

Unlocking Insights from Logbooks using AI at DESY and BESSY

A. Sulc, A. Eichler, G. Hartmann, T. Wilksen

Geneva, 15.02.2024

Rule 1



Rule 1

Don't talk about machine learning



Rule 1

Don't talk about machine learning (unless it is not helpful)!



Towards Data Science

<https://towardsdatascience.com> > ai-diagnoses-disease-... ⋮

AI Now Diagnoses Disease Better Than Your Doctor, Study ...

29 Sept 2020 — Creative **machine learning** is now **more** accurate in diagnosing disease than general practitioners, especially for rare diseases.



arXiv

<https://arxiv.org> > pdf PDF ⋮

ChipGPT: How far are we from natural language hardware ...

by K Chang · 2023 · Cited by 8 — **ChipGPT** serves as an EDA frontend framework, which aims to assist humans in compiling chip specifications to logic design. It takes chip ...



Google DeepMind

<https://deepmind.google> > discover > blog > funsearch... ⋮

FunSearch: Making new discoveries in mathematical ...

14 Dec 2023 — We introduce FunSearch, a method for **searching** for “functions” written in computer code, and **find** new solutions in **mathematics** and computer ...



Challenges & Improvements of Logbook

Challenges

Improvements



Challenges & Improvements of Logbook

Challenges

- > Different languages. (solved)

Improvements

Challenges & Improvements of Logbook

Challenges

- > Different languages. (solved)
- > Subjectivity, errors - normalization.
(in progress)

Improvements

Challenges & Improvements of Logbook

Challenges

- > Different languages. (solved)
- > Subjectivity, errors - normalization. (in progress)
- > Multimodality - often only screenshots are available. (in progress)

Improvements

23.01.2024 15:02 XFEL xfeloper XFEL_L3_StatusDisplay.xml: save actual readings



Challenges & Improvements of Logbook

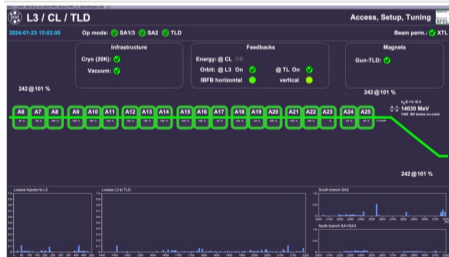
Challenges

- > Different languages. (solved)
- > Subjectivity, errors - normalization. (in progress)
- > Multimodality - often only screenshots are available. (in progress)

Improvements

- > FAIR: **F**indable, **A**ccessible, **I**nteroperable, **R**eusable. (solved)

23.01.2024 15:02 XFEL xfeloper XFEL_L3_StatusDisplay.xml: save actual readings



Challenges & Improvements of Logbook

Challenges

- > Different languages. (solved)
- > Subjectivity, errors - normalization. (in progress)
- > Multimodality - often only screenshots are available. (in progress)

Improvements

- > FAIR: **F**indable, **A**ccessible, **I**nteroperable, **R**eusable. (solved)
- > Factuality, references (solved)

23.01.2024 15:02 XFEL xfeloper XFEL_L3_StatusDisplay.xml: save actual readings



Challenges & Improvements of Logbook

Challenges

- > Different languages. (solved)
- > Subjectivity, errors - normalization. (in progress)
- > Multimodality - often only screenshots are available. (in progress)

Improvements

- > FAIR: **F**indable, **A**ccessible, **I**nteroperable, **R**eusable. (solved)
- > Factuality, references (solved)
- > Automation (DAQ). (in progress)

23.01.2024 15:02 XFEL xfeloper XFEL_L3_StatusDisplay.xml: save actual readings



Challenges & Improvements of Logbook

Challenges

- > Different languages. (solved)
- > Subjectivity, errors - normalization. (in progress)
- > Multimodality - often only screenshots are available. (in progress)

23.01.2024 15:02 XFEL xfeloper XFEL_L3_StatusDisplay.xml: save actual readings



Improvements

- > FAIR: **F**indable, **A**ccessible, **I**nteroperable, **R**eusable. (solved)
- > Factuality, references (solved)
- > Automation (DAQ). (in progress)
- > Root Cause Analysis, Problem-Solution Pipelines. (open)

