

SEBA Lab Kickoff WS2023/24

Felix Hoops

October 2023

Chair of Software Engineering for Business Information Systems (sebis)
Department of Computer Science
School of Computation, Information and Technology (CIT)
Technical University of Munich (TUM)
www.matthes.in.tum.de

Organization Details

Project Evaluation

Project Proposals

Next Steps



Prof. Dr. Florian Matthes
Head of sebis



Juraj Vladika
NLP



Tim Schopf
NLP & Knowledge Graphs



Phillip Schneider
NLP



Anum Afzal
NLP



Stephen Meisenbacher
Privacy & NLP



Mahdi Dhaini
NLP & Explainable AI



Nektarios Machner
NLP



Oliver Wardas
NLP & NLawP



Felix Hoops
Blockchain



Burak Öz
Blockchain

ALMA PHIL

Allianz 

 breathment

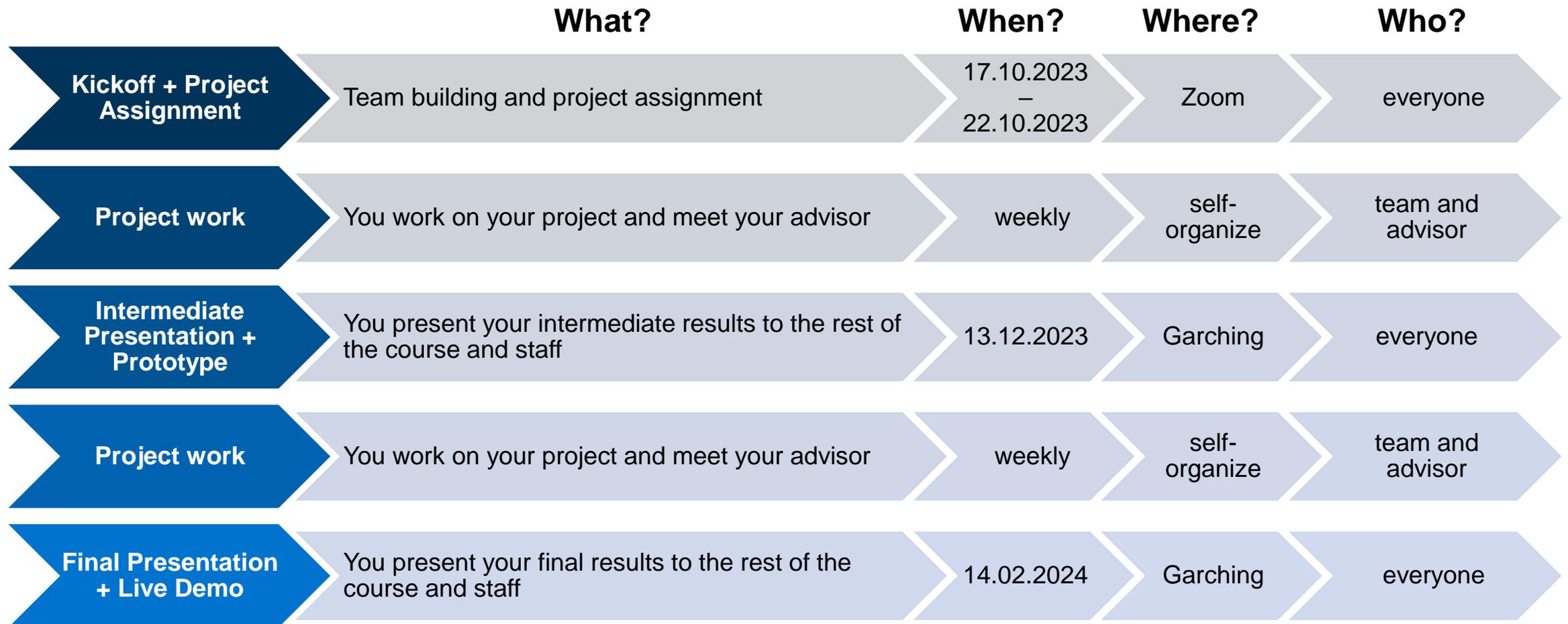
 fusionbase

 EclipseSource

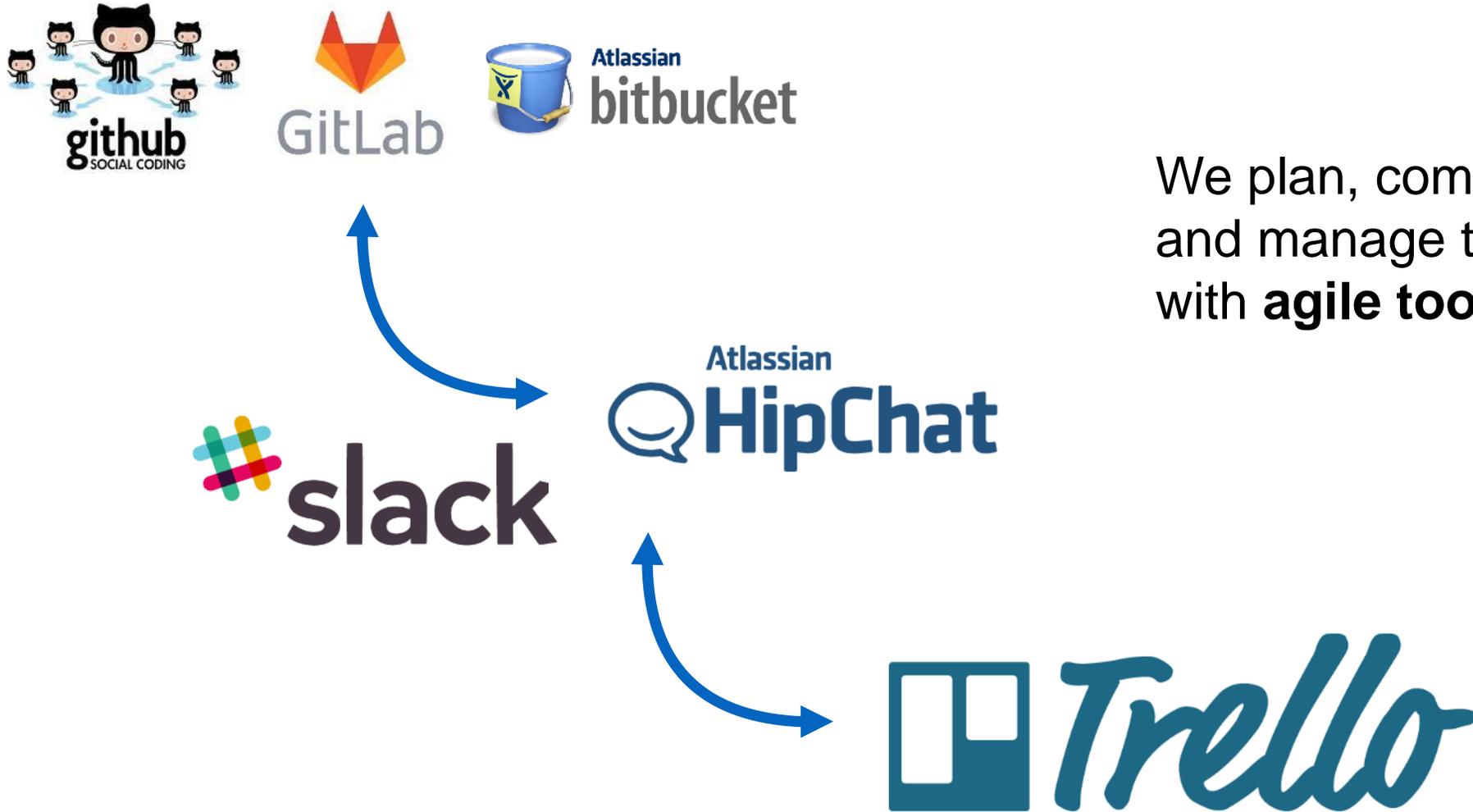
 MOTIUS
WE R&D.

 interhyp

Organization and Timeline



Attendance is mandatory for every student!



