

# A Step Towards the Interactive Shared Task

User Simulation Infrastructure at Touché RAD and TREC iKAT

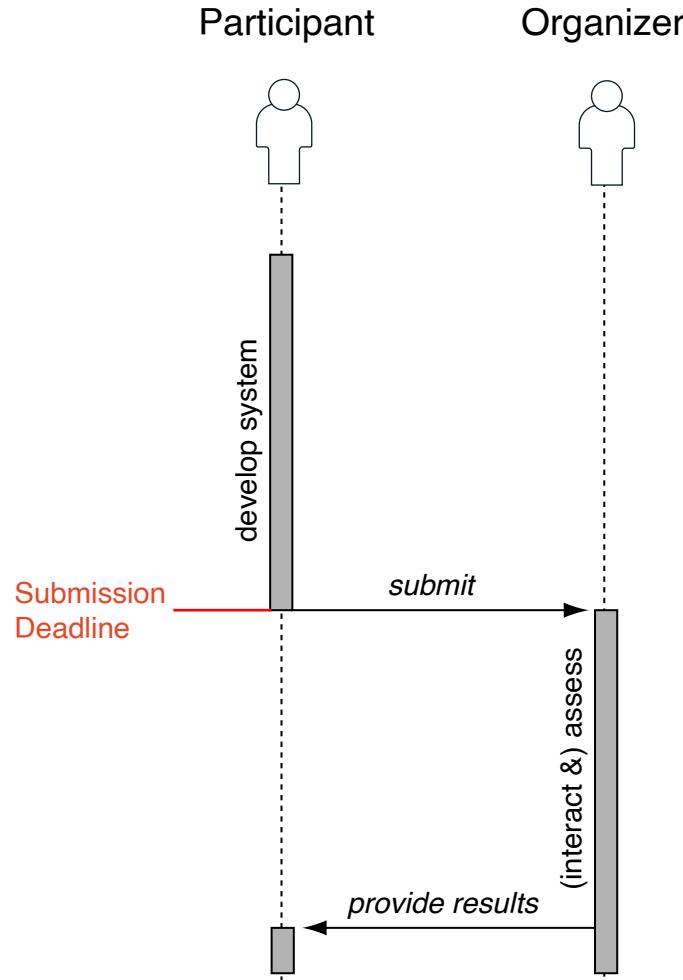


Marcel Gohsen

Bauhaus-Universität  
Weimar

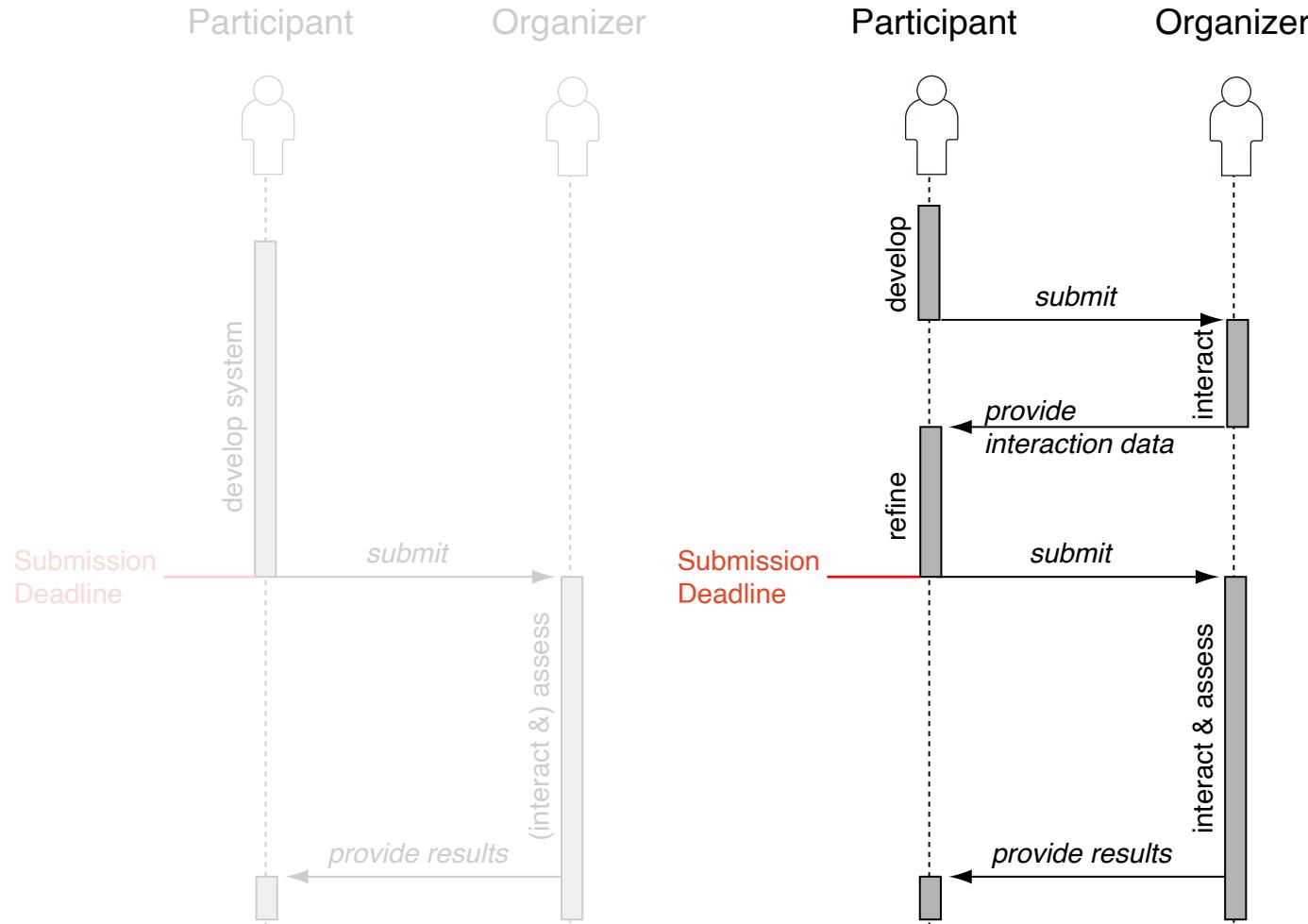
Research Group Intelligent Information Systems [webis.de]

