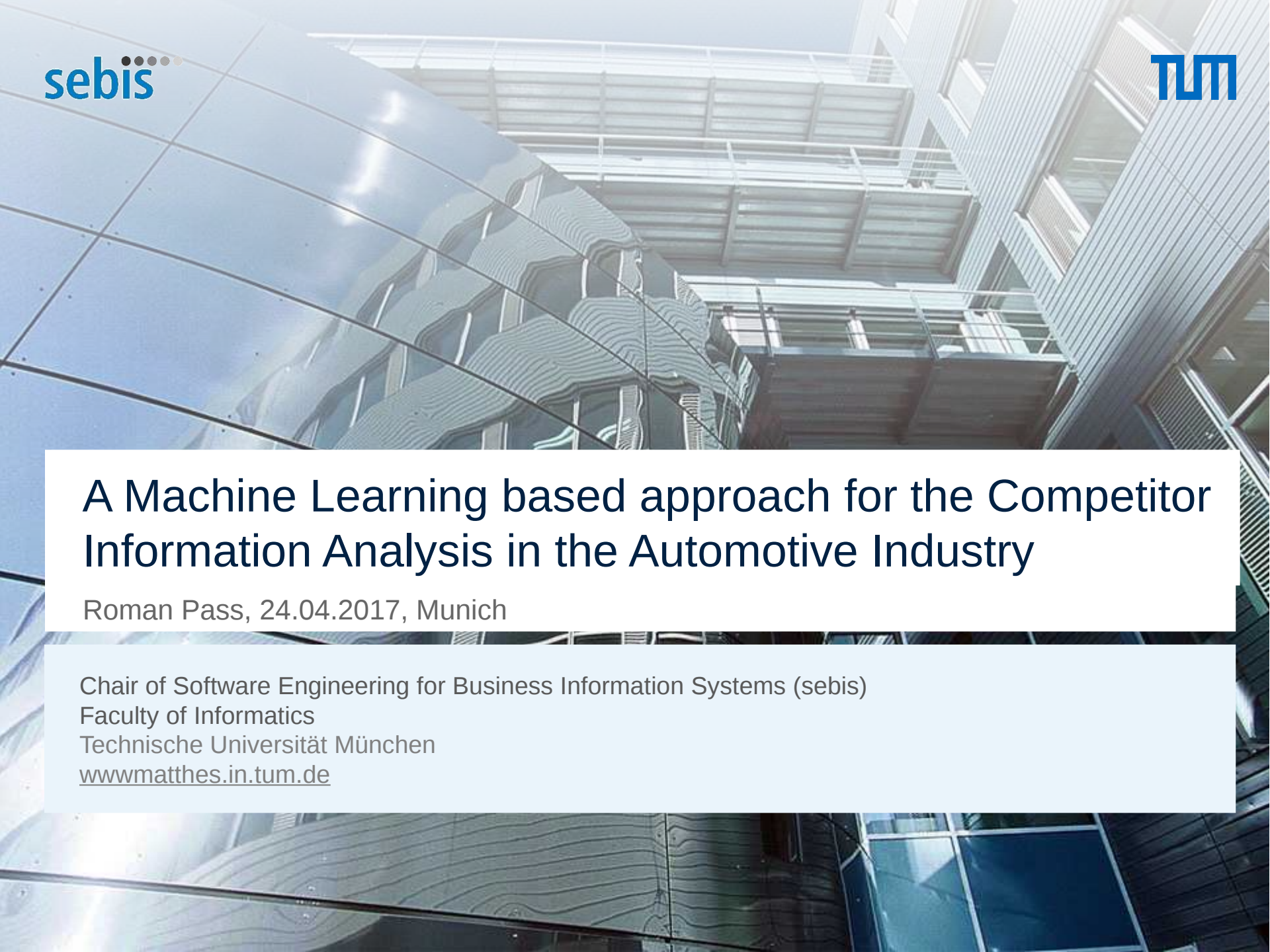


A Machine Learning based approach for the Competitor Information Analysis in the Automotive Industry

Roman Pass, 24.04.2017, Munich

Chair of Software Engineering for Business Information Systems (sebis)
Faculty of Informatics
Technische Universität München
wwwmatthes.in.tum.de