We plan, communicate, and manage the project with **agile tools**

Agile Project Management

Each week, we “sprint“ to the next

Daily Meeting/Scrum

Organize yourself and discuss the current state
Everyone talks shortly (timeboxed) about his achievements and challenges

Weekly meeting (approx. 60 minutes)

Discuss in team and with your Scrum Master (Advisor)
the current work in progress

1. Sprint Planning

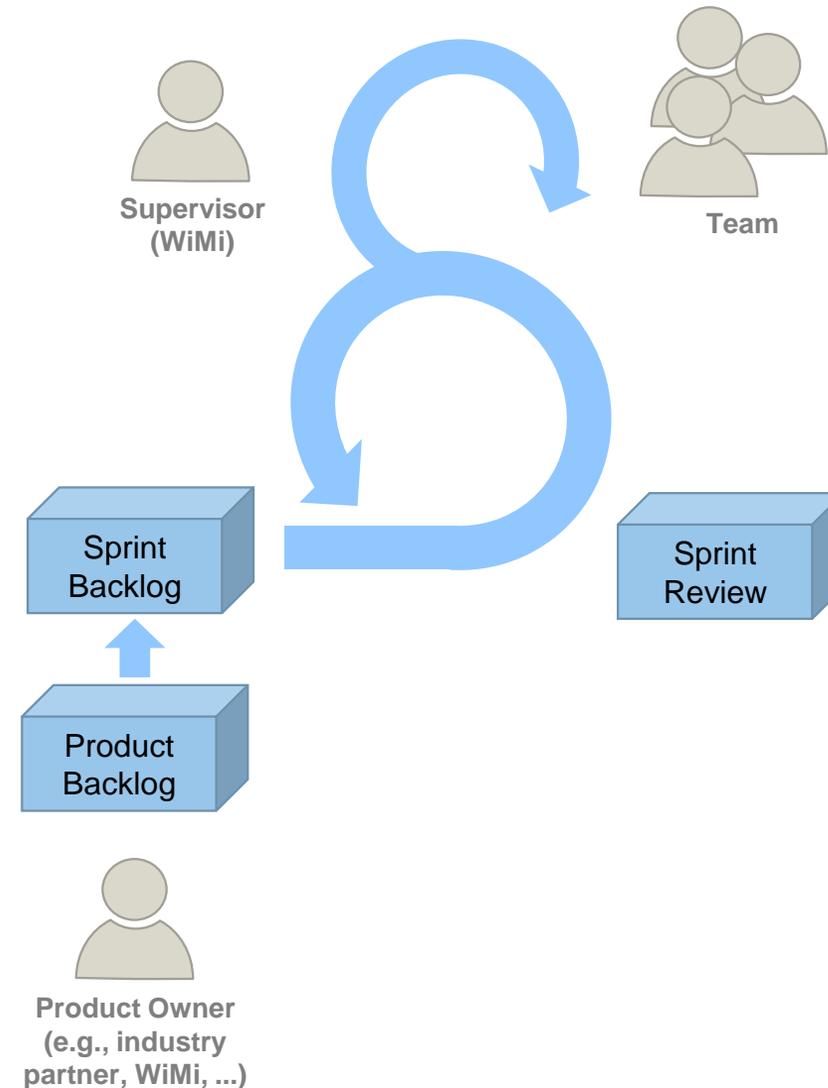
Discuss and define work items for the next sprint

2. Sprint Review

Present the current prototype to the product owner

3. Sprint Retrospective

Reflect upon the progress and effort estimates



How to deal with issues?

In case you encounter a technical or personal issue (e.g. team communication):

1. Try to solve the issue **within your team**. In Scrum every team is **self-organized**.
2. If the issues persists, talk to **your advisor** (Scrum Master or Product Owner).
3. If the issue persists, talk to the **course organizer**: Felix, felix.hoops@tum.de

Visit the course web page regularly

<https://wwwmatthes.in.tum.de/pages/18witnulbiwl6/SEBA-Lab-Course>

News

- Updates will be posted here

Contact

- For questions or any feedback on the course, please contact Felix Hoops.

Registration

- Registration per [Matching System](#) and [Survey](#) between the 14th of July 2023 and the 19th of July 2023.
(Note: The SEBA Lab course is listed as "Entwicklungspraktikum Software Engineering für betriebliche Informationssysteme (IN2106, IN2129)" in the matching system.)
- Only for Master's students!

Content and teaching goals

The Master Lab Course Web Applications is an opportunity for students to work on interesting projects in the field of web applications. Students will collaborate in small teams and implement a web application over the course of the term. Each team is advised by one teaching assistant of the chair.

The goals of the lab course are:

- Deepen your knowledge from the SEBA Master course
- Get familiar with new technologies such as blockchain, web3/web5, NoSQL databases, Cloud Computing, REST APIs, React.js, NLP, ML, etc.
- Get more practical experience in application development
- Collect teamwork experience and practice presentation skills in English
- Participate in current research projects at sebis and collaborate with industry partners

Schedule

Date	Time	Place	Topic	Attendance
13.07.2023	10:00 - 11:00	Zoom	Preliminary Discussion	voluntary
17.10.2023	10:00 - 12:00	Zoom (link will be on TUMonline)	Kickoff Meeting - Project Proposals	mandatory
Weekly meetings (on project team level)				
13.12.2023	10:30 - 16:30	02.5901.013	Intermediate Presentations	mandatory
Weekly meetings (on project team level)				
14.02.2024	9:30 - 16:30	00.04.011, MI Hörsaal 2	Final Presentations	mandatory



Outline



Organization Details

Project Evaluation

Project Proposals

Next Steps



Application

75%

- User-Centered Design
- Documentation (10-30 pages)
- Code Quality
- Team Work & Consistency



Presentation

25%

- Content
- Structure & Style
- Time Management
- Quality of the Answers

For successful completion of the course, both examination modules have to be passed!

Outline

Organization Details

Project Evaluation

Project Proposals

Next Steps

Project Proposals



#	Project	Advisor	sebis	Industry Partner	Advisor	Industry Partner	Email
1	ALPHA-KI: Health Intelligence Platform	Phillip		Alma Phil	Lutz Frick		lutz.frick@almaphil.com phillip.schneider@tum.de
2	Enhancing IVR Systems with LLM Integration	Nektarios, Phillip		Allianz	Ömer Uludag Daniel Faisst		nektarios.machner@tum.de omer.uludag@allianz.de daniel.faisst@allianz.de
3	An AI assistant for web-based IDEs for project-specific assistance	Anum		EclipseSource	Jonas Helming		anum.afzal@tum.de jhelming@eclipsesource.com
4	Breathment: Web-based Teletherapy Application	Burak		Breathment	Elçin Can Çavuşoğlu		elcin.cavusoglu@breathment.com burak.oez@tum.de
5	CD4AI: The Web App	Tim, Stephen		fusionbase	Patrick Holl		tim.schopf@tum.de stephen.meisenbacher@tum.de patrick.holl@fusionbase.com
6	Natural Language Processing Knowledge Graph	Tim		-	-		tim.schopf@tum.de
7	ChatHyp: Providing mortgage information in an easy way	Anum		Interhyp	Francisco De las Casas Young Felipe Wieman		anum.afzal@tum.de francisco.delascasasyoung@interhyp.de felipe.wieman@interhyp.de
8	Cliq – The Social Network	Felix		Motius	Zied Bahrouni Christoph Kipfer		felix.hoops@tum.de zied.bahrouni@motius.de christoph.kipfer@motius.de
9	A Data Exploration Tool for Blockchain-based Systems	Burak, Filip		-	-		burak.oez@tum.de filip.rezabek@tum.de
10	AI Employment Contract Analysis	Oliver		SYLVENSTEIN Rechtsanwälte	Dr. Dominik Herzog Domenic Böhm		oliver.wardas@tum.de
11	XNLP – Explanation Tool for NLP	Mahdi		-	-		mahdi.dhaini@tum.de
12	Synthesizing Evidence-Based Answers	Juraj		-	-		juraj.vladika@tum.de
13	Privacy Analytics with Differentially Private Text Rewriting	Stephen		-	-		stephen.meisenbacher@tum.de