## Non-Interactive Shared Task



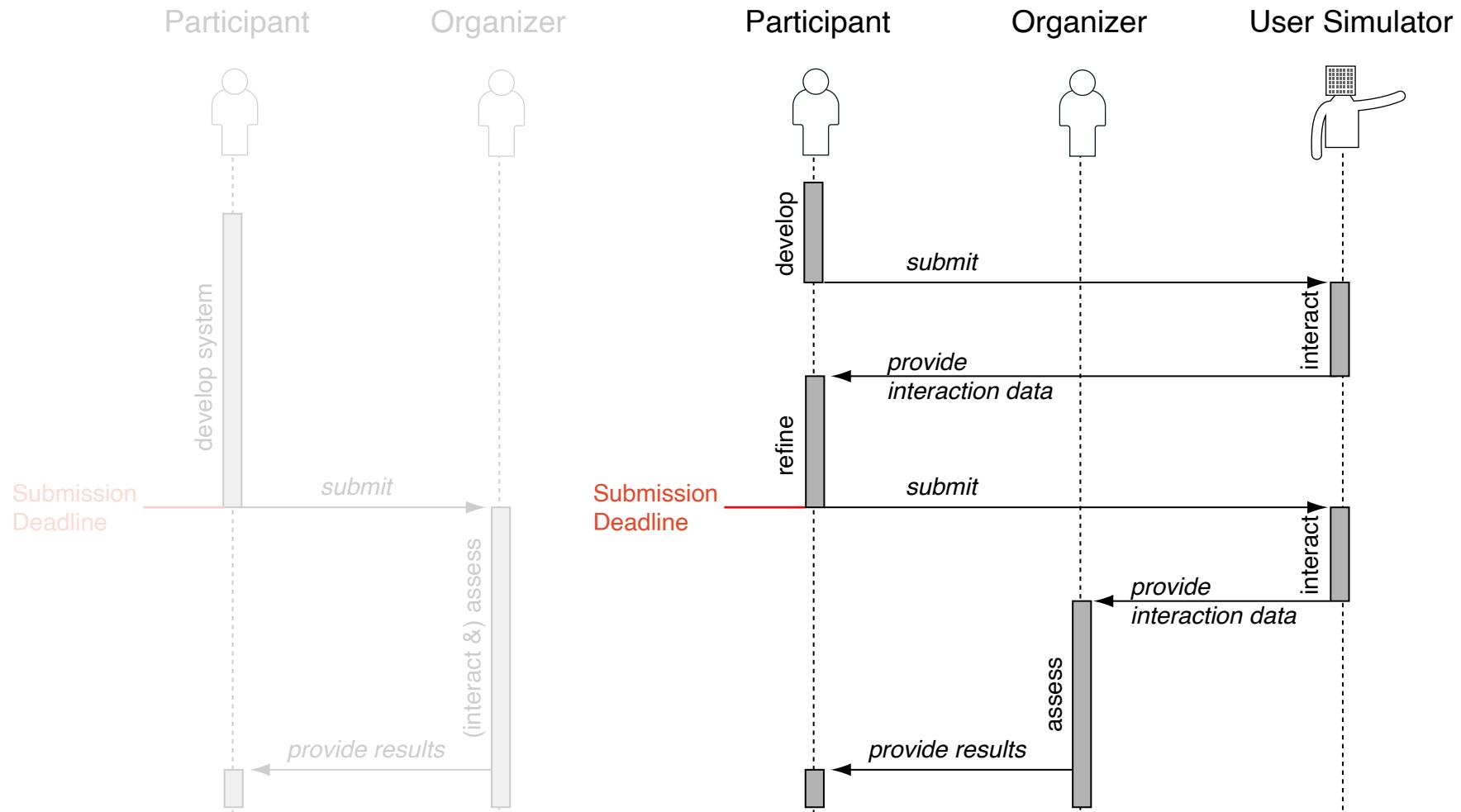
# Interactive Shared Task

(variant a)



# Interactive Shared Task

(variant b)



# Recent Interactive Shared Tasks (variant b)

## Touché 2025

Track: Retrieval-Augmented Debating (RAD).

Organizers: Marcel Gohsen, Nailia Mirzakhmedova, Harriscen Scells,  
Mohammad Aliannejadi, Maik Fröbe, Johannes Kiesel, Benno Stein.

<https://touche.webis.de/>

## Text REtrieval Conference (TREC) 2025

Track: Interactive Knowledge Assistance Track (iKAT).

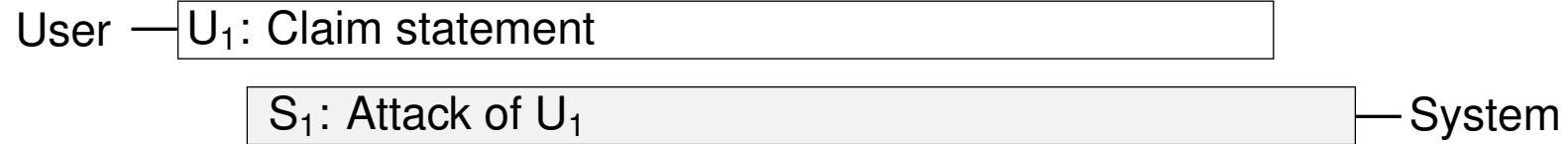
Organizers: Mohammad Aliannejadi, Zahra Abbasiantaeb, Simon Lupart,  
Nailia Mirzakhmedova, Marcel Gohsen, Johannes Kiesel, Jeff Dalton.

