

# Perspectives on Evaluating Diverse Open Web Search Applications

---



IR Lab, March 22, Padua, Italy

Maik Fröbe

Friedrich-Schiller-Universität Jena

# Evaluating Diverse Open Web Search Applications with **TIREx**

## Motivation

strumenti per replicabilità esperimenti IR

### Michael Granitzer

Leiter OpenWebSearch.eu



"I want to  
choose my  
search engine  
like my daily  
newspaper"

# Evaluating Diverse Open Web Search Applications with TIREx

What is Hindering a **Diverse and Vibrant Search Ecosystem?**

Sociological factors:

“Competition is just one click away.”

Eric Schmidt, former CEO of Google

# Evaluating Diverse Open Web Search Applications with TIREx

## What is Hindering a Diverse and Vibrant Search Ecosystem?

Sociological factors:

“Competition is just one click away.”

Eric Schmidt, former CEO of Google

**Tech @ Cliqz**

search privacy web monopoly

**Google—Competition is just one click and  
27 billion US dollars away**

# Evaluating Diverse Open Web Search Applications with TIREx

## What is Hindering a Diverse and Vibrant Search Ecosystem?

Sociological factors:

“Competition is just one click away.”

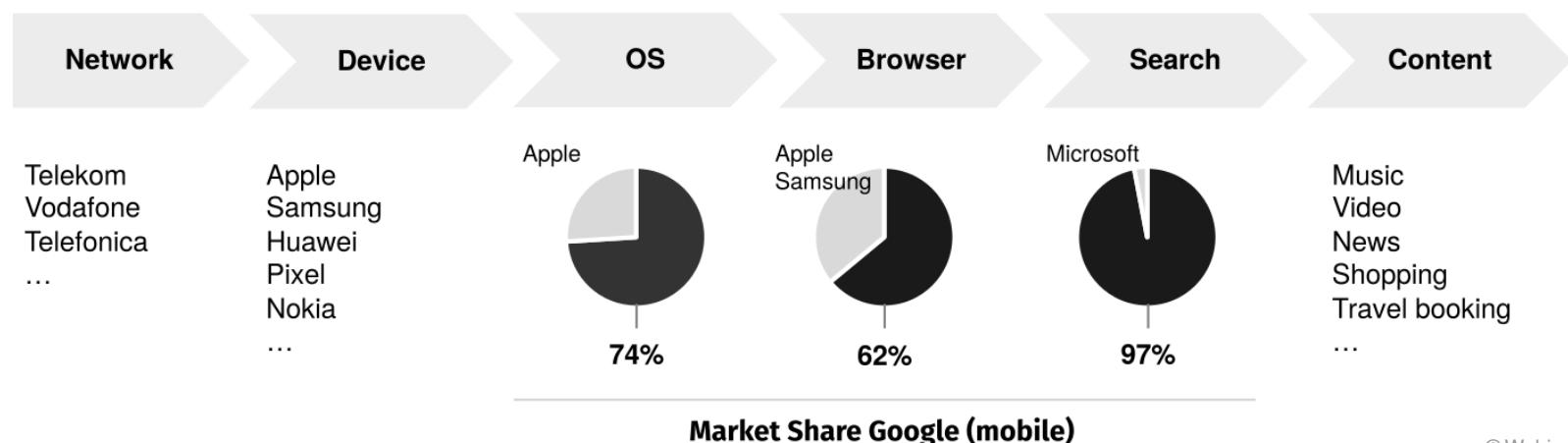
Eric Schmidt, former CEO of Google

### Tech @ Cliqz

search privacy web monopoly

**Google—Competition is just one click and 27 billion US dollars away**

The internet value chain:



# Evaluating Diverse Open Web Search Applications with TIREx

What is Hindering a Diverse and Vibrant Search Ecosystem?

Sociological factors:

“Competition is just one click away.”

Eric Schmidt, former CEO of Google

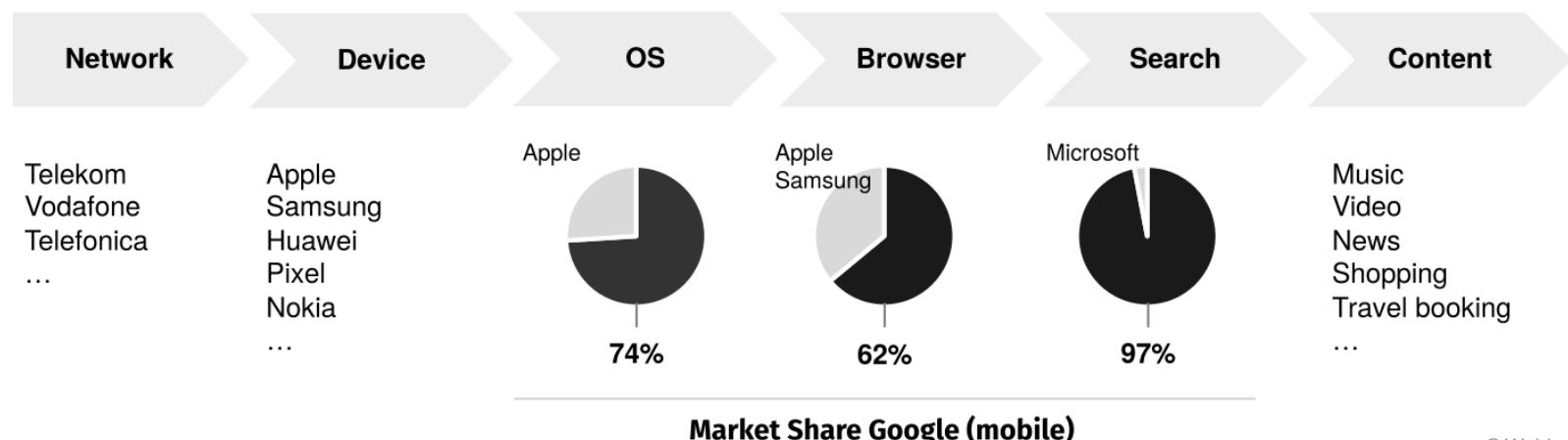
**Solution not within our scope**  
**Still, positive examples exist**

## Tech @ Cliqz

search privacy web monopoly

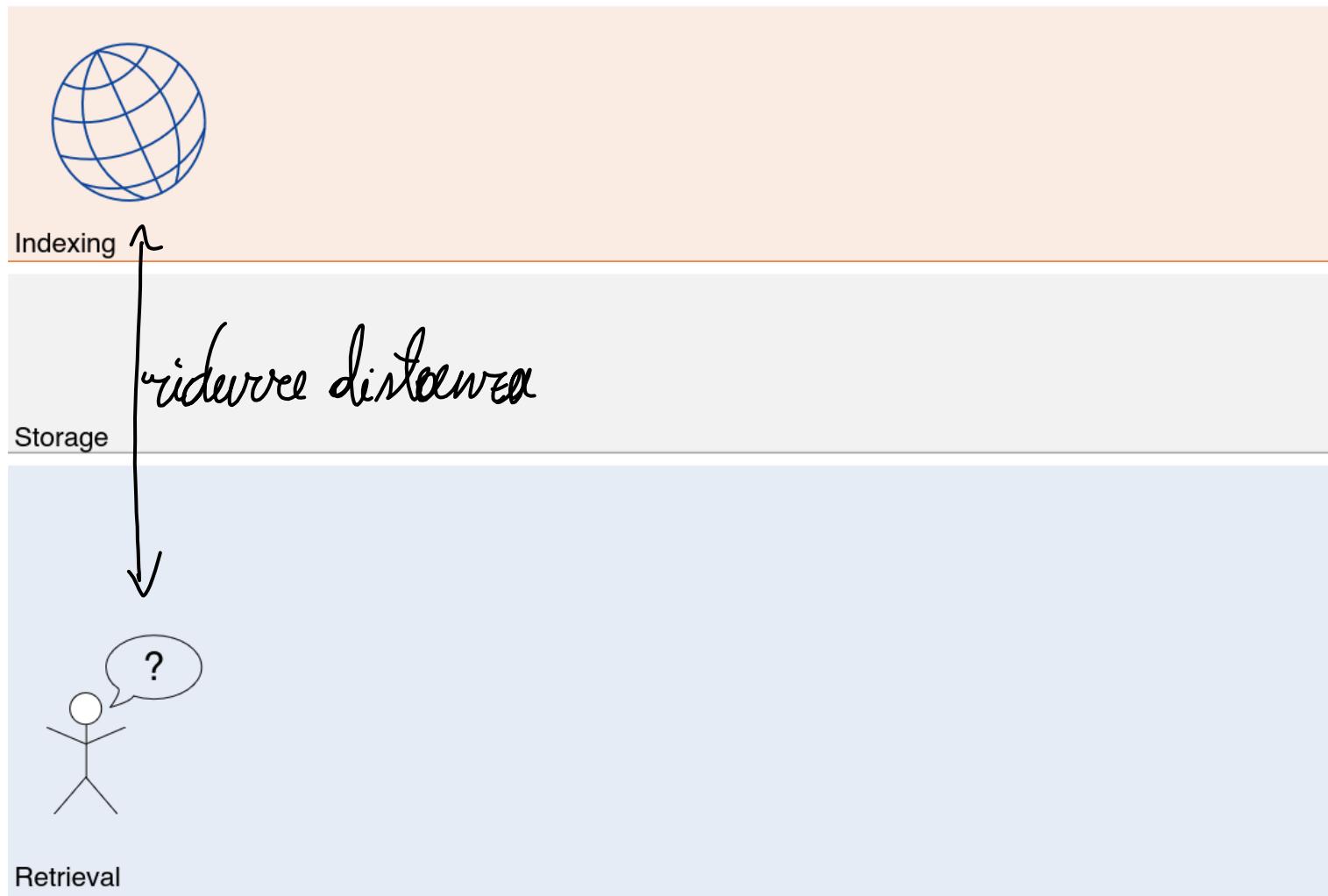
**Google—Competition is just one click and 27 billion US dollars away**

The internet value chain:



