die Qualität der digitalen Reproduktionen (v.a. Lesbarkeit) der Texte und Kartenbilder in digitaler Form.

Aus der Projekt-Definition haben wir für das Projekt eine Dreiphasen-Struktur entwickelt

- Analyse
- kognitiv-linguistische und historische Interpretation
- Synthese (Skizzenrekonstruktion)

## 1.1 Forschungsfragen

Die in diesem Bericht beschriebenen Arbeiten verstehen sich als Beitrag der "Digital" bzw. "Computational Humanities" zum Projekt und sind wegen der primär geographischen/kartographischen Ausrichtung auch in den Bereich der "Spatial Humanities" einzuordnen. Zu allgemeinen Forschungsfragen kann an dieser Stelle nur auf die mittlerweile recht umfangreiche einschlägige Literatur verwiesen werden. Eine Besonderheit unserer Arbeiten ist die klare kognitiv (-linguistische) Ausrichtung, die sie von früheren eher philologisch orientierten Studien unterscheidet (v.a. Clavuot: Biondos "Italia Illustrata" - Summa oder Neuschöpfung?, aber auch die Einleitung von Pontari in seiner Edition der "Italia Illustrata").

Da wir es vornehmlich mit geographischen/kartographischen Daten zu tun haben, waren - vor allem im vergangenen Jahr - die Arbeiten der geographischen Forschungsgruppe um Werner Kuhn (früher U. Münster, jetzt UCSB) und Simon Scheider (früher U. Münster, jetzt U. Utrecht) wegen ihres klaren Bezugs auf Kognition und Wissensrepräsentation/verarbeitung sehr einflussreich. Ohne dass wir an dieser Stelle auf eine detaillierte Diskussion eingehen können, seien Kuhns für allgemeine geographische Informationsverarbeitung formulierte [Forschungsfragen](#) kurz genannt, da sie grundsätzlich auch für unsere historisch ausgerichteten Arbeiten einen einschlägigen Rahmen abstecken:

1. What questions does spatial information answer?
2. How are hypotheses made spatial?
3. What is the ontology of observations in space and time?
4. What are spatial data workflows?
5. What makes reference systems operational?
6. What is a general theory of spatial referencing?
7. Can all spatial information be encoded as grids and graphs?
8. How does semantics without typing work?

Die für das Projekt gewählte Vorgehensweise wurde in verschiedenen Publikationen beschrieben; kompakt und für diesen Bericht hinreichend in Görz, Seidl, Thiering: Linked Biondo: Modelling Geographical Features in Renaissance Texts and Maps. (In: Boutoura et al. (Ed.): Proceedings 14th ICA Conference Digital Approaches to Cartographic Heritage, Thessaloniki, 2019; International Cartographic Association, Commission on Cartographic Heritage into the Digital)

## 1.2 Analytische Phase

Die Analysephase des Projekts bestand aus Datensammlung, Aufbereitung, elementarer Textstatistik, Konkordanzherstellung, morphosyntaktischem Tagging und Dependenzparsing sowie der Erstellung von geographischen Annotationen. Mit dem Werkzeug [Recogito](#) wurden die Toponyme in Biondos Latium-Buch sowie in 27 der ausgewählten zeitgenössischen Karten (zumeist Latium und Kampanien, in einigen Fällen sogar ganz Italien) annotiert. Die Auswahl der Karten umfasst genuine Italienkarten aus dem 14. und 15. Jh., Ptolemaeus-Karten (traditionelle und "novae") sowie Portolane. Weiterhin sollten die im Text ausgedrückten räumlichen Relationen nach kognitiv-linguistischen Kriterien annotiert werden, vor allem die Tripel *figure--spatial\_indicator--ground*. Diese Phase ist bis auf den letzten Schritt abgeschlossen und

wurde in einem Poster (Abb. 1) dokumentiert (präsentiert bei Linked Pasts 2018, Mainz, und der DHD-Konferenz 2019, Frankfurt).

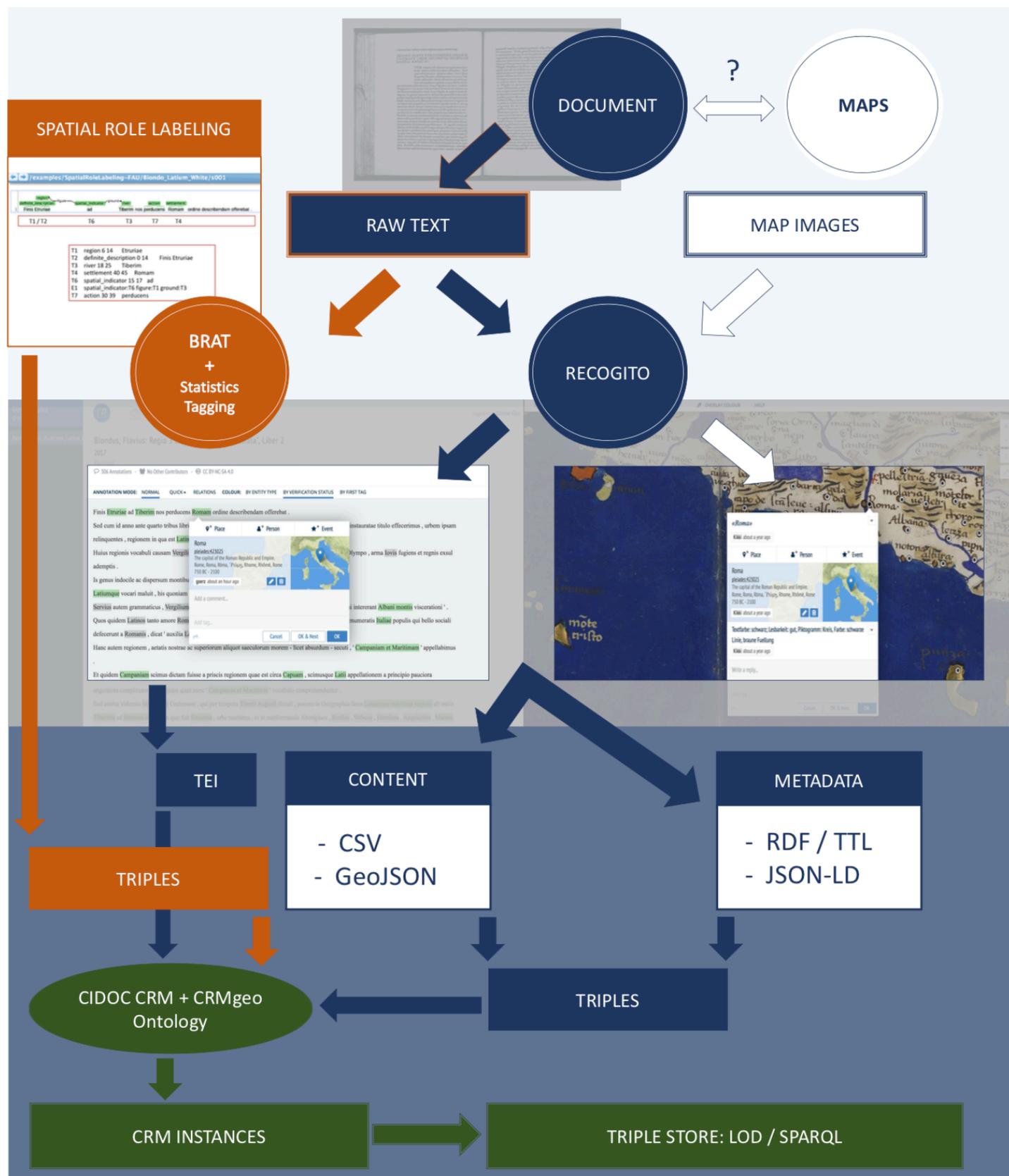


Abb. 1. Poster

Sämtliche bisher erarbeiteten Forschungsdaten (Text- und Karten-Annotationen) wurden mithilfe einer

Anwendungsontologie für historische Texte und Karten sowie annotierte Inhalte semantisch repräsentiert und in der Virtuellen Forschungsumgebung [WissKI](#) (s. 5.) gespeichert; damit stehen sie als Linked Open Data zur Publikation im Semantic Web zur Verfügung. Weitere Forschungsdaten, insbesondere die linguistischen Analysen, sind in einem Repository im [Drupal](#)-Rahmensystem von WissKI abgelegt.

### 1.3 Interpretationsphase

Im Zentrum des Projekts steht die kognitiv-linguistische und kunst-/historische Interpretation der Daten, d.h. vor allem des Textes, und dies im Vergleich zu zeitgenössischen Karten. Der Forschungsgegenstand sind also "kognitive Karten" im weitesten Sinn zur Beantwortung von Fragen nach dem historischen Verständnis des geographischen Raums und seines Wandels in der Vormoderne.

Entscheidend dabei ist eine Operationalisierung der Forschungsfragen, um diese als Fragen an die Datenauswertung richten zu können. Zugleich ist im Interesse der Projektorganisation eine Priorisierung der Forschungsfragen geboten. Hilfreich können standardisierte Anfrageformen im Sinne von (kulturhistorischen) "Query Patterns" sein (vgl. Constantopoulos et al.: Query Patterns: Foundations and Analysis. Technical Report AUEB/ISDB/TR/2008/01. Department of Informatics - Athens University of Economics and Business). Mit der Aufstellung von SPARQL-Anfragemustern in [Research Space](#) (British Museum, metaphacts) ist eine ähnliche Vorgehensweise in Arbeit.

Die Interpretationsphase ist nicht Gegenstand dieses Berichts. Erste Ergebnisse einer kognitiv-linguistischen Analyse liegen vor (Berthele/Thiering). Eine Abschätzung des Arbeitsaufwands kann erst gegeben werden, wenn eine Planung für diesen Projektteil vorliegt.

### 1.4 Synthetische Phase

Von Anfang an war eine Rekonstruktion von Kartenskizzen aus den Daten langfristig angedacht. Dazu wären folgende Arbeitsschritte erforderlich:

- Tripel-Extraktion aus dem Spatial Role Labeling (SpRL)
- Relationsgraphen-Erzeugung (s.u., vgl. Vasardani et al.: From Descriptions to Depictions: A Conceptual Framework. In: Tenbrink et al. (Eds.): COSIT 2013, LNCS 8116, pp. 299–319, 2013. Algorithmus 1, 2)
- Skizzen auf einem Gitternetz (s. Tobler: Cellular Geography. In: Gale and Olsson (eds.), Philosophy in Geography, 379-386, 1979)
  - Basis: Umrisskarte von Italien mit Koordinatenraster
  - Initiale Belegung aus Koordinaten
  - Belegung nach Tripeln des Relationsgraphen
  - Evaluation anhand der vorhandenen Geovisualisierung
- Nebenthema: Topologie und (historische) Regionengrenzen. Da nur in Ausnahmen (traditionelle Ptolemaeus-Karten) Regionengrenzen historisch gegeben sind, wäre ggf. eine künstliche Zerlegung in Gebiete als Voronoi-Diagramm (Delaunay-Triangulierung) zu erproben.

**Arbeits-Aufwand:** ca. 6 PM (= Personen-Monat, full time); eventuell ist für die anstehenden Programmierarbeiten mehr Zeit zu veranschlagen.

## 1.5 Stand der Dinge und Bedarf

### 1.5.1 Bestandsaufnahme

Mit den bisher erhobenen Daten sind erste Voraussetzungen für die Bearbeitung der Forschungsfragen und damit für den Wissenserwerb gegeben. Aktuell kann kaum von "Wissen" gesprochen werden, denn es liegen bisher vor allem Listen von Toponymen, also Daten vor, die auf einer modernen Karte bzw. mit Google Earth visualisiert werden können. Wenig aufwendig wäre es, die Nennungen der Toponyme in der im Text gegebenen Reihenfolge als Pfad abzubilden. Die Annotation von räumlichen Annotationen im lateinischen Text steht noch aus - die lückenhaften Annotationen der englischen Übersetzung sind nur von heuristischem Wert - und mittelfristig wäre auch zu überlegen, räumliche Relationen auf ausgewählten Karten zu annotieren. Ausgewertet ist damit allerdings noch lange nichts; ein erster Versuch wurde mit einem Vergleich von drei ptolemaeischen Tabulae Novae unternommen (Chiara Seidl).

