

Das Erfurter FactGrid, eine Wikibase-Instanz außerhalb des Wikidata environments

Unter der Adresse <http://factgrid.de> implementierten wir unter einer Kooperationsvereinbarung vom August 2017 Wikimedia Deutschland, WMDE, und die Universität Erfurt Anfangs 2018 eine Wikibase-Instanz (https://database.factgrid.de/wiki/Main_Page).

Die Instanz wurde in einem ersten Testlauf mit Daten der Hallenser und Gothaer Illuminatenforschung befüllt. Das Projektziel ist es im Feld historischer Forschung offen eine kollaborativ zu nutzende Plattform zu schaffen. Software-Entwicklungen sollen hier wie Forschungsdaten vor Inselfösungen geschützt werden. Die Datenlizenzierung erfolgt zu diesem Zweck unter CC0.

Die Entscheidung für die Wikibase-Software war die Entscheidung für eine leistungsstarke Software, die sich in rapider Entwicklung befindet - hier liegt gleichzeitig eine immense Herausforderung, da die nun, vorerst von Berlin aus Datenbank-betreut als externe Instanz im Geflecht der Entwicklungen gehalten werden muss.

Die interessanten Seiten der Wikibase Software sind offensichtlich: Benutzer können auf der Plattform in der **flachen Hierarchie** selbst Korrekturen und Erweiterungen vornehmen. Sie agieren **öffentlich transparent** (jeder einzelne Edit ist **versioniert** und rücksetzbar). Wikibase ist im selben Moment einzigartig **flexibel** darauf vorbereitet, dass Projekte ihre eigenen Forschungsfragen mit spezifischen Statements riskieren. Die Software ist dabei auf das Internet zugeschnitten.

Problemstellen zeichnen sich gleichzeitig ab: Die Wikibase Software ist von sich aus **nicht bibliothekstauglich**; sie kann eher eine Plattform unter eigenen Regeln generieren, von der aus dann Daten in eine hierarchisch betreute Datenbank fließen könnten. Wikibase **fehlt ein Benutzer-Interface**. Was sich hier machen lässt, zeigt Magnus Manskes "Reasonator" (<https://tools.wmflabs.org/reasonator/?&q=1339>) auf: Aus der Wikidata-Daten lassen sich strukturierte Datenblätter in beliebiger Sprache generieren, das Tool ist jedoch nicht Teil der von Wikimedia betriebenen Software-Entwicklung. Die SPARQL Datenbankabfrage setzt ein immenses Wissen über die jeweilige Datenstruktur voraus - **Suchschablonen** fehlen, ebenso steht die Arbeit mit **Eingabeschablonen** noch ganz am Anfang.

Eine Kooperation zwischen der GND/DNB, Wikimedia Deutschland und dem Gotha-Erfurter Forschungsverbund

Die **Gothaer/Erfurter Projektpartner** gewannen mit der Datenbank eine Plattform, die sich für gezielt für Projektkooperationen nutzen lässt.

Die GND stand hier früh als die Interessanteste Lieferantin einer "Vorbefüllung" zur Diskussion. Die Daten sind frei und Alltagsgeschäft der Projektnutzer. Mit dem selektiven Datenimport entfielen für die Benutzer der jeweils erste Recherche-Schritt - sie könnten sich weit klarer auf die Augmentierung und die Verknüpfung von Daten ausrichten. Mit dem GND-Datensatz kämen im selben Moment die Standards in die Plattform, die ohne Normen setzenden Eingabeschablonen vorläufig fehlen.

Im **Wikimedia**-Projektgefüge eröffnet die externe Plattform bislang die Chance, die Wikibase-Software im externen Betrieb zu testen und hier den Aufbau einer Plattform zu begleiten auf der gezielt "Original Research" betrieben wird und von der sich Daten im Verlauf wie von einer externen Quelle beziehen lassen, die ihren eigenen Spielregeln unterliegt.

Auf Seiten der **GND/ DNB** wurde die Wikibase als Kandidat zur Erweiterung und Modernisierung der GND-Infrastruktur ausgemacht. Ziel des GND-Entwicklungsprogramms ist unter anderem, eine weitreichende Öffnung der GND für die Partizipation durch Kultureinrichtungen aller Art und auch für eine nahtlose Zusammenarbeit mit privaten Initiativen und offenen Community-Projekten wie wikipedia und wikidata. Für einen solchen Prozess wird auch eine Infrastruktur benötigt, welche für die Zusammenarbeit zwischen unterschiedlichen Communities unterstützt - also das Thema "Kollaborative Datenpflege" im Herzen trägt. Hierfür könnte Wikibase der richtige Baustein sein. Daher sind die GND-Partner an einer Evaluation der Software für die Integration in die GND-Umgebung interessiert:

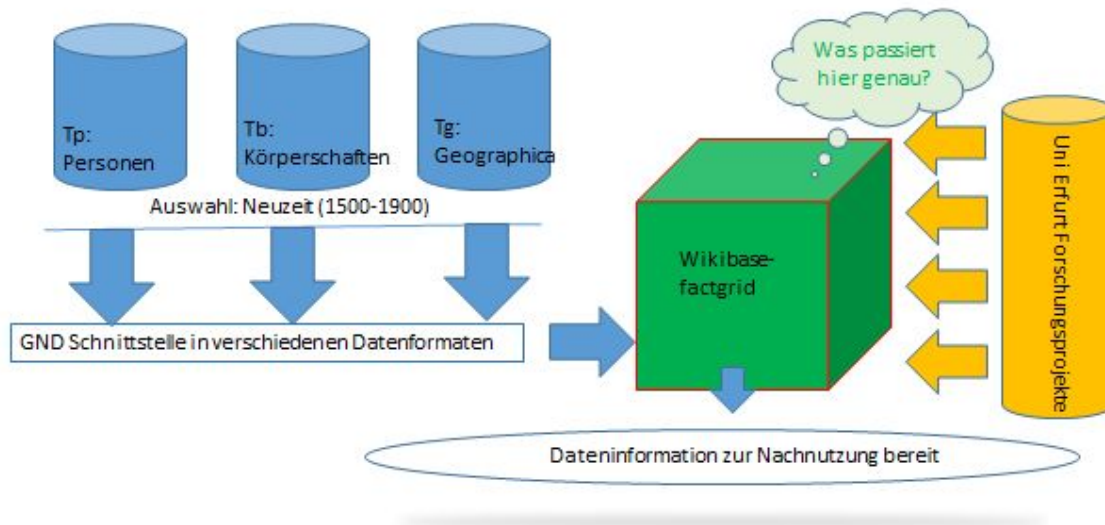
- Wir wollen lernen, welche Aufwände entstehen, um GND Datensätze in Wikibase einzuspielen
- Uns interessiert, wie die Oberfläche /Usability eines Wikibase Interface sich für GND Redakteure darstellt
- Wir wollen wissen, wie mittels Wikibase eine Erweiterung der Entity-Relation-Modells der GND um Zeit- und *Multilingualitäts*aspekte realisiert werden kann.
- Welche Anforderungen entstehen durch den Bedarf eines Daten-Reimports aus wikibase in die bisherige GND Datenbankstruktur?
- Last but not least sind wir an einer umfassenden Evaluation des Vorhabens interessiert, als Entscheidungshilfe für eine potentielle Wikibase-Installation für den GND-Betrieb.

Die vorliegenden Erwägungen gelten den Möglichkeiten einer Kooperation aller drei Beteiligten. Gotha gewänne mit der DNB den Projektpartner, der weitaus präziser die eigenen Datensätze kennt. Die DNB gewänne mit der Plattform eine von Berlin aus betreute Installation - in der gegenwärtigen Entwicklung ein wichtiges Moment. Für das Wikidata-Team eröffnete sich hier im selben Moment die Chance den größten komplementären Datensatz gegenüber dem eigenen auf der Wikibase-Software erkunden zu können.

Gegenüber der niederschweligen und pragmatischen Projektformulierung werden wir gleichzeitig die Chancen des offen inszenierten Projektes nutzen können - die Chancen einer Kooperation von der aus sich eigene weitere Schritte unternehmen lassen. Für das FactGrid spielten mir mit Gedanken einer weit größeren Datenbank des frühmodernen Buchangebots wie mit Kooperationen mit größeren Archiven wie dem Geheimen Staatsarchiv preußischer Kulturbesitz, das breite Teile der ersten in der Datenbank erfassten Dokumente verwaltet.

Überlegungen zu einem selektiven Datenbestand

GND im Erfurter Factgrid



Von Unmittelbarem Interesse sind vor allem GND-Daten zu

- Personen.
- Körperschaften
- georeferenzierten Objekten

Die Gothaer Projektbeteiligten überlegten im ersten Zug einen Zeitschnitt etwa für die Jahre 1400 bis 1900 anzustreben. Forschung in Gotha, wesentlich getragen auch von externen Stipendaten, die hier Gastaufenthalte absolvieren, gilt einerseits der Frühen Neuzeit und andererseits mit spezifischem Blick auf die Kartographie und die hier lagernden Bestände des Perthes-Verlags der Globalisierung im 19. Jahrhundert.

Das breitere Projekt einer Kollaboration mit der DNB und dem Ziel kollaborativ editierbarer GND-Daten wird die Beschränkung auf den Zeitschnitt weit weniger nahelegen. Hier würden gerade die aktuellen Personendaten spannend, die hier den erfassten Personen jetzt selbst editiert und verknüpft werden könnten.

Zur technischen Realisation des Dateninputs

Frage, ob es neben QuickStatements eine Schnittstelle gibt

Frage, wer in der GND ein Interesse an der Software-Erprobung hat, die hier geschehen würde.

Frage, wer von Wikimedia aus einen Dateninput dieses Ausmaßes betreuen würde

Zeitplan

Wir visierten einen Dateninput für das Jahr 2018 an, Hendrikje Carius und Wolfgang Ruschke hatte ihrerseits mit einer solchen Termin im eigenen Projekt gespielt und sie für von Gotha aus machbar erachtet.