

NLP Application Project @ FAST AI Movies

October 05th 2023

Alejandro Bravo de la Serna, George Maged Elfayoumi, Parag Bamel, Jinyu Lee, Mohamed Hesham Ibrahim Abdalla

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

What is an Application Project?

Project Overview

System Components

Main Goal

Timeline & Task Distribution

Project for Data Engineering & Analytics Students



Interdisciplinary Thinking: Data Analytics within the Application Domain



In Cooperation with Chairs and/or Start-Ups



Workload of 10 ECTS ~ 300 hours (distributed over 1-2 Semesters)



Teams of up to 5 DEA Students

Project for Data Engineering & Analytics Students



Interdisciplinary Thinking: Data Analytics within the Application Domain



In Cooperation with Chairs and/or Start-Ups



Workload of 10 ECTS ~ 300 hours (distributed over 1-2 Semesters)



Teams of up to 5 DEA Students

Similar to IDP

What is an Application Project?

Project Overview

System Components

Main Goal

Timeline & Task Distribution

NLP Application Project @ FAST AI Movies



→ Young, TUM-affiliated Start-Up FAST AI Movies

→ Software-as-a-Service (SaaS) for **e-Learning Video Generation**



Retrieving Information about **Structures and Semantic Relationships**
from **high-dimensional Textual Data** for **Visual Information Representation**

Tim Schopf (SEBIS), Philipp Gabriel (FAST AI Movies)



What is an Application Project?