Aktuell stünde eine Planung der nächsten Schritte an, d.h., was wie ausgewertet werden soll. Man kann z.B. aus den exportierten Annotationen (Tabellen) des Textes die Toponyme herausziehen und mit annotierten Karten vergleichen (s. 3.3.1). Im Prinzip besteht durch eine Untersuchung großer Datenbestände die Chance, vollständigere Ergebnisse als Clavuot zu erreichen, der nur Stichproben berichtet. Allerdings bezieht sich Clavuot auf den gesamten Text der "Italia Illustrata", während wir nur Daten über das Latium-Buch vorliegen haben. Stichproben zeigen, dass die verschiedenen Bücher der "Italia Illustrata" aber durchaus unterschiedliche Strukturen aufweisen. Wichtig für die Planung wäre, das Oberziel der Raumwahrnehmung und ihres Wandels in den Mittelpunkt zu stellen.

Sämtliche Rohdaten wurden zunächst auf dem Citrix-Server der Hertziana im sog. Laufwerk "V" im Verzeichnis "/Forschung/Biondo" abgelegt und die noch aktuellen werden nunmehr persistent im Drupal-Rahmen des *WissKI*-Systems gespeichert (s. 5.2; in *WissKI* erreichbar über die Startseite bzw. den Reiter "Inhalt"). Weiterhin sind Kartenbilder im Verzeichnis "/Forschung/Medioevo/Projekte/Cartografia/" abgelegt, wobei es sich vor allem um aktuell nicht mehr benötigte Bilddaten, z.B. aus der Akquisition, handelt. Das letztere Verzeichnis ist über einen Link von "/Forschung/Biondo" aus erreichbar; ein direkter Link "index-Bilder" verweist auf das Inhaltsverzeichnis der im Projekt bearbeiteten Kartenbilder.

Inhalt von "V:/Forschung/Biondo/":

#### 1. Textdaten und -Analysen des Latium-Buchs:

- Biondo\_OCR\_Dateien (alle Bücher, unkorrigiert)
- VERSION1\_2016\_White (englisch: Textanalysen; obsolet)
- VERSION2\_2017\_Geus (englisch: Textanalysen mit vielen Zwischenständen inkl. Dependenzparsing und brat-Annotationen, mit index-Datei ; ggf. lateinische Fassung zu bearbeiten)
- VERSION3\_2019\_White\_la (aktuelle lateinische Fassung, Zugriff auf die Text- und Analysedaten)

über eine index-Datei)

- *White\_vol2\_2016\_Italy\_Illuminated* (OCR, alle Bücher, Seiten nach Sprachen getrennt, unkorrigiert)

## 2. Linguistische Annotationen mit brat:

- *Biondo\_Latium\_White\_la.zip* (pro Satz eine Datei, Annotationsdateien noch leer)
- *Biondo\_Lazio\_Geus.zip* (englische Übersetzung, pro Satz eine Datei, mit Annotationsdateien, Annotationen unvollständig; erscheint redundant, s.o. 1. VERSION2...)
- *Biondo\_Lazio\_White\_en.zip* (englische Übersetzung, pro Satz eine Datei, mit Annotationsdateien, Annotationen unvollständig)

## 3. Kartenbilder

- *Karten\_15Jh* (Auswahl von Kartenbildern, redundant)
- *Karten-annotiert* (aktuelle Annotationsdaten der bearbeiteten Karten)
- *Karten-annotiert-ARCHIV* (Zwischenzustände der Annotation; obsolet)
- *Karten-bearbeitet-karo* (von Karolina Wagner mit Photoshop bearbeitete Kartenbilder)

## 4. ontologies: *crmgeo\_1.4.owl* und *hmap\_v1.owl*

## 5. Literatur

- Artikel (eigene Veröffentlichungen)
- *Flavio\_Biondo\_Italia\_illustrata\_Literatur* (Sekundärliteratur)
- *Generelle\_Literatur*
- *Geographiehistorische\_Texte*
- *Grundlagentexte\_Raumkognition*
- *Harley-Woodward.History\_of\_Cartography*

Zur Verarbeitung und Publikation der erarbeiteten Forschungsdaten wurde ein Konzept erarbeitet. Dieses sieht auch vor, die Daten als Linked Open Data (LOD) im Semantic Web zu publizieren, primär durch eine SPARQL-Schnittstelle und Triple-Store-Export. Hierzu wurde zunächst eine formale Ontologie *hmap* (s. 4.3) für historische Karten und (geographische) Texte und deren annotierte Inhalte auf der Basis von *CIDOC CRM* und *CRMgeo* (in OWL-DL) erarbeitet. Als kurzfristig verfügbare Plattform wurde die Virtuelle Forschungsumgebung [WissKI](#) gewählt (s. 5.). Die Metadaten für Karten, Atlanten, Texte, Bilder, Bildfolgen, Annotierte Orte und Annotierte Verbindungen wurden als Ontologie-Pfade (Tripelfolgen) modelliert und die Daten in die damit bereitgestellte semantische Datenbank eingegeben; die Metadaten händisch und die aus Recogito exportierten Annotationsdaten via Tabellenimport.

Diese Systemlösung ist kompatibel mit der am MPIWG erarbeiteten Datenarchitektur (Abb. 4), die drei grosse Komponenten vorsieht: Respository, Knowledge Graph, und local project environment. Die lokale Arbeitsumgebung besteht in unserem Fall aus Recogito und der WissKI-Datenbank; letztere implementiert auch Funktionalitäten des Repository. Da WissKI im Moment einen eigenen Triple-Store angeschlossen hat ([GraphDB](#), auch [blazegraph](#) wäre möglich), ist ein Anschluss an die Knowledge-Graph-Komponente unmittelbar möglich (Triple Store-Kopie oder Synchronisation).

**Da keine weiteren Ressourcen zur Verfügung stehen, sehen wir uns gezwungen, die Arbeiten auf**

dem hier beschriebenen Stand Ende 2019 einzustellen.

### **Wesentliche Schritte zur Vervollständigung der analytischen Phase wären**

- Bilinguale Glossierung des lateinischen Textes (Latium), s. 2.4.2
- Kognitiv-linguistische Annotation der räumlichen Relationen im lateinischen Text (Spatial Role Labeling, SpRL), s. 2.4.5
- Abgleich von SpRL und Abhängigkeitsstrukturen, s. 2.4.4
- Import der "Annotated Connections" (SpRL-Ergebnisse) in WissKI, s.4.3.3 und 5.1.9

### **1.5.2 Langfristige Perspektive 1: Erfassung weiterer Regionen**

Eine Ausweitung auf Kampanien war von Anfang an angedacht. Wenn über Biondo insgesamt belastbare Aussagen getroffen werden sollen, wäre eine Untersuchung auch der restlichen Bücher erforderlich. Es müssten dann auch die Auszeichnungen relevanter Karten vervollständigt werden. In den meisten Fällen ist bereits Süditalien ausgezeichnet, einige Karten sind komplett (u.a. B1.04 Cotton, B3.20N L20n, B3.23N L23n, B3.26N L26n, B4.05 Ulm1486, B4.06 und B4.06N Pirckheimer).

#### **Aufwand**

- für Aufbereitung und Recogito-Annotation der Texte (ohne Glossierung und kognitiv-linguistische Annotation): mindestens 12 PM
- für Kartenauszeichnung mindestens 12 PM

### **1.5.3 Langfristige Perspektive 2: Epistemischer Raum für historische Geodaten Italiens**

Von den vorhandenen Daten dieses Projekts und anderer Projekte des Instituts ausgehend wäre es möglich, ein Repository historischer geographischer Daten Italiens aufzubauen. Von Kurt Guckelsberger liegt das Angebot vor, sein Archiv kartometrischer Daten beizusteuern (dessen Zukunft ungesichert ist).

### **1.5.4 Räumliches (qualitatives) Schließen**

Da die Ontologien in einer beschreibungslogischen Sprache (OWL-DL) formuliert sind, bietet sich grundsätzlich an, auch eine Erweiterung der Möglichkeiten für räumliches Schließen zu untersuchen. Die grundsätzliche Funktionalität wurde bereits Anfang des letzten Jahrzehnts im Rahmen der Forschungen zum Behaim-Globus an der FAU gezeigt. Allerdings hat eine solche Studie für das vorliegende Projekt nur einen Sinn, wenn das räumliche Schließen tatsächlich einen Beitrag zur Lösung von Forschungsfragen liefern kann. Der grundsätzliche Rahmen ist durch die Idee der "Kognitiven Karten" gegeben; durch die logische Formalisierung ergibt sich eine Anhebung von der kognitiven auf die epistemische Ebene, in der eine Kombination mit symbolischem Schließen möglich wird.

Zu überarbeiten wäre in diesem Fall die formale (qualitative) Repräsentation von

- (abstrakten) Regionen (zellular im Sinne von Tobler, s. 1.4),

- ihren relativen Positionen,
- Orientierung,
- Entfernung.

Mit dem "Region Connection Calculus" (RCC-8), dessen Verbindung mit beschreibungslogischem Schließen (hybrides Schließen, z.B. "Pellet Spatial" für Topologie) gezeigt wurde, liegt eine elementare topologische Theorie des qualitativen räumlichen Schließens, vor, die direkt übernommen werden kann. Notwendig wären dann Erweiterungen für Entfernung und Orientierung, wozu auch einschlägige Literatur vorliegt.

**Arbeits-Aufwand:** Masterarbeit

## 2. Organisation und Datenerzeugung: Text

---

### 2.1 Textauswahl und Bereitstellung, Transkriptionsrichtlinien

Das Projekt sollte beginnen mit einer Bearbeitung des Latium-Buchs der "Italia Illustrata". Von den drei Textausgaben der "Italia Illustrata" (der auf Handschriften beruhenden lateinischen "Nationalausgabe" von Pontari, der auf der editio princeps beruhenden zweisprachigen von White und der auf dem Basler Druck von 1559 beruhenden zweisprachigen von Castner) war die die von White ausgewählt worden. Zunächst wurde der Text der relativ freien englischen Übersetzung bearbeitet, die jedoch wegen diverser Mängel und der zu großen Distanz zum lateinischen Text für die Forschung nur bedingt geeignet ist. Daher wurde von Klaus Geus die lateinische Textbasis des Latium-Buchs von White verbessert und mehr wörtlich ins Englische übersetzt, wobei diese Übersetzung zunächst für weitere Untersuchungen diene. Es zeigte sich dabei aber, dass es unumgänglich ist, den lateinischen Text zu bearbeiten. Wegen der langfristigen Perspektive, auch weitere Bücher der "Italia Illustrata" einzubeziehen, wurde 2018 entschieden, trotz diverser Mängel auf den lateinischen Text nach White zurückzugehen. Ein systematischer Vergleich mit der lateinischen Fassung von Geus wäre wünschenswert.

Die Texte wurden eingescannt und zahlreiche Scanfehler korrigiert. Die Bearbeitung der OCR-Scans des Textes in der Fassung von White erfolgte nach folgenden Richtlinien:

- Satzzeichen bleiben wie in der Edition gegeben; alle Satzzeichen werden durch einen Zwischenraum abgesetzt, am Satzende wird ein Zeilenvorschub eingefügt, so dass jeder Satz auf einer eigenen Zeile steht. Da Doppelpunkte in der Regel Zitate einleiten, gilt auch ein Doppelpunkt als Satzendezeichen.
- Anführungszeichen (Zitate) werden durch ` und ' (linksgeneigter Apostroph und senkrechter Apostroph) ersetzt. Besteht ein Zitat aus mehreren Sätzen, wird jeder Satz separat geklammert.
- Indirekte Rede wird nicht gesondert markiert.
- Alle Fußnoten und weitere Paratexte werden gelöscht.

Der lateinische Text liegt als "Rohtext"-Datei (utf-8) vor und diene in dieser Form als Eingabe für verschiedene Analysewerkzeuge (Wortlisten, Statistik, Konkordanz, Dependenzparsing, etc.) sowie die Auszeichnung der Latium-Toponyme durch Recogito. Für das "Spatial Role Labeling" wurde diese Datei

satzweise in insgesamt 626 Dateien zerlegt. Die englische Übersetzung ist damit synchronisiert, kann aber nur zu heuristischen Zwecken dienen; technische Analysen wurden mit dieser Fassung nicht durchgeführt (wohl aber mit inkompatiblen früheren Versionen, die sich jedoch von geringem Nutzen erwiesen).

Sollten künftig auch weitere Bücher der "Italia Illustrata" untersucht werden - zunächst war Kampanien in der Diskussion - müssen die OCR-Scans bearbeitet werden, u.a. sprachspezifische Seitentrennung, Korrektur, etc.