Normalization and Multimodality of Logbook Entries (1)

23.01.2024 15:02 XFEL

xfeloper

XFEL_L3_StatusDisplay.xml: save actual readings



Normalization and Multimodality of Logbook Entries (1)

Status display of L3/CL/TLD from xfeloper on 23.01.2024 15:02 where all cavities are O.K, cryo is O.K, vacuum is O.K and ... was 14030 MeV

23.01.2024 15:02 XFEL xfeloper XFEL_L3_StatusDisplay.xml: save actual readings



Normalization and Multimodality of Logbook Entries (2)

23.01.2024 15:01 Software

Main linac parameters

shift documentation

Timing

Num. bunches total 242

Req. num. bunches type 1 2

Req. num. bunches type 2 1

Req. num. bunches type 3 0

Main rep rate 10 Hz

User-defined pattern ... yes

Pattern sequence [A]

Pattern description 2257 kHz, 100xD, 30x2, 100xD, 40x1, 633x3, 0x0

Photoinjector Laser

Laser pulses at gun 242

Attenuator position 591

Laser pulse energy (gun) 56.75 nJ

Spot X (near, SP/RBV) .. 690/681.525

Spot Y (near, SP/RBV) .. 564/566.40

Spot X (far, SP/RBV) ... 596/597.92

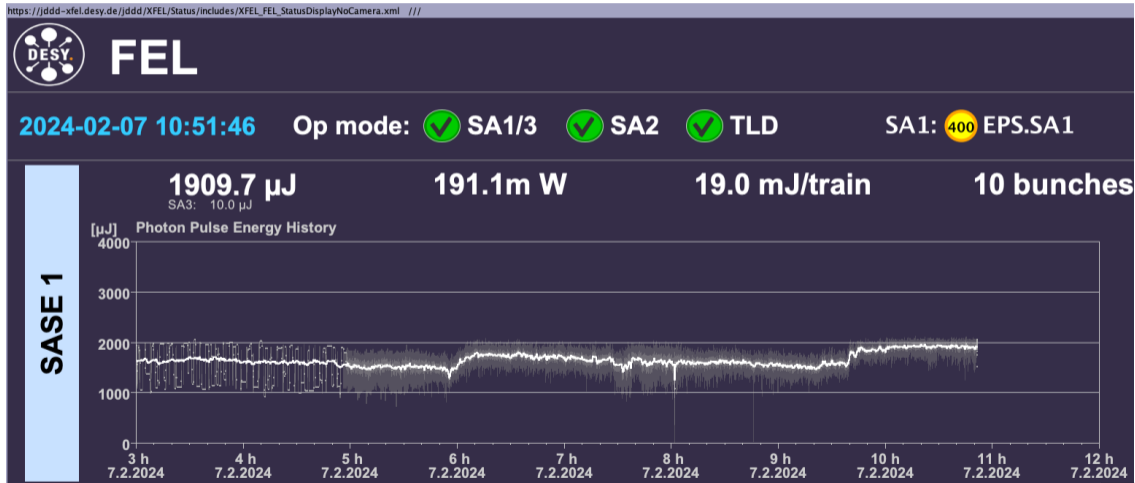
Spot Y (far, SP/RBV) ... 602/602.10

Normalization and Multimodality of Logbook Entries (2)

The shift on 23.01.2024 had following parameters: The total number of bunches was 242, main repetition rate was 10 Hz with user defined pattern...

23.01.2024 15:01 Software	Main linac parameters
shift documentation	
Timing	Photoinjector Laser
Num. bunches total 242	Laser pulses at gun ... 242
Req. num. bunches type 1 2	Attenuator position ... 591
Req. num. bunches type 2 1	Laser pulse energy (gun) 56.75 nJ
Req. num. bunches type 3 0	Spot X (near, SP/RBV) .. 690/681.525
Main rep rate 10 Hz	Spot Y (near, SP/RBV) .. 564/566.40
User-defined pattern ... yes	Spot X (far, SP/RBV) ... 596/597.92
Pattern sequence [A]	Spot Y (far, SP/RBV) ... 602/602.10
Pattern description 2257 kHz, 100x0, 30x2, 100x0, 40x1, 633x3, 0x0	

Multimodality - Why & How?



