Inhaltsverzeichnis

[1. Wissensmodellierung im Semantic Web 2](#_Toc204069603)

[1.1 Ontologie 6](#_Toc204069604)

[2. W3C Standards 6](#_Toc204069605)

[3. WikiData 7](#_Toc204069606)

[4. Methode 7](#_Toc204069607)

[Normdaten 9](#_Toc204069608)

[XML 10](#_Toc204069609)

[RDF 11](#_Toc204069610)

[Wikidata 12](#_Toc204069611)

## Formalisierung von Wissen: Semantic Web und Ontologie

In diesem Kapitel werden zentrale theoretische Grundlagen zur Repräsentation und Organisation von Wissen in digitalen Systemen vorgestellt. Anschließend werden die grundlegenden Konzepte des Semantik Web erläutert und zentralen Begriffe aufgegriffen und erläutert. Ein besonderer Fokus liegt dabei auf der graphbasierten Darstellung und Strukturierung von Wissen.

### 2.1 Wissensmodellierung im Semantic Web

Das menschliche Gedächtnis setzt sich aus zwei zentralen Komponenten zusammen: dem semantischen Teil für allgemeines, kulturell geteiltes Wissen und dem episodischen Teil für persönlich Erlebtes. Diese Kombination ermöglicht es uns, Bedeutung zu erkennen und Erinnerung sinnvoll einzuordnen. Wörter und Phrasen können verschiedene Bedeutungen haben und ohne Kontext bleibt die Bedeutung vieler Ausdrücke unklar. Selbst strukturierte Informationen wie Zeitangaben entfalten ihre Bedeutung nur, wenn man die dahinterliegende Syntax kennt und weiß, wie diese Information zu lesen ist. Syntax gibt die Regeln vor, während Semantik ihre inhaltliche Bedeutung bezeichnet. Pragmatik geht über die Bedeutungszuweisung hinaus und fragt nach dem Zweck und der Absicht sprachlicher Äußerungen. Sie ist entscheidend für eine vollständige Interpretation von Sprache im Handlungszusammenhang. Während Menschen über kognitive Fähigkeiten verfügen, um sprachliche Bedeutung intuitiv zu erfassen, muss diese Fähigkeit bei Maschinen künstlich nachgebildet oder kompensiert werden (vgl. Dengel 2012: 10–11, 13).

Damit Computer Informationen erkennen und verarbeiten können, muss Wissen in strukturierter Form repräsentiert werden. Es reicht nicht aus, Wissen als eine Ansammlung isolierter Fakten zu betrachten, vielmehr als ein dynamisches Netz miteinander verbundener Informationen (vgl. Werner 2020: 247). Die formale Repräsentation dieses Wissen erfolgt in sog. Wissensorganisationssystemen (WOS). Diese Systeme dienen dazu, Informationen systematisch zu ordnen, zu speichern und auffindbar zu machen. Dabei handelt es sich um Systeme, die Metadaten in strukturierter Form darstellen. In digitalen Anwendungen spielt die Strukturierung von Informationen durch WOS eine zentrale Rolle. Die Qualität von Suchergebnissen hängt davon ab, wie gut Inhalte organisiert sind. Dabei ist zu berücksichtigen, dass jedes WOS eine spezifische Sichtweise auf ein Fachgebiet abbildet: Begriffe können je nach Perspektive unterschiedlich modelliert werden. Damit Benutzer Informationen effizient wiederfinden können, muss die Struktur des Systems deren Denkweise und Begriffswelt widerspiegeln (vgl. Geyer-Hayden 2009: 128–129).

In unterschiedlichen wissenschaftlichen Disziplinen wird zunehmend anerkannt, dass Objekte viele Bedeutungen und Verbindungen aufweisen (vgl. Werner 2020: 247). Dies macht deutlich, dass Daten ohne ihren Kontext kaum aussagekräftig sind (vgl. Dengel 2012: 5). Durch technologische Entwicklungen ist es möglich, Metadaten systematisch mit externem Wissen zu vernetzen (vgl. Werner 2020: 247). Dabei entsteht ein semantischer Raum, in dem Daten nicht isoliert, sondern relational verstanden und genutzt werden (vgl. Dengel 2012: 5). Diese Anforderung bildet die Grundlage für Konzepte wie Semantic Web.