**Arbeits-Aufwand** inklusive Transkription nach den o.g. Regeln je Region (bei 14 Regionen) ca. 1 PM (s. 1.5.2).

## 2.2 Konzeption Spatial Role Labeling ("SpRL")

Als Grundlage der kognitiv-linguistischen Analyse des Textes dient die satzweise Auszeichnung der räumlichen Relationen zwischen Orten - gegeben durch Toponyme oder definite Beschreibungen - gemäß den von Martin Thiering vorgeschlagenen Parametern. In der einfachsten Form handelt es sich um Tripel der Form *figure (trajector) - spatial\_relation - ground (landmark)* wie in Abb. 2.

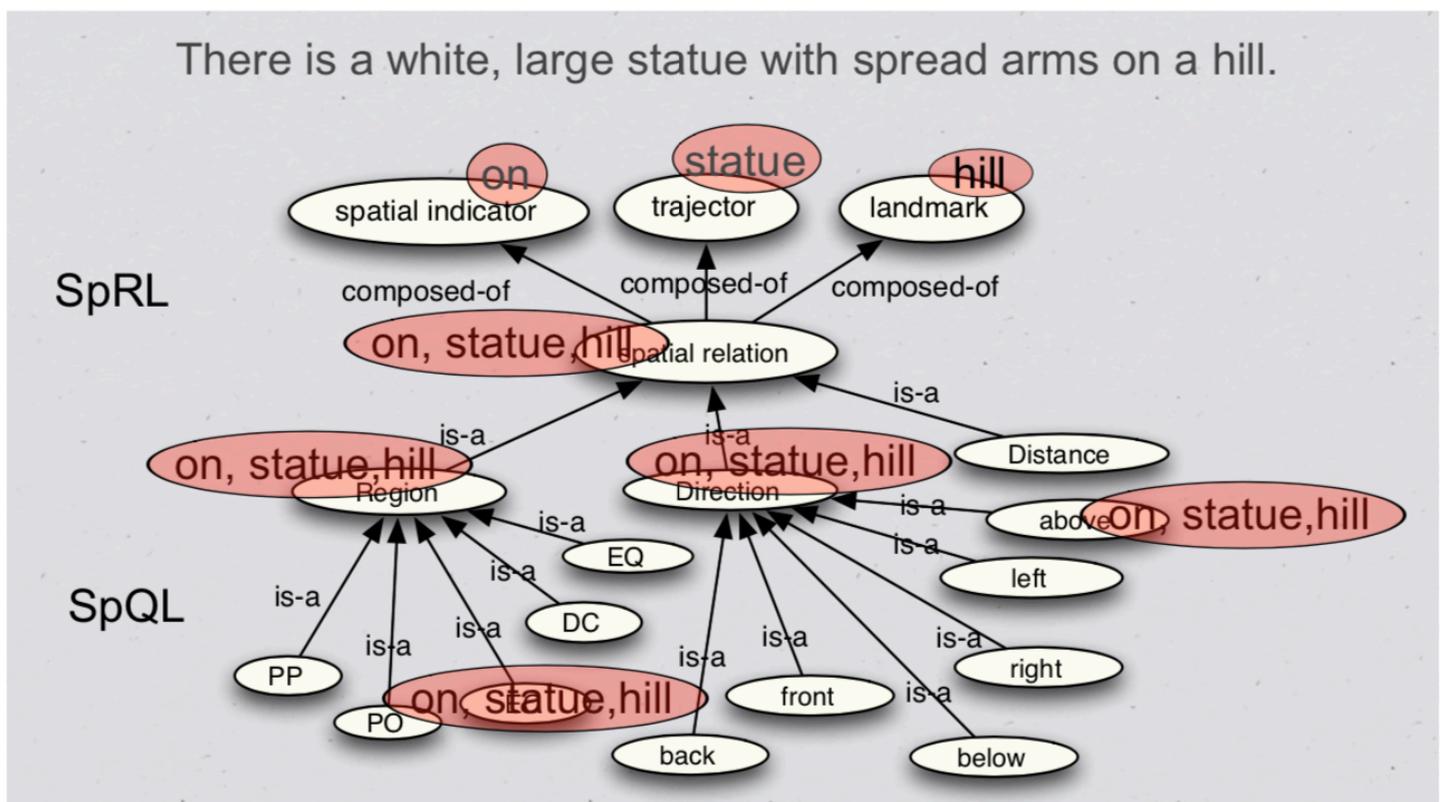


Abb. 2 Spatial Role Labeling

Die Konzeption ist abgeschlossen und zur Anwendung bereit (s.u. 2.4.5)

## 2.3 Konzeption Spatial Qualitative Relations ("SpQL")

In einem weiteren Schritt sollten dann Topologie (Regionen), Richtungen und Entfernungen annotiert werden. Hierfür ist noch konzeptionelle Arbeit zu leisten; es sind Definitionen zu erarbeiten, die

gleichermaßen auf Text und Karten(bild) anwendbar sind und die auch in die Basisontologie (4.2) und ggf. Anwendungsontologie *hmap* (4.3) einzubringen wären.

**Aufwand:** ca. 12 PM

## 2.4 Bearbeitung des Latium-Buchs nach White

Der lateinische Text nach White und seine englische Übersetzung stehen nach den o.g. Transkriptionsrichtlinien bereit. Früher erstellte Fassungen der englischen Übersetzungen, für die teilweise auch Textanalysen vorliegen, genügen diesen Richtlinien nicht; eine Umstellung wäre aus Vergleichsgründen nur für die lateinische Textfassung und die Übersetzung des Latium-Buchs von Geus jedoch wünschenswert. Zur Synchronisation müsste eine Umformatierung nach den o.g. Transkriptionsrichtlinien erfolgen.

**Aufwand** für die Umformatierung: 0.5 PM

Die Annotationen (Toponyme) des lateinischen Textes mit Georeferenzierung sind semantisch modelliert in der WissKI-Datenbank als "Annotated Place" bzw. "Annotated Connection" gespeichert, aber auch zusammen mit allen anderen Analyseergebnissen als "Primary Research Data" im Drupal-Rahmen des WissKI-Systems als Textdateien in unterschiedlichen Formaten.

### 2.4.1 Textanalyse

Die Vorverarbeitung ist abgeschlossen und umfasst folgende Schritte:

- Alphabetische Wort(formen)-Liste mit Häufigkeiten (nur Kleinschreibung)
- Rückwärts (Endungs-) sortierte Wort(formen)-Liste mit Häufigkeiten
- Aufsteigende Wort(formen)-Liste mit Häufigkeiten (nur Kleinschreibung)
- Absteigende Wort(formen)-Liste mit Häufigkeiten (nur Kleinschreibung)
- Konkordanz (KWIC) mit Index
- Wortliste, [Tree-tagged](#) (STTS), mit Lemmata
- Wortliste, Tree-tagged (STTS), mit Häufigkeiten, nach Tags sortiert
- [Collatinus](#) (ohne Akzente)

Diese Daten stehen als Textdateien zur Verfügung und sind im Drupal-Rahmen der WissKI-Datenbank von der Startseite über den Link "Primary Research Data" erreichbar.

Angedacht für eine fokussierte Betrachtung ist eine Auswahl der Sätze mit räumlichen Bezügen aus der historischen Erzählung.

**Aufwand** für den letztgenannten Schritt: 0.3 PM

### 2.4.2 Bilinguale Glossierung

Die bilinguale Glossierung ist ein Desiderat für die (kognitions)linguistische Bearbeitung des Textes. Die

Erstellung der Glossierung sollte unter Verwendung der Ergebnisse von TreeTagger und Collatinus. In diesem Zusammenhang wäre auch ein manueller Eintrag satzübergreifender Koreferenzen sinnvoll vorzunehmen.

**Aufwand:** mindestens 3 PM

### 2.4.3 Recogito Annotation

Die Annotation der genuinen (Latium) Toponyme in Whites lateinischer Textfassung ist abgeschlossen.

Recogito bietet folgende Exportformate für die Annotationsdaten von Texten und Bildern an, die sämtlich heruntergeladen und auf dem Server sowie in Drupal/WissKI gespeichert wurden:

1. goreferenzierte Daten
  - csv (Tabellen mit Georeferenzierung)
  - json (GeoJSON mit Georeferenzierung)
  - kml (für Anzeige mit Google Earth), geobrowser.kml (für Anzeige mit dem DARIAH geobrowser)
2. Annotations-Metadaten ohne Georeferenzierung
  - ttl (RDF im Turtle-Format)
  - jsonld (JSON-Format für Linked Data)
3. tei.xml: TEI-Datei (nur für Texte)

Die Datenformate der Gruppen 1 und 2 werden in gleicher Weise für die Annotationen von Bildern bereitgestellt; die csv-Tabellen spielen dabei für weitere Analyse- und Verarbeitungsschritte aktuell die wichtigste Rolle. Ihre Einträge sind semantisch als "Annotated Place" modelliert und werden direkt in die WissKI-Datenbank via Tabellenimport eingebracht. GeoJSON und KML dienen derzeit nur für Anzeigewecke. Die Details der Einträge der Annotationen in den csv-Tabellen sind in 3.3 aufgelistet. Da die Annotation im aktuellen Projekt in der Regel nicht als Gruppenprozess organisiert ist, sind die Daten der zweiten Gruppe (mit Angaben über Annotator und Annotationszeitpunkt) z.Zt. wenig relevant und werden deshalb, auch in Drupal, als Block abgespeichert; bei Bedarf könnte zu einem späteren Zeitpunkt eine detailliertere Modellierung nach dem Vorschlag für CORDH ([Data Model OLD](#)) nachgereicht werden.

### 2.4.4 Dependenzparsing

Unter den verschiedenen Formen grammatischer Analyse bilden, gerade auch für das Lateinische, Dependenzstrukturen eine Brücke zur semantischen Repräsentation. Seit relativ kurzer Zeit steht von der Karls-Universität Prag mit dem [UDPipe Dependency Parser UD 2.4](#) und dem Universal Dependencies-Sprachmodell *latin-ittb-ud-2.4-190531* ein leistungsfähiger Dependenzparser zur Verfügung. Mit ihm wurde das Latium-Kapitel satzweise analysiert, die Ergebnisse jedoch noch nicht qualitativ überprüft. Die Dependenzbäume sind in einer Datei im tabellarischen CONLLU-Format gespeichert (auch in Drupal) und können mit einem speziellen auf der Prager Website verfügbaren Visualisierungsprogramm graphisch angezeigt werden. Am Ende sollte ein Abgleich der Dependenzanalysen mit den SpRL-Annotationen (s.

2.4.5) erfolgen, sobald letztere vorliegen.

**Aufwand:** ca. 3 PM

## 2.4.5 Spatial Role Labeling

Da für diesen Zweck kein automatisches Analysewerkzeug zur Verfügung steht, wurde das browser-basierte Werkzeug [brat](#) zur interaktiven händischen Annotation ausgewählt. Dieses Werkzeug sollte auch **auf dem Hertziana WissKI-Server installiert** werden; bisher wurde es nur in lokalen Installationen (Mac OS) bzw. an der [FAU](#) benutzt.

Eine neue mit der Ontologie *hmap* kompatible *brat*-Konfiguration wurde erstellt; sie sollte ggf. kurz revidiert werden - auch im Hinblick auf mögliche Erweiterungen für SpQL.

**Die Annotation des lateinischen Textes steht noch aus.**

**Aufwand:** mindestens 4 PM, insofern eine volle bilinguale Glossierung vorliegt. Für diese Aufgabe sind neben den linguistischen auch sehr gute Lateinkenntnisse erforderlich.

Die Annotationen werden von *brat* in einem eigenen "standoff"-Textformat gespeichert. Zur weiteren Verarbeitung muss dieses in ein kompatibles (tabellarisches) Tripelformat konvertiert werden, wozu noch ein Algorithmus zu fertigzustellen ist.

**Aufwand:** 1 PM

Anschliessend wäre noch eine Verknüpfung der Toponyme in den SpRL-Tripeln mit den geographischen Informationen aus der Recogito-Annotation einzurichten (semantische Anreicherung aus den Toponym-Annotationen) - dies kann unmittelbar geschehen, denn es handelt sich ja in beiden Fällen um denselben Text.

Für den Import in die WissKI-Datenbank (s. 5.) ist bereits eine semantische Modellierung der *figure-spatial\_indicator-ground* Tripel als "Annotated Connection" vorhanden.

In früheren Projektphasen wurden Annotationen der englischen Übersetzungen von White und Geus vorgenommen; diese sind jedoch unvollständig, zudem ist der Umbruch (Satzgrenzen) nicht kompatibel und die zugrundegelegte *brat*-Konfiguration veraltet.

