

Question Answering: Das Retrieval der Zukunft?

Der Artikel geht der Frage nach, ob und inwieweit Informations- und Recherchesysteme von der Technologie natürlichsprachlicher Frage-Antwortsysteme, so genannter Question Answering-Systeme, profitieren können. Nach einer allgemeinen Einführung in die Zielsetzung und die historische Entwicklung dieses Sonderzweigs der maschinellen Sprachverarbeitung werden dessen Abgrenzung von herkömmlichen Retrieval- und Extraktionsverfahren erläutert und die besondere Struktur von Question Answering-Systemen sowie einzelne Evaluierungsinitiativen aufgezeichnet. Zudem werden konkrete Anwendungsfelder im Bibliothekswesen vorgestellt.

To what extent can information and research systems profit from the technology of natural language queries, or so-called question-answering systems? After a general introduction to the goals and the historical development of this special area of machine language processing, the author elucidates how it can be distinguished from traditional information retrieval and extraction procedures and elaborates on the particular structure of question answering systems (QA) as well as various evaluative endeavours. Furthermore, the author presents concrete areas of application in the library world.

Wie werden die Benutzerschnittstellen zukünftiger Auskunftssysteme und Retrievalsysteme aussehen? Werden wir weiterhin einzelne Suchbegriffe eingeben, in der Hoffnung, dass diese ausreichen, um die gewünschte Information zu erhalten? Oder werden wir in der Lage sein, natürlichsprachliche Fragen zu formulieren, die das System eindeutig versteht oder auf die es entsprechende Rückfragen stellt, und werden wir als Resultat dieser Anfrage eine einzige und richtige Antwort zu erhalten?

Diese Fragen sind von großer Relevanz für die Welt der Bibliotheken und Archive, hält man sich vor Augen, dass die zunehmende Digitalisierung der Bibliotheken sowie die Digitalisierungsinitiativen führender Suchmaschinen längerfristig dazu führen werden, dass Recherchen weniger auf das Aufsuchen relevanter Dokumente und deren Ranking beschränkt bleiben, sondern dass gleichsam auch die Suche nach den relevanten Informationsbestandteilen *innerhalb* der Dokumente Aufgabe des Recherche- und Retrievalprozesses sein wird.

Dies verlangt vom Benutzer, dass er sein Informationsbedürfnis genauer spezifiziert, beispielsweise in einer oder mehreren vollständig ausformulierten natürlichsprachlichen Fragen. Gleichzeitig steigen die Anforderungen an die maschinelle Verarbeitung der Benutzereingabe, den Retrievalvorgang und die Präsentation der Antwort, die idealerweise ein neues Aggregat aus unterschiedlichen Textfragmenten darstellt. Stellt man der Suchmaschine Ask (www.ask.com) die Frage »Where was Rousseau born?« (siehe Abbildung 1)

erhält man nicht nur die Antwort »Jean-Jacques Rousseau was born in Geneva, Switzerland.«, sondern gleichfalls ein Bild Rousseaus und Hinweise auf weiterführende Informationsquellen.

Question Answering, ein Sonderzweig der maschinellen Sprachverarbeitung, der gelegentlich auch als Answer Mining bezeichnet wird, beschäftigt sich mit dieser Aufgabe und weitet die Suche nach dem relevanten Informationsträger durch eine Suche *im* Informationsträger aus. Es geht folglich um die Erstellung offener Frage-Antwortsysteme, die in der Lage sind, Fragen jeglicher Art mit genau *einer* präzisen und ausreichend kurzen Antwort zu beantworten, die aus einer unübersichtlich großen, elektronischen Datensammlung extrahiert wird.

»You ought to be able to type in »Why is the sky blue?«, and get an answer – not just a bunch of people who want to sell you blue paint.« (Bender 2004). Mit diesen Worten wies niemand anderes als Bill Gates seinen Entwicklern den Weg und postulierte im gleichen Atemzug die Lösung dieser Aufgabe als eine der großen Grenzen, die die Informatik noch zu überschreiten hat. Dabei sind bei genauerem Hinsehen nicht allein die allgemeine Informatik, sondern vielmehr ihr verwandte oder beigeordnete Disziplinen wie etwa die Computerlinguistik oder die Künstliche Intelligenz (KI) an der Beschreitung dieses Wegs beteiligt. Gleichzeitig messen auch die Marktführer der Suchmaschinenteknologie dem Bereich des Question Answering immer größere Bedeutung bei, zumal das Benutzerinteresse an Online-Enzyklopädien, die eine Hauptquelle bei der Beantwortung natürlichsprachlicher Fragen sind, immer weiter zunimmt. Da es sich hierbei auch um einen Bereich der Wissensvermittlung handelt, der traditionell den Bibliotheken vorbehalten war, soll zunächst vorgestellt werden, welches Potenzial diese Technologie darstellt und in welchen Bereichen der Informationsvermittlung reelle Einsatzmöglichkeiten bestehen.



René Schneider

offene Frage-Antwort-Systeme



Jean-Jacques Rousseau was born in Geneva, Switzerland

Mostly self-educated in Switzerland, Jean-Jacques Rousseau ended up in Paris, France in the 1740s and became acquainted with Voltaire and Denis Diderot. Rousseau published *Discourse on the Origin and Foundations of Inequality* in 1754, arguing that the natural... [More »](#)

Abb. 1: Where was Rousseau born?
(Quelle: Ask.com 24.10.2006)