

Multilinguale Sprachtechnologie und Intelligente Informationsextraktion

Analyse und Auswertung von Kundeninteraktionen

PD Dr. Günter Neumann

Forschungsbereich Sprachtechnologie

DFKI – Deutsches Forschungszentrum für Künstliche Intelligenz

Saarbrücken

Die DFKI GmbH mit Standorten in Kaiserslautern, Saarbrücken, Bremen, Berlin und Osnabrück



Standort
Saarbrücken

DFKI GmbH
Campus D3 2
Stuhlsatzenhausweg 3
D-66123 Saarbrücken

Tel.: 0681 85775 5252
Fax.: 0681 85775 5341
email: info@dfki.de
www.dfki.de



Standort
Kaiserslautern
(Firmensitz)

DFKI GmbH
Trippstadter Straße 122
D-67663 Kaiserslautern

Tel.: 0631 205 75-0
Fax.: 0631 205 75-5030
email: info@dfki.de
www.dfki.de



Standort
Bremen

DFKI GmbH
Robert-Hooke-Straße 5
D-28359 Bremen

Tel.: 0421 178 45-4100
Fax.: 0421 178 45-4150
email: info@dfki.de
www.dfki.de



Projektbüro
Berlin

DFKI GmbH
Alt-Moabit 91c
D-10559 Berlin

Tel.: 030 23895 1800
Fax.: 030 23895 1810
email: info@dfki.de
www.dfki.de



Robotics
Innovation Center
Außenstelle
Osnabrück

DFKI GmbH
AVZ-Gebäude
Albrechtstraße 28
49076 Osnabrück

Tel.: 0541 969 3956
Fax.: 0541 969 2799
email: info@dfki.de
www.dfki.de

Die Forschungsbereiche und -gruppen des DFKI

Innovative
Fabriksysteme

Eingebettete
Intelligenz

Agenten
und
Simulierte
Realität

Cyber-physische
Systeme

Institut für
Wirtschafts-
informatik
im DFKI



Robotics
Innovation Center

Intelligente
Benutzer-
schnittstellen

Sprach-
technologie

Erweiterte
Realität

Wissens-
management

Sprachtechnologie am DFKI

FB Sprachtechnologie:

Leitung: Prof. Josef van Genabith & Prof. Hans Uszkoreit

R&D Staff: 30 MA
 Projekte: DE, EU, Industrie
 Umsatz: ~ 3.5 MEUR
 Standorte: Saarbrücken & Berlin

F&E Schwerpunkte:

Multilinguale Technologien
 Textbasiertes Informationsmanagement
 Mensch-Roboter-Interaktion

The screenshot shows the DFKI website homepage. At the top, it reads 'Deutsches Forschungszentrum für Künstliche Intelligenz' and 'German Research Center for Artificial Intelligence'. Below this is a navigation menu with links for 'STARTSEITE', 'AKTUELLES', 'OVER US', 'FORSCHUNG', 'BEMERKUNGSANZEIGEN', 'LIVING LABS', 'KOMPETENZZENTREN', 'TRANSFERZENTRUM', 'PRESSE', 'KONTAKT', and social media links for 'DFKI @ TWITTER', 'DFKI @ YOUTUBE', and 'DFKI @ FACEBOOK'. The main content area features a large image of a robot and a person, followed by the heading 'Intelligente Lösungen für die Wissensgesellschaft'. Below this, there is a paragraph about DFKI's location and research focus, followed by a list of current projects and their leaders. The bottom section highlights various research clusters and centers, including 'Software-Cluster', 'M-C-I', 'ert', 'software campus', 'academy cube', and 'Smart Data Innovation Lab'. A footer section mentions DFKI's cooperation with other research centers.