Mittelfristig wäre das SpRL zu ergänzen um Annotationen der "Spatial Qualitative Relations" ("SpQL", s. 2.3), d.h. von Topologie, Richtungen und Entfernungen. Solche Auszeichnungen wären auch auf Kartenbildern wünschenswert, was de facto einen Schritt in die Kartometrie bedeutet. Für beide Fälle wäre die Ontologie *hmap* (s. 4.3) auf der Basis von CRMgeo zu erweitern.

**Aufwand** für Konzeption, Ontologie-Erweiterung und Anwendung auf das Latium-Buch (ohne Kartometrie):  
12 PM

## 2.4.6 Reihenfolge der Latiumorte

Für die Bearbeitung der übergeordneten Forschungsfrage wäre es eventuell sinnvoll, die Reihenfolge der Erwähnung (Ersterwähnung?) der Toponyme im Text darzustellen und zu visualisieren, ggf. mit einer dynamischen Anzeige von Pfaden. In diesem Zusammenhang sollte auch festgehalten werden, ob benutzte Wege erwähnt werden (hodologische Perspektive), oder ob die Beschreibung dem Verlauf von Flüssen oder Küstenlinien folgt.

Die Anordnung könnte leicht durch Einführung einer Pseudo-Zeitachse vorgenommen werden; eine Darstellung auf modernen Karten wäre mit dem DARIAH Geobrowser/PLATIN bzw. KML (Google Earth) möglich.

In diesem Kontext sollten neuere Entwicklungen von [topotime](#) und [GeoJSON-T](#) bzw. [Linked Traces](#) von Karl Grossner im Auge behalten werden.

## 2.4.7 "Semantisches Netzwerk" mit WORDij

Im Zusammenhang mit einfachen Statistiken ist uns das Program [WORDij](#) aufgefallen (Zitate stammen von dieser Website): "WORDij is a suite of data science programs that automates many aspects of natural language processing..."

WORDij baut auf der Idee von signifikanten Wort-Kookkurrenzen auf und führt i.W. linguistisch uninformierte Berechnungen über benachbarten Wörtern durch, die in *semantisch* interessanten Verknüpfungen resultieren: " WordLink is the initial text processor. It moves a sliding window through text, centering on each word and tabulating directed word pair bigrams appearing three words before and three words afterward. By treating proximate words, WORDij is more precise than "bag of words" programs that treat all words in a document as related. WORDij preserves the order of words in a bigram, which embeds syntax effects. Directed bigrams are useful for the Opticomm software in WORDij, whose strings linking a seed word and target word then lead to near-grammatical statements."

Ein erster Versuch zur Analyse des (lateinischen) Textes ergab viel versprechende Ergebnisse (s. Abb. 3); weitere Versuche sollten mit verschiedenen Parametrisierungen unternommen werden. Mit dem Werkzeug [gephi](#) zur dynamischen Visualisierung von Graphen können semantische Cluster und ihre Zusammenhänge exploriert werden. Hierfür wäre die Erarbeitung ein Auswertungskonzept für die semantischen Cluster erforderlich.

**Aufwand:** ca. 3 PM

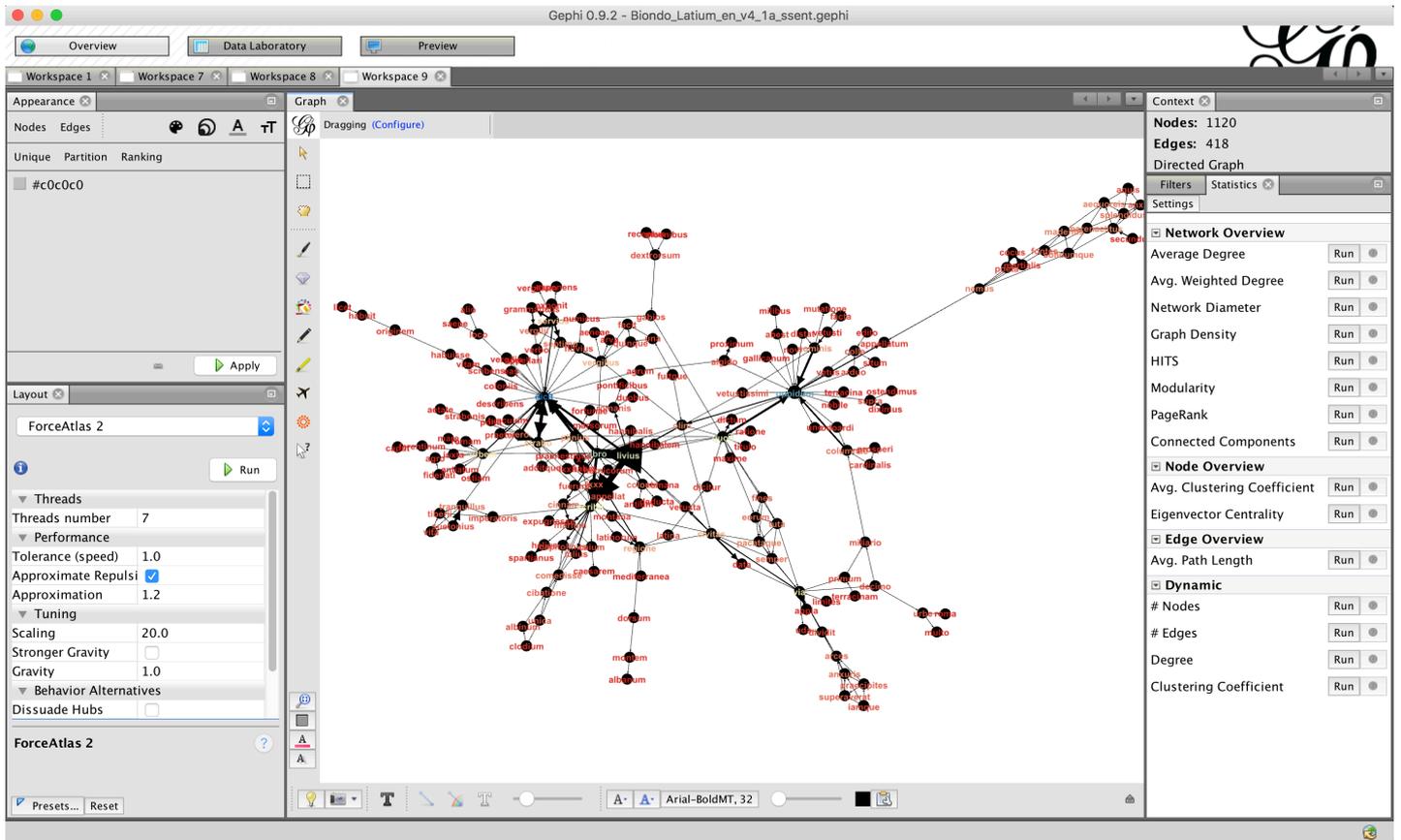


Abb. 3: WORDij Netzwerk, Anzeige mit gephi

## 3. Organisation und Datenerzeugung: Bild

### 3.1 Kartenauswahl und Bearbeitung

Biondo erwähnt die Benutzung von Karten, macht über diese jedoch keine näheren Angaben. Daher sollten im Projekt zunächst verschiedene Italienkarten aus dem 14./15.Jh. untersucht und dabei insbesondere die Toponyme (beginnend mit Latium) mit Recogito 2 annotiert werden. Ob noch weitere Merkmale der Karten (kartometrische, ikonographische,...) für das Projekt relevant sind, wurde vorerst offen gelassen.

Ausgewählt wurden

1. einzelne Italienkarten aus dem 15.Jh. nach der Liste von Milanese,
2. Ptolemaeus-Karten - "traditionelle" und "moderne" (N),
3. Sonstige - vor allem Seekarten (Portolane) wegen der Küstenorte.

Es besteht Konsens, dass die Karten der Gruppen (1) und (3) Priorität haben. Für die Auswahl der Karten der ersten Gruppe war vor allem massgeblich M. Milanese, *Antico e moderno nella cartografia umanistica: le grandi carte d'Italia nel Quattrocento* ("Geographia antiqua" 16-17, 2007-2008, 153-176). Eventuell kommen noch weitere Regionalkarten in Frage, sofern sie bekannt werden. Die Auswahl der Seekarten sollte noch einmal mit Fokus auf italienischen Portolanen anhand der chronologischen Tabelle von [Campbell](#) kritisch überprüft werden.

(Hinreichend) annotiert wurden die folgenden Karten (s. vollständige Kartenliste im Anhang A.1):

- B1.04, B1.08, B1.09
- B2.02
- B3.01, B3.03, B3.18a, B3.20, B3.20N, B3.23, B3.23N, B3.26N, B3.29, B3.31, B3.32N, B3.34N, B3.37
- B4.01, B4.03, B4.03N, B4.05, B4.06, B4.06N
- B5.01, B5.02, B5.05, B5.06, B5.07, B5.08

Wie in 2.4.3 ausgeführt ist, können aus Recogito die Annotationen in verschiedenen Formaten exportiert werden. Aktuell bevorzugen wir Tabellen (csv), die zunächst mit einschlägigen Programmen (Numbers, Excel) weiter verarbeitet werden können. Um die Annotationsergebnisse auswerten zu können, war eine Vereinheitlichung und Konsolidierung der vorliegenden Annotationen, vor allem bei (2) und (3), notwendig.

Zunächst lag der Fokus im Projekt auf Latium; zu möglichen Erweiterungen s. 1.5.2

## 3.2 Annotationsrichtlinien und Recogito

Für die Kartenannotation gelten die folgenden Transkriptionsrichtlinien:

Toponyme werden diplomatisch unter Beachtung von Gross- und Kleinschreibung transkribiert, aber

- langes s wird durch kurzes s ersetzt, langes z (Unterlänge) durch kurzes z,
- Ligaturen und Worttrennungen werden ignoriert,
- Diakritika, z.B. e mit Tilde, werden aufgelöst, also entsprechend: en (hierzu: Abbreviationes-Programm oder Cappelli, A.: *Lexicon Abbreviatorum - Dizionario di Abbreviature Latine ed Italiane*. Milano: Hoepli 1990/1999).
- Insbesondere dürfen KEINE Anführungszeichen, Kommas und Semikolons benutzt werden!

*Hinweise zur Benutzung von Recogito 2 befinden sich im Anhang A.2.*

## 3.3 Bearbeitung der exportierten Annotationsdaten, Tabellenverarbeitung

Eine detaillierte Beschäftigung mit den spätmittelalterlichen Karten, wie sie insbesondere durch die Annotation der Toponyme erforderlich ist, weist auf einen klaren Wechsel im geographischen Wissen im 14. und 15. Jh. hin. Dies zeigen zum einen die teilweise sehr detailreichen genuinen Italienkarten - etwa von Paolinus Minorita zur Cotton-Karte -, als auch der Unterschied zwischen traditionellen und "modernen" Ptolemaeuskarten, aber auch die Portolane.

Jede Annotation (für Bilder und Texte gleichermaßen) im csv-Tabellenformat - und analog in GeoJSON - enthält folgende Angaben:

- UUID, ein eindeutiger Identifikator der betreffenden the Annotation,
- FILE, der Name der annotierten Datei,
- QUOTE\_TRANSCRIPTION, die Transcription der Inschrift,

- ANCHOR, die Position des annotierten Elements im Text (Zeichenposition) oder Bild (Pixelkoordinaten),
- TYPE, entweder PLACE oder PERSON oder EVENT,
- URI, die eindeutige Webadresse des Orts in einem Gazetteer, wenn der Ort geographisch verifiziert werden kann,
- VOCAB\_LABEL, die Angabe des Namens im Gazetteer inklusive Schreibvarianten,
- VOCAB\_TEMPORAL\_BOUNDS, die zeitliche Geltung des Namens,
- LAT, geographische Breite des Orts,
- LNG, geographische Länge des Orts,
- PLACE\_TYPE, ein Deskriptor aus einem standardisierten Vokabular wie dem Pleiades-Vokabular, z.B. "settlement", "river", "mountain", etc.
- VERIFICATION\_STATUS, entweder geographisch VERIFIED, oder NOT\_IDENTIFIABLE, wenn der Ort in keinem der erreichbaren Gazetteers enthalten ist, oder UNVERIFIED, wenn der Verifikationsschritt nicht abgeschlossen wurde,
- TAGS, ein freies Feld, das für "place types" bei nicht geographisch verifizierten Orten benutzt werden kann,
- COMMENTS.

