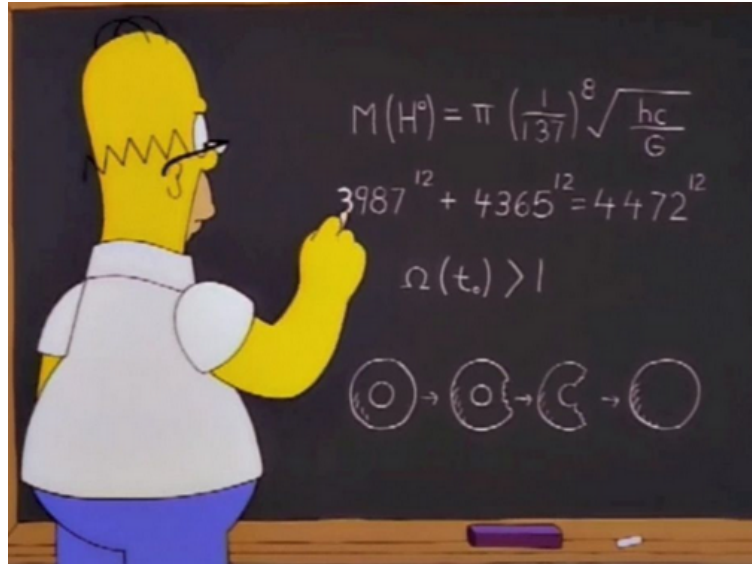


Prototyping Open Web Search Applications with TIRA: A Case Study in Research-oriented Teaching



Maik Fröbe, Theresa Elstner, Harrisen Scells, Benno Stein, Martin Potthast
OSSYM 11–13 October, CERN 2023

@webis_de

www.webis.de

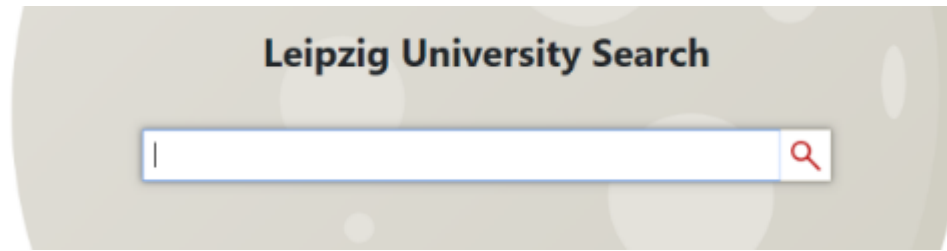
Prototyping Open Web Search in Research-oriented Teaching

an IR course
A long time ago in a ~~galaxy~~ far,
far away....

Prototyping Open Web Search in Research-oriented Teaching

an IR course
A long time ago in a ~~galaxy~~ far,
far away....

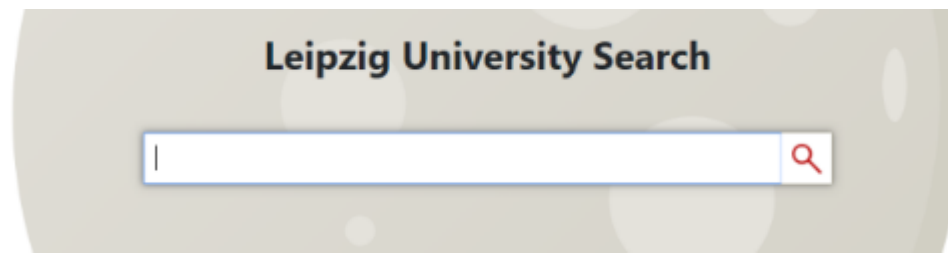
IR Exercise in 2017: Build a Search Engine for a Domain of your Choice



Prototyping Open Web Search in Research-oriented Teaching

an IR course
A long time ago in a ~~galaxy~~ far,
far away....

