



HRSM-Projekt

KONDE - Kompetenznetzwerk digitale Edition

1. Jahrestreffen, Salzburg, 13.04.2018



Das 2. Jahrestreffen 2019 findet in Klagenfurt statt, in der 2. Woche nach Ostern.



Sitzungsplan

13:00-14:30

- Bericht der Projektleitung
- Berichte der Standorte
- Berichte der ArbeitsgruppenleiterInnen

14:30-15:00: Pause

15:00-17:00

- Diskussion "Netzwerkstrategien"



Bericht der Projektleitung

- Jahresbericht 2017 vom Ministerium akzeptiert
- 1. Ausschüttung Sept. 2017 erfolgt und an die Partnerorganisationen weitergeleitet
- Website: Daten für Website eingelangt, Projektbeschreibung vollständig online.
- Google Drive: unspezifische Probleme in der Verwendung

Google Drive:

- Neue MitarbeiterInnen müssen bekannt gegeben werden, damit sie zum freigegebenen Ordner eingeladen werden können
- die Benutzung der Daten muss auf Eigeninitiative erfolgen

Bericht der Standorte

- Musil-Institut/Kärntner Literaturarchiv
 - Fortführung der Musil-Hybrid-Ausgabe
 - Kontinuierlicher Ausbau von MUSIL ONLINE (seit 1/2018 z.B. Musils Texte in historisch-kritischer Erschließung)
 - Beginn der Erstellung eines digitalen interdiskursiven Musil-Kommentars
- Institut für Germanistik
 - Grundlegende Überarbeitung des E-Mail-Korpus zur Alltagssprache (CODE-Alltag)
 - Anonymisierung und Pseudonymisierung individuenidentifizierender Merkmale
 - Entwicklung von Annotationsguidelines auf Basis der PHI-Kriterien
- Institut für Slawistik
 - Digitalisierung des überwiegenden Teils der Schülertexte des Projekts „Zweisprachiger Spracherwerb“
 - Transkription mit Transkribus, Auszeichnung slowenischer Texte mit Obeliks, deutscher Texte mit WebLicht; Anpassung an TEI P5 Standard

KONDE an der Alpen-Adria-Universität Klagenfurt

Bericht im Plenum

Kurze Zusammenfassung der KONDE-Aktivitäten an der AAU Klagenfurt von April 2017 bis April 2018 (max. 500 Wörter)

Am **Musil-Institut/Kärntner Literaturarchiv** wurde die Musil-Hybrid-Ausgabe fortgeführt, die bereits 2015 startete. Sie orientiert sich an der früheren Gesamtwerk-Edition der *Klagenfurter Ausgabe* (2009). Die Hybrid-Ausgabe besteht aus allen Schriften Robert Musils in Buchform, hg. von Walter Fanta (Verlag Jung und Jung, 2016-2022) und komplementär dazu der Internetplattform [MUSIL ONLINE](#). Parallel zur Buchausgabe, von der bereits 5 Bände erschienen sind, wird MUSIL ONLINE kontinuierlich ausgebaut. Im Bereich [Textwiedergabe](#) bietet das Portal seit Jänner 2018 exemplarisch Musils Texte in historisch-kritischer Erschließung an: Einzelne Nachlass-Mappen in hochauflösenden Scans mit Transkription werden laufend zur Verfügung gestellt und können im Format XML/TEI heruntergeladen werden.

Prototypisch wurden anhand eines Manuskriptkonvoluts aus dem Nachlass ([Mappe VI6](#)) erste Visualisierungen zur Repräsentation der Manuskripte (in HTML) erprobt und sind nun über MUSIL ONLINE abrufbar.

Mit März 2017 ist außerdem ein zweites, eng mit MUSIL ONLINE verbundenes FWF-Projekt angelaufen: die Erstellung eines digitalen interdiskursiven Musil-Kommentars für Forscher und Literaturvermittler als Zielgruppe. Mit ihm wird eine weitere tragende ‚Säule‘ von MUSILONLINE aufgebaut.

Da die in MUSILONLINE erarbeiteten best practices auf weitere digitale Editionen des Instituts angewendet werden sollen, ist begonnen worden, mit Hilfe von Transcribus Bestände zu Josef Winkler zu transkribieren.

Innerhalb des **Instituts für Germanistik** wurde mit einer grundlegenden Überarbeitung des bestehenden E-Mail-Korpus zur Alltagssprache (CODE_Alltag) begonnen. Sie betrifft zunächst die Anonymisierung und Pseudonymisierung individuenidentifizierender Merkmale. Dafür wurden Annotationsguidelines für die Anonymisierung auf der Grundlage der PHI-Kriterien (Personal Health Information) entwickelt. Diese bildeten die Basis für den Annotationsprozess, der von drei unabhängig voneinander arbeitenden Annotatoren durchgeführt wurde. Mit den auf diese Weise gewonnenen Golddaten soll ein Erkennungssystem für personenidentifizierende Merkmale mit Methoden des Maschinellen Lernens entwickelt werden, dessen Ausgabe in späterer Folge als Input für eine automatische Pseudonymisierung dient. Damit werden einerseits die Grundlagen für eine unrestringierte Weitergabe der Korpusdaten für CODE_Alltag gelegt, andererseits die Prozessierung des wesentlich größeren Archivkorpus ermöglicht. Gleichzeitig bildet dieses Korpus den Ausgangspunkt für die geplanten stilometrischen Untersuchungen. Darüber hinaus beschäftigt sich das Institut der Germanistik mit der Evaluierung der bestehenden technischen Rahmenbedingungen in Hinblick auf den Fortbestand und der Anpassung der Inhalte im Rahmen des Projekts „Transdisciplinarity in Austrian literature, art and culture of the Interwar Period / Transdisziplinäre Konstellationen in der Österreichischen Literatur, Kunst und Kultur der Zwischenkriegszeit“. Dabei geht es um die Entwicklung und Realisierung adäquater Lösungen für die Überführung bisher erhobener Daten in handhabbare Datenbankstrukturen und um die Entwicklung digitaler Instrumente zur Erschließung des in Aufbau befindlichen

Textarchivs.

Am **Institut für Slawistik** wurden zwei Projekte verfolgt:

1. Für das Projekt "Zweisprachiger Spracherwerb: Longitudinalstudie anhand schriftlicher Texte der Hermagoras-Volksschule: Digitalisierung, Auszeichnung, Auswertung" wurden Schulaufsätze einer zweisprachigen Schulklasse in deutscher und slowenischer Sprache von der zweiten bis zur vierten Schulstufe gesammelt. Bisher wurden davon die Texte der zweiten und dritten Schulstufe vollständig digitalisiert (eingescannt) und mit Transkribus transkribiert sowie für die weitere Auswertung mit Hilfe der Annotationswerkzeuge Obeliks (für die slowenischen Texte) und WebLicht (für die deutschen Texte) ausgezeichnet. Wegen der für Schülertexte typischen Abweichungen von der Standardsprache wurde der manuellen Korrektur der Auszeichnungen entsprechend Zeit und Aufmerksamkeit gewidmet. Dabei ergaben sich ähnliche Fragen wie bei der Editierung historischer Texte, wie z.B. die Auszeichnung uneindeutiger Textpassagen. Im Zuge dieser Korrekturen wurde zur Vereinheitlichung der beiden Auszeichnungsvarianten eine Anpassung an den TEI P5-Standard vorgenommen. Während bei den slowenischen Texten dafür nur geringe manuelle Anpassungen nötig waren, mussten die deutschen Texte mit einem selbstgeschriebenen Programm konvertiert werden. Die so vereinheitlichten Dateien stellen die Grundlage für die geplanten computergestützten Auswertungen zur Sprach- und Textkompetenz der Schülerinnen und Schüler dar.

2. Das vom ÖAD geförderte bilaterale Projekt „Unknown and Little-Known Manuscripts and Printed Texts of Older Slovenian Literature in the Wider Slovenian, Regional Austrian-Carinthian, and Austrian Context, and Applying Research Findings“ ist mit 01.01.2018 angelaufen und sieht eine Digitalisierung von Kärntner slowenischen Quellen vor.

Bericht: ASI

- Erstgespräche ZIM
 - Nachlass Karl Wiesinger
 - Tagebuchedition
- Sichtung des Materials
 - Beschlagwortung, Personenverzeichnis
- Umsetzungsstrategie

Tagebuch mit ca 700 Typoskriptseiten ist ein facettenreiches zeitgeschichtliches Dokument, das sich mit seinen unterschiedlichen Themenfeldern gut für eine Hybrid-Edition eignet, in der eine vertiefende Kontextualisierung digital erfolgt.