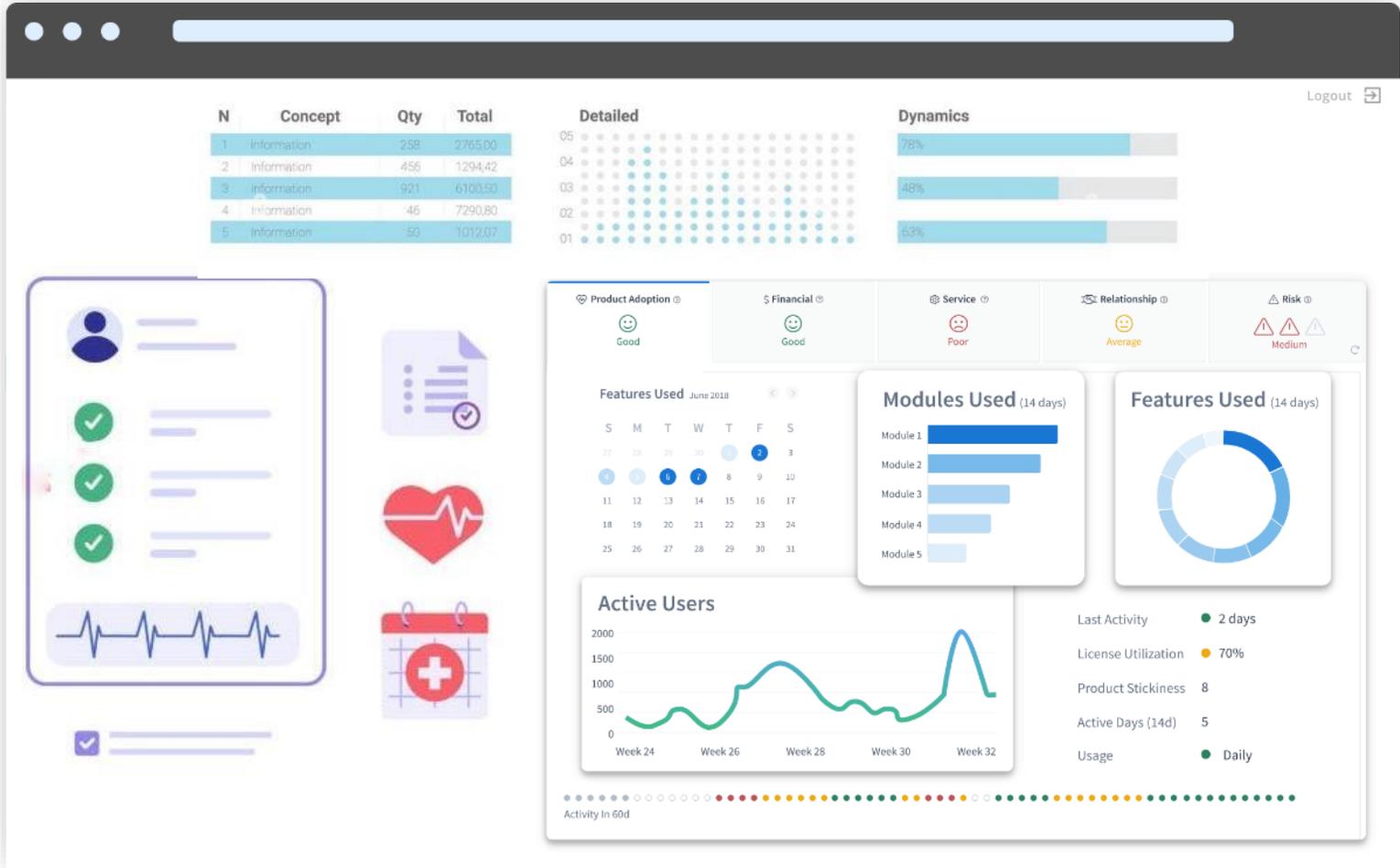
ALPHA-KI: Health Intelligence Platform

Advisor: Phillip Schneider

Product Owner: Lutz Frick

ALMA PHIL

ALPHA-KI: Health Intelligence Platform – Connecting Dots: From Data to Discovery

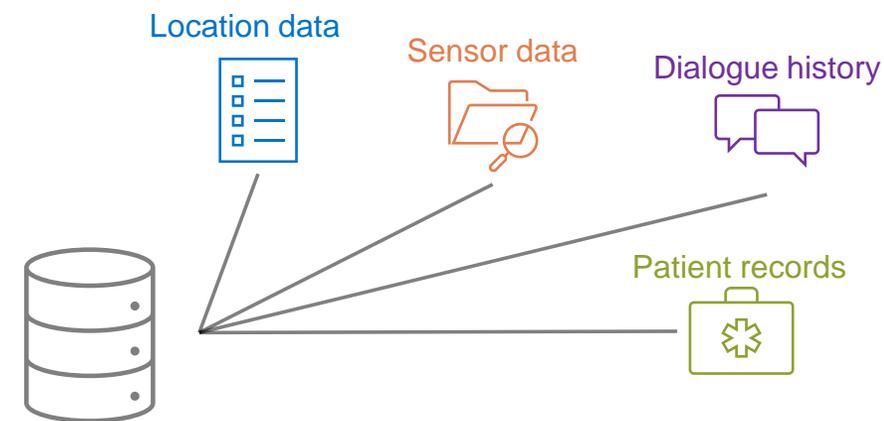


Outcome: Platform that supports:

- ✓ Intelligent integration of health data
- ✓ Multi-level information visualization
- ✓ NLP-Based generation of health insights: patient- and problem-centered



(technology stack is flexible)



Pitch: The goal of the project is to develop a health intelligence dashboard for analyzing multi-level patient data extracted from various sources from a distributed platform that implements a digital health assistant.

Basic functional requirements:

- Extract and integrate patient data from various sources (e.g., wearables, health records, voice assistant) without comprising scalability and performance
- Perform visual analyses of patient data at different levels, (e.g., patient- and technology-centered)
- Detection of trends or abnormalities and automatic summarization of health reports with large language models

Basic non-functional requirements:

- User-friendly UI, allowing developers and health professionals to intuitively access relevant information
- Modular system that is easy to maintain, with well-organized code and documentation
- Platform must adhere to industry regulations and standards related to healthcare data privacy and security

Expected prior knowledge

- Knowledge in Flask, React, JavaScript, HTML, CSS
- Strong programming skills in Python and basic knowledge in Natural Language Processing
- Knowledge in data processing and ETL
- Experience in visualization libraries like Plotly, Dash, or Grafana

Bonus

- Gain knowledge about conversational agents and the future of digital healthcare
- Get guidance and feedback from experienced industry professionals

Contact:

lutz.frick@almaphil.com

phillip.schneider@tum.de

Enhancing IVR Systems with LLM Integration

Advisor: Nektarios Machner
Phillip Schneider

Product Owner: Ömer Uludag
Daniel Faisst



Enhancing Interactive Voice Response (IVR) Systems with Large Language Model (LLM) Integration

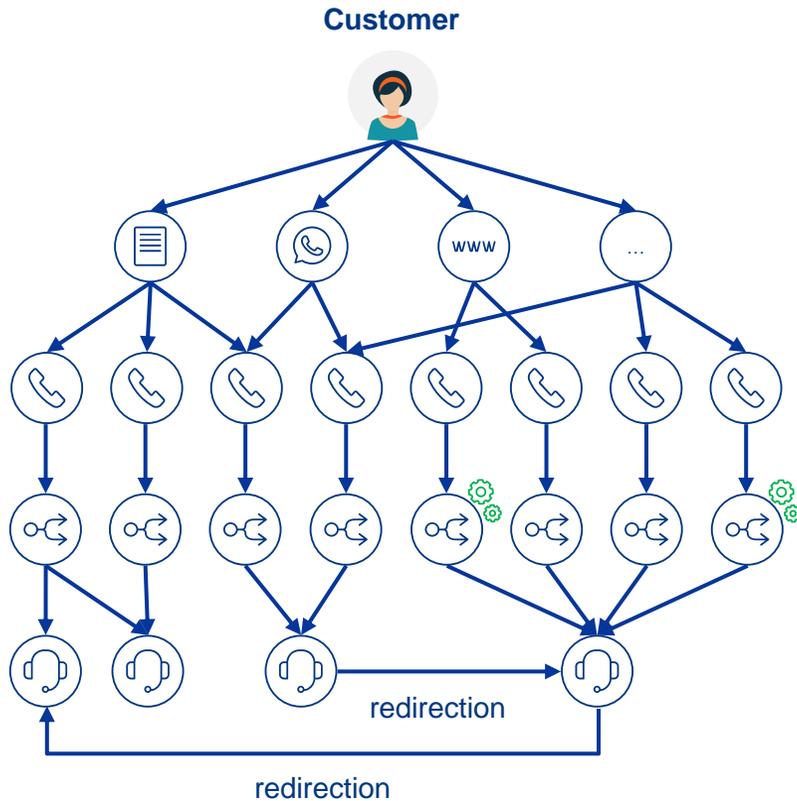
Today

Customer channels

Call types

IVR systems

Allianz agents



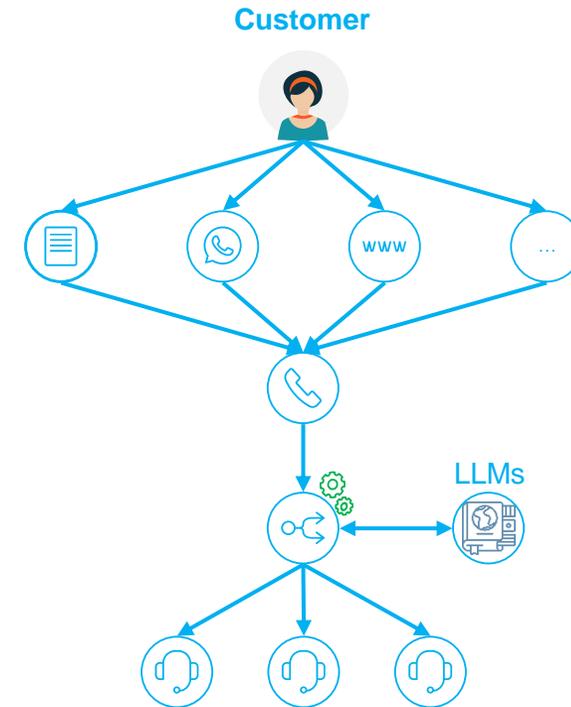
Our Vision

Customer channels

Call type

Intelligent routing

Allianz agents



Outcome:
 ✓ A concept and a prototype that demonstrates a LLM-powered IVR enhancement

Technologies:

Pitch: The goal of the project is (i) to explore the status quo of the usage of large language models (LLMs) in industry and at the Allianz group and (ii) to build a concept and prototype based on LLMs to improve the customer experience in Allianz's customer care services by simplifying/replacing the existing interactive voice response (IVR) landscape.