Deutsches Forschungszentrum für Künstliche Intelligenz
 German Research Center for Artificial Intelligence

STARTSEITE
 AKTUELLES
 OVER US
 FORSCHUNG
 BEMERKUNGSANZEIGEN
 LIVING LABS
 KOMPETENZZENTREN
 TRANSFERZENTRUM
 PRESSE
 KONTAKT
 DFKI @ TWITTER
 DFKI @ YOUTUBE
 DFKI @ FACEBOOK
 IMPRESSUM

2014-04-28 Innovation
 Verhängnisvoller Fehler:
 Torwart als Feldspieler

2014-04-27 Uni-Golubair
 German Open 2014: Bremer
 Roboterfußballer verteidigen
 Meistertitel

www.dfkj.de

Intelligente Lösungen für die Wissensgesellschaft

Das Deutsche Forschungszentrum für Künstliche Intelligenz (DFKI) mit den Standorten Kaiserslautern, Saarbrücken, Bremen (mit Außenstelle Osnabrück) und einem Projektbüro in Berlin ist auf dem Gebiet innovativer Softwaretechnologien die führende Forschungseinrichtung in Deutschland. In der internationalen Wissenschaftswelt zählt das DFKI zu den wichtigsten "Centers of Excellence" und ist derzeit, gemessen an Mitarbeiterzahl und Drittmittelvolumen, das weltweit größte Forschungszentrum auf dem Gebiet der Künstlichen Intelligenz und deren Anwendungen.

Die Geschäftsführung der DFKI GmbH bilden Prof. Dr. Dr. h.c. mult. Wolfgang Wahlster (Technisch-wissenschaftlicher Leiter und Vorsitzender der Geschäftsführung) und Dr. Walter Olthoff (Kaufmännischer Geschäftsführer).

Die aktuell laufenden Projekte werden durchgeführt in den DFKI
 Forschungsbereichen, Forschungsgruppen und Living Labs:

- Wissensmanagement (Prof. Dr. Prof. h.c. Andreas Dengel)
- Cyber-Physical Systems (Prof. Dr. Ralf Drechsler)
- Multilinguale Technologien (Prof. Dr. Josef van Genabith)
- Plattformbasierte Robotereinsparung (Prof. Dr. Joachim Hertzberg)
- Robotics Innovation Center (Prof. Dr. Frank Kirchner)
- Innovative Retail-Laboratory (Prof. Dr. Antonio Krüger)
- Institut für Wirtschaftsinformatik (Prof. Dr. Peter Löss)
- Eingebettete Intelligenz (Prof. Dr. Paul Lukowicz)
- Intelligente Netze (Prof. Dr.-Ing. Hans Schotten)
- Agenten und Simulierte Realität (Prof. Dr. Philipp Susaike)
- Erweiterte Realität (Prof. Dr. Dieter Stricker)
- Sprachtechnologie (Prof. Dr. Hans Uszkoreit)
- Intelligente Benutzerschnittstellen (Prof. Dr. Dr. h.c. mult. Wolfgang Wahlster)
- Innovative Fabricsysteme (Prof. Dr.-Ing. Dr. h.c. Detlef Zühlke)

Das DFKI kooperiert sehr eng mit seinem wissenschaftlichen Umfeld an den jeweiligen Standorten.

Software-Cluster
 M-C-I
 ert
 software campus
 academy cube
 Smart Data Innovation Lab