Im Netz der Zeitgeschichte. Die Linzer Protokolle des Autors Karl Wiesinger (1923-1991)

Am OÖ. Literaturarchiv/Adalbert-Stifter-Institut des Landes Oberösterreich wird seit 2015 der Nachlass des Linzer Autors und Publizisten Karl Wiesinger betreut. Sowohl in seinen literarischen als auch in seinen journalistischen Arbeiten beschäftigte sich Wiesinger intensiv mit der österreichischen Zeitgeschichte seit dem Bürgerkrieg im Jahr 1934, den er als Kind unmittelbar miterlebte. Auch seine weitere Biographie blieb eng mit den politischen Krisen und Entwicklungen verknüpft: Bereits in seiner Jugendzeit sympathisierte er mit der kommunistischen Idee; im 2. Weltkrieg als Soldat zur Wehrmacht

eingezogen, leistete er Agitationsarbeit gegen das NS-Regime und verbüßte deshalb eine längere Haftstrafe. Nach dem Krieg entwickelte er sich in Linz zum Schriftsteller und politischen Mitstreiter im KP-Umfeld.

Im Nachlass befindet sich ein unpubliziertes Tagebuch (ca. 700 Typoskriptseiten) aus den Jahren 1961 bis 1973, das Wiesingers Arbeit sowie seine Position im Linzer Kunstschaffen dokumentiert. Dieses Protokoll eines literarischen und politischen Außenseiters stellt ein facettenreiches zeitgeschichtliches Zeugnis dar, dessen Kontextintensität sich für eine digitale Edition in mehrerlei Hinsicht eignet. Nach einer ersten Auswertung haben sich einige zentrale Themenkreise herauskristallisiert, die im Rahmen der digitalen Edition eine verknüpfende Synopsis in Form von Einträgen, Kommentaren und Links entfalten sollen.

Themen/Felder:

- Schreiben: a) Hinweise und Komm. zur eigenen Arbeit; b) literarische/künstlerische Öffentlichkeit (Kollegen, Vereine, Institutionen); c) Einflüsse (Exzerpte, Zitate)
- Politik/Zeitgeschichte: a) Innenpolitik; b) internationale Politik (Kalter Krieg; imperialistische Konflikte)
- Biographie: a) persönliche Entwicklung; b) Linz als biographisches Umfeld (Orte, Szenen, Medien); c) erotische/sexuelle Begegnungen
- Reisen: Berichte über längere Aufenthalte in der UdSSR, CSSR, DDR, Jugoslawien, Frankreich



Bericht: IBK

- Brenner Archiv
 - Edition: Gesamtbriefwechsel Ludwig von Ficker, Familienkorrespondenz Ferdinand I., Tagebuchedition Joseph Maria Baernreither
 - literaturtirol.at, hier speziell Literaturlandkarte Tirol/Südtirol
 - Planung von Kooperation mit MEKNES-Projekt (Uni Bozen)
- READ/Transkribus
 - Ausarbeitung HowToGuides
 - Vorbereitung Webinare
 - Technische Betreuung Editionen/Transkriptionen (Nosceumus, Ambraser Heldenbuch)
 - Testprojekte: Robert Musil, Andreas Okopenko
- Ambraser Heldenbuch
 - Transkription
 - Einbindung in Lehrveranstaltung (Germanistik)
 - Webinterface (Prototyp)

Kurzbericht

Brenner-Archiv

Für die Editionsprojekte, die im Brenner-Archiv betreut werden, wird v.a. daraufhin gearbeitet, wie Editorinnen und Editoren in effizienter Weise edieren können. Neben der genuinen Editionsarbeit wurde v.a. an der DOCX zu TEI Konvertierung, die von TEI-Consortium entwickelt wurde, gearbeitet, welche individuell für jedes Projekt adaptiert wird. Hierzu gehören v.a. die semi-automatische Auszeichnung von Personen- und Ortsnamen, und die computer-gestützte Register-Erstellung. Daneben gilt es, Front- und Backend der Publikationsplattform zu programmieren und zu betreuen. Eine besondere Herausforderung stellt hierbei die Geo-Referenzierung literaturwissenschaftlicher Daten dar. Während die Einbindung von Normdaten (geonames, gnd und WikiData) bereits gut funktioniert, und die Darstellung der Geo-Referenzen auf einer Landkarte gelingt, haben wir unterschiedliche Darstellungsmöglichkeiten ausprobiert, die zumeist nicht zufriedenstellend sind, weil Benutzerinnen und Benutzer entweder nicht verstehen, was auf den Karten dargestellt werden, oder weil die Landkarten mit Daten überfrachtet werden, sodass

das Ziel einer übersichtlichen Darstellung verfehlt wurde.

DEA - Transkribus

Im Projekt wurden die folgenden Schwerpunkte gesetzt:

- Ausarbeitung neuer HowTo Guides für die Transkribus Plattform

Die bisherigen Darstellungen waren veraltet und das Wiki hat sich als weniger praktisch erwiesen, sodass von Johanna Walcher eine Serie von HowTo Guides erstellt wurde, die die wesentlichen Schritte beim Gebrauch von Transkribus erläutern:

- [How To Prepare Test Projects with Transkribus - for Archives and Libraries](#)
- [How To Transcribe Documents with Transkribus - Simple Mode](#)
- [How To Transcribe Documents with Transkribus - Advanced Mode](#)
- [How To Use Existing Transcriptions to train a Handwritten Text Recognition model](#)
- [How To Process Tables in Transkribus](#)
- [How To Train a Handwritten Text Recognition Model in Transkribus](#)
- [How To Transcribe Documents with Transkribus - Keyword Spotting](#)
- [How To Export documents from Transkribus](#)

Deutsche Übersetzungen sowie Videos sind in Vorbereitung.

- Unterstützung laufender Editionsprojekte an der Universität

Im Rahmen des ERC Advanced Grants NOSCEMUS wurden Mitarbeiter eingeschult, sodass sie nun selbstständig Modelle zur Druck- und Handschriftenerkennung trainieren können.

- Testprojekte

Mit vorhandenen Transkriptionen aus der Musil-Edition, sowie der Okopenko-Edition, die uns dankenswerterweise von Walter Fanta und Desiree Hebenstreit zur Verfügung gestellt wurden, konnten wir diverse Tests ausführen. Bei beiden Handschriften konnten sehr gute Ergebnisse erzielt werden.

Ambraser Heldenbuch

Im Projekt werden, nachdem die Transkriptionsrichtlinien erstellt wurden, die Texte des Ambraser Heldenbuchs in Transkribus transkribiert. Bei der Erarbeitung der Sonderzeichenauswahl und der Tags wurde darauf geachtet, dass das generierte Datenmaterial auf möglichst vielfältige Weise später genutzt werden kann. Eine erste mögliche Weiterverwendung des Datensets wird derzeit mit der Programmierung eines Webinterface erstellt, das die Transkription in verschiedenen Ausgabeformaten darstellen kann. Dazu werden die Transkription und das Manuskript nebeneinander angezeigt, welche miteinander verknüpft sind und bei Auswahl einer Zeile der

Transkription bzw. des Manuskripts die entsprechende Stelle in der anderen Ansicht auch hervorgehoben wird. Es kann zwischen Darstellungen der allographischen oder normalisierten Transkription sowie zwischen Manuskript- und Versumbrüchen gewechselt werden. Des Weiteren können Abbriviaturen aufgelöst werden und außerdem werden Initialen, Lombarden und Rubrizierungen grafisch in der Transkription dargestellt.