Basic functional requirements:

- Investigate the status quo regarding the usage of LLMs in industry, preferably in the context of customer care services, and at Allianz
- Evaluate the identified LLMs based on various aspects (adoption scenarios, costs, legal, language, etc.)
- Analyze Allianz's IVR landscape and develop a concept for using LLMs for an improved routing of Allianz customers
- Develop a prototype that implements the designed concept

Basic non-functional requirements:

- Development of an easy-to-understand concept for the integration of LLMs to simplify/replace the existing interactive voice response (IVR) landscape.
- Development of a reusable prototype (based on a build-measure-learn-approach)
- Excellent code documentation

Expected prior knowledge

- Good programming skills and fundamental understanding of machine learning and its various approaches
- Basic understanding of Large Language Models (LLMs) and related APIs (e.g., OpenAI)
- Experience with common ML frameworks such as PyTorch or TensorFlow.
- Solid knowledge of at least one programming language such as Python or Javascript.
- Experience with Vector Databases (Weaviate / Pinecone) and Cognigy beneficial but optional.

Bonus:

- Work on a real-world use case with a high relevance and impact for the Allianz group
- First-hand experience in software engineering with LLMs

Contact:

omer.uludag@allianz.de

daniel.faisst@allianz.de

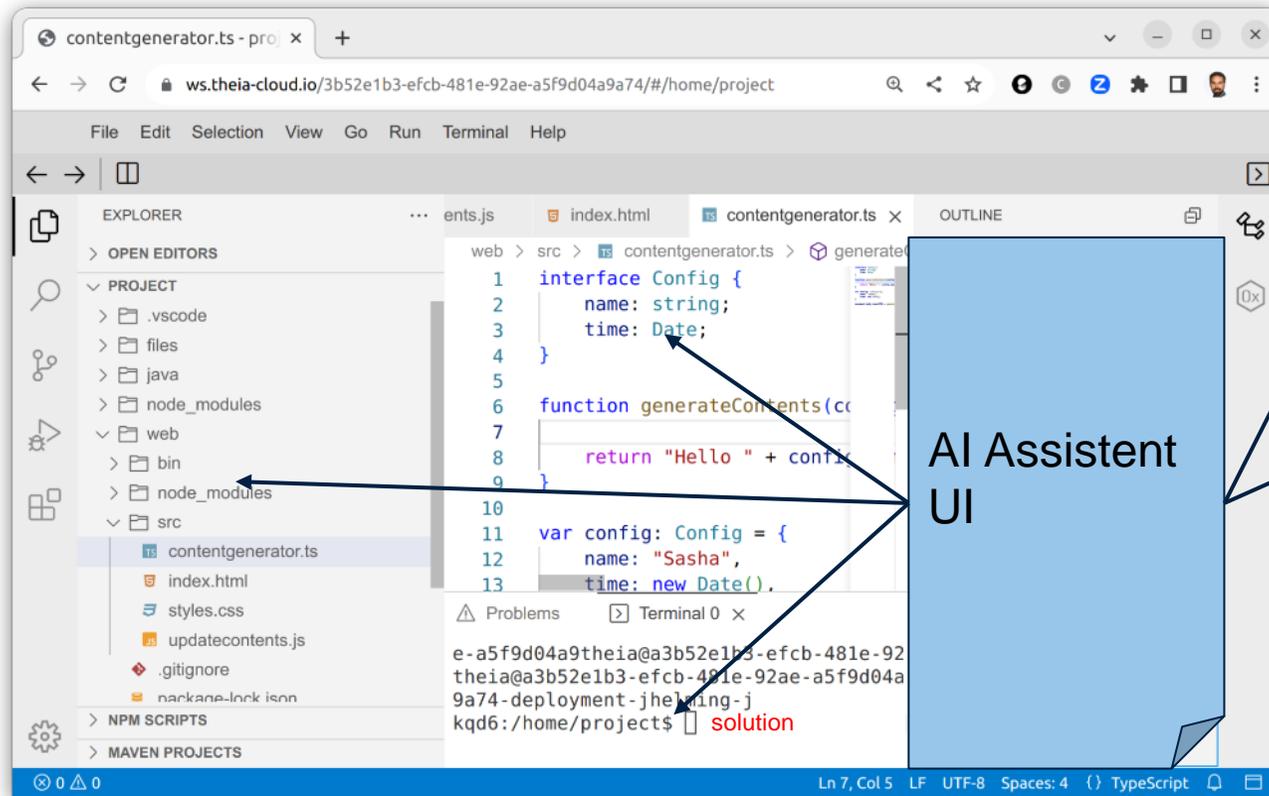
An AI assistant for web-based IDEs for project-specific assistance

Advisor: Anum Afzal

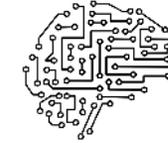
Product Owner: Jonas Helming



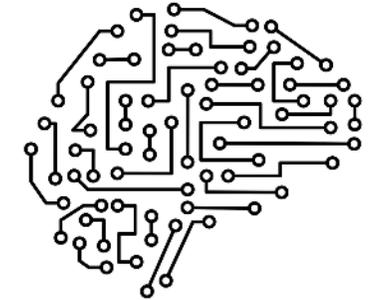
An AI assistant for web-based IDEs for project-specific assistance



Project context



Generic AI Model



User context



AI Assistant UI

Outcome: App that supports:

- ✓ Supporting developers in the IDE via AI
- ✓ Allow experts to fine tune “on the fly”

Technologies:



Pitch: The goal of the project is to integrate AI into a web-based IDE so that developers get assisted in their development tasks and can also fine-tune the AI with project-specific knowledge.

Basic functional requirements:

- Get answers to contextual questions
- Get help with errors (e.g. in the terminal)
- Get assisted in the IDE, e.g. by typing commands
- Improve AI by providing project-specific knowledge

Basic non-functional requirements:

- All code will be contributed under an Open Source license (EPL+MIT)
- Abstraction layer for the underlying AI
- Reproducible build and set-up

Expected prior knowledge

- Knowledge in HTML, CSS, TypeScript and React
- Basic Knowledge in AI
- Basic understanding of web- and cloud based tools (e.g. VS Code)
- Standard “toolbox” including Git and VS Code

Contact:

jhelming@eclipsesource.com

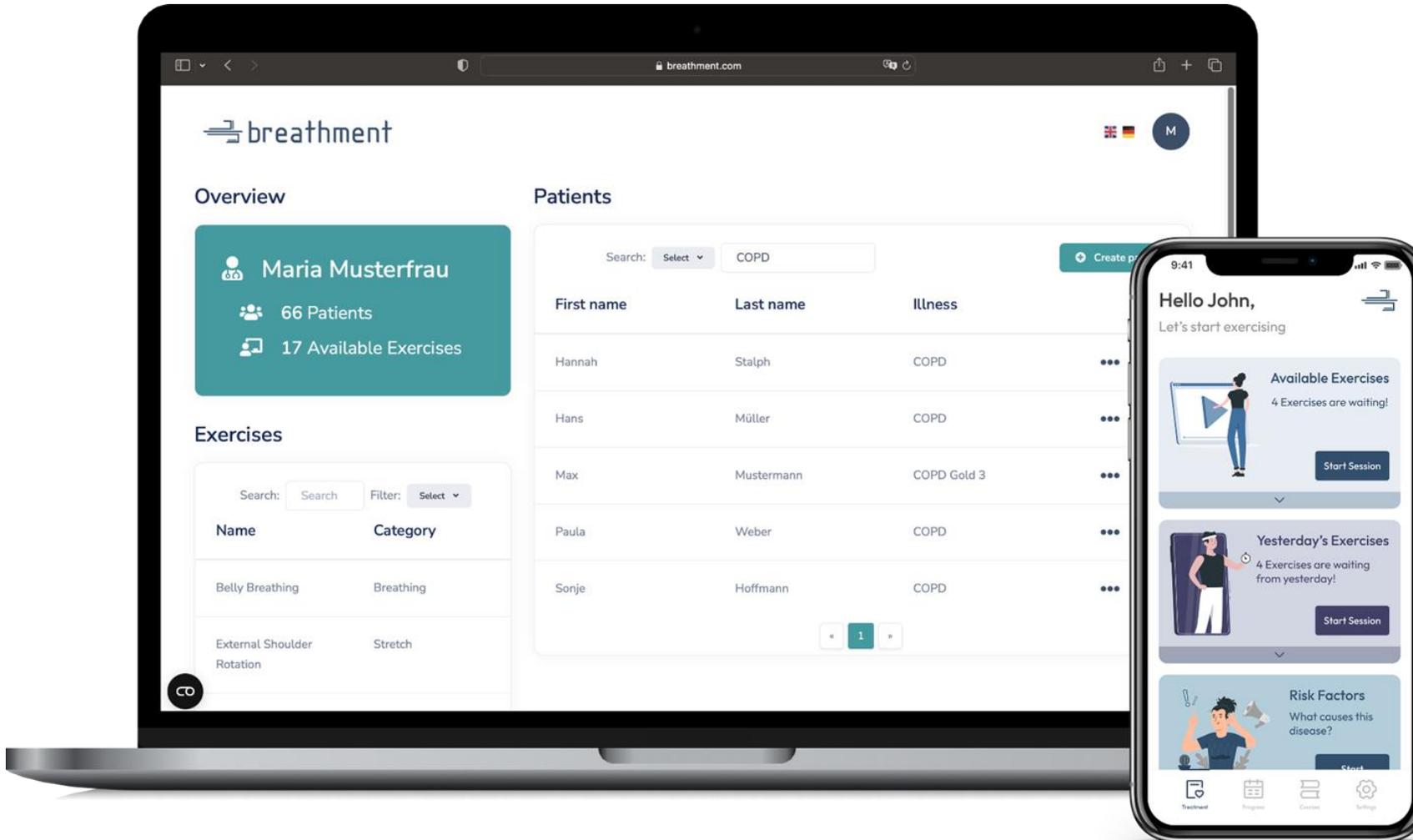
Breathment – Remote Patient Monitoring Application

Advisor: Burak Öz

Product Owner: Elçin Can Çavuşoğlu



Breathment – Remote Patient Monitoring Application



Outcome: Integrated video-based vital sign monitoring for a web & mobile app that supports:

- ✓ Detection of physiological signs
- ✓ Visualization of these measures
- ✓ Alert notifications in case of critical conditions
- ✓ Comparison & progress over time

Technologies:

 ANGULAR



React Native



Pitch: The goal of the project is to build an integrated component to an existing web & mobile application that provides video-based vital sign monitoring.

