

Wie wissende Computer den Menschen besser verstehen

Passen Mähdrescher und Ackerschlagkartei zusammen?

Bei allem technischen Fortschritt kommen uns Computer manchmal ziemlich dumm vor. Offensichtliche Zusammenhänge, die Menschen intuitiv erkennen, bleiben Computern oft verborgen. Informiert zum Beispiel ein Mitglied eines Opel-Fan-Clubs seine Vereinskameraden, er habe sich einen neuen Manta gekauft, wird wohl keiner der Kameraden davon ausgehen, er sei Aquarist geworden und habe sich gerade einen Rochen gekauft. Computeranwendungen wie Suchmaschinen fällt es aber immer noch schwer zu entscheiden, wonach ein Nutzer gerade suchen möchte.

Homonyme (auch als Teekesselchen bezeichnet, wie im Manta-Beispiel) und Synonyme bereiten Computern weiterhin Probleme, da dem Computer notwendiges Hintergrundwissen zu dem jeweiligen Kontext fehlt. Doch natürlich wird auch an diesem Problem intensiv geforscht und es wurden bereits beachtliche Fortschritte erzielt. Um aber zu verstehen, wie Computer Wissen erlangen, müssen einige grundlegende Begriffe voneinander abgegrenzt werden.

Während man in der Frühzeit der Computerisierung überwiegend von elektronischer Datenverarbeitung (kurz EDV) sprach, ist heutzutage der Begriff der Infor-

werden. Zeichen sind grafische Symbole wie Buchstaben, Zahlen oder auch Sonderzeichen wie zum Beispiel Smileys. Zeichen in einer syntaktisch korrekten Zusammen-

kolonnen (Daten) wurden zum Beispiel in der Lohnbuchhaltung aufsummiert oder sortiert. Heutzutage steht der Anwendernutzen stärker im Vordergrund, sodass wir von Informationstechnologie (IT) sprechen. Damit aber ein Computer souverän mit den oben erwähnten Homonymen oder Synonymen umgehen kann, benötigt auch er Wissen über die Welt oder zumindest über einen Teilbereich der Welt, eine sogenannte Domain. Wer Weltwissen besitzt, muss nicht immer alles erneut erklärt bekommen, da man eine Wissensbasis voraussetzen und darauf aufbauen kann. Im Opel-Club ist es eben nicht notwendig, explizit darauf hinzuweisen, dass mit Manta kein Fisch, sondern ein mittlerweile historisches Sportcoupé gemeint ist.



Mähdrescher und Ackerschlagkartei müssten sich besser verstehen, wenn Computer über landwirtschaftliches Hintergrundwissen verfügen.

mationstechnologie (kurz IT) geläufiger. Und in der Tat muss zwischen Zeichen, Daten, Informationen und Wissen unterschieden

setzung ergeben Daten wie Zahlen, Wörter oder auch Sätze. So ist 123,4 eine syntaktisch korrekte Zahl, 1,23,4 jedoch nicht, da jede Zahl nur maximal ein Kommazichen enthalten darf. Fügt man Daten eine für Menschen relevante Bedeutung hinzu, erhält man Informationen. Um Informationen einzuordnen und bewerten zu können, bedarf es einer Wissensbasis, wobei Wissen wiederum aus vielen miteinander verknüpften und somit in Beziehung stehenden Informationen besteht. Studenten erkläre ich den Zusammenhang gerne anhand der drei Zeichen: ,17. In der Reihenfolge 1,7 bilden die drei Zeichen eine syntaktisch korrekte Zahl. Sage ich einem Studenten, dass dies seine Note im Fach Electronic Business sei, wird dieses Datum für ihn zu einer Information. Kennt er auch die Noten der Kommilitonen (Wissen), kann er einordnen, wie gut seine Note wirklich ist.