<https://www.trecikat.com/>

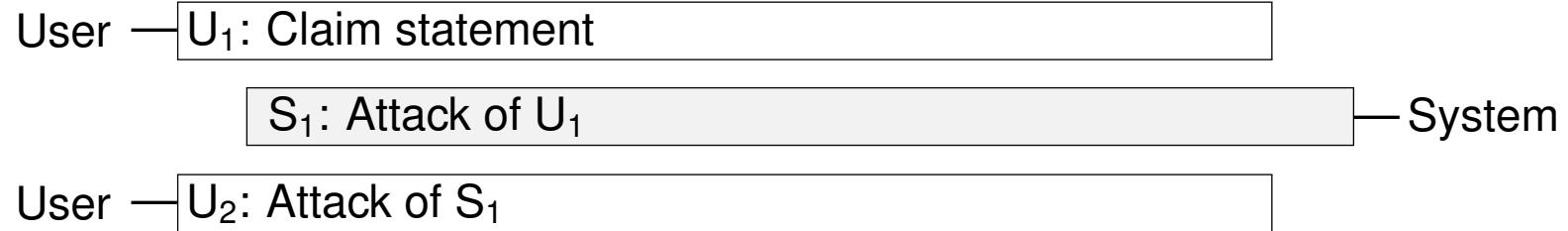
# Touché RAD

User — U<sub>1</sub>: Claim statement

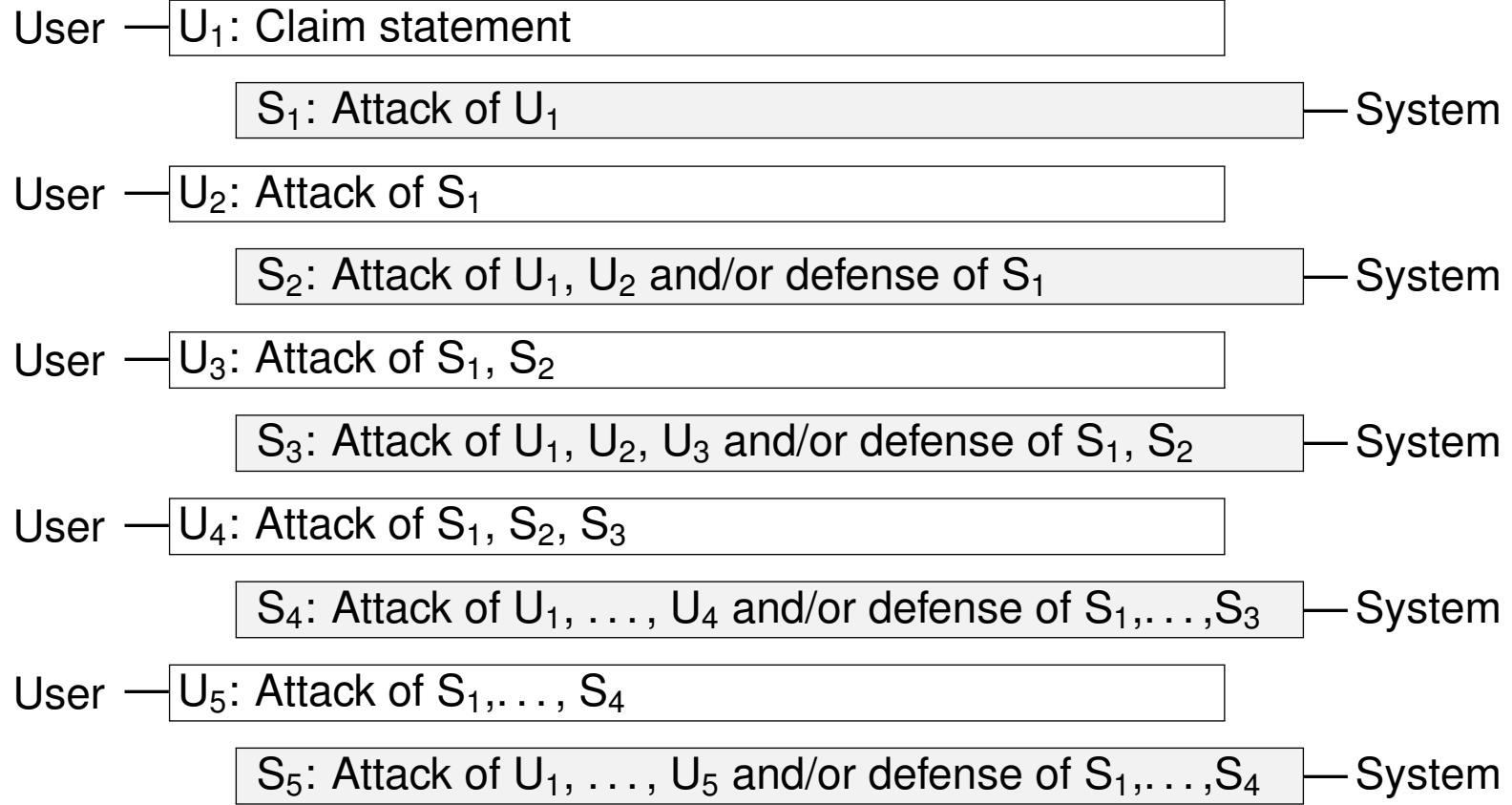
# Touché RAD



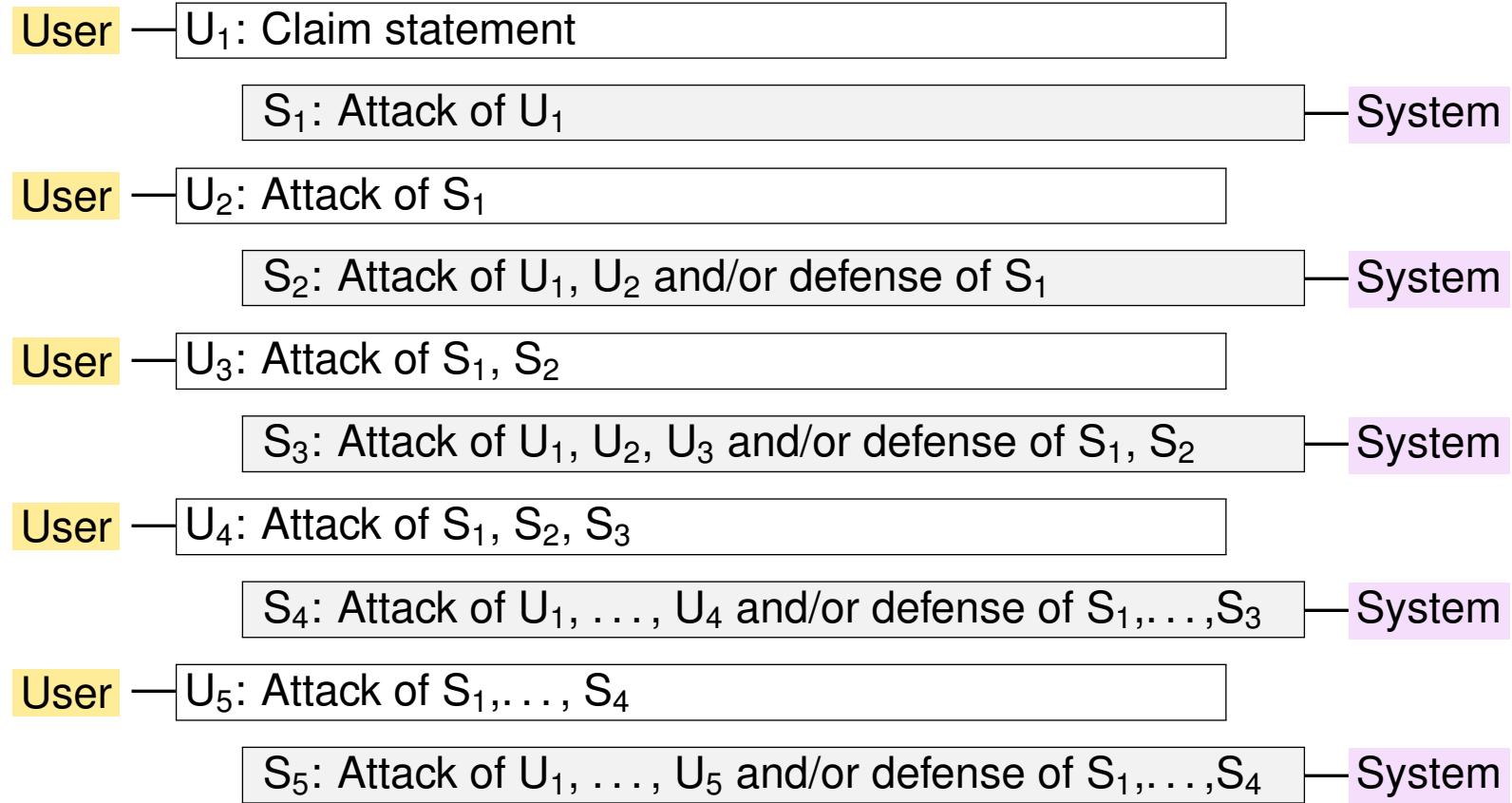
# Touché RAD



# Touché RAD



# Touché RAD



participant

organizer

# Touché RAD

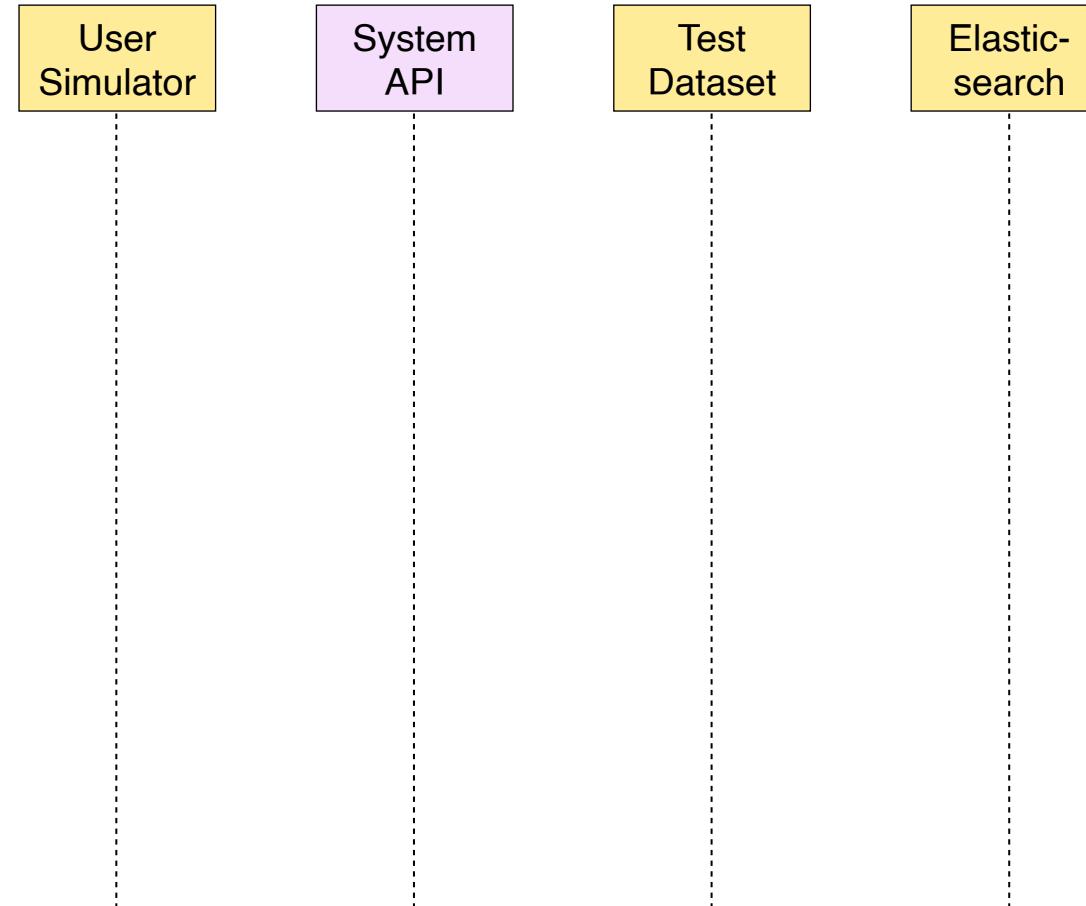
## Requirements of participant submissions

- ❑ Dockerized debate system
- ❑ System that implements predefined HTTP-API
- ❑ Runs inside TIRA