Outline

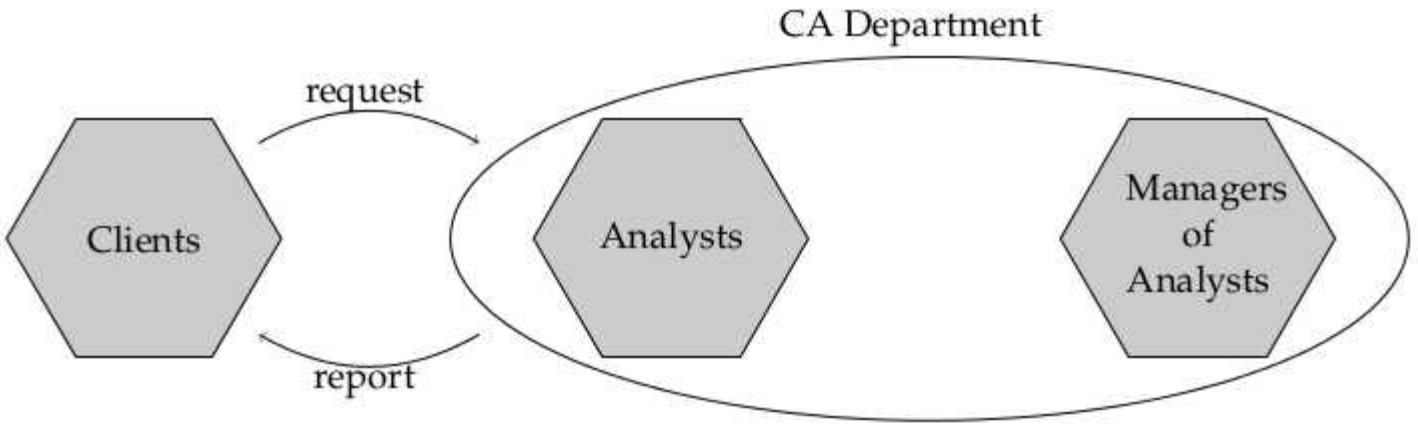
1. Introduction
   - Competitor Analysis
   - Problems
   - Research questions
2. Requirements
3. System Design
4. Evaluation
   - System Usability
   - Feedback
5. Future Work

Introduction

Competitor Analysis (CA)

- Identify Strength, Weakness, Opportunities, Threats (SWOT)
- Identify
  - All competitors
  - Basic competitors
  - Primary competitors
- Support decisions in management and development

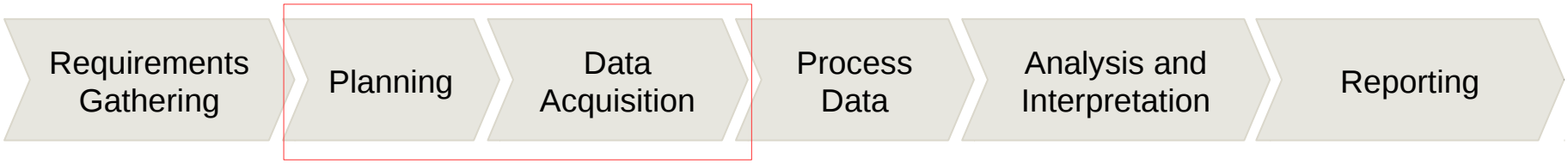
→ Spread Competitive Intelligence (CI) within the company



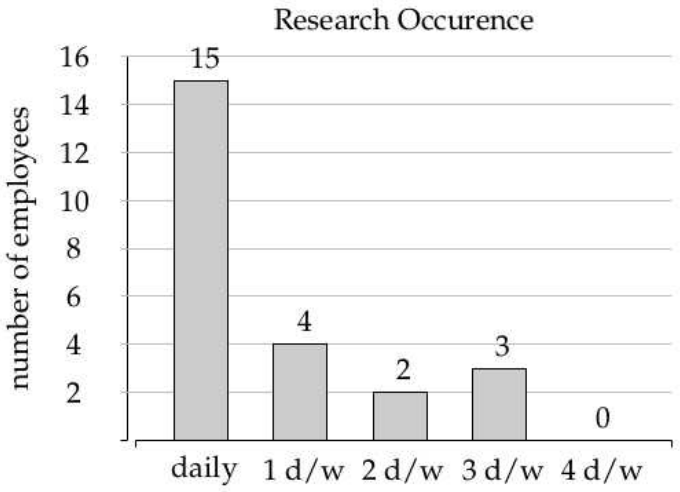
Introduction

Competitive Intelligence Process

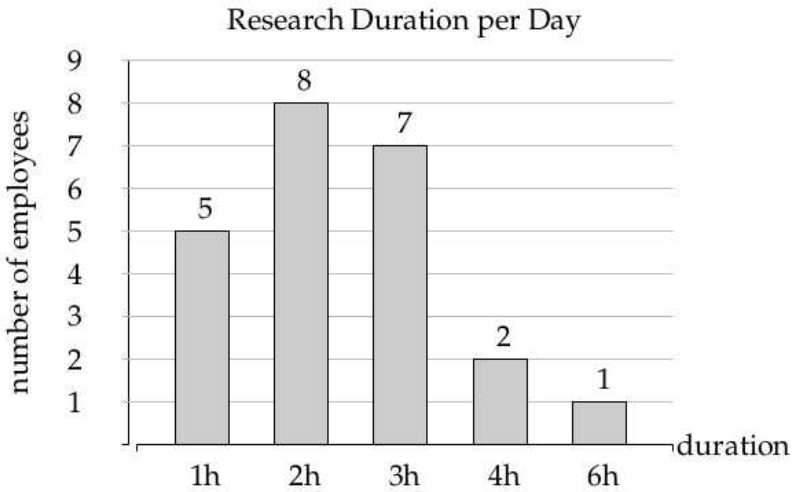
50 % of process time



Source: Rainer Michaeli. Competitive Intelligence: Strategische Wettbewerbsvorteile erzielen durch systematische Konkurrenz-, Markt- und Technologieanalysen. Springer-Verlag Berlin Heidelberg, 2006.