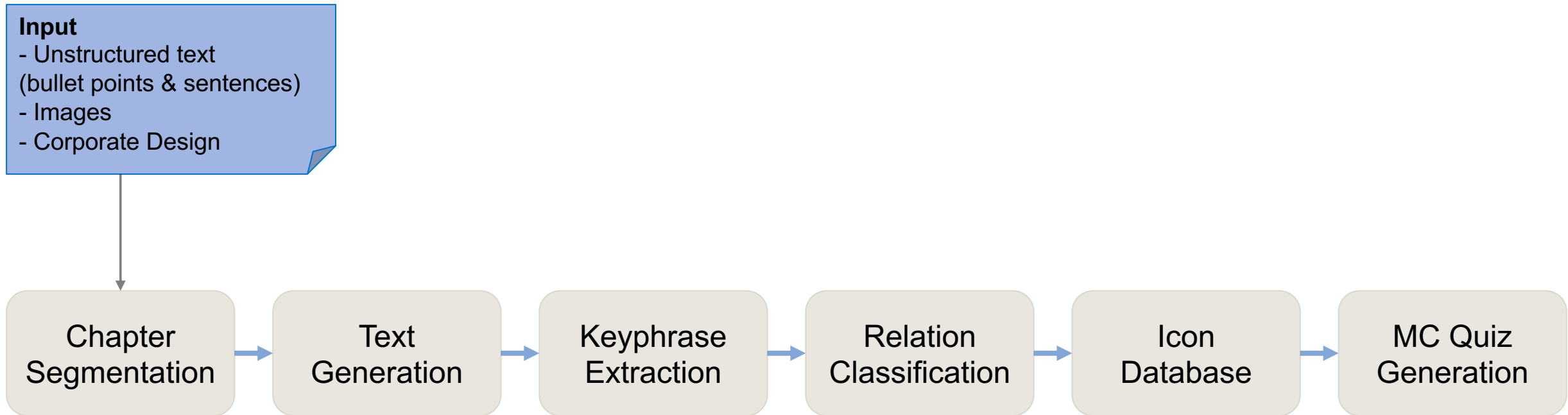
Project Overview

System Components

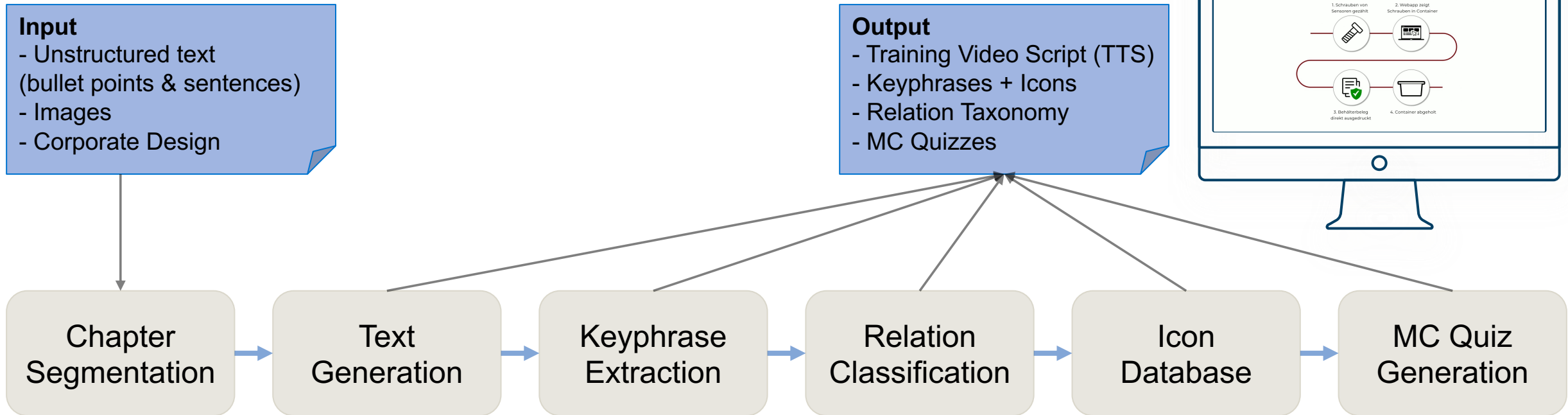
Main Goal

Timeline & Task Distribution

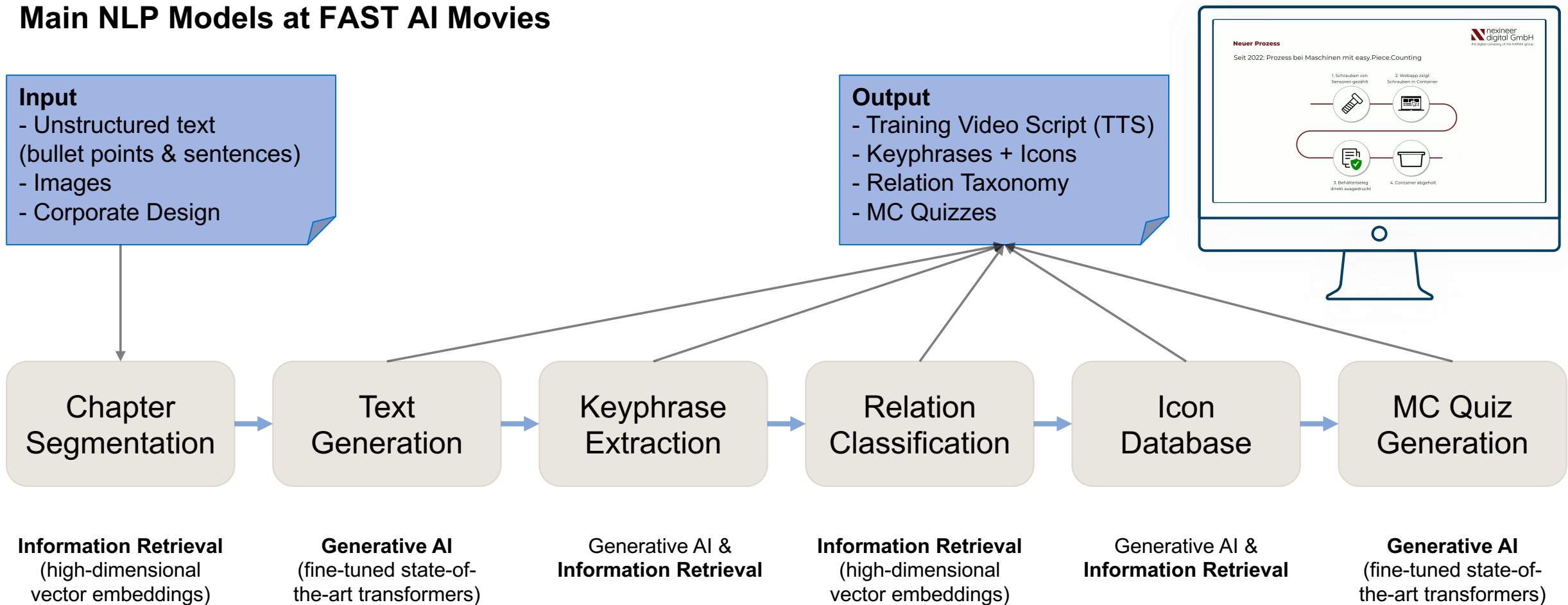
