



## Wünsche und Anforderungen an geografische Datensätze in der GND

Das Herder-Institut für historische Ostmitteleuropaforschung erfasst in seinen Arbeitsbereichen und Projekten Metadaten zu Sammlungsbeständen und sonstigen Objekten unterschiedlichster Art. Dabei kommen unter anderem Erfassungssysteme aus den Bereichen Bibliothek, Bibliografie, Bilddokumentation, Kunstgeschichte, Kartografie, Archiv- und Verlagswesen zum Einsatz. Aktuell bauen wir bestandsübergreifende Onlinerecherchemöglichkeiten für einen weltweiten Nutzerkreis auf und streben hierfür eine Angleichung der Erschließung von personellen, raumzeitlichen und thematischen Bezügen an. Insbesondere hinsichtlich der Erfassung historischer geografischer Namen und Gebiete stehen wir bereits im Austausch mit unterschiedlichen Partnerinstitutionen.

Die indirekte Verortung stützt sich dabei auf die geografischen Datensätze der Gemeinsamen Normdatei. Für diese schlagen wir Anpassungen in drei Prioritätsstufen vor:

- 1) Um die Nutzbarkeit der GND weiter zu verbessern und die geplanten Recherchemöglichkeiten umsetzen zu können, wünschen wir uns:

### ***Validation der Individualisierung***

Unter den Geografika in der GND sind zahlreiche Datensätze ohne identifizierende Namenszusätze beziehungsweise ohne Bezug zu übergeordneten Geografika zu finden. Beim Neuanlegen oder Ändern eines Datensatzes wäre daher eine erzwungene Individualisierung wünschenswert.

### ***Namenszusätze zu allen mehrdeutigen Geografika***

Die Bekanntheit einer geografischen Einheit bei verschiedenen Personen unterscheidet sich stark. Auch im Sinne einer Eindeutigkeit sämtlicher bevorzugter Namen sollten daher alle gleichnamigen Geografika identifizierende Zusätze erhalten, inklusive den jeweils vermeintlich bekannteren.

### ***Sprach- und Schriftangabe als Standard***

Sämtliche Namensformen, inklusive der bevorzugten, sollten eine Angabe zur jeweiligen Sprache und Schrift erhalten (§U, §L). Dann wäre es zudem möglich, in einem Katalog vom jeweiligen Nutzenden wählen zu lassen, in welcher Sprache und Schrift die bevorzugten Namen der Geografika möglichst ausgegeben werden sollen. Möglicherweise wäre es hierfür sogar sinnvoll, analog zu *Geonames*, für jede Sprache eine bevorzugte Namensalternative in der GND zu markieren. Vielleicht können die Kategorien 451 und 751 langfristig auch zusammengelegt werden?

### **Auswertbare Zeitangaben**

Für alle Zeitangaben ist eine Aufschlüsselung in drei getrennte Angaben sinnvoll: Anfangsdatum, Enddatum (beide numerisch, mindestens Jahre) und Verbale Datierung. Alle drei Angaben sollten vorhanden sein. Die verbale Angabe könnte optional aus den numerischen Angaben generiert werden. Beispiele:

A: 1920 E: 31.01.1931 V: 1920 bis 31.01.1931

A: 1301 E: 1500 V: 14. bis 15. Jahrhundert

### **Zusätzliche Koordinatenwerte**

Neben dem Mittelpunkt und Rahmenkoordinaten (vier Werte) sollte es möglich sein, die Koordinaten eines Streckenverlaufes oder ein Grenzpolygon zu erfassen (034 \$s, \$t). Die genauere Angabe des Bezugsgebietes eines Geografikums ermöglicht es, bei der koordinatenbasierten Suche nach Geografika oder direkt verorteten Sammlungsbeständen (z.B. Karten, Luftbildern) sinnvolle Polygonverschneidungen durchführen zu können.

- 2) Darüber hinaus würden folgende Änderungen die Arbeit mit der GND erleichtern beziehungsweise die Datenqualität erhöhen:

### **Aktive Webverweise**

Wenn im Erfassungssystem alle URIs als aktive Weblinks ausgegeben würden, kann man zum Beispiel die Datenquelle von Koordinaten (034 \$u, Geonames-Datensatz etc.) direkt aufrufen und sich damit die Lage des Geografikums einfach in einer Karte anzeigen lassen.

### **Markierung offizieller Namen**

Für verschiedene Anwendungen ist es nützlich, jene Namen einer Gebietskörperschaft kennzeichnen zu können, die die offizielle, landessprachliche Namensform in einem bestimmten Zeitraum darstellen. Dies wäre über den Beziehungscode *nauv* (Name in unveränderter Form) möglich. Dieser ist bisher aber nur im Feld 451 (\$4) verfügbar. Abgesehen von einer Freitextangabe im Bemerkungsunterfeld ist außerdem keine Zeitangabe möglich.

### **Determiniertheit von Standardnummern**

Analog zur Determiniertheit von DDC-Notationen (083 \$d) könnte auch zu anderen Standardnummern (024) der Grad der Übereinstimmung (vollständig, hoch, mittel, gering) angegeben werden.

### **Direkte Quellenangaben**

Es wäre wünschenswert, dass man einzelne Merkmale, wie abweichende Namen oder Zeiträume, direkt mit einer eigenen Quellenangabe versehen kann. Das würde zu eindeutigeren Quellenverweisen und einer leichteren Überprüfbarkeit führen.

3) Im Zuge eines weiteren Ausbaus der GND sind folgende Optionen denkbar:

#### ***Abgleich mit amtlichen Verzeichnissen***

In den meisten Ländern werden von zentraler Stelle umfassende Verzeichnisse mit geografischen Namen und Mittelpunkt- oder Grenzkoordinaten gepflegt. Das deutsche Bundesamt für Kartographie und Geodäsie (BKG) realisiert zudem in der Vereinigung *EuroGeographics* gemeinsame Projekte mit den anderen europäischen Katasterverwaltungen. Um zumindest für aktuelle Daten die aufwändige Einzelerhebung von Koordinaten zu vermeiden, wäre ein zentraler Transfer der amtlichen Daten in die GND sehr hilfreich. Im Rahmen eines regelmäßigen Abgleichs könnten zukünftige Veränderungen sogar systematisch nachgehalten werden. Insofern die Daten nicht ohnehin frei verfügbar sind, wären entsprechende Nutzungsvereinbarungen anzustreben. Für die Datenaufbereitung könnte bei Bedarf eine Förderung durch die Deutsche Forschungsgemeinschaft (DFG) beantragt werden.

#### ***Aussprachebeispiele***

In Webdatenbanken wie *Forvo* ([de.forvo.com](http://de.forvo.com)) werden Aussprachebeispiele für Wörter, inklusive geografische Namen, in zahlreichen Sprachen gesammelt. Man könnte also in Betracht ziehen es zu ermöglichen, zu den einzelnen Namensformen eines Geografikums in einem weiteren Unterfeld einen entsprechenden URI zu hinterlegen.

#### ***Grafische Benutzeroberfläche***

Für die einfachere Pflege der GND-Datensätze wäre eine grafische Benutzeroberfläche nützlich. In diese ließen sich auch leicht Auswahllisten für kontrolliertes Vokabular sowie halbautomatische Zuweisungen einbinden (z.B. Anlegen der Bezüge zu den Elementen des Namenszusatzes zum bevorzugten Namen).