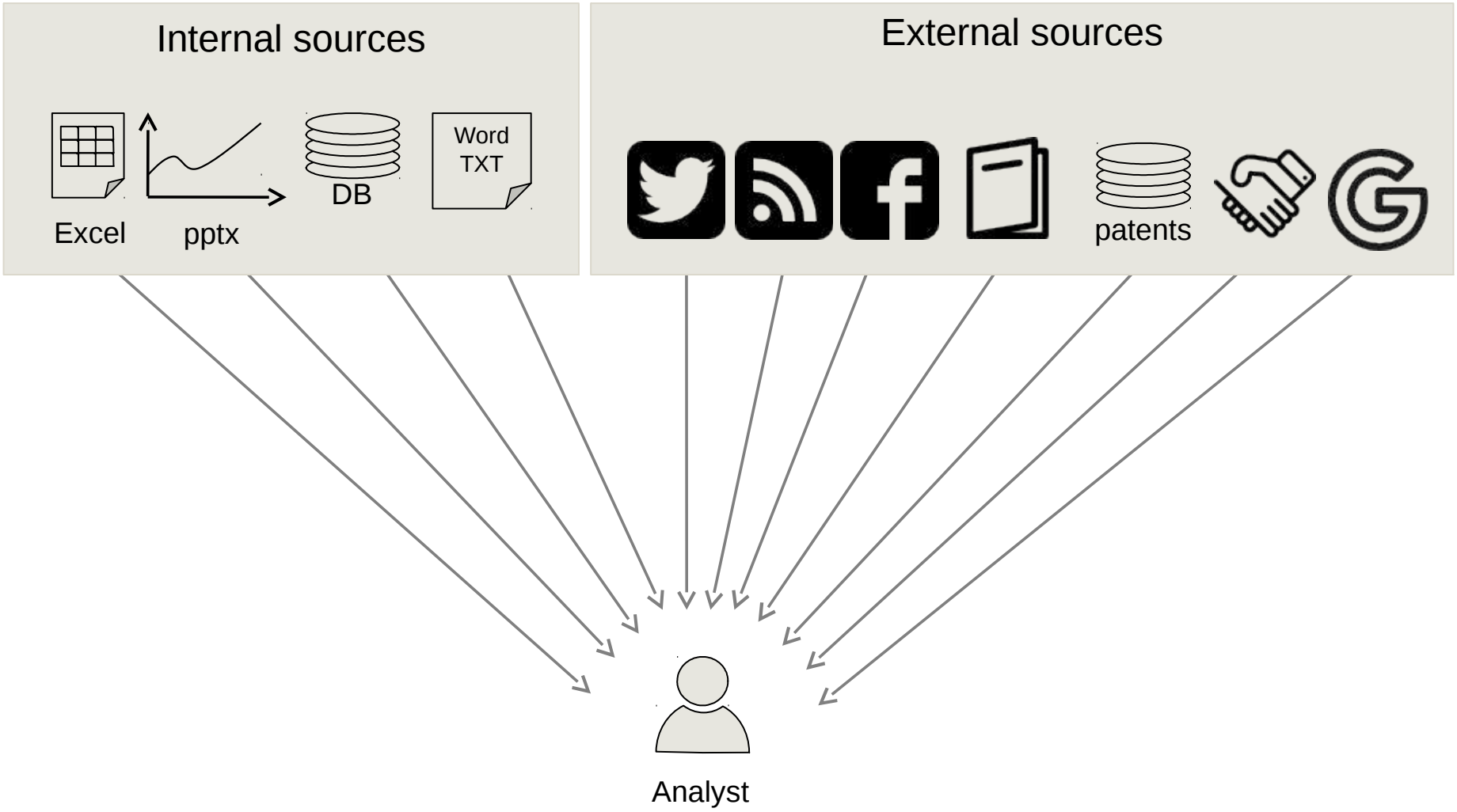
(a) Research Occurence per Week



(b) Research Duration per Day

Introduction

Competitive Intelligence Process – Data Acquisition



Introduction

Competitive Intelligence Process – Problems

- Too many information sources for manual analysis
- Long duration and high occurence of analysis
- Destinguish between relevant and irrelevant information
- Retrieve processed information

Introduction

Research Questions

- How to improve the existing competitor information analysis process in the automotive industry?
- What kind of documents and document sources are used for competitor information analysis in the automotive domain?
- Which categories/topics do these documents belong to?
- Which supervised classification algorithm is suitable for automatic document classification to support competitor information analysis in the automotive domain?

