

**Fraunhofer**  
Dresden      IAIS

# Knowledge Graphs

## In Conversational AI Platforms

Dr. Mikhail Galkin

Fraunhofer IAIS & Smart Data Analytics  
**Dresden**

# Outline

Background

NLP, NLU, KGs & Conversational AI

- The Mushroom Effect

The SPEAKER Platform



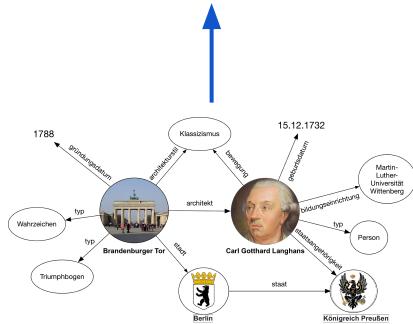
# Background

We build conversational AI platforms



# Background

We build conversational AI platforms

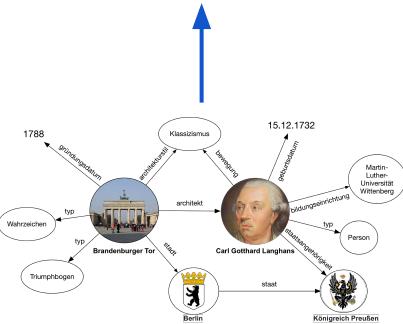


Powered by knowledge graphs

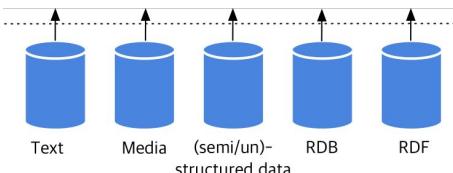


# Background

We build conversational AI platforms



Powered by knowledge graphs

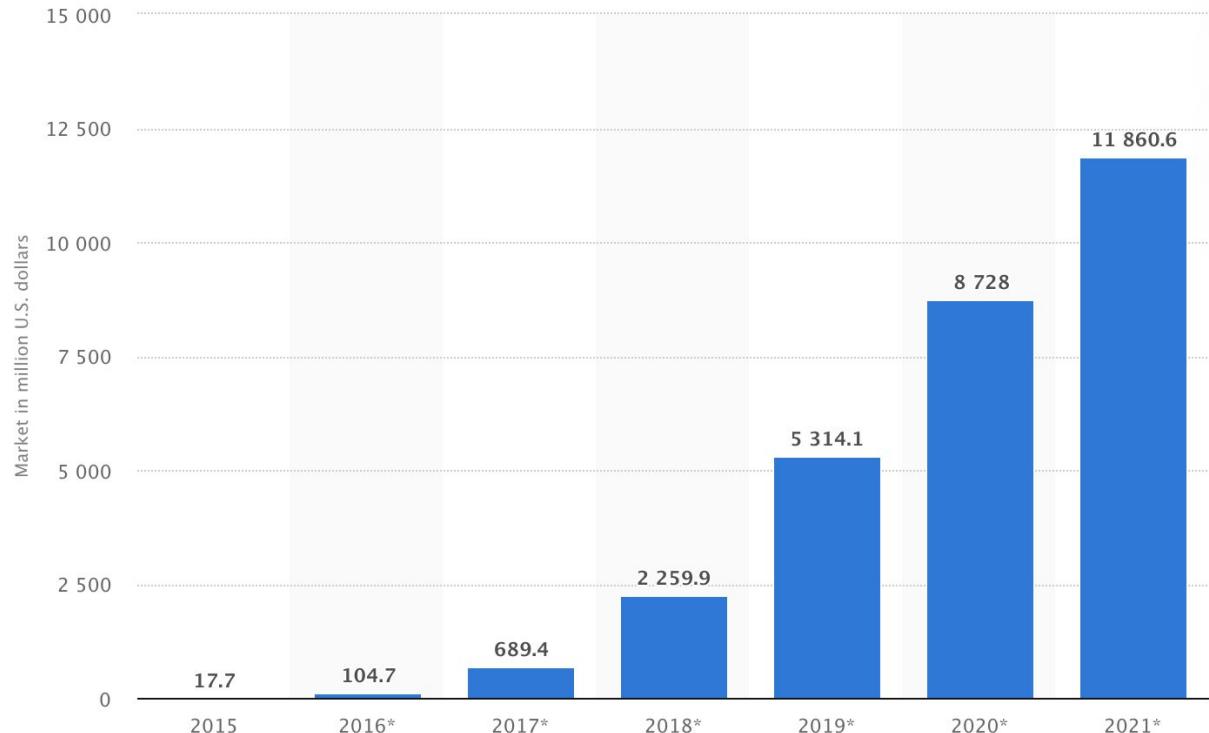


Obtained by integrating heterogeneous data

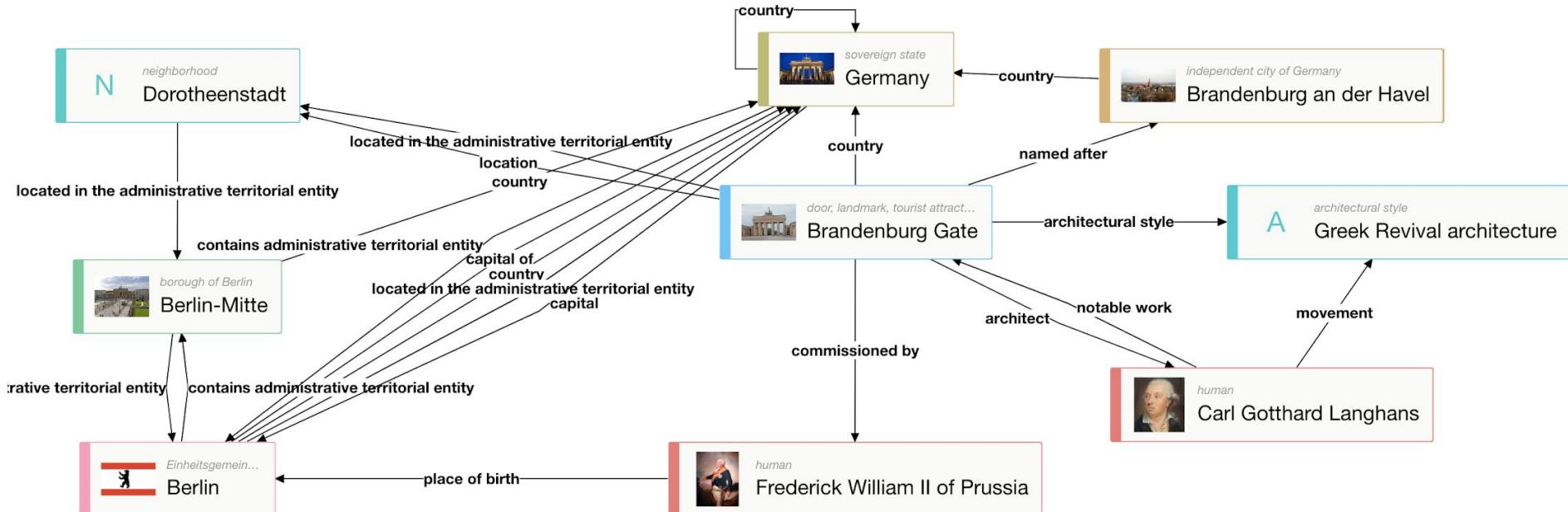
# Background

- Speech and text are the most natural interfaces
- Dialogue systems become a standard communication interface