Das Konzept Semantic Web geht auf Tim Berners-Lee zurück und zielt darauf ab, Informationen im World Wide Web automatisch und maschinell interpretierbar zu machen (vgl. Dörpinghaus 2022). Dadurch können Beziehungen zwischen beliebigen digitalen Objekten herzustellen – von Fotos über Finanztransaktionen bis hin zu abstrakten Konzepten ([vgl. Feigenbaum ebd. 2007](http://thefigtrees.net/lee/sw/sciam/semantic-web-in-action#page2)). Das Semantic Web erweitert das bestehende Web, indem es dessen globale Systeme zur Indexierung und Benennung nutzt. Es ist sowohl mit herkömmlichen Browsern als auch mit semantisch-fähiger Software zugänglich (vgl. Domingue/Fensel/Hendler 2011: 5). Dabei soll jeder Begriff eine eindeutige Kennung erhalten, um Identitäten klar aufzulösen (vgl. Domingue/Fensel/ Hendler 2011: 5). So müssen bspw. Unterschiedliche Bedeutungen eines Begriffs jeweils durch eigene URIs eindeutig gekennzeichnet werden ([vgl. Berners-Lee/ Hendler/ Lassila 2001](https://web.archive.org/web/20070713230811/http:/www.sciam.com/print_version.cfm?articleID=00048144-10D2-1C70-84A9809EC588EF21)). Ein zentrales Element ist dabei die Repräsentation von Zusatzinformationen (Metadaten) in einer Form, die von Maschinen interpretiert werden kann (vgl. Grütter 2008: 75). Semantische Metainformationen dienen dabei als Verknüpfung der Daten, die idealerweise öffentlich zugänglich sind. Dies bildet die Grundlage für ein maschinenlesbares Wissensnetzwerk (vgl. Dörpinghaus 2022). Solche semantischen Netze organisieren Wissen durch strukturierte, graphische Darstellung, meist in Form von Knoten und Kanten: Graphen (vgl. Dengel 2012: 34).

Graphen dienen als mathematische Modelle zur formalen Beschreibung netzartiger Strukturen. Solche Systeme zeichnen sich typischerweise durch eine Zusammensetzung von zwei strukturell unterschiedlichen Objektklassen. In der Terminologie der Graphentheorie werden diese beiden Elementen als Knoten (Objekte) und Kanten (Verbindungen bzw. Beziehungen) bezeichnet (vgl. Tittmann 2019: 11), wobei die Knoten Entitäten eines bestimmten Bereichs wie bspw. eine Person oder ein Produkt repräsentieren. Die Kanten verknüpfen diese Entitäten miteinander und geben die Art der Verbindung zwischen ihnen an (vgl. Barrasa/ Webber 2023: [Introducing Knowlege Graphs](https://learning.oreilly.com/library/view/building-knowledge-graphs/9781098127091/ch01.html#idm45681067535008)).

In der Graphentheorie (Euler 1736) werden zunächst die rein topologischen Eigenschaften der Netzstruktur untersucht. Ein Graph G = (V, E) wird formal durch die Knotenmenge (V) und Kantenmenge (E) definiert. Zwei Knoten u, v ∈ V gelten als benachbart oder als Nachbarn, wenn es eine Kante {u, v} ∈ E gibt, die sie miteinander verbindet. Man unterscheidet grundsätzlich zwischen *gerichteten* (Kanten haben eine Richtung haben (u, v) ≠ (v, u)) und *ungerichteten* (Kanten haben keine Richtung (u, v) ∈ E ⇔ (v, u) ∈ E) Graphen (vgl. Tittmann 2019: 11; vgl. Dörpinghaus 2022).

Knoten können für die Personen, Ereignisse oder Eigenschaften stehen, während Kanten deren assoziative Beziehungen abbilden. Dabei wird Wissen nicht isoliert, sondern über Verknüpfungen zugänglich gemacht (vgl. Dengel 2012: 34). Der Satz *Lena fährt Fahrrad* lässt sich durch ein einfaches Semantisches Netz darstellen (Abbildung 2).

fährt

Abb. 2 Semantisches Netz für den Satz *Lena fährt Fahrrad*

Wissen wird aber nicht nur in einfache Form eines Tripels (Subjekt – Prädikat – Objekt) dargestellt, kann auch komplexe Konstruktionen, wie Sätze mit mehreren Beteiligten oder Rollen abgebildet werden (vgl. Dengel 2012: 34).