Outline

1. Introduction
   - Competitor Analysis
   - Problems
   - Research questions

2. Requirements

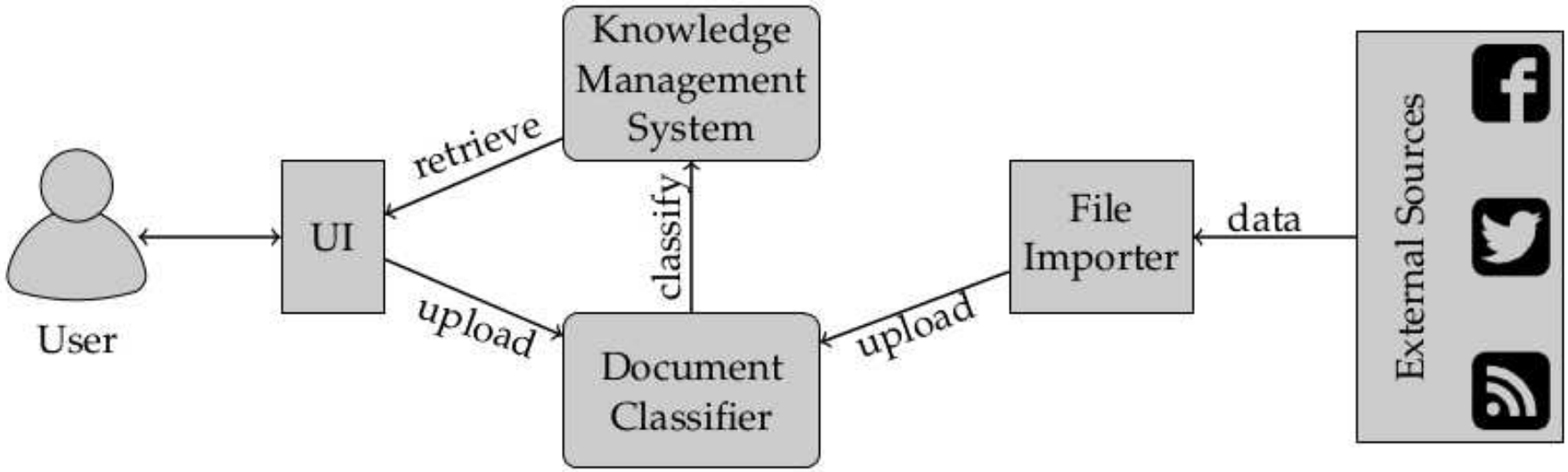
3. System Design

4. Evaluation
   - System Usability
   - Feedback

5. Future Work

Requirements

- One stream of information
- Search engine as UI
- Automated import of data from external sources
- Knowledge management system with the support for storing both the internal and external data
- Topic structure within knowledge management system

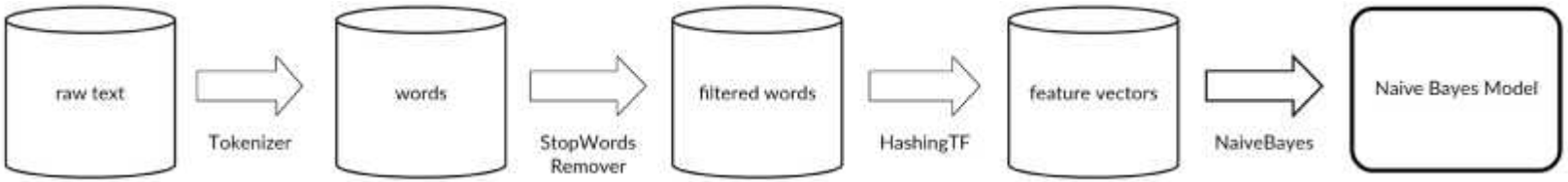


Automatic Document Classification

- Labels defined by department's topics
  - Prognosis
  - Exhibitions
  - General Information
  - Initial Evaluation
  - Press Evaluation
  - Garbage

- Training documents assigned manually to each topic

- Supervised Machine Learning (ML) algorithms compared
  - Support Vector Machine (SVM): low prediction accuracy
  - K-Nearest Neighbours (k-NN): low prediction performance
  - Naive Bayes (NB): good prediction performance and accuracy