IR Exercise in 2017: Build a Search Engine for a Domain of your Choice



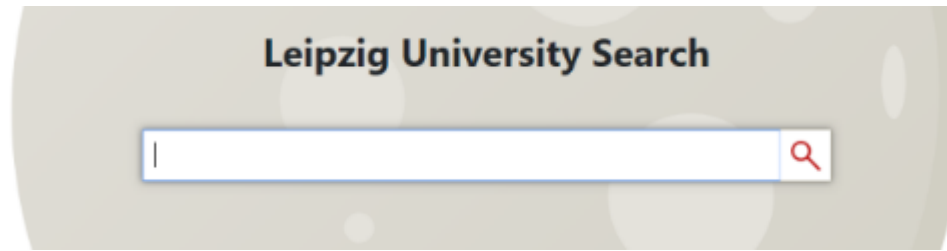
Pro

- ❑ Praxis project in a group
- ❑ Interesting technologies
 - Nutch, Docker, Lucene, JS
- ❑ Cranfield style evaluation
 - Evolved into research

Prototyping Open Web Search in Research-oriented Teaching

an IR course
A long time ago in a ~~galaxy~~ far,
far away....

IR Exercise in 2017: Build a Search Engine for a Domain of your Choice



Pro

- ❑ Praxis project in a group
- ❑ Interesting technologies
 - Nutch, Docker, Lucene, JS
- ❑ Cranfield style evaluation
 - Evolved into research

Con

- ❑ Each group works on a different problem
 - No/Limited shared learning
- ❑ Created resources are barely reusable
 - Much too small, unclear quality, etc.

Prototyping Open Web Search in Research-oriented Teaching

IR Exercise WiSe 2023

Shared Task by Students for Students

Prototyping Open Web Search in Research-oriented Teaching

IR Exercise WiSe 2023

Shared Task by Students for Students

Scenario: Searching the IR Anthology

IR Anthology



Prototyping Open Web Search in Research-oriented Teaching

IR Exercise WiSe 2023

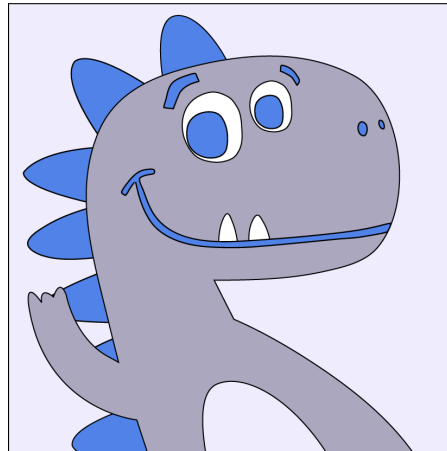
Shared Task by Students for Students

Scenario: Searching the IR Anthology

IR Anthology



Realization with TIREx (= an integration of existing tools)



TIREx = TIRA + ir_datasets + PyTerrier

Prototyping Open Web Search in Research-oriented Teaching

IR Exercise WiSe 2023


Shared Task by Students for Students

Lets look at the results:

`https://www.tira.io/`

Prototyping Open Web Search in Research-oriented Teaching

IR Exercise WiSe 2023: Backup

 TIRA Admin Forum 🔍 ☰ 👤

IR Lab Jena/Leipzig SoSe 2023

by [ir-lab-rose23](#) (contact)

The lab for the information retrieval exercise in sommer semester 2023.

📅 2023 👤 34 📁 0

Submissions

Dataset: [iranthology](#)

<input type="checkbox"/>	Team	Approach	Run	P@10 ↓	RR	nDCG@10	
<input type="checkbox"/>	ir-lab-rose-2023-chirp	MS3-TFIDF_PL2_DPH	2023-06-27-11-39-11	0.582	0.812	0.639	⌵
<input type="checkbox"/>	ir-lab-rose-2023-chirp	CHIRP-MILESTONE3_TFIDF-DPH	2023-06-21-15-49-07	0.581	0.814	0.641	⌵
<input type="checkbox"/>	ir-lab-rose-2023-chirp	chirp-milestone02-tfidf	2023-06-18-18-32-11	0.579	0.792	0.63	⌵
<input type="checkbox"/>	ir-lab-rose-2023-dogument-retriever	tfidf-DPH	2023-06-27-17-32-38	0.578	0.808	0.639	⌵
<input type="checkbox"/>	ir-lab-rose-2023-dogument-retriever	wicker-trunk	2023-06-27-18-31-43	0.578	0.808	0.639	⌵
<input type="checkbox"/>	ir-lab-rose-2023-dogument-retriever	generative-hydrolysis	2023-06-27-23-17-38	0.578	0.806	0.636	⌵
<input type="checkbox"/>	ir-lab-rose-2023-dogument-retriever	single-3	2023-06-27-23-26-14	0.578	0.808	0.639	⌵
<input type="checkbox"/>	ir-lab-rose-2023-irbbb	tfidf-dph	2023-06-26-21-02-34	0.575	0.79	0.63	⌵

