

Abbildung von Webformularen auf aktive Ontologien

Masterarbeit

Wasim Said

Betreut von Martin Blersch und Mathias Landhäußer

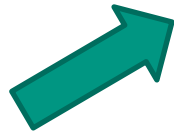
IPD Tichy, Fakultät für Informatik



Motivation

Aktive Ontologie

EASIER



Data Entry Form

Project Name

Billing Code

Date

Notes

A Web Page

http://

Name

Street

City

Country

email

phone no

Registration Details

University:

Institute:

Branch:

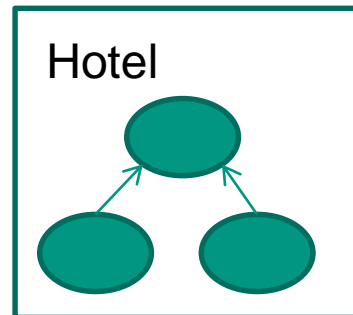
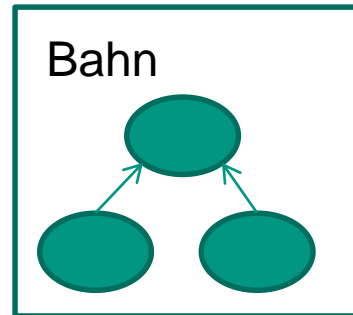
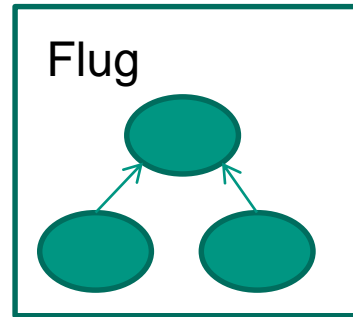
Degree: Pursuing Completed

Average CPI: Upto Th Semester

Experience: Years

Your Website Or Blog:

1 Step 2



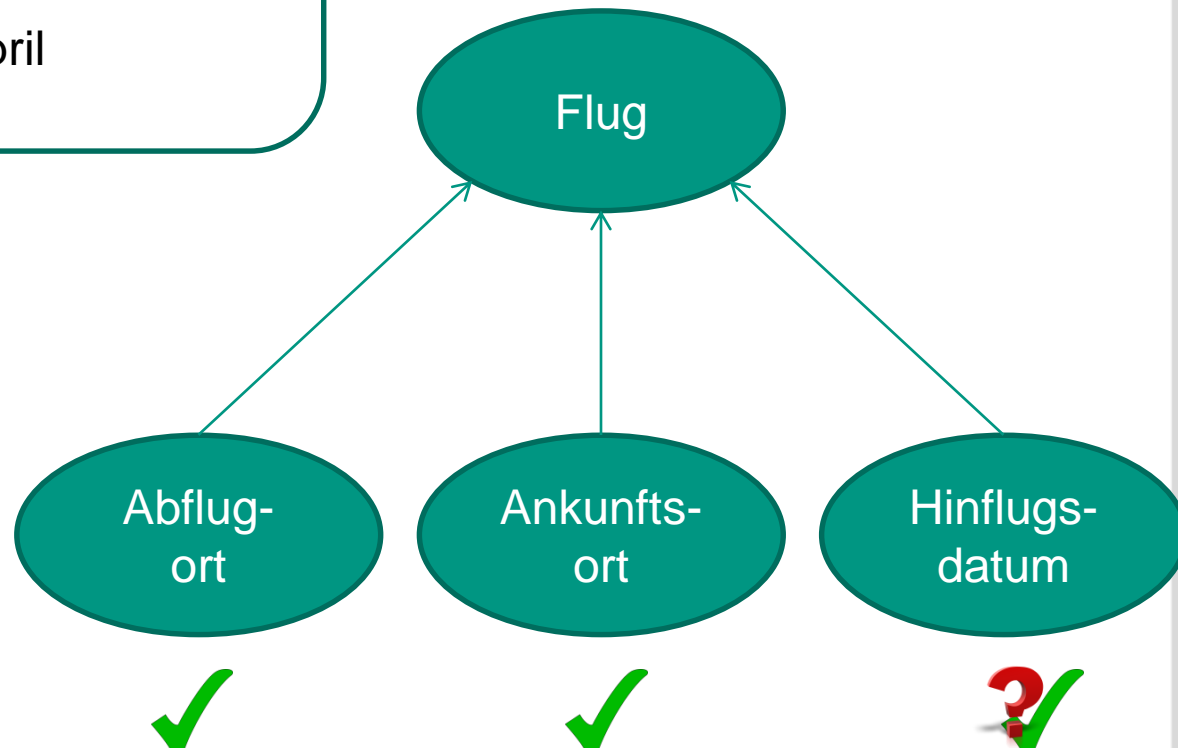
Motivation

Dialog Manager

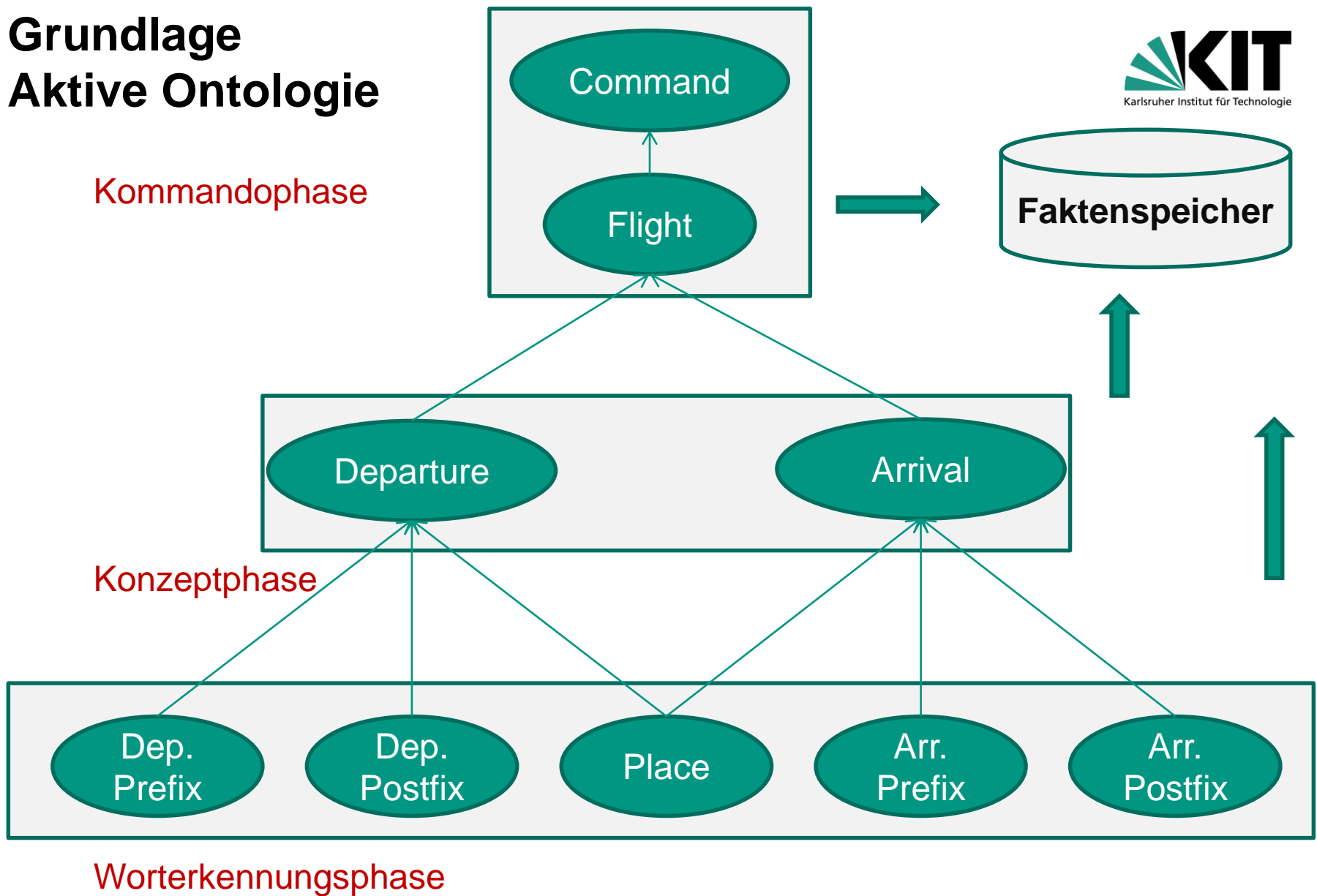
User: Book me a flight from
Karlsruhe to Berlin

System: When do you want to fly?