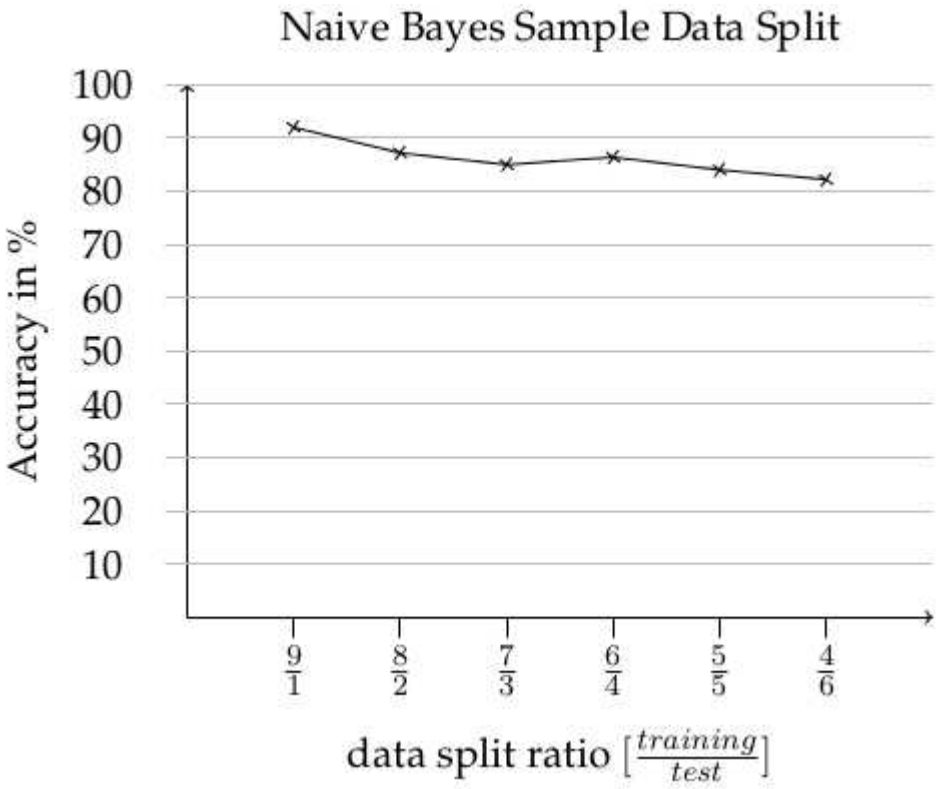
Automatic Document Classification

Naive Bayes

Training Phase:

- n-Fold cross validation (n=10)

- Different data split ratios (training : test) were tested



Outline

1. Introduction
   - Competitor Analysis
   - Problems
   - Research questions

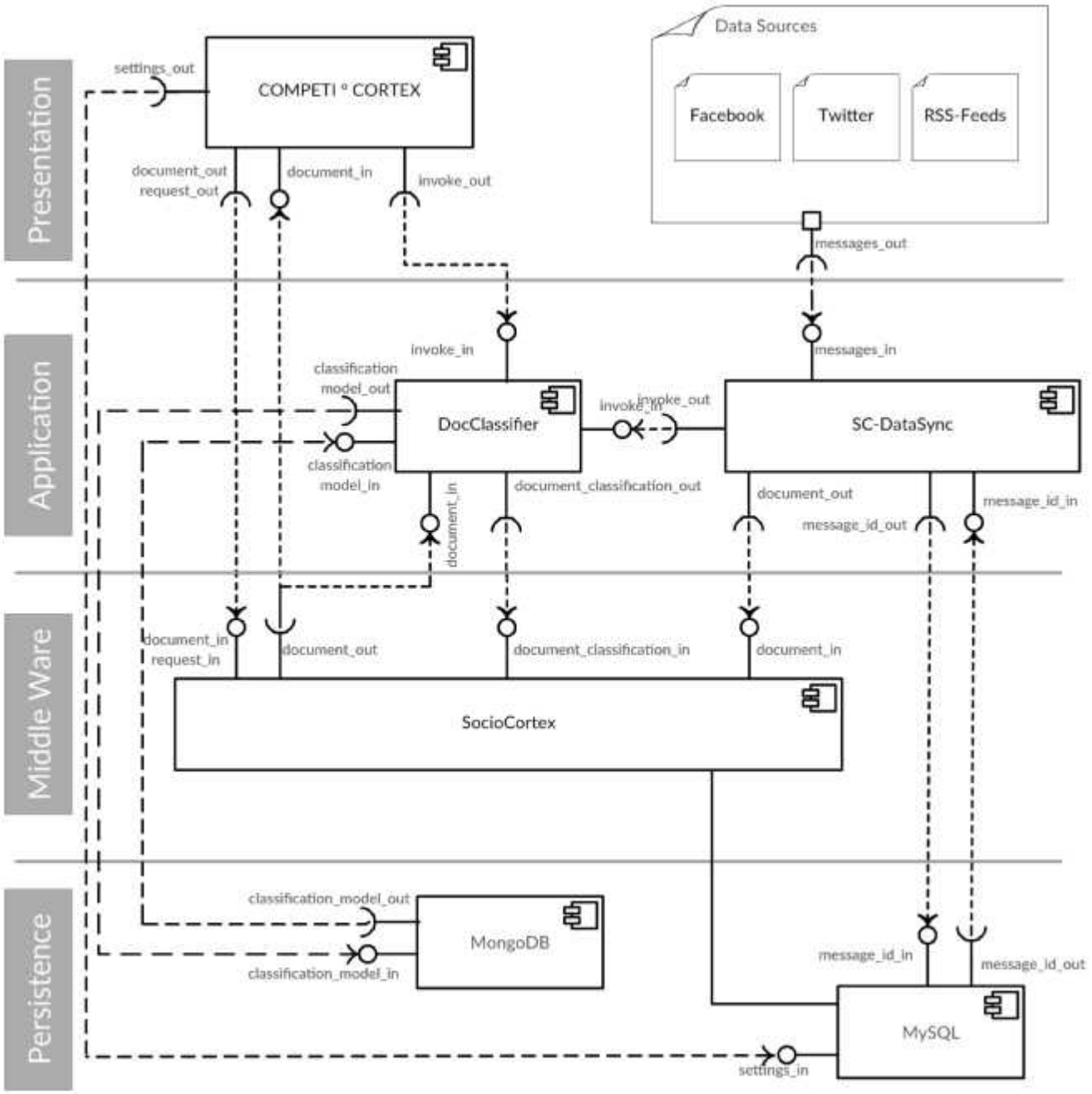
2. Requirements

3. System Design

4. Evaluation
   - System Usability
   - Feedback

5. Future Work

System Design



COMPETI•CORTEX

Demo

Outline

1. Introduction
   - Competitor Analysis
   - Problems
   - Research questions

2. Requirements

3. System Design

4. Evaluation
   - System Usability
   - Feedback

5. Future Work

Evaluation

- A subset of internal documents and messages of external sources used
- 9 test subjects (managers and analysts)
- Introduce system to every person individually
- Test the system
- Evaluation questionnaire (System Usability and Feedback)