Multimodality - Why & How?

How: We have metadata! OCR (tesseract) + ML (text prediction)

https://jddd-xfel.desy.de/jddd/XFEL/Status/includes/XFEL_FEL_StatusDisplayNoCamera.xml //



FEL

2024-02-07 10:51:46

Op mode: ✔ SA1/3 ✔ SA2 ✔ TLD

SA1: 400 EPS.SA1

1909.7 μ J

SA3: 10.0 μ J

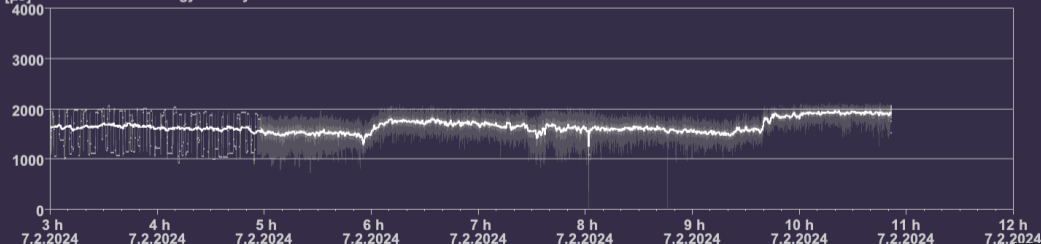
191.1m W

19.0 mJ/train

10 bunches

[μ J] Photon Pulse Energy History

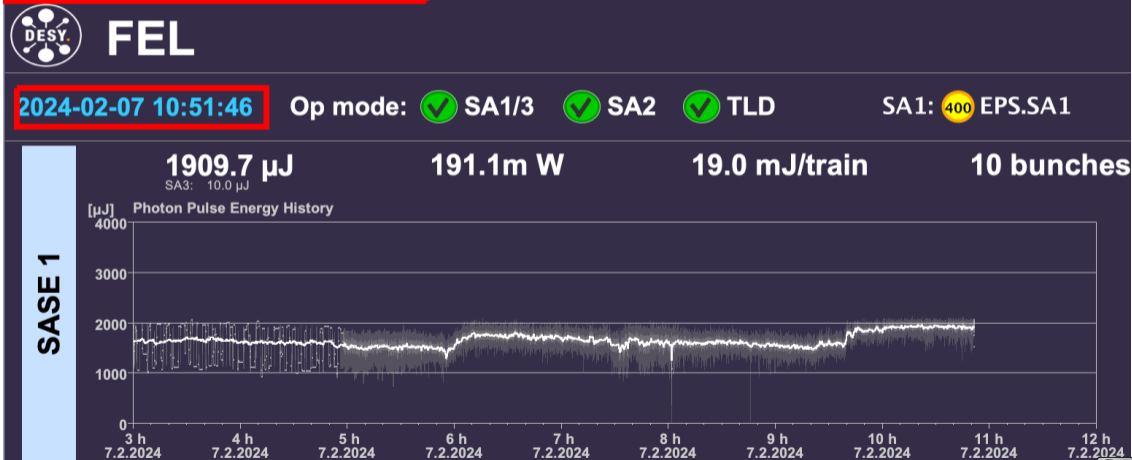
SASE 1



Multimodality - Why & How?