Main NLP Models at FAST AI Movies



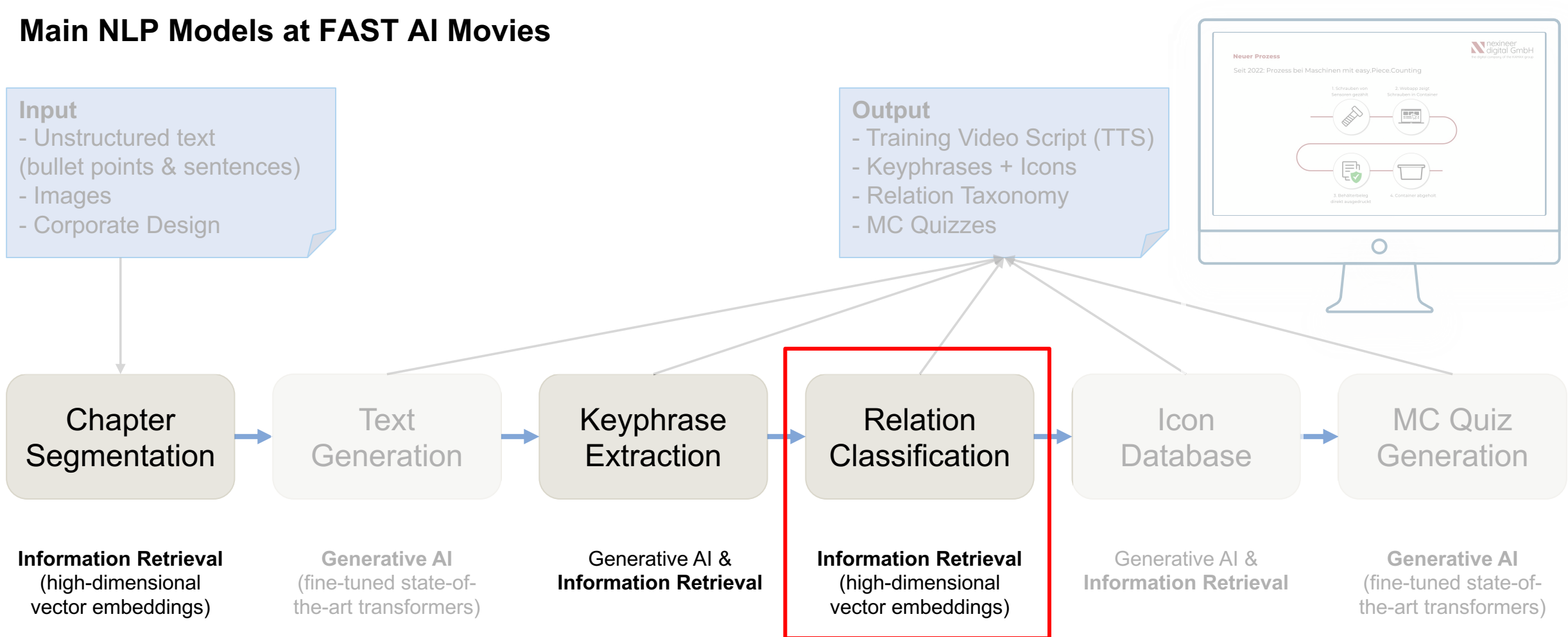
Main NLP Models at FAST AI Movies



Main NLP Models at FAST AI Movies



Main NLP Models at FAST AI Movies



Outline

What is an Application Project?