Size of the consumer virtual digital assistant (VDA) market worldwide from 2015 to 2021  
*(in million U.S. dollars)*



# Knowledge Graph



<https://wikidata.metaphacts.com>

## Knowledge Graphs

# What's available

**Open-domain:** Wikidata, DBpedia

**Biomed:** Drugbank, SNOMED-CT, Bio2RDF

## Ontologies

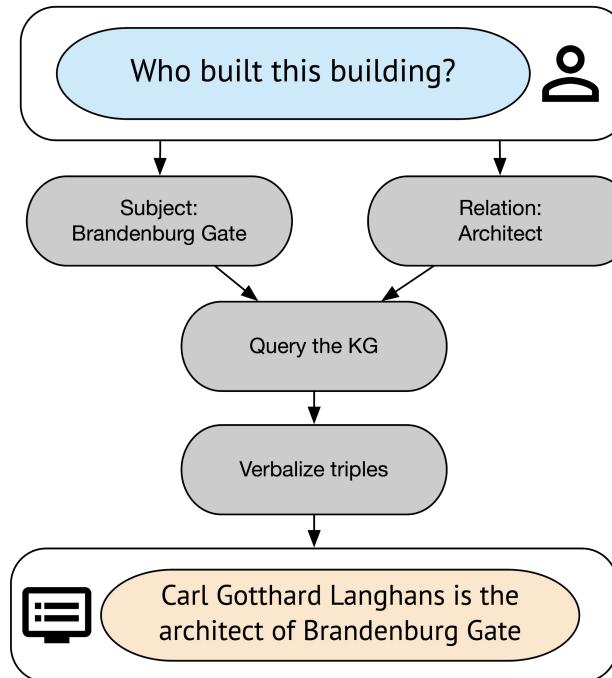
**Industry 4.0:** RAMI

**Finance:** FIBO, FRO, XBRL, FinReg

etc

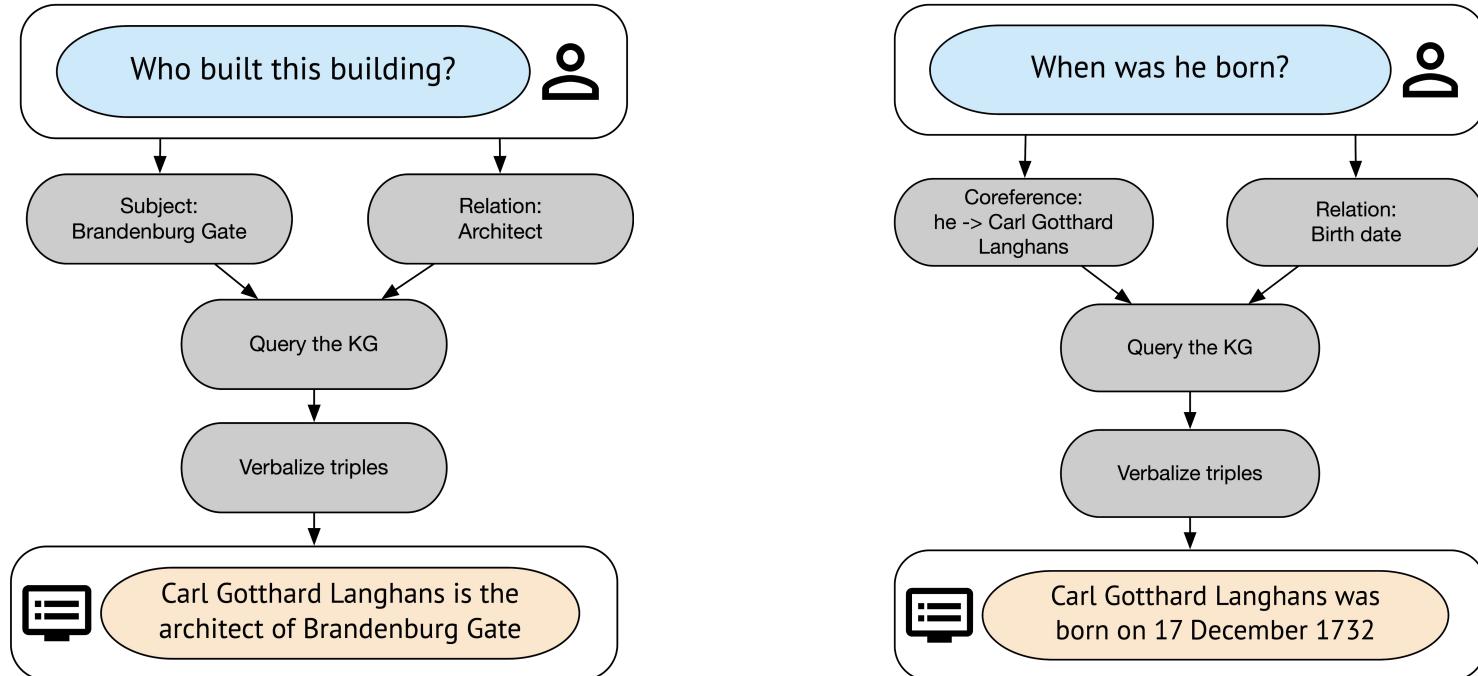
<https://wikidata.metaphacts.com>

# What's possible - Question Answering



<https://wikidata.metaphacts.com>

# What's possible - Dialogues



# Knowledge graph-based dialogue systems

Relation linking	#
Predicates	131
Training examples	52122

- Knowledge graph schema can be small, but training data has to be large ( $> 10\ 000$  data points)
  - Less data -> rule-based systems are preferred

# Knowledge graph-based dialogue systems

Relation linking	#
Predicates	131
Training examples	52122

- Knowledge graph schema can be small, but training data has to be large ( $> 10\ 000$  data points)
  - Less data -> rule-based systems are preferred

We signed a contract with  
X last week.

Great

How many employees  
work there?

- Dialogue  $\neq$  QA
  - Goal-oriented and particular context
  - Dialogue state (context) has to be memorized

# Mushrooms?

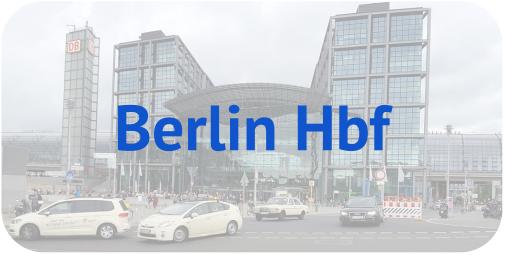


Berlin Hbf

What is this building?

Q

# Mushrooms?



What is this building?

Q

A

This is Berlin Hauptbahnhof

# Mushrooms?



Berlin Hbf

What is this building?

Q

A

This is Berlin Hauptbahnhof

What is its architectural style?