## Assessment of submissions

- ❑ Simulation based on secret simulator
- ❑ Simulation based on secret test set

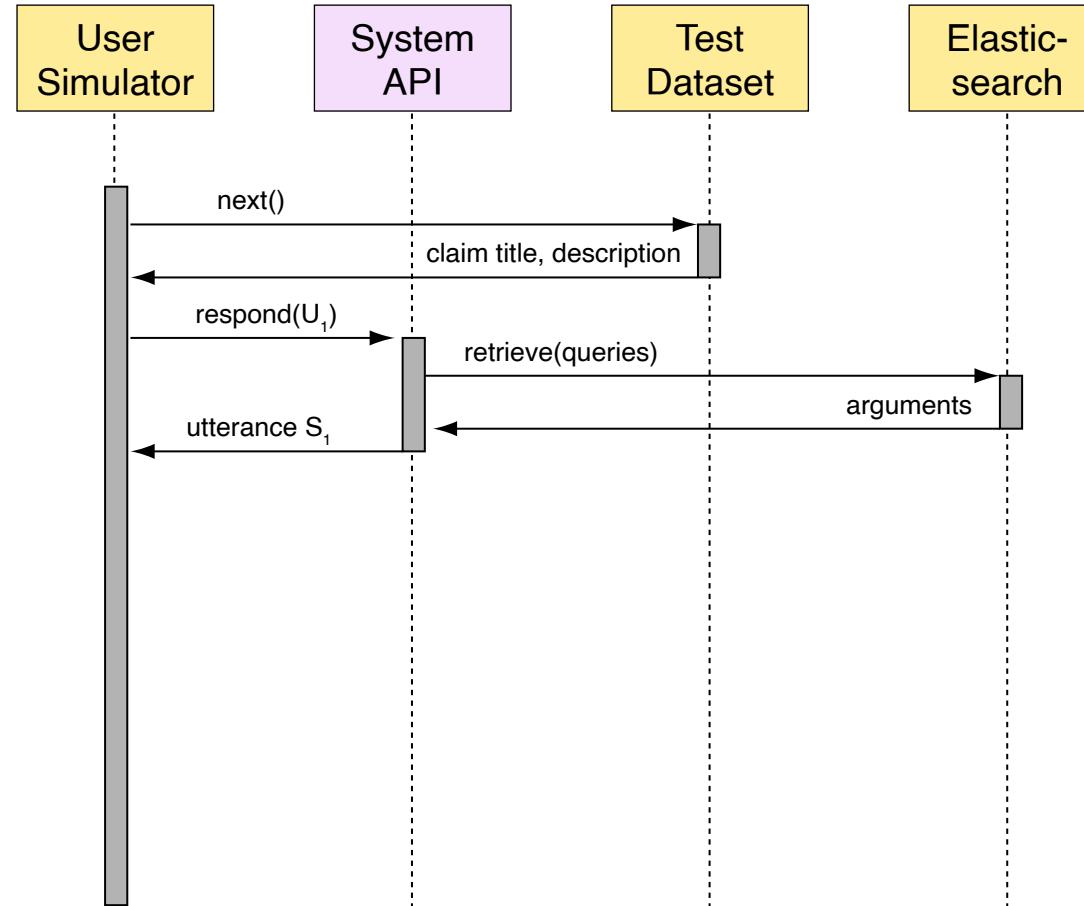
# Touché RAD



participant

organizer

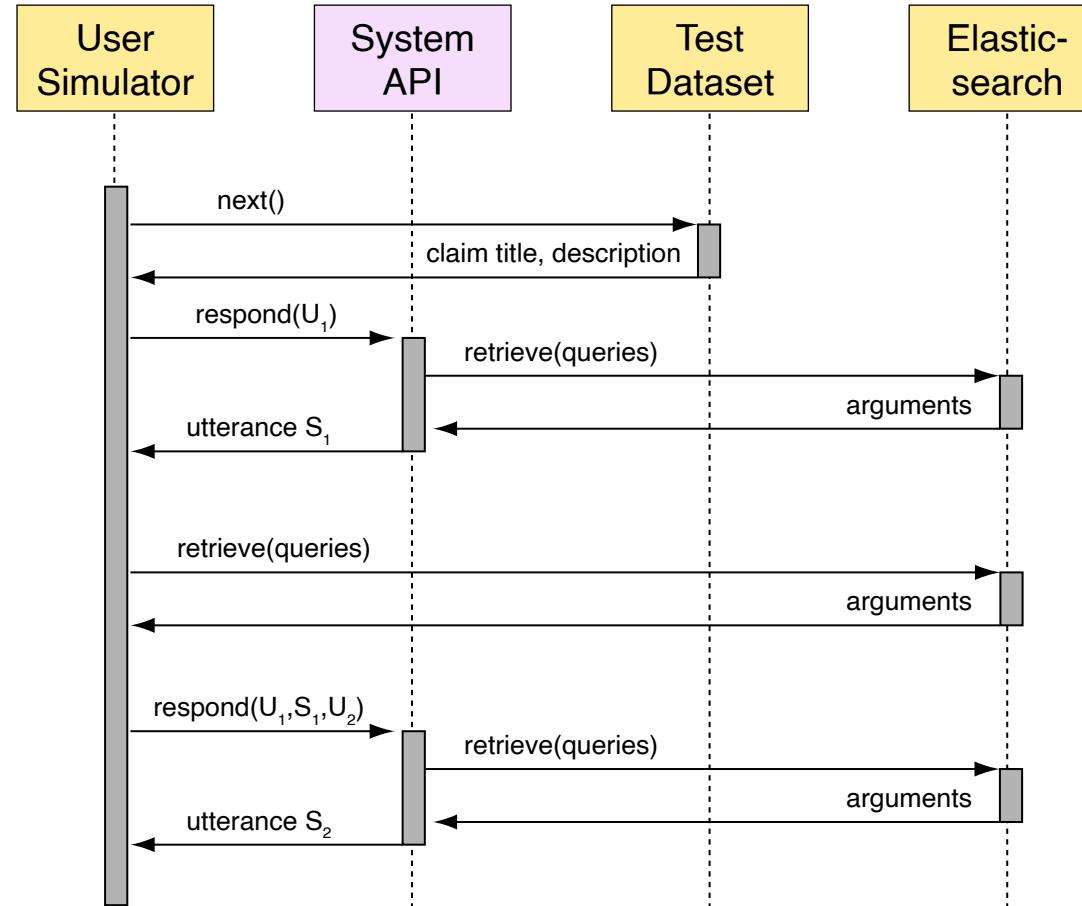
# Touché RAD



participant

organizer

# Touché RAD



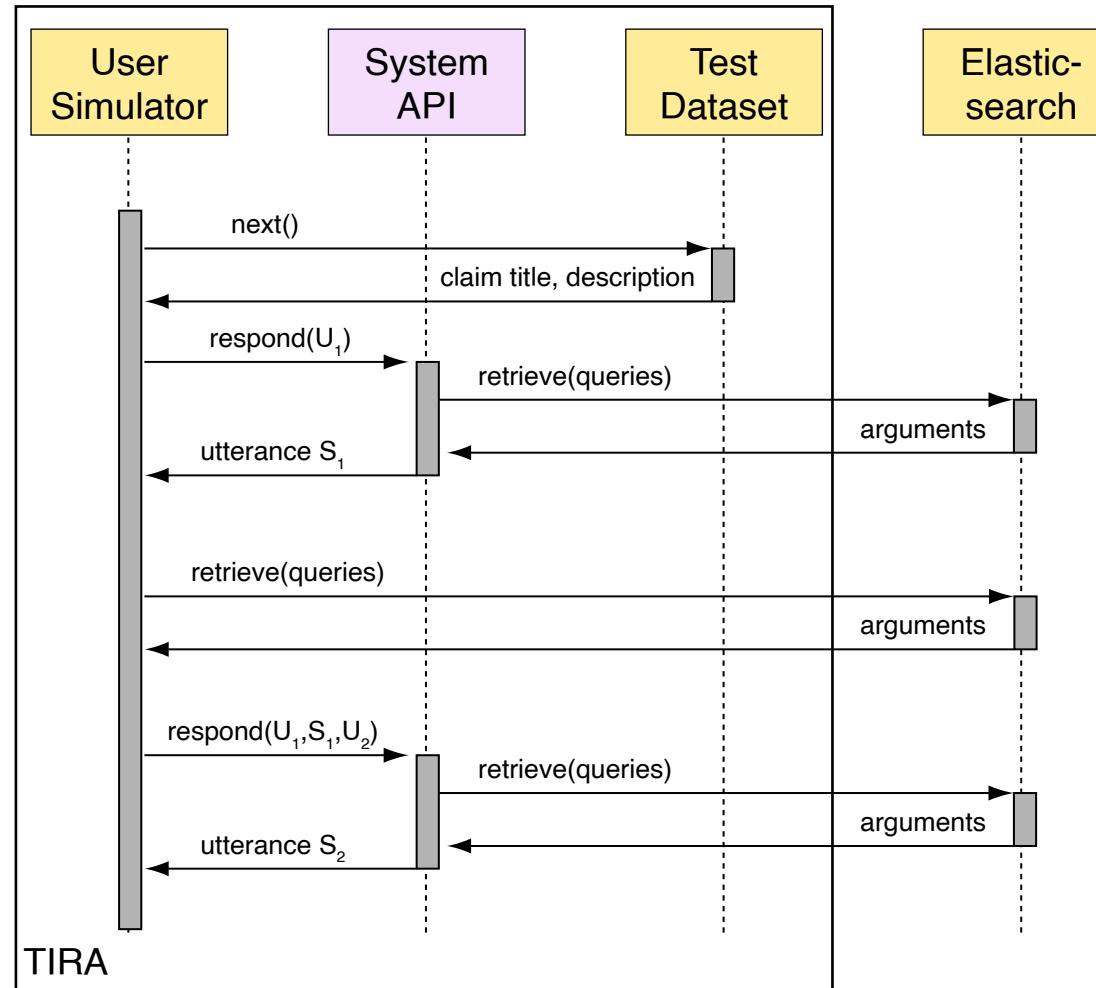
participant

organizer

# Touché RAD

participant

organizer



# Touché RAD

## Participants

- + Low software stack limitations
- + No hardware requirements
- High development complexity
- Risk of run failures on unseen data

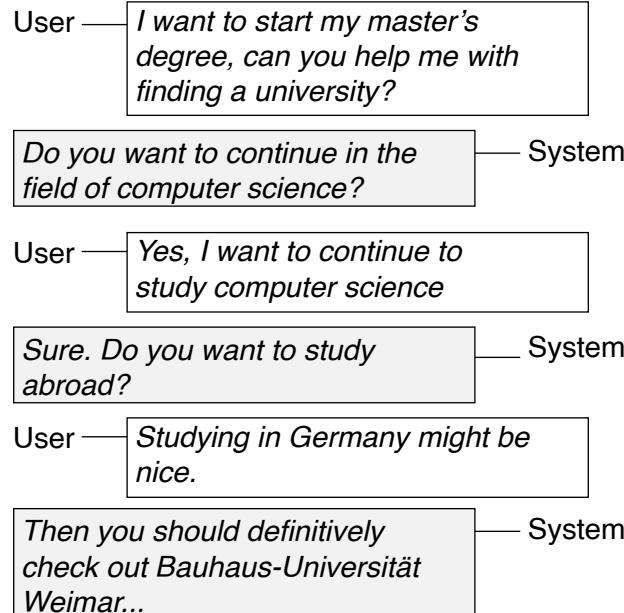
## Organizer

- + Reusable software artifacts
- + Reproduce runs on different datasets
- + Reproduce runs with different simulators
- + Measure and collect meta information
- + Allows secret test datasets
- High support and maintenance effort

## User A ( $t_1$ )

- has a bachelor's degree of computer science.
- lives in the Netherlands
- studied at Tilburg university