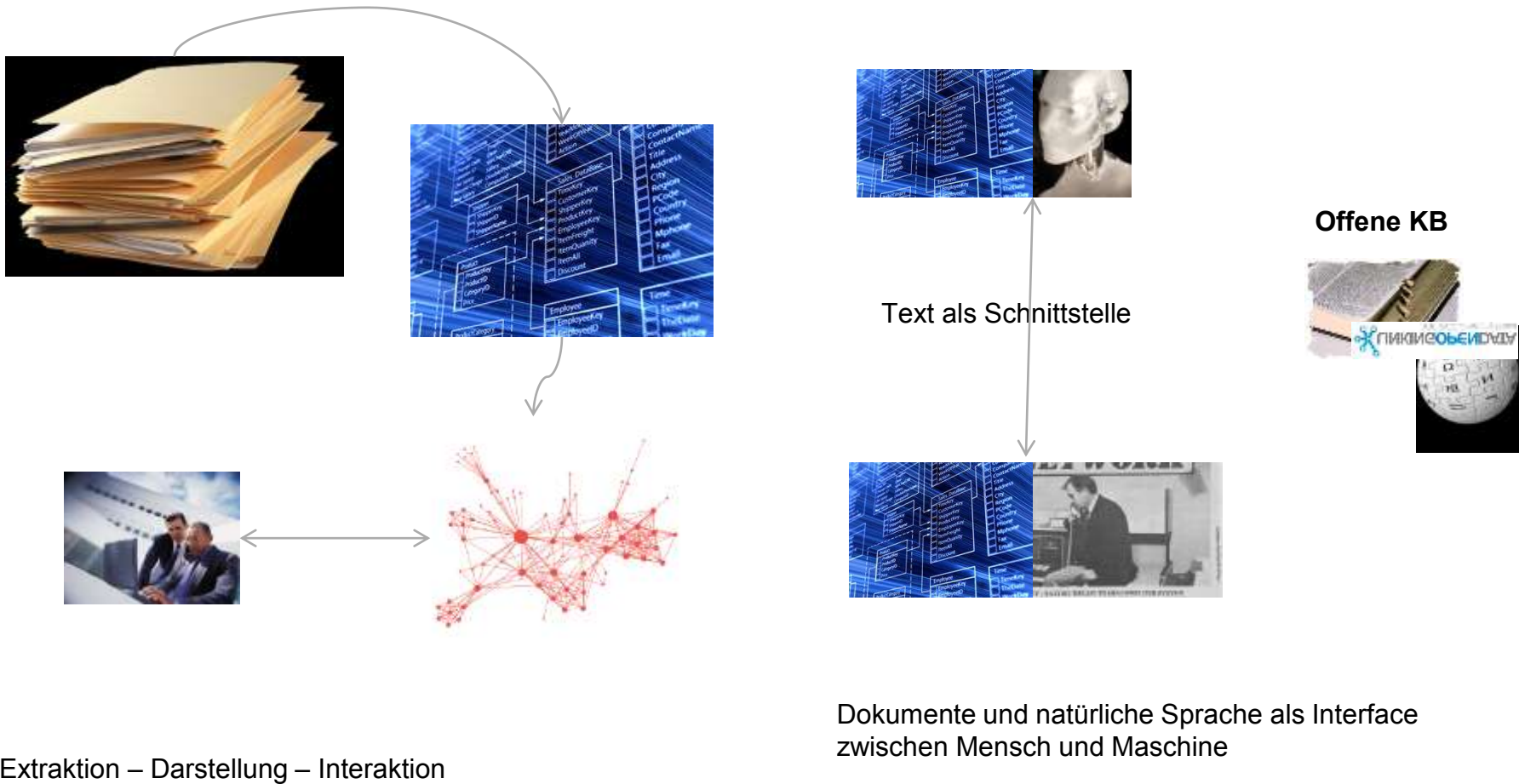
Das DFKI ist als einziges deutsches Informationsinstitut an allen drei Spitzenforschungszentren beteiligt und Mitbegründer des Software Campus, des Academy Cube und des Smart Data Innovation Lab.

Deutsch Englisch
 HANNOVER MESSE
 INDUSTRIE 4.0 - From Vision to Reality
 Prof. Wahlster
 Übersicht Hermes Award 2014 bei der HANNOVER MESSE
 German Open 2014:
 Bremer Roboterfußballer verteidigen Meistertitel
 DFKI News 01/2014
 CeBIT 2014: DFKI zeigt ersten Affenroboter mit beweglicher Wirbelsäule und fühlenden Fingern
 CeBIT
 HANNOVER
 DFKI auf der CeBIT

Multilinguale Sprachtechnologien

- Sprachtechnologien:
Die Umsetzung sprachtheoretischer Grundlagen in praxisrelevante Software-Lösungen und Anwendungen.
- Multilingual:
Sprachtechnologien, die „schnell“ auf verschiedene Sprachen angepasst werden können.
- Crosslingual:
Sprachtechnologien, die verschiedene Sprachen „verbinden“ können.