### 3.3.1 Tabellenauswertungen

Aus den (operationalisierten) Forschungsfragen könnten nun zu diesen Angaben unterschiedliche Sortierungen und Filterungen der Daten vorgenommen werden, z.B.

- Auswertung der place-types (welche kommen vor?); Siedlungen im Text vs. andere Orte im Text
- Vergleiche Text - Karten
  - Vergleichende Übersicht der vorkommenden Toponyme: Text | Ptolemaeus traditionell | Ptolemaeus modern | andere Italienkarten (insbes. Cotton und Hamilton) | Portolane; d.h. insbesondere
    - Welche im Text explizit genannten *antiken* Orte kommen in Ptolemaeus traditionell / modern vor?
    - Für antike Orte: Peutingeriana (Datenbank) einbeziehen?
    - Welche der Orte im Text kommen speziell in modernen Ptolemaeus-Karten vor?
    - Orte auf Karte(n), die *nicht* im Text vorkommen (nach Kartentyp)
    - Inwieweit sind Schreibvarianten der Toponyme interessant? Verifizierte Orte: mit gleichen / mit verschiedenen Transkriptionen
- Für die geographisch verifizierten Orte (antike Toponyme, soweit vorhanden: Pleiades; moderne Toponyme: geonames, evtl. DARE) kann zum Vergleich die Visualisierung auf einer heutigen Italienkarte hilfreich sein (mit Recogito oder <http://geojson.io> oder Leaflet in WissKI). Zur Anzeige von Koordinaten auf verschiedenen modernen Karten mit zusätzlichen Informationen kann evtl. auch "[GeoHack](#)" dienen.
- Evtl. Anzeige der Pixelkoordinaten in einem Rechteck; ggf. mit schematischer Umrisslinie Italiens?
- Vergleich von Distanzen im Text zu euklidischer Distanz (Kartometrie)

- Graphen-Darstellung Ort-Kante-Ort. Technische Möglichkeiten zur (web-basierten) Darstellung z.B. mit "Data Driven Documents" [D3.js](#), "Plotly JavaScript Open Source Graphing Library" [plot.ly](#) mit Python Framework [dash](#).

### Besonderheiten bei ptolemaeischen Karten: Weitere Möglichkeiten

In der Ptolemaeischen Geographie liegen - im Unterschied zu anderen Karten - Ortslisten mit Koordinaten vor (Italien: EU06; im Ptolemaeus-Text sind Küste und das Landesinnere getrennt). Im Rahmen der Kartenbearbeitung wurde auch die Ortstabelle Italiens (EU06) bei Ptolemaeus (aus der Edition Stückelberger et al.) mit Recogito annotiert, so weit das möglich ist. Damit existiert für fast alle Toponyme ein Link zu einem Gazetteer-Eintrag (v.a. Pleiades und DARE; Export-Datei: ptolemaios.EU06.c4.csv.txt).

Damit werden Vergleiche der Ptolemaeus-Karten miteinander und mit dem Ortskatalog möglich.

Die ptolemaeische Ortstabelle ist nach Völkern (Regionen) organisiert, was mittelfristig Möglichkeiten für Vergleiche der ptolemaeischen Ortsliste EU06 mit noch zu bestimmenden Regionen in traditionellen Ptolemaeus-Karten eröffnet:

- explizit eingezeichnet und benannt
- implizit (Völkersymbole)
- Vergleich der Regionen bei Plinius und Ptolemaeus.

Weiterhin steht ein Vergleich der Koordinaten Ptolemaeus' "bedeutender Städte" Italiens (im Ergänzungsband der Stückelberger-Ausgabe) aus; diese liegt noch nicht als csv-Tabelle vor, aber es gibt Varianten (Dateien: "Ptolemy Table of Important Cities-AWMC" Ptolemy Directory.csv und ptolemaeus.themelioi.juergenheess.EU06.csv).

Im Unterschied zu den traditionellen Karten gibt es für die **Tabulae Novae** keine Orts-Koordinatenliste; in der Regel sind nicht einmal Markierungen für Längen- und Breitengrade am Rand eingetragen.

Vergleichsmöglichkeiten:

- Moderne Karten von Donnus Nicolaus Germanus - die Karten von seiner Hand weisen einen hohen Grad an Übereinstimmung auf - vs. Karten von anderen Händen
- Welche Orte sind auf welchen Karten?
- Welche Orte aus traditionellen Ptolemaeus-Karten sind auf den modernen?
- Ist ein detaillierter Vergleich der "modernen" Karten (Tabellen) untereinander zielführend (Vorarbeiten von Chiara Seidl)?

### 3.4 Erweiterung um Kartometrie

Die Untersuchung der kartographischen Wiedergabe räumlicher Relationen, d.h. der Topologie, aber auch von Richtungen und Entfernungen ist ein Schritt in Richtung historische Kartometrie. Zunächst wäre eine systematische Literatur-/Softwarestudie über geeignete methodische Ansätze und Werkzeuge erforderlich.

- Positionsunterschiede auf verschiedenen Karten sinnvoll?

- Vergleich moderner Ptolemaeus-Karten mit traditionellen und anderen Karten
- Streuung der Positionen von Orten auf verschiedenen Karten
- Sind weitere kartometrische Untersuchungen erwünscht? (Guckelsberger)
  - Pixelkoordinaten ausmessen vs. Rückrechnen in Geo-Koordinaten
  - Ptolemaeische Koordinatenumrechnung für ganz Italien (Bezugsort Rom) ?? Nach Geus (persönliche Mitteilung) problematisch.
  - Karten ausmessen bzgl. Entfernungen?

## 4. Semantische Modellierung

---

Parallel zur Bearbeitung der Texte und Kartenbilder, die Primärdaten in der Form kompatibler Exportformate, v.a. CSV erzeugt, geht es in diesem Teil um die Erarbeitung einer semantischen Repräsentationsebene, in der die Daten aus den Annotationen - inkl. Quellenverweisen, geographischen Angaben, etc. - und die aus den kognitiv-linguistischen Annotationen zusammengeführt und in einer Anwendungsontologie auf der Basis CRM+CRMgeo dargestellt und als Linked Open Data publiziert werden können.

Die grundsätzliche Vorgehensweise zeigt Abb. 1.

### 4.1 Vokabular (Thesaurus)

In Pleiades ist eine Klassifikation von geographischen Orten ([place-types](#)) als flache Liste vorgegeben und wird in dieser Form benutzt (automatisch in Recogito).

Für künftige Arbeiten wäre u.U. eine leichte Hierarchisierung zweckmässig, z.B. in Typen der physischen und der Humangeographie. Zur Darstellung bietet sich das SKOS-Format an. Für eine Kombination mit der Anwendungsontologie könnten die Thesaurus-Einträge als CRM-Types ("E55") vorgehalten werden (beide Optionen - OWL und SKOS - wie im Beispiel [geosparql\\_vocab\\_all.rdf](#)). Ggf. sollte ein Abgleich mit anderen einschlägigen Vokabularen erfolgen.

Eine strukturierte Darstellung in OWL mit Eintrag der SKOS-Konzepte aus der flachen Liste (wie in [geosparql\\_vocab\\_all.rdf](#)) (für Spalten "place\_type" und "tags") liegt nahe. Zukünftige Erweiterungen könnten nach Vorlagen in "Ontology Design Patterns" aus den in 4.2 zitierten Arbeiten von Gkadolou (Athen) und Scheider et. al. (Münster) vorgenommen werden.

### 4.2 Basisontologie

Für die semantische Modellierung wurde als Basisontologie das CIDOC Conceptual Reference Model [CRM](#) mit der Erweiterung [CRMgeo](#) zur Integration von Geo-Informationen (inkl. [geosparql](#)) gewählt (s.a. Hiebel, Doerr, Eide. CRMgeo: A spatiotemporal extension of CIDOC-CRM. International Journal on Digital Libraries 18 (2017), 271–279). Aktuell benutzen wir die OWL-DL Implementation [Erlangen CRM](#) und eine (eigene) OWL-Version von CRMgeo, [ecrmgeo](#), aktuell lauffähig in der Version 1.4.

Optionen für künftige Erweiterungen der Modellierung:

- angekündigte Revision von CRMgeo (2020?)
- Die Darstellung der Zeit- und Lageabhängigkeit der benannten Orte kann noch weiter detailliert werden (CRM "Spacetime Volume"), z.B. wie in Görz, Albers: Representing Place in Space and Time - Methodological Aspects in Modelling the Provenance of Cultural Heritage Knowledge (Proc. CIDOC 2018, Heraklion, 8pp.).
- Einführung weiterer kartographischer/geometrischer Konzepte und "object properties" (Topologie, Orientierung, Entfernung), s. Salas, Hart: [NeoGeo Vocabulary](#) - Defining a shared RDF representation for GeoData. Dies ist auch grundlegend für die SpQL-Annotation (2.3).
- Personen im Text.
- Events (benannte Ereignisse) im Text.
- Zeitangaben und zeitliche Relationen im Text, s. [HeidelTime](#) ("a multilingual, domain-sensitive temporal tagger").
- Ein Ereignisbezug im Sinne von CRM könnte z.B. zentral durch ein MOVE-Ereignis modelliert werden.
- CRM/CRMgeo-Modellierung der OpenAnnotation-Metadaten, wie sie aus Recogito exportiert werden (RDF). Hierfür besteht aktuell keine Notwendigkeit, sie könnte aber zukünftig bei Bedarf vorgenommen werden.

### 4.3 Anwendungs-Ontologie fuer historische Karten (*hmap*)

Über der Basisontologie haben wir eine an unseren Anforderungen ausgerichtete Anwendungsontologie für historische Karten *hmap* definiert. Einflüsse hierfür kamen von einer in einem früheren Erlanger Projekt ("mappae") erarbeiteten Anwendungsontologie sowie von Arbeiten von Gkadolou, Stefanakis und Chalkias (Harakopio Univ., Athen) und von der Gruppe um [Scheider](#) (früher Univ. Münster). Zu Letzteren, allerdings ohne Verwendung von CRM, s.a. Scheider et al.: Encoding and Querying Historical Map Content (in Huerta et al. (Ed.): Connecting a Digital Europe Through Location and Place. Berlin, 2014, 251-273).

*hmap* umfasst derzeit 45 Klassen und 47 "object properties", mithilfe derer dann Metadaten definiert wurden für Karten (Map), Kartenwerke/Atlanten (Work), Kartenbilder (Image), Bilderserien (Image series), (geographische) Texte (GeoText), Literaturangaben (Bibliography, [Zotero](#)-kompatibel) und (annotierte) Karten-Inhalte, in unserem Fall annotierte Orte (Annotated Place) und Relationen zwischen Orten (Annotated Connection).

#### 4.3.1 Metadaten Karte, Kartenwerk/Atlas

Für Karten (Map) gibt es folgende Elemente:

ID, Title Map, Theme/Kind, Full Map, Map Parts, Owner, Current Location, Catalog Number (map), Catalog Number (work), Page, Language, Shape, Type, Material, Dimension (Width, Height), Production (Cartographer, Creator, Publisher, Sources, Template, Technique, Production Place, Production Date), Scale, Orientation, Reference System, Projection, Comment, Condition, Preservation Comment, Images, Bibliography

Für Kartenwerke/Atlanten (Work):

ID, Creator (work), Title, Theme, Catalog number (work), Contains map, Bibliography

### **4.3.2 Metadaten Kartenbild, Bilderserie**

Für Kartenbilder (Image):

ID, Recogito-ID, Title, Theme, Image-URL, Map, Pixel Dimension (Width, Height), Description, Series, Rights (Rights Owner, License, License URL), Image Source, Recogito-iframe, Depiction

Für Bilderserien (Image series): ID, Title, Images contained

### **4.3.3 Metadaten Text**

Für GeoText:

ID, Author, Title, Theme, Text-URL, Text Dimension (Characters, Sentences), Description, Rights (Rights Owner, License, License URL), Text Source, Depiction, Bibliography

### **4.3.4 Daten: Text- und Kartenbild-Inhalt**

Für Annotated Place: s. 3.3, zusätzlich "Image"-Link

Für Annotated Connection: ID, Label (spatial\_indicator), Source (figure), Target (ground)

## **4.4 Tripel-Folgen (Pfade) zur semantischen Modellierung der Metadaten/Formulare**

Die Metadaten- und Datenfelder werden (weitestgehend) in der Anwendungsontologie formuliert als Folgen von RDF-Tripeln, sog. Ontologie-Pfaden. Durch die Wahl von WissKI (s. 5.) kann dies interaktiv mit optimaler Unterstützung in dessen "Pathbuilder"-Modul vorgenommen werden. Daraus werden automatisch Eingabe- und Suchformulare erzeugt.