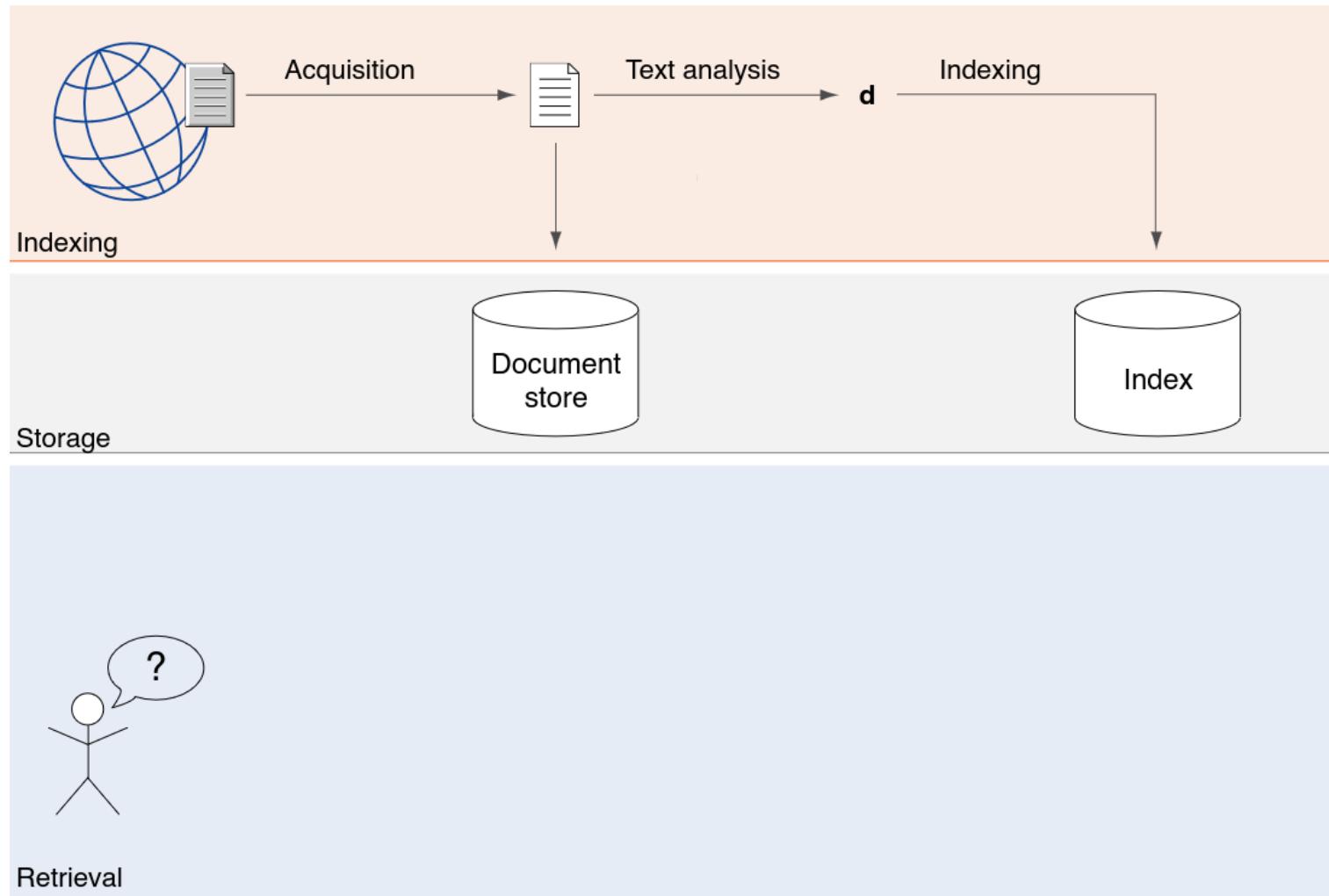
# Evaluating Diverse Open Web Search Applications with TIREx

What Technology Enables a Diverse and Vibrant Search Ecosystem?



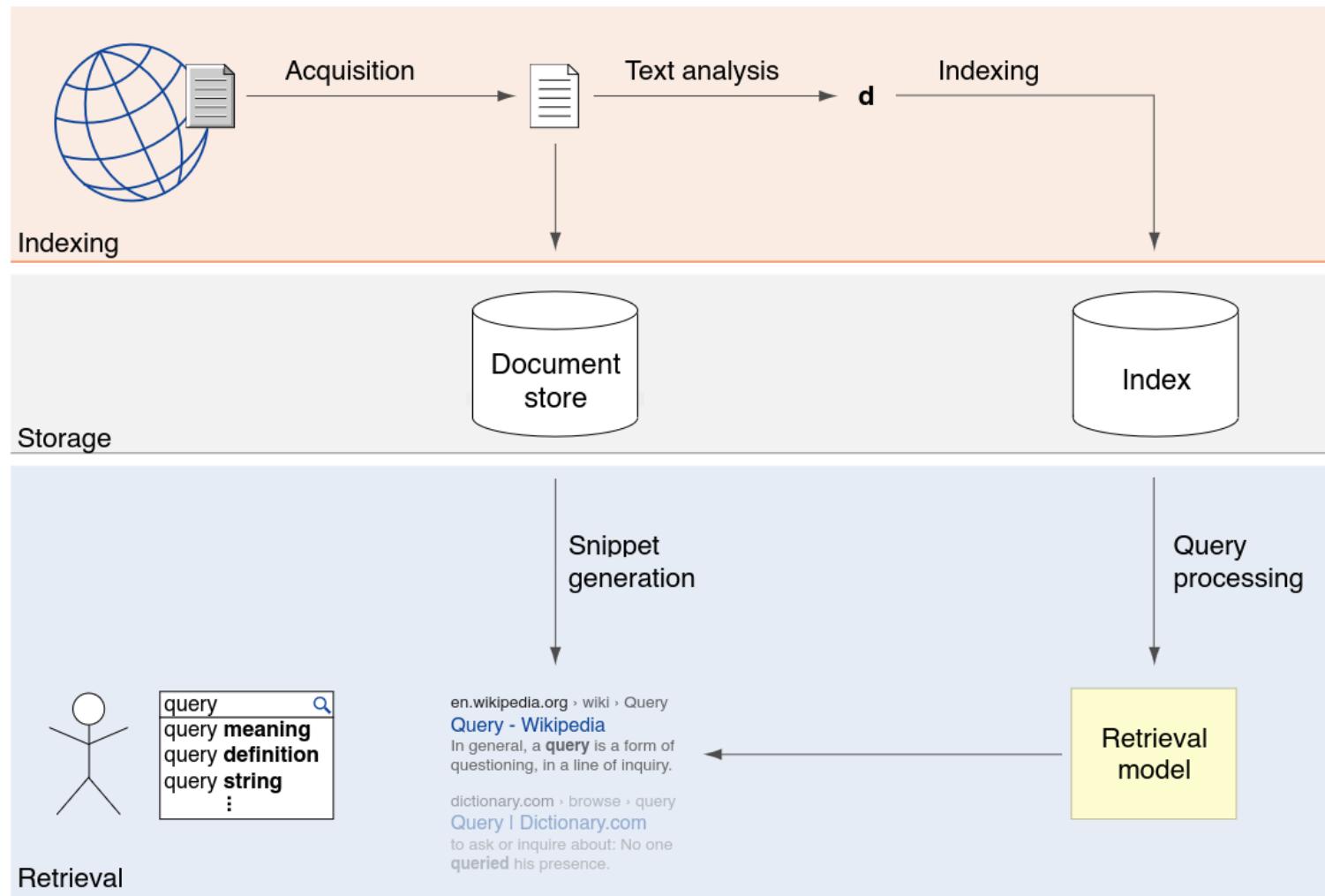
# Evaluating Diverse Open Web Search Applications with TIREx

What Technology Enables a Diverse and Vibrant Search Ecosystem?



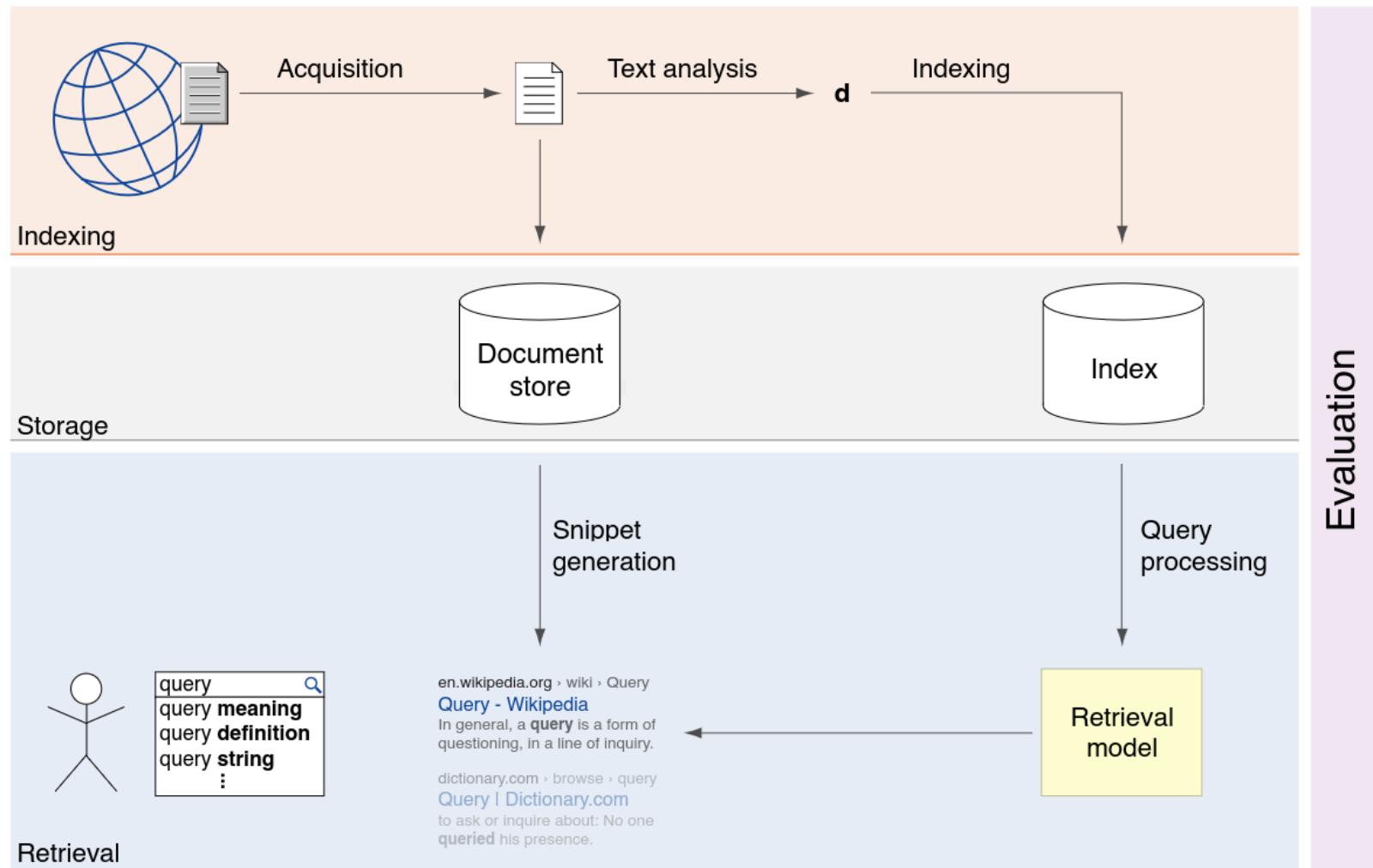
# Evaluating Diverse Open Web Search Applications with TIREx

