

Computer-Linguistik in der Industrie

Anwendungsbereiche

Systeme und Technologien

Berufsbilder

Praxisbeispiel

Kader Diagne

Die Terminologie im Kontext der Präsentation

Viele Bezeichnungen für die gleiche Sache?

Computational Linguistics, Human Language Technology, Language Engineering, Language Technology (***Sprachtechnologie***), Natural Language Processing, Natural Language Understanding, ...

Worum geht's eigentlich?

Es geht darum, den Computer dazu zu bringen, etwas Nützliches mit natürlicher Sprache zu tun, egal ob gesprochen oder geschrieben; diese wird dabei automatisch verarbeitet (Analyse und Generierung), und zwar in einer Art und Weise, die eine gewisse Sprachkompetenz voraussetzt; u.a. Wissen über die Funktion von Sprache im Allgemeinen bzw. bestimmter ihrer Teilbereiche.

Die Anwendungsgebiete sind so vielfältig, dass sie nur partiell aufgezählt und selektiv betrachtet werden können

Der Schwerpunktmässige Anwendungskontext ist Informations- und Kommunikationsmanagement

Im Unternehmenskontext geht es primär um die Optimierung der Geschäftsabläufe und die Verbesserung der Entscheidungsfindung

Im Internet ist Information Retrieval insb. Suche die dominierende Anwendung

Eine Reihe von Basissystemen bereiten den Weg für viele Anwendungsmöglichkeiten

- Textverarbeitung (Worttrennung, Schreibfehlerkorrektur, Grammatik-, und Stilüberprüfung)
- OCR (Optical Character Recognition)
- Textanalyse (Kategorisierung, Informationsextraktion inkl. Question Answering)
- Textgenerierung
- Maschinelle Übersetzung
- Spracherkennung (Diktiersysteme, Sprachsteuerung)
- TTS

Im Unternehmensbereich sind Anwendungsmöglichkeiten der Sprachtechnologie allgegenwärtig, da 80%+ der Daten unstrukturiert

- DM (Dokumenten-Management)
- CM (Content-Management)
- ECM (Enterprise Content Management)
- CRM (eCRM, e-Services, Call-Center)
- BI (Text Mining)
- KM (Knowledge Management)
- Enterprise Search
- PLM (Product Life Cycle Management)

Die Verarbeitung sehr großer Datenmengen, hohe Performanz und Anpassbarkeit, Skalierbarkeit und Robustheit bestimmten den Technologie-Schwerpunkt und dienen zugleich als Technologie-Filter

- Effiziente Algorithmen und Datenstrukturen
- Datenkomprimierung
- Hybride Ansätze (Flache linguistische Analyse, statistische Verfahren / Maschinelles Lernen)
- Dedizierte Dateisysteme und intelligentes Speichern
- Verteilte Verarbeitung
- Erstklassiges Software-Engineering

Industriell erfolgreiche Computer-Linguistik setzt fortgeschrittenes Software-Engineering und einen stärkeren Anwendungsfokus voraus

Entsprechend gestalten sich der typische Berufseinstieg in der Industrie

- Forschung & Entwicklung
- Produkt- und Anwendungsentwicklung
- Pre-Sales Consulting
- Produktmanagement

Eine Disziplin im Trend und mit großen Zukunftschancen

Wie kommt die Technologie vom Labor auf den Markt

Beispiel XtraMind Technologies GmbH

Xtramind Lösungsanbieter für effizientes Kommunikations- und Informationsmanagement auf der Basis von KI-Technik und Sprachechnologie; insbesondere Komplettsysteme

- Zur Optimierung des email-basierten Kundenservice
- Zur Erleichterung der Recherche (Personalisierte Informationsbeschaffung und -verwaltung)
- Und Bausteine für Textanalyse und Informationsmanagement

Danke für Ihre Aufmerksamkeit

Nun ... Fragen und Antworten / Diskussion