Why: A lot of data, like timestamp,

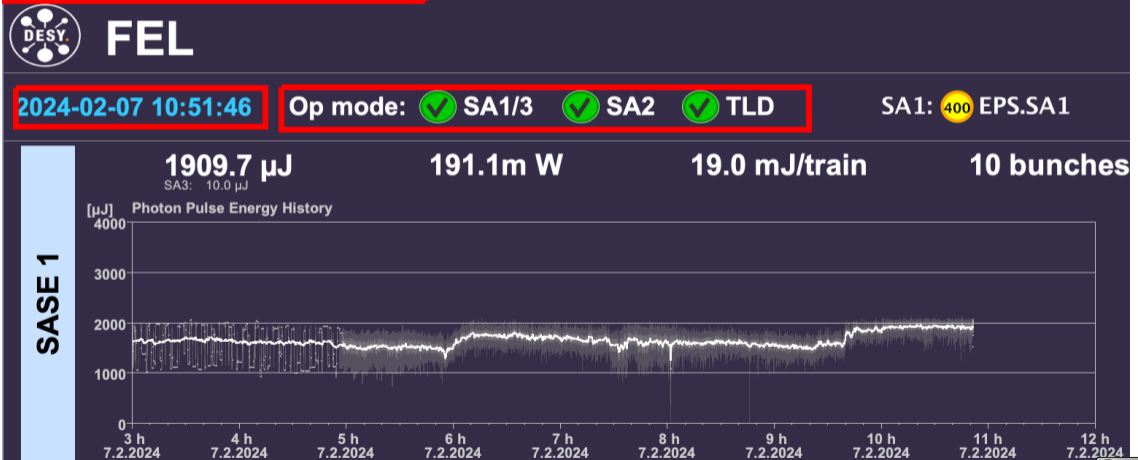
https://jddd-xfel.desy.de/jddd/XFEL/Status/includes/XFEL_FEL_StatusDisplayNoCamera.xml



Multimodality - Why & How?

Why: A lot of data, like timestamp, status,

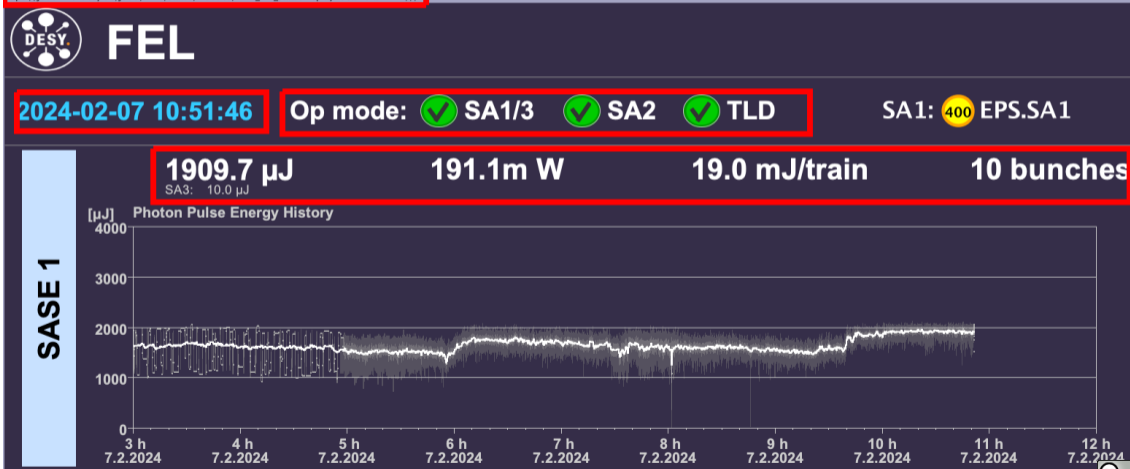
https://jddd-xfel.desy.de/jddd/XFEL/Status/includes/XFEL_FEL_StatusDisplayNoCamera.xml //



Multimodality - Why & How?

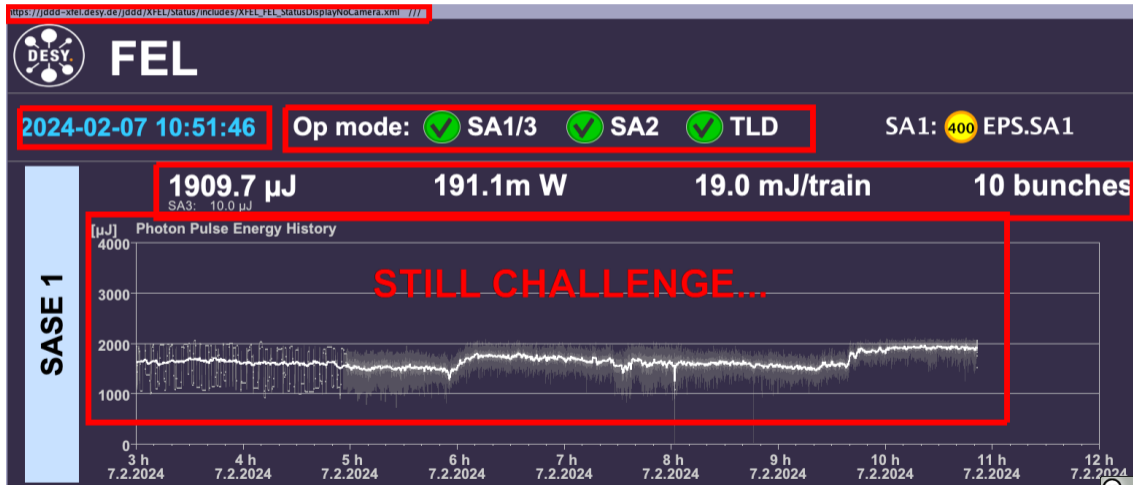
Why: A lot of data, like timestamp, status, values,