What Technology Enables a Diverse and Vibrant Search Ecosystem?



# Evaluating Diverse Open Web Search Applications with TIREx

What Technology Enables a Diverse and Vibrant Search Ecosystem?



# Evaluating Diverse Open Web Search Applications with TIREx

The Open Web Index To the Rescue?



# Evaluating Diverse Open Web Search Applications with TIREx

## The Open Web Index: Overview of Partners

14 partners + 3rd party calls

- Research, infrastructure, industry, and NGOs

### Research

Information Retrieval, AI, HCI, Geo-spatial Data Processing



### Infrastructure Organisations

Data Storage, HPC, Services and Scientific Computing



Associations  
for a future Web /Internet



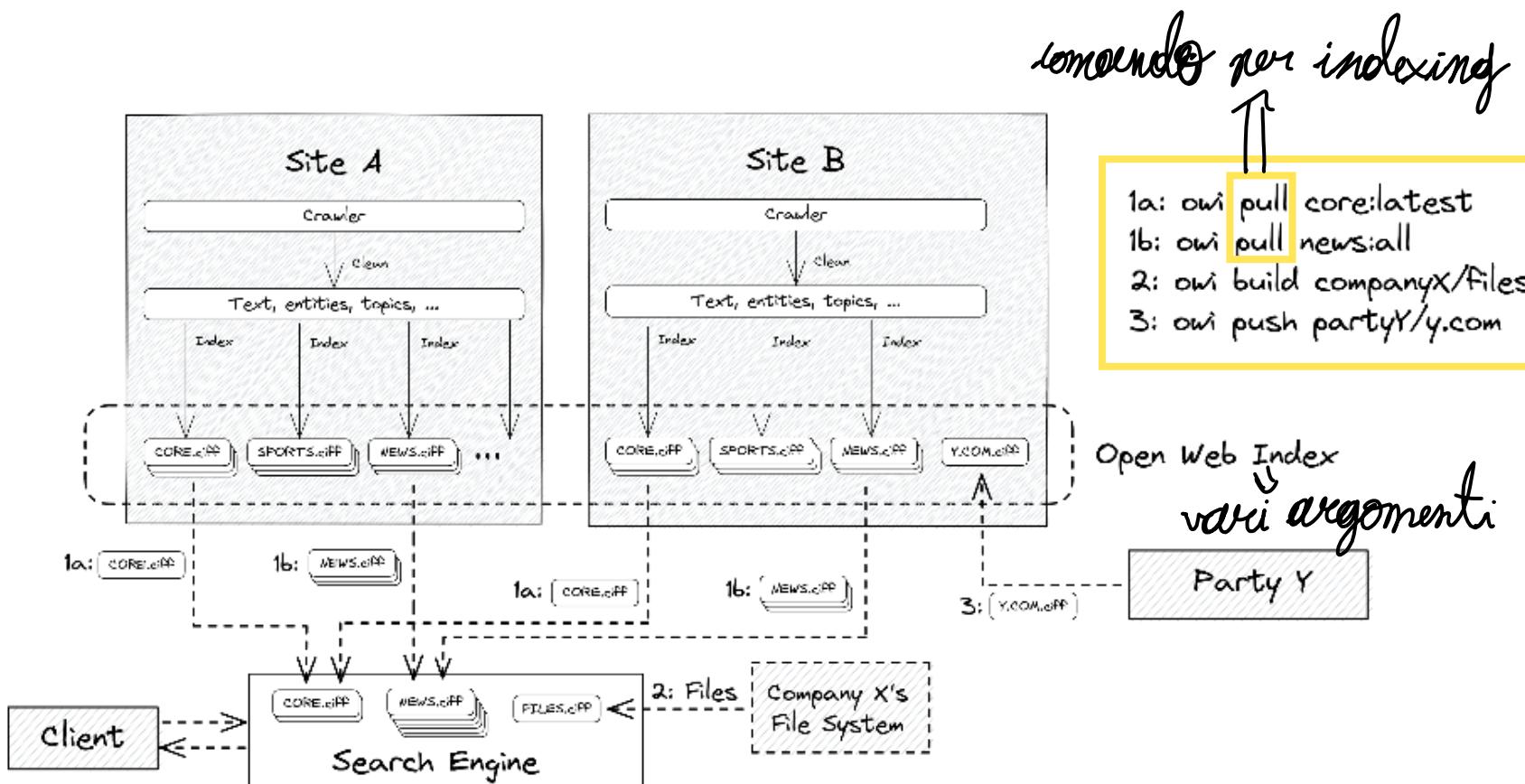
# Evaluating Diverse Open Web Search Applications with TIREx

## The Open Web Index: Overview of Partners



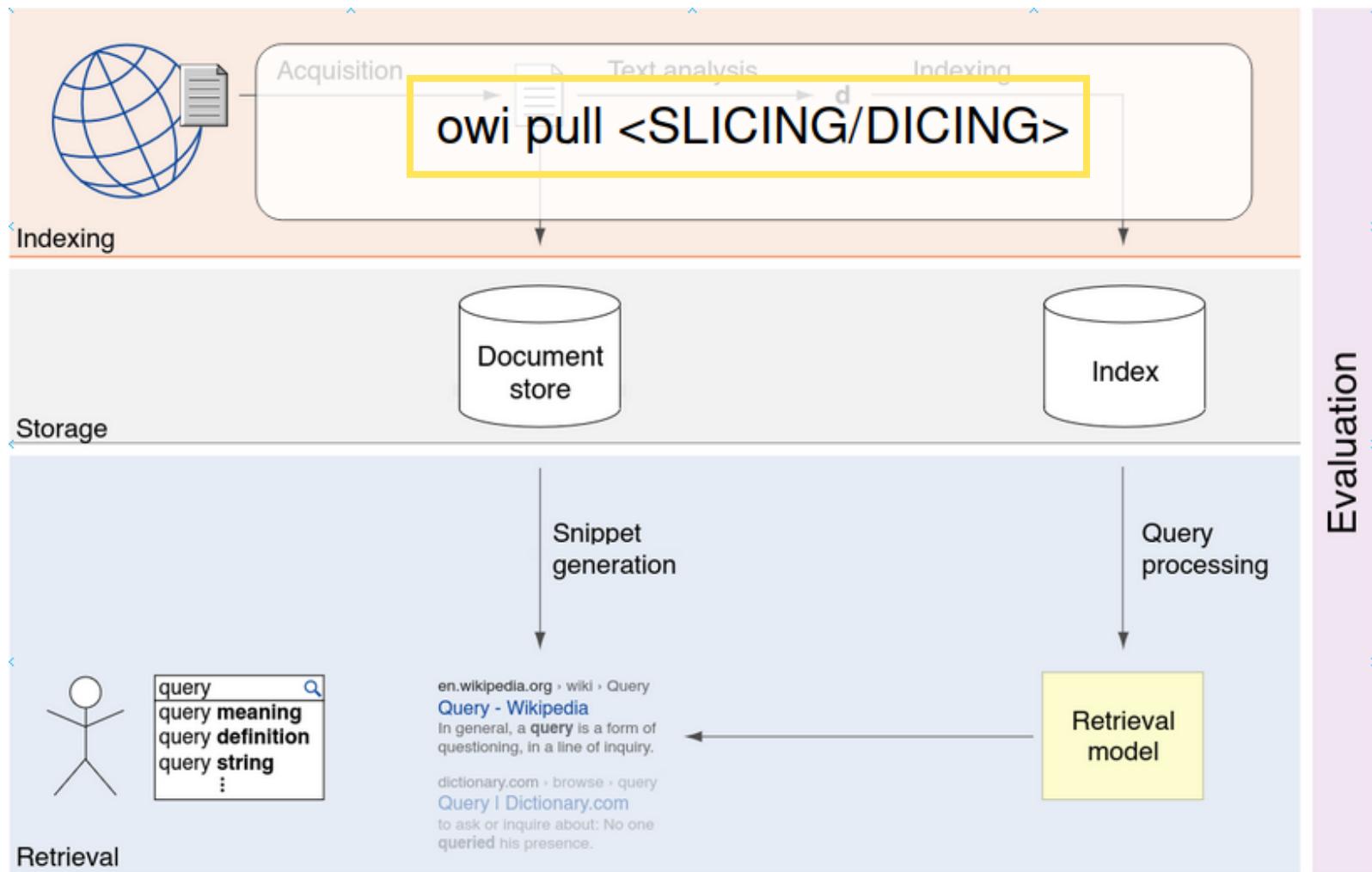
# Evaluating Diverse Open Web Search Applications with TIREx