### Topic 1: Finding a university



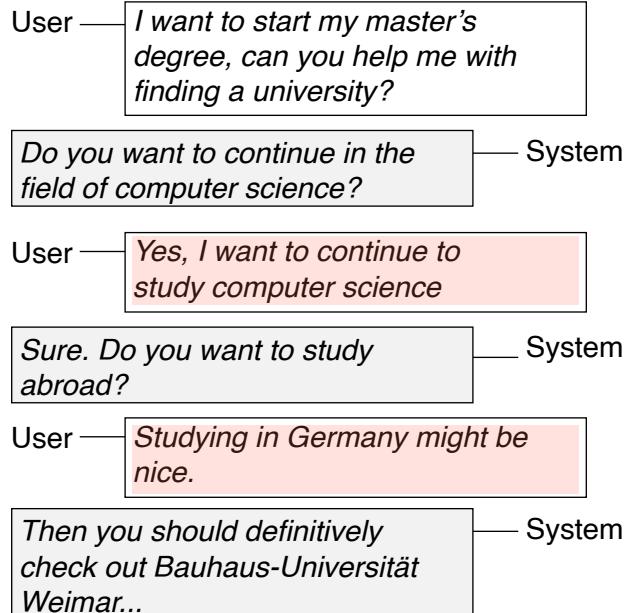
## User A ( $t_1$ )

- has a bachelor's degree of computer science.
- lives in the Netherlands
- studied at Tilburg university

## User A ( $t_2$ )

- has a bachelor's degree of computer science.
- lives in the Netherlands
- studied at Tilburg university
- wants to continue to study computer science.
- wants to study in Germany

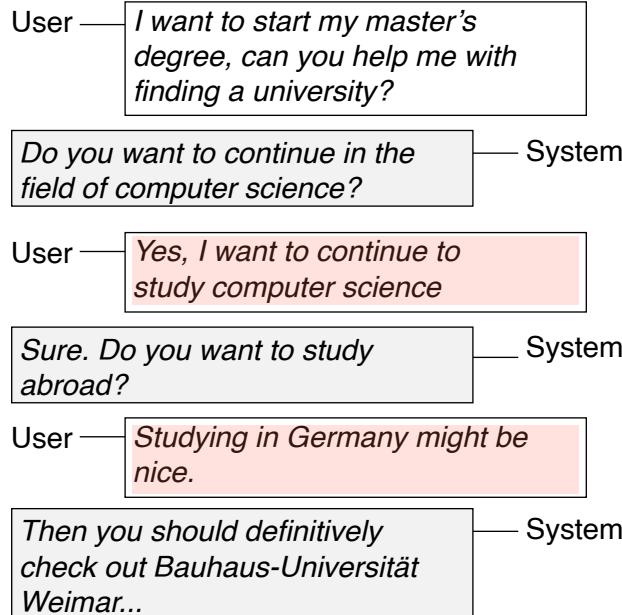
### Topic 1: Finding a university



## User A (t<sub>1</sub>)

- has a bachelor's degree of computer science.
- lives in the Netherlands
- studied at Tilburg university

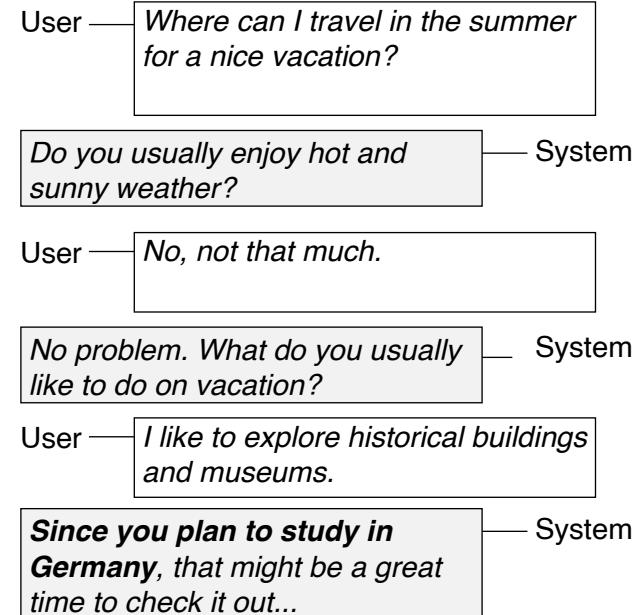
### Topic 1: Finding a university



## User A (t<sub>2</sub>)

- has a bachelor's degree of computer science.
- lives in the Netherlands
- studied at Tilburg university
- **wants to continue to study computer science.**
- **wants to study in Germany**

### Topic 2: Finding a vacation destination



## User A ( $t_1$ )

- has a bachelor's degree of computer science.
- lives in the Netherlands
- studied at Tilburg university

### Topic 1: Finding a university

User — *I want to start my master's degree, can you help me with finding a university?*

*Do you want to continue in the field of computer science?* — System

User — *Yes, I want to continue to study computer science*

*Sure. Do you want to study abroad?* — System

User — *Studying in Germany might be nice.*

*Then you should definitively check out Bauhaus-Universität Weimar...* — System

## User A ( $t_2$ )

- has a bachelor's degree of computer science.
- lives in the Netherlands
- studied at Tilburg university
- wants to continue to study computer science.
- wants to study in Germany

### Topic 2: Finding a vacation destination

User — *Where can I travel in the summer for a nice vacation?*

*Do you usually enjoy hot and sunny weather?* — System

User — *No, not that much.*

*No problem. What do you usually like to do on vacation?* — System

User — *I like to explore historical buildings and museums.*

*Since you plan to study in Germany, that might be a great time to check it out...* — System