https://jddd-xfel.desy.de/jddd/XFEL/status/includes/XFEL_FEL_StatusDisplayNoCamera.xml

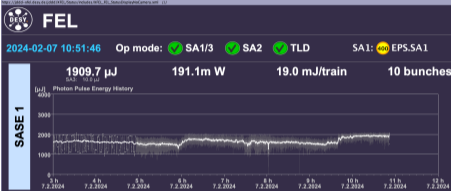


Multimodality - Why & How?

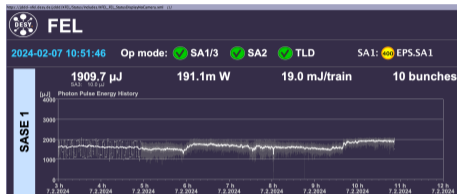
Why: A lot of data, like timestamp, status, values, histories, charts etc



Multimodality - Screenshots



Multimodality - Screenshots

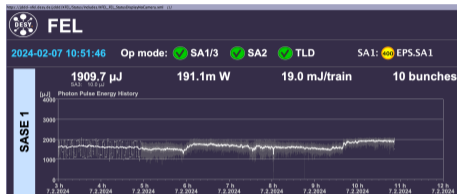


images

```
{
  date : "2024-02-07 10:51:46"
  Op mode-SA1/3 : OK
  :
  Photon Pulse Energy History : [1035.,...]
}
```

metadata (text, vectors)

Multimodality - Screenshots



```
{  
  date : "2024-02-07 10:51:46"  
  Op mode-SA1/3 : OK  
  :  
  Photon Pulse Energy History : [1035.,...]  
}
```

images

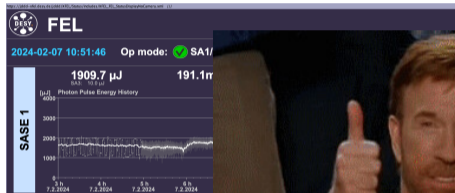


Machine Learning



metadata (text, vectors)

Multimodality - Screenshots



image

```
      : "2024-02-07 10:51:46"  
      : OK  
      :  
Photon Pulse Energy History : [1035.,...]
```

metadata (text, vectors)

ng



Why Normalization?

LLMs are good in answering questions and normalization enables **F**indability.



Why Normalization?

LLMs are good in answering questions and normalization enables **F**indability.

> Q: *When was the last day when number of bunches was above 200*

Why Normalization?

LLMs are good in answering questions and normalization enables **F**indability.

- > Q: *When was the last day when number of bunches was above 200*
- > A: The shift on 23.01.2024.

Why Normalization?

LLMs are good in answering questions and normalization enables **F**indability.

- > Q: *When was the last day when number of bunches was above 200*
- > A: The shift on 23.01.2024.

Where do you take the confidence that this is a correct answer?

Why Normalization?

LLMs are good in answering questions and normalization enables **F**indability.

- > Q: *When was the last day when number of bunches was above 200*
- > A: The shift on 23.01.2024.

Where do you take the confidence that this is a correct answer?

Factuality, RAGs!