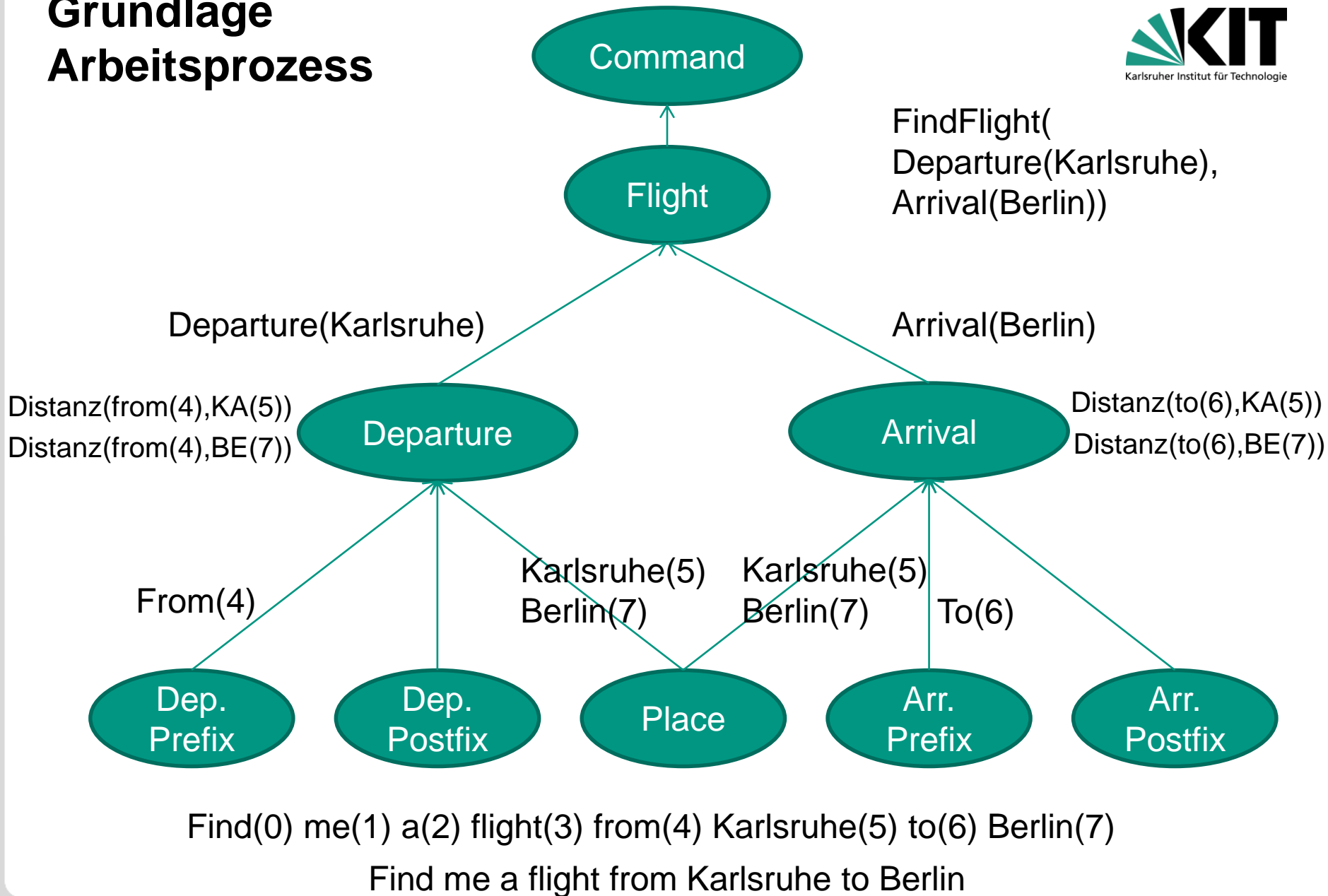
User: On the 15th of April



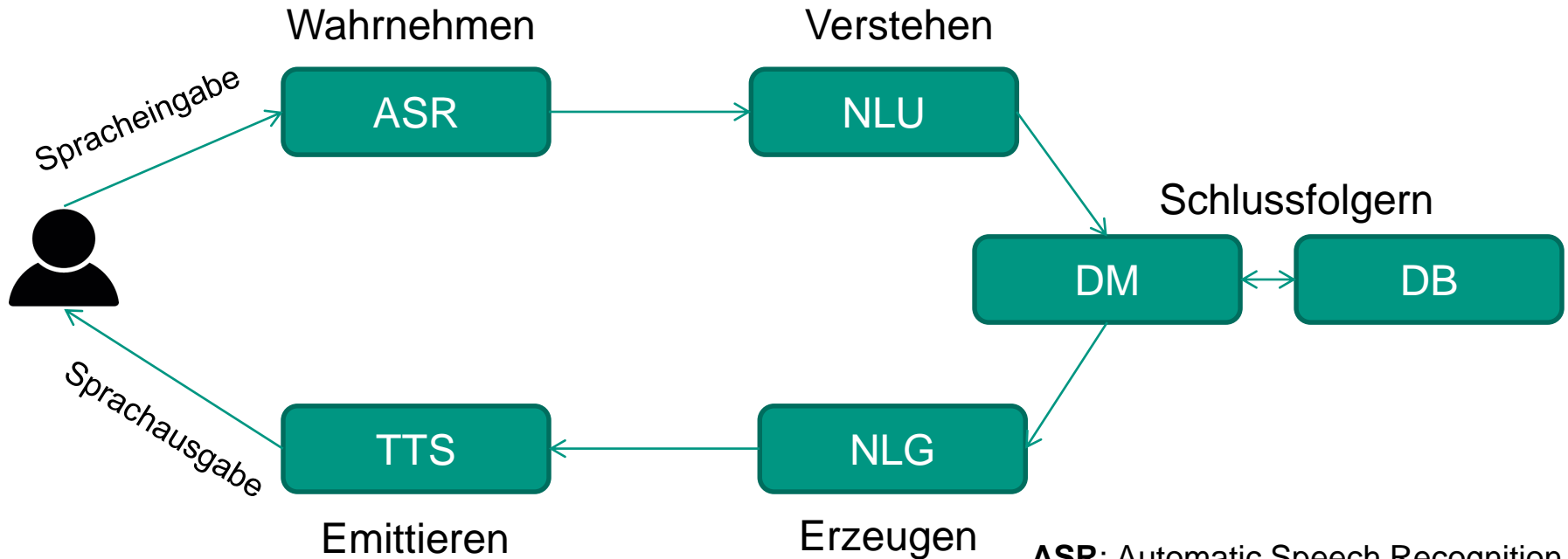
Grundlage Aktive Ontologie



Grundlage Arbeitsprozess



Grundlagen Sprachdialog-Systeme



ASR: Automatic Speech Recognition
NLU: Natural Language Understanding
DM: Dialog Manager
DB: Database
NLG: Natural Language Generation