Bericht: KFU

- Projekte mit universitären PartnerInnen
 - Transkriptionsworkflow für allographische Transkription (Institut für Germanistik)
 - FERCAN (Institut für Alte Geschichte)
 - Roth Foto-Vorlass (Franz-Nabl-Institut)
- Projekte am ZIM
 - Fedora Migration
 - Kollaboration mit TUG: Graphenvisualisierung der Briefe des Hl. Patrick
 - Kofler Kommentar
 - Cantus
 - MWS
 - Illuminierte Urkunden
 - Stefan Zweig Nachlass Rekonstruktion
 - Medialität diplomatischer Kommunikation
 - Werktagebücher Hartmut Skerbisch

Germanistik: Transkriptionsworkflow für mal. Texte, der Micro-/ allografische Transkription berücksichtigt

Nabl: Erfassung und Annotation von Fotovorlass als Basis für zukünftige Editionen (Datenerfassung, Beschlagwortung mit Normdaten)

Alte Geschichte: Edition lateinischer Inschriften aus der Provinz Germania Inferior (Leidener Klammersystem vs. EPIDOC-TEI, Webpräsentation, erscheint auch in Druck)

Fedora-Migration: Umstellung der GAMS-Infrastruktur von Fedora 3 auf Fedora 4

Kollaboration mit TUG: Vergleich von Textvarianten durch Distant- und Close Reading



Bericht: KUG

- Erstgespräche
- Notenedition
- Fokus: MEI
- Kooperation mit dem ZIM: GAMS als Infrastruktur



Fuer eine Kooperation mit der Kunst Uni Graz wurden Erstgespraeche durchgefuehrt. Während der Kooperation könnte eine Edition mit MEI umgesetzt werden. Vertiefende Gespraechе mit dem ZIM sind fuer den Sommer geplant.



Bericht: OEAW

- Workflows
 - Named Entity Annotation Plugins: MS Word, Oxygen, Atom
 - XML Tokenizer- Webapplikation released
 - eXgit package (git-Integration in eXist-db)
- NLP
- Spacy Webapp - Wrapper für spacy.io NLP Library
 - TEIhencer – Named Entity Curation auf Basis von APIS
- Visualisierung
- Archivierung - Launch von ARCHE (Fedora 4-basiert)
- Projekte: Schnitzler Entity Hub, Ministerratsprotokolle Digital, Repertorium.at
- Mitarbeit in den AGs 3, 7, 10

1. Workflows & NLP-Processing

Angeregt durch die Diskussionen in AG 3 (Natural Language Processing, siehe auch unten), wurden bestehende Bearbeitungsworkflows weiterentwickelt bzw. wo notwendig um neue Komponenten erweitert. Folgende Ergebnisse sind dabei hervorzuheben:

Entwicklung von *eXgit* einer Erweiterung für die XML-Datenbank *eXist-db* zur Verwendung des Versionskontrollsystems *git* in XQuery. *eXgit* ermöglicht es, direkt in einer XML-Datenbank zu editieren und dennoch den Bearbeitungsverlauf unter

Versionskontrolle zu stellen; auch der umgekehrte Weg - programmatischer Import von XML-Dokumenten aus einem git-Repository - ist möglich.

Release eines konfigurierbaren XSLT-basierten Tokenizers für XML-Dokumente, der bestehendes Markup erhält und auf die Verarbeitung der Besonderheiten von TEI-Markup hin zugeschnitten ist. Der Tokenizer kann als XSLT-Library eingebunden werden oder über das *eXist-db*-basierte Webservice *xtx* (<http://xtx.acdh.oeaw.ac.at>) verwendet werden.

Mehrere Plugins für populäre Editoren zur Integration von externen Tools in die Arbeit mit XML-Dokumenten:

- Der *atom-entity-linker* erweitert den beliebten Texteditor *Atom* für die Annotation/Referenzierung von Entitäten wie z.B. Personen oder Orten
- *wEditor*, ein Plugin für Microsoft Word, übernimmt eine ähnliche Aufgabe (derzeit in der Beta-Phase)
- *acdh-oxygen-toolbox*: Ein generisches Oxygen-Plugin zur automatischen (linguistischen) Annotation von XML-Dokumenten mittels Webservices

spacyapp: ein Django-basiertes Webservice, das die Funktionen der Python-Library *spacy.io* (Tokenisierung, Lemmatisierung, Part-of-Speech-Tagging und Named Entity Recognition) als Webservice verfügbar macht. Das *nlp-module* prozessiert TEI-Dokumente in *eXist-db* mit *xtx* und *spacyapp*.

TEIHencer Web App: Eine auf APIS, dem *Austrian Prosopographical Information System* basierende Webapplikation (<https://teihencer.acdh.oeaw.ac.at>) zur

automatischer Auflösung von Ortsnamen gegen *geonames* und GND, ihrer Disambiguierung in einem Web-Frontend und Export als `<tei:listPlace>`.

2. Visualisierung

Generische Viewer-Applikationen zur Darstellung von spezifischen Datentypen wurden entwickelt, basierend auf *dsebaseapp*: Network-Vis, GeoVis, Timeline, Diagramme

3. Archivierung

Im Projektzeitraum wurde ARCHE, das Repository des ACDH-ÖAW für DH-Daten gelauncht und als erstes Repository in Österreich, dem die Fedora Commons 4 zugrunde liegt, mit dem *Core Trust Seal* und CLARIN B Centre akkreditiert. Der Code der selbstentwickelten Komponenten ist auf github verfügbar.

ARCHE wird ergänzt durch eine wachsende Reihe von spezialisierten Dissemination-Services, die zur Darstellung/Visualisierung von verschiedenen Datentypen dienen. Ein Beispiel ist der *TEIminator* (<https://teiminator.acdh.oeaw.ac.at/pages/index.html>) für die Darstellung von TEI-kodierten Dokumenten.



Bericht: OENB

- Evaluation verschiedener DE-Systeme
- Aufbau einer nachhaltigen Infrastruktur für digitale Editionen an der ÖNB
- Implementierung und Anpassung einer GAMS3 Installation
- NER mit SpaCy
- Transformationen für online Präsentation und Druckausgabe
- Erste Entwürfe zum Benutzerinterface
- Technische Unterstützung des FWF-Kooperationsprojektes zur Hybridedition von Tagebüchern aus dem Nachlass von Andreas Okopenko
- Entwicklung Workflow für Zusammenarbeit
- Aktive Teilnahme an fünf KONDE AGs (1,2,4,7,10)

KONDE an der ÖNB

Bericht zum Projektjahr 1

Stand 2018-04-05

An der Österreichischen Nationalbibliothek (ÖNB) wird innerhalb der Strategieperiode 2017-2021 im Rahmen der Strategie Digital Humanities eine nachhaltige Infrastruktur für Digitale Editionen aufgebaut. KONDE bietet den Rahmen, die Strategiegruppe personell zu unterfüttern. Ziel ist, eine technische Lösung für alle digitalen Editionen an der ÖNB bereit zu stellen. Dies gilt sowohl für hauseigene Projekte als auch für Kooperationen mit Forschungseinrichtungen, die mit Beständen der Österreichischen Nationalbibliothek arbeiten. Die nachhaltige Infrastruktur für digitale Editionen enthält sowohl eine umfassende Komponente für die Erstellung als auch eine Komponente für die langfristige Präsentation digitaler Editionen.

Da Nachnutzung vorhandener Ressourcen im Vordergrund unserer Strategie stehen, wurden zu Projektbeginn zunächst verschiedene Digitale Editions-Systeme hinsichtlich technischer und strategischer Aspekte evaluiert, mit dem Ziel eine Entscheidung für ein passendes System zu treffen. Evaluiert wurden folgende Backendsysteme: [cr-xq-mets](#), [eXist](#), [GAMS](#), [generic-de-](#)

[web-app](#), [Kiln](#), [SADE](#), [TextGrid](#). Die Entscheidung für Backend und Repository ist auf das Geisteswissenschaftliche Asset Management System (GAMS) der Universität Graz gefallen.

Die Erstellungskomponente der Infrastruktur gliedert sich in eine Erfassungsumgebung für TEI-XML Dokumente mit Modulen für die Bearbeitung (Editor, Anbindung zu Normdaten) und Vorschau. Weiters wird über die Erfassungsumgebung ein möglichst einfacher Zugang zu Quelldigitalisaten an der ÖNB ermöglicht sowie eine Versionierung via GIT für das XML-Quellmaterial gewährleistet.

Einschulungen zu Aufbau und Nutzung der Infrastruktur sowie Beratung zu Metadatenstandards als auch technische Unterstützung runden das Portfolio der Infrastruktur ab.