Evaluation

System Usability Scale (SUS)

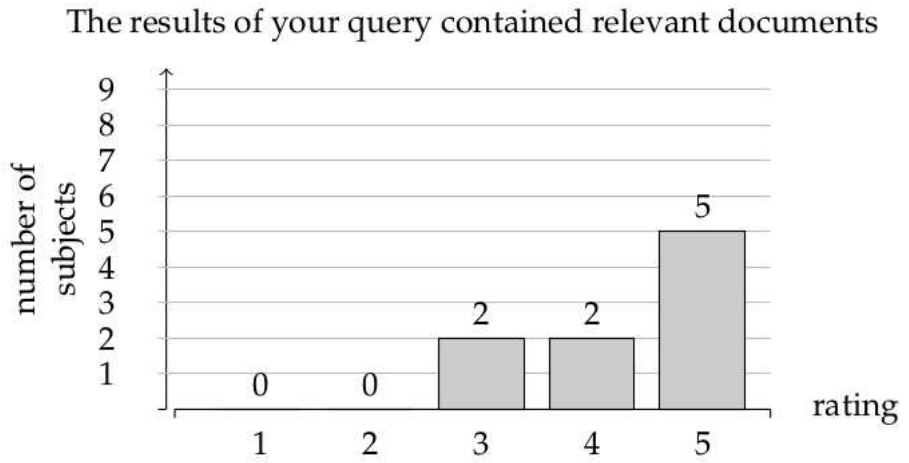
- Ten-item scale [1]
- Illustrating subjective assessments of system usability
- Positional value for each question (1 – 5)
- Calculating SUS-score (0% – 100%)

➔ Exceeds the recommended value of 68% [2]

Source: [1] John Brooke. Sus: A quick and dirty usability scale, 1996.
[2] John Brooke. Sus: A retrospective. Journal of Usability Studies, 8(2):29–40, 2013.

Feedback (1)

- What do you like most about the application?
  - Usability, simplicity & attractiveness of user views (intuitive system handling)
- What should be improved in the current system?
  - Transparency in results' relevance
- Did the result of your query contain relevant documents?



Feedback (2)

- Could you retrieve your uploaded file?
  + PDF, Word, Powerpoint
  - Excel

- Assess whether the efficiency of the information acquisition can be optimized by using the system
  + 7 subjects agree, if document pool is extended and external message currentness is granted
  - 1 subject: Only a subset of internal information sources can be included

