

Beitrag zu: *Das aktuelle Schlagwort*

## Ontologie(n)

Ontologie ist ein überlieferter Begriff aus der Philosophie und steht dort für die *Lehre vom Sein* - genauer von den Möglichkeiten und Bedingungen des *Seienden*, ist also eng verwandt mit der *Erkenntnistheorie*, die sich mit den Möglichkeiten und Grenzen menschlichen Wahrnehmens und Erkennens auseinandersetzt.

In der *Informatik* steht man in vielen Bereichen vor der Aufgabe, *Erkanntes* oder *Erdachtes* zu repräsentieren und *Wissen* zu kommunizieren, z.B. über Fakten, Sachverhalte oder Regeln in einem technischen Anwendungsbereich, in einem Geschäftsprozess oder in einem juristischen Verfahren oder über die Inhalte von Dokumenten oder Webseiten. *Menschen* können sich gespeichertes Wissen zunutze machen, indem sie auf ihr Grund- und Kontextwissen des jeweiligen Wissensbereichs zurückgreifen, Lehrbücher, Regelwerke, Lexika und Schlagwortregister verwenden und mit den gespeicherten Inhalten verbinden. Sollen dagegen *Automaten* Such-, Kommunikations- und Entscheidungsaufgaben in Bezug auf das gespeicherte Wissen übernehmen oder Daten austauschen, die selbst Information darüber enthalten, wie sie zu strukturieren und zu interpretieren sind (sog. *Metadaten*), so benötigen sie dazu eine *Repräsentation* der zu Grunde liegenden Begriffe und derer Zusammenhänge. Dafür hat sich in einigen Zweigen der Informatik in den letzten Jahren der Begriff *Ontologie* eingebürgert. Der wohl bekannteste Definitionsversuch stammt von T. Gruber. Dieser bezeichnet Ontologie als "*explizite formale Spezifikation einer gemeinsamen Konzeptualisierung*" (orig.: *shared conceptualization*).

In diesem Sinne beschreibt eine Ontologie also einen *Wissensbereich* (*knowledge domain*) mit Hilfe einer standardisierenden Terminologie sowie Beziehungen und ggf. Ableitungsregeln zwischen den dort definierten Begriffen. Das gemeinsame Vokabular ist in der Regel in Form einer Taxonomie gegeben, die als Ausgangselemente (*modelling primitives*) Klassen, Relationen, Funktionen und Axiome enthält. Da es viele Wissensbereiche - mit einer jeweils eigenen oder sogar mehreren konkurrierenden Terminologie(n) - gibt, macht hier (im Gegensatz zur Philosophie) auch der Gebrauch des Plurals ("*Ontologien*") Sinn.

Wozu werden Ontologien in der Informatik verwendet? Gruninger und Lee [G-L 02] unterscheiden drei Anwendungsfelder: *Kommunikation*, *Automatisches Schließen* und *Repräsentation sowie Wiederverwendung von Wissen*. Sollen zwei Programme (z. B. Web-Suchmaschinen oder Software-Agenten) miteinander kommunizieren, so müssen sie entweder selbst die Interpretationsvorschrift für die Daten in sich tragen (sind also datenabhängig) oder aber sie liefern diese in Form von Metadaten aus einer beiden Seiten zugänglichen Ontologie mit. Beim *Automatischen Schließen* können Programme logische Schlüsse schon aufgrund der per Ontologie bekannten Ableitungsregeln ziehen - diese müssen also nicht stets von Neuem übermittelt werden. Ähnlich ist es bei der *Wissens-Repräsentation und -Wiederverwendung*: (vgl. dazu die ausführliche Darstellung in [Sta 02]).

Ontologien sind also in allen mit Wissen befassten Bereichen der Informatik von Bedeutung - wie etwa *Künstliche Intelligenz*, *Datenbanken* und *Informationssysteme* (im weitesten Sinne, einschließlich des globalen Informationssystems WWW). Dazu kommen angrenzende Bereiche wie *Softwaretechnik* und *Multimedia-Kommunikation* sowie *Anwendungsgebiete* wie z.B. Medizin, Rechtswesen und Wirtschaftsinformatik.