Schlüssel-Technologie: Verarbeitung unstrukturierter Informationen



Kundeninteraktionsanalyse

Kundendienste

Betriebliche Effizienz







Marketing & Sales

Vertragswesen

Innovationsmanagement



**Sprach und Text
Analytics**

	Phone
	Email
	Phone Survey
	Social Media
	Chat
	Web



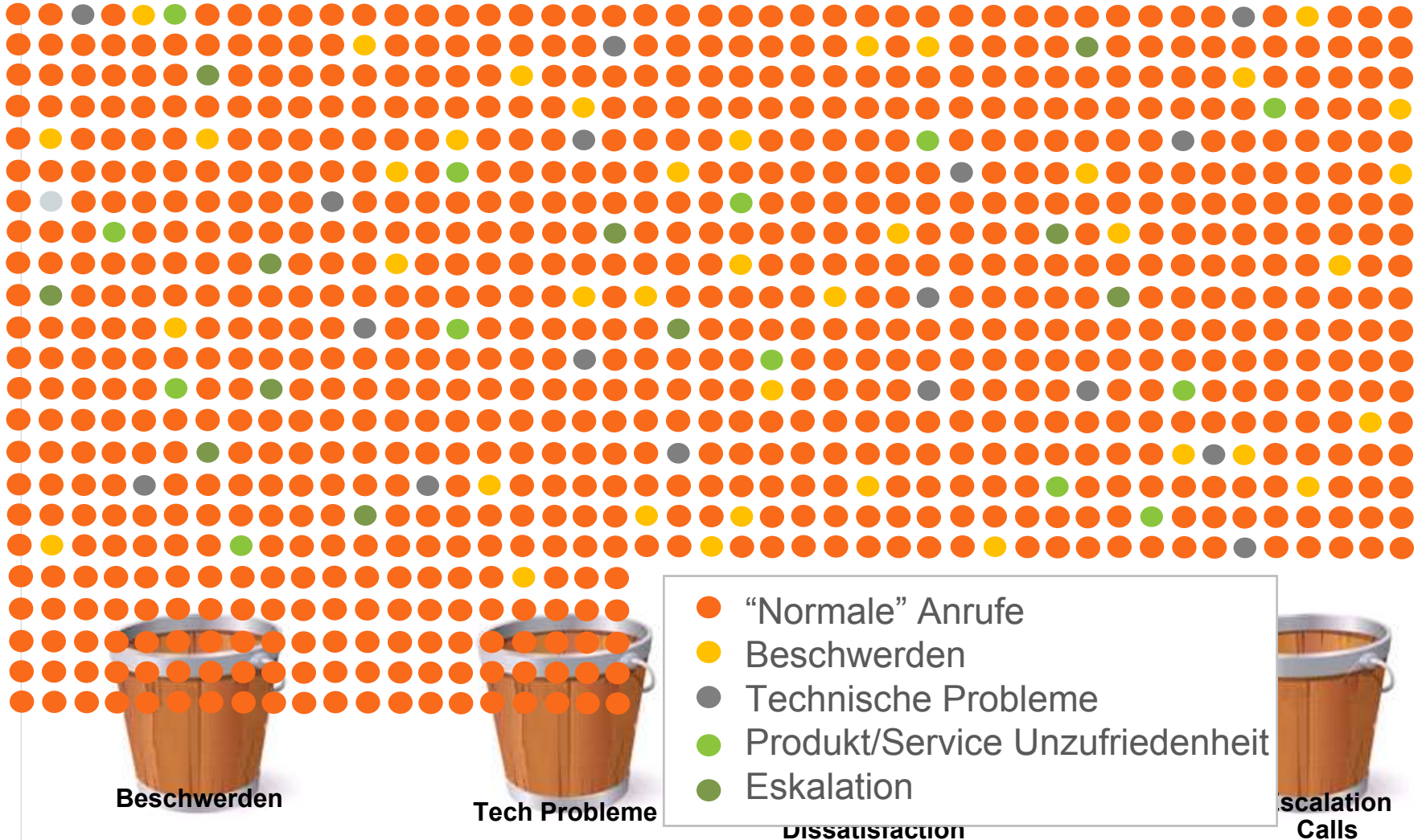
Motivation

- ✦ Weg vom Telefon als zentrales Medium für Kundeninteraktionen hin zu Mehr-Kanal-Kommunikation