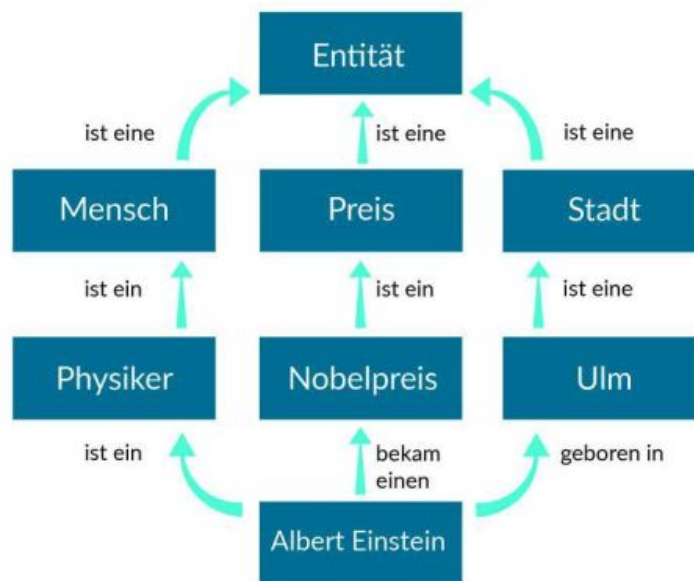


Programmcode preisgekrönter Wissensdatenbank ist nun für jedermann nutzbar

31.08.2017, 13:00 | Wissenschaft, Forschung, Bildung

Pressemitteilung von: *Universität des Saarlandes*

Presseagentur: *Universität des Saarlandes*



Durch automatisch generierte Entitäten und deren Verbindungen können Computer selbst komplexe Suchanfragen beantworten. (MPI)

Im Internet hat fast jedes Wort mehr als eine Bedeutung, was die Suche oft mühsam macht. Daher verwenden moderne Suchmaschinen sogenannte Wissensbanken. Diese enthalten Filme, Firmen, Personen und Produkte. Eine der ersten Wissensbanken ist Yago, entwickelt von Wissenschaftlern am Max-Planck-Institut für Informatik in Saarbrücken und der Télécom ParisTech in Paris. Vergangene Woche wurde Yago von der wichtigsten wissenschaftlichen Zeitschrift für Künstliche Intelligenz, dem Artificial Intelligence Journal (AIJ), ausgezeichnet. Die Forscher veröffentlichen nun auch den Programmcode, sodass jeder die Wissensbank anpassen und erweitern kann.

„Wenn ich zum Beispiel nach dem Begriff Allianz suche, dann ist das für die jeweilige Suchmaschine nur eine Ansammlung von Buchstaben, ohne Bedeutung“, erklärt Professor Gerhard Weikum, wissenschaftlicher Direktor am Max-Planck-Institut für Informatik in Saarbrücken. „Mithilfe einer Wissensbank kann diese Buchstabenkette auf mögliche Bedeutungen abgebildet werden, zum Beispiel auf den Allianz-Versicherungskonzern oder die Rebellen-Allianz aus den Star-Wars-Filmen“. Dieses Hintergrundwissen ist heute bei Suchmaschinen nicht mehr wegzudenken, nur durch Wissensbanken kann beispielsweise Google neben den Ergebnissen der Web-Suche auch Börsenkurse, Logos und den Geschäftsführer der Allianz-Versicherung einblenden.

Die Wissensbanken begannen als akademische Forschungsprojekte. „Vor allem Yago und nur wenig später ‚DBpedia‘ haben in diesem Feld Pionierarbeit geleistet“, sagt Weikum. Das Yago-Projekt war 2007 Thema der Dissertation von Fabian Suchanek am Max-Planck-Institut für Informatik in Saarbrücken. Mehr und mehr Forscher beteiligten sich an dem Projekt. Heute ist Yago eine Zusammenarbeit des Max-Planck-Instituts und der Télécom ParisTech Universität, wo Suchanek inzwischen eine Professur innehat, sowie der Max-Planck-Ausgründung Ambiverse. Yago enthält das Wissen von Wikipedia und anderen Quellen in für den Computer lesbarer Form. So weiß das System beispielsweise, dass die

Allianz ihren Hauptsitz in München hat, aber auch, dass die Rebellen-Allianz im Star-Wars-Universum kämpft.

Da momentan viele Programme in den unterschiedlichsten Industriezweigen mit Hilfe von Künstlicher Intelligenz effizienter und vor allem einfacher bedienbar gemacht werden sollen, ist Yago vielfältig im Einsatz. Anwendungen können mit Hilfe von Yago in mehreren Sprachen suchen oder Fakten sowohl räumlich als auch zeitlich einordnen. Suchanfragen wie „Nenne mir alle Wissenschaftler, die im 20. Jahrhundert gelebt haben und mit einem Nobelpreis ausgezeichnet wurden, sowie im weiteren Umkreis von Stuttgart geboren wurden!“ sind somit möglich. Ein Beispiel für die Nutzung liefert Primal.com, ein kanadisches Start-up, das andere Unternehmen mit Hilfe von Yago dabei unterstützt, die Interessen von individuellen Kunden zu verstehen, um somit Inhalte und Produkte besser vorschlagen zu können. Das prominenteste Beispiel der vergangenen zehn Jahre ist wohl, dass Yago von IBM als Bestandteil des Watson Systems verwendet wurde, welches 2011 die Quiz-Show „Jeopardy!“ gewann.