Q

# What is its architectural style?



# What is architectural style?



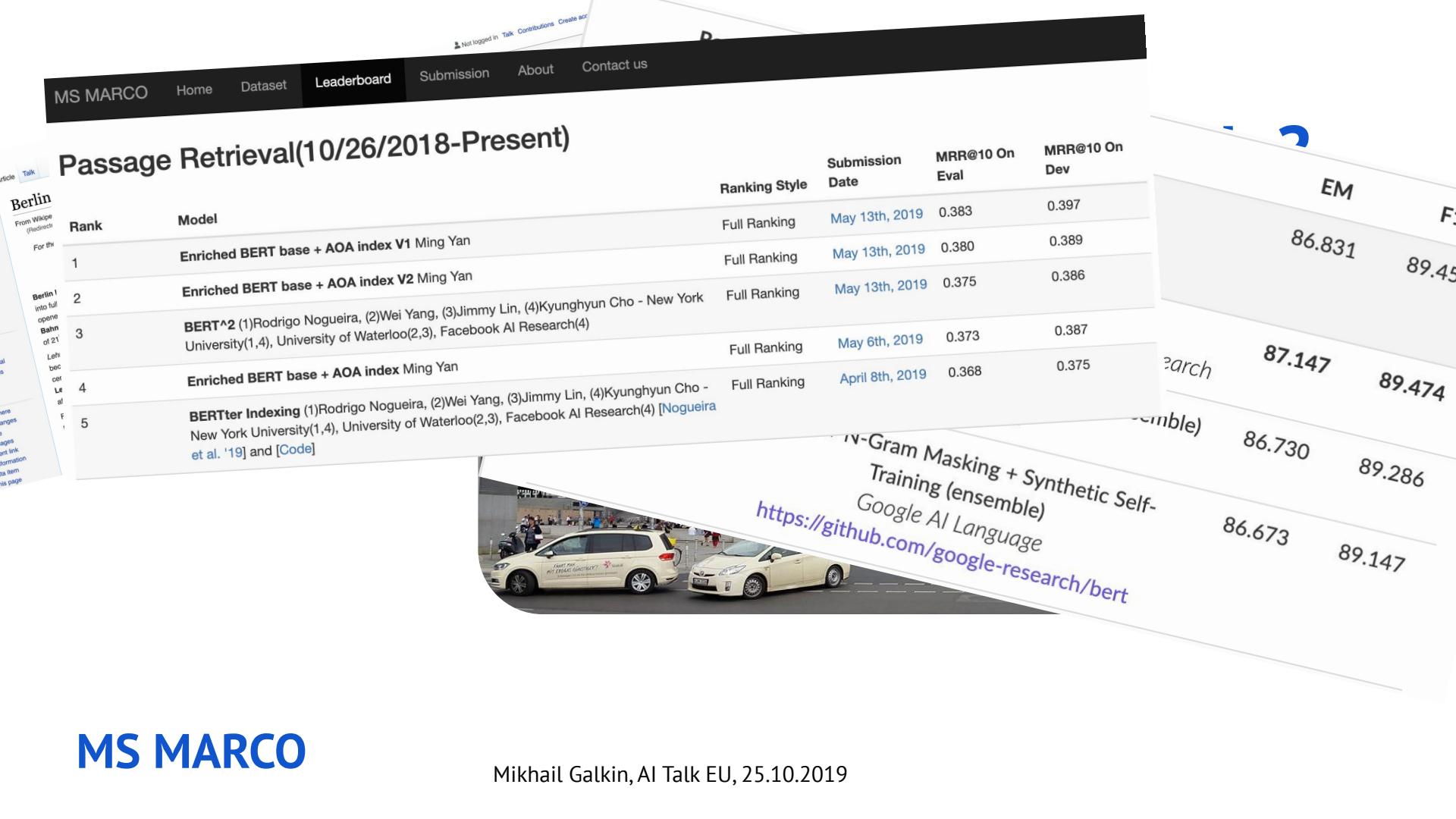
Rank	Model	P	R	F1
1	Human Performance Stanford University (Rajpurkar & Jia et al. '18)	86.831	89.452	
2	BERT + DAE + AoA (ensemble) Joint Laboratory of HIT and iFLYTEK Research	87.147		89.474
3	BERT + ConvLSTM + MTL + Verifier (ensemble) Layer 6 AI	86.730		89.286



## Rank

Mikhail Galkin, AI Talk EU, 25.10.2019

# SQuAD



# Passage Retrieval(10/26/2018-Present)

Rank	Model
1	Enriched BERT b...
2	Enriched BERT t...
3	BERT <sup>^2</sup> (1)Rod...
4	Enriched BERT
5	BERTter Inc...

From Wikipedia (Redirect)  
For the  
Berlin 1 into full  
openne  
Bahn  
of 21  
Leh  
bec  
cer  
Le  
al  
F ,  
here  
anges  
tags  
link  
formation  
a item  
us page

## Leaderboard (Fullwiki Setting)

Ranking Style	Submission Date	MRR@10 On Eval	MRR@10 On Dev
Full Ranking	May 13th, 2019	0.383	0.397
	May 13th, 2019	0.380	0.389
		0.386	

In the *fullwiki* setting, a question-answering system must find the answer to a question in the scope of the entire Wikipedia. Similar to in the distractor setting, systems are evaluated on the accuracy of their answers (Ans) and the quality of the supporting facts they use to justify them (Sup).

Model	Code	Ans		Sup		Joint	
		EM	F <sub>1</sub>	EM	F <sub>1</sub>	EM	F <sub>1</sub>
Cognitive Graph QA (single model) Tsinghua KEG & Alibaba DAMO Academy Ding et al., ACL'19	Feb 21, 2019	37.12	48.87	22.82	57.69	12.42	34.92
MUPPET (single model) Anonymous	Mar 5, 2019	30.61	40.26	16.65	47.33	10.85	27.01
GRN + BERT (single model) Anonymous	Apr 7, 2019	29.87	39.14	13.16	49.67	8.26	25.84
GRN (single model) Anonymous	Mar 4, 2019	27.34	36.48	12.23	48.75	7.40	23.55
Mikhail Galkin, ArXiv 2019							

# Mushrooms!



What is this building?

Q

A

This is Berlin Hauptbahnhof



What is its architectural style?

Q

A

Mushroom

Q

# Mushrooms!

What is its architectural style?

## Function [ edit ]

The Berlin Hauptbahnhof is part of the **mushroom** concept that was being made in Berlin, in which the station forms as a connecting point for converging and intersecting lines, of different modes of public transport there.

## Planning the new station [ edit ]

Soon after the fall of the **Berlin Wall** in 1989, city planners began work on a transport plan for reunified Berlin. One element of this became the "Pilzkonzept" (**mushroom** concept), in which a new north-south railway line intersecting the Stadtbahn was to be constructed. The name derived from the shape formed by the new line and existing lines, which vaguely resembles a **mushroom**.

Q

# Mushrooms!

What is its architectural style?

## Function [edit]

The Berlin Hauptbahnhof is part of the **mushroom** concept that was being made in Berlin, in which the station forms as a connecting point for converging and intersecting lines, of different modes of public transport there.

