



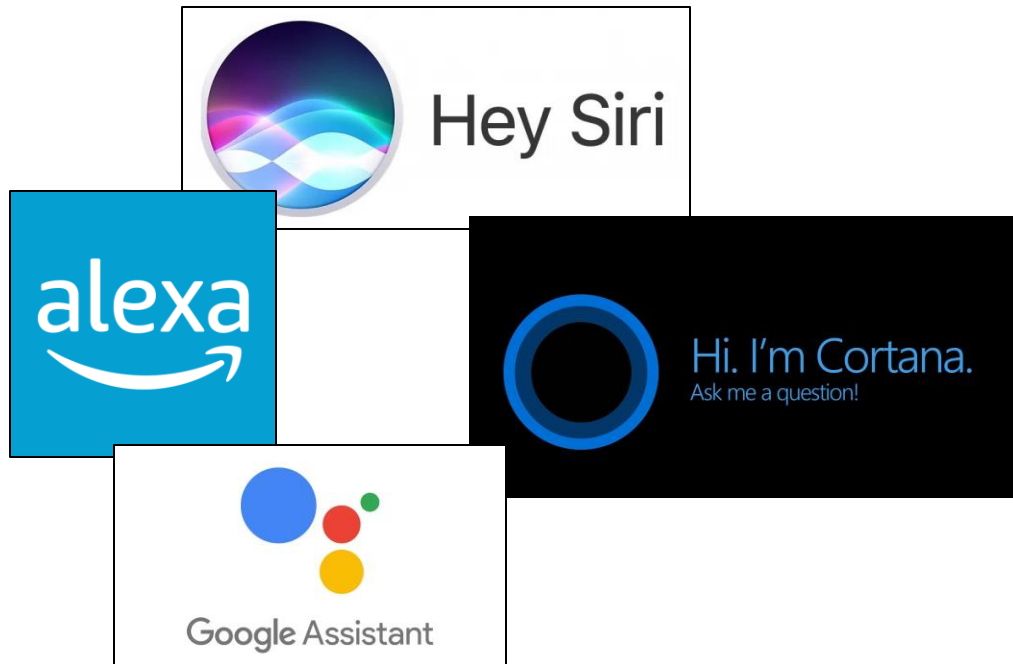
Building an Intelligent Conversational Agent

ATAI 2023, Prof. Abraham Bernstein

Rosni Vasu, Athina Kyriakou





Conversational Agents



Who directed Batman 1989

Batman was directed by Tim Burton.

 KNOWLEDGE

 **Tim Burton**
American filmmaker and artist


Timothy Walter Burton is an American filmmaker and artist. He is known for his gothic fantasy and horror films such as Beetlejuice, Edward Scissorhands, The Nightmare Before Christmas, Ed Wood, Sleepy Hollow, Corpse Bride, Sweeney Todd: The Demon Barber of Fleet Street, and Dark Shadows.

Wikipedia

Date of birth	August 25, 1958
Place of Birth	Burbank >

Where was the Director of Batman 1989 born


Here's what I found.

 WEBSITES

Batman (1989 film) - Wikipedia
After Burton was hired as director in 1986, Steve Englehart and Julie Hickson wrote film treatment...
en.wikipedia.org

Tim Burton | Batman Wiki - Fandom
Timothy Walter Burton is an American film director, animator and artist, known for his frequent collabo...
batman.fandom.com

Batman (1989) - IMDb
Batman: Directed by Tim Burton. With Michael Keaton, Jack Nicholson, Kim Basinger, Robert...
imdb.com

Show Google Results 



1. Project Overview

Goal: Implement a conversational agent that can hold conversations with humans and answer natural language questions (in a **team of 2**)

Make sure you register your team in OLAT
by **27.09.2023!**

Types of questions:

- **Closed:** questions that have a specific answer given the available data (e.g., “*Who is the director of Titanic?*”)
- **Open:** questions that have multiple valid answers (e.g., “*Which movie should I watch tonight if I like documentaries?*”)