Parallel zur Entwicklung der Infrastruktur wird gegenwärtig für das laufende [FWF-Kooperationsprojekt "Andreas Okopenko: Tagebücher aus dem Nachlass \(Hybridedition\)"](#) gemeinsam mit der Universität Wien der technische Aufbau dieser Hybridedition konzipiert und entwickelt. Die geleistete Arbeit umfasst zum einen die Beratung des Editionsteams zu Fragen zur TEI-Kodierung, zum anderen hinsichtlich der Durchführung von automatisierten Abläufen (z.B. Anreicherung/Synchronisierung der TEI-XML Dateien mit Normdaten), als auch der Entwicklung von Mockups und Prototypen für die digitale Edition einschließlich der Transformationen via XSLT-Stylesheets für die verschiedenen Präsentationsformen der online Edition sowie der Vorstufe für die Druckedition in Absprache mit dem Editionsteam. Für das reibungslose Zusammenspiel der verschiedenen asynchron verlaufenden Prozesse zwischen Editions- und Entwicklerteam wurde ein passender Workflow konzipiert und umgesetzt.

Darüber hinaus wurde für eine digitale Edition des Bibliothekskatalogs von Hugo Blotius, dem ersten offiziellen kaiserlichen Hofbibliothekar der Wiener Hofbibliothek, aus dem Jahr 1576, Beratungen zur TEI-Kodierung des Quelltextes durchgeführt und erste Konzepte zur Gestaltung der hoch vernetzten digitalen Edition eines (historischen) Bibliothekskataloges diskutiert.

ÖNB-MitarbeiterInnen haben sich in folgende KONDE-Arbeitsgruppen aktiv eingebracht: AG 1 Digitalisierung, AG 2 Transkription, AG 4 Interfaces, AG 7 Archivierung, AG 8 Kommentar, AG 10 Hybridedition. Letztere wird von einer ÖNB-Mitarbeiterin moderiert.

Liste der Projekte, die im Rahmen von

KONDE erarbeitet werden

- Aufbau einer nachhaltigen Infrastruktur für digitale Editionen an der ÖNB
- Technische Unterstützung und Metadatenberatung für das FWF-Kooperationsprojekt „Andreas Okopenko: Tagebücher aus dem Nachlass (Hybridedition)“
 - NER, GAMS, TEI-Beratung
- Technische Realisierung und Metadatenberatung der Digitalen Edition des Blotiuskataloges
- Erste Kooperationsgespräche zu Musil-Hybridedition
- Unterstützung in technischen Belangen für Anträge zu DE von an der ÖNB beherbergten Objekten



Bericht: SBG

- Aufbau einer Plattform für den sicheren und dauerhaften Betrieb von DH Projekten
- Ausschreibung und Besetzung der Stelle für Plattform-Administrator/Entwickler
- Organisatorische Schritte und Projektmanagement
- Projekt MHDBDB
 - Datenmodellierung
- Projekt „Stefan Zweig Digital SZD“
 - Datenmodellierung
 - Archivierung

Ziel der Teilnahme der Universität Salzburg PLUS an KONDE ist es, mit Hilfe der HRSM Mittel eine Plattform aufzubauen, die den sicheren und dauerhaften Betrieb von Digital Humanities Projekten der Universität gewährleistet. Die Entwicklung der Plattform soll anhand von 3 Pilotprojekten erfolgen: Ebner Online, Mittelhochdeutsche Begriffsdatenbank und Stefan Zweig Digital.

Der Schwerpunkt der Aktivitäten im ersten Projektjahr lag in der Schaffung und Besetzung einer Stelle an den IT-Services der PLUS mit einer Person, die diese Arbeiten durchführen kann und in weiterer Folge für den Betrieb der Plattform zuständig sein wird. Nach 2 maliger Ausschreibung konnte die Stelle mit Daniel Schlager besetzt werden, der seinen Dienst am 1.12.2017 antrat.

Nach dieser erfolgreichen Personalentscheidung war der nächste Schritt die Vergabe des Namens „dhPLUS“ für die Plattform und die Entscheidung für eine von insgesamt drei Logo-Varianten.

Zugleich wurde mit der Arbeit an dem Lastenheft für die Plattform begonnen, das aus 2 Teilen besteht: einem allgemeinen Teil für die Plattform selbst und einem darauffolgenden Teil für

die einzelnen Pilotprojekte. Eine erste Version des allgemeinen Teils des Lastenheftes wurde am 22.3.2018 präsentiert und unter Auflage von einigen Änderungen akzeptiert.

Parallel dazu erfolgte von Seite des MHDBDB Projektes eine erste Fassung seines Lastenheftes.

Um diese zentralen Aktivitäten herum gab es organisatorische Aktivitäten wie die Vergabe der Rollen und Aufgaben im Projektteam, es wurde an der für die Migration bzw. Entwicklung erforderlichen Datenerstellung und Datenmodellierung einzelner Pilotprojekte gearbeitet, es wurden Tools für die Dokumentation des Projektes eingerichtet, es gab die Teilnahme von Projektmitarbeitern an diversen, für die Plattform interessanten Veranstaltungen und Treffen und es gab die Entscheidung, die nächste DHA an der PLUS durchzuführen und sie zu organisieren.

KONDE-Aktivitäten

1. Projektjahr

•Ziel der Teilnahme der PLUS an KONDE: Aufbau einer Plattform für den sicheren und dauerhaften Betrieb von Digital Humanities Projekten der Universität Salzburg PLUS.

•Aktivitäten dazu im ersten Projektjahr:

–Stellenausschreibung für Plattform-Administrator/Entwickler

–Bewerbungsgespräch

–Besetzung der Stelle mit Daniel Schlager

–Rollenvergabe für die Mitarbeit an KONDE PLUS

–Vergabe eines Namens für die Plattform: dhPLUS

–Erstellen eines Logos für dhPLUS

–Erstellen einer ersten Version für das Lastenheft der Plattform

–Erstellen einer ersten Version für das Lastenheft des Pilotprojektes „Mittelhochdeutsche Begriffsdatenbank MHDBDB“

–Datenmodellierung für MHDBDB

–Datenmodellierung für das Pilotprojekt „Stefan Zweig Digital SZD“

–Archivierung des Stefan Zweig Teilnachlasses am LAS für SZD

–Titelaufnahme von Nachlassmaterialien für SZD in ALEPH/ALMA

–Einarbeiten in die für dhPLUS relevanten IT-Systeme der IT-Services

–Projektmanagement und -dokumentation anhand eines Projekthandbuches

–Weitere Dokumentation mit Confluence Tool der ITS und RedMine

–Organisation und Durchführung von 12 Meetings, darunter ein kick off Meeting nach der Stellenbesetzung und eines zur Präsentation des Lastenheftes

–Teilnahme am Wiener Arbeitsgespräch der AG2 „Annotation versus Kommentar“ mit einem

Impulsreferat über Ebner Online

- Teilnahme an der Veranstaltung der Uni Wien „Herausforderungen für Repositorien“
- Teilnahme an der DHA Konferenz in Innsbruck
- Teilnahme am Vernetzungstreffen von Repositoriumsmanagern an der Uni Wien
- Teilnahme an der DhD in Köln
- Entscheidung für die Ausrichtung der DHA 2018 in Salzburg
- Rollenvergabe für die Organisationarbeit an der DHA 2018
- Beginn der Organisationsarbeit für die DHA 2018



Bericht: TUG



- Kooperationsgespräche mit dem ZIM -ACDH
- Feature Engineering und Analyse von Daten in sozialen Medien
- Graphvisualisierung in digitalen Editionen
 - Beispiel der Briefe des Heiligen Patrick
 - Entwicklung eines Prototyp
- Visualisierung von Korrespondenznetzwerken (CorrespSearch, Rollett Briefedition)
- Dramaanalyse mit Daten aus der Ödön von Horváth digitalen Edition

KONDE Bericht ISDS TU Graz, 2017

Im ersten Jahr von KONDE wurden am Institute for Interactive Systems and Data Science (ISDS) zwei Teilprojekte behandelt. Zum einen wurde in Zusammenarbeit mit dem Zentrum für Informationsmodellierung (ZIM) ein Visualisierungstool entwickelt, das bei der Darstellung von Unterschieden in Texten von mehreren Textzeugen helfen soll. Hierfür fand Ende 2017 ein erster Datenaustausch statt und Anfang 2018 wurde mit der Entwicklung des Visualisierungstools begonnen. Die erhaltenen Daten beinhalten zwei unterschiedliche Texte, welche beide von 7 Textzeugen transkribiert wurden. Neben einer Liste mit Wörtern pro Zeitzzeuge sind noch andere Attribute vorhanden (z. B. ob ein Wort eine Abkürzung ist oder nicht). Geplant ist eine Visualisierung unter Zuhilfenahme von Graphen, die einerseits auf bestehenden Methoden aufbaut und andererseits auch neue Informationen in die Visualisierung aufnehmen. Ende Februar 2018 wurde ein erster Prototyp fertiggestellt. Die Bedienung kann über alle modernen Webbrowser erfolgen. Zu einem späteren Zeitpunkt soll das Tool in die bestehende Infrastruktur des ZIM aufgenommen werden und in einer Publikation vorgestellt werden.