Items per page: 1-10 of 189 ⏪ ⏩

- ❑ 189 submissions of Docker images
- ❑ Shared leaderboard

Prototyping Open Web Search in Research-oriented Teaching

IR Exercise WiSe 2023: Backup

The screenshot shows the TIRA search interface. At the top, there is a navigation bar with the TIRA logo, 'Admin Forum', and search icons. Below this is a search bar containing the query 'DiffIR'. A dropdown menu shows the selected query 'retrieval system improving effectiveness'. Below the dropdown is a form with 'query_id 1' and 'title retrieval system improving effectiveness'. The search results are displayed as a list of four items, each with a rank, relevance score, and a snippet of text. The results are:

- 1 Rel: 1 Score: 31.8347
default_text: ...tion **retrieval** systems. Each **system** demonstrated improved **effectiveness** with the use of blind feedback, in which the results of a preliminary **retrieval** step were used to augment the efficacy of a seco...
- 2 Rel: 1 Score: 26.8159
default_text: ...tion **retrieval** systems AbstractNowadays, access to information requires managing multimedia databases effectively, and so, multi-modal **retrieval** techniques (particularly Images **retrieval**) have become ...
- 3 Rel: 1 Score: 26.2787
default_text: ...ment **retrieval** behaviour. In particular relevance feedback query-expansion methods, which are often effective for **improving** electronic text **retrieval**, are observed to be less reliable for **retrieval** of...
- 4 Rel: 1 Score: 26.2697
default_text: ... for **improving** web **retrieval effectiveness** AbstractThis paper talks about several schemes for **improving retrieval effectiveness** that can be used in the named page finding tasks of web Information **retr...**

- Looking into SERPs rendered with DiffIR allows to formulate hypothesis

Prototyping Open Web Search in Research-oriented Teaching

IR Exercise WiSe 2023: Backup

Submissions

Dataset
iranthology

<input type="checkbox"/>	Team	Approach	Run	P@10 ↓	RR	nDCG@10	
<input type="checkbox"/>	ir-lab-2023-document-retriever	PL2-KL	2023-06-24-16-12-04	0.39	0.66	0.438	↓
<input type="checkbox"/>	ir-lab-2023-memory	Multi-Field	2023-06-21-03-15-35	0.387	0.722	0.459	↓
<input type="checkbox"/>	ir-lab-2023-order-google-scholar	IN-EXPC2	2023-06-25-23-09-51	0.387	0.691	0.447	↓
<input type="checkbox"/>	ir-lab-2023-order-google-scholar	lemurf_idf	2023-06-25-23-05-08	0.36	0.674	0.42	↓
<input type="checkbox"/>	ir-lab-2023-order-google-scholar	lemurf_idf-2	2023-06-26-21-20-32	0.36	0.674	0.42	↓
<input type="checkbox"/>	ir-lab-2023-tutors	<u>IR-Anthology 2022</u>	2023-06-19-18-48-35	0.31	0.664	0.371	↓
<input type="checkbox"/>	ir-lab-2023-tutors	ANCE	2023-06-18-15-47-12	0.271	0.496	0.294	↓
<input type="checkbox"/>	ir-lab-2023-memory	bm25-sdm	2023-06-20-13-28-17	0.265	0.68	0.366	↓

Items per page: 10 171-180 of 189 |< < > >|

- ❑ Almost all student teams outperformed our production system

Prototyping Open Web Search in Research-oriented Teaching

IR Exercise WiSe 2023

Shared Task by Students for Students

Lets look at the results: <https://www.tira.io/>

Prototyping Open Web Search in Research-oriented Teaching

IR Exercise WiSe 2023

Shared Task by Students for Students

Lets look at the results: <https://www.tira.io/>

Realization + Short Hands-On

- ❑ Milestone 1: Data creation + Pooling + Judgment
- ❑ Milestone 2: Create software submissions
- ❑ Milestone 3: Write up + fill judgment holes