Slicing and Dicing the Open Web Index for a Diverse Search Ecosystem



# Evaluating Diverse Open Web Search Applications with TIREx

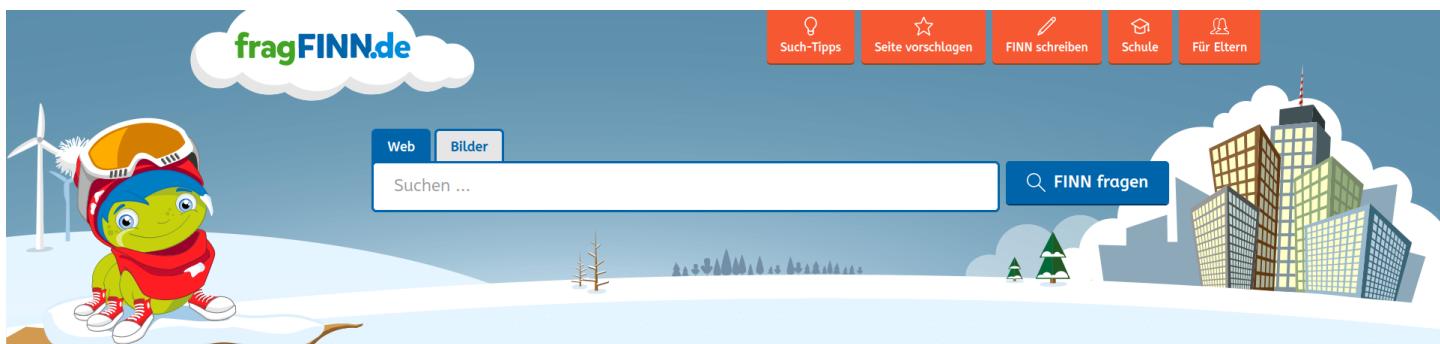
Slicing and Dicing the Open Web Index for a Diverse Search Ecosystem



# Evaluating Diverse Open Web Search Applications with TIREx

## A (biased) Selection of Use-Cases of Slicing and Dicing the Open Web Index

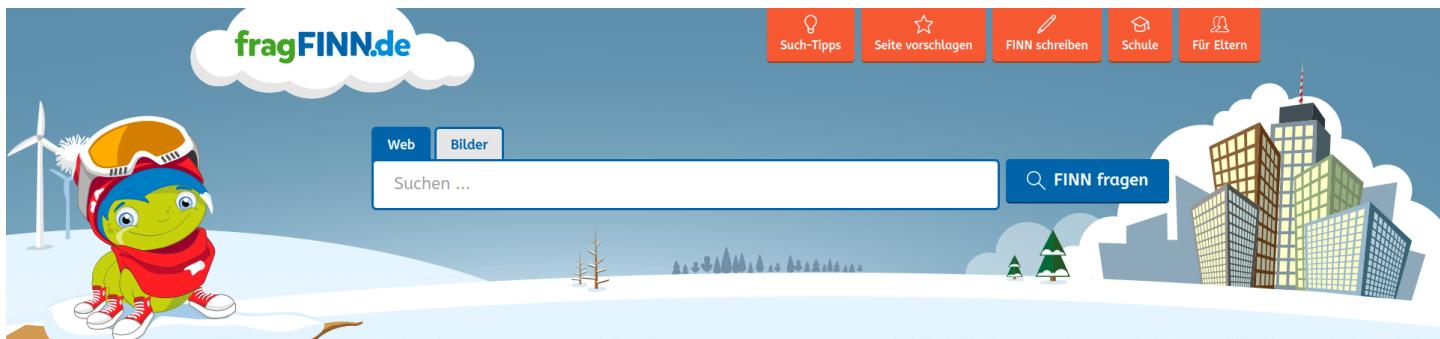
Search engine for kids:



# Evaluating Diverse Open Web Search Applications with TIREx

## A (biased) Selection of Use-Cases of Slicing and Dicing the Open Web Index

Search engine for kids:



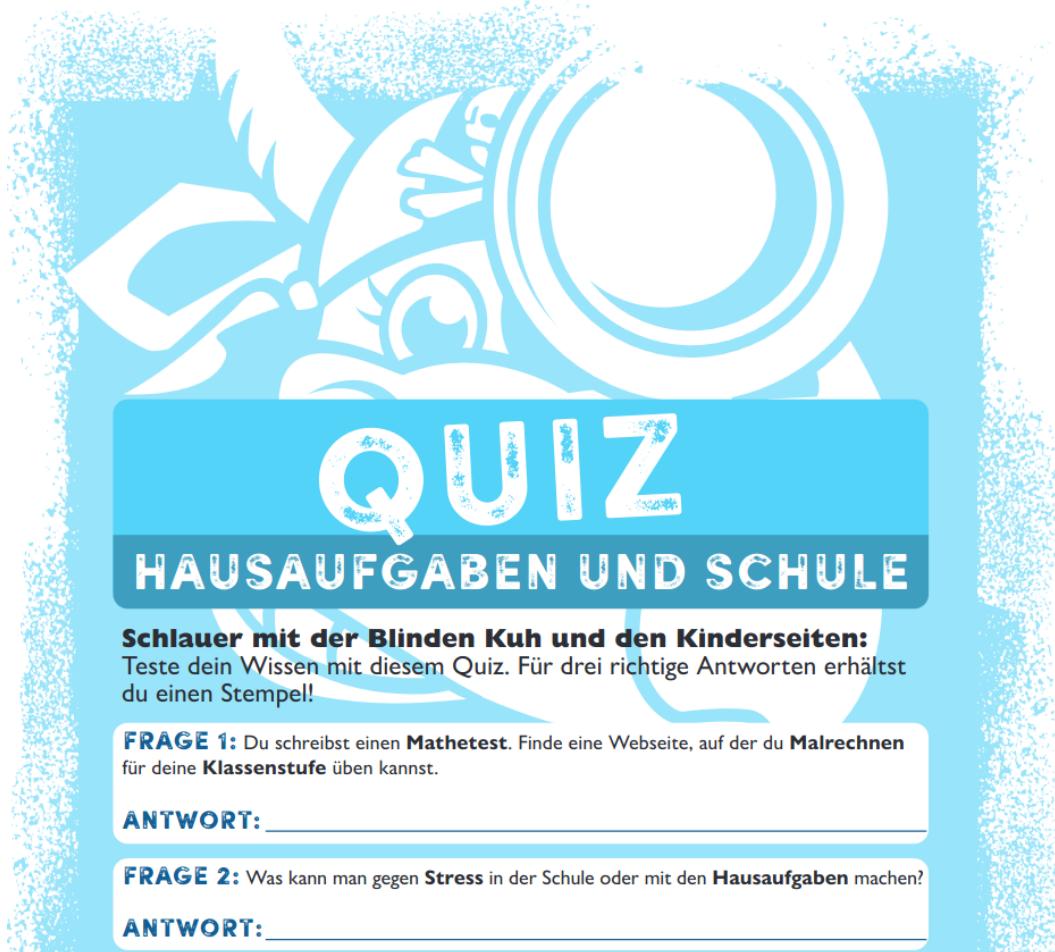
Substantial differences to commercial search:

- ❑ Index size: 1GB
- ❑ Encoded values, guidance, manual curation, ...

# Evaluating Diverse Open Web Search Applications with TIREx

## A (biased) Selection of Use-Cases of Slicing and Dicing the Open Web Index

Search engine for kids:



- We currently build a test collection

# Evaluating Diverse Open Web Search Applications with TIREx

## A (biased) Selection of Use-Cases of Slicing and Dicing the Open Web Index

More Ideas than time

- ❑ Build your own Pokemon search engine, or Harry Potter, etc.
- ❑ A climate change search engine (upcoming MANILA'24 workshop at SIGIR)
- ❑ A search engine to support financial experts, lawyers, etc.

# Evaluating Diverse Open Web Search Applications with TIREx

## A (biased) Selection of Use-Cases of Slicing and Dicing the Open Web Index

More Ideas than time

- ❑ Build your own Pokemon search engine, or Harry Potter, etc.
- ❑ A climate change search engine (upcoming MANILA'24 workshop at SIGIR)
- ❑ A search engine to support financial experts, lawyers, etc.

## Use Cases: The Web as Searchable Resource for AI

Web data drives innovation beyond search, particularly in AI

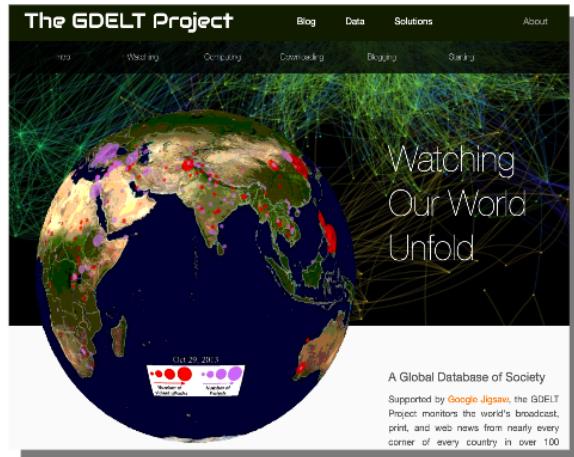


GPT-3: 80% of training data was  
Web data

# DALL-E 2

DALL-E 2 is an AI system that can create realistic images and art from a description in natural language.

Not only AI:  
GDELT: Web-scale Event Analytics



# Evaluating Diverse Open Web Search Applications with TIREx

## Summary

- ❑ The [OpenWebSearch.eu](#) project aims to provide an [Open Web Index](#)
- ❑ The Open Web Index aims to enable a diverse search ecosystem

# Evaluating Diverse Open Web Search Applications with TIREx

## Summary

- ❑ The OpenWebSearch.eu project aims to provide an Open Web Index
- ❑ The Open Web Index aims to enable a diverse search ecosystem

## Evaluation is Key

- ❑ Different organizations with different goals interact on the same fundamental

“ You don’t want that the client calls you at night.”