Einen weiteren Anwendungsfall liefert Ambiverse selbst. Vor wenigen Monaten analysierte das Spin-off des Max-Planck-Instituts die Panama Papers mittels Yago. Innerhalb weniger Stunden konnte Ambiverse so neue Erkenntnisse über die Inhaber der Panama-Konten gewinnen, die man manuell nur mit größtem Aufwand hätte finden können. Eine solche Analyse war nur möglich, weil Yago alle Personen in eine Struktur einordnet. Bis dato konnten Computer zwar Unmengen an Daten speichern, jedoch diese weder einordnen noch verstehen. Mit Yagos Struktur kann der Computer hingegen beispielsweise zwischen Gerd Müller, Fußballweltmeister aus dem Jahr 1974 und „Bomber der Nation“, und Gerd Müller, CSU-Entwicklungsminister, unterscheiden. „YAGO ordnet Personen das Umfeld zu, sodass ganz einfach festgestellt werden kann, ob nun mehr Sportler oder mehr Politiker Konten in Panama haben“, erläutert Johannes Hoffart, Geschäftsführer von Ambiverse. Solche Strukturen wurden ursprünglich per Hand erstellt. Dies ist jedoch sehr aufwendig – sowohl in puncto Produktion als auch Überprüfung.

Für Yago haben die Forscher ein Verfahren entwickelt, das dieses Dilemma auf geschickte Art umgeht. Sie schürfen systematisch in der Wissensgrube Wikipedia. Nicht nur das Wissen, ob es sich um Sportler oder Politiker handelt, wird so in maschinenlesbare Form gebracht, sondern auch Beziehungen zwischen diesen. Die Informatiker bezeichnen diese als „Relationen“. So verbindet die Relation „hat Hauptsitz in“ die Entität „Allianz-Versicherung“ mit der Entität „Stadt München“. In Yago wird jeder Wikipedia-Artikel zu einer Entität der Wissensbank. Auf diese Weise schaffen es die Forscher, in Yago fast 17 Millionen Entitäten und 150 Millionen Beziehungen zwischen diesen abzuspeichern.

Unter anderem dafür wurden sie in der vergangenen Woche mit dem Prominent Paper Award des Artificial Intelligence Journal (AIJ), der wichtigsten wissenschaftlichen Zeitschrift für Künstliche Intelligenz, ausgezeichnet. Der Preis würdigt herausragende Publikationen der vergangenen sieben Jahre, die in ihrer Bedeutung und ihrem Einfluss außergewöhnlich sind. Die Wissenschaftler nahmen ihn in Melbourne entgegen, wo mit der IJCAI die bedeutendste wissenschaftliche Konferenz im Bereich Künstliche Intelligenz stattgefunden hat.

Heute veröffentlichen die Forscher den Quellcode ihrer Wissensdatenbank auf der Plattform GitHub unter der Open Source Lizenz GNU GPL v3. Diese Softwarelizenz stellt sicher, dass jeder den damit geschützten Programmcode ausführen, studieren, verändern und teilen darf. „Mit Yago bekommt die Entwickler-Gemeinde eine Wissensbank von hoher Qualität“, erklärt Professor Fabian Suchanek, Gründer des Projekts. „Von der Veröffentlichung erhoffen wir uns nicht nur weitere Anwendungen von Yago, sondern auch Beiträge von der Entwickler-Gemeinde“. Den zugehörigen Quellcode kann sich nun jedermann unter dem folgenden Link herunterladen: <https://github.com/yago-naga/yago3>

Weitere Informationen:

Projektseite „YAGO“

<https://www.mpi-inf.mpg.de/de/departments/databases-and-information-systems/research/yago-naga/yago/>

The Era of Cognitive Systems – An inside look at IBM Watson and how it works

<http://www.redbooks.ibm.com/redpapers/pdfs/redp4955.pdf>

Von den Panama Papers zum eigenen Nachrichtendienst für das Unternehmen

<https://idw-online.de/de/news660809>

Weitere Fragen beantworten:

Professor Dr. Gerhard Weikum
Max-Planck-Institut für Informatik
Saarland Informatics Campus
Tel.: +49 681 9325-5000
E-Mail: weikum@mpi-inf.mpg.de

Professor Dr. Fabian M. Suchanek
Télécom ParisTech
Tel.: +33 14581 8062
E-Mail: suchanek@telecom-paristech.fr

Dr. Johannes Hoffart
Ambiverse GmbH
Saarland Informatics Campus
Tel.: +49 681 3720 309-1
E-Mail: johannes@ambiverse.com

Redaktion:
Gordon Bolduan
Kompetenzzentrum Informatik Saarland
Saarland Informatics Campus E1.7
Tel.: +49 681 302 70741
E-Mail: gbolduan@mmci.uni-saarland.de

Hinweis für Hörfunk-Journalisten: Sie können Telefoninterviews in Studioqualität mit Wissenschaftlern der Universität des Saarlandes führen, über Rundfunk-Codec (IP-Verbindung mit Direktanwahl oder über ARD-Sternpunkt 106813020001). Interviewwünsche bitte an die Pressestelle (0681/302-3610).

Quelle: idw

Portrait

-

News-ID: 966225 • Views: 588 (Stand: 30.10.2026)

Link zur Pressemitteilung:

<https://www.openpr.de/news/966225/Programmcode-preisgekroenter-Wissensdatenbank-ist-nun-fuer-jedermann-nutzbar.html>