

Peters, Isabella

Book Chapter — Published Version

Ontologien - oder: Eine formalisierte Repräsentation des Weltwissens

Suggested Citation: Peters, Isabella (2021) : Ontologien - oder: Eine formalisierte Repräsentation des Weltwissens, In: Smeddinck, Ulrich (Ed.): Transdisziplinäre Entsorgungsforschung am Start: Basis-Texte zum transdisziplinären Arbeitspaket "DIPRO – Dialoge und Prozessgestaltung in Wechselwirkung von Recht, Gerechtigkeit und Governance", TU Clausthal, Clausthal-Zellerfeld, pp. 105-109,
<https://doi.org/10.21268/20210609-0>

This Version is available at:
<http://hdl.handle.net/11108/512>

Kontakt/Contact

ZBW – Leibniz-Informationszentrum Wirtschaft/Leibniz Information Centre for Economics
Düsternbrooker Weg 120
24105 Kiel (Germany)
E-Mail: info@zbw.eu
<https://www.zbw.eu/de/ueber-uns/profil-der-zbw/veroeffentlichungen-zbw>

Standard-Nutzungsbedingungen:

Dieses Dokument darf zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden. Sie dürfen dieses Dokument nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen. Sofern für das Dokument eine Open-Content-Lizenz verwendet wurde, so gelten abweichend von diesen Nutzungsbedingungen die in der Lizenz gewährten Nutzungsrechte.

Terms of use:

This document may be saved and copied for your personal and scholarly purposes. You are not to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public. If the document is made available under a Creative Commons Licence you may exercise further usage rights as specified in the licence.



<https://creativecommons.org/licenses/by/4.0/>

19. Ontologien – oder: Eine formalisierte Repräsentation des Weltwissens¹

Isabella Peters

Stellen Sie sich vor, Sie müssten das Wissen der Welt – also alles, was die Menschheit weiß, und kennt – in eine Ordnung bringen. Wie würden Sie vorgehen? Würden Sie es alphabetisch sortieren, so wie Sie es aus Enzyklopädien wie Wikipedia kennen? Das würde dann wahrscheinlich eine lange Liste an Wörtern samt Erläuterungen werden. Können Sie mit dieser Ordnungsform auch die Zusammenhänge und Bezüge zwischen den Wörtern darstellen? Im Web geht das einfach über Links, auf Papier muss man ganz schön viel blättern und querlesen.

Vielleicht lassen sich die Bezüge also besser über eine hierarchische Anordnung darstellen, die ähnlich einer Baumstruktur verschiedene Äste (oder auch: Schubladen) zur Verfügung stellt. Derartige Taxonomien sind Ihnen vielleicht aus dem Biologie-Unterricht vertraut, wo Lebewesen und Pflanzen in Schubladen wie Säugetier – Raubtier – Katze eingeordnet werden. Die hierarchische Beziehung zwischen den Wörtern ist klar: Eine Katze ist ein spezielles Raubtier und damit ein spezielles Säugetier. Solche Schubladensysteme sind immer sehr ordentlich, weil jedes Stückchen Wissen der Welt einer Schublade zugeordnet wird und zum Schluss alles verstaut ist. Doch was machen Sie, wenn das Wissen nicht nur in eine Schublade passt (z.B. sind Erdbeeren die Früchte von Rosengewächsen, aber sie ließen sich auch in die Schublade Obst einsortieren)?

Wissensrepräsentation und Begriffsordnungen

Mit Fragen wie diesen beschäftigen sich u.a. die Bibliotheks- und Informationswissenschaft sowie die Informatik. Es geht dabei um die angemessene Ordnung und Repräsentation von Wissen mit dem Ziel, Wissen in seinen Kontexten zu beschreiben und dieses für die Anwendung nutzbar zu machen. In der Informationswissenschaft dient die Wissensrepräsentation der Abbildung und Einordnung von Informationsressourcen (Bücher, Bilder, Websites, etc.) mittels Ordnungssystemen, die früher vor allem als Aufstellordnung in Bibliotheken, und nun als Hilfsmittel der Modellierung von Wissen in der Künstlichen Intelligenz-Forschung genutzt wurden. Das Wiederauffinden von Ressourcen ist primäres Ziel der Wissensrepräsentation: Bibliothekskund*innen erwarten, dass Bücher zum gleichen Thema auf demselben Regalbrett stehen – oder dass sie auch Treffer für ‚Sonabend‘ erhalten, selbst wenn sie nur nach ‚Samstag‘ gesucht haben.