Für 2018 ist geplant das Tool in einer Publikation im Bereich Digital Humanities, voraussichtlich in einem Journal, vorzustellen.

Im zweiten Teilprojekt beschäftigen wir uns mit der Analyse von kontroversen Themen in sozialen Medien. In einer ersten Analyse und Evaluierung berechneten wir

unterschiedliche Features, die es uns ermöglichen sollen kontroverse Nutzerkommentare von nicht kontroversen unterscheiden zu können. Als Datengrundlage diente das soziale Netzwerk Reddit.com, das seinen Benutzer das Teilen von jeglichen digitalen Inhalten erlaubt. Unsere Features berechnen wir auf unterschiedliche Arten und wir können sie daher in drei Kategorien aufteilen. Zuerst erstellen wir Features die auf textuellen Eigenschaften beruhen. Diese enthalten einfache Metriken, wie zum Beispiel die Anzahl der Worte oder ganzer Sätze in einem Kommentar, aber auch kompliziertere Metriken, wie zum Beispiel die Lesbarkeit von Kommentaren oder die Emotionen die damit übermittelt werden. Weiters studierten wir zeitliche Zusammenhänge und analysieren Features, wie z. B., die Wochentage an denen mehr kontroverse Kommentare entstehen als an anderen Tagen. Die dritte Features-Kategorie beinhaltet strukturelle Eigenschaften, die entstehen, wenn Benutzer miteinander interagieren.

Wir planen diese ersten Ergebnisse auf der ASONAM 2018 einzureichen und zu veröffentlichen. Nach erfolgreichem Abschluss dieser ersten Analyse planen wir weitere Veröffentlichungen, die uns weiteres Verständnis über die Benutzer im Zusammenhang mit kontroversen Themen in sozialen Netzwerken liefern sollen. Abschließend sollen kontroverse Themen automatisch erkannt werden.



Bericht: VIE

- Stemmaweb
- Online-Kommentar zur Kofler Gesamtausgabe
- Pez Project
- Reassembling the Republic of Letters
- Mitarbeit an den AG 5, 8

Stemmaweb & Stemmarest: Graph-based digital edition frameworks

Prof. Tara Andrews (Institute for History)

Stemmaweb, a front-end editing interface using graph visualisations to allow editors to work across all witnesses of a transcribed digitised text, and Stemmarest, a neo4j based graph database system for text storage, interface together to form a flexible system for edition editing and stemma analysis. Stemmaweb provides a word by word diagram of the text, allowing editors to see where witnesses agree on a particular word and where they diverge, with tools for tagging words, selecting lemmas, and splitting or merging nodes. Potential future additions include expanding functionality for stemma analysis, and using the graph data functionality of Stemmarest to provide tag-linkages from the text to dates or prosopographical information to allow for more developed commentaries.

The full Stemmaweb/Stemmarest system is being developed with the intention of initially using it to work towards a complete digital edition of the Chronicle of Matthew of Edessa, an Armenian historian writing in the early to mid 12th century.

The Annotated Complete Works of Kofler: A Hybrid Edition

Dr. Claudia Dürr, Dr. Wolfgang Straub (Institute for Germanic Studies)

This hybrid edition aims to prepare Kofler's complete prose works in two related editions – as a book and as a website. The primary text published in book form will ensure a compact, affordable and easily readable complete edition of Kofler's works. In addition to the primary text and commentary, the print version includes extensive explanatory postscripts.

The digital utilisation of the commentary makes new possibilities for access to Werner Kofler's work possible. Owing to the unlimited space, the website can contain extensive image material and thereby offer low-threshold access to Kofler's work: archive material from the estate will be assigned to individual commentaries in order to make the context in which the texts were written "visible". Sources and their utilisation in literary text are linked, which facilitates conclusions on Kofler's working methods. The information can be processed by the recipients using various tools and thus provides access to the Koflerish oeuvre for readers and visitors as well as for literary scholarship.

The Pez Project & Reassembling the Republic of Letters

Dr. Thomas Wallnig (Institute for History)

The Pez project aims to construct a hybrid edition of the complete letters of the eighteenth century brothers Bernhard and Hieronymus Pez, both Benedictine monks, both incorporating a digital text edition and digitised papers in a virtual research environment. The current project, "Benedictines, Church Reform and the State in Austria: 1720-40", has run since 2016 and is a follow-up to the preceding "Monastic Enlightenment and the Benedictine Republic of Letters".

Dr. Wallnig is also vice-chair of Reassembling the Republic of Letters, which aims to link up and envisage open-access digital infrastructure for letter collections from the period 1500-1800.

Bericht der AG LeiterInnen



Bericht: AG 1 - Klug

- Standards für Quelledigitalisierung
- Metadaten und Metadatenstandards
- Leitfaden Digitalisierung

In Ausblick auf das Weißbuch sollen Richtlinien zur Digitalisierung und für die Digitalisate begleitende Metadaten gesammelt werden. Dabei steht das Sammeln von bestehenden Leitfäden im Mittelpunkt. Es werden Informationen, Adressen und Angebote von Digitalisierungstellen gesammelt. Ein spezielles Thema, weil vom bibliothekarischen Zugang abweichend, ist der Umgang von Archiven mit Digitalisierung.

Ideen für Weißbuchinhalte:

- Angebot an Digitalisierungseinrichtungen/Angeboten österreichweit
- Umgang mit den (privaten) Digitalisierungsangeboten in zB. Wien
- Best-Practice Digitalisierungsworkflow (Qualität, Benennung, ...)
- Bibliothekssystem und Archivsystem erhalten zu versuchen
- Bildmetadaten (exif)
- Embedded-Librarian-System einbeziehen
- Bewusstseinsbildung
 - Open Science
 - Ressourcenschonung
 - Open data
- Österreichbezogen: Was muss ich beim Umgang mit der Institution X beachten?
 - Leitfäden / Checkliste

- Checkliste Edition
- Arbeitsweise Archiv/Bibliothek vs. Wissenschaftler
- Transformation in Standards (RNA/RNAb, HS-Beschreibung, ...)
- Use-Cases (corema, Stefan Zweig, Musil-Archiv)
- Hilfestellung für Anträge / Richtlinien für Fördergeber (sollte von KONDE politisch durchgesetzt werden! U.a. Datenmanagementplan)
 - Finanzierung und Umfang von digitaler Umsetzung
 - Wer sollte informiert werden: Rektorenkonferenz, FWF (Fr. Maruska: Was gibt es dann da für Vorgaben), FFG, Nationalbank, Länder, ÖAW, Tiroler Wirtschaftsförderungsfond... andere Fördergeber
- Checkliste Digitalisierung (Abläufe)
- Austauschbarkeit der Digitalen Daten
- Datenmanagementplan



Bericht: AG 2 - Zeppezauer-Wachauer

- Bisher zwei Treffen der AG (bei der DHa in Innsbruck und heute vor dem Jahrestreffen)
- "Schaukasten" mit Kodierungsbeispielen
- AG2 erarbeitet gemeinsam einen Fragebogen fuer eine Umfrage zum Thema "Texttranskription und Textauszeichnung"
- Die Ergebnisse der Umfrage werden als Grundlage für das Weißbuchkapitel dienen.

Erhebung zu Transkription und Textauszeichnung mithilfe eines Fragebogens:

AG2 erstellt gemeinsam einen Fragebogen, der anschließend an die Evaluierung/Emendation bei der AG-Besprechung am Freitag 13.4. an alle KONDE-Mitglieder ausgeschickt werden und als Grundlage für das Weißbuch Kapitel dienen soll.