Arjen de Vries, OWS.eu

# Evaluating Diverse Open Web Search Applications with TIREx

## Summary

- ❑ The OpenWebSearch.eu project aims to provide an Open Web Index
- ❑ The Open Web Index aims to enable a diverse search ecosystem

## Evaluation is Key

- ❑ Different organizations with different goals interact on the same fundamental
  - “You don’t want that the client calls you at night.”  
Arjen de Vries, OWS.eu
- ❑ Good evaluation = good sleep :)

# Evaluating Diverse Open Web Search Applications with TIREx

## Summary

- ❑ The OpenWebSearch.eu project aims to provide an Open Web Index
- ❑ The Open Web Index aims to enable a diverse search ecosystem

## Evaluation is Key

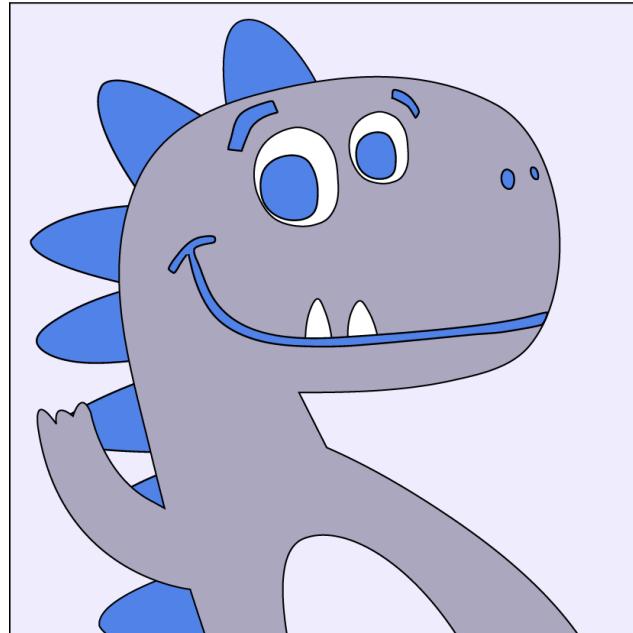
- ❑ Different organizations with different goals interact on the same fundamental
- “ You don’t want that the client calls you at night.”  
Arjen de Vries, OWS.eu
- ❑ Good evaluation = good sleep :)

## Strong Collaborations to Evaluation Campaigns

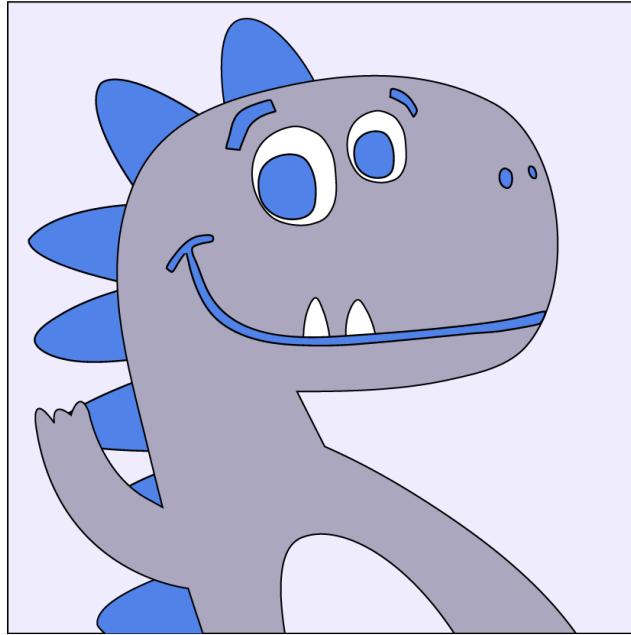
Example: Workshop on Open Web Search at ECIR 2024 (next week)

- ❑ We collaboratively develop and evaluate retrieval systems
- ❑ Collect standardized retrieval components
- ❑ Collect standardized evaluations (new test collection for Mastodon)
- ❑ Ensures diverse use cases are possible

# TIREx for Evaluation



## TIREx for Evaluation



TIREx = TIRA + ir\_datasets + PyTerrier

### TIRA

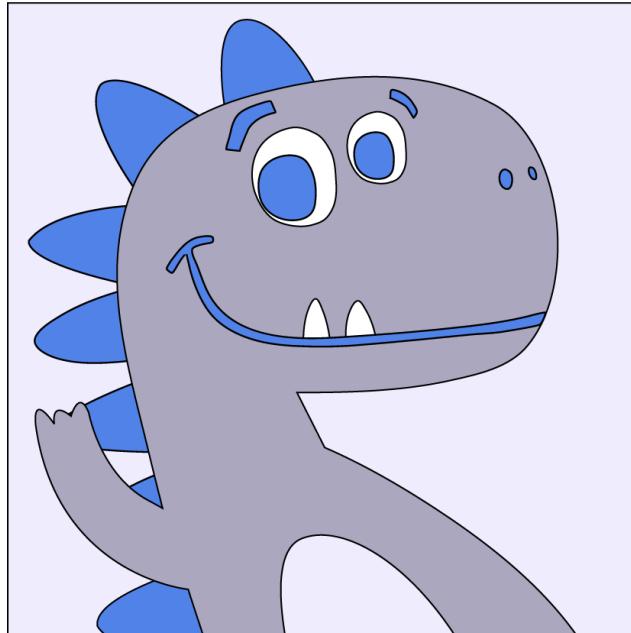
- ❑ Reproducible shared tasks: software submissions + blinded experiments

### ir\_datasets

- ❑ Unified + random data access: documents + queries + judgments

### PyTerrier

- ❑ Declarative pipelines



TIREx = TIRA + ir\_datasets + PyTerrier

TIRA

Reproducibility + Caching

- Reproducible shared tasks: software submissions + blinded experiments

Use Confidential Data

ir\_datasets

Including new Evaluations

- Unified + random data access: documents + queries + judgments

PyTerrier

- Declarative pipelines

# TIREx: Overview

- Each approach implemented in Docker image
- Executed in a sandbox  $\Rightarrow$  evitare che mancano dipendenze  
→ blinded experimentation + caching + reproducibility
- More than 50 transferrable retrieval models
- 35 test collections (18 public + 17 private test)

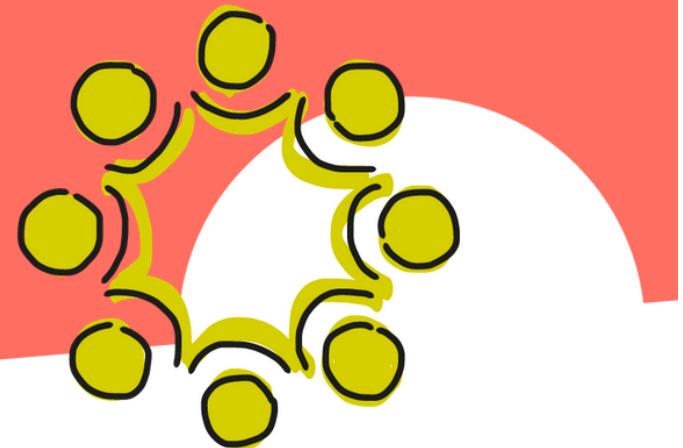
↓  
implementati con ir\_datasets  
possono essere trasferiti

# TIREx: Overview

- ❑ Each approach implemented in Docker image
- ❑ Executed in a sandbox
  - blinded experimentation + caching + reproducibility
- ❑ More than 50 transferrable retrieval models
- ❑ 35 test collections (18 public + 17 private test)

Next Week at ECIR

1st International  
Workshop on Open  
Web Search  
**#wows2024**



- ❑ We hope that the collected resources are useful to you

# WOWS 2024: Overview Retrieval Components

Type	Team	Retrieval Component	
		Description	#
Re-Ranking	h2oloo	Citadell (In progress)	1
	naverlabseurope	Splade	1
	tu-dresden-02	Snippet Extraction	2
Query	QPPTK	Query Performance Prediction	12
	qspell	Spelling Correction	4
	salamander	Comparative Query Classification	1
	tu-dresden-03	LLM Query Expansion	9
	OWS	Query Segmentation (baseline)	6
	marcel-gohsen	Eintity Linking / Query Interpretation	2
	dossier	Query Intent Prediction	2
Document	fschlatt	Health Classification	2
	seanmacavaney	Corpus Graph	1
	seanmacavaney	DocT5Query	1
	tu-dresden-01	Genre Classification	3
	tu-dresden-04	Readability/Quality/Coherence Features	2
Σ	12		42

# WOWS 2024: Overview Retrieval Components

Type	Team	Retrieval Component	#
		Description	#
Re-Ranking	h2oloo	Citadell (In progress)	1
	naverlabseurope	Splade	1
	<b>tu-dresden-02</b>	<b>Snippet Extraction</b>	<b>2</b>
Query	QPPTK	Query Performance Prediction	12
	qspell	Spelling Correction	4
	salamander	Comparative Query Classification	1
	<b>tu-dresden-03</b>	<b>LLM Query Expansion</b>	<b>9</b>
	OWS	Query Segmentation (baseline)	6
	marcel-gohsen	Eintity Linking / Query Interpretation	2
	dossier	Query Intent Prediction	2
Document	fschlatt	Health Classification	2
	seanmacavaney	Corpus Graph	1
	seanmacavaney	DocT5Query	1
	<b>tu-dresden-01</b>	<b>Genre Classification</b>	<b>3</b>
	<b>tu-dresden-04</b>	<b>Readability/Quality/Cohherence Features</b>	<b>2</b>
$\sum$	12		42

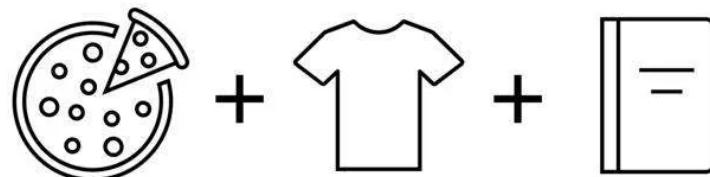
# WOWS 2024: Student Hackathon at TU Dresden

- 12 Students registered, covering Bachelor and Master students
- Students get credits + pizza + t-shirt
- Workflow: 1 week to read paper + 1 week hackathon + 0.5 weeks writing
- Big thanks to Anja Reusch, Julius Gonsior, and Wolfgang Lehner!