An den Ontologie-Begriff knüpfen sich viele Zusammensetzungen wie z.B. *Ontologie-Entwurf* (*ontology design*) und *Ontologie-Technik* (*ontological engineering*). Beide Begriffe machen selbstverständlich nur beim Begriffsverständnis der Informatiker Sinn - in philosophischer Deutung könnten dies höchstens angenommene Vorgehensweisen eines Demiurgen, also eines metaphysischen höheren Wesens sein. *Ontological engineering* umfasst - analog zum *Software engineering* - alles, was zur Unterstützung des *Ontologie-Lebenszyklus* dienen kann [G-L 02] - dafür wäre eigentlich der Begriff *Ontology engineering* angebrachter. *Ontologie-Entwurf* kann grundsätzlich durch einen *induktiven* Ansatz (Bildung größerer Ontologien aus mehreren kleinen "leichtgewichtigen" durch Verbindung - *merging*) oder durch einen *deduktiven* Ansatz (Festlegung allgemeiner Konzepte und Regeln durch ein Gremium oder Konsortium, Überprüfung, Standardisierung und anschließende Spezialisierung für Teilbereiche) geschehen [H-S 02].

Der Wert einer Ontologie steht und fällt mit dem Umfang der Anerkennung und Zustimmung (*ontological commitment*), die diese in der betreffenden Fachwelt erfährt. Im allgemeinen ist diese Zustimmung umso leichter zu erreichen, je mehr Entscheidungsträger und Betroffene am Entwurfsprozess beteiligt sind. Andererseits steigt der Aufwand in der Regel mit der Zahl der am Entwurf beteiligten Personen.

Besonderes, weltweites Interesse finden Ontologien in jüngster Zeit aufgrund der *Semantic Web-Initiative* des WWW-Schöpfers Tim Berners-Lee und seiner Kollegen [BHL 01]. Es beruht auf der Grundidee, Web-Dokumente (jeglicher Größenordnung) mit "Semantik" in Form von *Metadaten* ("*tags*") zu versehen, die ihren Inhalt näher beschreiben und diese durch Ableitungsregeln ("*inference rules*") miteinander zu verknüpfen. Damit sollen Suchmaschinen und andere elektronische Mechanismen wie Agenten unterstützt werden, benötigte Informationen gezielt und effizient zu finden und miteinander zu verbinden. Ontologien dienen dazu, die dafür benötigten Grundlagen an Metadaten und Verknüpfungsregeln zu liefern. Das bedeutet, dass sich z.B. zwei Agenten über ihre Aufgaben und Ergebnisse mit Hilfe einer beiden verfügbaren Ontologie verständigen können.

Für die Entwicklung und Prüfung von Ontologien sind in den letzten Jahren eine Reihe von Sprachen, Methoden und Werkzeugen entstanden bzw. verfügbar gemacht worden. Im Zusammenhang mit dem Semantic Web-Ansatz sind hier vor allem *XML* (*Extensible Markup Language*) und *RDF* (*Resource Description Framework*) zu nennen: XML für die Annotierung und Strukturbeschreibung von Daten und Dokumenten, RDF für die Möglichkeit, Ressourcen durch Eigenschaften zu beschreiben und diesen Werte, darunter auch Verweise auf andere Ressourcen zuzuordnen. Dieser Ansatz fußt auf der bekannten Grundidee, semantische Netze als Graphen aufzufassen. Die Mächtigkeit dieser Sprachmittel und die Lesbarkeit von Dokumenten lässt sich durch den Einsatz von *Schema-Definitionen* (*XSD - XML Schema Definition* und *RDFS - RDF Schema*) beträchtlich steigern.

Daneben hat die US-Behörde DARPA *DAML* (*DARPA Agent Markup Language*) als Agenten-Verständigungssprache - "an Esperanto for Machines" [DAML 02] definiert. Diese wurde im Hinblick auf Ontologien mit *OIL* (ursprünglich: *Ontology Inference Layer*, neuerdings: *Ontology Interchange Language*) zu *DAML+OIL* kombiniert und wird vom WWW-Consortium als Standard für die Repräsentation von Metadaten und Ontologien vorgeschlagen. Dieser Ansatz beruht auf den von der KI her bekannten *Frames* sowie auf *Description Logics* zur Beschreibung von Semantik und logischen Verknüpfungen. Für nähere Hinweise dazu und zu Werkzeugen wie *Ontologie-Editoren* und Entwicklungsumgebungen vgl. [Sta 02]

Ontologien sind bereits für verschiedene Wissensgebiete entwickelt worden, so z.B. für Entscheidungsunterstützungs-Systeme (Decision Support Systems, DSS vgl. dazu [H-S 02]) oder für das Wissensmanagement (vgl. [Sta 02]). Im kommerziellen Bereich spielen sie - z.B. als Grundlage für e-business-Systeme - bereits eine wichtige Rolle. Für den Bereich der Informationssysteme hat die IFIP-Arbeitsgruppe FRISCO (Framework of Information System Concepts) ein umfassendes konzeptuelles Rahmenwerk vorgeschlagen, das auch als allgemeine Basis für weitere, spezialisierte Ontologien dienen könnte [FHL+ 98]. Einen damit verwandten Ansatz haben Y. Wand und R. Weber entwickelt [W-W 90].