Aber auch heutzutage findet noch Datenverarbeitung statt. Die modernen Begriffe Data Mining oder Data Lakes verwenden zu Recht den Begriff der „Daten“ und nicht den der „Informationen“. Beim Data Mining werden Rohdaten nach Mustern durchforscht, die anschließend durch einen Menschen interpretiert und erst dadurch zu Informationen werden. Data Lakes, also Datenseen, nehmen in Big-Data-Szenarien sehr große Mengen an Rohdaten schnell, aber unsortiert auf – ähnlich wie ein See große Wassermengen nach einem Wolkenbruch. Auch hier werden die Daten erst später interpretiert und nach sinnvollen Informationen in einem spezifischen Kontext durchsucht. Auch in der Landwirtschaft produzieren Maschinen und Sensoren bereits heute riesige Datenmengen, die hoffentlich später einmal zu wertvollen Informationen werden.

EDV oder Informationstechnologie?

In der Zeit der EDV wurden somit meist nur „dumme“ Daten verarbeitet. Lange Listen mit Zahlen-

Ontologien repräsentieren Wissen

Um Computern „Wissen“ beizubringen, bedient man sich sogenannter Ontologien. Während der Begriff der Ontologie in der

91 Das Bauernblatt erreicht 91 % der landwirtschaftlichen Betriebe in Schleswig-Holstein.

Quelle: Mediaanalyse agriMa 2019



Wenn Sie mehr wissen wollen:
bauernblatt.com

bauernblatt

Philosophie „das Seiende“ bedeutet, also beschreibt, was es wirklich auf der Welt gibt, steht er in der Informatik für die digitale Repräsentation von Wissen. Das Wissen wird häufig in Form von Beziehungsnetzen, sogenannten semantischen Netzen, gespeichert, in denen Begriffe miteinander in Beziehung stehen. Jeder Manta ist ein Auto, aber nicht jedes Auto ist ein Manta. Somit gelten für jeden Manta alle Eigenschaften, die auch für ein Auto gelten, nicht aber umgekehrt. Hat man für ein Auto definiert, dass es vier Räder hat, muss auch jeder Manta vier Räder haben. Der Informatiker spricht davon, dass sich die Eigenschaften der höheren Kategorie „Auto“ auf die niedrigere Kategorie „Manta“ vererbt haben. Man kann mit diesem Wissen folgern, dass jedes Fahrzeug mit nur drei Rädern unmöglich ein Manta sein kann.


Semantische Netze sind nicht neu. In der Biologie werden Tiere und Pflanzen über Taxonomien seit Jahrhunderten in Beziehung gesetzt. Zur Klasse der Säugetiere gehört auch die Familie der Katzen. Somit ist jede Katze ein Säugtier und „erbt“ damit alle gemeinsamen Eigenschaften der Säugetiere. Darüber hinaus haben Katzen aber auch spezifische Eigenschaften, die sie zum Beispiel von Hunden abgrenzen. Zu jeder Kategorie gehören Individuen, die ebenfalls die Eigenschaften der jeweiligen Kategorie erben. Zur Art Homo sapiens gehören diverse verstorbene und noch lebende Individuen

(also Menschen), die spezifische Eigenschaften aufweisen. Die Beziehungen der Menschen untereinander werden durch Social Graphs in

der Menschheitsgeschichte stellt Wikipedia dar. Auch hier sind Begriffe über Hyperlinks mit anderen Begriffen in Beziehung gesetzt. Es

ANZEIGE

Mehr Infos unter:
www.kws.de/extras



Sozialen Netzen wie Facebook und Xing oder durch Stammbäume re-präsentiert. Beziehungen können durchaus vielfältig sein und haben

verwundert daher kaum, dass Wissenschaftler daran arbeiten, diese Wissensbasis zu formalisieren und Computern verfügbar zu machen.

ANZEIGE

INITIO BIRD PROTECT

Mit gezielter Beizung den Maisbestand sichern!

www.kws.de/beizung-mais



Ausprägungen wie „A kennt B“, „B arbeitet für C“, „C ist Kind von D“ et cetera.