## IR Hackathon @ APB

- Team-Projekt (2 – 5 Leute)
- Eine kleine IR Komponente entwickeln
  - Dauer: 1 – 4 Tage je nach Lust und Zeit ☺



Pizza

T-Shirt

Anrechenbar  
im Studium



Kontakt: Anja Reusch &  
Julius Gonisor

# WOWS 2024: Student Hackathon at TU Dresden

Some OWSome impressions of the WOWS hackathon at Dresden:

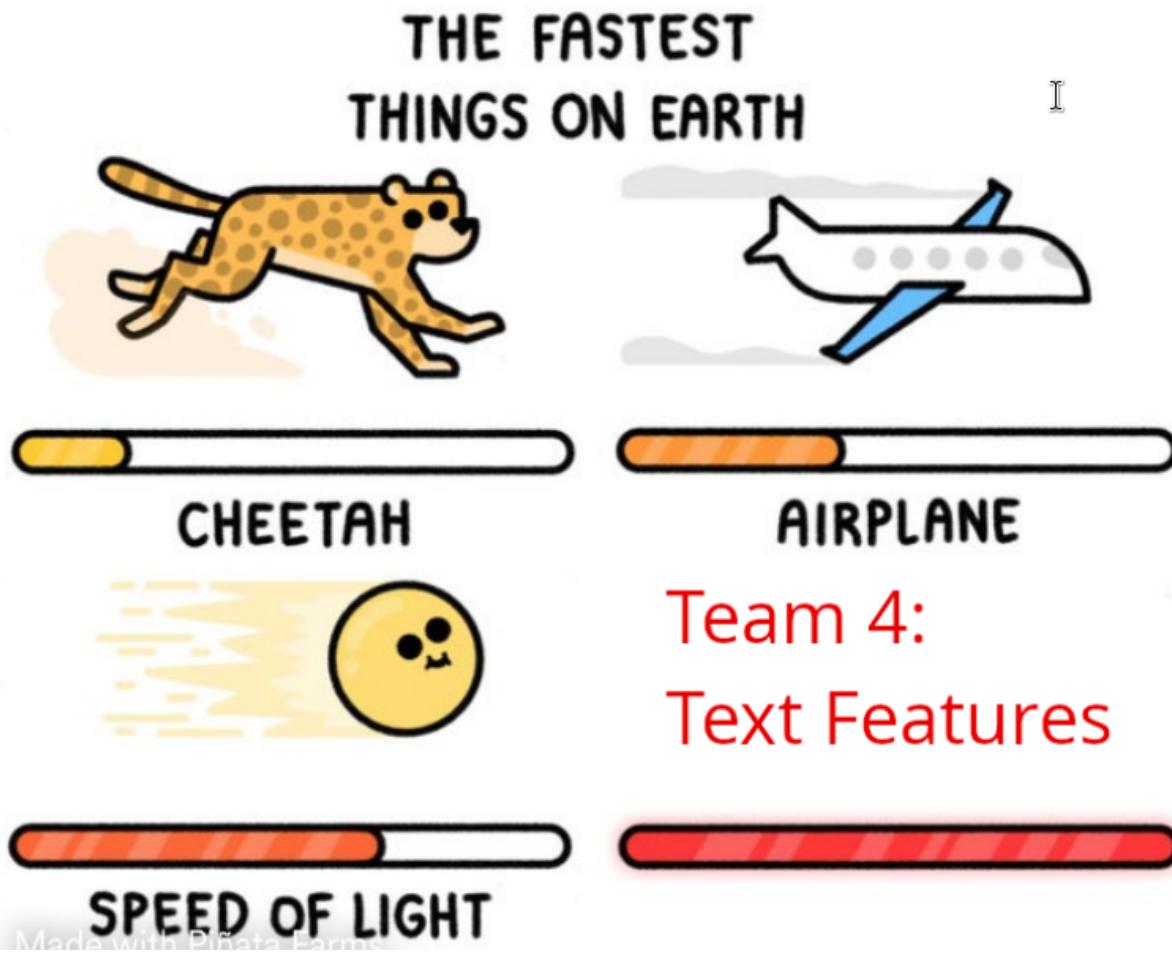


# WOWS 2024: Student Hackathon at TU Dresden

We had four teams, so we had four prizes...

# WOWS 2024: Student Hackathon at TU Dresden

We had four teams, so we had four prizes...



Team tu-dresden-04: First working submission on Day 1

# WOWS 2024: Student Hackathon at TU Dresden

We had four teams, so we had four prizes...



Team tu-dresden-03: 39,542 additions and 272 deletions

# WOWS 2024: Student Hackathon at TU Dresden

We had four teams, so we had four prizes...

Most commits :)



Team tu-dresden-02: 153 commits in one week

# WOWS 2024: Student Hackathon at TU Dresden

We had four teams, so we had four prizes...

## Most Working Submissions



Team tu-dresden-01: 9 working submissions

# WOWS 2024: Overview Components

Some OWSome impressions of the WOWS hackathon at Dresden:



# WOWS 2024: Overview Components

Some OWSome impressions of the WOWS hackathon at Dresden:



# WOWS 2024: Overview Components

## Some Example Components

Most fundamental: Loading some index :)



```
pt_dataset = pt.get_dataset('irds:antique/test')
```

[5]

```
index = tira.pt.index('ir-benchmarks/tira-ir-starter/Index (tira-ir-starter-pyterrier)', pt_dataset)
```

[6]

```
... Download from Zenodo: https://zenodo.org/records/10743990/files/2023-01-07-13-40-04.zip?download=1
Download: 100%[██████████] 31.1M/31.1M [00:02<00:00, 12.8MiB/s]
Download finished. Extract...
Extraction finished: /root/.tira/extracted\_runs/ir-benchmarks/antique-test-20230107-training/tira-ir-starter
```

# WOWS 2024: Overview Components

## Some Example Components

Most fundamental: Loading some index :)

```
▶ pt_dataset = pt.get_dataset('irds:antique/test')
[5]

[6] index = tira.pt.index('ir-benchmarks/tira-ir-starter/Index (tira-ir-starter-pyterrier)', pt_dataset)
... Download from Zenodo: https://zenodo.org/records/10743990/files/2023-01-07-13-40-04.zip?download=1
Download: 100%|██████████| 31.1M/31.1M [00:02<00:00, 12.8MiB/s]
Download finished. Extract...
Extraction finished: /root/.tira/extracted\_runs/ir-benchmarks/antique-test-20230107-training/tira-ir-starter
```

- Load a PyTerrier index created in TIRA from Zenodo
- The final “owi pull” might substantially differ, still:
  - Comparison to many alternative indexing pipelines important
  - Test collections rarely change, this index can be used for years
  - Open and transparent evaluations ⇒ informed decision

*permette valutazioni aperte e trasparenti*

# WOWS 2024: Overview Components

## Some Example Components: Genre Classification

```
pt_dataset = pt.get_dataset('irds:clueweb09/en/trec-web-2009')
topics = pt_dataset.get_topics('query')

bm25 = tira.pt.from_submission('ir-benchmarks/tira-ir-starter/BM25 Re-Rank (tira-ir-starter-pyterrier)', pt_dataset)
```

✓ 1.0s

```
genre_mlp_classifier = tira.pt.transform_documents('ir-benchmarks/tu-dresden-01/genre-mlp', pt_dataset)
```

✓ 4.8s

Download: 2.66MiB [00:00, 19.7MiB/s]

Download finished. Extract...

Extraction finished: [/root/.tira/extracted\\_runs/ir-benchmarks/clueweb09-en-trec-web-2009-20230107-training/tu-dresden-01](#)

```
(bm25 >> genre_mlp_classifier)(topics[topics['qid'] == '21']).head(5)
```

✓ 0.0s

	qid	query	rank	docno	predicted_label	probability_Shop	probability_Linklists	probability_Protrait non private
0	21	volvo	1	clueweb09-zh0015-47-11207	Shop	0.430557	0.050686	0.303563
1	21	volvo	2	clueweb09-en0035-03-39670	Shop	0.566059	0.035091	0.255652
2	21	volvo	3	clueweb09-zh0033-92-44184	Protrait non private	0.225574	0.050969	0.553595
3	21	volvo	4	clueweb09-ja0009-84-31373	Shop	0.525040	0.071335	0.173629
4	21	volvo	5	clueweb09-en0028-06-13844	Shop	0.429051	0.043311	0.389856

# WOWS 2024: Overview Components

## Some Example Components: Query Expansion with LLMs (Only the gist)

```
# Baselines
bm25 = pt.BatchRetrieve(index, wmodel="BM25")
bm25_rm3 = bm25 >> pt.rewrite.RM3(index) >> bm25
bm25_kl = bm25 >> pt.rewrite.KLQueryExpansion(index) >> bm25

# llm expansions with gpt
gpt_cot = tira.pt.transform_queries('workshop-on-open-web-search/tu-dresden-03/qe-gpt3.5-cot', dataset)
gpt_sq_fs = tira.pt.transform_queries('workshop-on-open-web-search/tu-dresden-03/qe-gpt3.5-sq-fs', dataset)
gpt_sq_zs = tira.pt.transform_queries('ir-benchmarks/tu-dresden-03/qe-gpt3.5-sq-zs', dataset)

# llm expansions with llama
llama_cot = tira.pt.transform_queries('ir-benchmarks/tu-dresden-03/qe-llama-cot', dataset)
llama_sq_fs = tira.pt.transform_queries('ir-benchmarks/tu-dresden-03/qe-llama-sq-fs', dataset)
llama_sq_zs = tira.pt.transform_queries('ir-benchmarks/tu-dresden-03/qe-llama-sq-zs', dataset)

# llm expansions with flan
flan_cot = tira.pt.transform_queries('ir-benchmarks/tu-dresden-03/qe-flan-ul2-cot', dataset)
flan_sq_fs = tira.pt.transform_queries('ir-benchmarks/tu-dresden-03/qe-flan-ul2-sq-fs', dataset)
flan_sq_zs = tira.pt.transform_queries('ir-benchmarks/tu-dresden-03/qe-flan-ul2-sq-zs', dataset)

pt.Experiment(
    [bm25, bm25_rm3, bm25_kl, pipeline_gpt_cot, pipeline_gpt_sq_fs, pipeline_gpt_sq_zs, pipeline_llama_cot, pip
    names=['BM25', 'BM25+RM3', 'BM25+KL', 'BM25+GPT-COT', 'BM25+GPT-SQ-FS', 'BM25+GPT-SQ-ZS', 'BM25+Llama-COT',
    topics=pt_dataset.get_topics('query'),
    qrels=pt_dataset.get_qrels(),
    eval_metrics=['recall_1000'],
    verbose=True
)
```

	name	recall_1000
0	BM25	0.751156
1	BM25+RM3	0.799385
2	BM25+KL	0.793911
3	BM25+GPT-COT	0.846802
4	BM25+GPT-SQ-FS	0.759494
5	BM25+GPT-SQ-ZS	0.770243
6	BM25+Llama-COT	0.810467
7	BM25+Llama-SQ-FS	0.761412
8	BM25+Llama-SQ-ZS	0.778425