## Planning the new station [edit]

Soon after the fall of the **Berlin Wall** in 1989, city planners began work on a transport plan for reunified Berlin. One element of this became the "Pilzkonzept" (**mushroom** concept), in which a new north-south railway line intersecting the Stadtbahn was to be constructed. The name derived from the shape formed by the new line and existing lines, which vaguely resembles a **mushroom**.

- ✓ its = Berlin Hauptbahnhof
- ✓ In Berlin
- ✓ Concept ~ Style

- ✗ Applicable to the city concept, not architectural style of the station
- ✗ Close, but not correct



# With KG flavors

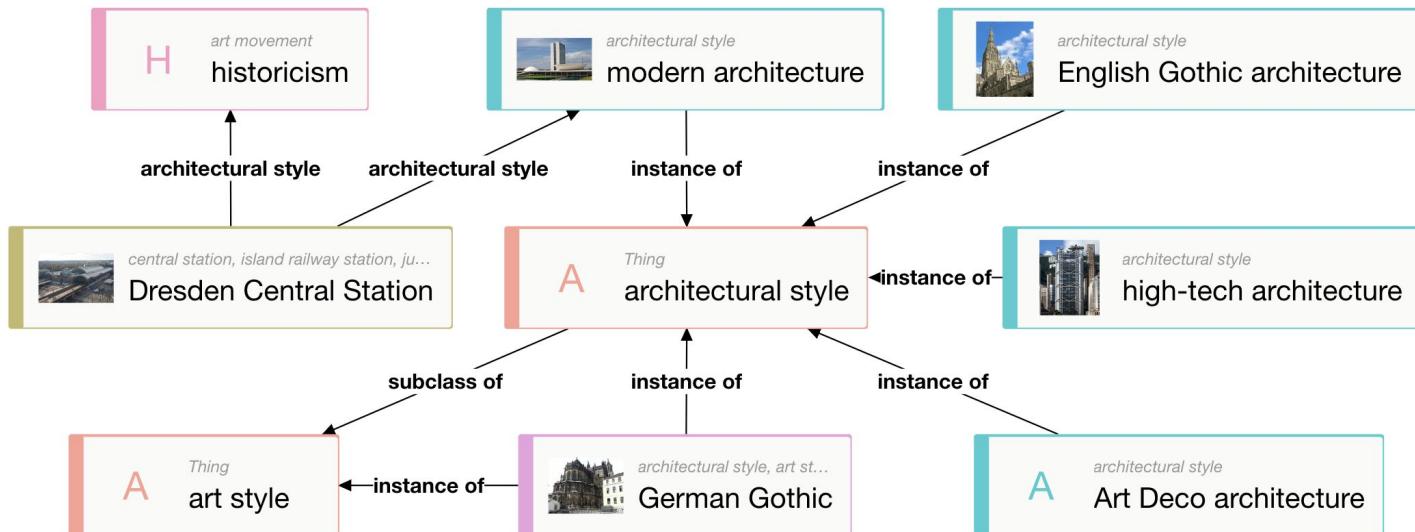
Can mushrooms be  
an architectural  
style?



Can mushrooms be  
an architectural  
style?

Probably not

# With KG flavors



# So why you need graphs?

How many children  
does Berlin Hbf have?

Implicit or explicit constraints on produced answers

# So why you need graphs?

How many children  
does Berlin Hbf have?

Train stations  
don't have kids

Implicit or explicit constraints on produced answers

- reduce candidates space
- help to fight the mushroom effect
- **ontologies help**

# So why you need graphs?

How many children  
does Berlin Hbf have?

Train stations  
don't have kids

What is the busiest  
train station in  
Germany?

Implicit or explicit constraints on produced answers

- reduce candidates space
- help to fight the mushroom effect
- **ontologies help**

Complex QA via (sub)graphs aggregations

# So why you need graphs?

How many children  
does Berlin Hbf have?

Train stations  
don't have kids

What is the busiest  
train station in  
Germany?

Hamburg Hbf

Implicit or explicit constraints on produced answers

- reduce candidates space
- help to fight the mushroom effect
- **ontologies help**

Complex QA via (sub)graphs aggregations

```
select ?station ?visits where {  
?station wdt:P31 wd:Q18543139 .      # central stations  
?station wdt:P17 wd:Q183 .              # in Germany  
?station wdt:P1373 ?visits .            # daily visits  
} ORDER BY DESC(?visits) LIMIT 1        # sort
```

# So why you need graphs?

Takeaway 1

Graphs significantly improve reasoning  
compared to sole natural language inference

# So why you need graphs?

Takeaway 1

Graphs significantly improve reasoning compared to sole natural language inference