Basic functional requirements:

- Detection of physiological signs using device camera
- Detection of physiological signs using hardware devices
- Visualization of physiological measures
- Alert notifications in case of critical conditions
- Comparison of different dates
- Progress of patient over time

Basic non-functional requirements:

- Real time client – server communication
- Design of modular components

Expected prior knowledge

- Knowledge in a JavaScript frontend framework (preferably Angular 2+ and React Native), JavaScript/TypeScript, HTML and CSS
- Basic knowledge in rest APIs
- Basic knowledge in NoSQL databases (preferably MongoDB)

Contact:

<https://breathment.com>

elcin.cavusoglu@breathment.com

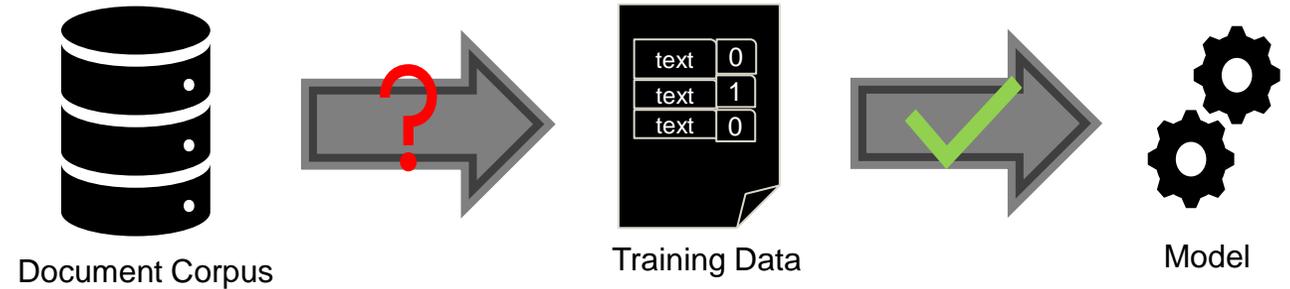
CD4AI: The Web App

Advisor / Product Owner: Tim Schopf
Stephen Meisenbacher
Patrick Holl



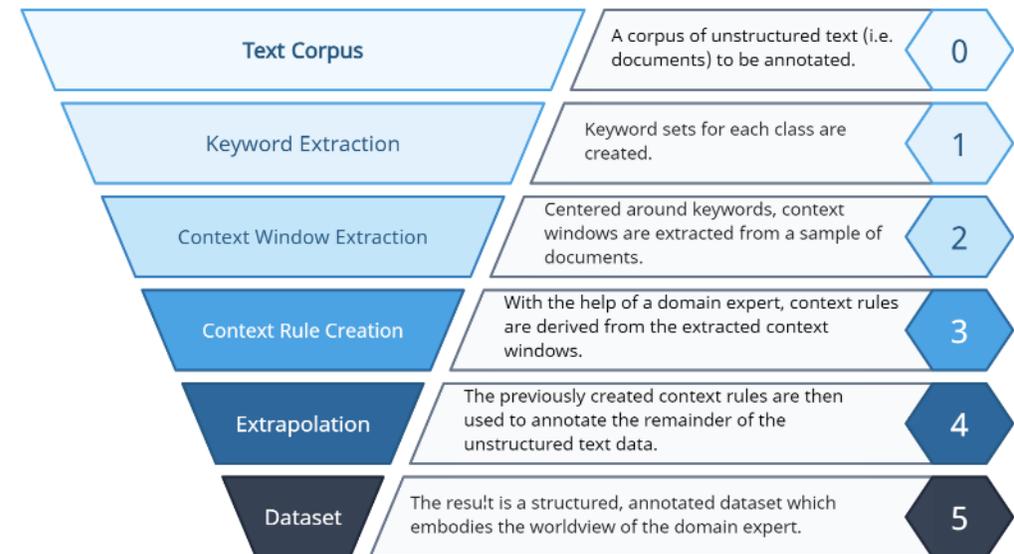
CreateData4AI - Motivation

- Data is today's currency
- AI models (think LLMs) are data hungry
- But ~80% of data is unstructured (e.g., text!)
- How do we obtain meaningful annotated data from unstructured text in an efficient, yet accurate way?



CreateData4AI project at sebis

- In year 1/3
- Performed in collaboration with 
- Goal:
 - Created structured, annotated datasets from unstructured text corpora
- Approach:
 - Proposed pipeline to the right



Introduce a query

Predictions Annotations Status Metadata Sort NEW

LANGUAGE [1] ORDINAL [2] FAC [3] DATE [4] TIME [5] LOC [6] CARDINAL [7] PERSON [8]

user: Select options

Search...

annotator1 (33)

annotator2 (33)

annotator3 (33)

Cancel Filter

Find similar ...

Mode

Metrics

Refresh

EXAMPLE

Discarded

Would I wish to send up my name now ? Again I declined , to the polite astonishment of the concierge , who evidently considered me a queer sort of a friend . He was called to his desk by a guest , who wished to ask questions , of course , and I waited where I was . At a quarter to eleven **Herbert Bayliss** emerged from the elevator . His appearance almost shocked me . Out **late the night** before ! He looked as if he had been out **all night** for many nights .

Validate Discard Clear Reset

It might have been by Gad , but it was by **Worth** . **Four** shades of grey , and pearls . Mrs. **Beaumont** distinctly thought that this was not the sort of dress to dash into the faces of a quiet country party . It was like letting off

Records per page: 10

< Prev 1 2 3 ... 10 Next >

Outcome: App that supports:

- ✓ Domain experts in defining class-specific keywords and descriptions
- ✓ Recommendation of keywords and enhanced descriptions
- ✓ Encapsulation of the CD4AI project

Technologies:

Pitch: The goal of the project is to develop a web application that supports domain experts in defining keywords and descriptions in order to conceptualize classes for structured datasets.

Basic functional requirements:

- Domain experts can create profiles and define classes using keywords and textual descriptions
- Domain experts get recommendations for further keywords and suggestions for enhanced description based on their initial definitions

Basic non-functional requirements:

- Interactive, user-friendly, and responsive
- Low latency with API calls / LLM outputs
- Consistent color scheme with CD4AI project
- Clean and commented code base, allowing for future extensions and improvements

Expected prior knowledge

- Knowledge in React.js, Vue.js, and Python
- Basic knowledge in NLP beneficial
- Experience with data pipelines a plus

Contact:

tim.schopf@tum.de

stephen.meisenbacher@tum.de

patrick.holl@fusionbase.com

Natural Language Processing Knowledge Graph (NLP-KG)

Advisor / Product Owner: Tim Schopf



Natural Language Processing – Knowledge Graph

NLP-KG

Search Publications, Researchers, Fields of Study... Search

Top Fields of Study

Sort by Publication

35656 Semantic Text Processing 25976 Text Generation 21808 Language Models 21686 Responsible & Trustworthy NLP 21668 Information I

Top Publications

Sort by Citation

Attention Is All You Need
Ashish Vaswani, Noam M. Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, Illia Polosukhin • arXiv • 12 June 2017
TLDR: We propose a new simple network architecture, the Transformer, based solely on attention mechanisms, dispensing with recurrence and convolutions entirely.
Show Abstract
58,331 arXiv Semantic Text Processing Language Models Machine Translation Text Generation Multilinguality

BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding
Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova • Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers) • 01 June 2019
TLDR: We introduce a new language representation model called BERT, which is designed to pre-train deep bidirectional representations from unlabeled text by jointly conditioning on both left and right context in all layers.
Show Abstract
52,660 DOI Semantic Text Processing Natural Language Interfaces Language Models Question Answering

Top Researchers

Sort by Citation

Christopher D. Manning
#publications 248
h-index 109
citations 113666

Tomas Mikolov
#publications 24
h-index 19
citations 85541

Noam M. Shazeer
#publications 24
h-index 20
citations 73998

Lukasz Kaiser
#publications 17
h-index 16
citations 69645

Kenton Lee
#publications 39

Outcome: App that supports:

- ✓ Personalized profiles for organization and sharing of research papers
- ✓ Personalized recommendation engine for relevant papers

Technologies:

Pitch: The goal of the project is to extend an existing web application for NLP paper search with personalized profile and recommendation features.