Prototyping Open Web Search in Research-oriented Teaching

IR Exercise WiSe 2023

Shared Task by Students for Students

Lets look at the results: <https://www.tira.io/>

Realization + Short Hands-On

- ❑ Milestone 1: Data creation + Pooling + Judgment
- ❑ Milestone 2: Create software submissions
- ❑ Milestone 3: Write up + fill judgment holes

Pro

- ❑ Standardization + reusability
- ❑ Did scale very well
- ❑ We enrich “something” with complete experiment setups
 - Data + dockerized software

Con

- ❑ Quality of resulting data is not clear
 - Not yet clear if observations from previous work transfer
[Sakai'22,Schaer'12]
- ❑ I hoped that student teams build more on top of each other

Prototyping Open Web Search in Research-oriented Teaching

IR Exercise WiSe 2023

Shared Task by Students for Students

Lets look at the results: <https://www.tira.io/>

Realization + Short Hands-On

- ❑ Milestone 1: Data creation + Pooling + Judgment
- ❑ Milestone 2: Create software submissions
- ❑ Milestone 3: Write up + fill judgment holes

Pro

- ❑ Standardization + reusability
- ❑ Did scale very well
- ❑ We enrich “something” with complete experiment setups
 - Data + dockerized software

Con

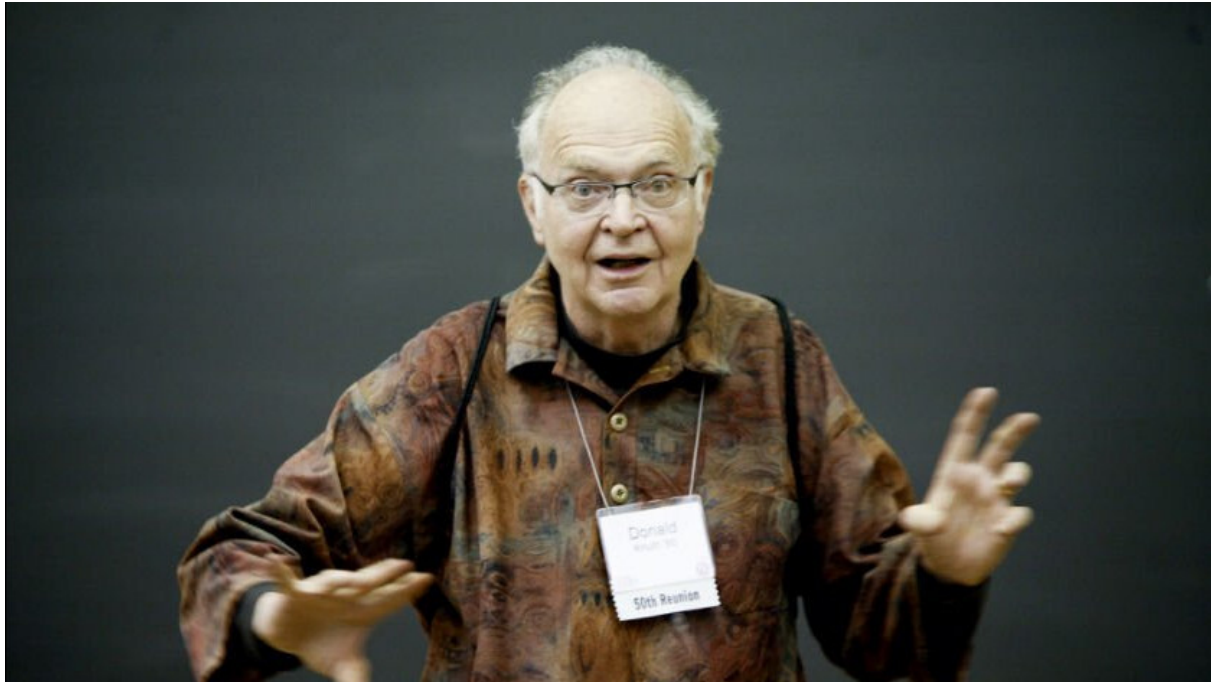
- ❑ Quality of resulting data is not clear
 - Not yet clear if observations from previous work transfer
[Sakai'22,Schaer'12]
- ❑ I hoped that student teams build more on top of each other