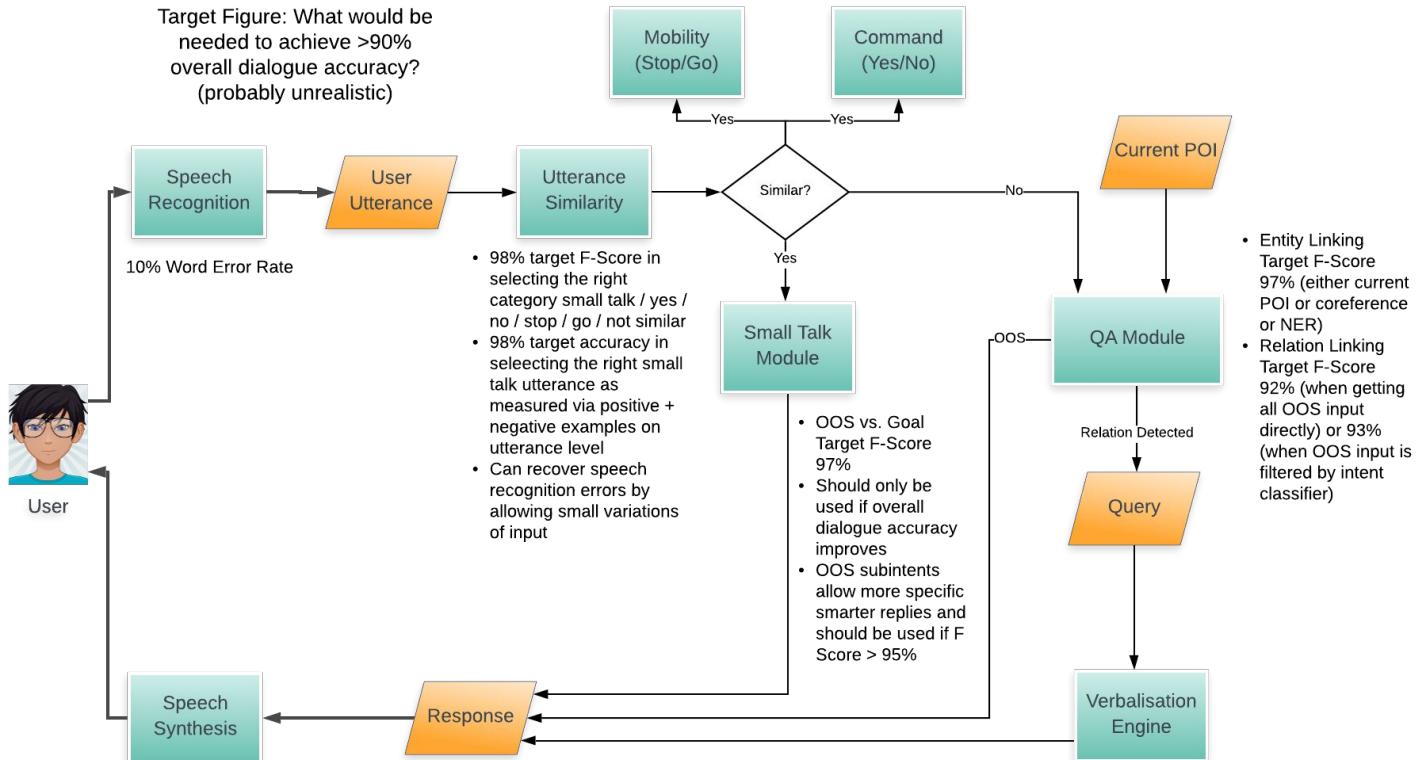
Takeaway 2

Reasoning outcomes are  
**explainable** and traceable

# KDDS

Knowledge-driven  
in-car dialogue  
system (EN/DE)

Full DBpedia  
2019 (wikidata  
branch)  
> 50M entities  
> 4B triples



# KDDS @ Hannover-Messe

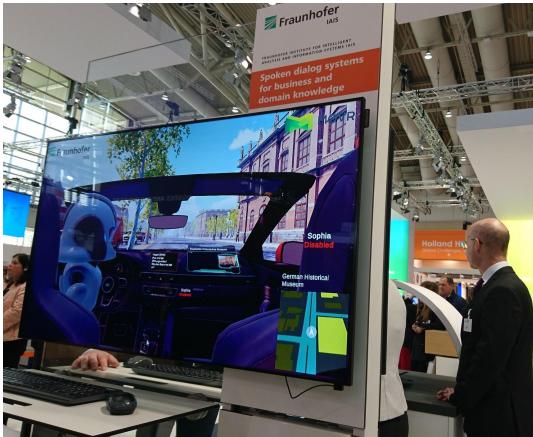
Building KGs from enterprise  
sources is still a challenge



# KDDS @ Hannover-Messe

Building KGs from enterprise sources is still a challenge

Depth of knowledge vs  
variety of domains



# KDDS @ Hannover-Messe

Building KGs from enterprise sources is still a challenge

Depth of knowledge vs variety of domains

Explainability is crucial



# Fraunhofer IAIS

## Standort Dresden

### ML2R

National Competence Center for  
Machine Learning Rhein-Ruhr

### Fraunhofer Center for Machine Learning

IAIS-led part of Fraunhofer  
Cluster of Excellence  
Cognitive Internet  
Technologies

### CEE AI

Center for Explainable and Efficient AI  
Technologies with TU Dresden



**Fraunhofer-Alliance Big Data AI**  
The biggest Fraunhofer alliance  
with > 30 institutes led by IAIS

### AI4EU

EU Lighthouse  
Project for AI



### International Data Spaces Association

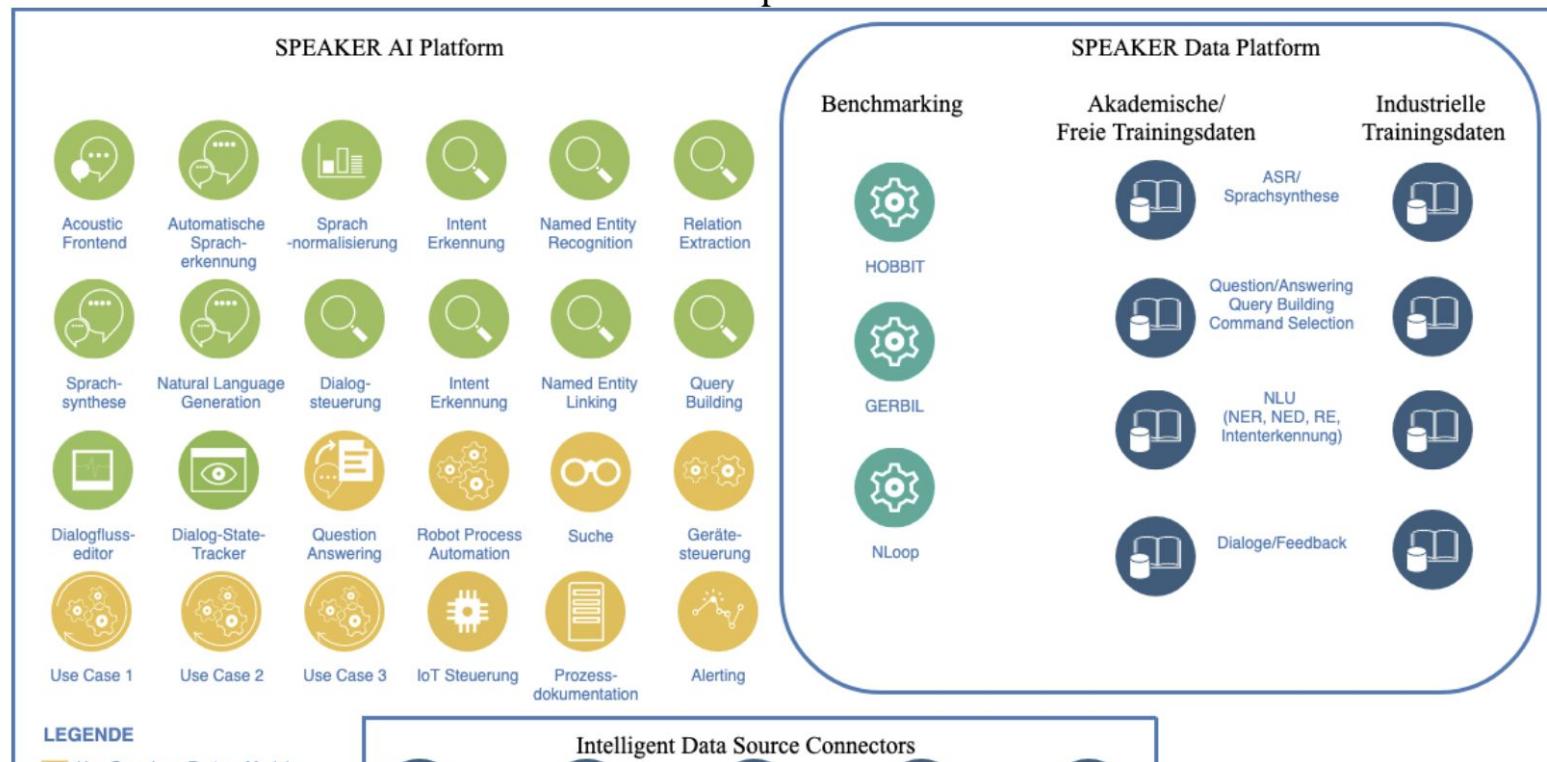
Data sovereignty for Big  
Data and AI, 100+  
companies

# SPEAKER

- ★ Digital sovereignty
  - Independent from Google, AWS
  - Compliant with German enterprise and user data protection regulations
  
- ★ B2B-oriented

# SPEAKER

- ★ Conversational AI
  - Acoustic frontend
  - Speech recognition in DE & EN
  - Focus on integrating enterprise knowledge
  - Modular design



#### LEGENDE

- Use Case bzw. Partner Modul
- SPEAKER Modul
- Benchmarking
- Datenkonnektoren

#### Intelligent Data Source Connectors



Industrial Data Sources



Datenbanken

Tabellen  
(CSV, Excel,...)