# **5. Implementation der WissKI-Plattform und Publikation von Linked Open Data**

---

Als Plattform für Speicherung, Zugriff und Publikation (Linked Open Data) der semantisch angereicherten Forschungsdaten des Projekts wurde Anfang 2019 die Virtuelle Forschungsumgebung [WissKI](#) gewählt. Grundsätzlich wäre wohl auch die für den Einsatz an der Bibliotheca Hertziana geplante "Research Space"-Umgebung geeignet. Sie wird im Rahmen eines Architekturkonzepts im MPIWG erprobt (Abb. 4), steht jedoch noch nicht zur Verfügung; zahlreiche Entwicklungen sind im Rahmen des CORDH-Verbands noch in Arbeit. Zusammen mit Recogito bildet WissKI das "Working Environment" in Abb. 4.

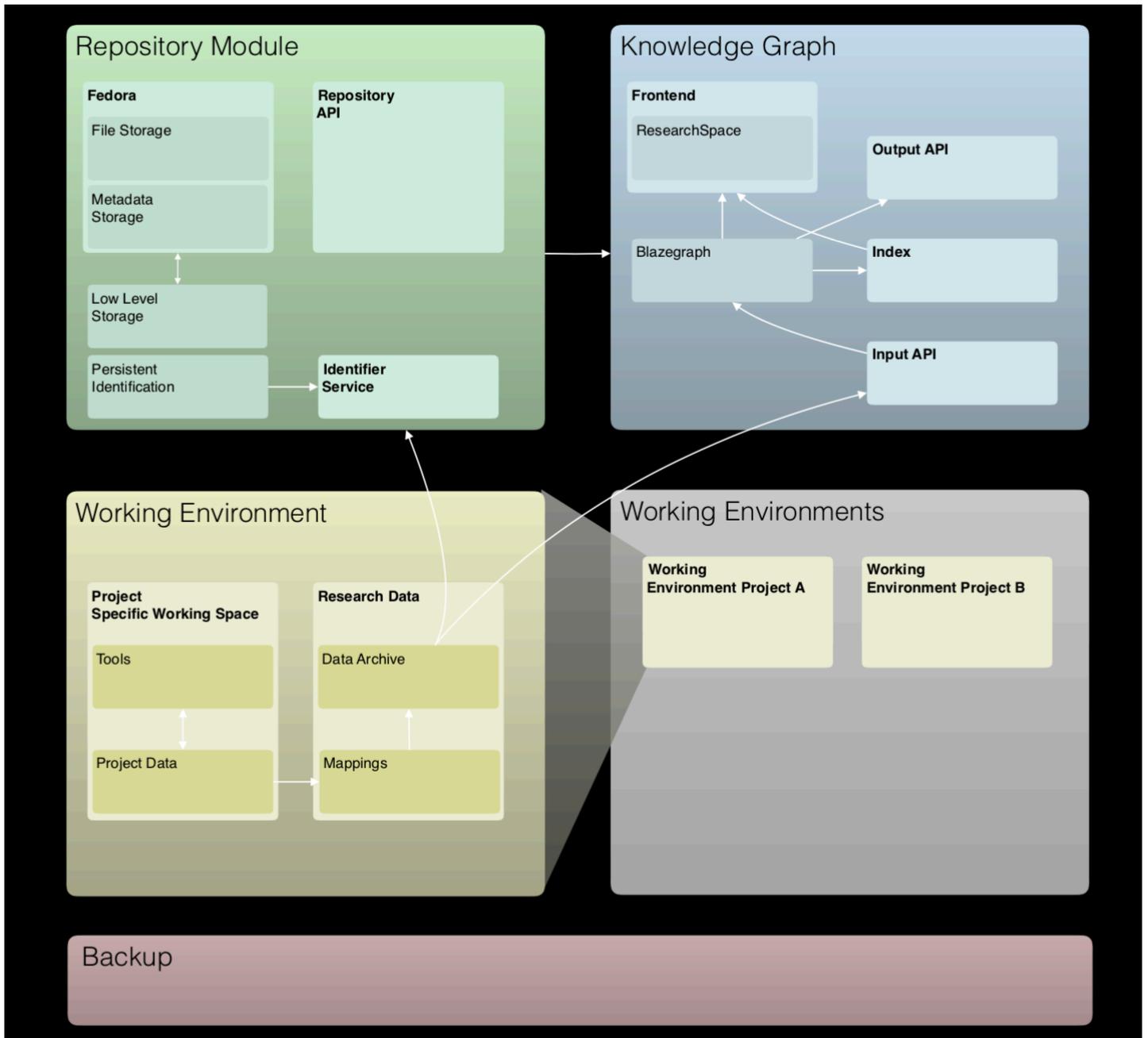


Abb. 4: Infrastruktur-Diagramm (MPIWG)

Gründe für die Wahl von WissKI waren:

- die gewünschte Funktionalität ist vorhanden bzw. konnte mit relative geringem Aufwand realisiert werden: semantische Modellierung (inkl. Disambiguierung) mit CIDOC CRM und Erweiterungsontologien, flexible Erweiterbarkeit, Unterstützung bei der Dateneingabe, Tabellenimport, IIIF-basierte Bildanzeige, Kartenanzeige (Leaflet), Triple Store (OntoDB, blazegraph möglich), SPARQL-Schnittstelle.
- Vertrautheit mit WissKI: Es ist eine eigene DFG-geförderte Entwicklung in Kooperation mit dem Germanischen Nationalmuseum Nürnberg; dort wird das System auch gewartet und erweitert (in Zusammenarbeit mit der FAU, UB Heidelberg und dem Herder-Institut Marburg). WissKI ist stabil und wird in weit über 120 Anwendungen genutzt, u.a. zur Objektdokumentation (Museen, Archive), für

Editionen (Hans Posse am GNM, Behaim-Globus), zur Präsentation (Gesichter des deutschen Kunstarchivs), Ausstellungsvorbereitung (u.a. Der frühe Dürer, Henlein, ...).

- Ein [Demonstrationssystem](#) konnte direkt von Martin Raspe and der Hertziana aufgesetzt werden und ist einsatzbereit.
- Sämtliche Primär-Forschungsdaten können im selben Rahmensystem (CMS Drupal) vorgehalten werden, das auch eine automatische Versionsverwaltung bietet. Damit sind die Voraussetzungen zur Reproduzierbarkeit der Ergebnisse durch Dritte gegeben.

## 5.1 Semantische Datenbank / LOD

Die zentrale Website zu [WissKI](#) enthält eine umfassende [Dokumentation](#), Informationen über aktuelle Entwicklungen und verschiedene Anwendungen. In der [Dokumentation](#) des Projekts "[Objekte im Netz](#)" gibt es weitere Handbücher, die jedoch auf spezifische Anforderungen dieses Projekts hin gestaltet sind.

### 5.1.1 Installation und Einrichtung Triple-Store

Eine Instanz des WissKI-Systems wurde für das Projekt installiert: <http://wisski.biblhertz.it>. In ihr sind sämtliche "WissKI Entities" mit den in 4.3 genannten Metadaten implementiert (s.a. 5.1.3).

Für Testzwecke ist eine Benutzerkennung guest/biblhertz eingerichtet, mit der nur lesend (Navigate, Find) zugegriffen werden kann. Voll berechtigte Benutzer können mit Create auch neue Instanzen ("WissKI Entities") erzeugen.

Als Triple Store wurde eine Instanz von GraphDB eingerichtet. Für SPARQL-Anfragen gibt es ein [Suchfenster](#).

Beispiel:

```
PREFIX ecrm: <http://erlangen-crm.org/170309/>
PREFIX hmap: <http://digilib.biblhertz.it/ontologies/hmap/>
select * where {
  ?s hmap:A32_has_title ?o
} limit 100
```

Alternativ sind solche Anfragen auch an den Triple Store [direkt](#) möglich.

Selbstverständlich ist eine SPARQL-Anfrageschnittstelle primär eine Schnittstelle für maschinellen Zugriff; für eine benutzerfreundliche Exploration des Datenbestandes und die Formulierung von Anfragen gemäss den Dimensionen Raum und Zeit sowie weiteren semantischen Konzepten ist Entwicklungsarbeit zu leisten. An dieser Stelle kann nur exemplarisch auf die experimentelle Anfrageschnittstelle von [Research Space](#) sowie Scheider et al.: Exploratory querying of SPARQL endpoints in space and time (Semantic Web Journal, vol. 8, 2017, no. 1, pp. 65-86) verwiesen werden.

## 5.1.2 WissKI-Konfiguration

Es wurden drei "Pathbuilder" angelegt:

- italia15: dort sind die Pfade definiert für die Gruppen Annotated Place, Annotated Connection, Bibliography, GeoText, Image, Imag Series, Map, Work.
- SparqlGraphBuilder als Interface zur SparqlGraphDB.
- zotero mit der Gruppe Bibliography als Interface zu dem Zotero-Bibliographiesystem; in Zotero wurde eine [Bibliographie für das Projekt](#) angelegt.

## 5.1.3 Eingabe Ontologien

Als Anwendungsontologie (für SparqlGraphDB) wurde *hmap\_v1.owl* eingegeben, die ihrerseits die Ontologien *crmgeo\_1.4.owl* und dann *ecrm* von <http://erlangen-crm.org> einliest.

## 5.1.4 Einrichtung Tabellenimport

Der Import der csv-Tabellen (Annotated Place) kann am einfachsten mit PHPMyAdmin vorgenommen werden. Dazu wurde unter <http://wisski.biblhertz.it/phpMyAdmin> ein Benutzerkonto eingerichtet und ein Importskript für WissKI angelegt, das die Abbildung der Tabellenstruktur auf die Gruppe "Annotated Place" definiert.

## 5.1.5 Dateneingabe

Durch die Definition der in 4.3 angegebenen Metadaten-Schemata werden automatisch Eingabefelder erzeugt.

Für alle für das Projekt ausgewählten Karten (s. 3.1 und A.1) wurden Metadatensätze angelegt; der Vollständigkeit halber sogar für sämtliche lateinischen Ptolemaeuskarten, auch wenn von vielen keine Bilder vorliegen. Jeder Karte wurde genau ein Bild zugeordnet (Metadaten und Bilddatei), das annotiert wurde bzw. werden soll. Liegen für eine Karte weitere Bilder vor, z.B. solche, die durch verschiedene Bildverarbeitungsverfahren abgeleitet wurden, um etwa die Lesbarkeit zu verbessern, wurden diese bisher nicht in der WissKI-Datenbank gespeichert, sondern als Rohdaten in deren Drupal-Rahmen.

Die zu den annotierten Karten aus Recogito als csv-Tabellen exportierten Annotationen wurden als Annotated Places importiert.

Weiterhin wurde eine Möglichkeit geschaffen, direkt die Anzeige von Recogito in WissKI als sog. Recogito-iframe einzubetten. Dies ist allerdings nur möglich, wenn für das Bild eine freie Lizenz besteht, was allerdings für fast alle der aktuell annotierten Bilder nicht der Fall ist.

## 5.2 Repository für Forschungs-Primärdaten (Drupal)

Sämtliche vorhandenen Forschungsrohdaten des Projekts, wie in 1.5.1 beschrieben, sind auf dem Citrix-

Server des Instituts gespeichert. Das Repositorium der aktuellen Daten ist in Drupal in der Form von "Book Pages" abgelegt und über einen Link auf der Startseite erreichbar. Die folgenden Abschnitte beschreiben kurz den Inhalt des Repositoriums.

### 5.2.1 Texte

1. Rohtexte (lateinische Fassung von White, s. 2.4)
2. Analysen inkl. Parsing (2.4.1, 2.4.4)
3. Annotationen (Recogito Exportformate) (2.4.3)
4. Spatial Role Labeling

### 5.2.2 Bilder

Die annotierten Bilddateien werden in der WissKI-Datenbank gespeichert. Zusätzliche, z.B. mit Photoshop bzgl. Lesbarkeit bearbeitete, werden als Forschungsdaten abgelegt.

1. Einer Karte zugeordnete weitere Bilddateien
2. Geo-Annotationen (in allen Recogito Exportformaten)
3. Annotations-Metadaten im Open-Annotation-Format als *eine* RDF-Datei

### 5.2.3 Ontologien und Vokabulare

Hier sind gespeichert die Anwendungsontologie *hmap\_v1.owl* sowie *crmgeo\_1.4.owl*. *crmgeo* ist außerdem als [ECRMGEO](#), ebenso wie [ECRM](#) in der benutzten Version 170309 verlinkt.