Hallucinations & Expertise

Hallucinations:

Q: *How man SRF Cavities does European XFEL have?*

Hallucinations & Expertise

Hallucinations:

Q: *How man SRF Cavities does European XFEL have?*

A: *... 98 superconducting radiofrequency (SRF) cavities ...*

Or lack of expertise

Hallucinations & Expertise

Hallucinations:

Q: *How man SRF Cavities does European XFEL have?*

A: *... 98 superconducting radiofrequency (SRF) cavities ... wrong*

Or lack of expertise

Q: *I have a cavity, what should I do?*

Hallucinations & Expertise

Hallucinations:

Q: *How man SRF Cavities does European XFEL have?*

A: *... 98 superconducting radiofrequency (SRF) cavities ...* **wrong**

Or lack of expertise

Q: *I have a cavity, what should I do?*

A: *I'm not a dentist, but I can offer some general advice.* **good advice, but ...**

Hallucinations & Expertise

Hallucinations:

Q: *How man SRF Cavities does European XFEL have?*

A: *... 98 superconducting radiofrequency (SRF) cavities ... wrong*

Or lack of expertise

Q: *I have a cavity, what should I do?*

A: *I'm not a dentist, but I can offer some general advice. good advice, but ...*

Q: *What if it is from niobium?*

Hallucinations & Expertise

Hallucinations:

Q: *How man SRF Cavities does European XFEL have?*

A: *... 98 superconducting radiofrequency (SRF) cavities ... wrong*

Or lack of expertise

Q: *I have a cavity, what should I do?*

A: *I'm not a dentist, but I can offer some general advice. good advice, but ...*

Q: *What if it is from niobium?*

A: *It's highly unlikely that a cavity in your teeth is directly related to niobium.*

Hallucinations & Expertise

Hallucinations:

Q: *How man SRF Cavities does European XFEL have?*

A: *... 98 superconducting radiofrequency (SRF) cavities ...* wrong

Or lack of expertise

Q: *I have a cavity, what should I do?*

A: *I'm not a dentist, but I can offer some general advice.* good advice, but ...

Q: *What if it is from niobium?*

A: *It's highly unlikely that a cavity in your teeth is directly related to niobium.*

Good, but I had something else in my mind ...

Hallucinations & Expertise

Hallucinations:

Q: *How man SRF Cavities does European XFEL have?*

A: *... 98 superconducting radiofrequency (SRF) cavities ...* **wrong**

Or lack of expertise

Q: *I have a cavity, what should I do?*

A: *I'm not a dentist, but I can offer some general advice.* **good advice, but ...**

Q: *What if it is from niobium?*

A: *It's highly unlikely that a cavity in your teeth is directly related to niobium.*

Good, but I had something else in my mind ...



Hallucinations & Expertise

Hallucinations:

Q: *How man SRF Cavities does European XFEL have?*

A: *... 98 superconducting radiofrequency (SRF) cavities ...* wrong

Or lack of expertise

Q: *I have a cavity, what should I do?*

A: *I'm not a dentist, but I can offer some general advice.* good advice, but ...

Q: *What if it is from niobium?*

A: *It's highly unlikely that a cavity in your teeth is directly related to niobium.*

Good, but I had something else in my mind ...



Fine-tuning!

Expertise : Fine-tuning

- > Fine-tuned LLM for particle accelerators trained on technical texts like books, conferences, and preprints. **expertise.**
- > Automates data collection and question generation to create training data without human involvement. **up-to-date.**

PACuna: Automated Fine-Tuning of Language Models for Particle Accelerators

Antonin Sulc*
DESY,
Hamburg, Germany
antonin.sulc@desy.de

Raimund Kammering
DESY,
Hamburg, Germany

Annika Eichler
DESY,
Hamburg, Germany

Tim Wilksen
DESY,
Hamburg, Germany

Problem: Hard to keep updated with logbook

Factuality - Retrieval-Augmented Generation (RAG)

Question  Answer

LLM

Factuality - Retrieval-Augmented Generation (RAG)

Question

RAG

Factuality - Retrieval-Augmented Generation (RAG)

Question → Retriever



RAG

Context (Logbook)



Factuality - Retrieval-Augmented Generation (RAG)