kommt\_aus

fährt\_zu

fährt\_mit

Abb. 2 Semantisches Netz für den Satz *Lena fährt mit dem Fahrrad zu einem Freund, der gerade aus Berlin zurück ist.*

Die Verbindungen haben immer eine definierte Richtung, die von einem Knoten zu einem anderen verläuft und können ebenfalls keine oder mehrere Eigenschaften besitzen. Dabei hat jede Beziehung immer einen Start- und einen Endknoten (vgl. Barrasa/ Webber 2023: [Introducing Knowlege Graphs](https://learning.oreilly.com/library/view/building-knowledge-graphs/9781098127091/ch01.html#idm45681067535008)). Die Beziehungen sind oft bidirektional, werden aber häufig in einer bestimmten Richtung benötigt (vgl. Dengel 2012: 34).

Durch ihre flexible Struktur können Graphen verschiedenartige Datenquellen verknüpfen und unterschiedliche Beziehungstypen und Anwendungsbereiche abzudecken (Heist et al. 2021: 128). Unterschiedliche Visualisierungen desselben Graphen können dabei verschiedene strukturelle Merkmale sichtbar machen oder in den Vordergrund rücken. Bei der Darstellung gerichteter Graphen wird die Richtung der Relationen zwischen den Knoten üblicherweise durch Pfeile gekennzeichnet (vgl. Dörpinghaus 2022). Diese Struktur dient dazu, Objekte und deren Eigenschaften mit verschiedenen Formen von Relationen zu verbinden, um sie besser zu verstehen (vgl. Dengel 2012: 35). In diesem Kontext stellt sich auch die Frage, wie Wissen systematisch organisiert und zugänglich gemacht werden kann.

WOS leisten hierzu einen wichtigen Beitrag: Sie werden entwickelt, um die Inhalte für Benutzer besser auffindbar zu machen, aber auch um das Wissen eines bestimmten Themengebiets zu analysieren und abzugrenzen. Im Verlauf der Modellierung entsteht dabei ein gemeinsames Verständnis über die Struktur des jeweiligen Wissensbereichs (vgl. Geyer-Hayden 2009: 134–135). Frühere WOS waren häufig zentral organisiert und setzten einheitliche Begriffsdefinitionen voraus. Dies schränkte ihre Skalierbarkeit und Flexibilität stark ein. Komplexe oder offene Fragestellungen ließen sich in solchen Systemen oft nur begrenzt abbilden. Das Semantic Web verfolgte hingegen einen offenen Ansatz, der auch Ambiguitäten zulässt, um eine höhere Ausdrucksstärke und Anpassungsfähigkeit zu ermöglichen. Ziel war die Entwicklung einer Sprache, die sowohl Datenstrukturen als auch logische Schlussregeln beschrieb und die Integration bestehender Wissenssysteme ins Web erlaubte ([Berners-Lee/ Hendler/ Lassila 2001](https://web.archive.org/web/20070713230811/http:/www.sciam.com/print_version.cfm?articleID=00048144-10D2-1C70-84A9809EC588EF21)).

Das World Wide Web Consortium (W3C) entwickelte die Grundprinzipien des Semantic Web und stellte standardisierte Sprachen bereit (vgl. Davies/ Studer/ Warren 2006: 2). Mit den Technologien des Semantic Web können Nutzer:innen Datenspeicher im Web aufbauen, Vokabulare entwickeln und Regeln zur Datenverarbeitung festlegen (vgl. Kamzelak 2018: 424). Auf diese Grundlage entstehen innovative Dienste, die den Austausch und die gemeinsame Nutzung von Wissen vereinfachen (vgl. Dengel 2012: 9). Technologien[[1]](#footnote-1) wie Resource Description Framework (RDF), Web Ontology Language (OWL) oder Protocol and RDF Query Language (SPARQL) ermöglichen es, komplexe Sachverhalte durch semantisch angereichte Beziehungen zwischen Daten dazustellen Aus diesen Daten bildet sich ein Informationsnetz, der durch seine offene Struktur eine vielseitige Verknüpfung erlaubt. Jede Information kann über eine Uniform Resource Identifier (URI) eindeutig angesprochen werden, wodurch sie mit anderen externen, offenen und öffentlich zugänglichen Datenquellen im Web verbunden werden kann (vgl. Andraschke/Wagner 2020: 118–119).