TTS: Text To Speech

Verwandte Arbeiten

■ Erstellung von Ontologien aus den Webformularen

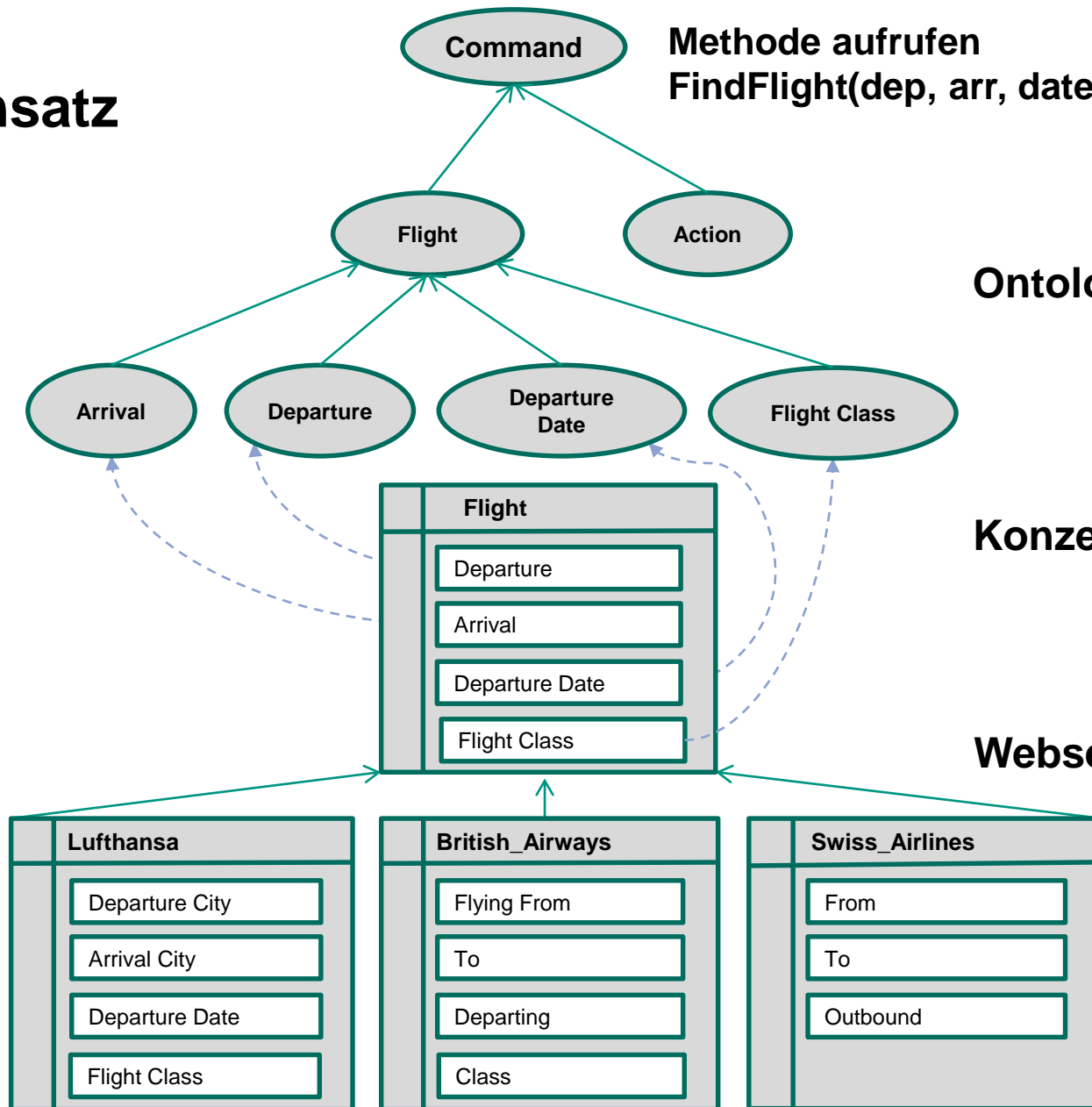
- Faeton: Form analysis and extraction tool for ontology construction [BJNS10]
- Ontobuilder: Fully automatic extraction and consolidation of ontologies from web sources using sequence semantics [RG06]