via Emails, Chats, SMS, Apps für Smartphones, Blogs, Foren, sozialen Netzwerken*
- ✦ Interaktive Analytics als notwendiges Werkzeug, das Unternehmen dabei hilft, Einblicke und Erkenntnisse von einer großen Anzahl an Kunden zu erlangen und diese für Entscheidungen im Geschäftsprozess zu nutzen.



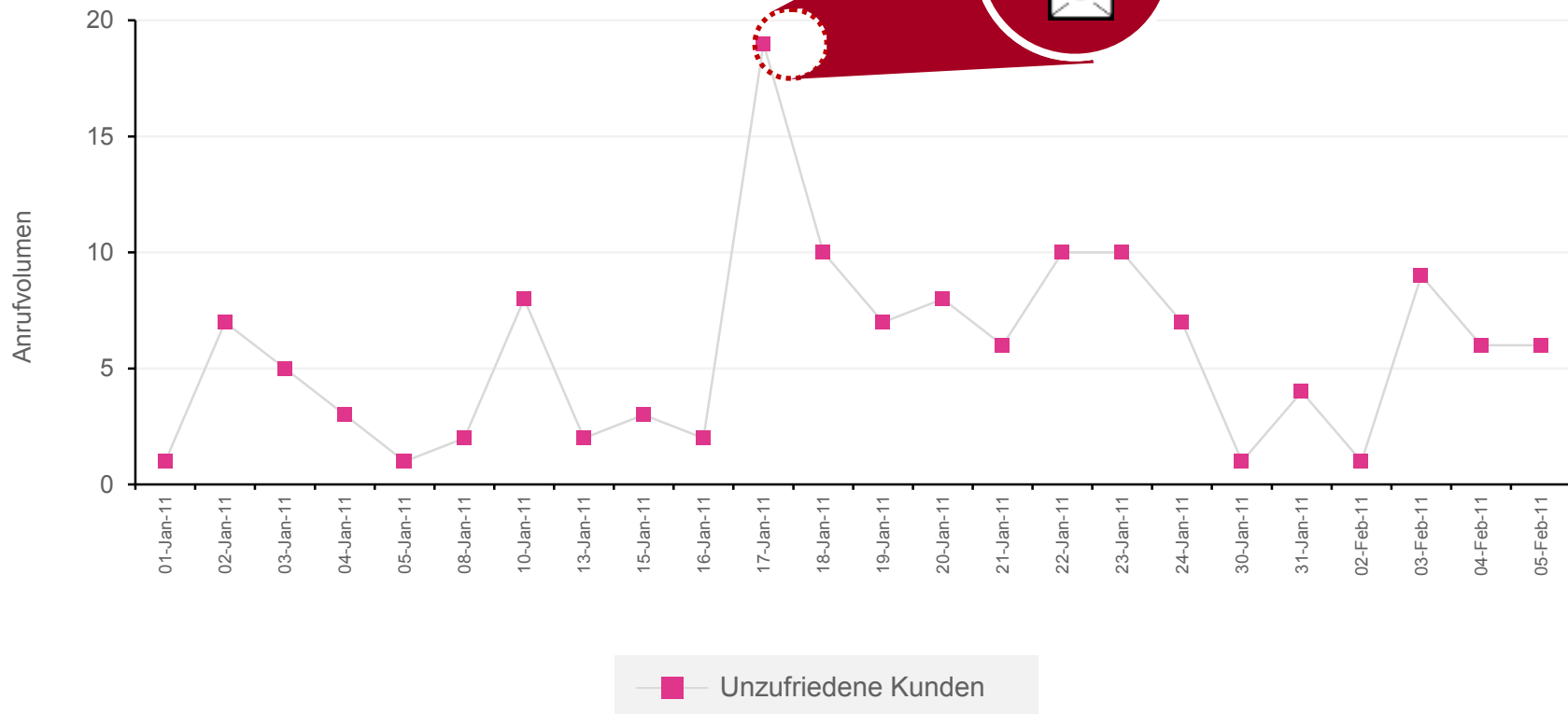
* Vor 2009: 80% vs. 20%, Nach 2009: 44% vs. 56% (vgl. Dimitrova et al. , 2011, S.16)