Schaukasten:

https://drive.google.com/drive/u/0/folders/1ln0YeFKJvn2taaXX_S2zfNQ2SCikpEm1



Bericht: AG 3 - Vogeler

- Arbeit an Annotationsbewahrendem Tokenizer (ACDH-ÖAW)
- Evaluation von NLP-Methoden für lateinische Texte (VIE)
- Sammlung von NLP-Methoden für historische Sprachstände (KFU)
- Erkennen von "Personen identifizierenden Angaben" mit Hilfe neuronaler Netze (AAU)

In der AG 3 wird an einem annotationsbewahrendem Tokenizer gearbeitet, NLP Methoden fuer lateinische Text werden evaluiert.
NLP-Methoden fuer historische Sprachstaende gesammelt.
Weiters wird an der Erkennung von Personen identifizierenden Angaben mit Hilfe neuronaler Netze gearbeitet.



Bericht: AG 4 - Bleier

Bisher zwei Treffen der AG (bei der DHa in Innsbruck und heute vor dem Jahrestreffen)

Themen:

- Sammlung und Evaluierung von existierenden Interface Lösungen
- Testen von generischen Lösungen zur Implementation von User Interfaces (ÖAW, OenB, ZIM)
- Usability und User-studies (MHDBDB, Transkribus)
- Kollaborative Interfaces (MHDBDB)
- Interface und Kommentar (Musil Online)

19

In der AG4 wurde mit dem Sammeln und Evaluieren von existierende Loesungen fuer User Interfaces und Elemente in User Interfaces begonnen.
Weitere Themen die diskutiert wurden und in Hinblick auf das Weissbuch interessant sein koennten sind Fragen von Usability und User-studies, kollaborative Interfaces und User Interface und Kommentar

1) Struktur des Kapitels User Interfaces im Weißbuch

Erweiterung von oXygen/Plugins

User Testing: Interface App Testing

User Group:

Software Testing:

Entwicklergruppe hat automatisierte Tests

User Gruppen: Usability Studien

Transkribus:

Nicht jeder Entwickler ist happy gleich mit Usern zu

reden

Neuer Entwickler: er redet zuerst mit Usern

Johannes Stigler;

Forschungsprojekte: Usability muss anders gedacht
werden

Projektleiter ist der Hauptuser

Anderer Zugang als in der Wirtschaft

Wissenschaftlicher Zugang, Editor als Hauptuser




Bericht: AG 5 - Andrews

- Arbeitsgruppe ist im Aufbau



Bericht: AG 6 - Helic

Aktuelle Arbeitsthemen:

- Feature Engineering und Analyse von Daten in sozialen Medien
- Graphvisualisierung in digitalen Editionen am Beispiel der Briefe des Heiligen Patrick

Geplante Arbeitsthemen:

- Visualisierung von Korrespondenznetzwerken mit Daten von CorrespSearch oder der Rollett Edition (ZIM)
- Dramaanalyse mit Daten aus der Ödön von Horváth digitalen Edition

In der AG 6 wird gegenwaertig an feature engineering und analyse von daten in sozialen Medien und an Graphvisualisierung gearbeitet. Fuer die Zukunft ist geplant sich naeher mit der Visualisierung von Korrespondenznetzwerken und Dramaanalyse zu beschaeffigen.



Bericht: AG 7 - Stigler

Bisher zwei Treffen in Innsbruck bzw. Salzburg

Themen:

1. Bestandsaufnahme: Standortbezogene Lösungen dominieren
2. Grundsatzdiskussionen zu
 - Gesamtösterreichische Lösung denkbar?
 - Metadatenstandards
 - APIs für Digitale Editionen
3. ÖAW: Lösungen auf Basis Fedora 4 bzw. eXist DB
4. Salzburg: entwickelt Plattform auf Basis MongoDB
5. ÖNB nutzt GAMS (Fedora 3) nach
6. ZIM: Migration der Plattform nach Fedora 4 läuft

22

AG Treffen gestern hat starkes Interesse an einer Art Gesamtlösung ergeben: die in unterschiedlich intensiven Stufen realisiert werden könnte:

- gemeinschaftlicher Katalog von Editionen
- gemeinsame und daher kompatible Archivarchitektur
- aufbauend: vernetzter Datenpool

Bericht: AG 8 - Straub, Dürr

- AG Treffen in Wien (Sept. 2017)
- Bisher wurde die Herausforderungen der Kommentierung diskutiert
 - Definition des Kommentierungsbedürftigen
 - Zielgruppen
 - 'interdiskursiver'/offener Kommentar
- Nächsten Schwerpunkt werden wir das Thema Textkritik
 - Fokus: Gegenwartsliteratur
 - Visualisierung von Textgenese
 - Einstiegsmöglichkeiten in Textkorpora
 - Theoretische Konzepte (z.B. Critique Genetique, etc.)

Die Arbeitsgruppe Kommentar & Textkritik hat sich bislang mit Herausforderungen der Kommentierung

befasst und diese in zwei Arbeitsgesprächen am Wiener Institut für Germanistik diskutiert, etwa:

Definition des Kommentierungsbedürftigen, Frage der Zielgruppe(n), Möglichkeiten ‚interdiskursiver‘/offener Kommentierung. Fazit bislang: Wichtig ist Kooperation (bis hin zur

Zusammenführung von Kommentaren), Zurückgreifen auf Vorleistungen, gemeinsame Standards für Datenbanken.

Als nächsten Schwerpunkt werden wir das Thema Textkritik behandeln. Auch hier gilt es wieder, auf die

spezifischen Gegebenheiten der Edition von Gegenwartsliteratur einzugehen (etwa, welches Gewicht

haben Entscheidungen des Autors; juristische Fragen, ...) und digitale Vermittlungsmöglichkeiten zu

diskutieren (etwa: wie kann man Textgenese visualisieren, welche Einstiege in das Textkorpus bietet

man, welche theoretischen Konzeptionen können fruchtbar gemacht werden (Critique Genetique), ...).




Bericht: AG 9 - Vogeler

- Strategische Ziele von KONDE
- Diese Gruppe wird zu einem späteren Zeitpunkt aktiv werden



Bericht: AG 10 - Fritze

- AG-Treffen in Innsbruck und Salzburg
- Telefonischer und schriftl. Kontakt zu allen aktiven AG-Mitgliedern
- Sammlung österr. Hybrid-Editionen
- Praxis-Beispiel
 - (Pez-Editionsprojekt vorgestellt von Dr. Victor Wang, Böhlau Verlag)



Zusammenfassung und Ausblick

- Blog
 - Soll aktiv betrieben werden
 - Liste mit "Publikationsterminen" folgt
- Twitter #kondeAT
- Berichtlegung 2017/18:
 - Aufforderung dazu erfolgt gesondert
- Workshop "Digitale Edition: Ecosystem & Workflow"
 - Ende September in Graz
- Konzeptionierung Weißbuch
 - Auf Basis der heutigen Diskussion und des Workshops im Juni

Blog: Excel mit Terminen vorbereiten



Arbeitsauftrag Kaffeepause

- Nimm 4 Aufkleber mit der Farbe, die wir deiner Institution zugeordnet haben.
- Klebe 2 Aufkleber auf jene Bereiche der Poster, die dich in Bezug auf Angebot am ehesten betreffen bzw. 2, die dich in Bezug auf Bedarf am ehesten betreffen.
 - Angebot: Das machen wir mit unserer Edition, da können wir unterstützen.
 - Bedarf: Das würden wir in unserer Edition gerne machen, brauchen aber Tipps und Unterstützung.

2 Listen mit Partnerinstitutionen => Farbpunkte dazukleben
2x A4: Angebot, Bedarf

KAFFEPAUSE

**Wahre Mitteilung findet nur
unter Gleichgesinnten,
Gleichdenkenden statt.**

Novalis (1772 - 1801)

—

29

Diskussion: Input-Statement

Es gibt zarte Pflänzchen der Zusammenarbeit

Wie kanns im nächsten Jahr besser gemacht werden?

Diskussion: Standortbestimmung

- Angebot: Publikation, Transkription, Textkommentierung u. -erschließung
- Bedarf: Publikation, Archivierung, Nachnutzung, Annotation-Metadaten

Aspekte des klass. Edierens wenig präsent

AG für Textkritik notwendig, geht aus AG Kommentar hervor

AG 2 überhaupt noch notwendig? Ja, man kann zB die Werkzeuge der TEI verbessern (zB. Beispiele)

Fokus auf Text um andere Medien ausweiten? Erhebung notwendig ...

30

Wo sind Schwerpunkte?

=> Poster

=> Punkteverteilung

=> Schwerpunkte bei Angebot, Bedarf => kommentierende Statements

Was wird gebraucht, Was soll gemacht werden?

