

neonion - Kollaboratives Annotieren zur Erschließung von textuellen Quellen

Müller-Birn, Claudia

clmb@inf.fu-berlin.de
Freie Universität Berlin, Deutschland

Breitenfeld, Andre

andre.breitenfeld@fu-berlin.de
Freie Universität Berlin, Deutschland

Im Rahmen einer Posterpräsentation stellen wir die kollaborative Annotationssoftware neonion vor, dessen Entwicklung inspiriert wurde von der Vision des Memex. Vannevar Bush führt in seinem Artikel dazu aus, dass “[a] record if it is to be useful to science, must be continuously extended, it must be stored, and above all it must be consulted.” (Bush 1945: 37). Ein solcher “record” kann beispielsweise ein historisches Dokument sein. Am Anfang des Forschungsprozesses, ist noch wenig darüber bekannt, aber das Wissen um dieses Dokument wächst kontinuierlich durch die Forschungsarbeit der Wissenschaftler_innen. Mit Hilfe von *neonion* sollen Forschende Dokumente gemeinschaftlich mit Hilfe von Annotationen erschließen können (Müller-Birn et al. 2015). Die Software wird mit dem Ziel entwickelt, als Forschungsumgebung zu fungieren, um kollaborative Wissensgenierung im Rahmen von Annotationsprozessen in der Forschungsarbeit zu analysieren.

Die grundsätzlichen funktionalen Anforderungen an die Software *neonion* wurde basierend auf den Erkenntnissen einer Interviewstudie durchgeführt. Insgesamt wurden sechs Interviews durchgeführt. Diese Interviews waren semi-strukturiert und wurden am Arbeitsplatz durchgeführt, um auch Informationen über das Arbeitsumfeld zu erlangen und einen direkten Einblick in die genutzten Softwarewerkzeuge und Abläufe zu erhalten. Die Interviews dauerten eine bis anderhalb Stunden und wurden anschließend transkribiert. Alle Teilnehmer_innen waren Mitglieder der gleichen Forschungsreinrichtung (für weitere Informationen s. Müller-Birn et al. 2015). Das Ziel dieser Studie war es, den Kontext der Forschungsarbeit in den Geisteswissenschaften besser zu verstehen. Solche Interviews sind zentral bei der nutzerzentrierten Softwareentwicklung, einem Ansatz, der bei uns konsequent verfolgt wird. Die Ergebnisse der Interviews wurden nach vier Gesichtspunkten ausgewertet: Form, Funktion, Wert und Status (in Anlehnung an Marshall 1998). Wir nutzen diese Ergebnisse im Folgenden, um die grundsätzlichen Funktionen von *neonion* vorzustellen.

Der Bereich Form setzt sich mit der Struktur von Annotationen auseinander. Die Mehrzahl der Befragten gab an, vor allem basierend auf Textdokumenten zu forschen. Daher wurde entschieden, *neonion* zunächst für Textdokumente zu verwenden. Die Softwarearchitektur wurde mit Webframeworks umgesetzt. Hierzu findet Django im Bereich des Back-Ends und vorrangig Bootstrap und AngularJS im Front-End Anwendung. Die erzeugten Annotationen werden zur dauerhaften Aufbewahrung zum einen in den AnnotationStore, der auf Grundlage von Elasticsearch arbeitet, gespeichert, zum andern in den Sesame Triple Store eingespeist.

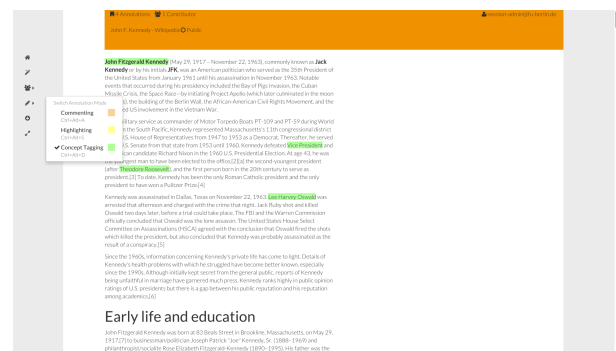


Abb. 1: Auswahl des Annotationsmodus im Annotator

Im Bereich Funktion wurde der Frage der Verwendung nachgegangen. So wurde ersichtlich, dass unterschiedliche Annotationsmodi (s. Abbildung 1) notwendig sind. In *neonion* werden daher drei Arten der Annotation unterschieden: die Markierung für Zitate, der Kommentar für Paraphrase und semantische Tags für das semantische Erschließen von Dokumenten (z. B. basierend auf einer zugrundeliegenden Ontologie). Auch wenn diese drei Arten von Annotationen aus den Interviews entstanden sind, können diese Annotationsmodi je nach Anwendungszweck sehr variabel eingesetzt werden. Ein Einsatz von *neonion* im Bereich der Linguistik wäre beispielsweise möglich, aber ein praktischer Anwendungsfall fehlt bisher.