Outline

1. Introduction
   - Competitor Analysis
   - Problems
   - Research questions

2. Requirements

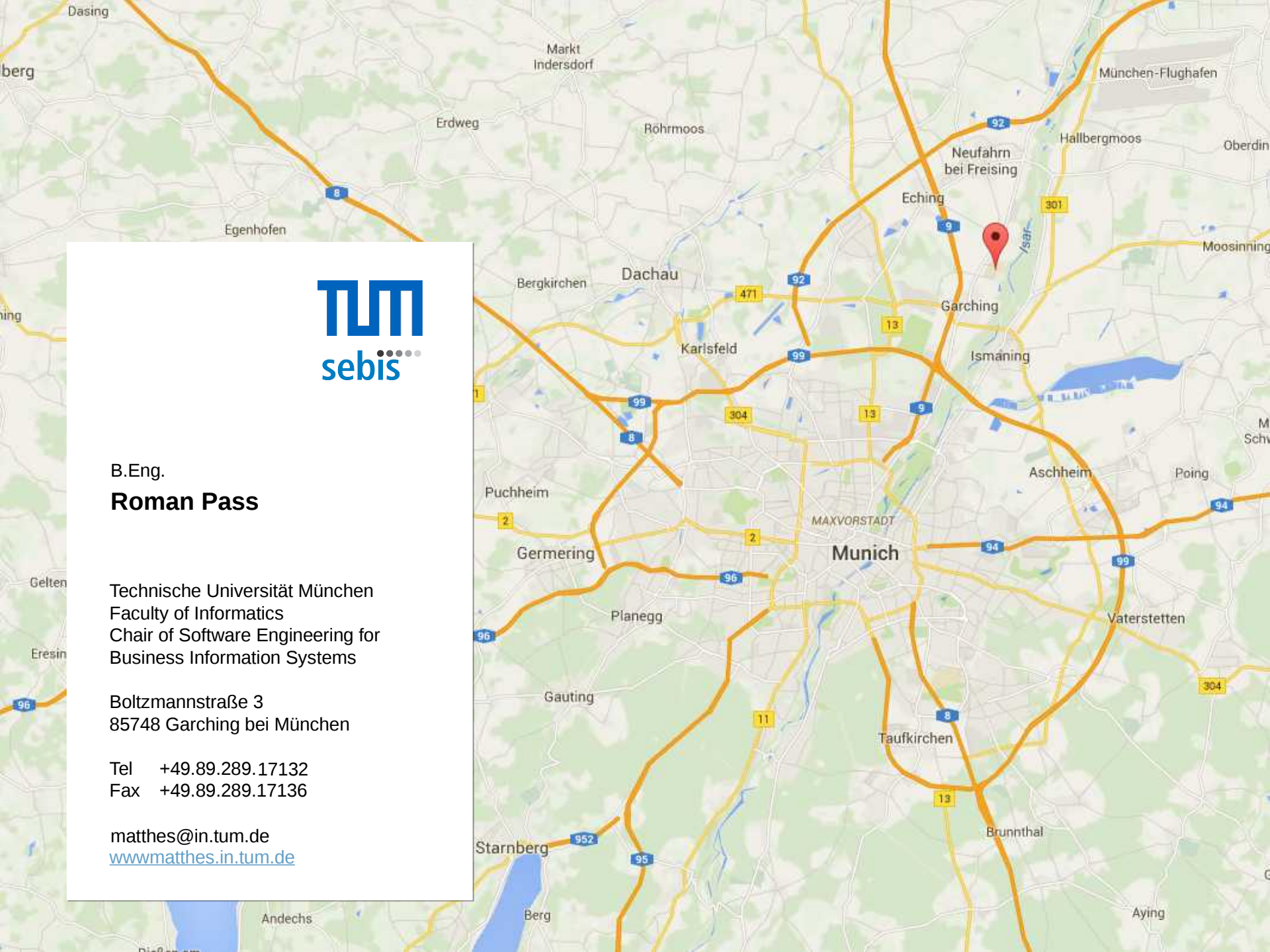
3. System Design

4. Evaluation
   - System Usability
   - Feedback

5. Future Work

Future Work

- Extend information sources → improve automatic document classification
- Provide individuality (user accounts)
- Dictionary
  - Synonyms
  - Translations
- Provide Optical Character Recognition (OCR) for scanned documents



B.Eng.

Roman Pass

Technische Universität München
Faculty of Informatics
Chair of Software Engineering for
Business Information Systems

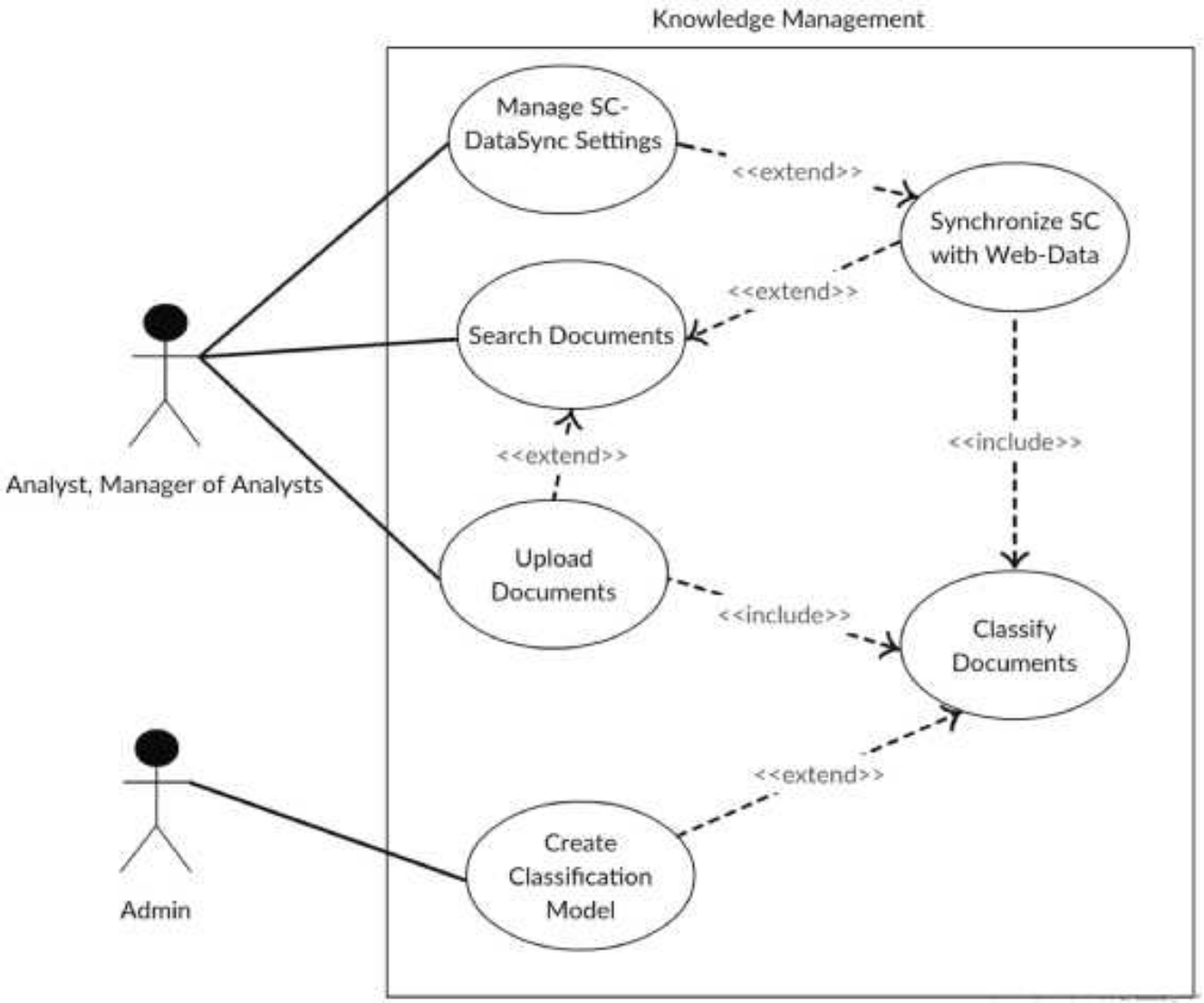
Boltzmannstraße 3
85748 Garching bei München