Techniques: knowledge-graph, embeddings, recommendation, multi-modal, crowdsourcing



2. Project Evaluation Overview

- 20% of the final course grade
- You **need to pass** the project to pass the overall course!
- Evaluation in 4 phases of incremental complexity
- All deadlines are 9 am Zurich time
- At least 1 team member needs to be present on each evaluation event

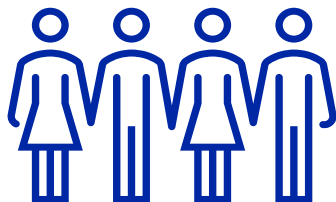
Evaluation	Submission Deadline	Evaluation Event	% of Project Grade
1st intermediate	no submission	11.10.2023	5%
2nd intermediate	30.10.2023	06.11.2023	10%
3rd intermediate	22.11.2023	29.11.2023	20%
Final	11.12.2023	13.12.2023	65%



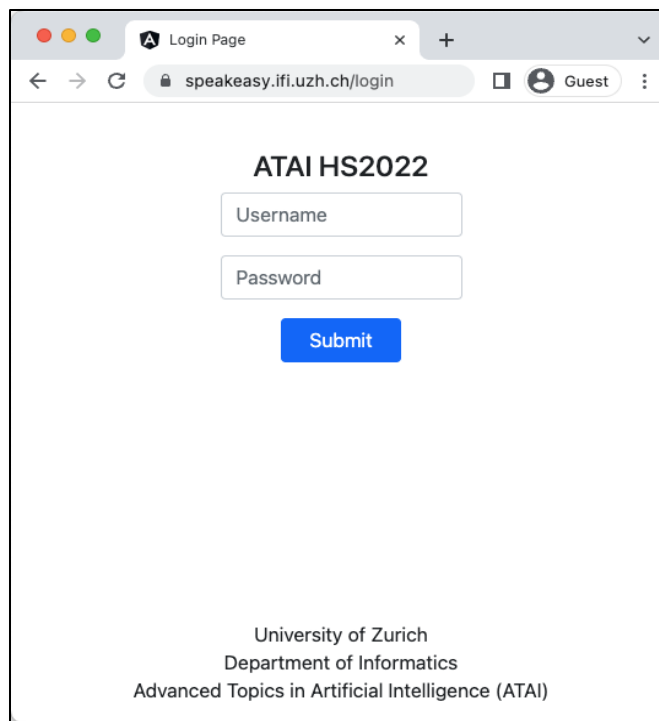
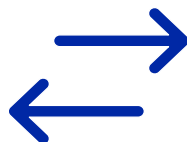
3. Infrastructure & Provided Resources

Speakeasy: a web-based infrastructure to connect your agent and interact with the agents of others'

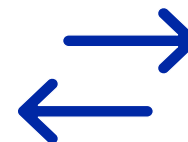
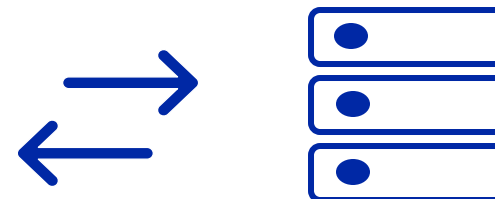
User-side:



Users



Web Pages

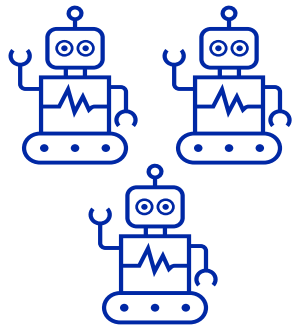


<https://speakeasy.ifi.uzh.ch/>

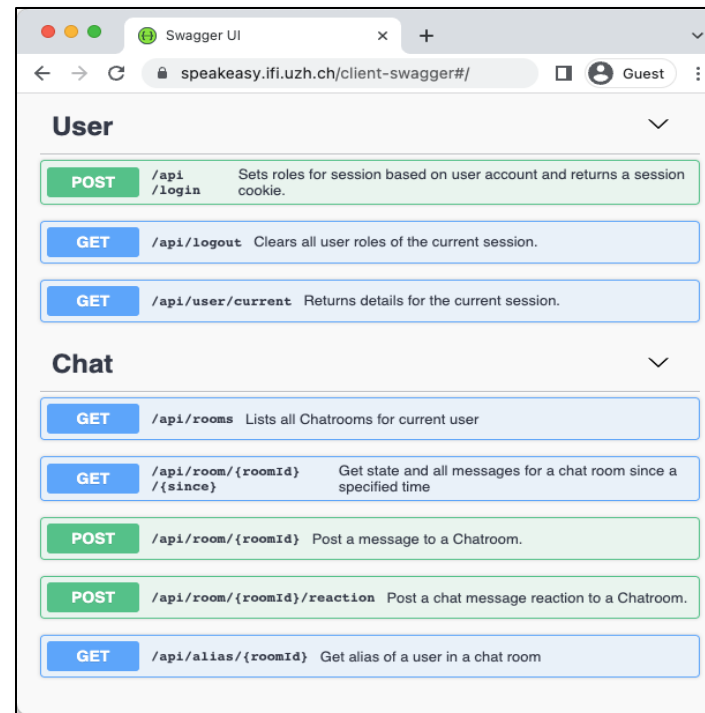
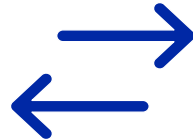


3. Infrastructure & Provided Resources

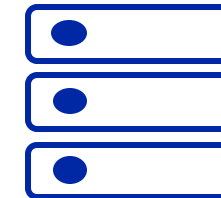
Agent-side:



Agents



HTTP Requests



<https://speakeasy.ifi.uzh.ch/>

For Speakeasy we provide:

- A user account per student
- An agent account per team
- A python client library (OLAT > Materials > Tutorials > Speakeasy_Project)

If you don't receive the accounts **by 01.10.2023**, please contact us immediately!






3. Infrastructure & Provided Resources

Data: we provide a collection of datasets that your agent **must use** for answering questions in the evaluation event:

- a **knowledge graph** (*Tutorial Available*)
- a **multimedia** dataset
- pre-trained **knowledge graph embeddings**
- a **crowdsourcing** dataset

Index of /ddis/teaching/2023/ATAI/dataset

<u>Name</u>	<u>Last modified</u>	<u>Size</u>	<u>Description</u>
 Parent Directory		-	
 crowd_data/	2022-11-15 20:37	-	
 ddis-graph-embeddings.zip	2021-10-12 19:03	145M	
 ddis-movie-graph.nt.zip	2021-09-21 17:09	13M	
 ddis-movie-graph.tsv.zip	2021-09-21 17:09	27M	
 imdb-top-250.t	2021-10-26 10:37	2.4K	
 movienet/	2021-09-21 14:29	-	
 plots.csv	2021-09-22 15:00	54M	
 user-comments.csv	2021-09-22 15:00	20M	

<https://files.ifi.uzh.ch/ddis/teaching/2023/ATAI/dataset/>



4. Intermediate Evaluations

1st: answer SPARQL queries using the provided knowledge graph

2nd: answer **closed questions** using the provided data

3rd: answer **open questions** using the provided data



5. Grading of Intermediate Evaluations

Grading for each intermediate evaluation is based on your agent's capabilities on that specific week. You are free to change your implementation until the final submission!

Submission: latest version of your code + 2 page report

No submission + no
presentation session for
the 1st intermediate
evaluation!

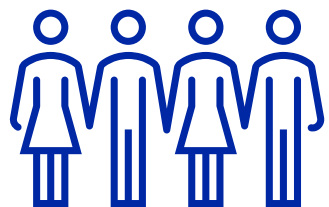
Evaluation Event:

- i. Evaluation Round (45'): online evaluation of your agent
- ii. Presentation Session (45'): 4 teams will be selected to give a 10' presentation (including Q&A) of their implementation. Selected teams will be informed 2 days before the evaluation event.