Kommunikations-  
kanäle

Business  
Processes

IoT/Cloud  
Speicher

# SPEAKER

## Consortium

WDR<sup>®</sup>

OTTO

VOLKSWAGEN  
AKTIENGESELLSCHAFT



+ 31 Assoziierte Partner

SAP DB

comma soft<sup>®</sup>  
THE KNOWLEDGE PEOPLE

OMNIBOT<sup>®</sup>  
ONE FOR ALL

mojin robotics  
moving intelligence

INNOVO CLOUD<sup>®</sup>

COD cognitive commerce

QIZY.NET  
AI APPLY-IT-YOURSELF NETWORK

SERVICE MEISTER

+ 5 KI-Innov. Proj.

retresco

SCOPEVISO

ONSEI

PlanQK

ProSeMo

+ 20 Verbundpartner

SIEMENS

DFKI

DATEV

Konsortialführung

Fraunhofer  
IAIS

Fraunhofer  
IIS

Ministerium für Wirtschaft, Innovation,  
Digitalisierung und Energie  
des Landes Nordrhein-Westfalen



Wolters Kluwer

EB  
Elektrorit

UNIVERSITÄT  
DUISBURG  
ESSEN

UNIVERSITÄT  
SAARLANDES

COMMERZBANK

BOSCH  
Invented for life

NRW.BANK

SENSAPE

YellowMap

Orchestrating a brighter world  
NEC

LAN<sup>1</sup>

eco  
VERBUND DER  
INTERNETWIRTSCHAFT

INTERNATIONAL DATA  
SPACES ASSOCIATION

UNIVERSITÄT  
PADERBORN  
Die Universität der Informationstechnik

Technische  
Universität  
Braunschweig

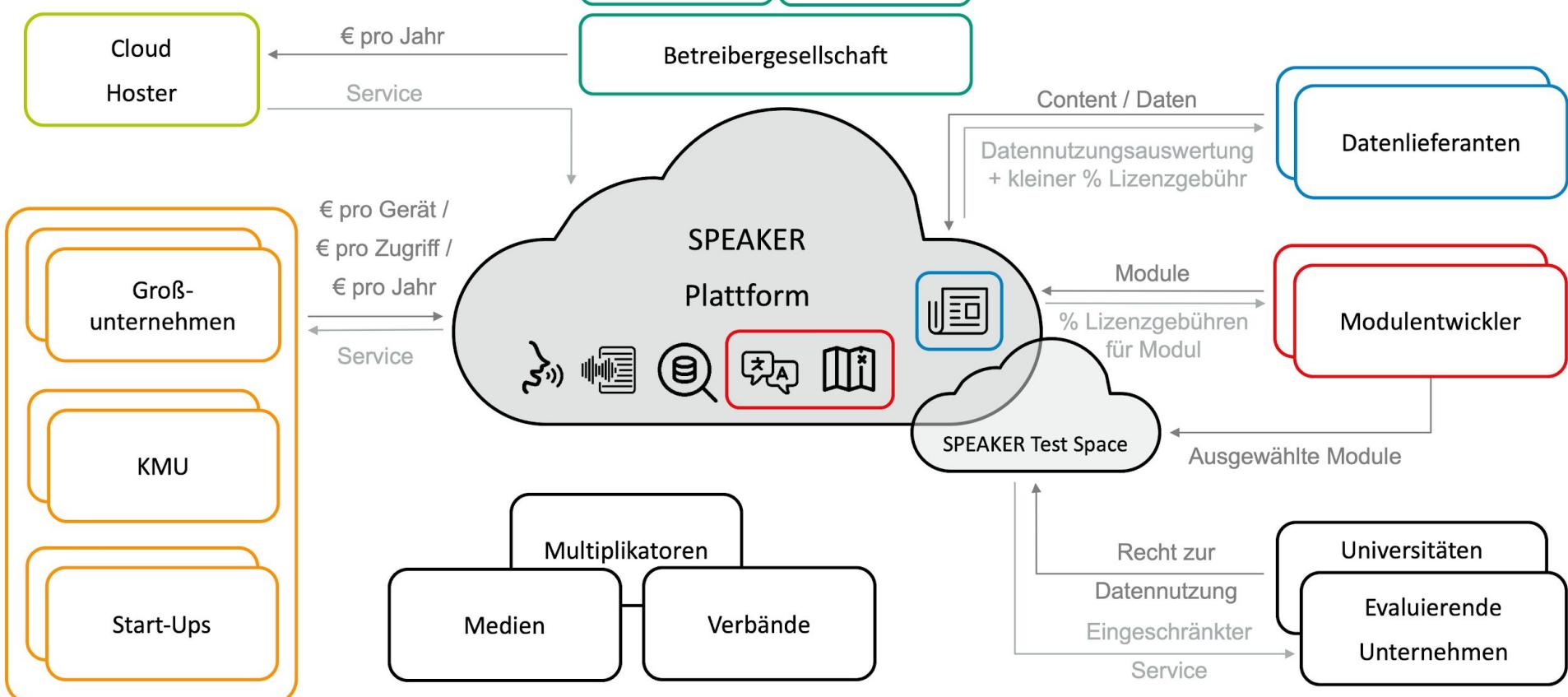
LMU  
LUDWIG-MAXIMILIANS-  
UNIVERSITÄT MÜNCHEN

WIKIFRIEND

CARMEQ.<sup>®</sup>

comevis

# SPEAKER Platform



# SPEAKER Use-cases

## Automotive

- Easy access to information about surroundings and landmarks
- Voice input & output increases driver safety
- Interactive city guides via natural dialogues