# TREC iKAT

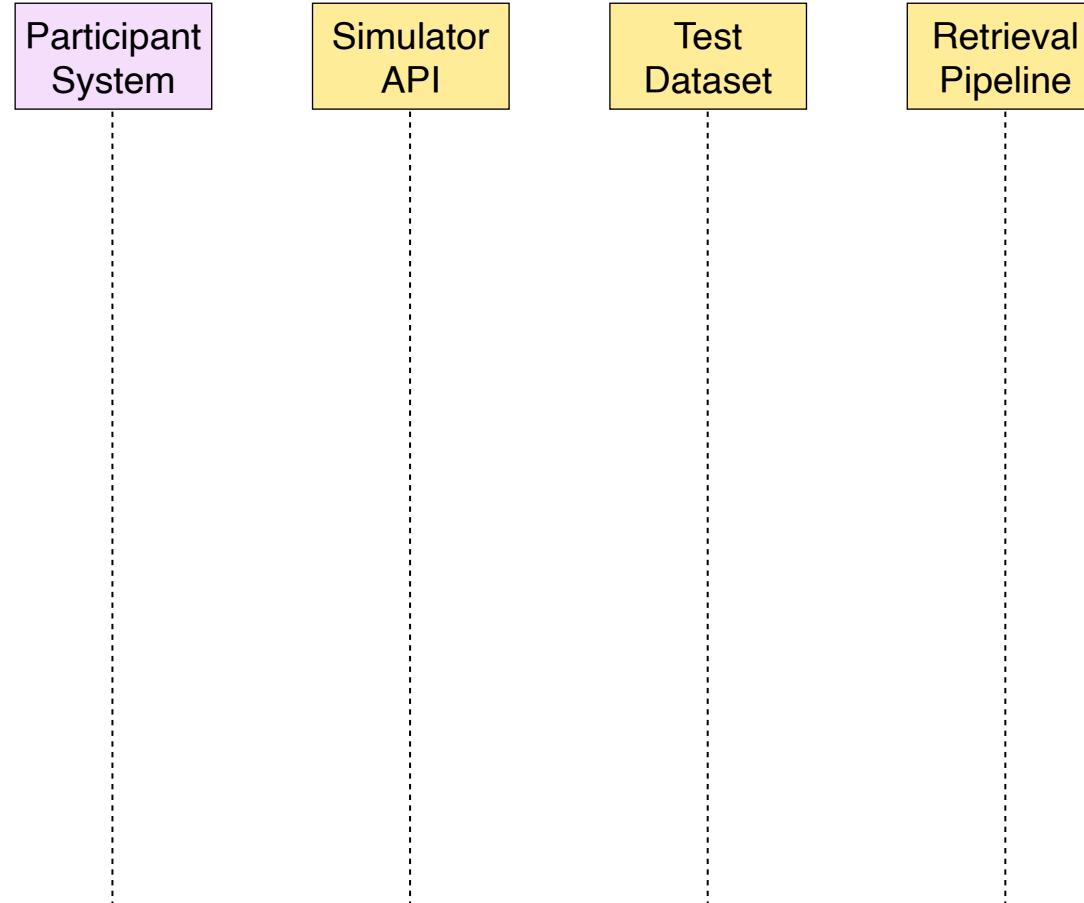
## Submission of participants

- Run files (offline track)
- Submission through API interaction (interactive track)