¹ Teile dieses Kapitels sind aus Peters 2009 übernommen.

Um dies zu gewährleisten, nutzt sie nicht die Volltexte bzw. die Wörter, Terme² oder Bezeichnungen der zu erschließenden Ressourcen allein, sondern zusätzlich Begriffe als Repräsentanten des Inhalts. Ein Begriff ist dabei die Zusammenfassung von Objekten, Ideen oder Vorstellungen unter einer Klasse, d.h. unterschiedliche Wörter können denselben Begriff versprachlichen: Die Wörter „Samstag“ und „Sonntag“ beziehen sich auf den gleichen Begriff „6. Tag der Woche“. Begriffe sind damit unabhängig von der konkreten Sprachverwendung. Die Ressource wird also durch „Stellvertreter“ oder „Metadaten“ in der Datenbasis repräsentiert und auch darüber zugänglich gemacht.

Die Aufgabe der Wissensrepräsentation ist es, Eindeutigkeit im Hinblick auf den Informationsinhalt der Ressource zu erreichen und dadurch die Unsicherheit der Nutzer – zum einen bei der Bewertung der Ressource (relevant oder nicht?) und zum anderen bei der Formulierung ihrer Suchanfrage – zu vermindern.

Wie wir oben an den Beispielen gesehen haben, bedürfen das Ordnungssystem bzw. die Metadaten dabei unbedingt einer Struktur, um effektiv den Inhalt der Informationsressourcen wiedergeben zu können. Die Methoden der Wissensrepräsentation werden daher durch verschiedene Regelwerke normiert.

In Begriffsordnungen werden Begriffe über Relationen miteinander verbunden und so in ein semantisches Netz eingeordnet. Wenn man das Katzen-Beispiel von oben zur Hilfe nimmt, weiß man, dass Katze, Raubtier und Säugetier in einer hierarchischen Beziehung zueinander stehen, und dass das bedeutet, dass eine Katze alle Merkmale eines Raubtiers (z.B. ein spezielles Gebiss) und alle Merkmale eines Säugetiers (z.B. gleichwarme Körpertemperatur) aufweist. Eine Katze lässt sich also im Gesamtgefüge der Lebewesen genau verorten oder in genau eine Schublade stecken. Damit weiß man auch in welcher hierarchischen Beziehung die Katze zu anderen Lebewesen, z.B. Fischen, steht. Derartige Ordnungen lassen sich auf andere Bereiche übertragen, z.B. auf Krankheiten (z.B. International Classification of Diseases³), technische Erfindungen (z.B. Internationale Patentklassifikation IPC⁴) oder auch auf nukleare Unfälle (z.B. Nuclear Accident Knowledge Taxonomy⁵, siehe Abbildung 1).

6. ACCIDENT CONSEQUENCES		
6.1. Consequences for environment	6.1.1. Radioactivity releases to environment	6.1.1.1. Atmosphere (gaseous) releases
		6.1.1.2. Liquid discharge
		6.1.1.3. Radioactivity fallout/deposition
		6.1.1.4. Environment contamination (Soil, air, aquatic)
	6.1.2. Impact on environment	6.1.2.1. Impact on fauna
		6.1.2.2. Impact on flora
		6.1.2.3. Impact on aquatic environments
	6.1.3. Impact on food (<i>Link to 4.2.3.8. Food and commodities restrictions</i>) (Milk, fish, mushrooms)	

Abbildung 1: Ausschnitt aus der Nuclear Accident Knowledge Taxonomy.

² Der Begriff Term wird umgangssprachlich für alles verwendet, das eine Bedeutung trägt.

³ <https://www.who.int/classifications/icd/icdonlineversions/en/>

⁴ <https://www.dpma.de/recherche/klassifikationen/patenteundgebrauchsmuster/internationalepatentklassifikation/index.html>

⁵ https://www-pub.iaea.org/MTCD/Publications/PDF/Pub1730_web.pdf

Um das Wissensnetz spannen zu können, bedarf es Relationen, die zwischen den Begriffen bestehen. Generell lassen sich zwei Arten von Relationen unterscheiden: a) paradigmatische Relationen sind fest verdrahtete Verbindungen grundsätzlicher Natur, b) die Begriffe sind über festgeschriebene Relationen der Begriffsordnung miteinander verbunden, und c) syntagmatische Relationen sind „ad hoc“-Relationen, d.h. die Verbindung der Begriffe kommt durch ihr gemeinsames Auftreten in einer Ressource als „Co-Okkurrenz“ zustande, etwa wenn eine Beziehung zwischen Freitag und Samstag aufgebaut wird, weil beide Wörter immer gemeinsam in einem Dokument erwähnt werden.