Adjustments for SoSe 2023/24 to Milestone 3 resolves the cons

Mini components + submission to LongEval

Prototyping Open Web Search in Research-oriented Teaching

Vision: Make many “mini stones” available to students

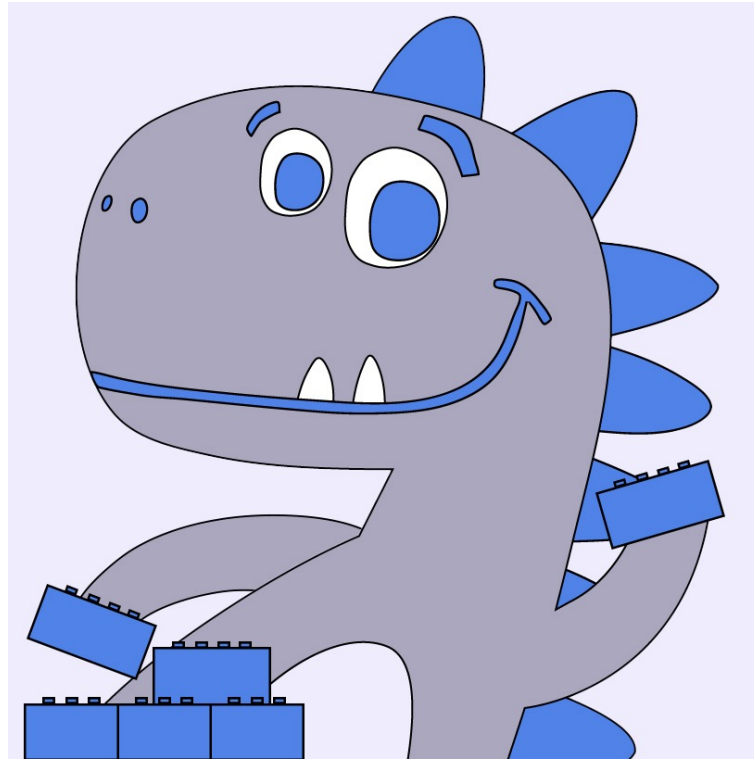


“People think that computer science is the art of geniuses but the actual reality is the opposite, just many people doing things that build on each other, like a wall of mini stones.”

Donald Knuth

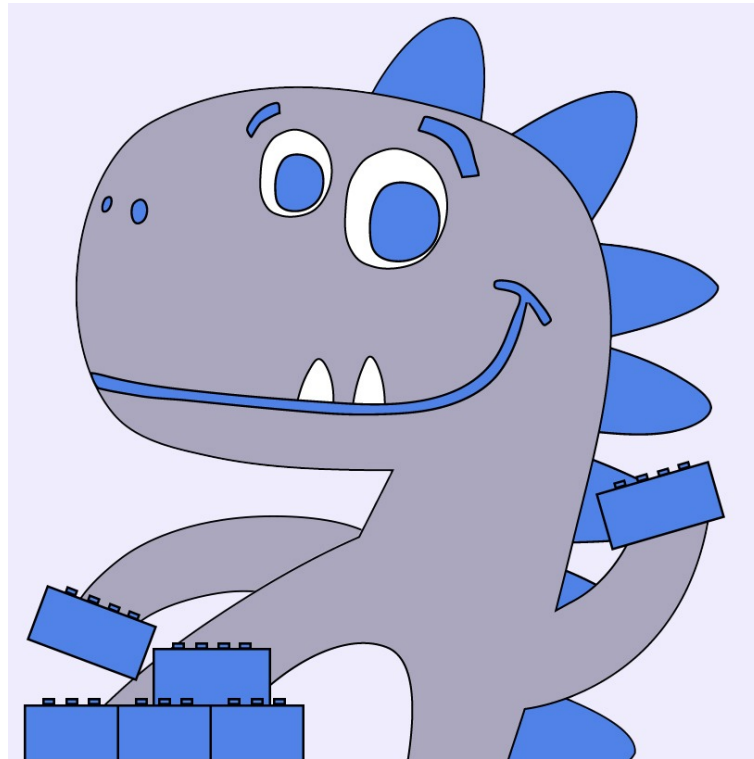
Prototyping Open Web Search in Research-oriented Teaching