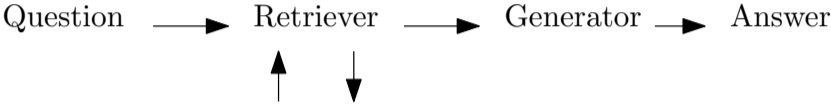
Question → Retriever → Generator



RAG

Context (Logbook)

Factuality - Retrieval-Augmented Generation (RAG)



RAG

Context (Logbook)



Factuality - Retrieval-Augmented Generation (RAG)

Q1

R

Towards LLMs as Operational Copilots for Fusion Reactors

Viraj Mehta¹, Joe Abbate², Allen M. Wang³, Andy Rothstein^{2,4}, Ian Char¹, Jeff Schneider¹, Egemen Kolemen^{2,4}, Cristina Rea³, and Darren T. Garnier³

¹Carnegie Mellon University, Pittsburgh, PA, USA

²Princeton Plasma Physics Laboratory, Princeton, NJ, USA

³MIT Plasma Science and Fusion Center, Cambridge, MA, USA

⁴Department of Mechanical and Aerospace Engineering, Princeton University, Princeton, NJ, USA

CONTEXT (LOGBOOK)

Rule 2

If you don't know how to solve it, wait for a few months,



Rule 2

If you don't know how to solve it, wait for a few months,

someone else will solve it for you

Toolformer: Language Models Can Teach Themselves to Use Tools

Timo Schick Jane Dwivedi-Yu Roberto Dessì[†] Roberta Raileanu
Maria Lomeli Luke Zettlemoyer Nicola Cancedda Thomas Scialom

Meta AI Research [†]Universitat Pompeu Fabra

Abstract

Language models (LMs) exhibit remarkable abilities to solve new tasks from just a few examples or textual instructions, especially at

The New England Journal of Medicine is a registered trademark of [QA("Who is the publisher of The New England Journal of Medicine?") → Massachusetts Medical Society] the MMS.

Rule 2

If you don't know how to solve it, wait for a few months,

someone else will solve it for you

Toolformer: Language Models Can Teach Themselves to Use Tools

awesome-hallucination-detection

 awesome License: Apache 2.0

Citing this repository

```
@misc{MinerviniAHD2014,  
  author = {Pasquale Minervini and others},  
  title = {awesome-hallucination-detection},  
  year = {2014},  
  publisher = {GitHub},  
  journal = {GitHub repository},  
  howpublished = {\url{https://github.com/EdinburghNLP/awesome-hallucination-detection}}  
}
```

Rule 2

If you don't know how to solve it, wait for a few months,

someone else will solve it for you

Toolformer: Language Models Can Teach Themselves to Use Tools

awesome-hallucination-detection

Textbooks Are All You Need

Suriya Gunasekar Yi Zhang Jyoti Aneja Caio César Teodoro Mendes

Allie Del Giorno Sivakanth Gopi Mojan Javaheripi Piero Kauffmann

Gustavo de Rosa Olli Saarikivi Adil Salim Shital Shah Harkirat Singh Behl

Xin Wang Sébastien Bubeck Ronen Eldan Adam Tauman Kalai Yin Tat Lee

Yuanzhi Li

Microsoft Research

Abstract

We introduce **phi-1**, a new large language model for code, with significantly smaller size than competing models: **phi-1** is a Transformer-based model with 1.3B parameters, trained for 4 days on 8 A100s, using a selection of “textbook quality” data from the web (6B tokens) and synthetically generated textbooks and exercises with GPT-3.5 (1B tokens). Despite this small scale, **phi-1** attains **pass@1** accuracy 50.6% on HumanEval and 55.5% on MBPP. It also displays surprising emergent properties compared to **phi-1-base**, our model *before* our finetuning stage on a dataset of coding exercises, and **phi-1-small**, a smaller model with 500M parameters trained with the same pipeline as **phi-1** that still achieves 45% on HumanEval.