Basic functional requirements:

- Users can manage their profiles to organize and share lists of NLP papers, assisted by ChatGPT
- Development of a personalized recommendation engine of papers based on the interest of users
- Additional features for paper exploration

Basic non-functional requirements:

- Intuitive profile design
- Useful recommendations
- Realtime client – server communication
- Design of modular components

Expected prior knowledge

- Knowledge in Next.js, TypeScript, and Python
- Basic Knowledge in Neo4j and vector databases beneficial

Contact:

tim.schopf@tum.de

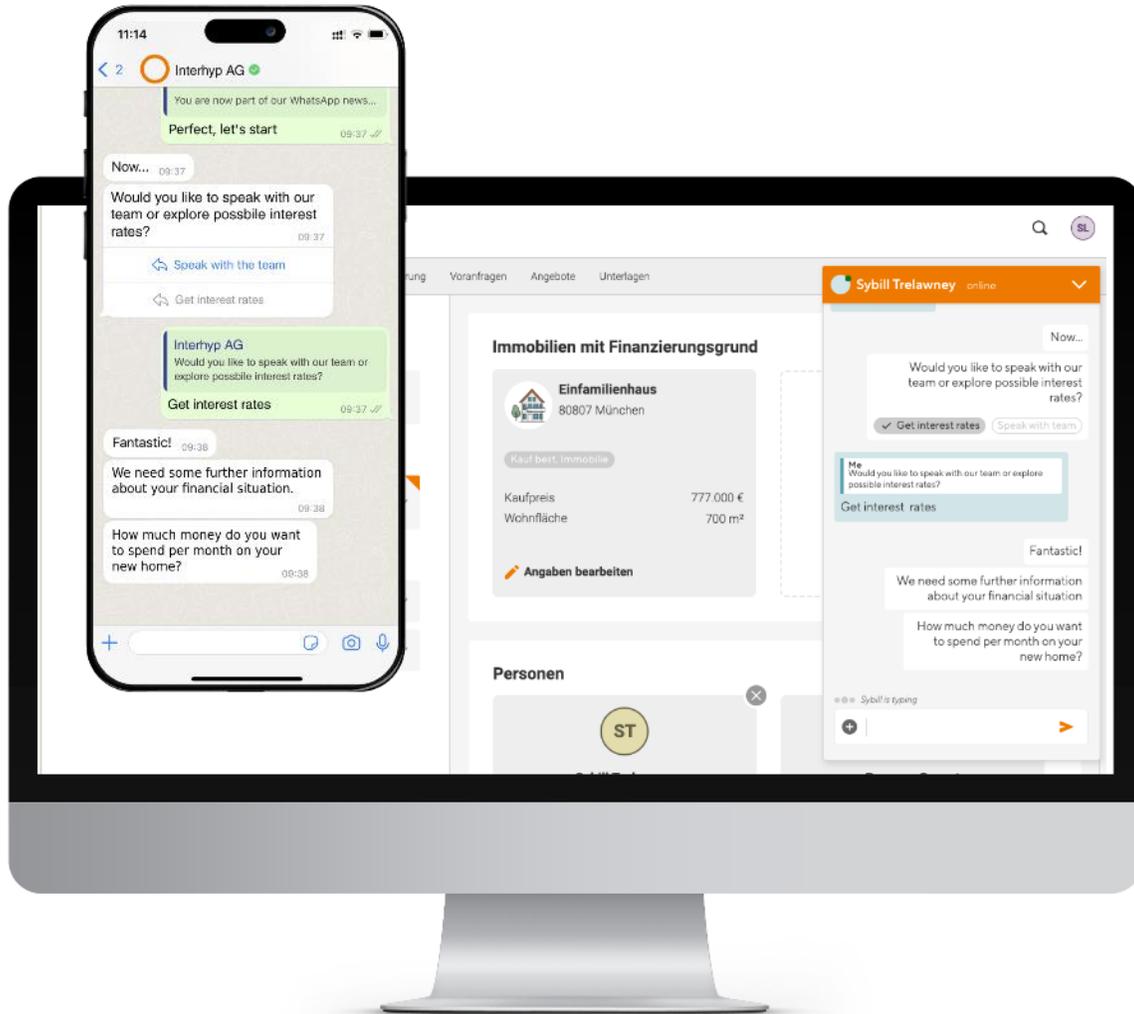
ChatHyp: Providing mortgage information in an easy way

Advisor: Anum Afzal

Product Owner: Francisco De las Casas Young
Felipe Wieman



ChatHyp - Providing mortgage information in an easy way



Outcome – Web platform that supports:

- ✓ Integration of communication and financing API
- ✓ Finding the mortgage twin for each customer
- ✓ Realtime chat function with more features than just texting
- ✓ Reactivate customer through chat - CRM
- ✓ Personalized mortgage recommendations

Technologies:



Pitch: The goal of the project is to develop a web platform to provide mortgage information and services in an easy way to our customers via WhatsApp.

Basic functional requirements:

- Personalized recommendations for mortgages based on our collected data (Machine Learning)
- Enabling our consultants to chat with the customer via the web platform (Twilio)
- Communication with customer via WhatsApp (Appointment, CRM, Document Upload)

Basic non-functional requirements:

- Intuitive, performant and real time user interface
- Design of modular and reusable components
- Security (2FA, Access token)

Expected prior knowledge

- Knowledge in Frontend Technologies
 - HTML, CSS, JavaScript, TypeScript, React
- Knowledge in Backend Technologies
 - Java, Node, Kotlin, Spring Boot
- Bonus: Knowledge of Twilio API and Machine Learning

FYI: If you want to choose another technology that is not listed here, just let us know! :)

Contact:

francisco.delascasayoung@interhyp.de
felipe.wieman@interhyp.de

CLIQ – The Social Network

Advisor: Felix Hoops

Product Owner: Zied Bahrouni

Christoph Kipfer



Cliq - social capital network

Motius

5+

Downloads

12

USK: Ages 12+ ⓘ

Install

Share

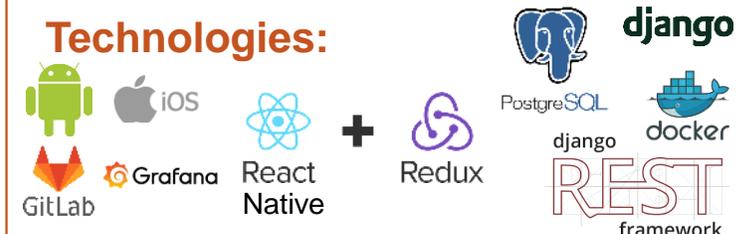
Add to wishlist



Outcome: App that supports:

- ✓ Forwarding of requests
- ✓ Real-time processing & transactions
- ✓ Network path-finding
- ✓ Conceptualizing suitable user flow

Technologies:



Pitch: The goal is to enable the "cross-clique" functionality, allowing users to act as "brokers" by mediating transactions between two cliques.

Basic functional requirements:

- Enable forwarding of Requests (either Favor or Call to Actions) from one clique to another.
- Allow users to mediate transactions between two cliques.
- Conceptualizing a suitable user flow and designing frontend elements

Basic non-functional requirements:

- Real-time forwarding of requests.
- Scalability to handle multiple transactions across various cliques.
- Reliable "Path Finding" mechanism to efficiently determine the best route for requests.

Expected prior knowledge

- Knowledge in Django, Django Rest Framework, PostgreSQL, Docker.
- Experience in working on published apps and an existing code-base would be beneficial.
- Understanding of network pathfinding algorithms.
- Experience in real-time data processing and transaction handling.