Kategorisierung von Interaktionen - Relevante Anrufe bestimmen



Überwachung von Trends und Automatischer Alarm bei zu großer Unzufriedenheit

Anzahl Anrufe in Periode 05 to Feb 05



Kunden-Interaktions-Analyse - Technologie-Stand Heute und Morgen

Heute

- ✦ Kategorisierung von Interaktionen – Schlüsselwort-basierte Methoden
- ✦ Nachteil: manuelle Überwachung – die spezifischen Problemfälle müssen vorab konfiguriert werden



Morgen

- ✦ Lösungen auf der Basis von textueller semantischer Inferenz
- ✦ Wichtig: Ebene von Aussagen (nicht Keywords)
- ✦ Vorteil: unüberwachtes Aufdecken von Problemfällen und Visualisierung von Verbindungen zw. Interaktionen durch semantische (entailment) Relationen



Ganz kurz – Was ist Textuelle Inferenz ?

- „Eine Hypothese H folgt aus einem Text T, wenn ein Leser von T mit hoher Wahrscheinlichkeit annehmen würde, dass H wahr ist“ („anwendungsorientiertes“ logisches Schliessen)
- Textuelles Schliessen umfasst auch Folgerungsbeziehungen, die nicht zwingend sind, aber die ein Mensch ziehen würde, weil sie sehr plausibel sind („common sense reasoning“). Beispiel:
 - Text: Durch den Klimawandel schmilzt immer mehr Eis in der Arktis.
 - Hypothese: Die Arktis wird wärmer.
- Angenommen T stimmt, dann ist H außerordentlich plausibel. Dass es Umstände geben könnte, unter denen T wahr und F falsch ist (wenn z.B. das Eis durch Sonneneinstrahlung schmilzt, das Wasser aber im Durchschnitt kälter wird), spielt keine Rolle.

Anwendungsnahes Forschungsprojekt EXCITEMENT



EXCITEMENT

EXploring Customer Interactions through Textual EntailMENT

MEMBERS LOGIN

Search...

GO

HOME

ABOUT

NEWS AND EVENTS

PUBLICATIONS

PARTNERS

RESULTS

LINKS

SAB

CONTACT US

EXCITEMENT

EXploring Customer Interactions through
Textual EntailMENT

It's interesting what we care

About the EXCITEMENT project

There are two interleaved high-level goals for this project. The first is to set up, for the first time, a generic architecture and a comprehensive implementation for a multilingual textual inference platform and to make it available to the scientific and technological communities.

The second goal of the project is to develop a new generation of inference-based industrial text exploration applications for customer interactions, which will enable businesses to better analyze and make sense of their diverse and often unpredicted client content. These goals will be achieved for three languages – English, German and Italian, and for three customer interaction channels – speech (transcriptions), email and social media.



Recent News & Events

The EXCITEMENT project will be present at the [HLT Projects Village](#) which will be held during LREC 2014, 26-31 May 2014, Reykjavik, Iceland.

In particular, the EXCITEMENT booth will present the [EXCITEMENT Open Platform \(EOP\)](#).