Die semantisch angereichten Daten werden in sog. RDF-Tripeln modelliert. In der Praxis werden solche strukturierte Daten als Linked Open Data (LOD) bezeichnet (vgl. Andraschke/Wagner 2020: 119).

### 2.2 Linked Open Data

### Ontologie

Zusätzliche Stärke erhält das Semantic Web durch Ontologien, die es ermöglichen Beziehungen zu klassifizieren, abzufragen und darüber zu schlussfolgern ([vgl. Feigenbaum ebd. 2007](http://thefigtrees.net/lee/sw/sciam/semantic-web-in-action#page2)). Datenbanken verwenden oft unterschiedliche Begriffe für dasselbe Konzept. Ontologien helfen, diese semantischen Unterschiede durch bspw. Konvertierungsdienste zu überbrücken ([Berners-Lee/ Hendler/ Lassila 2001](https://web.archive.org/web/20070713230811/http:/www.sciam.com/print_version.cfm?articleID=00048144-10D2-1C70-84A9809EC588EF21)). Sie ermöglichen es, Information formal zu strukturieren und dadurch maschinelles Schlussfolgern zu erlauben. (vgl. Davies/ Studer/ Warren 2006: 4–5).

Ontologien bestehenaus vier zentralen Komponenten: Konzepten (Klassen), Relationen (Eigenschaften), Instanzen und Axiomen sog. 4-Tupel <C, R, I, A> (vgl. Davies/ Studer/ Warren 2006: 4–5).

## W3C Standards

Ein Bild, das Text, Screenshot, Design enthält.

KI-generierte Inhalte können fehlerhaft sein.

Abb. The Semantic Web Layer Cake – 1. (Domingue/Fensel/Hendler 2011: 18)

### 3.1 XML/ TEI

Eine Grundvoraussetzung für die Erstellung des Semantic Web sind strukturierte und maschinenlesbare Daten. Um diesen Anforderungen gerecht zu werden, wurde XML als Teil der Standardized Generalized Markup Language (SGML) entwickelt (vgl. Jeckle 2002: 11). XML ermöglicht es, eine individuelle Auszeichnungssprache zu erstellen und Elemente sowie Attribute nach Bedarf zu definieren (vgl. Ray 2001: 157). Diese Flexibilität erlaubt es, Daten in einem einheitlichen, maschinenlesbaren Format zu erstellen (vgl. Jeckle 2002: 11).

Obwohl XML als universelle Meta-Sprache gilt, bietet es keine Unterstützung zur Beschreibung der Semantik der Daten. Die Bedeutung der Beziehungen zwischen geschachtelten Elementen bleibt meist implizit und hängt von der jeweiligen Anwendung ab. Dieses Defizit wird durch Einsatz von RDF kompensiert (vgl. Baumeister/ Seipel 2005: 40).

XML erlaubrt die Strukturierung von Informationen, gibt aber keine Bedeutung vor ([vgl. Berners-Lee/ Hendler/ Lassila 2001](https://web.archive.org/web/20070713230811/http:/www.sciam.com/print_version.cfm?articleID=00048144-10D2-1C70-84A9809EC588EF21)).

### 3.2 RDF

RDF ist eine domänenunabhängige Sprache zur semantischen Beschreibung von Datenmodellen (vgl. Baumeister/ Seipel 2005: 40). Damit werden sowohl Datenobjekte als auch deren Relationen auf eine Weise, die Maschinen verstehen und verarbeiten können ([vgl. Feigenbaum ebd. 2007](http://thefigtrees.net/lee/sw/sciam/semantic-web-in-action#page2)).

Damit lassen sich Informationen über *Dinge* (Aussagen) darstellen, die im Web eindeutig identifiziert werden können. Jede Aussage wird in Form eines Tripels dargestellt: Subjekt, Prädikat und Objekt. Diese Tripel bilden zusammen einen Graphen aus Knoten und Kanten (vgl. Dengel 2012: 110).

RDF nutzt Triple-Strukturen zur Repräsentation von Bedeutung (Subjekt-Prädikat-Objekt). Jede Aussage wird in Form eines Tripels modelliert, wobei Subjekt und Objekt durch URIs eindeutig identifiziert werden ([vgl. Berners-Lee/ Hendler/ Lassila 2001](https://web.archive.org/web/20070713230811/http:/www.sciam.com/print_version.cfm?articleID=00048144-10D2-1C70-84A9809EC588EF21)).