Contact:

- zied.bahrouni@motius.de
- christoph.kipfer@motius.de

TUMChainBook - A Data Exploration Tool for Blockchain- based Systems

Advisor / Product Owner: Burak Öz
Filip Rezabek



TUMChainBook - A Data Exploration Tool for Blockchain-based Systems

jupyter tutorial Last Checkpoint: 3 minutes ago (autosaved) Logout

File Edit View Insert Cell Kernel Widgets Help Trusted Python 3

PyCon 2018: Using pandas for Better (and Worse) Data Science

GitHub: <https://github.com/justmarkham/pycon-2018-tutorial>

```
In [1]: import matplotlib.pyplot as plt
import pandas as pd
pd.__version__
```

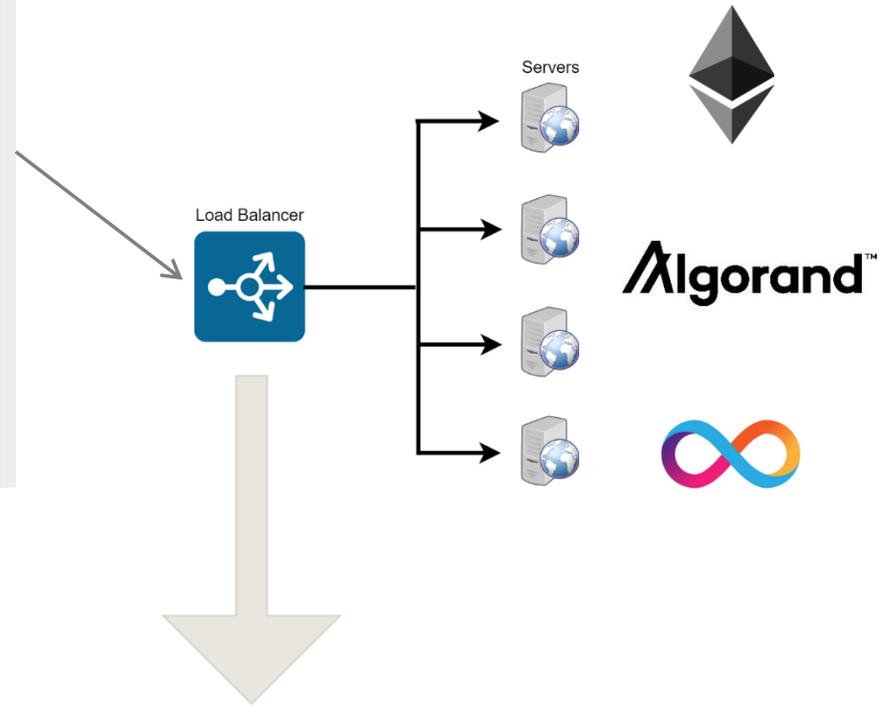
Out[1]: '0.24.1'

Dataset: Stanford Open Policing Project (video)

```
In [2]: # ri stands for Rhode Island
ri = pd.read_csv('police.csv')
```

```
In [3]: # what does each row represent?
ri.head()
```

	stop_date	stop_time	county_name	driver_gender	driver_age_raw	driver_age	driver_race	violation_raw	violation	search_
0	2005-01-02	01:55	NaN	M	1985.0	20.0	White	Speeding	Speeding	
1	2005-01-18	08:15	NaN	M	1965.0	40.0	White	Speeding	Speeding	
2	2005-01-23	23:15	NaN	M	1972.0	33.0	White	Speeding	Speeding	
3	2005-02-20	17:15	NaN	M	1986.0	19.0	White	Call for Service	Other	



Caching of results
 Storing additional data
 If not present, contact the server endpoints

- Outcome:** App that supports:
- ✓ Data Visualization
 - ✓ Realtime result change
 - ✓ Complex Event Processing
 - ✓ Data aggregation and querying

Preferred Technologies:



Pitch: Our project aims to create a **comprehensive platform for executing and analyzing on-chain and off-chain data in blockchain systems**, with an initial emphasis on Algorand and Ethereum blockchains.

Basic functional requirements:

- Setup a Jupyter notebook server and introduce authentication mechanisms
- Connect the notebook with backend indexers containing blockchain data
- Setup a proxy functionality to load balance the requests (e.g., in an additional database)

Basic non-functional requirements

- Reusable data queries
- Realtime client – server communication
- Caching of results in the server
- Reusable codebase components

Expected prior knowledge

- Knowledge in Python and Jupyter or similar notebook setups
- Basic Knowledge working with REST APIs, RPCs
- Basic Knowledge of Socket.io or similar realtime client – server communication frameworks
- Knowledge of databases and deployment infrastructure
- Knowledge of blockchains is a plus

Resources

Dune Analytics - <https://dune.com/home>

Hex.tech - <https://hex.tech/>

Contact

Burak Öz | burak.oez@tum.de

Filip Rezabek | filip.rezabek@tum.de

AI Employment Contract Analysis

Advisor / Product Owner: Oliver Wardas
Dr. Dominik Herzog
Domenic Böhm



SYLVENSTEIN Rechtsanwälte

AI Employment Contract Analysis

Arbeitsvertrag

§ 4 Vergütung

Sofern die aktuelle Vergütung oberhalb der jeweils gültigen Beitragsbemessungsgrenze der gesetzlichen Rentenversicherung liegt, sind sämtliche Tätigkeiten des Arbeitnehmers aus diesem Vertrag inklusive Überstunden und Mehrarbeit abgegolten.

§ 5 Kündigung/Beendigung

Kündigungen müssen schriftlich erfolgen. Eine außerordentliche Kündigung aus wichtigem Grund gilt im Falle ihrer etwaigen Unwirksamkeit hilfsweise vorsorglich als ordentliche Kündigung zum nächst zulässigen Termin.

§ 6 Pfändung/Abtretung

Die Arbeitnehmerin/der Arbeitnehmer darf ihre/seine Vergütungsansprüche weder verpfänden noch abtreten.

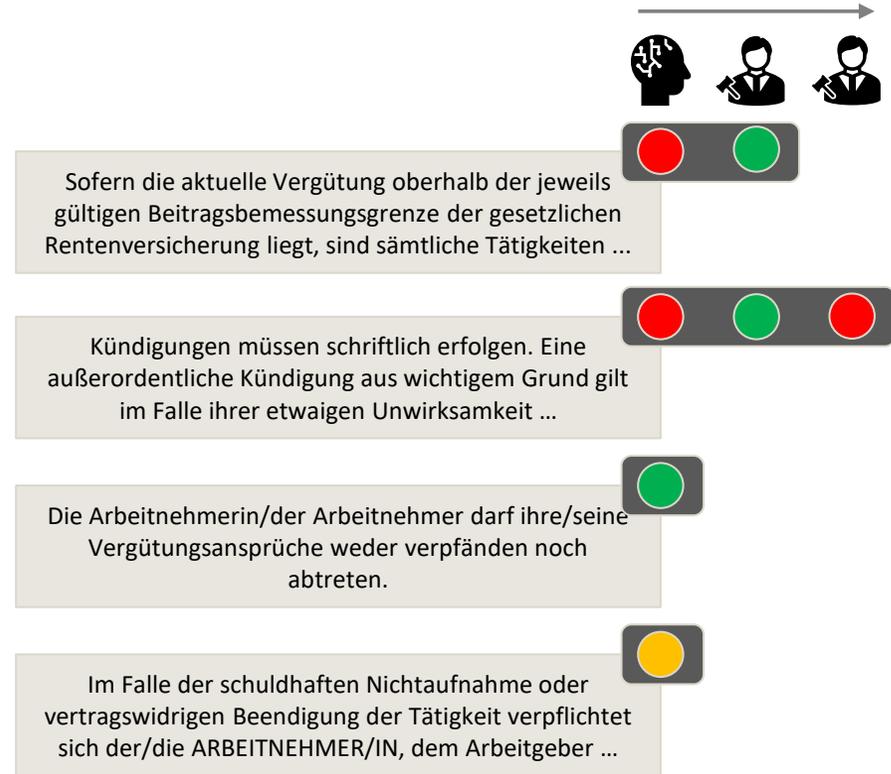
§ 7 Vertragsstrafe/Wettbewerbsverbot

Im Falle der schuldhaften Nichtaufnahme oder vertragswidrigen Beendigung der Tätigkeit verpflichtet sich der/die ARBEITNEHMER/IN, dem Arbeitgeber eine Vertragsstrafe in Höhe eines Gesamtmonatseinkommens zu bezahlen. Die FIRMA ist berechtigt, einen weitergehenden Schaden geltend zu machen.

§ 8 Leistungen

Die Arbeitnehmerin hat die ihr obliegenden Aufgaben sorgfältig und gewissenhaft nach Maßgabe der Gesetze und der ihm vom Arbeitsgeber erteilten Weisungen zu erfüllen.

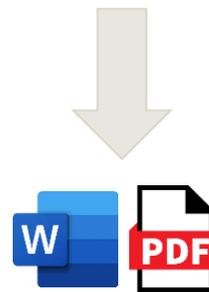
OCR



Kündigungen müssen schriftlich erfolgen. Eine außerordentliche Kündigung aus wichtigem Grund gilt im Falle ihrer etwaigen Unwirksamkeit ...

Kündigungen müssen schriftlich erfolgen.

Kündigungen können nur in Schriftform erfolgen. Eine außerordentliche Kündigung aus wichtigem Grund gilt im Falle ihrer etwaigen Unwirksamkeit ...



Outcome: App that supports:

- ✓ Upload of PDF documents
- ✓ OCR 3rd party API calls
- ✓ AI & DB backend service
- ✓ User-friendly interface

Technologies:

Pitch: The goal of the project is to develop a web app for AI analysis of employment contracts, supporting document OCR and segmentation, AI legal reviews, management of conflicting reviews (AI & human) and new contract creation.

