



Professor Dr. Adrian Paschke

hat sein Studium der Informatik an der Friedrich-Alexander-Universität Erlangen-Nürnberg absolviert und seine Promotion an der Technischen Universität München erlangt.

Sein wichtigstes Projekt ist das „Corporate Semantic Web“, sein zweites Zuhause – neben Berlin – quasi das Internet. Bereits als Grundschüler schrieb er Programme für den elterlichen Computer und fragt sich seitdem: „Was sind Information und Wissen überhaupt?“ Er kommt dieser Frage bei seiner Arbeit täglich näher.

Professor Paschke

was ist das „Semantic Web“?

„Nehmen wir einmal an, Sie suchen einen Begriff wie Jaguar. Da gibt es schließlich mehrere Möglichkeiten: Zum einen, klar, kommen die Einträge für eine Raubkatze. Dann gibt es Hinweise auf Süd- und Mittelamerika als Verbreitungsgebiet und das goldgelbe Fell mit den schwarzen Tupfen. Gut. Aber Sie wollen noch mehr darüber wissen, z.B. die bevorzugten Beutetiere. Falls Sie einen Survivaltrip ins Amazonasgebiet planen und Ihr zweiter Name „Vorsicht“ ist, dann möchten Sie auch noch wissen, wie der Pfotenabdruck des Tieres aussieht, oder? Dann geben Sie vermutlich in eine Online-Suchmaschine Jaguar ein. Und was fällt dem World Wide Web dazu ein?

Unglaublich viel! Ihr Computer gibt den Befehl, alle verfügbaren Dokumente nach dem Schlüsselbegriff Jaguar zu durchsuchen. Die Treffer werden in einer schier endlosen Liste zusammengefasst. Fertig. Doch das Systematisieren und Interpretieren ist nun Ihre eigene Aufgabe. Sie finden verschiedenste Informationen zur amerikanischen Raubkatze, darunter auch ihre Beutetiere. Allerdings auch massenhaft Daten zu einem britischen Fahrzeughersteller, zu weiteren Firmen, Organisationen ... und der Tatzendruck? Den finden Sie möglicherweise gar nicht – nämlich dann, wenn diese Informationen zu Jaguar dazu in wissenschaftlichen Publikationen enthalten sind, die statt Jaguar allein den lateinischen Begriff *Panthera onca* verwenden.

Das ist ärgerlich und liegt daran, dass das World Wide Web prinzipiell dumm ist. Ihm fehlt jedes semantische Hintergrundwissen,

also die Fähigkeit, die Bedeutung von Begriffen zu verstehen. Doch jetzt kommt das Semantic Web ins Spiel. Das Semantic Web erweitert das ursprüngliche World Wide Web mit dem Ziel, die Informationen des Webs für Maschinen verständlich und damit besser verarbeitbar zu machen. Inhalte werden so angereichert, dass eine Automatisierung auch auf der Bedeutungsebene möglich wird. Doch was heißt das konkret?

Ein semantischer Suchassistent wird Sie zunächst fragen, ob Sie die Raubkatze, den Fahrzeughersteller oder etwas anderes suchen. Entscheiden Sie sich für die Raubkatze, bekommen Sie nicht nur alle Dokumente angezeigt, die das Wort Jaguar enthalten und sich auf die Raubkatze beziehen, sondern auch solche zu *Panthera onca* – dafür aber keine Fahrzeuge mehr!

Das Semantic Web nützt dem Amazonas-Abenteurer – und es nützt Unternehmen. Wenn große Mengen verschiedenster Daten anfallen, ist die automatische Verarbeitung ein entscheidender Wettbewerbsvorteil. In Berlin forschen wir zusätzlich am Corporate Semantic Web – der Anwendung von Semantic Web Technologien in Unternehmen. Dies umfasst die Unterstützung von kollaborativen Arbeitsabläufen und Prozessen, die Kommunikation und das Wissensmanagement durch eine auf semantische Technologien aufbauende IT Infrastruktur eines Unternehmens. Neben den technologischen Aspekten umfasst Corporate Semantic Web auch den pragmatischen Einsatz des Semantic Web in Unternehmen. ■