Was wird gemacht, Was wird geboten?

Man muss unterscheiden zwischen Material und Personal und Technologien
Beispiel Projektdokumentation

Bedarf im Bereich Publikation:

Kommentar: Man muss unterscheiden zwischen digitalen Bedarf, Print, etc. (wird eigentlich vom Poster abgebildet)

Analyse von digitalen Editionen:

Fazit: es wird doch sehr viel auch analysiert und gearbeitet mit Digitalen Editionen.
Aber: Daten exportieren und mit anderen Mitteln analysieren, nicht mehr primäre Aufgabe der Edition selbst (Wallnig)

Arbeitsgruppen:

Es gibt zwei Gruppen der Textkodierung, AG 2 und AG 5. Frage: Sollte eine aufgelöst werden? Dafür müsste die AG 5 aktiv werden: Treffen der AGs zum Gedankenaustausch

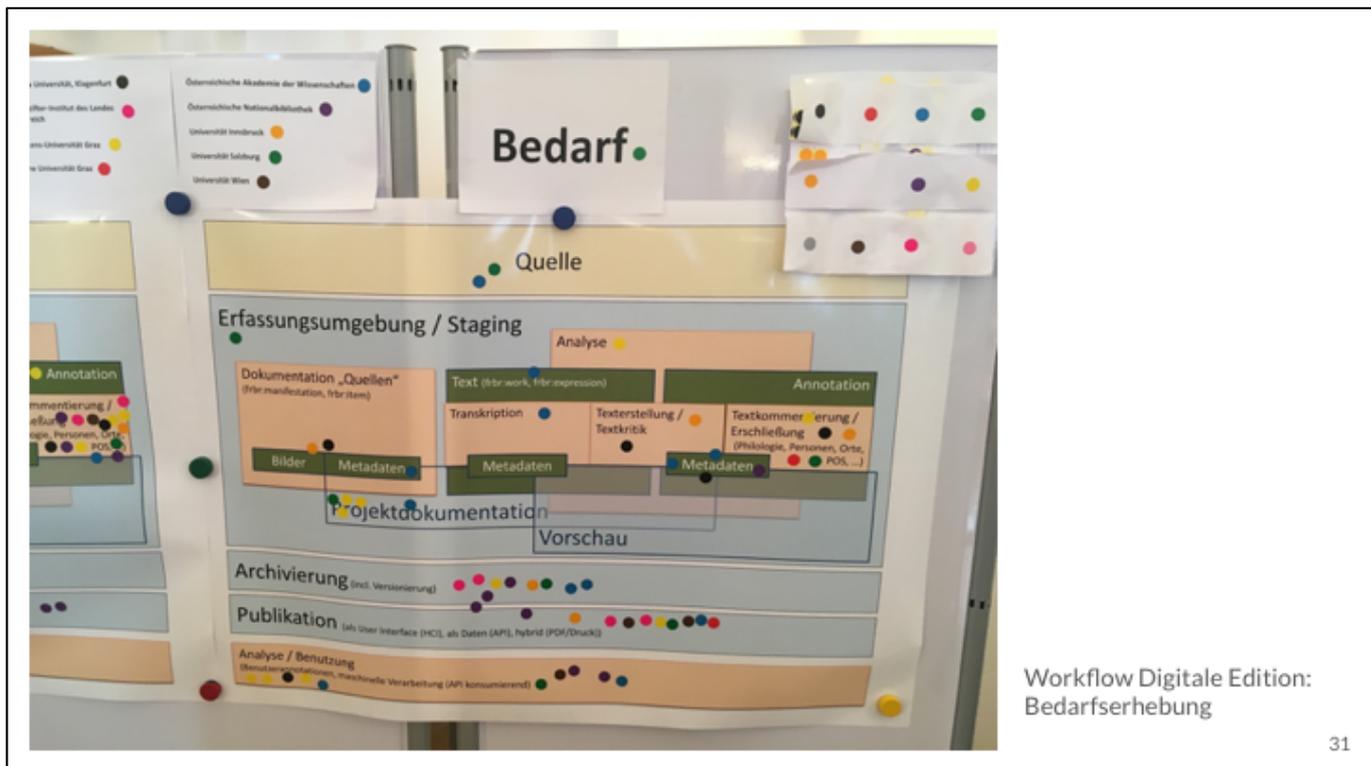
Fanta: Es gibt genug Diskussionsbedarf über die Leistungsfähigkeit von TEI. TEI als Standard kann nicht in Frage gestellt werden. Dafür ist es zu wichtig.

=> Verbesserung der TEI Guidelines durch Beispiele von Musil, 2-tägiger Workshop im März 2019 in Klagenfurt zu diesem Thema - Andrews und Bailey sollten unbedingt mitmachen (Vogeler)

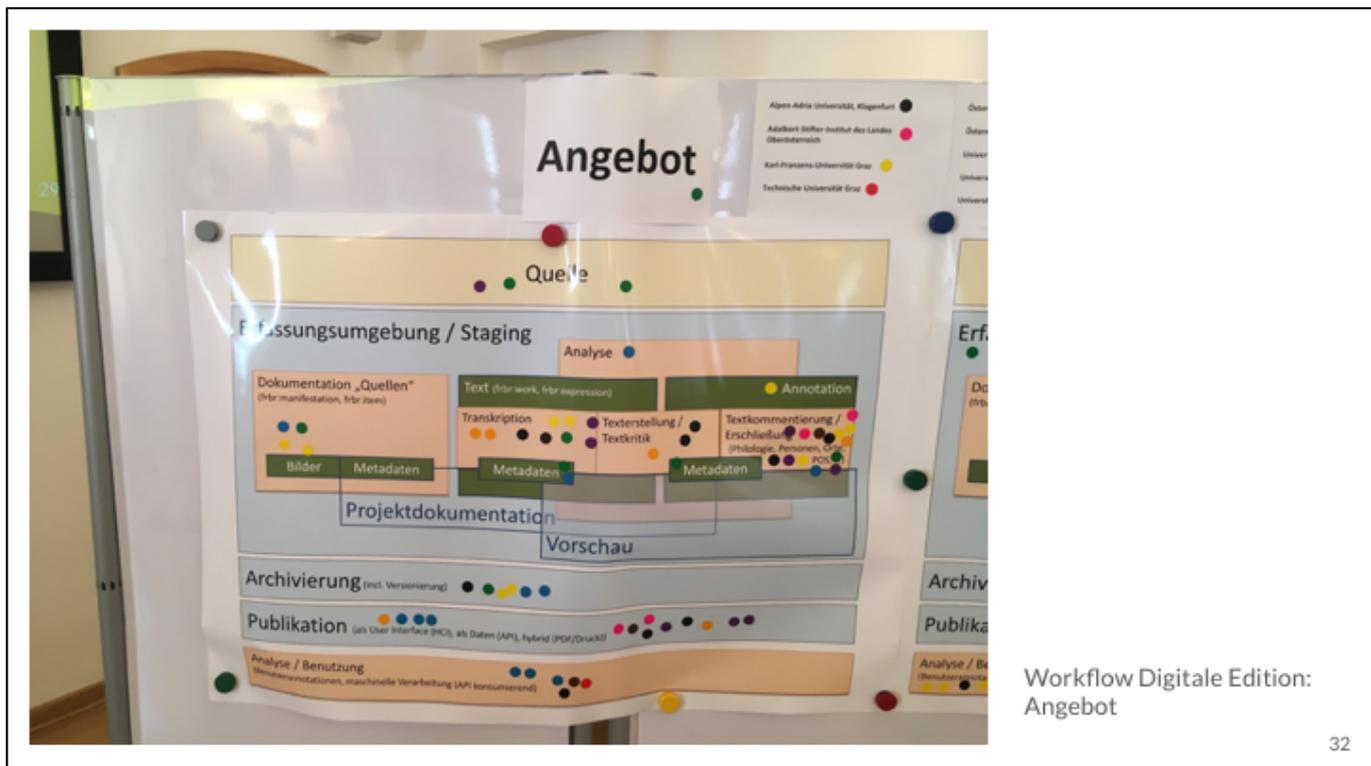
Diskussion in AG 2: alle Beispiele sind sehr textlastig

Kontakte herstellen und Erheben von Leuten, die andere Objekte editieren, Umfrage geplant (Zeppezauer); aber Umfrage sinnvoll, damit jene die es machen kommunizieren können (Vogeler); aber: Projektfokus Text, nicht zu sehr ausweiten (Stigler)