### 5.2.4 Dokumentation

Diese Seite ist vorgesehen für die Ablage von Veröffentlichungen,

1. Zeitschriften-, Buch- und Konferenzbeiträge
2. "Folien"

Die Bibliographie wird in [Zotero](#) erstellt und ist direkt an WissKI angeschlossen (Pathbuilder "zotero").

### 5.2.5 Verschiedenes

### 5.2.6 Nützliche Links

zu verschiedenen Werkzeugen, Datenbeständen, etc.

## Technischer Anhang

---

### A.1 Ausgewählte Italien-Karten des 14. und 15.Jh. für das BIONDO-Projekt

Alle Karten sind durch eine "laufende Nummer" identifiziert ("B" gefolgt von der Gruppennummer und einer laufenden Zählung; bei Tabulae Novae wurde N angehängt). Die in eckigen Klammern angegebenen Nummern dienen der eindeutigen Identifikation der annotierten Kartenbilder. Sofern sie ein X enthalten, beziehen sie sich auf digital bearbeitete Bilder (K. Wagner) zur Unterstützung der Annotation, die aber selbst nicht annotiert werden. Diese werden auch in der im Aufbau befindlichen WissKI-Datenbank benutzt.

Handschriften und Drucke des Nicolaus Germanus hat Babicz (1980) zusammengestellt, ebenso die Redaktionen anderer Bearbeiter. Die für Handschriften benutzte L-Nummerierung, die auch hier verwendet wird, geht auf Fischer zurück. Der Vollständigkeit halber enthält die Liste auch Angaben zu Ptolemaeus-Karten, von denen keine Bilder vorliegen. (Literatur: J. Babicz, *Donnus Nicolaus Germanus - Probleme seiner Biographie und sein Platz in der Rezeption der ptolemäischen Geographie*. In: Koeman, C. (Hg.): *Land- und Seekarten im Mittelalter und in der frühen Neuzeit*. Wolfenbütteler Forschungen, Bd. 7, 1980, 9-42.)

Eine chronologische Liste von Portolanen gibt Campbells "[Census](#)", dessen Nummerierung (Cam) wir hier folgen.

## 1. Einzelne Italien-Karten

- B1.01 (Anonym) Tabula Peutingeriana Tafel 5 [B1.01.1] und Tafel 6 [B1.01.2]. 1507, Wien, ÖNB Cod. 324, H 34cm
- B1.02 Paulinus Minorita: Chronologia Magna, 1328-1343. Vatikan, Vat.lat.1960: Italien [B1.02.1], Norditalien [B1.02.2], Süditalien [B1.02.3], für Lesbarkeit optimierte zusammengesetzte Version [B1.02.X01].
- B1.03 (Anonym) Italienkarte [B1.03.1], 1401-1450. Firenze, Archivio di Stato, Carte Nautiche e Topografiche n.10, 63x111 cm; für Lesbarkeit optimierte Version [B1.03.1].
- B1.04 (Anonym) Italiae provinciae modernus situs [B1.04.1], 1350-1410, London, British Library, Cotton roll XIII 44, 64x137cm; für Lesbarkeit optimierte Version [B1.04.1X1]. Mit besserer Auflösung in drei Teilen: Oben (Süd-Italien) [B1.04.2], Mitte (Mittel-Italien) [B1.04.3], Unten (Nord-Italien) [B1.04.4]. Für Lesbarkeit optimierte Varianten [B1.04.2X01-2X07, B1.04.3X01-3X02, B1.04.4X01-4X05]
- B1.05 (Anonym) Italia settentrionale e centrale [B1.05.1], 1440-1450. Strasbourg, Bibliothèque Nationale e Universitaire, Ms. 1.816, 65,5x84 cm
- B1.06 (Anonym) Italienkarte [B1.06.1], 1449, provenienza Cicogna; Venezia, Civico Museo Correr, Dep.19, 68x118 cm Für Lesbarkeit optimierte Varianten [B1.06.X01-X14]
- B1.07 (Anonym) Italienkarte Norden [B1.07.1] und Süden [B1.07.2], 1450-1500, provenienza Boni; Modena, Biblioteca Estense e Universitaria, C.G.A.7, 65,5x104 cm. Für Lesbarkeit optimierte Varianten [B1.07.1X01-1X03, B1.07.2X01-2X03]
- B1.08 (Anonym) Italienkarte [B1.08.1], 1450-1500, Berlin, Staatsbibliothek, Codex Hamilton, ms.108
- B1.09 Mittelmeer [B1.09.1] in Insularium, Laurent.Plut.29.25. Firenze, Biblioteca Nazionale

## 2. Griechische Ptolemaeus-Karten, Handschriften

- B2.01 Codex Florentinus Laurentianus Graecus Plut.28.49, Norden [B2.01.1] und Süden [B2.01.2]. B-

Redaktion, 14.Jh.

- B2.02 Codex Florentinus Laurentianus Graecus Conv.Soppr.626 [B2.02.1]. A-Redaktion, 15.Jh.
- B2.03 Codex Venetus Marcianus Graecus GR 516 [B2.03.1]. A-Redaktion, ca. 1450
- B2.04 Codex Vindobonensis historicus graecus 1, Norden [B2.04.1] und Süden [B2.04.2]. A-Redaktion, 1454

### **3. Ptolemaeus-Karten, Handschriften**

Hier ist der Vollständigkeit halber die gesamte Liste der lateinischen Handschriften nach Fischer aufgenommen; Bilder sind nur zu Einträgen mit Angaben in eckigen Klammern vorhanden.

#### **Älteste Manuskripte, entsprechen der 1. Redaktion von Donnus Nicolaus Germanus (ausgenommen L2); keine Tabulae Novae**

- B3.01 L01 Codex Vaticanus latinus 5698, Gesamt [B3.01.1]. ca. 1415, Roma, Biblioteca Vaticana.
- B3.01a L01a Codex Cracoviensis latinus 7805, 1400-1450
- B3.02 L02 Codex Nanceanus 441, Gesamt [B3.02.1], Norden [B3.02.2] und Süden [B3.02.3]. 1427
- B3.03 L03 Atlas Laurentianus Aediliuni 175 [B3.03.1]. ca. 1430, Firenze, Biblioteca Nazionale
- B3.04 L04 Codex Parisiensis latinus 4801 [B3.04.1]. Mitte 15.Jh., Paris, Bibliotheque Nationale
- B3.05 L05 Codex Neapolitanus latinus V F33
- B3.06 L06 Codex Oratorianus latinus Pil.IX,2
- B3.07 L07 Codex Mediolanensis Ambrosianus B52 inf.
- B3.08 L08 Codex Parisiensis latinus 4803
- B3.09 L09 Codex Parisiensis latinus 4804 [B3.09.1]. 1485, Paris, Bibliotheque Nationale
- B3.10 L10 Codex Lond. Harlecianus latinus 7195
- B3.11 L11 Codex Lond. Harlecianus latinus 7182
- B3.12 L12 Codex Parisiensis latinus 15184
- B3.13 L13 Codex Zeitzianus latinus, Hus.fol.497
- B3.14 L14 Codex St. Petersburgensis latinus [verloren nach 1917]

#### **Erste Redaktion von Donnus Nicolaus Germanus, 1460-1466, keine Tabulae Novae**

- B3.15 L15 Codex Neapolitanus latinus V.F. 32. Napoli, Biblioteca Nazionale, Pergament, 44x297mm
- B3.16 L16 Codex Ebnerianus latinus [B3.16.1]. ca.1460, New York Public Library, Lenox Division, 417x280mm
- B3.16a L16x Codex Parisiensis latinus 10764 [B3.16a.1]. 1490, Paris, Bibliotheque Nationale, Pergament, 265x146mm
- B3.17 L17 Codex Valencianus latinus [B3.17.1]. Biblioteca Universidad de la Valencia, Pergament, 455x314mm
- B3.18 L18 Codex Mutinensis Estensis latinus 463.a.x.1.3 [B3.18.1]. 1466, Modena, Biblioteca Nazionale, Pergament, 455x305mm
- B3.18a L18a Codex Parmensis latinus 1635 [B3.18a.1]. Parma, Biblioteca Palatina, Papier, 417x295mm

- B3.19 L19 Codex Parisiensis latinus 4805. Paris, Bibliotheque Nationale, Pergament, 452x800mm

### **Zweite Redaktion von Donnus Nicolaus Germanus, 1466-1468, drei Tabulae Novae**

- B3.20 L20 Codex Florentinus Laurentianus Plutei, XXX,3, alte [B3.20.1] und B3.20N neue [B3.20N.2] Italienkarte. Firenze, Biblioteca Laurentiana, Pergament, 436x400mm. Dedikation für Fürst Borso d'Este
- B3.21 L21 Codex Zamommuskianus latinus, alte [B3.21.1] und B3.21N neue [B3.21N.2] Italienkarte. 1467, Warszawa, Biblioteka Narodowa, Pergament, 420x310mm. Dedikation für Papst Paul II.
- B3.22 L22 Codex Urbinas latinus 274. Roma, Biblioteca Vaticana, Pergament, 445x310mm. Dedikation für Papst Paul II.
- B3.23 L23 Codex Urbinas latinus 275. Alte [B3.23.1] und B3.23N neue [B3.23N.2] Italienkarte. 1467, Roma, Biblioteca Vaticana, Pergament, 422x291mm. Dedikation für Papst Paul II.

### **Dritte Redaktion von Donnus Nicolaus Germanus, 1468-1482, fünf Tabulae Novae**

- B3.24 L24 Codex Wolfeggianus latinus, 1468, Karten nach 1474. Schloß Wolfegg, Württemberg, Pergament, 443x238mm. Dedikation für Papst Paul II.
- B3.24a L24a Codex Monacensis Misc. Latinus 10691. München, Staatsbibliothek, Papier, 125x63mm
- B3.24b L24b Codex St. Gallensis Vadianus latinus. Sankt Gallen, Stadtbibliothek
- B3.24c L24c Codex Bruxellensis latinus 14887. Bruxelles, Bibliotheque Royale, Pergament, 651x450mm
- B3.24d L24d Codex Monacensis latinus 29. München, Staatsbibliothek, Papier, 422x300mm
- B3.24e L24e Codex Monacensis Misc.latinus 388. München, Staatsbibliothek
- B3.24f L24f Codex Newberiensis latinus. Chicago, The Newberry Library, Papier, 279x197mm
- B3.24g L24g Codex Parisiensis latinus 8536 [B3.24g.1]. 1475-1500, Paris, Bibl. Arsenal
- B3.24h L24h Codex Parisiensis latinus 11523. nach 1486, Paris, Bibl. Arsenal
- B3.25 L25 Codex Laurentianus latinus Plutei XXX,4. Alte [B3.25.1] und B3.25N neue [B3.25N.2] Italienkarte. Nach 1467, Firenze, Biblioteca Laurentiana, Pergament, 420x290mm. Dedikation für Fürst Borso d'Este
- B3.26 L26 Codex Vaticanus latinus 3811. Alte [B3.26.1] und B3.26N neue [B3.26N.2] Italienkarte. Nach 1468, Roma, Biblioteca Vaticana, Pergament, 422x292mm; Textteil ist Vat.lat. 3810

### **Handschriften anderer Bearbeiter, entsprechend der 3. Redaktion des Donnus Nicolaus Germanus**

#### *Pietro del Massaio und Hugo Comminelli, 1469-1482, 7 Tabulae Novae*

- B3.27N L27 Codex Vaticanus latinus 5699. Neue [B3.27N.1] Italienkarte. 1469, Roma, Biblioteca Vaticana
- B3.28 L28 Codex Urbinas latinus 277. Alte [B3.28.1] und B3.28N neue [B3.28N.2] Italienkarte. 1472, Roma, Biblioteca Vaticana Für Lesbarkeit optimierte Varianten [B3.28N.2X01-2X08] der neuen Italienkarte.
- B3.29 L29 Codex Parisiensis latinus 4802. Alte [B3.29.1] und neue B3.29N Italienkarte, Norden

[B3.29N.2] und Süden [B3.29N.3]