# WOWS 2024: Overview Components

## Some Example Components: Query Expansion with LLMs (Only the gist)

```
# Baselines
bm25 = pt.BatchRetrieve(index, wmodel="BM25")
bm25_rm3 = bm25 >> pt.rewrite.RM3(index) >> bm25
bm25_kl = bm25 >> pt.rewrite.KLQueryExpansion(index) >> bm25

# llm expansions with gpt
gpt_cot = tira.pt.transform_queries('workshop-on-open-web-search/tu-dresden-03/qe-gpt3.5-cot', dataset)
gpt_sq_fs = tira.pt.transform_queries('workshop-on-open-web-search/tu-dresden-03/qe-gpt3.5-sq-fs', dataset)
gpt_sq_zs = tira.pt.transform_queries('ir-benchmarks/tu-dresden-03/qe-gpt3.5-sq-zs', dataset)

# llm expansions with llama
llama_cot = tira.pt.transform_queries('ir-benchmarks/tu-dresden-03/qe-llama-cot', dataset)
llama_sq_fs = tira.pt.transform_queries('ir-benchmarks/tu-dresden-03/qe-llama-sq-fs', dataset)
llama_sq_zs = tira.pt.transform_queries('ir-benchmarks/tu-dresden-03/qe-llama-sq-zs', dataset)

# llm expansions with flan
flan_cot = tira.pt.transform_queries('ir-benchmarks/tu-dresden-03/qe-flan-ul2-cot', dataset)
flan_sq_fs = tira.pt.transform_queries('ir-benchmarks/tu-dresden-03/qe-flan-ul2-sq-fs', dataset)
flan_sq_zs = tira.pt.transform_queries('ir-benchmarks/tu-dresden-03/qe-flan-ul2-sq-zs', dataset)

pt.Experiment(
    [bm25, bm25_rm3, bm25_kl, pipeline_gpt_cot, pipeline_gpt_sq_fs, pipeline_gpt_sq_zs, pipeline_llama_cot, pip
    names=['BM25', 'BM25+RM3', 'BM25+KL', 'BM25+GPT-COT', 'BM25+GPT-SQ-FS', 'BM25+GPT-SQ-ZS', 'BM25+Llama-COT',
    topics=pt_dataset.get_topics('query'),
    qrels=pt_dataset.get_qrels(),
    eval_metrics=['recall_1000'],
    verbose=True
)
```

	name	recall_1000
0	BM25	0.751156
1	BM25+RM3	0.799385
2	BM25+KL	0.793911
3	BM25+GPT-COT	0.846802
4	BM25+GPT-SQ-FS	0.759494
5	BM25+GPT-SQ-ZS	0.770243
6	BM25+Llama-COT	0.810467
7	BM25+Llama-SQ-FS	0.761412
8	BM25+Llama-SQ-ZS	0.778425

Note to self: use for-loops :)

# WOWS 2024: Overview Components

Some Example Components: DocT5Query =>

*doc e modello di trasformazione  
in input*

```
bm25 = tira.pt.from_submission('ir-benchmarks/tira-ir-starter/BM25 Re-Rank (tira-ir-starter-pyterrier)', dataset)
docT5Query = tira.pt.transform_documents('ir-benchmarks/seanmacavaney/DocT5Query', dataset)
[7] ✓ 29.2s
...
Download: 2.15MiB [00:00, 14.8MiB/s]
Download finished. Extract...
Extraction finished: /root/.tira/extracted_runs/ir-benchmarks/antique-test-20230107-training/tira-ir-starter
Download from the Incubator: https://files.webis.de/data-in-production/data-research/tira-zenodo-dump-preparation/doc-t5-query/2024-03-19-18-06-23.zip
This is only used for last spot checks before archival to Zenodo.
Download: 100%|██████████| 16.8M/16.8M [00:00<00:00, 47.5MiB/s]
Download finished. Extract...
Extraction finished: /root/.tira/extracted_runs/ir-benchmarks/antique-test-20230107-training/seanmacavaney
```

- Saves between a few hours and multiple days GPU compute
- Exploration of diverse pipelines now easy:
  - How to combine DocT5Query with other components?
  - On which corpora?
  - For which queries?
  - For which retrieval models?

*Riduce distanza utente - doc*

# Workshop on Open Web Search

## The Short Term Perspective: Teaching Initiatives

- Wilhelm von Humboldt: Good teaching is cutting edge research
- Super strong story for the OpenWebSearch.eu project:
  - Hackathon at TU Dresden in Germany
  - Two weeks later, potential re-use of components: IR Lab at the University of Padua for LongEval at CLEF 2024
- Accessibility (main challenge):  
<https://tira-io.github.io/teaching-ir-with-shared-tasks/>

# Workshop on Open Web Search

## The Short Term Perspective: Teaching Initiatives

- Wilhelm von Humboldt: Good teaching is cutting edge research
- Super strong story for the OpenWebSearch.eu project:
  - Hackathon at TU Dresden in Germany
  - Two weeks later, potential re-use of components: IR Lab at the University of Padua for LongEval at CLEF 2024
- Accessibility (main challenge): <https://tira-io.github.io/teaching-ir-with-shared-tasks/> *↗ tasks condive per insegnare IR*

## The Mid Term Perspective

For each new Open Web Search scenario, we re-run all components and combinations to make informed decisions how a suitable search engine would look for this scenario.

- The “client” has full transparency to re-use or re-implement components

# Workshop on Open Web Search

## Conclusions

- We collected 42 re-usable retrieval components
- We gained experience in collaboratively building retrieval systems
- Upcoming teaching initiatives with potential for creative exploration:
  - How to combine components?
  - What is missing, what is popular?
  - Fast research-oriented prototyping: minutes instead of hours/days
  - Connected to shared tasks => *e.g. LongEnd*

# Workshop on Open Web Search

## Conclusions

- ❑ We collected 42 re-usable retrieval components
- ❑ We gained experience in collaboratively building retrieval systems
- ❑ Upcoming teaching initiatives with potential for creative exploration:
  - How to combine components?
  - What is missing, what is popular?
  - Fast research-oriented prototyping: minutes instead of hours/days
  - Connected to shared tasks

## Future Work

- ❑ Integrate OpenWebSearch.eu scenarios: search for children is upcoming
- ❑ Teaching initiatives super much fun, makes sense to grow this
- ❑ Finalize archival of cached public outputs to Zenodo
- ❑ Monthly reproducibility checks with ReproduceMeGit
- ❑ Joint SIGIR Forum paper: experiences, concepts, and perspectives

# Fast Forward: Your IR Lab at LongEval@CLEF

Now we get our fingers dirty...



# Fast Forward: Your IR Lab at LongEval@CLEF

Now we get our fingers dirty...



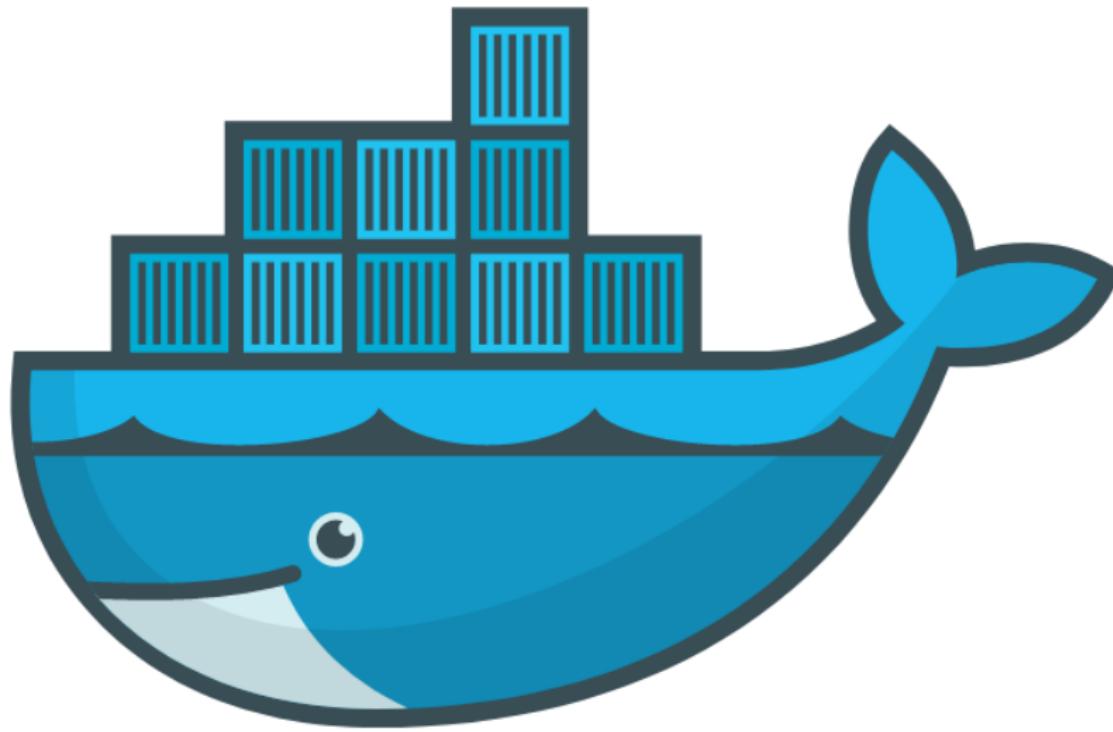
# IR Lab: Hands-on Session

## Aganda

- ❑ Docker
- ❑ Devcontainer
- ❑ ir\_datasets
- ❑ Re-using outputs of TIREx components