Mit RDF lassen sich Datenfelder aus unterschiedlichen Datenbanken eindeutig identifizieren und miteinander verknüpfen ([vgl. Berners-Lee/ Hendler/ Lassila 2001](https://web.archive.org/web/20070713230811/http:/www.sciam.com/print_version.cfm?articleID=00048144-10D2-1C70-84A9809EC588EF21)).

RDF besteht aus zwei Komponenten: dem RDF-Modell (einer Menge von Aussagen) und der RDF-Syntax zur Serialisierung und damit zum Austausch (vgl. ebd.: 118). Das anwendungsspezifische Vokabular eines RDF-Datenmodells wird im RDF-Schema (RDFs) und in einer Ontologie (OWL) festgelegt. Im Unterschied zu XML-Schema, das die Struktur von XML-Dokumenten durch Einschränkungen definiert, beschreibt das RDFs die Eigenschaften von Objekten sowie deren Beziehungen zueinander und zu ihren Wertebereichen (vgl. Baumeister/ Seipel 2005: 40–41).

### 3.3 OWL

## WikiData

WikiData ist eine freie und frei zugängliche Datenbank, die strukturierte Daten sammelt. Die Sammlung von Wikidata besteht aus Datenobjekten, die durch ein *Q* und eine Nummer, dem sog. eindeutigen Bezeichner, identifiziert werden (vgl. [Wikidata:Introduction](https://www.wikidata.org/wiki/Wikidata:Introduction/de)). Jedes Datenobjekt enthält eine Aussage (Statement), die die Eigenschaften (property, P) und die dazu gehörigen Werte (values, V) eines Objekts beschreibt. Für Personen können bspw. Information wie Ausbildung hinzugefügt werden (vgl. [Help:Statements](https://www.wikidata.org/wiki/Help:Statements/de)).

Die Datenobjekte im Graph-Format (Subjekt-Prädikat-Objekt) enthalten Informationen, die über sämtliche Wikimedia-Portale erfasst wurden. Dies erfolgt durch die Verknüpfung einer Eigenschaft mit mindestens einem Wert. Die Aussagen können weiterentwickelt, mit Anmerkungen ergänzt oder durch zusätzliche Werte und optionale Qualifikationen erweitert sowie in einen spezifischen Kontext eingeordnet werden. Zudem dienen sie dazu, Objekte miteinander zu verknüpfen. Dies führt zu einer vernetzten Datenstruktur (vgl. [Help:Statements](https://www.wikidata.org/wiki/Help:Statements/de)).

## Methode

## Visualisierung

Netzwerkanalyse biete neue Möglichkeit zur Visualisierung und Analyse von Sammlungswissen. Ihre zentralen Ziele sind die Identifikation wichtiger Verbindungen und Strukturen. (vgl. Werner 2020: 248–249).

Für geistes- und sozialwissenschaftliche Forschung ist die Verbindung zwischen Netzwerken und Visualisierung von großer Bedeutung. Die graphische Visualisierung von abstrakten oder komplexen Strukturen macht die Zusammenhänge, die sich auf rein sprachlichem Wege nur schwer oder unzureichend vermitteln lassen, greifbar. Digitale Werkzeuge bieten hier neue Möglichkeiten, große Datenbestände nicht nur zu verarbeiten, sondern auch visuell aussagekräftige Formen überführen können. Daraus ergibt sich einen Anknüpfungspunk für die Verwendung netzwerkbasierter Modelle. Für die biographische Forschung eröffnet der Netzwerkansatz die Perspektive, individuelle Lebensläufe nicht isoliert zu betrachten, sondern als Teil größer Beziehungsgeflechte (vgl. Dörpinghaus 2022).