Basic functional requirements:

- Upload, OCR and segmentation of Contract Docs.
- Using AI service for classifying clauses
- Presenting (conflicting) review annotations (+Input)
- Vector DB search for clause alternatives + user select (optional)

Basic non-functional requirements:

- UI with loading indicators and user feedback
- Intuitive controls for uploading, annotating etc.
- Design of modular components and services

Expected prior knowledge:

- Good Knowledge in React, NodeJS
- Basic Knowledge in Python
- Basic NLP understanding

Beneficial but NOT expected:

- NLP/Machine Learning programming experience

Contact:

oliver.wardas@tum.de

XNLP – Explanation Tool for NLP

Advisor / Product Owner: Mahdi Dhaini



XNLP – Explanation Tool for NLP

EXAMPLE

XNLP Explanation Tool for NLP Hj, Mahdi! Settings

Dashboard **Datasets** **Models** **Reports**

Input Text

The movie is awful for kids

Output

<u>Prediction</u>	<u>Confidence</u>
negative	0.98

Model roBERTa-base **Dataset** imdb

Actions: Example text from dataset Misclassified example Get prediction Generate explanation

Counterfactual Explanations

	<u>Prediction</u>
The movie is not awful for kids	positive
The movie is awful for kids	negative
A movie is awful for kids	negative
The movie is great for kids	negative
The movie is awful for adults	negative

Generate Different Explanations **Negation**

Saliency Maps

SHAP LIME

The movie is awful for kids

Word	Weight
kids	0.28
for	0.06
awful	0.30
is	0.04
movie	0.20
The	0.05

Human Feedback

Are the explanations helpful?

Generate Report

Pitch: The goal of the project is to develop an interactive analysis and explainability tool for the behavior and predictions of natural language processing models.

Basic functional requirements:

- Creation of a graphical user interface for input and navigation
- Interactive explanations of model output based on different inputs
- Visualization of different feature-based and counterfactual explanations
- Managing different models and datasets.
- Generating reports summarizing the models' analysis

Basic non-functional requirements:

- Realtime processing of predefined and new queries
- Adaption of the result set in real-time
- Modular and reusable components
- Scalable system architecture

Expected prior knowledge

- Knowledge and skills in web development.
- Good knowledge of Python programming language
- Desirable:
 - Understanding of NLP methods and applications.
 - Knowledge or interest in Explainable AI methods.

Outcome: App that supports:

- ✓ Realtime explanation of models
- ✓ Visualized explanations

Contact:

Mahdi.dhaini@tum.de

Technologies Pool:



Synthesizing Evidence-Based Answers

Advisor / Product Owner: Juraj Vladika



SEBA: Synthesizing Evidence-Based Answers

EXAMPLE

The screenshot shows the SEBA web application interface. At the top, there is a search bar with the query "Is sitting for long periods of time bad for your health?". To the right of the search bar is a "Filter:" section with three toggle switches for "- year", "- domain", and "- score". Below the search bar, there are three result cards. Each card contains a title, a "Practitioner Summary" snippet, a citation, and a thumbs-up/down icon with a "Yes" or "No" label. To the right of the results is a progress bar with three segments: "Yes" (green), "Maybe" (yellow), and "No" (red). Below the progress bar is a "Summary" section with a paragraph of text.

SEBA

Is sitting for long periods of time bad for your health?

Filter:
- year
- domain
- score

The immediate effect of the abdominal drawing-in maneuver technique on stature change in seated sedentary workers with chronic low back pain.

Practitioner Summary: Prolonged sitting seemingly harms sedentary workers' health, particularly affecting the lower back. (...) **Yes**

P. Saiklang et al. | 2020 | Citations: 48 | *Journal of Ergonomics* (Impact Score: 7.8)

Sitting patterns at work: objective measurement of adherence to current recommendations

Long uninterrupted sedentary periods, independent of total sedentary time, are risk factors for poor health (...) **Yes**

C. Ryan et al. | 2011 | Citations: 190 | *Journal of Accident Analysis and Prevention* (Impact Score: 6.3)

Sedentary behaviour and risk of mortality from all-causes and cardiometabolic diseases in adults: evidence from the HUNT3 population cohort

However, prolonged sitting in specific contexts (ie, watching TV, at work) do not adversely impact health in the same timeframe. (...) **No**

J. Chau et al. | 2013 | Citations: 135 | *British Journal of Sports Medicine* (Impact Score: 18.3)

Yes

Maybe

No

Summary

Some studies suggest that prolonged sitting is negatively associated with health, affecting the lower back and increasing cardiovascular disease risk factors, while other studies do not support the hypothesis that occupational sitting is associated with health problems.

- Outcome:** App that supports:
- ✓ Extensive document search of research publications
 - ✓ Evidence detection and argument mining
 - ✓ Synthesizing and summarizing the results

Technologies:

The image shows logos for various technologies: Python, CO, HTML5, JS, and others.

Pitch: The goal of the project is to develop a tool that for a given scientific question aims to find evidence and arguments in a database of scientific publications and provide an overview and analysis of the synthesized results, based on ML & NLP technology.

Basic functional requirements:

- Preparation and processing of textual data and documents
- Creation of a graphical UI for input and navigation
- Synthesis and visual presentation of discovered results
- Construction and configuration of machine learning pipelines with underlying NLP models

Basic non-functional requirements:

- Realtime processing of predefined and new queries
- Efficient retrieval of documents from a large database
- Modularity and reusability of components

Expected prior knowledge

- Knowledge of HTML, JavaScript, CSS
- Skills in Python, Flask / Django, Angular / React
- Integration with databases, cloud deployment

Desirable:

- Understanding of NLP methods & models
- Knowledge of frameworks for LLM-based apps (like LangChain) and vector databases (like Weaviate)

Contact:

juraj.vladika@tum.de

PATER

Privacy Analytics with Differentially Private Text Rewriting

Advisor / Product Owner: Stephen Meisenbacher



☰
PATER
Welcome, Stephen!

-
-
-
-
-

Framework

Noise

Initial Distribution:

Noise Type:

Privacy Budget

Epsilon

Utilization

ORIGINAL TEXT

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. heart Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

PRIVATE TEXT

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. liver Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

PRIVACY WORKBENCH

heart (/hɑ:t/)

noun
noun: heart; plural noun: hearts

1. a hollow muscular organ that pumps the blood through the circulatory system by rhythmic contraction and dilation. In vertebrates there may be up to four chambers (as in humans), with two atria and two ventricles.

verb **INFORMAL**
verb: heart; 3rd person present: hearts; gerund or present participle: hearting; past tense: hearted; past participle: hearted

1. like very much; love.

Word Information	
POS in context	NOUN
Information Content	10
Perturbed Word	liver
Theoretical budget spent	1.43 (1.4%)
Actual budget spent	0.42 (0.42%)
Budget Gap	1.01
Privacy Score ©	42

Budget	Word	Score
Change word: heart		
To: <input type="text" value="cardiovascular"/>		
New budget spend: 0.2		
New budget utilization: 89.7%		
Thus saving: 0.22 (0.22%)		

SEBA Lab WS 23/24, Meisenbacher

Goal: to develop an interface that allows users not only to privatize their text via Differential Privacy mechanisms, but also to tailor the privatization to individual privacy risk tolerances.

Outcome: App that supports:

- ✓ Interactive dashboard
- ✓ Real-time calculations
- ✓ Complex data visualization
- ✓ Rich analytics backend

Technologies:

231017 SEBA Lab Kickoff Hoops

52

Pitch: The goal of this project is to develop an interface that allows users not only to privatize their text via Differential Privacy (DP) mechanisms, but also to tailor the privatization to individual privacy risk tolerances.

Basic functional requirements:

- Creation of an interactive text privatization dashboard
- Adaptation of calculations to user-given parameters
- Display of results with interactive text and adaptable parameters
- “Gamification” of privatization to incentivize use
- Providing global statistics on user privacy preferences

Basic non-functional requirements:

- Real-time execution of DP mechanisms
- Real-time capturing of user preferences and production of corresponding visualizations
- Design of custom components to facilitate the UI
- Code readability and quality

Expected prior knowledge:

- Knowledge in Flask, React, Javascript, HTML, CSS
- Strong programming skills in Python
 - Useful: high-performance computing / data engineering skills
- Basic knowledge in Natural Language Processing
- Solid knowledge in Probability & Statistics
- Genuine interest in (data) privacy and Privacy-Enhancing Technologies!
- Experience in JavaScript / Python visualization libraries like D3, Chart, Plotly, etc.

Contact:

stephen.meisenbacher@tum.de

Outline

Organization Details

Project Evaluation

Project Proposals

Next Steps

Next Steps

1. If you have any questions about a particular project, **get in touch with the according advisor or industry partner.**
2. **Submit your preferences** via e-mail to felix.hoops@tum.de **until 22.10.2023**

Subject:

SEBA Lab 23 - Preferences - #Your Last Name#, #Your First Name#

Body:

Prio 1: #Project Name#

Prio 2: #Project Name#

Prio 3: #Project Name#

Prio 4: #Project Name#

Prio 5: #Project Name#

Preferences for fellow team members (max. team size of 4):

#Last Name#, #First Name#

#Last Name#, #First Name#

#Last Name#, #First Name#

3. You will receive your final team and project information via e-mail **latest 25.10.2023**



M.Sc.

Felix Hoops

Technical University of Munich (TUM)
TUM School of CIT
Department of Computer Science (CS)
Chair of Software Engineering for Business
Information Systems (sebis)

Boltzmannstraße 3
85748 Garching bei München

+49.89.289.17126

felix.hoops@tum.de

www.matthes.in.tum.de