■ Dialog-Manager

- The Philips automatic train timetable information system [AOSS95]
- Partially observable markov decision processes for spoken dialog systems [WY07]

Ansatz

Methode aufrufen
FindFlight(dep, arr, date)



Ansatz

Konzepte Identifizieren

Airline	Departure	Arrival	Dep. Date	Flight Class
Lufthansa	Departure City	Arrival City	Dep. Date	Flight Class
British Airways	Flying From	To	Departing	Class
Swiss Airlines	From	To	Outbound	---

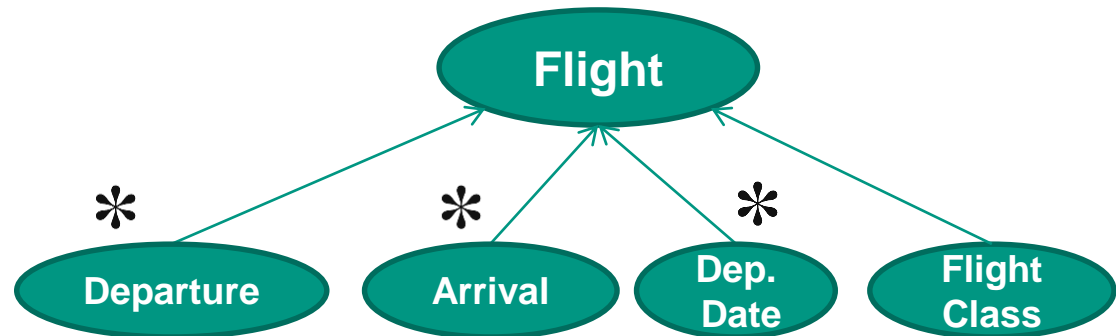
Book a flight * Mandatory

Departure *

Arrival *

Dep. Date *

Flight Class



■ 3 Fragen pro Kategorie

1. Ohne Hilfestellung
2. Bilder von Webformularen
3. Weitere Optionen



■ 40 Probanden

- Sprachniveau (20 Muttersprache, 20 Zweitsprache)
- Onlinebuchungserfahrung (24 mit Erfahrung, 16 ohne Erfahrung)
- Alter (von 15 bis 58)
- Bildungsniveau

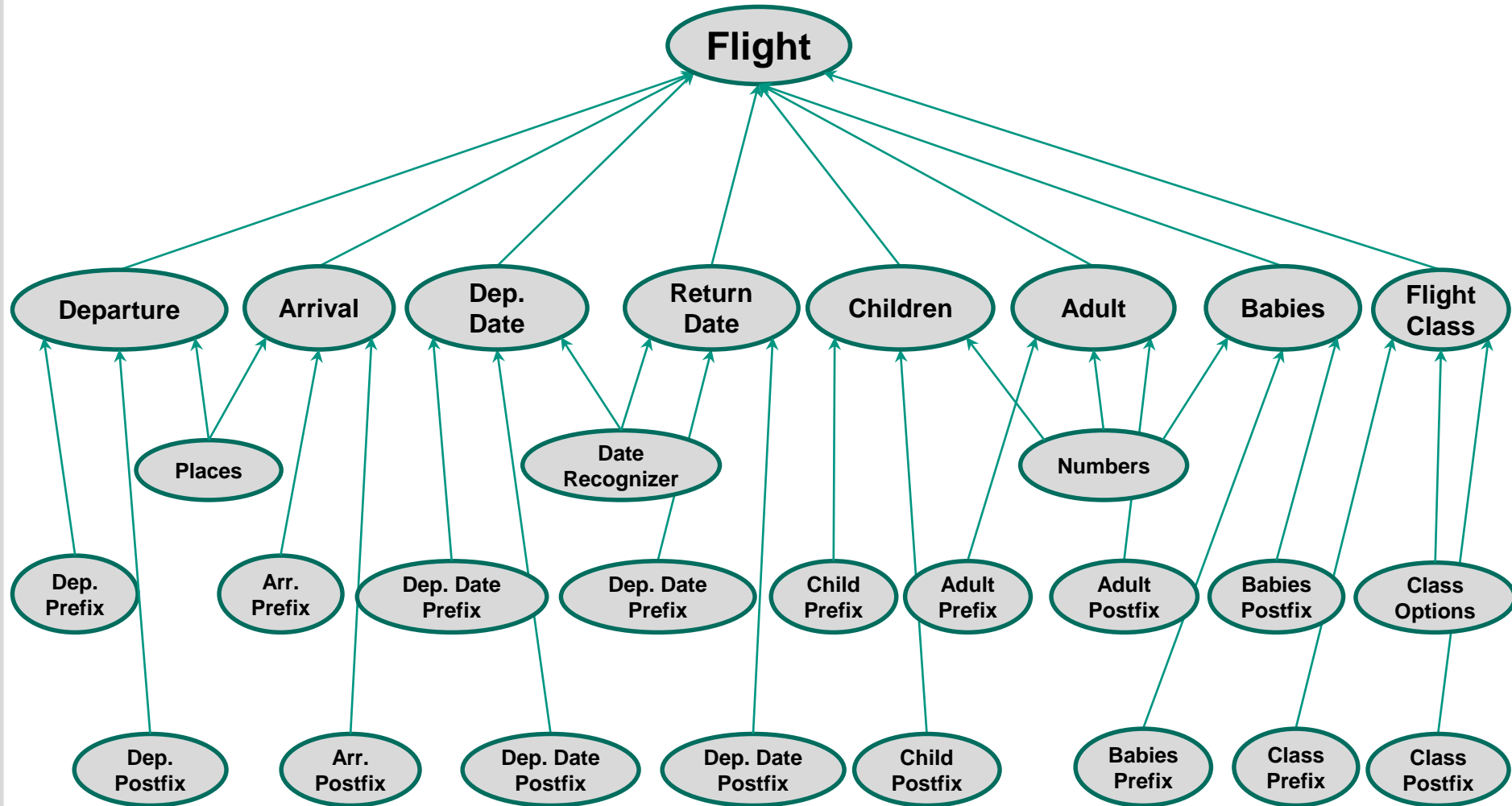
■ 50% für die Entwicklung, 50% für die Evaluation