Wenden wir uns nun den Besonderheiten von Ontologien als Methode der Wissensrepräsentation zu.

Ontologien

Ontologien sind die detaillierteste Art, Wissen zu beschreiben und abzubilden und sollen im Sinne des „Semantischen Webs“ vor allem der Interaktion von Mensch und Computer sowie von Computern untereinander dienen (Guarino, Oberle, & Staab, 2009). Ontologien werden zur Darstellung von Begriffen und ihren Beziehungen untereinander genutzt mit dem Zweck, eine begrenzte Wissensdomäne vollständig abzubilden. Die klassische Definition von Ontologie lautet in der Informatik: „An ontology is an explicit specification of a conceptualization“ (Gruber, 1993).

Ontologien machen sich die Hilfsmittel der Informatik (z.B. formalisierte Sprachen), zu Nutze, um auch dem Computer das abgebildete Domänenwissen zur Verarbeitung bereitstellen zu können. Die formalisierten Relationen zwischen den Begriffen sowie definierte Regeln für ihre Anwendbarkeit sollen es dem Computer ermöglichen, automatisch Schlussfolgerungen über bestimmte Sachverhalte ziehen und sich damit der menschlichen Denk- und Arbeitsweise annähern zu können: Wenn eine Katze ein Raubtier ist, dann ist sie auch ein Säugetier, wobei auf die letzte Aussage automatisch geschlossen werden kann, wenn das dazugehörige semantische Netz von oben vorliegt. Weil Ontologien auch sog. „Instanzen“ (d.h. individuelle Objekte der realen Welt) aufnehmen können, verbinden sie die beschreibende Metaebene der Wissensrepräsentation („Eine Katze ist ein Säugetier“) mit konkreten Vorkommen in der Welt („Mimi ist die Katze von Mara Musterfrau und Mimi ist ein Säugetier“). Ontologien sind damit auch ein wichtiger Bestandteil der Forschung zur Künstlichen Intelligenz.

Da Ontologien einen starken Gebrauch von Relationen machen und dabei insbesondere die allgemeine Assoziationsrelation in eine Vielzahl von spezifischeren Relationen (z.B. `hat_Urheber`, `is_located_in`, etc.) nach Notwendigkeit aufspalten, ist ihre Ausdruckskraft zwar enorm; sie kann sich bei besonderer Detailliertheit aber nur auf eine begrenzte Wissensdomäne beziehen (z.B. auf menschliche Geninformationen wie die Gene Ontology⁶). Die Assoziationsrelationen haben ja einen grundsätzlichen Anspruch und müssen gültig für den gesamten zu beschreibenden Bereich sein. Dies ist auch die besondere Herausforderung bei der Erstellung von Ontologien: Die Ontologie benötigt immer einen Konsens hinsichtlich ihres abgebildeten Wissens und den Bezeichnungen, um überhaupt strukturiert und universell genutzt werden zu können. Aus praktischen Gründen werden wir daher mit vielen

⁶ <http://www.geneontology.org>

domänenbezogene Ontologien konfrontiert, anstatt mit einer allumfassenden Ontologie für das Weltwissen.

Die Bezeichnung „Ontologie“ wird oftmals nicht einheitlich verwendet bzw. verallgemeinernd für alle strukturierten Begriffsordnungen genutzt. Stock und Stock (2008) empfehlen, Begriffsordnungen nur dann Ontologie zu nennen, wenn sie folgende Merkmale enthalten:

- Verwendung einer standardisierten Ontologie-Sprache (z.B. OWL),
- Möglichkeit des automatischen Schlussfolgerns unter Einsatz terminologischer Logik,
- Vorkommen von und Unterscheidung nach Allgemeinbegriffen und Instanzen (z.B. „Regisseur“ und „Steven Spielberg“),
- Verwendung spezifischer Relationen (neben den Hierarchierelationen).