26 March 2014 - Release 1.1.1 of the EXCITEMENT Open Platform (EOP) is available at the following URL:

<http://hlfbk.github.io/Excitement-Open-Platform/>

The EOP is a generic multi-lingual platform for textual inference made available to the scientific

Partners



OMQ[®]

Almawave



UNIVERSITÄT
HEIDELBERG
Zukunft. Seit 1386





Fondazione
Enrico Jucker

NICE[®]

Useful Links

Die zentralen Ziele von EXCITEMENT

- Die erste open-source Plattform für Textuelle Inferenz schaffen 
 - <http://hlfbk.github.io/Excitement-Open-Platform/>
- Multilinguale Textuelle Inferenz realisieren 
 - Deutsch, Englisch, Italienisch, ...
- Verschiedene Use-Cases realisieren auf der Basis neuartigen Datenstrukturen 
 - Entailment-Graphen
 - Validierung durch Integration in Produkte der Industriepartner

Use Case: Kundenabwanderung



★ Kunden-Interaktion 1

I need to [redacted] entdecken von Abwanderung

and instead of that you [redacted] is service.

yes

what am i mean i see [redacted] that they . yes extraktion der Gründe

it was just [redacted] and it was just . [redacted]

★ Kunden-Interaktion 2

mentality is avoiding it sent to me to provide excellent customer service what is it we can do for you to do . and

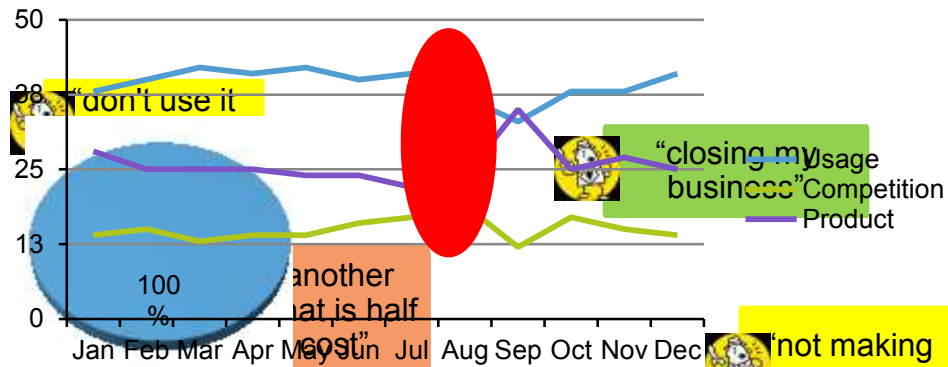
under to [redacted] Entdeckung von Abwanderung

[redacted] s you know m [redacted] extraktion der Gründe

i had see in [redacted] more that's the welcome week locating briefly probably several months . and especially the first month or two i don't [redacted]

gonna [redacted] beneficial to keep it going okay

Analyse der Zentralen Gründe – Semantische Inferenz



“the interface was so complicated”

“not making much use”

“the UI is user friendly”

Competition

“stopping the services”

Product

“expensive service”

“found a cheaper”

Usage

“inst. pr.”

Price

“found a different solution with more features”