# SPEAKER Use-cases

## Disaster Management

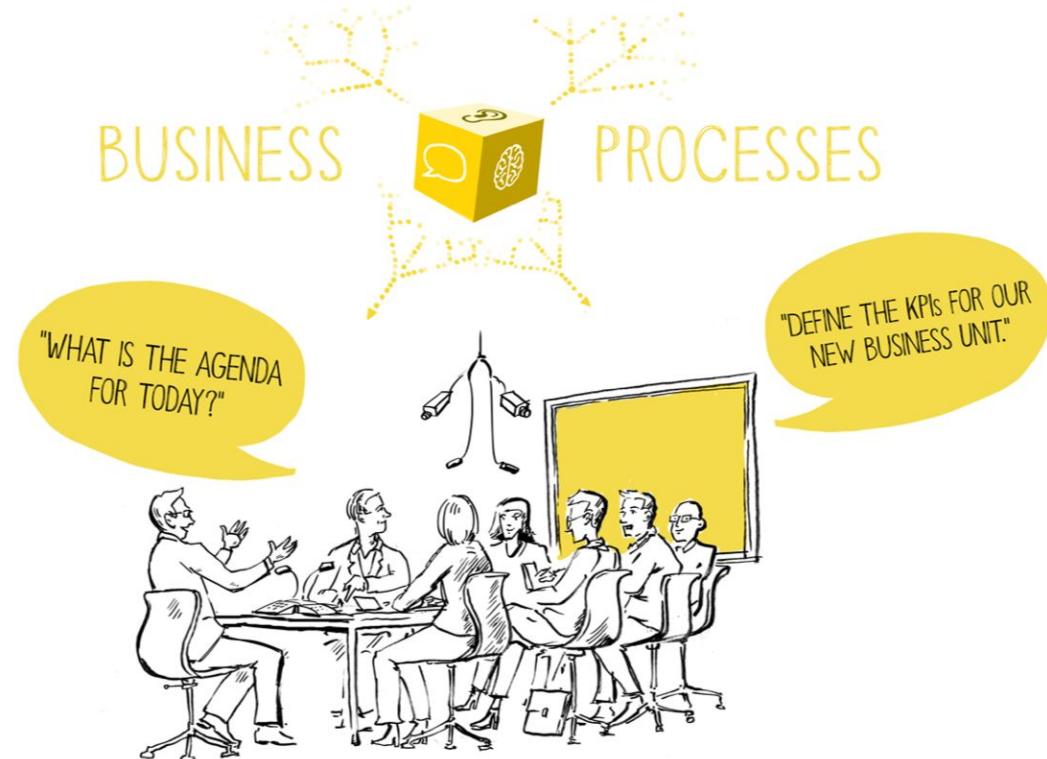
- Timeliness of response and remedy
- Optimization of resource planning and use
- Consistency of available information in real time



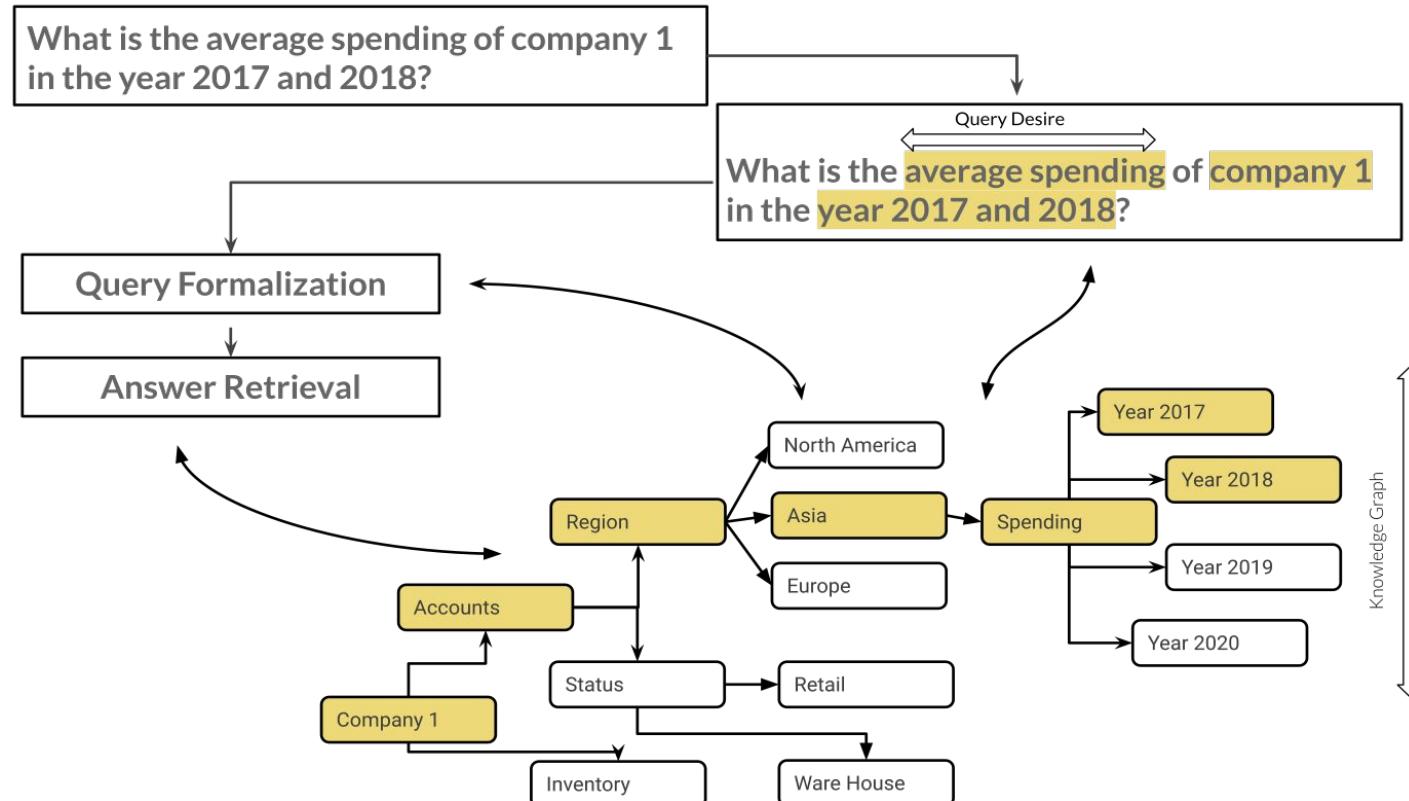
# SPEAKER Use-cases

## Business Processes

- Support and optimization for business and decision processes
- Comprehensive answers on the fly combining multiple data sources
- No need to search through files in the information system