Ontologie(n) für DIPRO

Welche Rolle spielen Ontologien nun in DIPRO? Es ist geplant, das Wissen zum Entsorgungspfad, bzw. einer noch auszuwählenden Teil-Domäne, mit der sich DIPRO beschäftigt, nach den Beispielen von oben systematisch zu repräsentieren – also eine Ontologie zu erstellen. Zum Beispiel könnten mittels Ontologien die Narrative um das Entsorgungsproblem⁷ systematisch beschrieben und als Wissensraum visualisiert werden. So könnte gezeigt werden ob und wenn ja wie Narrative auf öffentliche Debatten und Formen der Governance-Einfluss nehmen.

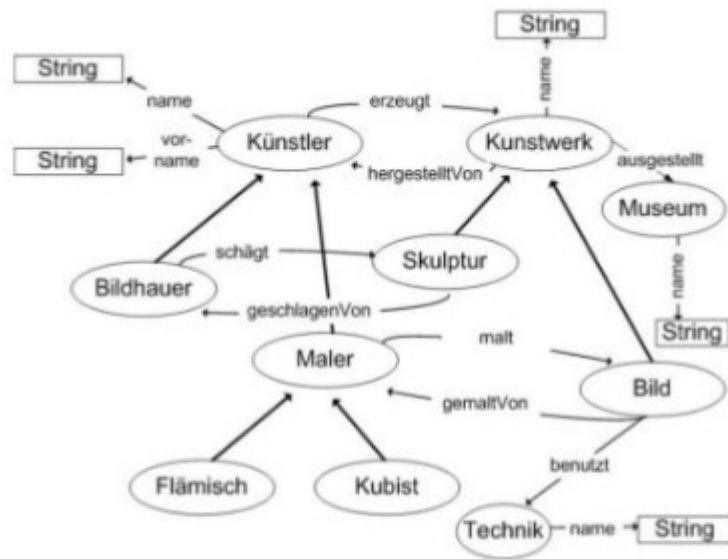
Die Erstellung der Ontologie erfolgt in drei Schritten:

1. Sammlung des Wissens anhand konkreter Dokumente, z.B. wissenschaftliche und journalistische Texte, Broschüren, Social Media-Inhalte,
2. Analyse des Wissens, z.B. im Hinblick auf vorhandene Begriffe und ihre Versprachlichung sowie die Relationen zu einander,
3. Systematisierung des Wissens und Übertragung in eine Ontologie, damit Dokumente automatisch verarbeitet werden können.

Das Ziel ist es, das (Teil-) Wissen zum Entsorgungspfad in seinen Bedeutungszusammenhängen darzustellen (siehe Abbildung 2 für eine Kunst-Ontologie). Neben einer manuellen Bearbeitung der Wissensbestände, bietet sich u.U. auch der Einsatz von automatischen Methoden an. Dies muss für die drei durchzuführenden Schritte geprüft werden.

Eine offene Frage für DIPRO ist, wie TD-Partner*innen in die Wissensrepräsentation, z.B. über einen kollaborativen Ontologieaufbau, einbezogen werden können, um dem transdisziplinären Anspruch des Projekts gerecht zu werden. Dies soll während der Laufzeit von DIPRO evaluiert werden.

⁷ Vgl. die Beiträge von Roßmann/Berg, Kap. 5, sowie Brunnengräber/Hirschhausen, Kap. 17, in diesem Bericht.



Quelle: [https://de.wikipedia.org/wiki/Ontologie_\(Informatik\)](https://de.wikipedia.org/wiki/Ontologie_(Informatik))

Abbildung 2: Beispiel für eine Ontologie in der Wissensdomäne „Kunst“.

Literaturverzeichnis

Gruber, T. R. (1993), A translation approach to portable ontology specifications. Knowledge acquisition, 5(2), 199-220.

Guarino, N., Oberle, D., Staab, S. (2009), What is an ontology? In: Staab, S., Studer, R. (Hg.), Handbook on ontologies (pp. 1-17). Springer, Berlin, Heidelberg. https://iaoa.org/isc2012/docs/Guarino2009_What_is_an_Ontology.pdf

Peters, I. (2009). Folksonomies. Indexing and retrieval in Web 2.0. Walter de Gruyter.

Stock, W. G., Stock, M. (2008), Wissensrepräsentation: Informationen auswerten und bereitstellen (Vol. 2). Oldenbourg Verlag.