## Assessment of submissions

- Simulation based on secret simulator
- Simulation based on public test set

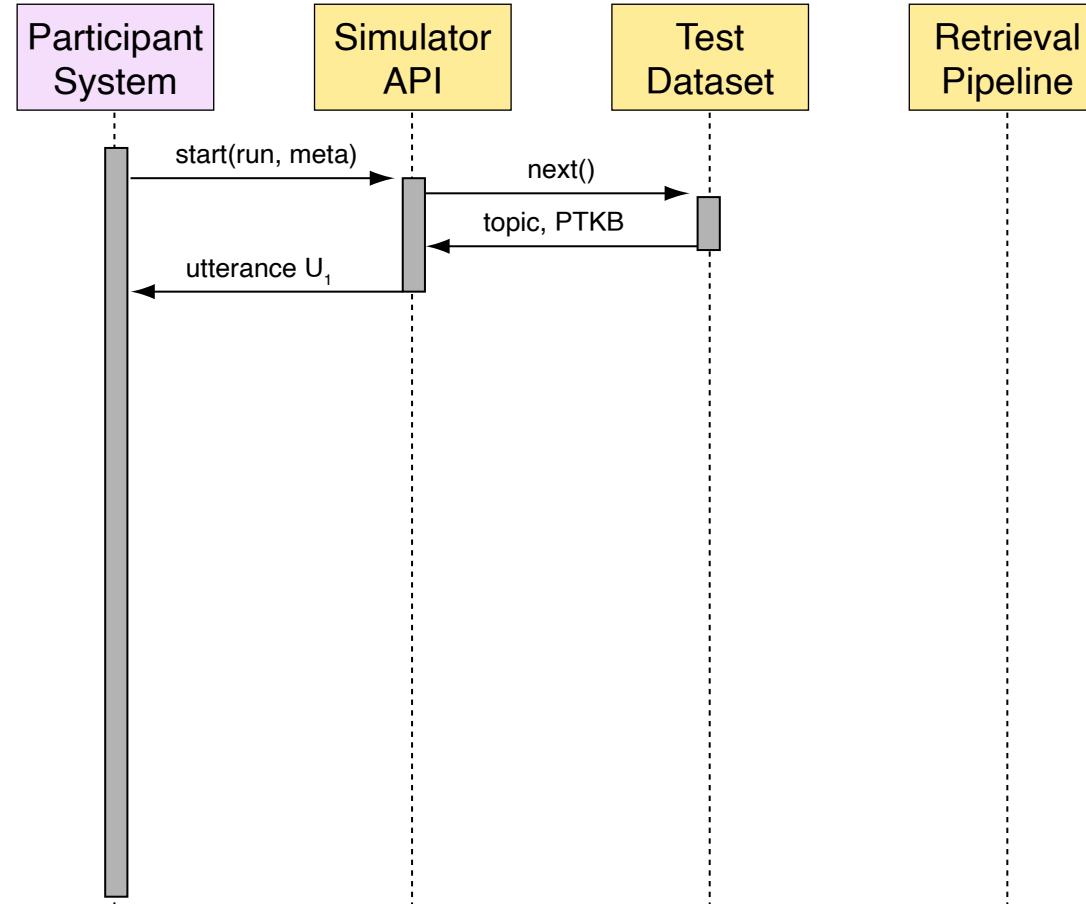
# TREC iKAT



participant

organizer

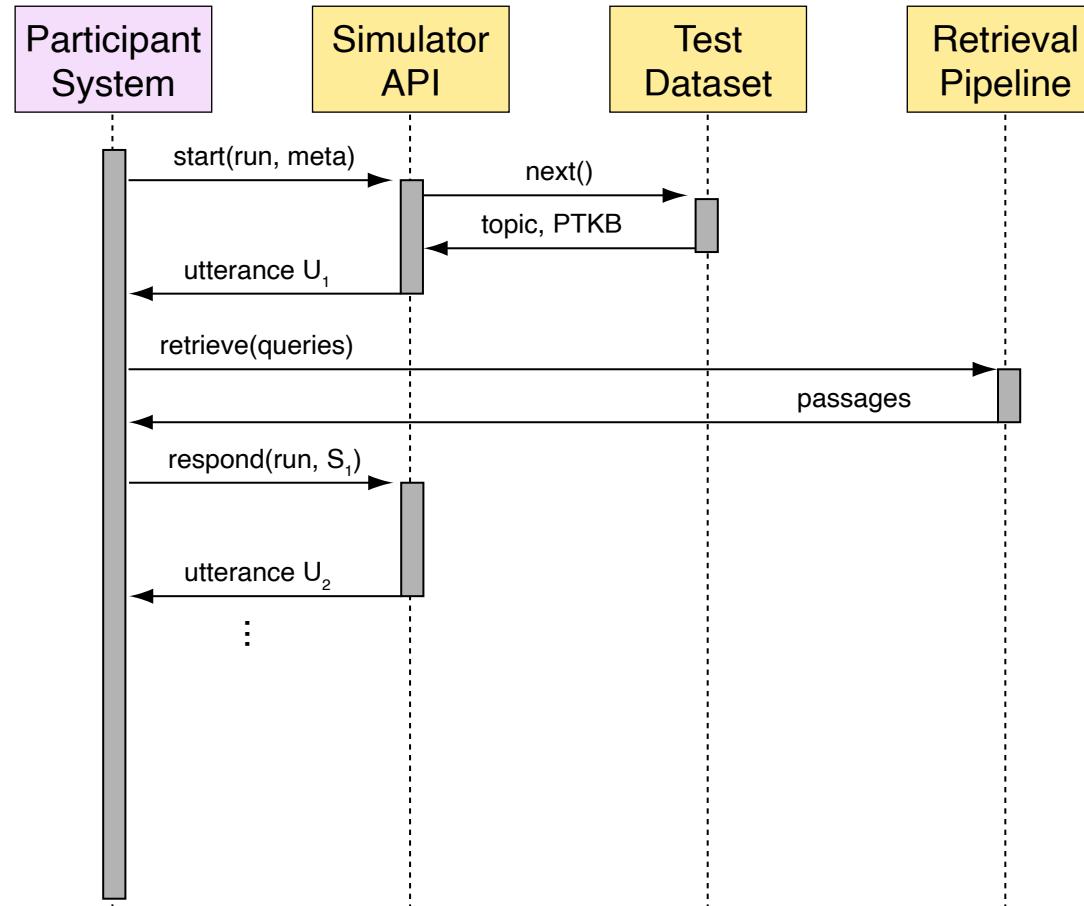
# TREC iKAT



participant

organizer

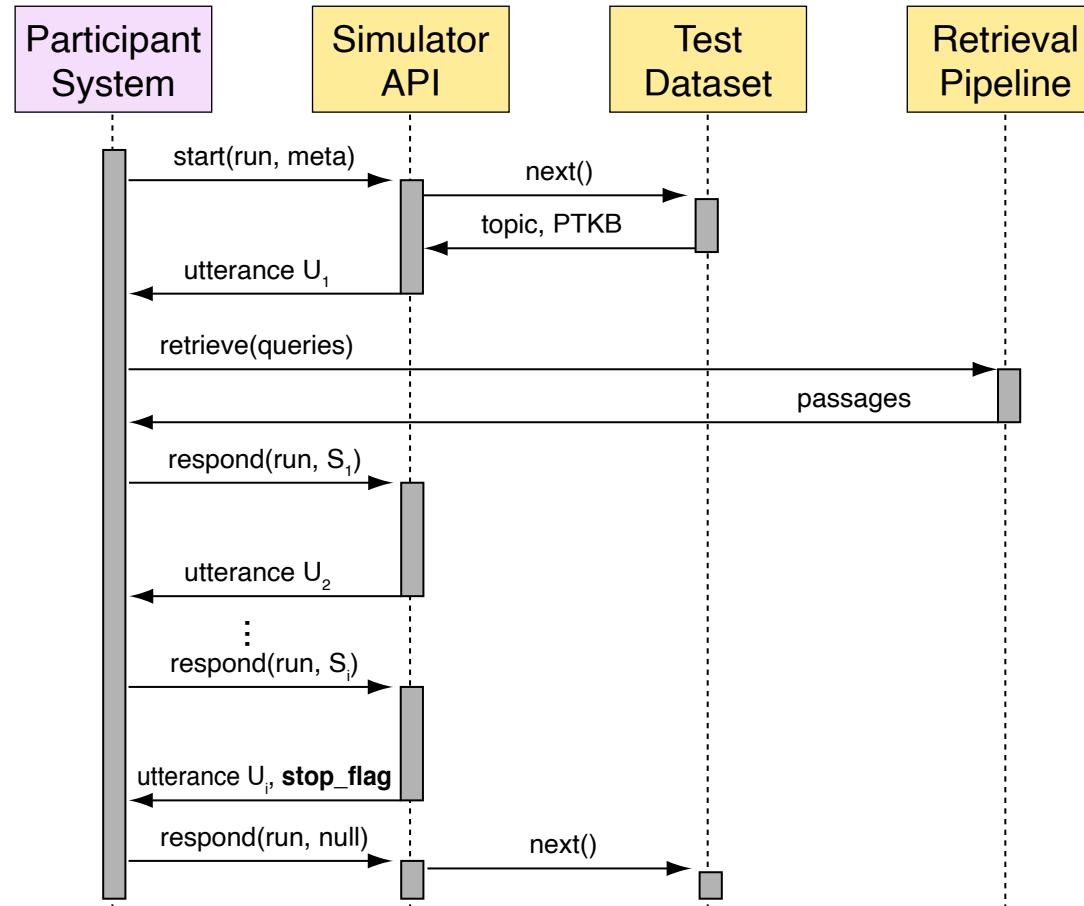
# TREC iKAT



participant

organizer

# TREC iKAT



participant

organizer

## Requirements for simulator API:

- ❑ **Session tracking**  
utterances, retrieved passages, meta information of each run, participating teams
- ❑ **Authentication**  
authenticate participating teams through access tokens
- ❑ **Debugging endpoints**  
allow teams to test their systems before run submission
- ❑ **Budgeting and rate limits**  
limit total number and rate of consecutive requests
- ❑ **Persistent and concurrency-safe storage**  
ACID-compliant storage based on a relational database
- ❑ **Failsafe and recoverable states**  
Recover and continue sessions after server reboot

## Participants

- + No software stack limitations
- + Low development complexity
- + Allows “manual” runs
- + No formatting or run failure problems
- Hardware requirements

## Organizer

- + (Presumably) low support effort