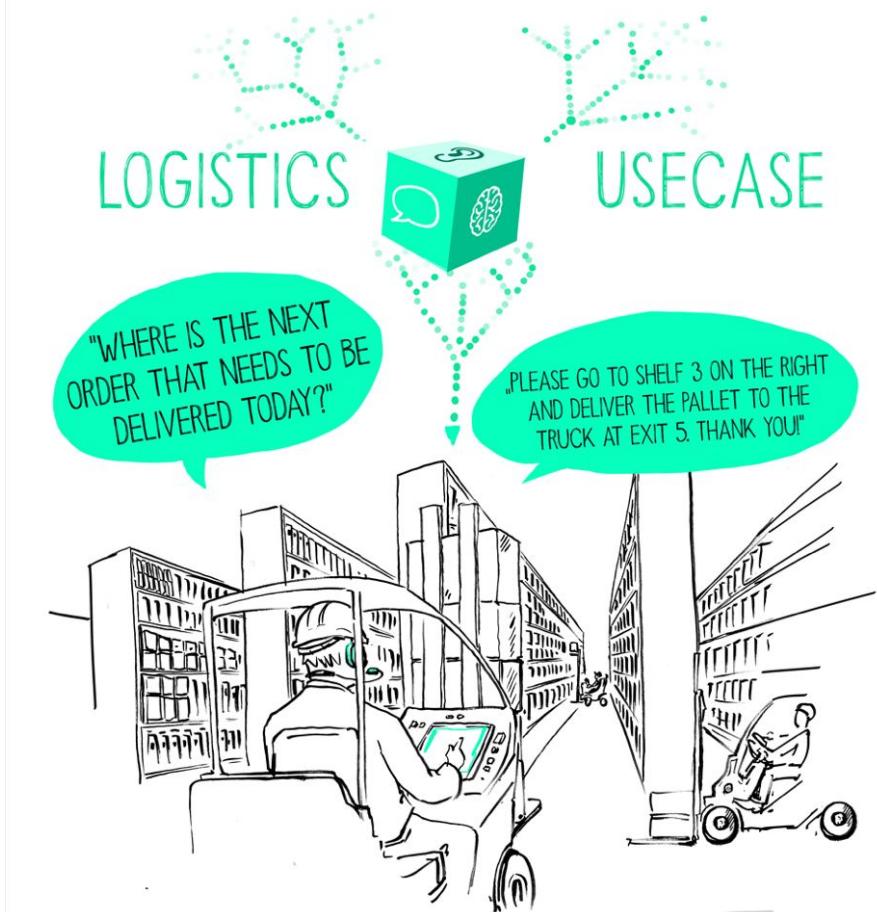
# SPEAKER Use-cases - Financial QA



# SPEAKER Use-cases

## Logistics

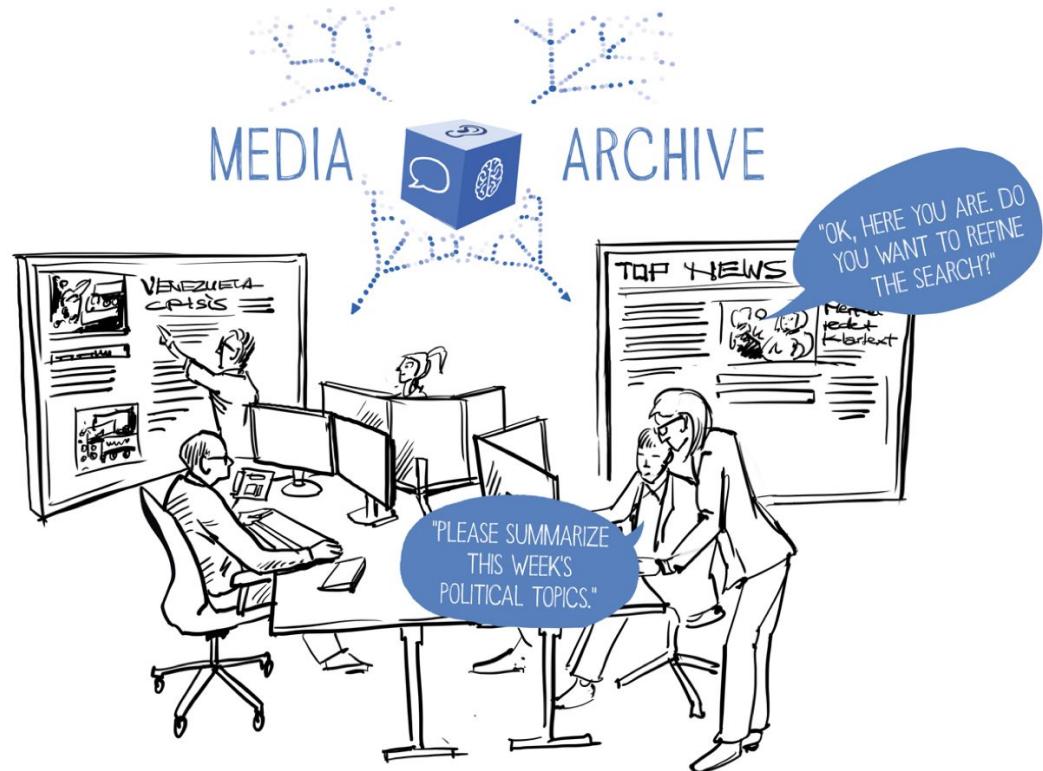
- Navigation through logistic centers or warehouses
- Low-barrier access to information and on the fly inferencing
- Time saving with less mistakes



# SPEAKER Use-cases

## Media Archive

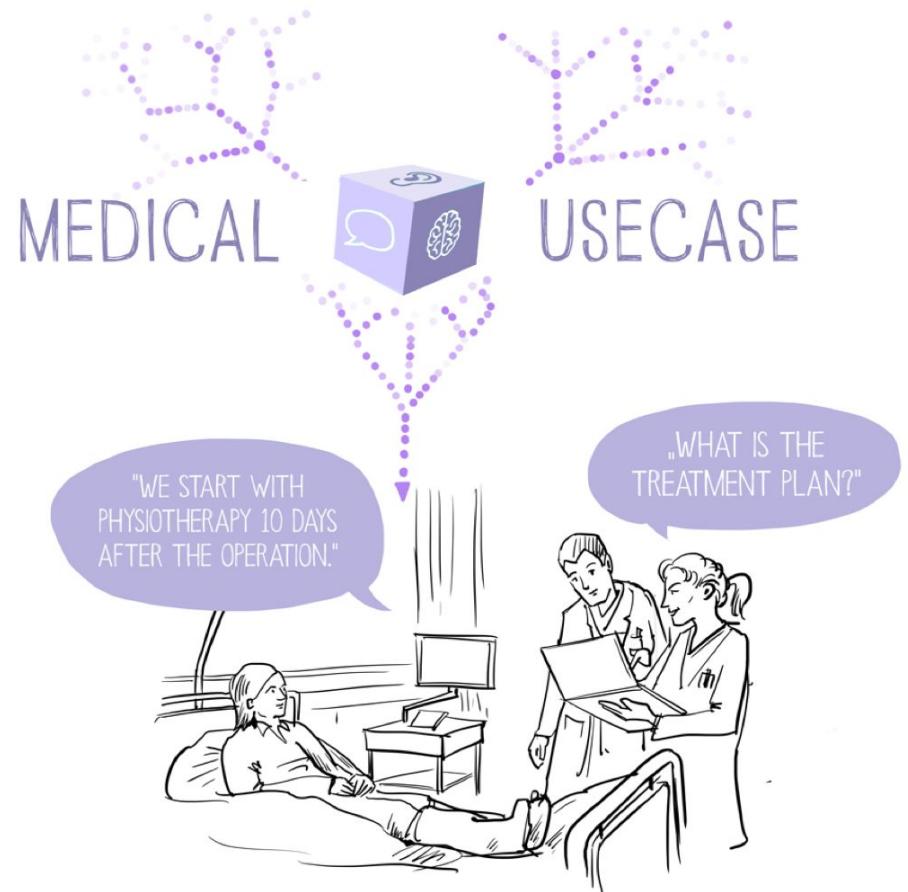
- Lightweight speech-based search for specific content
- Rapid orientation in massive amounts of media files
- Hands-free: intuitive and comfortable operation



# SPEAKER Use-cases

## Medical

- Reduced efforts for patient status reports
- Condensed natural language output combining multiple datasources
- Immediate access for efficient scheduling & treatment planning





**Fraunhofer**  
Dresden      IAIS

# Thank you!

[mikhail.galkin@iais.fraunhofer.de](mailto:mikhail.galkin@iais.fraunhofer.de)

[jens.lehmann@iais.fraunhofer.de](mailto:jens.lehmann@iais.fraunhofer.de)