# Normdaten

Da die Maschinen keine indirekten Bedeutungen erkennen können, daher müssen Relationen zwischen Informationen eindeutig sein. Normdaten wie [GND-Nummer](https://www.dnb.de/DE/Professionell/Standardisierung/GND/gnd_node.html) (Gemeinsame Normdaten) ermöglichen diese Eindeutigkeit. Wird eine Person über GND-Nummer identifiziert, können Inhalte automatisch mit bspw. [Wikipedia](https://www.wikipedia.de/)-Eintrag oder [Deutschen Biographie](https://www.deutsche-biographie.de/) verknüpft werden. Trotzt der Bemühungen der Deutschen Nationalbibliothek kann die Eindeutigkeit von Normdaten nicht in allen Fällen gewährgeleistet werden. Regelmäßig auftretende Probleme, wie etwa fehlerhafte oder doppelt vorhandene GND-Einträge, lassen sich häufig auf die dezentrale Struktur der Datenerfassung zurückführen, die im Rahmen kooperativer Arbeitsprozesse. In der Praxis wird aus Effizienzgründen oder unter Zeitdruck nicht selten darauf verzichtet, bereits bestehende Datensätze auf ihre Richtigkeit und Vollständigkeit hin zu prüfen. Stattdessen wird ein neuer Eintrag erzeugt, was wiederum zu Redundanzen sowie zu einer Verringerung der Datenkonsistenz und -qualität im Gesamtsystem führen kann. Die Erstellung und Pflege von Normdaten stellt auch die Herausforderung im Bereich von Körperschaften, geographische Orte, Werte, Periodika usw. Ihre eindeutige Identifikation erweist sich oftmals als deutlich komplexer. Insbesondere bei Orten treten spezifische Schwierigkeiten auf: Während georeferenzierende Systeme (GIS) in der Regel lediglich punktuelle Koordinaten erfassen, bleiben flächenhafte Ausdehnungen sowie historische Veränderungen unberücksichtigt. Die historische Dimension eines Ortes – etwa sich verändernde politische Zugehörigkeiten, Ortsnamen oder Grenzverläufe – kann daher nur unzureichend über bestehende textbasierte Normdaten erfasst werden und bedarf ergänzender, oft nicht-textueller Quellen (vgl. Kamzelak 2018: 426-427).

Was soll geschrieben werden:

* Was ist NER, wie sie erkannt werden
* Erstellung von Vokabular (einheitlich, orientier an Wikidata) – Normdaten als Hilfe?
* FAIR Prinzipien
* Visualisierung? Sollen unterschiedliche Darstellungen gezeigt werden?
* Sprachmodelle?

Wissensgraphen gewinnen zunehmend an Bedeutung, da sie geeignet sind, komplexe Wissensstrukturen abzubilden (Heist et al. 2021: 128).

Hierbei erweist sich die graphbasierte Netzwerkanalyse als ein zentrales Instrument zur systematischen Verknüpfung von Datenobjekten und deren Kontextualisierung (vgl. Werner 2020: 249–250).

Diese methodischen Entwicklungen knüpfen an grundlegende Konzepte der Graphentheorie.

Hier schließt die Graphentheorie als methodisches Fundament an: Sie ist ein junges Teilgebiet der Mathematik, das ursprünglich durch praktische Problemstellungen inspiriert wurde. Ihre Relevanz für die Informatik ergibt sich aus der Schnittstelle zwischen mathematischer Struktur und algorithmischer Verarbeitung. Während sich die Mathematik vor allem mit der Struktur und den Eigenschaften von Graphen befasst, konzentriert sich die Informatik auf deren Berechnung und Verarbeitung. Die Ideen der Graphentheorie gehen auf Leonhard Euler (1736) zurück, wobei der Begriff *Graph* erst 1878 eingefügt wurde. Mit der digitalen Entwicklung erhielt die Graphentheorie neue Impulse, insbesondere durch: (a) Reflexion wissenschaftlicher Methoden im Netzwerkumgang; (b) Übernahme von Linked Data-Ansätzen aus der Datenwissenschaft und (c) kritische Methodendiskussion in der Netzwerkanalyse (vgl. Dörpinghaus 2022). Die Graphentheorie eröffnet vielfältige Perspektive auf gesellschaftlichen Strukturen und Prozesse. Darüber hinaus lässt sich *Gesellschaft* selbst als ein komplexes, dynamisches Netzwerk modellieren, das durch graphentheoretische Modelle systematisch beschrieben und untersucht werden kann (vgl. Schmidt 2023: 5).