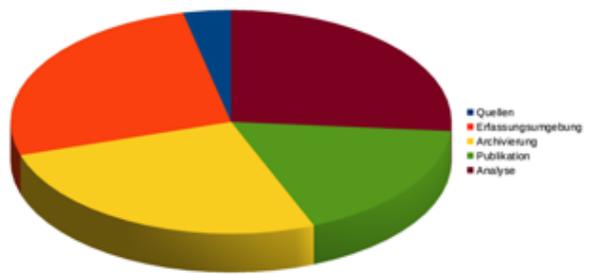
AG 8: Verschiebung des Fokus: Übergang von Kommentar zu Kritik, kein Diskussionsbedarf mehr bei Kommentar (Dürr)



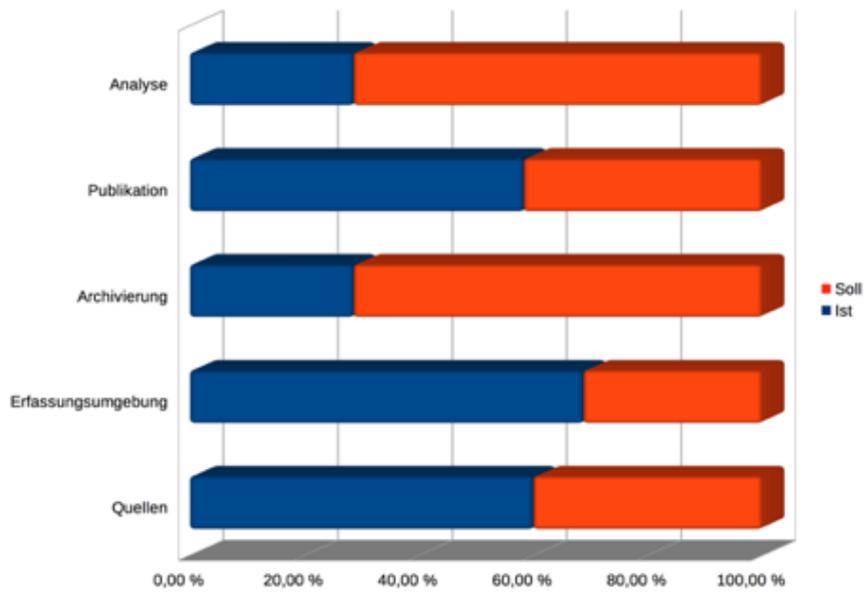
Workflow Digitale Edition: Bedarfserhebung



Workflow Digitale Edition: Angebot



Ecosystem "Digitale Edition": Ist- und Soll



Ecosystem "Digitale Edition": Ist- und Soll-Vergleich



Diskussion: (Technischer) Fokus, Werkzeuge

- automatisches Anreichern von Texten (NLP, NER): Treetagger, Spacy, NLTK, Weblight, UD-Pipe, ...
- Datenaustausch? TCF (Weblight) und Transformationen dazu, CIDOC-CRM
- Werkzeuge
 - Blickwinkel: anwenderorientiert vs. angebotsorientiert
 - Transkription: Oxygen, Transkribus,
 - Archivlösungen: Fedora, Versionierungstools (SVN, GIT)
 - Publikation: XML-DB (eXist), iiif, latex, Frameworks (bootstrap), Visualisierung, Zitation (Handle und DOI, URN, permalink),
 - Suchprozesse: Solr, elasticsearch ...

Diskussion über 'wichtige' Technologien

=> Vorschlag: Ontology in AG Texterschließung diskutieren (Vogeler)

Wallnig: Werkzeuge für audiovisuelles Material nötig

Stigler: Frage an ÖAW Apache Stanbol, wird nicht mehr weiterentwickelt

Durco: funktioniert, liefert RDF zurück



Diskussion cont'd

- Werkzeuge
 - Semantic Web: RDF, BlazeGraph, protégé, Apache Jena, Apache Stanbol
 - Analyse: Gephi, D3
 - Versionierungstools: **SVN**, **Github**
 - Organisation: Benutzerverwaltung, Projektdokumentation
 - Authorityfiles, LOD: GND, Wikidata, (VIAF), Geonames, (Handschriftencensus, Verfasserlexikon), historical Gazetteers (Pelagius), Germania Sacra, Cern Thesaurus

Fritze: Annotationstool für Nutzer (incl. User Management), spiegelt sich auch bis jetzt noch in keiner Arbeitsgruppe

Wallnig: Archivierung von Nutzerdaten

Tool für Projektdokumentation?

Gibt zwar Ticketsysteme für Projektmanagement

Redmine, Git → später nutzbar als Dokumentation

Wikis

Stigler: Dokumentation von Kodierungsrichtlinien wichtig

Vogeler: tag declaration im teilHeader eventuell dafür nutzbar

Fritze: Datenformate für Quelldokumente, Technologien wie man von nicht digitalen Daten zu digitalen Daten kommt.




Technischer Fokus ... auf Institutionen

... siehe Berichte der Institutionen

Detaillierte Ausarbeitung im Workshop im Herbst (Aviso: 11-12. Oktober)

Diskussion: Kooperationsvorhaben

- Bestehende Kooperationen: TU - ZIM, ZIM - KUG, AAU - OeNB, AAU - IBK, OEAW - VIE, OENB - VIE, IBK - ZIM
- Welche Methoden verwenden die Leute in KONDE?
- In welcher Domänen sind die einzelnen Mitglieder unterwegs?
- Methodologisches und inhaltliches Tagging
- Terminologie, Schlagwortsammlung, Definitionssammlung in AGs
 - Lexicon of Scholarly Editing: <http://uahost.uantwerpen.be/lse/>
 - Glossar DE (GDoc)
- Problemkatalog vs. Angebotskatalog

38

Welche Bedingungen sind notwendig, die Kooperation zu verstärken?

Wir haben jetzt die Bedürfnisse und Angebot festgestellt:

Wo gibt es Anknüpfungspunkte zw. Projekten und Standorten? Wie können diese Kooperationen gefördert werden?

Terminologie, Schlagwortsammlung, Definitionssammlung in AGs: Was ist eine Hybridedition?

Wo ist Kooperation möglich?

=> Kooperation mit anderen verwandten Projekten? => etwas fuer AG 9

Kooperation vorantreiben:

Kaiser: Technologieliste auf Institutionen mappen als online-Ressource beste Basis

Geiger: Methodologie: wer ist Experte für was?

Johannes: jede AG sollte grundlegende Begriffe als Konsens definieren und im Weißbuch festhalten

Wallnig: definitorische Arbeit, Glossar

Durco: Glossar als Teil des Weißbuches, Poster als Basis

Klug: Nutzen der Mailingliste für Austausch und Problemdiskussion, ev. auch Blog

Durco: Links [zu Projekten und Ressourcen] in Zotero, Nachteil: oft nicht öffentlich oder fertig

Diskussion Projektinhalt vs. Technik

Fanta: Projekt im Vordergrund, Technik eigentlich nebensächlich, Gefahr sich zu sehr von der formalen Ebene ablenken zu lassen, Kooperation zwischen konkreten Projekten intensivieren

Wallnig: Gemeinsamkeiten der Editionen formal, technisch usw sind in diesem Projekt im Fokus und nicht die inhaltlichen Fragestellungen der Projekte

Fanta: Diskussion über konkrete editorische Arbeit/Probleme und erst dann mit welchen Tools

Wang: Problemkatalog aufbauen



Diskussion: Nächste Schritte

Ideen für das Weißbuch

- Lobbying-Grundlage
- FWF-Richtlinien
- Publikationen
- Nutzung als Handreichung für neue Editoren
- Zotero mehr nutzen für Link-Dokumentation
- Achtung auf Crediting!

Individuelle Treffen der AGs und der einzelnen kooperierenden Projekte
Johannes: viele AGs treffen sich bei dha im November

Wallnig: Timeline Weißbuch? was passiert im nächsten Jahr?

Georg: das letzte Jahr ist reserviert für die Ausarbeitung des Weißbuchs / der einzelnen Beiträge

Helmut: vorletztes (also das laufende Jahr) dient der Vorbereitung: provisorisches IVZ

Johannes: nächstes Jahrestreffen zweite Woche nach Ostern in Klagenfurt