In der Softwaretechnik gewinnen Ontologien zur Zeit in zweierlei Hinsicht an Bedeutung:

(1) Ontologien als Hilfsmittel und Wissensfundus bei der Software-Entwicklung. Will man die Idee von *Web services*, d.h. im Web verteilten, wieder- und weiterverwendbaren *Anwendungssystemen* oder *-komponenten* verwirklichen, so müssen diese auf einem von allen potentiellen Benutzern getragenen gemeinsamen Struktur- und Begriffsverständnis des betreffenden Anwendungsgebiets beruhen. Ontologien können hier dabei helfen, den Anteil an wiederverwendbaren Ergebnissen (z.B. Konzepten oder Modellen) in den sog. "frühen" Phasen deutlich zu erhöhen.

(2) Softwaretechnik als Gegenstand einer Ontologie d.h. als standardisiert zu beschreibendes Wissensgebiet. Für eine solche *Software Engineering Ontology* hat die SWEBoK-Initiative (*Software Engineering Body of Knowledge*) wertvolle Vorarbeit geleistet. In diesem Zusammenhang sind auch deutschsprachige Arbeiten an einem "Begriffsnetz" für die Softwaretechnik zu nennen [HBB+ 94], [GI 02].

Die wichtigsten, noch ungelösten Probleme mit Ontologien betreffen aus heutiger Sicht die folgenden Punkte:

- Wie lassen sich gut verwertbare Metadaten für sehr große Ressourcen-Bestände erzeugen und konsistent weiterentwickeln? Lässt sich das Annotieren von Dokumenten sinnvoll automatisieren?

- Können Ressourcen klar und eindeutig klassifiziert werden, z.B. in Dokumente, Daten, Metadaten, physische und virtuelle Aktoren (*actors*), physische Einheiten?

- Wie lassen sich Metadaten in (möglicherweise überlappende und inkonsistente) Ontologien einordnen? Wie werden Synonyme, Homonyme und zirkuläre Definitionen behandelt?

- Macht es Sinn, nach einer gemeinsamen, allen Ontologien unterliegenden *top level*-, Universal- oder Meta-Ontologie zu suchen - oder führt dies zu ähnlichen Schwierigkeiten wie die Suche nach einem objektiven Weltbild in der Physik?

Nicht zuletzt stellt sich die Frage nach einer philosophischen Fundierung eines (informatischen) Ontologiebegriffs. Für Philosophen ist die Antwort auf die ontologische Grundfrage "Was ist?" keineswegs so trivial, wie es zunächst scheinen mag, da

(1) unsere Wahrnehmung der Wirklichkeit durch unsere Wahrnehmungsorgane gefiltert wird, wir also keineswegs sicher sein können, dass die Welt so beschaffen ist, wie wir sie wahrnehmen,

(2) die Beantwortung zwangsläufig mit sprachlichen Mitteln erfolgen muss, wir uns also bei der Beschreibung des Wahrgenommenen möglicherweise nochmals durch die Wahl unserer Begriffe, Worte und Bilder vom eigentlich "Seienden" entfernen.

Die sogenannte *traditionelle Ontologie* war von alters her dem Gebiet der *Metaphysik* zugeordnet Während sich die Physik mit der Existenz der (materiellen) Dinge, deren Bewegungen und der natürlichen Notwendigkeit der Materie (*res concretae*) befasst, sind es bei der Metaphysik die Gründe der Dinge und der dahinter stehenden Ideen (*res abstractae*). Schon Kant beurteilte die traditionelle Ontologie als "anmaßend" und ersetzt die (doktrinäre) Annahme "gegebener" Objekte durch eine Transzendental-Philosophie, die "die Bedingungen und ersten Elemente aller unserer Erkenntnis a priori enthält" [Rit 98, S. 1194] und damit die Erkenntnis als Filter für das "Seiende" einsetzt.

