

Semantic Web

Hinkelmanns, Peter; peter.hinkelmanns@sbg.ac.at

Mit dem *Semantic Web* wird das reguläre World Wide Web um semantische Informationen ergänzt, die einen globalen Informationsgraphen ergeben. Geprägt wurde die Idee des *Semantic Webs* durch Tim Berners-Lee (Berners-Lee/Fischetti 2000), eine ausführliche Einführung bieten etwa Dean Allemang und Jim Hendler (Allemang/Hendler 2011).

Der Informationsgraph besteht aus Aussagen zu den unterschiedlichsten Entitäten. Normdaten, etwa zu Personen, Orten oder Werken, stellen Projekte wie die Gemeinsame Normdatei (GND) (Gemeinsame Normdatei 2019) oder Wikidata (Wikidata 2019) bereit.

Eine Information besteht dabei immer aus einem Tripel, also einer Subjekt-Prädikat-Objekt-Beziehung. Beispielsweise:

- [Stefan Zweig] [ist] [Person]
- [Stefan Zweig] [ist Autor von] [Schachnovelle]
- [Schachnovelle] [ist] [Literarisches Werk]

Aus diesen Triples kann ein gerichteter Graph aufgebaut werden, sodass die Subjekte und Objekte die Knoten und die Prädikate die Kanten bilden. Die Ausrichtung der Kanten erfolgt dabei immer vom Subjekt auf das Objekt:

Abbildung 1: Graph aus drei Triples

Übergreifende Abfragen im *Semantic Web* funktionieren dann, wenn im gesamten Netz einheitliche Identifikatoren, genannt URI – *Uniform Resource Identifier*, genutzt werden (siehe auch *Persistent Identifier*). Um diese notwendige Einheit zu erzielen, ermöglichen *Ontologien* die Erstellung von Regeln für *Triples*. Im obigen Beispiel könnten etwa die Prädikate ‘ist’ und ‘ist Autor von’ und die beiden Objekte ‘Person’ und ‘Literarisches Werk’ in einer solchen Ontologie beschrieben werden. Eine Regel könnte sein, dass ‘ist Autor von’ nur von einer ‘Person’ zu einem ‘literarischen Werk’ zeigen darf. Mit diesen wenigen, definierten Eigenschaften können nun weitere Datensätze zu Autorinnen und Autoren und deren Werken erstellt werden. Nun könnte etwa abgefragt werden, welche Autorinnen und Autoren existieren und wie viele Werke sie jeweils verfasst haben.

Ein anderes Projekt könnte diese Daten ergänzen, indem es etwa die Lebensdaten Stefan Zweigs aufnimmt und sich dabei auf denselben Knoten ‘Stefan Zweig’ bezieht. Die Gesamtheit dieser Informationen wird als *Linked Data (LD)*

oder *Linked open Data* (LoD) bezeichnet. Wichtigstes Grundprinzip ist, dass im *Semantic Web* alle Teilnehmerinnen und Teilnehmer alles über alle Themen sagen können (AAA-Modell).

Literatur:

- Allemang, Dean; Hendler, Jim: What is the Semantic Web? In: Semantic Web for the Working Ontologist: 2011, S. 1-12.
- Berners-Lee, Tim; Fischetti, Mark: Weaving the Web: the original design and ultimate destiny of the World Wide Web by its inventor Weaving the Web. San Francisco: 1999.
- Gemeinsame Normdatei. URL: <https://www.dnb.de/gnd>

Verweise:

Semantic Web-Technologien, Linked Open Data, Normdaten, Persistent Identifier, Kontrollierte Vokabularien: GND, Kontrollierte Vokabularien: Wikidata, CIDOC, RDF

Themen:

Einführung, Annotation und Modellierung

Projekte:

Wikidata

Lexika

- Edlex: Editionslexikon

Zitervorschlag:

Hinkelmanns, Peter. 2021. Semantic Web. In: KONDE Weißbuch. Hrsg. v. Helmut W. Klug unter Mitarbeit von Selina Galka und Elisabeth Steiner im HRSM Projekt "Kompetenznetzwerk Digitale Edition". URL: <https://gams.uni-graz.at/o:konde.167>