- + Measure and collect meta information
- + Allows secret test datasets
- No reusable software artifacts
- No posthoc simulator or dataset changes

# Findings from User Simulation in Shared Tasks

Important user simulation criteria:

- ❑ Verbosity and clarity
- ❑ Reproducible “trajectory”
- ❑ System control and feedback
- ❑ Dialog act diversity

Organizer requirements:

- ❑ Fair and comparable runs
- ❑ Responsive technical support
- ❑ Simulator evaluation for final assessment

# Join TREC iKAT 2025!

Interactive Knowledge Assistance Track (iKAT).

Deadlines:

- Offline track: 23rd July, 2025.
- Interactive track: TBD.

Organizers:

Mohammad Aliannejadi

Zahra Abbasiantaeb

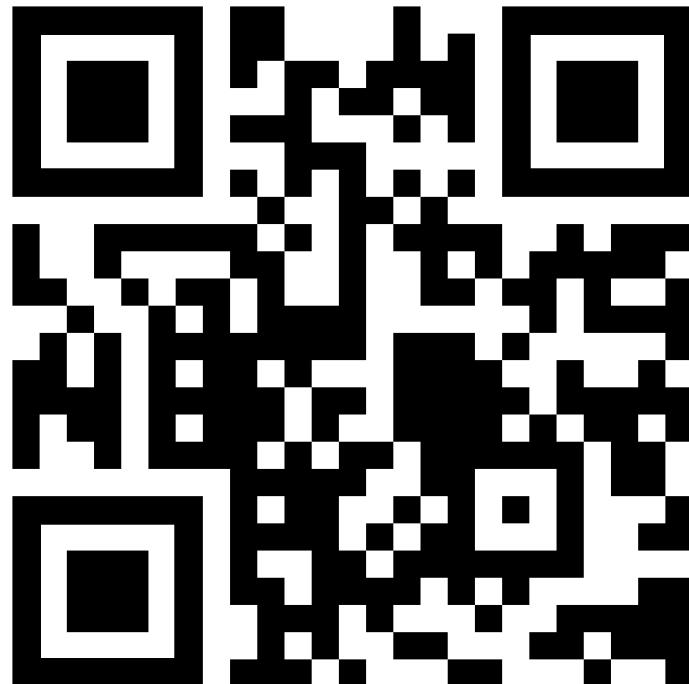
Simon Lupart,

Nailia Mirzakhmedova

Marcel Gohsen

Johannes Kiesel

Jeff Dalton



<https://www.trecikat.com/>

# Join TREC iKAT 2025!

Interactive Knowledge Assistance Track (iKAT).

Deadlines:

- Offline track: 23rd July, 2025.
- Interactive track: TBD.

Organizers:

Mohammad Aliannejadi

Zahra Abbasiantaeb

Simon Lupart,

Nailia Mirzakhmedova

Marcel Gohsen

Johannes Kiesel

Jeff Dalton



<https://www.trecikat.com/>

**Thank you!**

