

# Fusion von Nutzergenerierten Daten von Openstreetmap und Wikidata zur Erstellung von Karten

Mathias Gröbe

Keywords: Openstreetmap, Wikipedia, Wikidata, Datenaufbereitung, GIS, Geovisualisierung, Atlas, VGI

Im Forschungsfeld der Nutzergenerierten Daten und darüber hinaus sind Openstreetmap und Wikipedia allgemein bekannte und vielseitig genutzte Plattformen. Diese verwalten neben den gesammelten Daten ebenso Werkzeuge sowie Schnittstellen und entwickeln diese in der Regel nach ihren eigenen Bedürfnissen weiter. Die beiden Projekte haben zwar unterschiedliche Schwerpunkte aber genauso gegenseitige Anknüpfungspunkte. Im Sinne einer vernetzten Welt werden Objekte über Plattformen und Sprachen hinweg immer mehr miteinander verknüpft. Diese Verbindungen bieten den großen Mehrwert, dass Informationen einfacher zugänglich und damit besser nutzbar sind. Die unterschiedliche Ausrichtung der Projekte gestattet es unterschiedlichste semantische, geographische und geometrischen Informationen miteinander zu verknüpfen.

Auf diese Weise ist es möglich, über das Wikipedia angegliederte Wikidata Projekt Informationen zu geographischen Objekte in strukturiert Form mittels der Abfragesprache für Graphen „SPARQL“ abzufragen. Für einen Fluss sind beispielsweise Informationen zu durchflossenen Ländern, Nebenflüssen, Einzugsgebiet und Objektidentifikatoren gespeichert. Mit Hilfe des Openstreetmap Identifikators lässt sich über die Overpass-API die aktuelle Geometrie mit Attributen aus der Datenbank von Openstreetmap extrahieren. Diese Vorgehensweise ermöglicht beispielshalber die Fragestellung nach den längsten Flüssen oder nach Gebirgen mit ihren höchsten Bergen in Deutschland zu beantworten und kartographisch darzustellen.

Am Beispiel der Ableitung einer Übersichtskarte für Deutschland wurden die Daten von Openstreetmap und Wikidata kombiniert, aufbereitet und visualisiert. Für die Bewertung des Ergebnisses wurde der freie Datensatz von NaturalEarth verwendet. Der Vergleich verdeutlicht Unterschiede hinsichtlich Homogenität, Abdeckung und Modellierung. Vorteilhaft scheint die Nutzung der beschriebenen Vorgehensweise in punkto Aktualität, Korrektheit und Mehrsprachigkeit der Daten. Dabei treten aber auch Besonderheiten von Nutzergenerierten Daten hervor und sorgen teilweise für unerwartete Ergebnisse.