Zur Implementierung der Annotationskomponente kommt die quelloffene JavaScript Bibliothek *Annotator.js* zum Einsatz. Die Bibliothek der OKFN wurde zusätzlich durch eigene Plug-Ins, insbesondere zur Realisierung einer semantischen Annotation, erweitert.

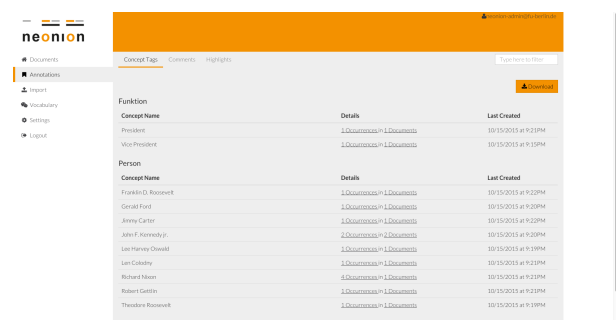


Abb. 2: Annotationsübersicht gefiltert nach semantischen Tags

Der längerfristige Wert der Annotation wurde im dritten Bereich untersucht. Hier wurde von allen Interviewteilnehmer_innen angegeben, dass eine Weiterverwendung der Annotationen in anderen Kontexten nicht möglich ist, da die verwendete Software den Export der Annotationen verhinderte. Diese Mangel sollte in neonion behoben werden. Alle Annotationen werden einerseits innerhalb eines standardisierten Datenmodells – dem Open Annotation Data Models (OADM) - gespeichert und andererseits haben Nutzende die Möglichkeit, alle ihre Annotationen nach unterschiedlichen Gesichtspunkten zu filtern und in ein Textdokument zur Weiterverarbeitung, z. B. in ein Textverarbeitungsprogramm zu exportieren (s. Abbildung 2). Mit Hinblick auf die Kollaboration besteht die Möglichkeit Annotationen innerhalb von Gruppen mit anderen Nutzer_innen zu teilen bzw. gemeinschaftlich Dokumente zu annotieren. Ebenfalls können die verwendeten strukturierten Vokabulare gemeinschaftlich erstellt werden. Es ist geplant, hier entsprechend benötigte Diskussionsfunktionen einzubauen.

Darüber hinaus wurde das bestehende Open Annotation-Datenmodell um die Möglichkeit erweitert, semantische Tags über eine typisierte Verbindung in Relationen zueinander zu setzen. Die semantischen Tags stellen in diesem Zusammenhang Instanzen von vordefinierten Konzepten mit eigener URI (Unified Resource Identifier) dar und ermöglichen durch die Beziehung der Instanzen zueinander eine mehrstufige Analyse von Annotationen.

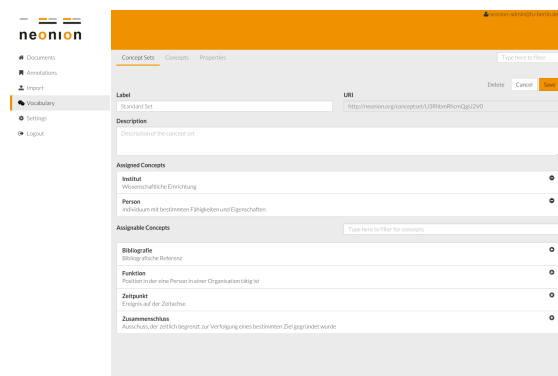


Abb. 3: Definition eines Begriffssystems als Concept Sets

Im vierten Bereich wurde der Frage nachgegangen, wie Annotationen inhaltlich geteilt werden. Unsere Interviewpartner führten aus, dass vor allem im Bereich der strukturierten Annotationen (semantische Tags basieren auf einem vordefinierten Begriffssystem) es sehr umständlich und zeitaufwändig ist, ein gemeinschaftliches Begriffssystem zu erstellen. In neonion können solche Begriffssysteme, die zu einer Ontologie weiterentwickelt werden können, einfach als sogenannte Concept Sets (s. Abbildung 3) hinterlegt werden. Diese Concept Sets

können dann auch wieder anderen Personen in neuen Forschungskontexten zur Verfügung gestellt werden. Aus technischer Sicht bietet das Backend von neonion verschiedene Dienste an, um Annotationen und Concept Sets mit unterschiedlichen Systemen über eine spezifizierte REST API oder SPARQL Endpoint auszutauschen.

Bibliographie

Bush, Vannevar (1945). „The atlantic monthly“. in: *As we may think* 176, 1: 101-108.

Marshall, Catherine C. (1998) „Toward an ecology of hypertext annotation“, in: *Proceedings of the ninth ACM conference on Hypertext and hypermedia: links, objects, time and space —structure in hypermedia systems* 40-49.

Müller-Birn, Claudia / Klüwer, Tina / Breitenfeld, Andre / Schlegel, Alexa / Benedix, Lukas (2015) „neonion: Combining Human and Machine Intelligence“, in: *Proceedings of the 18th ACM Conference Companion on Computer Supported Cooperative Work & Social Computing* 223-226.