- B3.30 L30 Codex Wilton-Huntingtonianus H.M.1902. Nach 1482

#### *Francesco Berlinghieri, 1480-1500, 4 Tabulae Novae*

- B3.31N L31 Codex Mediolanus Braidensis italicus A.N.XV.2, Novella Italia. [B3.31N.1] 1480, Milano, Braidense
- B3.32N L32 Codex Urbinas latinus 273. Neue [B3.32N.1] Italienkarte. Roma, Biblioteca Vaticana
- B3.33 L33 Codex Parisiensis latinus 8834
- B3.34 L34 Codex Florentinus Laurentianus latinus Plutei, XXX,1. Alte Italienkarte Norden [B3.34.1] und Süden [B3.34.2], B3.34N neue Italienkarte Norden [B3.34N.3] und Süden [B3.34N.4]. Nach 1480, Firenze, Biblioteca Nazionale

#### *Henricus Martellus Germanus, 1480-1500, 12 Tabulae Novae*

- B3.35 L35 Codex Magliab. latinus CL. XIII,16. Alte [B3.35.1] und B3.35N neue [B3.35N.2] Italienkarte. Ca. 1490, Firenze, Biblioteca Nazionale
- B3.36 L36 Codex Vaticanus latinus 7289 [B3.36.1]. 1480, Rom, Biblioteca Vaticana

#### *Unbekannte Autoren, Zusammenhang mit den ältesten Karten der ersten Redaktion des Donnus Nicolaus Germanus*

- B3.37 L37 Codex Laurentianus latinus Plutei XXX,2. Norden [B3.37.1] und Süden [B3.37.1]. Firenze, Biblioteca Nazionale
- B3.38 L38 Codex Escorialensis latinus, e.d.1
- B3.39 L39 Codex Constantinopolitanus Seragliensis latinus 44
- B3.40 L40 Codex Wilczek-Braunianus latinus. Nach 1482
- B3.xx Lxx Italienkarte [B3.xx.1], Codex Helsinki (La Sfera) p154081 140003

### **4. Lateinische Ptolemaeus-Karten, Drucke (Auswahl)**

- B4.01 Bologna 1477. Crivelli: Gesamt [B4.01.1], Norden [B4.01.2] und Süden [B4.01.3].
- B4.02 Rom 1478. Calderini/Sweynheym/Buckinck [B4.02.1]
- B4.03 Berlinghieri 1482. Alte Italienkarte [B4.03.1](#), B4.03N neue Italienkarte [B4.03N.2], koloriert [B4.03N.3]
- B4.04 Donnus Nicolaus Germanus, Holl: Ulm 1482 [B4.04.1]
- B4.05 Donnus Nicolaus Germanus, Reger: Ulm 1486. Alte Italienkarte [B4.05.1], B4.05N neue Italienkarte [B4.05.2]
- B4.06 Pirckheimer 1525. Ptolemaeus, Claudius; Pirckheimer, Willibald; Regiomontanus, Johannes: Clavdii Ptolemaei Geographicae Enarrationis Libri Octo, 1525. München, Bayerische Staatsbibliothek. E.Slg/2 A.gr.b. 991, VD16 P 5211. Alte Italienkarte Norden [B4.06.1] und Süden [B4.06.2], B4.06N Neue Italienkarte Norden [B4.06N.3] und Süden [B4.06N.4]

### **5. Seekarten (Portolan-Karten) bis 1465 (Auswahl)**

- B5.01 Carte Pisane, 1290 (Cam002). Paris, Bibliotheque Nationale, Cartes et Plans, Res. Ge B 1118. farbig [B5.01.1] und schwarz-weiss [B5.01.2]
- B5.02 Angelino Dulcert, 1339 (Cam019). Paris, Bibliotheque Nationale, Cartes et Plans, Res. Ge B 696. [B5.02.1]
- B5.03 Albertin de Virga, 1409 (Cam046). Paris, Bibliotheque Nationale, Cartes et Plans, Res. Ge D 7900 [B5.03.1]
- B5.04 Battista Beccari, 1426 (Cam062). München, Bayerische Staatsbibliothek, Cod. icon. 130 [B5.04.1] und schwarz-weiss (Kamal) [B5.04.2], Mapp. XXV, 1y ?
- B5.05 Andrea Bianco, Portolan-Atlas, 1436 (Cam068). Tafel 4 [B5.05.1]
- B5.06 Petrus Roselli, 1449 (Cam083). Karlsruhe, Badische Landesbibliothek, S 6 [B5.06.1]
- B5.07 Anonym (Campbell: Petrus Roselli), Mallorca, 1450-1460 (Cam141). Modena, Biblioteca Estense, C.G.A.5.b. [B5.07.1]
- B5.08 Anonym, Mallorca, 1450 (Campbell: 1471-1485) (Cam204). Modena, Biblioteca Estense, C.G.A.5.d. [B6.08.1]

## A.2 Hinweise zur Benutzung von Recogito 2

- Voreingestellt wird das Inhaltsverzeichnis immer in Portionen von 10 Einträgen angezeigt; wenn man mehr auf einmal sehen möchte, muss man im Browser oben den Parameter &size=100 (z.B.) anhängen.
- Zur besseren Identifikation der Ptolemaeus-Karten wurde in die Metadaten aufgenommen:
  - (P) am Anfang steht für gedruckte Ausgaben.
  - (Lnn) sind die Nummern von Fischer für die lateinischen Handschriften.
- Grundsätzlich sollte mit der Maus gearbeitet werden. Die Grundeinstellung bei TOOLS (direkt über dem Bild) ist MOVE, d.h. man kann den Bildausschnitt bei gedrückter linker Maustaste verschieben.
- "HELP" gibt die folgenden Informationen:
  - Keyboard Shortcuts
  - SPACE toggle between MOVE and current annotation tool
  - Map Navigation
    - Hold SHIFT : to drag a box and zoom to the area
    - Hold SHIFT + Alt : to rotate the map around the center of the screen
  - Die Rotation macht man mit der linken Maustaste; das ist sinnvoll, wenn eine Beschriftung sehr schräg oder senkrecht oder auf dem Kopf steht.
  - Also: Mit SPACE kommt man nach rechts zu POINT (Standardeinstellung) oder RECTANGLE oder TILTED BOX, je nachdem, was zuletzt benutzt wurde.
- Zuerst (!) sollten Punkte gesetzt werden; dabei ist auf eine möglichst genaue Platzierung zu achten; evtl. bietet die Karte einen Anhaltspunkt durch einen Punkt, eine Stadtsilhouette, eine Lage am Fluss,

etc.

- Wenn es sich um kompliziert angeordnete oder sich auf grosse Bereiche beziehende Toponyme ("Italia") oder andere Bezeichnungen wie Völkernamen handelt, sind Rähmchen (Boxes) geeignet.

Das Rähmchen-Zeichnen geht einfach: Im RECTANGLE-Modus an eine Stelle klicken, die eine Ecke des Rechtecks werden soll, und dann (ohne die Maustaste zu drücken) die Maus bewegen und damit das Rechteck aufziehen und mit einem weiteren Mausklick abschliessen.

- Soll ein Punkt gesetzt und zusätzlich ein Rähmchen gezeichnet werden, ist "linked box" auszuwählen. Der erste Klick (ohne Halten) setzt den Punkt, der zweite den Anfang der Box, der dritte das Ende der Box-Basis und der letzte Klick die Höhe. (Späteres Hinzufügen eines Punkts zu einer vorhandenen Box ist derzeit nicht möglich - es erzeugt eine NEUE Annotation!).
- Danach öffnet sich das Annotationsfenster:
  - Toponym eintragen (unbedingt Gross- und Kleinschreibung beachten!)
  - "Place" anklicken. Entweder wird dann eine Vorschlagsliste präsentiert, aus der man auswählen kann; bitte *immer* in der Prioritätsfolge P (Pleiades), D (DARE = Digital Atlas of the Roman Empire), G (geonames). Wenn ein Ort nicht identifiziert werden kann ("No automatic match found"), weil z.B. die Schreibweise auf der Karte dem Gazetteer nicht bekannt ist, im Annotationsfenster "search" drücken. Dann öffnet sich in einem eigenen Fenster eine Vorschlagsliste (auch hier die Prioritätsfolge beachten) und eine leere Karte. Im Suchfenster oben links kann der übernommene Name geändert oder ein vermuteter (u.U. moderner) Name eingegeben werden. Oft findet dann Pelagios oder ein anderer Gazetteer eine Lösung. Durch Anklicken des Felds mit den drei Linien in der Karte kann man zu Ancient Places (DARE) - oft sehr hilfreich -, Modern Places (OpenStreetMap) oder einer Satellitenaufnahme wechseln. Bei Orten, die man geographisch nicht verifizieren kann, den Ort "flaggen" ("flag as unidentifiable") und ggf. einen Kommentar wie "In der Nähe von 'IDENTIFIZIERTEM\_ORTSNAMEN'" einfügen. Bei Unsicherheit bitte so lösen anstatt eine unplausible oder zweifelhafte Zuschreibung zu machen.
- Wird ein Völkernamen annotiert, "Person" anklicken
- Im Kommentarfeld sollte stehen, durch Komma getrennt Schriftfarbe: (schwarz), ggf. Schriftart, Lesbarkeit: gut/maessig/schlecht, Piktogramm: (z.B. Stadtsilhouette mit ..., Turm, Kreis, ...), Farbe: ... Und bitte *unbedingt* Zeilenvorschübe (in jedem Feld) vermeiden.
- Mit "Add Tag" kann man Fachtermini bzw. Schlagwörter zuordnen. Bei PLACEs sollte das nur verwendet werden, wenn "Flag as unidentifiable" benutzt wurde. Wird der Pleiades-Gazetteer verwendet, werden automatisch die dort enthaltenen Tags eingesetzt (s. Pleiades-Vocabularies.Feature\_Place\_Categories-TERMS.txt). Bei eigener Zuordnung sollten auch die Termini dieser Liste verwendet werden.

- Mit OK bzw OK & Next abschliessen.

Zur Identifikation von Orten in unklaren Fällen ist am ehesten noch die ptolemaeische Ortstabelle Italiens (EU06) aus der Berner Edition hilfreich (vorhanden als Scan/pdf des Italien-Kapitels, durchsuchbar, oder die Tabelle EU6 aus der dazugehörigen Datenbank, die ebenfalls mit Recogito annotiert wurde, s. 3.3). Wenn es darin keine plausiblen Treffer gibt und auch die moderne Karte nichts Klares bringt, dann muss man eben so verfahren wie oben beschrieben. Von den Ptolemaeus-Koordinaten in den genannten Tabellen kann man, wenn überhaupt, eigentlich nur die Breite berücksichtigen; die Länge müsste wegen des von Ptolemaeus zugrundegelegten zu kleinen Erdumfangs und seines "Nullmeridians" bei den Kanarischen Inseln erst aufwendig umgerechnet werden (grobe Differenz für Italien: 25°).

Eine Revision der eingezeichneten Rechtecke ist momentan noch nicht möglich. Es bleibt nur, die Annotation zu löschen und neu zu zeichnen. Wenn man alle Elemente der bestehenden Annotation löscht (Kommentare, Tags, georesolution, etc.), verschwindet auch das Rechteck. Um die Transkription selber zu verbessern, im Annotationsfenster auf das Dreieck rechts oben klicken und "Edit" oder "Delete" auswählen (ebenso im Kommentarfeld).

Da in Recogito keine *Suchfunktion* vorhanden ist, bietet sich folgende Möglichkeit an, wenn eine bestimmte Annotation gesucht und ggf. geändert werden soll: Jede Annotation hat eine eindeutige Identifikation (ID). Diese ist zu sehen, wenn man die Annotationen exportiert - bevorzugt im csv-Format. In der exportierten Tabelle steht dann diese Identifikation in der ersten Spalte. Damit kann folgendermassen gesucht werden: Hat man den gesuchten Eintrag, z.B. über seine Transkription, gefunden, wird die ID kopiert und im Adressfeld des Browserfensters hinter die aktuelle URL im Anschluss an "edit" mit einem vorangestellten "#" angefügt und die Suche mit "reload" gestartet, z.B.

[https://recogito.pelagios.org/ ... /part/1/edit#2a4a0681-eee9-42bd-90f7-d835e3a6cb6f](https://recogito.pelagios.org/.../part/1/edit#2a4a0681-eee9-42bd-90f7-d835e3a6cb6f) .

Dann öffnet sich das Annotationsfenster und kann bearbeitet werden.

---

Abbildungen:

1. Poster Biondo Daten
  2. Spatial Ontology Roles Relations (Kordjamshidi Diss.)
  3. White Latium (Ia) wordij-gephi Analyse
  4. MPIWG Infrastruktur Diagramm
-