Project Overview

System Components

Main Goal

Timeline & Task Distribution

Improvement of the Relation Classification

Input

- Text
- Keyphrases in Text
["A", "B", "C", "D"]



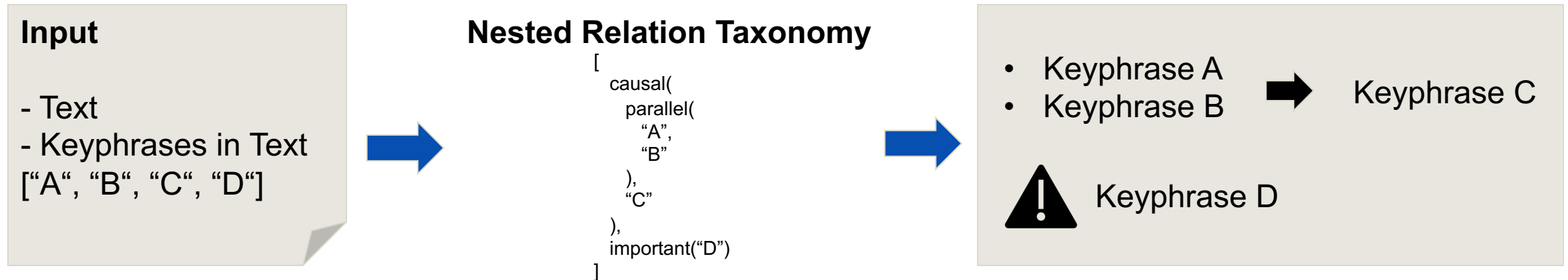
Nested Relation Taxonomy

```
[  
  causal(  
    parallel(  
      "A",  
      "B"  
    ),  
    "C"  
  ),  
  important("D")  
]
```



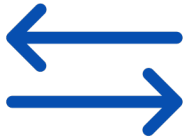
- Keyphrase A → Keyphrase C
- Keyphrase B → Keyphrase C
- ! Keyphrase D

Improvement of the Relation Classification



Goal: Remove/reduce human from the loop → Improve SaaS Capabilities

Challenges



Pair Relations



Transitive Taxonomy

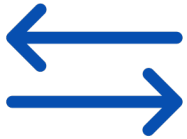


Root Relations



Few-Shot Approaches

Challenges



Pair Relations



Transitive Taxonomy



Root Relations



Few-Shot Approaches

→ Overlaps with Tim's Research

Outline

What is an Application Project?

Project Overview

System Components

Main Goal

Timeline & Task Distribution

Milestones

Task	Timeline	Assignes
1. Onboarding, Setup & Planning	Weeks 1-2	All
2. Data Inspection & Quality Assurance	Weeks 3-4	All
3. Understanding & Optimizing the Input Models 3.1 Segmentation (Chapterization) 3.2 Keyphrase Extraction	Weeks 5-7	Jinyu & Parag Alejandro & George
4. Relationship Classification 4.1 Pair Relations 4.2 Transitive Taxonomy 4.3 Root Relations	Weeks 8-13	All Jinyu & Alejandro George & Parag
5. Testing and Evaluation	Weeks 14-15	All
6. Documentation and Finalization	Weeks 15-16	All



Prof. Dr.

Florian Matthes

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

+49.89.289.17132
matthes@in.tum.de
www.matthes.in.tum.de