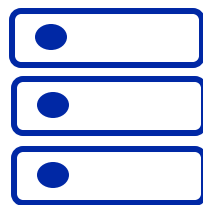
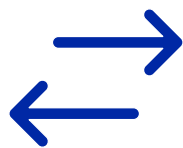


6. What is happening during an evaluation round?

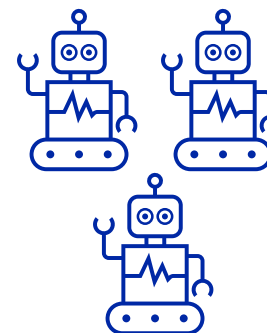
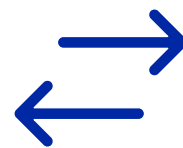
On each evaluation event there will be at least one evaluation round



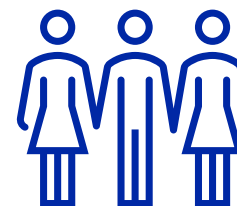
Students



<https://speakeasy.ifi.uzh.ch/>



Students' Agents



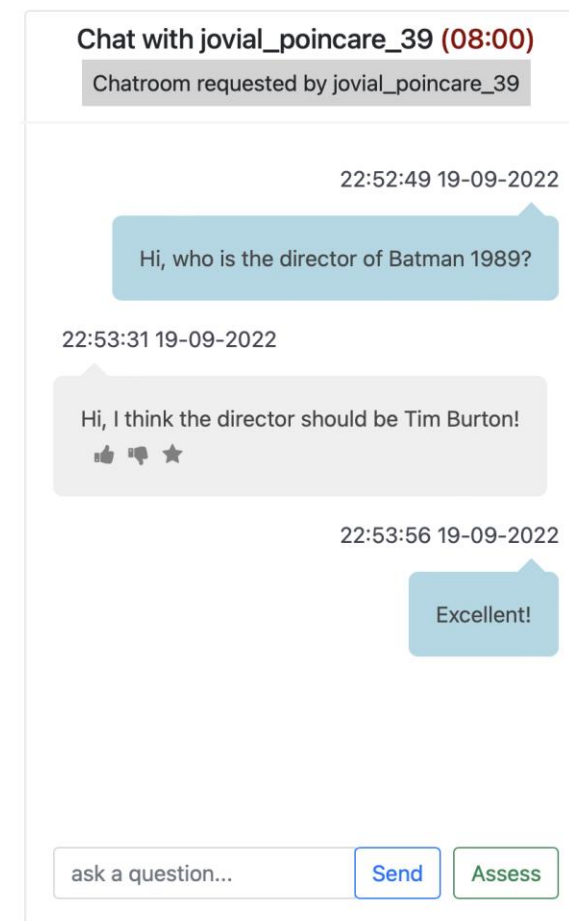
TAs + Instructors



6. What is happening during an evaluation round?

➤ Chatroom

- there is **one agent** from another student
- you need to have **a conversation** with the agent
- **prompts** about **what questions to ask** in the conversation will be given in the header of the chatroom
- the conversation will mainly be **you asking questions** and the **agent answering your questions**
- there is a **time limit** for each chatroom





6. What is happening during an evaluation round?

■ Rating

- when the conversation is finished, you need to fill out a form to **rate the agent**
- agents are assessed in terms of **accuracy**, **completeness**, **timeliness**, and **human likeness**
- there will be **several rounds of evaluation** to collect sufficient ratings for your agent from different and random people
- **all your conversations and ratings are logged in Speakeasy!**

Chat with jovial_poincare_39

Chatroom requested by jovial_poincare_39

22:52:49 19-09-2022

Hi, who is the director of Batman 1989?

22:53:31 19-09-2022

Hi, I think the director should be Tim Burton!