We want to follow this idea with TIREx



Prototyping Open Web Search in Research-oriented Teaching

We want to follow this idea with TIREx



Goal: Simplify to compose reproducible and declarative pipelines

- ❑ Workshop proposal under review at ECIR'24: Call for reusable components
- ❑ Many components must only be executed “once in a lifetime”
- ❑ Students can use expert components while focusing on the modeling part

Prototyping Open Web Search in Research-oriented Teaching

Example: Query Segmentation

WWW 2011 – Session: Query Analysis

March 28–April 1, 2011, Hyderabad, India

Query Segmentation Revisited

Matthias Hagen

Martin Potthast

Benno Stein

Christof Bräutigam

Prototyping Open Web Search in Research-oriented Teaching

Example: Query Segmentation

WWW 2011 – Session: Query Analysis

March 28–April 1, 2011, Hyderabad, India

Query Segmentation Revisited

Matthias Hagen

Martin Potthast

Benno Stein

Christof Bräutigam

- ❑ **Query:** hubble telescope achievements
- ❑ **Segments:** hubble telescope | achievements
- ❑ hubble telescope achievements \Rightarrow "hubble telescope" achievements

Prototyping Open Web Search in Research-oriented Teaching

Execute the 2011 query segmentation... Before submission to TIREx/TIRA

INSTRUCTIONS



Prototyping Open Web Search in Research-oriented Teaching

Execute the 2011 query segmentation... Before submission to TIREx/TIRA

INSTRUCTIONS



Prototyping Open Web Search in Research-oriented Teaching

Query segmentation using PyTerrier of component submitted to TIREx/TIRA

Setup Retrieval Pipeline

```
query_segmentation = tira.pt.transform_queries('webis-query-segmentation/hyb-a', dataset)

pipeline = query_segmentation >> rewrite_segments_to_phrases >> bm25
```

[5]

Python

Prototyping Open Web Search in Research-oriented Teaching

Query segmentation using PyTerrier of component submitted to TIREx/TIRA

Setup Retrieval Pipeline

```
query_segmentation = tira.pt.transform_queries('webis-query-segmentation/hyb-a', dataset)

pipeline = query_segmentation >> rewrite_segments_to_phrases >> bm25
```

[5]

Python

- ❑ Improves nDCG@10 from 0.00 to 0.31 for `hubble telescope achievements`
- ❑ Use cached outputs from TIREx if available, otherwise use Docker image

Conclusions

Summary

- ❑ Students develop IR systems in a shared task style scenario
- ❑ Last term: IR Anthology as corpus
- ❑ Next term: LongEval as corpus + optional submission to this shared task
- ❑ Some future term: Selective search against the OWS index?
 - We can prototype open web search applications that way

Conclusions

Summary

- ❑ Students develop IR systems in a shared task style scenario
- ❑ Last term: IR Anthology as corpus
- ❑ Next term: LongEval as corpus + optional submission to this shared task
- ❑ Some future term: Selective search against the OWS index?
 - We can prototype open web search applications that way

Future Work

- ❑ Hopefully ECIR workshop to collect more dockerized retrieval components
- ❑ ir_pad for teaching: Small standardized notebooks
- ❑ Pipeline search

Conclusions

Summary

- ❑ Students develop IR systems in a shared task style scenario
- ❑ Last term: IR Anthology as corpus
- ❑ Next term: LongEval as corpus + optional submission to this shared task
- ❑ Some future term: Selective search against the OWS index?
 - We can prototype open web search applications that way

Future Work

- ❑ Hopefully ECIR workshop to collect more dockerized retrieval components
- ❑ ir_pad for teaching: Small standardized notebooks
- ❑ Pipeline search

thank you!