In der Graphentheorie werden zunächst die rein topologischen Eigenschaften der Netzstruktur untersucht. Ein Graph G = (V, E) wird formal durch die Knotenmenge (V) und Kantenmenge (E) definiert. Zwei Knoten u, v ∈ V gelten als benachbart oder als Nachbarn, wenn es eine Kante {u, v} ∈ E gibt, die sie miteinander verbindet. Man unterscheidet grundsätzlich zwischen *gerichteten* (Kanten haben eine Richtung haben (u, v) ≠ (v, u)) und *ungerichteten* (Kanten haben keine Richtung (u, v) ∈ E ⇔ (v, u) ∈ E) Graphen (vgl. Tittmann 2019: 11; vgl. Dörpinghaus 2022).

Notwendigkeit

Die heutige Informationsgesellschaft produziert und verbreitet mehr Daten, als Einzelpersonen effektiv verarbeiten können. Durch das Internet – insbesondere das World Wide Web – ist eine nie dagewesene Informationsverfügbarkeit entstanden, die jedoch mit einer Überforderung der Nutzer einhergeht. Diese Entwicklung verlangt nach neuen, intelligenten Methoden der Informationsverarbeitung. (vgl. Davies/Studer/Warren 2006: 1)

Analyse

* Herausforderungen bei der Normdatenverknüpfung von Personen/ Körperschaften/ Orten in autobiografischen Quellen: doppelte und fehlerhafte Daten
* Falsche Relationen ()
* Kontrolle
* Regeln für maschinell verständliche Aussagen: Kamzelak 2018: 431-434
  + Prädikate im **Präsens**.
  + Prädikate im **Aktiv**

Andraschke, Udo/Wagner, Sarah (Hg.) (2020): Objekte im Netz: Wissenschaftliche Sammlungen im digitalen Wandel (= Digitale Gesellschaft). 1. Aufl. Bielefeld, Germany: transcript Verlag.

Baumeister, Joachim/Seipel, Dietmar (2005): Anfragesprachen für das Semantic Web. In: Informatik-Spektrum 28 (1), S. 40–44. https://doi.org/10.1007/s00287-004-0454-1.

Blumauer, Andreas/Pellegrini, Tassilo (Hg.) (2009): Social Semantic Web (= X.media.press). Berlin, Heidelberg: Springer.

Davies, John/Studer, Rudi/Warren, Paul (2006): Semantic Web Technologies Trends and Research in Ontology-based Systems.

Dengel, Andreas (Hg.) (2012): Semantische Technologien. Heidelberg: Spektrum Akademischer Verlag.

Domingue, John/Fensel, Dieter/Hendler, James A. (Hg.) (2011): Handbook of Semantic Web Technologies. Berlin, Heidelberg: Springer.

Dörpinghaus, Jens (2022): Wissensgraphen: Interdisziplinäre Perspektiven für Linked Data in den Geistes- und Sozialwissenschaften. https://doi.org/10.17175/2022\_011.

Grütter, Rolf (2008): Semantic Web zur Unterstützung von Wissensgemeinschaften. München: Oldenbourg.

Heist, Nicolas/Hertling, Sven/Ringler, Daniel/Paulheim, Heiko (2021): Wissensgraphen im Web. Wiesbaden: Springer Fachmedien.

Jeckle, Mario: XML, Standards und andere Aktivita¨ten zur Formierung des Semantic Web.

Kamzelak, Roland S. (2018): „Ei, dem alten Herrn zoll’ ich Achtung gern“. In: Richts, Kristina/Stadler, Peter (Hg.): Digitale Editionen im semantic web. Chancen und Grenzen von Normdaten, FRBR und RDF. 2016. Aufl. Allitera Verlag, München. https://doi.org/10.25366/2018.2.

Schmidt, Christin (2023): Graphentheorie und Netzwerkanalyse: Eine kompakte Einführung mit Beispielen, Übungen und Lösungsvorschlägen. Berlin, Heidelberg: Springer.

Tittmann, Peter (2019): Graphen. In: Graphentheorie. Carl Hanser Verlag GmbH & Co. KG. S. 11–28. https://doi.org/10.3139/9783446465039.001.

o. A. (2007): Scientific American: The Semantic Web. https://web.archive.org/web/20070713230811/http://www.sciam.com/print\_version.cfm?articleID=00048144-10D2-1C70-84A9809EC588EF21 (letzter Zugriff 22.07.2025).

1. Eine ausführliche Darstellung der zugrunde liegenden Technologien des Semantic Web erfolgt in Kapitel 2 dieser Arbeit. [↑](#footnote-ref-1)