Im 20. Jahrhundert gibt es Versuche zu einer logisch präzisen Fassung, z.B. bei E. Husserl als *formale Ontologie* mit Mitteln der "reinen Logik". Husserl unterscheidet explizit zwischen dem (materialen, in der Natur verankerten) *Ding schlechthin* und dessen *Sinn*, dem keine "realen Eigenschaften" zugeordnet werden können. Die Analyse von Wahrnehmungen und Erlebnissen im Hinblick auf ihren Sinn führt zur Erkenntnis verschiedener "Seinsregionen" und damit zu - jetzt im Plural auftretenden (!) *materialen* bzw. *regionalen Ontologien*. Modernere Ansätze laufen darauf hinaus, der Ontologie ihren metaphysischen Charakter zu nehmen. So rückt in der *Analytischen Philosophie* der sprachliche Gesichtspunkt ins Zentrum: Eine Ontologie ist sprecher-bezogen und gibt dem Hörer oder Leser die Existenzvoraussetzungen an, an die er sich durch das Akzeptieren einer Sprache oder Theorie bindet [Rit 98, S. 1199]. Eine radikal-formalistische Ausprägung dieses Ansatzes formuliert W. van O. Quine: "To be is to be the value of a bound variable" (zit. n. [Sce 99]).

Hier kommen sich zu guter Letzt Philosophie und Informatik doch noch nahe: Ontologie wird nicht mehr als metaphysische Deutung des Wesens aller Dinge, sondern als praxisgeleitete sprachliche Kategorisierung von Lebens- und Wissensbereichen verstanden. In diesem Sinne sollte es - trotz der damit verbundenen und von P. Janich mit Recht aufgezeigten Gefahren von Missverständnissen [Jan 01] - auch Informatikern erlaubt und für sie möglicherweise sogar nützlich sein, sich mit Ontologie(n) zu beschäftigen [H-B 01].

### **Literaturhinweise:**

- [BHL 01] T. Berners-Lee, J. Hendler, O. Lassila: The Semantic Web. Scientific American, [http://www.scientificamerican.com/2001/0501issue/0501\\_berniers-lee.html](http://www.scientificamerican.com/2001/0501issue/0501_berniers-lee.html) (May 2001)
- [FHL+ 98] E.D. Falkenberg, W. Hesse, P. Lindgreen, B.E. Nilsson, J.L.H. Oei, C. Rolland, R.K. Stamper, F.J.M. Van Assche, A.A. Verrijn-Stuart, K. Voss: FRISCO - A Framework of Information System Concepts - The FRISCO Report. IFIP WG 8.1 Task Group FRISCO. Web version: <ftp://ftp.leidenuniv.nl/pub/rul/fri-full.zip> (1998)
- [GI 02] Gesellschaft für Informatik (GI)-Arbeitskreise "Terminologie der Softwaretechnik" und "Begriffe für Vorgehensmodelle": Informatik-Begriffsnetz. <http://www.tfh-berlin.de/~giak/> (Stand v. 1.6.2002)

- [G-L 02] M. Gruninger, J. Lee: *Ontology - Applications and Design*. CACM 45.2, pp. 39-41 (Feb. 2002)
- [H-B 01] W. Hesse, H. v. Braun: *Wo kommen die Objekte her? Ontologisch-erkenntnistheoretische Zugänge zum Objektbegriff*. In: K. Bauknecht et al. (eds.): *Informatik 2001 - Tagungsband der GI/OCG-Jahrestagung*, Bd. II, S. 776-781. books\_372ocg.at; Bd. 157, Österr. Computer-Gesellschaft 2001
- [HBB+ 94] Hesse, W., Barkow, G., v. Braun, H., Kittlaus, H.B., Scheschonk, G.: *Terminologie der Softwaretechnik - Ein Begriffssystem für die Analyse und Modellierung von Anwendungssystemen*, Informatik- Spektrum 17.1, S. 39-47 und 17.2, S. 96-105 (1994)
- [H-S 02] C. W. Holsapple, K.D. Joshi: *A Collaborative Approach to Ontology Design*. CACM 45.2, pp. 42-47 (Feb. 2002)
- [Jan 01] P. Janich: *Wozu Ontologie für Informatiker? Objektbezug durch Sprachkritik*. In: K. Bauknecht et al. (eds.): *Informatik 2001 - Tagungsband der GI/OCG-Jahrestagung*, Bd. II, pp. 765-769. books\_372ocg.at; Bd. 157, Österr. Computer-Gesellschaft 2001
- [Rit 98] J. Ritter, K. Gründer: *Historisches Wörterbuch der Philosophie*, Bd. 6: Mo-O. Wiss. Buchgesellschaft Darmstadt 1998
- [Sce 99] P. Scheffe: *Softwaretechnik und Erkenntnistheorie*, Informatik-Spektrum 22.2, S. 122-135 (1999)
- [Sta 02] S. Staab: *Wissensmanagement mit Ontologien und Metadaten*. Informatik-Spektrum 25.3, pp. 194-209 (2002)
- [W-W 90] Y. Wand, R. Weber: *An ontological model of an information system*. IEEE Transact. on Software Eng. 16 (11), pp. 1282-1292 (1990)