Ansatz

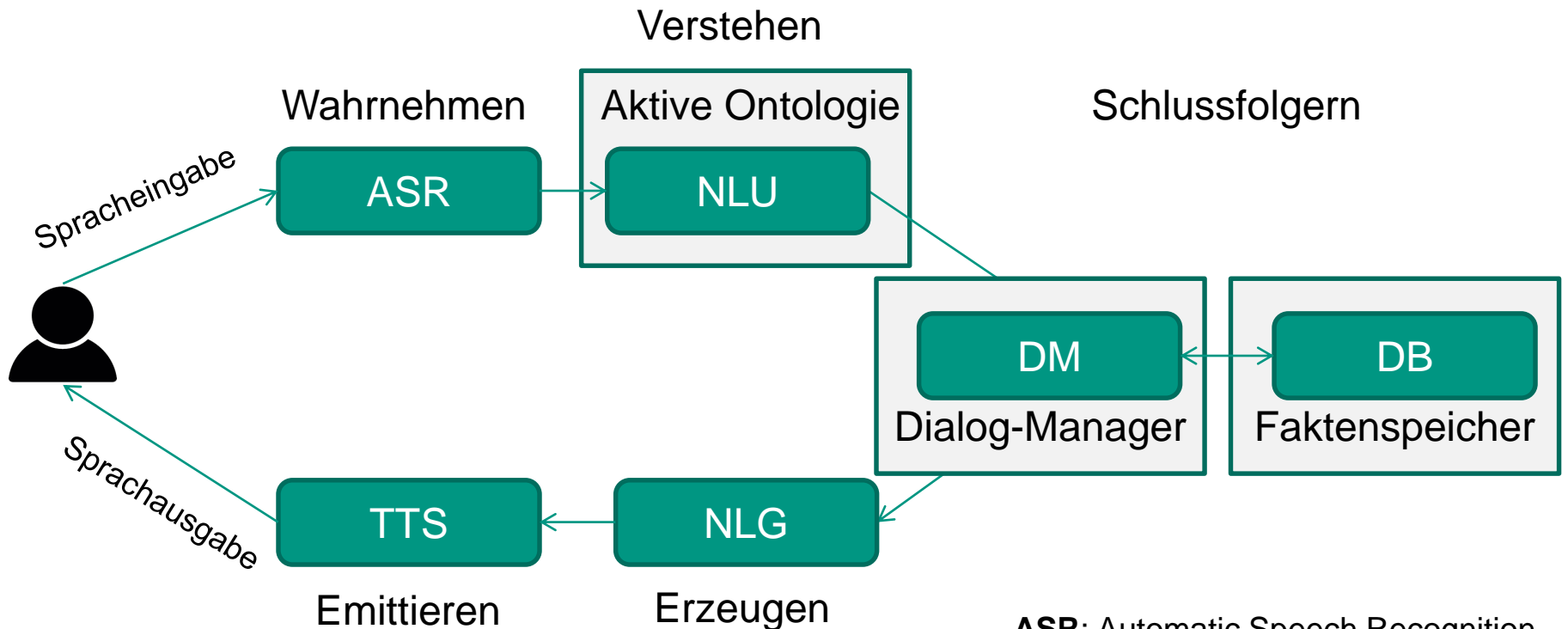
Erkennung der Konzepte

Konzept	Verwendete Technik	Beispiel
Departure Arrival	Präfixe Postfixe Wörterliste	From Karlsruhe Berlin is the destination Karlsruhe, Berlin
Dep. Date	Präfixe Postfixe reguläre Ausdrücke Bibliothek (SUTime - Stanford)	Fly on 11/12/2015 11/12/2015 is the departure date 11/12/2015 Next Monday
Flight Class	Präfixe Postfixe Wörterliste	Flight class is economy First class Economy