👍 🗨️ ★

22:53:56 19-09-2022

Excellent!

ask a question... Send Assess

Ratings for jovial_poincare_39

1. On average, the questions I asked in this conversation were answered accurately.

☐ Strongly Disagree ☐ Slightly Disagree

☐ Undecided ☐ Slightly Agree

☐ Strongly Agree

2. On average, the answers I received in this conversation were complete.

☐ Strongly Disagree ☐ Slightly Disagree

☐ Undecided ☐ Slightly Agree

☐ Strongly Agree

3. On average, the answers I received in this conversation were provided in a timely fashion.

☐ Strongly Disagree ☐ Slightly Disagree

Submit and Close



7. Final Evaluation

Goal: answer **open and closed** questions

Objectives:

- i. Implementation of suitable techniques to answer **5 types** of open and closed questions
- ii. Assessment of your agent's capability to interpret natural language queries and retrieve and articulate valid responses in terms of accuracy, completeness, timeliness, and human likeness



8. Required Functionalities for the Final Evaluation

Answer **factual questions** based on the **provided knowledge graph**

Who is the director of Good Will Hunting?	Gus Van Sant is the director of Good Will Hunting.
Who directed The Bridge on the River Kwai?	David Lean directed The Bridge on the River Kwai.
Who is the director of Star Wars: Episode VI - Return of the Jedi?	I think it is Richard Marquand.



8. Required Functionalities for the Final Evaluation

Answer **embedding questions** based on the **provided pre-trained embeddings**

Who is the screenwriter of The Masked Gang: Cyprus?	The screenwriter of The Masked Gang: Cyprus is Murat Aslan.
What is the MPAA film rating of Weathering with You?	The MPAA film rating of Weathering with You is PG-13.
What is the genre of Good Neighbors?	The genre of Good Neighbors is neo-noir.



8. Required Functionalities for the Final Evaluation

Answer **multimedia questions** based on the **provided multi-media dataset**

Show me a picture of Halle Berry.

What does Julia Roberts look like?

Let me know what Sandra Bullock looks like.





8. Required Functionalities for the Final Evaluation

Answer **recommendation questions** based on the **provided knowledge graph**

Recommend movies similar to Hamlet and Othello.	Adequate recommendations will be movies in the drama genre that are based on classic literature (e.g., Shakespeare, Dickens, or Jane Austen).
Given that I like The Lion King, Pocahontas, and The Beauty and the Beast, can you recommend some movies?	Adequate recommendations will be (2-D) animated movies or real-life remakes of Disney movies.
Recommend movies like Nightmare on Elm Street, Friday the 13th, and Halloween.	Adequate recommendations will be horror movies from the 1970s or 1980s or sequels to the movies.



8. Required Functionalities for the Final Evaluation

Answer **crowdsourcing questions** based on the **provided crowdsourcing dataset**

What is the box office of The Princess and the Frog?	The box office of The Princess and the Frog is 267000000. [Crowd, inter-rater agreement 0.236, The answer distribution for this specific task was 2 support votes, 1 reject vote]
Can you tell me the publication date of Tom Meets Zizou?	The publication date of Tom Meets Zizou is 2011-01-01. [Crowd, inter-rater agreement 0.040, The answer distribution for this specific task was 3 reject votes, 0 support votes]
Who is the executive producer of X-Men: First Class?	The executive producer is Sheryl Lee Ralph. [Crowd, inter-rater agreement 0.199, The answer distribution for this specific task was 2 support votes, 1 reject votes]



9. Grading of the Final Evaluation

Submission: latest version of your code + report

Evaluation Event: we expect

- i. Your agent to complete at least 3 rounds of evaluation
- ii. Personal active participation/completion in all the rounds of the final evaluation event
- iii. Both team members to be present



10. FAQs

- You can use any [programming language](#) you prefer - our materials will be predominantly in Python
- You can use [external libraries](#) and pre-trained models for solving specific problems
- There is [no support for programming problems](#) that do not relate to class-related content
- There are [no sessions for developing](#) your agent during the course
- Further details on [project overview](#), [intermediate evaluations](#), [final evaluation](#) in OLAT
- For further questions use the [Forum](#) in OLAT