Tel +49.89.289.17132
Fax +49.89.289.17136

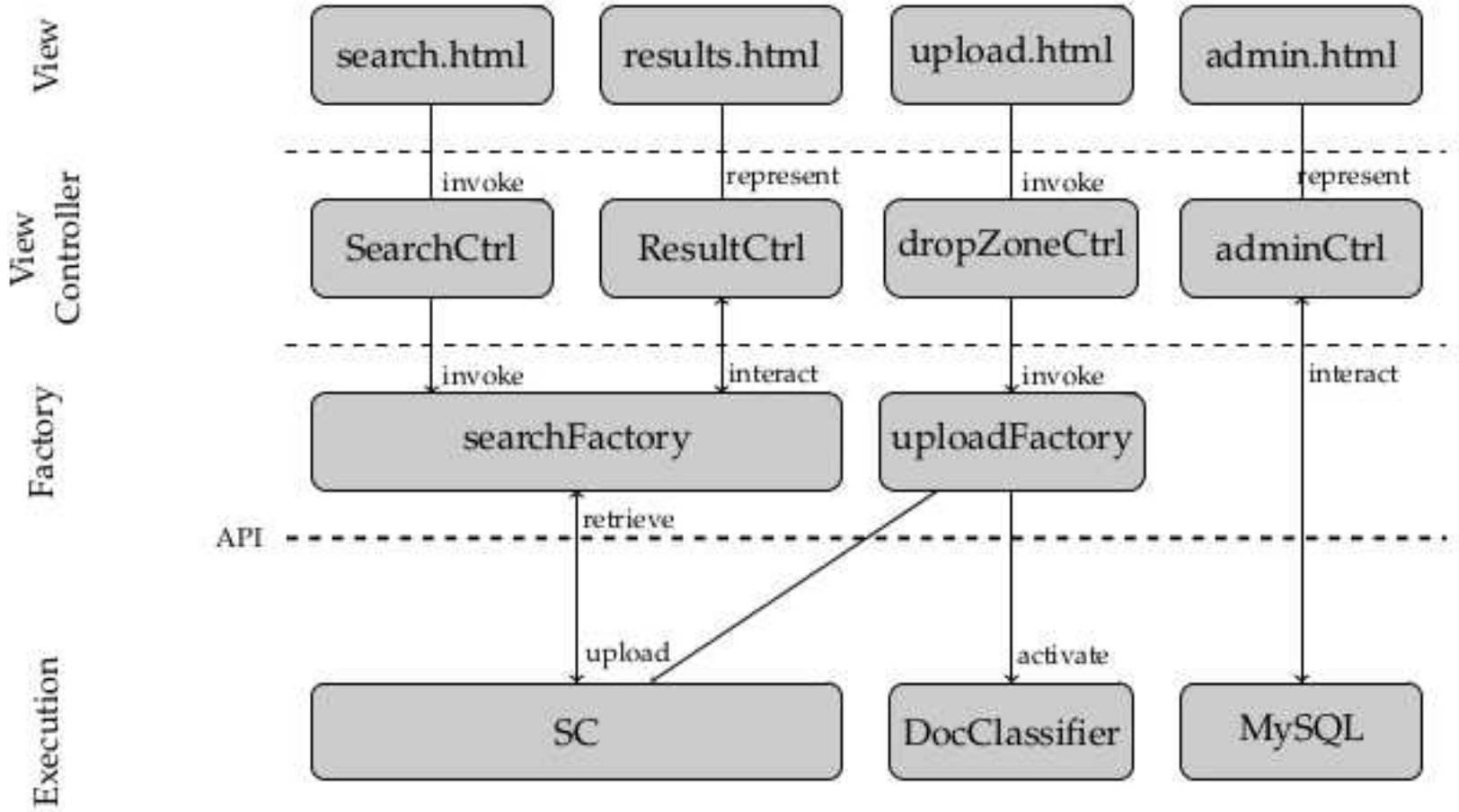
matthes@in.tum.de
wwwmatthes.in.tum.de

Backup

Use Case



CompetiCortex Layers



Comparison Between Classification Algorithms

Classification	accuracy	ø-precision	ø-recall
kNN (k=3)	91,28 %	88,12 %	85 %
Naive Bayes	92,77 %	89,46 %	88,02 %
SVM	47,66 %	7,94 %	16,67 %