Ergebnis Aktive Ontologie „Flug“



Entwurf Dialog Manager



ASR: Automatic Speech Recognition
NLU: Natural Language Understanding
DM: Dialog Manager
DB: Database
NLG: Natural Language Generation
TTS: Text To Speech

Entwurf Dialog Manager

Book me a flight
to Manchester

Flight

Departure

Arrival

Dep.
Date

Flight
Class

Dialogue

Mandatory

Optional

Fehlende obligatorische
Eingabe

- 1- Departure
- 2- Departure date

Factstore

Fehlende optionale
Eingabe

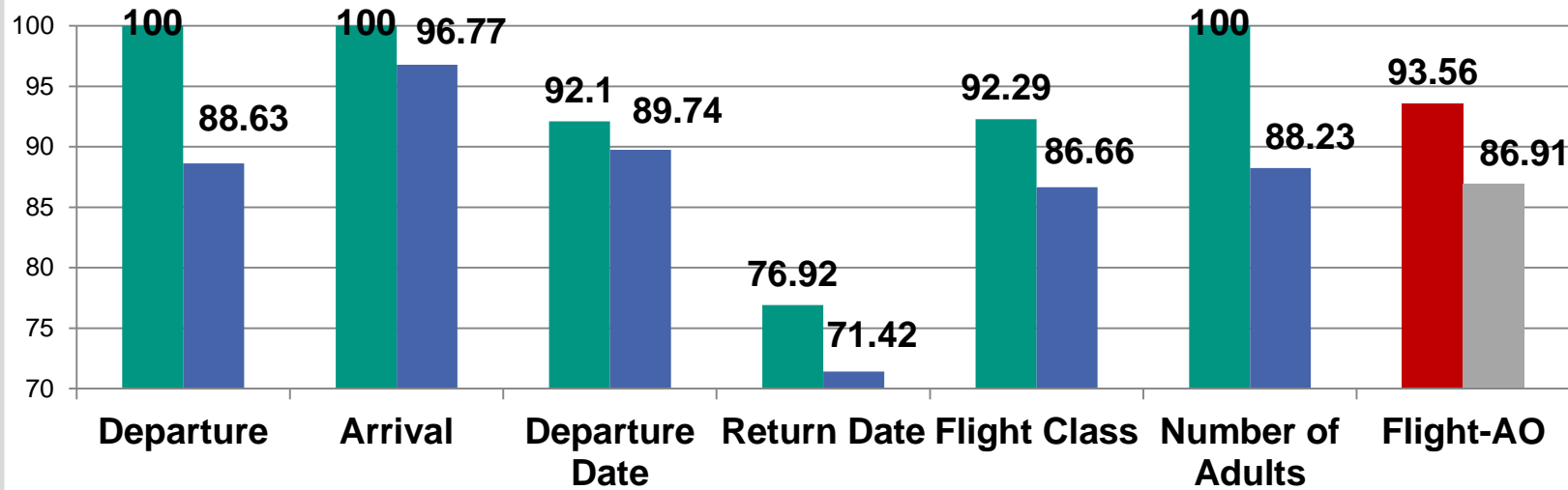
- 1- Flight class

Demo

Evaluation

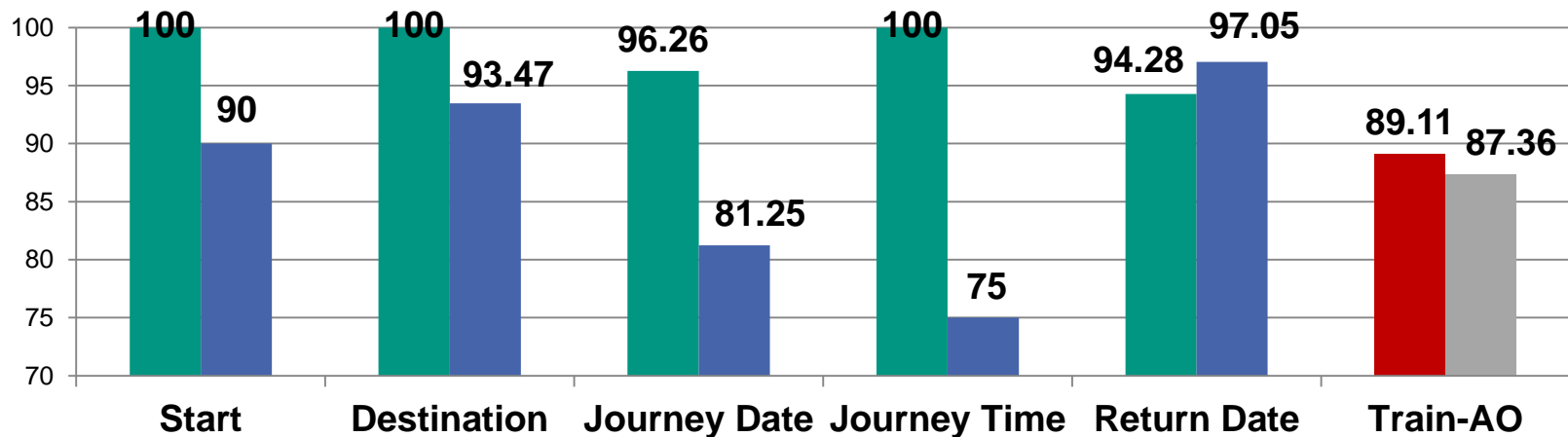
FlugAO

■ Präzision
■ Ausbeute

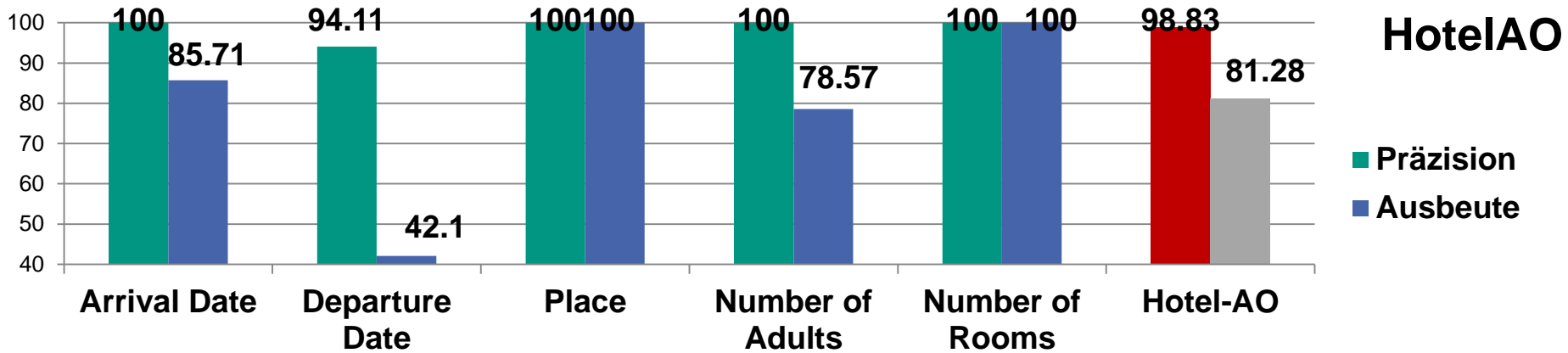


BahnAO

■ Präzision
■ Ausbeute



Evaluation



Eingabe	vollständig	
Erkennung	vollständig	unvollständig
Flug	21 (66%)	11 (34%)
Bahn	15 (68%)	7 (32%)
Hotel	14 (74%)	5 (26%)
Gesamt	50 (65%)	23 (35%)

■ Aktive Ontologien

- 30 formularbasierte Dienste aus drei verschiedenen Kategorien können verwendet werden

■ EASIER

- Neue Knotentypen (Datum, Zeit, Ziffern)
- Unterstützung von Ausdrücken die aus mehreren Wörtern bestehen

■ Dialog-Manager

- Nach fehlende Informationen fragen
- Ergebnisse zeigen

■ Ausblick

- Aktive Ontologien
 - Negation, Referenzen und weitere sprachliche Varianten werden unterstützt
- Dialog-Manager
 - Fragen vollautomatisch generieren
 - Aufruf externer Dienste

**Vielen Dank für Ihre
Aufmerksamkeit!**

Literatur