# Docker Tutorial

## Docker Basics



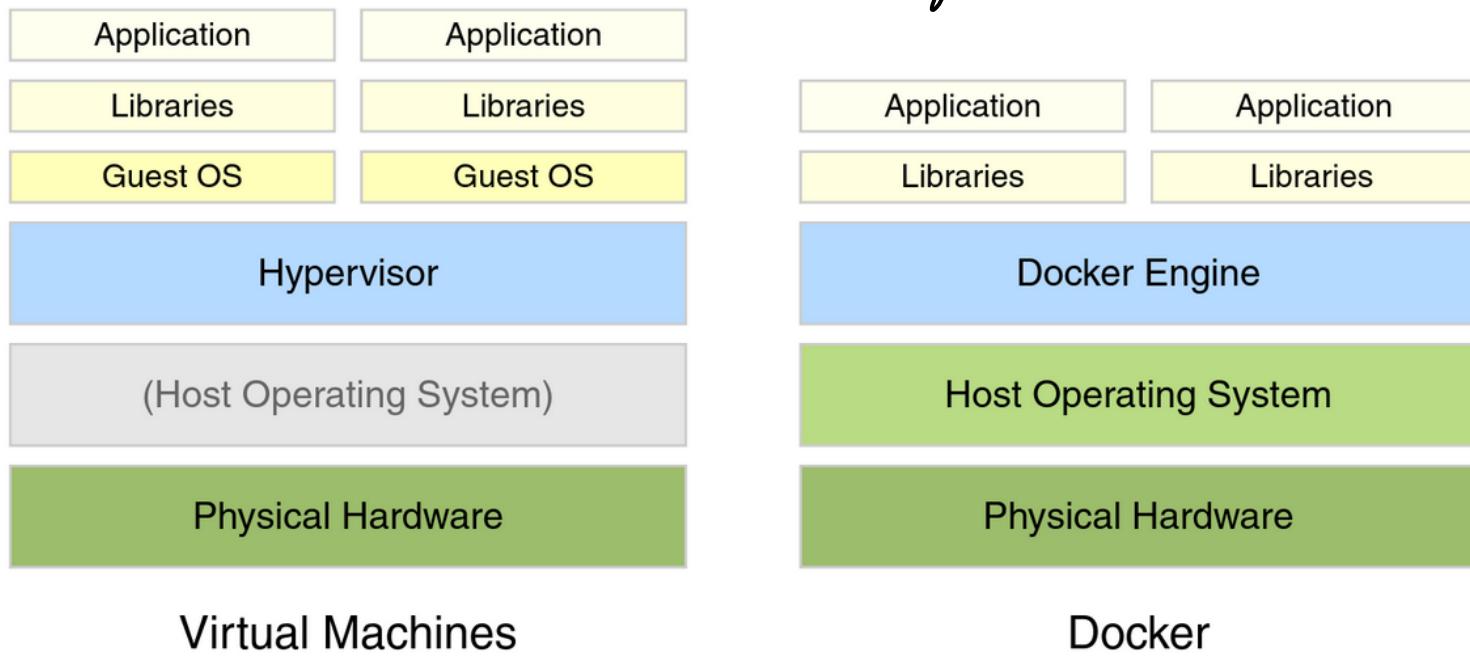
- ❑ Goal: Understand the basics behind dev containers
- ❑ <https://docs.docker.com/get-docker/>
- ❑ We will provide all required commands

# Docker Tutorial

## Use Cases for Docker

- ❑ Run guest systems as containers
- ❑ Shipping and running micro services as portable images
- ❑ Exploring and experimenting with new technologies
- ❑ Encapsulation mechanism to deploy applications in parallel without conflicts

Virtual Machines vs Docker  $\Rightarrow$  simile a VM + leggero



# Docker Tutorial

## Example Docker Commands

- Visit [hub.docker.com](https://hub.docker.com) => esempi di immagini
- We use the `bash` and `webis/ir-lab-wise-2023:0.0.4` images

  
shell

# Docker Tutorial

## Example Docker Commands

- Visit [hub.docker.com](https://hub.docker.com)
- We use the `bash` and `webis/ir-lab-wise-2023:0.0.4` images

### Bash Image

`docker run --rm -ti bash`  $\Rightarrow$  in terminal

- `--rm`: Remove container after completion *no standard I/O*
- `-ti`: Attach stdin and stdout  $\uparrow$  *finisce subito*
- ToDo: Run above command without `-ti`. What happens?
- ToDo: Write text to some file, restart the container. What happens?

*per vedere cosa sta andando: docker ps  
se confronto leah image e shell di sistema, cambiano  
programmi in esecuzione  
in uscita, perdo file creati con immagine*

# Docker Tutorial

## Example Docker Commands

- ❑ Visit [hub.docker.com](https://hub.docker.com)
- ❑ We use the `bash` and `webis/ir-lab-wise-2023:0.0.4` images

### Bash Image

```
docker run --rm -ti bash
```

- ❑ `--rm`: Remove container after completion
- ❑ `-ti`: Attach stdin and stdout
- ❑ ToDo: Run above command without `-ti`. What happens?
- ❑ ToDo: Write text to some file, restart the container. What happens?

### Bash Image With Volume Mounts

```
docker run --rm -ti -v $PWD:/bla bash
```

- ❑ `-v <HOST_PATH>:<CONTAINER_PATH>`: Mount the directory `<HOST_PATH>` on the system to the directory `<CONTAINER_PATH>` within the container
- ❑ ToDo: Write text to some file so that it is persistent.

# Docker Tutorial

Now We repeat this with Dev-Containers in VS Code

- Clone the repository

<https://bitbucket.org/frrncl/se-homework-template>

- Run both programs

- Show how to add new libraries

anche in repo di gruppo

avviso VScode: run in container



viene caricata immagine

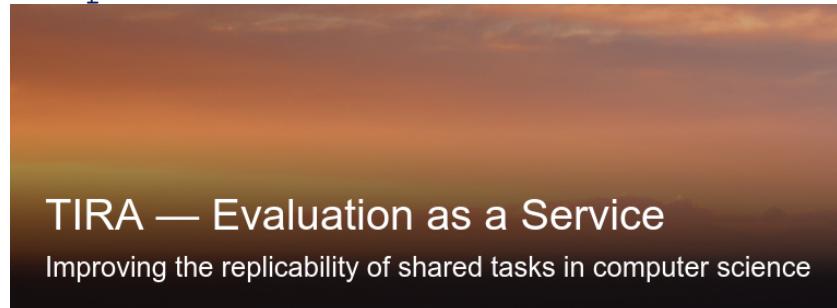
# Docker Tutorial

Some Suggestions on Using TIREx/WOWS components as additional inputs

- ❑ <https://github.com/tira-io/teaching-ir-with-shared-tasks/blob/main/tutorials/tutorial-data-access-from-java.ipynb>

# Submissions to TIRA

<http://tira.io>



## Step-by-Step Guide

- Step 1: Develop your System(s) on the training data *lqf, train.sh*
- Step 2: Get your system to run in a Docker image on your machine
- Step 4: Upload the working Docker image to TIRA
- Step 5: Execute your Approach in TIRA

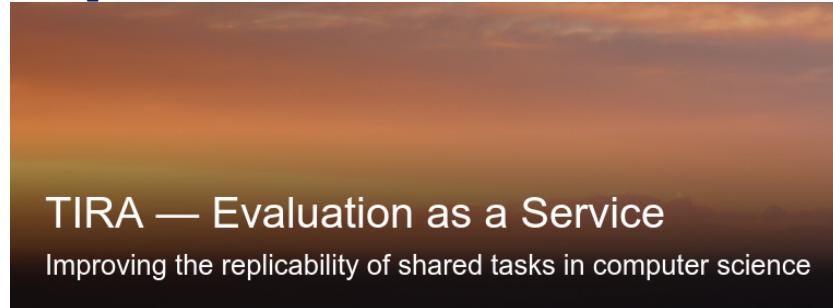
da DOCKERFILE dato, creare proprio -> va bene per TIRA  
in pagina per submit trovo istruzioni per creare immagini -> seguire quelle  
e le istruzioni in "cat DOCKERFILE"

PRIMA COSA: su hello-tipster, configurarlo per docker, provare a  
caricarlo in task

*docker build*  
*docker push*

# Submissions to TIRA

<http://tira.io>



## Step-by-Step Guide

- ❑ Step 1: Develop your System(s) on the training data
- ❑ Step 2: Get your system to run in a Docker image on your machine
- ❑ Step 4: Upload the working Docker image to TIRA
- ❑ Step 5: Execute your Approach in TIRA

## Typical workflow:

- ❑ Indexing: Documents  $\Rightarrow$  Index
  - ❑ Retrieval: Queries + Index  $\Rightarrow$  run
  - ❑ Re-Rankign: Query-Document pairs  $\Rightarrow$  run
- Tutorial: [github.com/tira-io/teaching-ir/tutorial-re-ranking.ipynb](https://github.com/tira-io/teaching-ir/tutorial-re-ranking.ipynb)

## Thats all, have fun!

- Please do not hesitate to contact me in case of problems or questions
- Next Steps: register to <https://www.tira.io/task-overview/ir-lab-padua-2024>
- I can provide some early feedback, maybe in two weeks?

# Thats all, have fun!

- Please do not hesitate to contact me in case of problems or questions
- Next Steps: register to <https://www.tira.io/task-overview/ir-lab-padua-2024>
- I can provide some early feedback, maybe in two weeks?



# Thats all, have fun!

- Please do not hesitate to contact me in case of problems or questions
- Next Steps: register to <https://www.tira.io/task-overview/ir-lab-padua-2024>
- I can provide some early feedback, maybe in two weeks?



**Thank you!**