Auch in den Sprachwissenschaften existieren schon lange Thesauri, die die Bedeutungen von Wörtern in Beziehung setzten. Einige Sprachwissenschaftler sind sogar der Meinung, dass Wörter ihre Bedeutung ausschließlich durch den gemeinsamen Gebrauch mit anderen Wörtern bekommen. Die derzeit größte Wissensrepräsentation

Um computerlesbare Ontologien zu erstellen, haben Informatiker formale Sprachen wie die Web Ontology Language (OWL) entwickelt, die technisch auf dem Resource Description Framework (RDF) basiert. Da das Erstellen einer Ontologie viel Fachwissen voraussetzt, ist es einzelnen Forschergruppen nicht möglich, eine einzige große und somit sehr komplexe Ontologie mit dem gesamten Weltwissen zu erstellen. Daher versucht man derzeit, ein eher loses Netz aus dezentralen spezialisierten Ontologien zu erstellen, die sich aber aufeinander beziehen und somit ergänzen. Auch für die Landwirtschaft hat es mit agroRDF erste Schritte gegeben, eine Wissensbasis zu erstellen. Eine Vorreiterrolle in Deutschland nimmt hierbei das Kuratorium für Technik und Bauwesen in der Landwirtschaft (KTBL) in Darmstadt ein.

Anwendungen von Ontologien

Um Informationen automatisiert für den jeweiligen Anwendungsfall richtig zu interpretieren, bedarf es also ausreichend guter Ontologien. Auch Google arbeitet natürlich an diesem Thema, um seine Anwendungen immer intelligenter zu machen. Die Inhalte von E-Mails und angehängten Dateien werden analy-

siert und automatisch richtig interpretiert. Bekommt man per E-Mail eine Reisebestätigung an seinen Google-Account, werden sehr wahrscheinlich automatisch die Flugdaten in den Google-Kalender eingetragen und der Ort des Hotels bei Google Maps aufgefördert markiert. In Zweifelsfällen können semantische Systeme gezielt nachfragen, zum Beispiel ob in einem speziellen Falle mit Manta ein Auto oder ein Fisch gemeint ist.

Mähdrescher und Ackerschlagkartei passen

In der Landwirtschaft könnten semantische Technologien helfen, das Schnittstellenproblem zwischen unterschiedlichen IT-Systemen, Anwendungen oder Landmaschinen zu lösen. Wenn man ausreichend gute Ontologien für die Domain Landwirtschaft erstellt hat, müssten sich auch Mähdrescher und Ackerschlagkartei besser verstehen. Stellt ein Mähdrescher einer Ackerschlagkartei ein Datenfeld namens „Getreidedurchfluss“ zur Verfügung, kann die Ackerschlagkartei die Bedeutung dieses Datenfeldes erkennen und somit richtig zuordnen, auch wenn das korrespondierende Datenfeld in der Ackerschlagkartei vielleicht „aktuelle Erntemenge“ heißt. Man müsste mit ausreichend präzisen Ontologien Schnittstellen zwischen beliebigen Systemen ad hoc automatisch generieren können. Ein Traum für jeden Wissenschaftler und jeden Landwirt, der heutige Precision-Farming-Technologien einsetzt.

Eine weitere Anwendung von Ontologien ist die Errichtung des sogenannten „Semantik Web“, also eines Internets, das seine Bedeutung kennt. Suchmaschinen könnten dadurch nochmals bessere Suchergebnisse generieren. Sucht ein Mechaniker nach einer „Mutter“ im Web, werden ihm künftig keine Windeln mehr angezeigt. Auch für einen automatisieren Handel wäre es notwendig, dass der Handelsroboter über Hintergrundwissen verfügt. Ansonsten würde er vielleicht doch einen preiswerten Rochen für den Opel-Club kaufen und „Manni“ wäre enttäuscht.

Prof. Michael Clasen
Hochschule Hannover
Tel.: 05 11-92 96-15 88
michael.clasen@hs-hannover.de



Auch in der Landwirtschaft produzieren Maschinen und Sensoren bereits heute riesige Datenmengen. Fotos: landpixel