Features

Lack of Usage

Relevancy

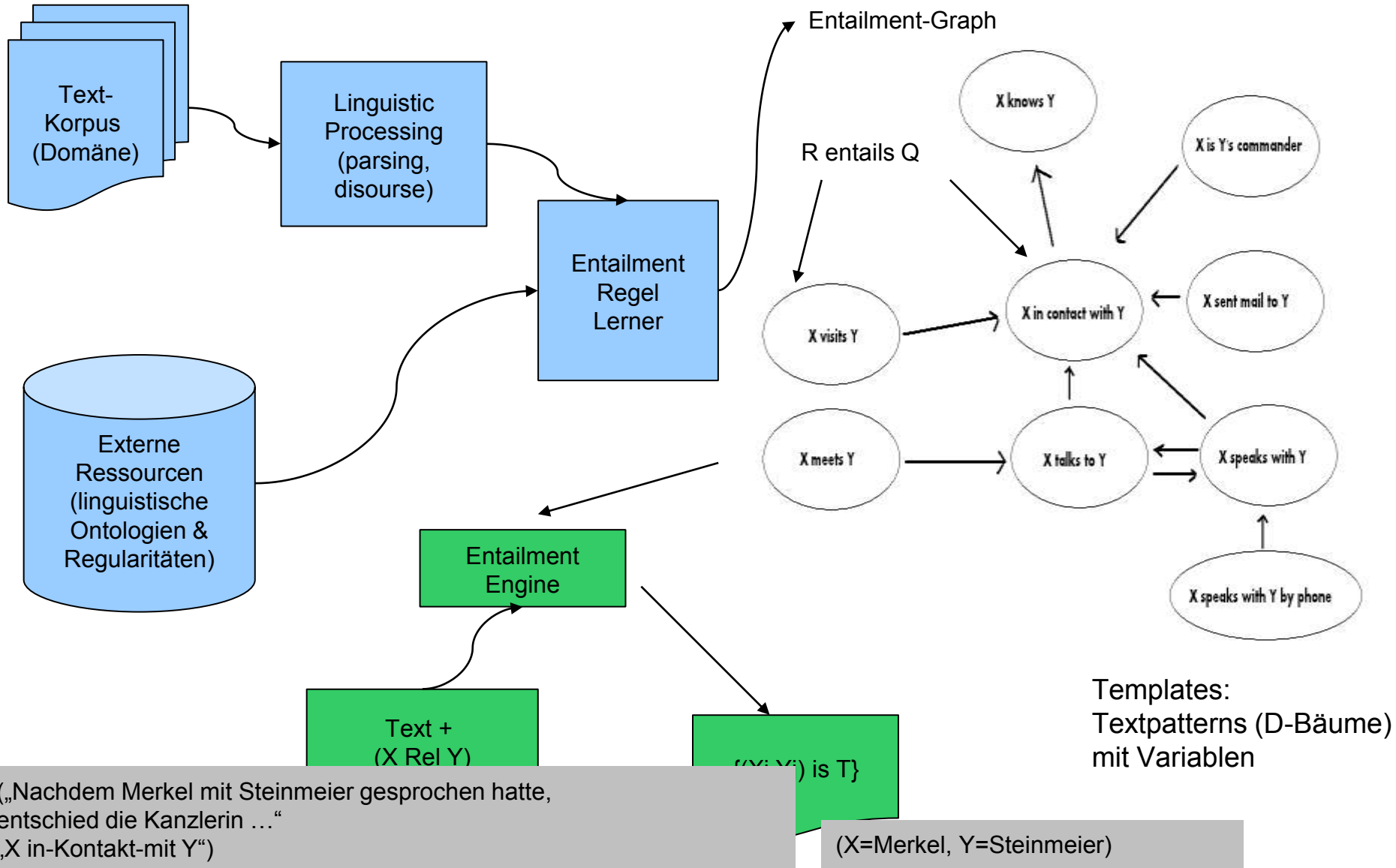
Installation

Interface

Entailment-Graphen

- Relationen zwischen Entitäten, z. B. „X ist in Kontakt mit Y“, „X spricht mit Y“, „X hat besucht Y“, „X ist angekommen in Y“
- Ziel: Relationen zwischen Entitäten automatisch extrahieren und verbinden
- Zwei Schritte:
 - Das Lernen von Taxonomien für Relationen, und
 - Das Inferieren des Auftretens solcher Relationen in unstrukturierten Texten.

Entailment-Graphen mittels Open Information Extraction



DFKI - Methoden und Lösungen

- Multilinguale Kernkomponenten für die Sprachtechnologie (POS-Tagger, Eigennamenerkennung, Morphologie, Syntax-Analyse, semantisches Role-Labeling, ...)
- Multilinguale textuelle semantische Inferenz (Open Source EXCITEMENT Plattform)
- Open Information-Extraktion & Maschinelles Lesen
- Ontologie-basierte Information-Extraktion
- Cross-linguale Frage-Antwort-Systeme & community QA (Yahoo!Research)
- Text Mining (Klassifikation, Sentiment/Opinion Mining)

Kontakt:

Dr. Günter Neumann

neumann@dfki.de

LT-Lab, DFKI GmbH
Stuhlsatzenhausweg 3,
66123 Saarbrücken