

Umwelt und Ontologie

Ein Forschungsprojekt fördert IT für die moderne Gesellschaft. Ein Teil des Vorhabens fokussiert auf semantische Technologien bei Geo- und Umweltdatenmanagement

Komplexität darf als das wesentlichste Merkmal unserer Gesellschaft angesehen werden. Sprach man in den 1990er Jahren noch von der Informationsgesellschaft, befinden wir uns heute am Übergang zur Wissensgesellschaft, in der Information nicht mehr technologisch, sondern als Grundlage von Beurteilung und Verstehen von Inhalten begriffen wird. Vor diesem Hintergrund ist auch das 2007 gestartete, rund 200 Millionen Euro teure Forschungsprogramm „Theseus“ zu sehen. Mit ihm sollen Basistechnologien und technische Standards für ein künftiges „Internet der Dienste“ entwickelt werden.

„Theseus hat das Ziel, innovative Technologien für ein höherwertiges Informations- und Dienstemanagement im Internet zu entwickeln“, sagt Dr. Andreas Goerdeler, Ministerialrat im Referat Entwicklung konvergenter IKT (Informations- und Telekommunikationstechnik) beim **Bundesministerium für Wirtschaft und Technologie** (BMWi). Das heißt, die Vernetzung der Alltagsgegenstände über das Internet (Stichwort intelligenter Kühlschrank) wird um Dienstleistungen und interaktive Angebote erweitert. Betroffen sind unter anderem technische Dienste, kooperative Geschäftsmodelle, Informationsdienste oder auch ortsbezogene Dienstleistungen.

Innerhalb des Programms existiert mit „Theseus Mittelstand 2009“ ein Unterprogramm, das zwölf zusätzliche Anwendungsprojekte umfasst. 65 kleine und mittelständische Unternehmen hatten dazu Forschungsprojekte eingereicht, die von einer Jury bewertet wurden. Eines dieser Projekte nennt sich „Hippolytos“ und wurde von dem Karlsruher Unternehmen **Disy** gewonnen, einem Spezialisten für datenbankgestützte Anwendungen im Geo- und Umweltbereich. Das **Fraunhofer-Insti-**

tut für Informations- und Datenverarbeitung (IITB) aus Karlsruhe ist als Unterauftragnehmer an Hippolytos beteiligt. Das Projekt hat eine Laufzeit von zweieinhalb Jahren. Das Teilprojekt beschäftigt sich mit Diensten im Umwelt- und Geodatenbereich. Auch hier geht es um die Frage, wie aus Daten Wissen erzeugt werden kann. Im Zentrum stehen sogenannte semantische Softwaretechnologien.

In der Softwaretechnik bedeutet Semantik soviel wie ein Instrument zur Reduktion von Komplexität mithilfe bestimmter Schemata. Semantische Technologien stehen daher in der Tradition der Forschung über künstliche Intelligenz. Auch hier werden Datenstrukturen auf ein höheres Abstraktionsniveau gestellt. Was theoretisch abstrakt klingt, ist in der Praxis mitunter einfach verständlich: Zeigen die Daten ein Bodensegment mit hohem Wasseranteil, großer Höhendifferenz und stark heterogener Bewachsung ist dies mit hoher Wahrscheinlichkeit ein Uferstreifen, also ein „geographisches Objekt“ auf höherer Abstraktionsniveau.

Projektleiter Wassilios Kazakos sieht darin die Chance für Geo- und Umweltthemen, „die häufig gepflegte Nischenexistenz zu verlassen und ein wichtiger Motor zur Entwicklung des Web 3.0 werden.“ Web 3.0, ein Synonym für das semantische Web, meint, dass bereits der Computer die Bedeutung von Informationen erkennen kann und so Wissen generiert.

Heutzutage liegen diese Daten meist noch in verschiedenen Fachsystemen verborgen. Dies erschwert die Beantwortung selbst häufig gestellter Fragen, zum Beispiel, wie sich die

Luftverschmutzung in verschiedenen Stadtteilen im Laufe der letzten Jahre entwickelt hat. Disy arbeitet daran, die Fülle von Umweltdaten jedem in einer leicht verständlichen Form zur Verfügung zu stellen.

Das Problem dabei ist, dass gerade Umweltdaten aus Sicht der Semantik sehr komplex sind. Heute sind noch kaum einheitliche Metadaten-Informationssysteme im Umweltbereich etabliert. Die Semantik, also die fachliche Bedeutung der Umweltdaten, aber auch die Syntax, sprich die Datenstrukturen, sind nicht einheitlich geprägt. Dies ist gerade für die Zusammenführung von Daten aus heterogenen Quellen von hoher Bedeutung, nicht nur bei der Integration von Geodaten selbst,

sondern auch bei der Verschneidung mit Sachdaten. „Gerade Umweltdaten haben zwar fast immer einen geographischen Bezug, bestehen aber im Wesentlichen aus Sachinformationen“, sagt Kazakos mit besonderem Blick auf die Inspire-Richtlinie, die die Verbindung von Umweltdaten anvisiert.

Umweltdaten haben viele Dimensionen, eine ist die geographische.

Im Projekt werden also Werkzeuge und Methoden entwickelt, mit denen strukturierte und unstrukturierte Umwelt- und Geoinformationen zusammengeführt, gemeinsam recherchiert und genutzt werden können. Disy spricht in diesem Zusammenhang von Ontologieverwaltung. Am Ende des Projekts soll ein Prototyp eines semantischen Umweltportals stehen, mit dem die intelligente Suche und Navigation in Umwelt- und Geodaten möglich ist. Die geographischen Komponenten der Umweltdaten sind dabei ein Anhaltspunkt, nachdem man die Daten visualisieren und analysieren kann – und so im Sinne der Wissensgesellschaft schneller fachliche Bedeutungszusammenhänge verstehen kann.

Solche Anforderungen verfolgen universelle Prinzipien, das erkennt man an der Einordnung von Hippolytos in das Theseus-Anwendungsszenario „Ordo“. Vor dem Hintergrund der drohenden Zunahmen an Unsicherheit und der kommunikativen Überlastung der Wissensgesellschaft sollen dabei Werkzeuge entwickelt werden, die Informationen automatisch ordnen und die Suche vereinfachen. Ordo hat allerdings einen Fokus auf Pharmaanwendungen und die Entwicklung neuer Medikamente, bei der ebenfalls große Datenmengen entstehen. (sg)

www.theseus-programm.de



Schützenswertes Landschaftsbild: Um die verfügbaren Umweltdaten intelligent auszuwerten, können sogenannte semantische Softwaretechnologien eingesetzt werden.