- [Guz08] Guzzoni, Didier: Active: a unified platform for building intelligent applications, Ecole Polytechnique Federale De Lausanne, PhD Thesis, Januar 2008

- [BL] Blersch, Martin; Landhäußer, Mathias: EASIER: An Approach to Automatically Generate Active Ontologies for Intelligent Assistants. In: International Symposium on Information Systems and Software Engineering: ISSE 2016

- [BJNS10] Berlanga, Rafael ; Jimenez-Ruiz, Ernesto ; Nebot, Victoria ; Sanz, Ismael: Faeton: Form analysis and extraction tool for ontology construction. In: International Journal of Computer Applications in Technology

- [RG06] Roitman, Haggai ; Gal, Avigdor: Ontobuilder: Fully automatic extraction and consolidation of ontologies from web sources using sequence semantics. In: Current Trends in Database Technology-EDBT 2006. Springer, 2006, S. 573-576

Literatur

- [DVNR03] Davalcu, H. ; Vadrevu, Srinivas ; Nagarajan, Saravanakumar ; Ramakrishnan, I.V.: OntoMiner: bootstrapping and populating ontologies from domain-specific Web sites. In: IEEE Intelligent Systems 18 (2003), September, Nr. 5, S. 24-33

- [AOSS95] Aust, Harald ; Oerder, Martin ; Seide, Frank ; Steinbiss, Volker: The Philips automatic train timetable information system. In: Speech Communication 17 (1995), November, Nr. 3-4, S. 249-262.

- [WY07] Williams, Jason D. ; Young, Steve: Partially Observable Markov Decision Processes for Spoken Dialog Systems. In: Comput. Speech Lang. 21 (2007), April, Nr. 2, S. 393-422.