Rule 2

If you don't know how to solve it, wait for a few months,

someone else will solve it for you

Toolformer: Language Models Can Teach Themselves to Use Tools

awesome-hallucination-detection

Textbooks Are All You Need

Article

Mathematical discoveries from program search with large language models

<https://doi.org/10.1038/s41586-023-06924-6> Bernardino Romera-Paredes^{1,4}✉, Mohammadamin Barekatin^{1,4}, Alexander Novikov^{1,4}, Matej Balog^{1,4}, M. Pawan Kumar^{1,4}, Emilien Dupont^{1,4}, Francisco J. R. Ruiz^{1,4}, Jordan S. Ellenberg², Pengming Wang¹, Omar Fawzi³, Pushmeet Kohli^{1,2} & Alhussein Fawzi^{1,4}✉

Received: 12 August 2023

Accepted: 30 November 2023

Published online: 14 December 2023

Open access

Check for updates

Large language models (LLMs) have demonstrated tremendous capabilities in solving complex tasks, from quantitative reasoning to understanding natural language. However, LLMs sometimes suffer from confabulations (or hallucinations), which can result in them making plausible but incorrect statements^{1,2}. This hinders the use of

searching in the function space), an evolutionary procedure based on pairing a pretrained LLM with a systematic evaluator. We demonstrate the effectiveness of

Rule 2

If you don't know how to solve it, wait for a few months,

someone else will solve it for you

Toolformer: Language Models Can Teach Themselves to Use Tools
awesome-hallucination-detection



Textbooks Are All You Need

Article
On the rheology of cats

IG Nobel Prize 2017 for Physics
”Can a Cat Be Both a Solid and a Liquid?”

In this letter I highlight some of the recent developments around the rheology of *Felis catus*, with potential applications for other species of the felidae family. In the linear rheology regime many factors can enter the determination of the characteristic time of cats: from surface effects to yield stress. In the nonlinear rheology regime flow instabilities can emerge. Nonetheless, the flow rate, which is the usual dimensional control parameter, can be hard to compute because cats are active rheological materials.

παντα ρει! Everything flows! This famous aphorism used to characterize Heraclitus' thought is also the motto of *rheology*. “Everything flows and nothing abides; everything gives way and nothing stays fixed.” A recipe for insubordination actually from Simplicius and Plato. Ev-

(a)  (b) 

Where are we?

- > I have been making really bold statements,



Where are we?

- > I have been making really bold statements,
- > We have already made some studies on our existing DESY logbooks.

Where are we?

- > I have been making really bold statements,
- > We have already made some studies on our existing DESY logbooks.
- > The results turned out to be **really bad** (bad language).



Where are we?

- > I have been making really bold statements,
- > We have already made some studies on our existing DESY logbooks.
- > The results turned out to be **really bad** (bad language).
- > That is why we talked about **fine-tuning** to enhance the logbook entries first.



Where are we?

- > I have been making really bold statements,
- > We have already made some studies on our existing DESY logbooks.
- > The results turned out to be **really bad** (bad language).
- > That is why we talked about **fine-tuning** to enhance the logbook entries first.
- > Then **factuality!**

Where are we?

- > I have been making really bold statements,
- > We have already made some studies on our existing DESY logbooks.
- > The results turned out to be **really bad** (bad language).
- > That is why we talked about **fine-tuning** to enhance the logbook entries first.
- > Then **factuality!**
- > Next step is **multi-modality** (images, databases, reports, charts) for **normalization** to improve **FAIR-ness**.



Where are we?

- > I have been making really bold statements,
- > We have already made some studies on our existing DESY logbooks.
- > The results turned out to be **really bad** (bad language).
- > That is why we talked about **fine-tuning** to enhance the logbook entries first.
- > Then **factuality!**
- > Next step is **multi-modality** (images, databases, reports, charts) for **normalization** to improve **FAIR-ness**.

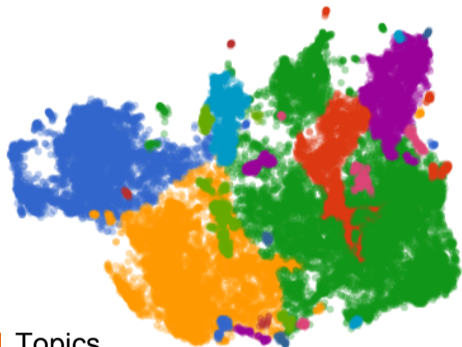
Start with metadata!



Previews - Trends, Topics, Interesting Findings

Dark Matter

Strong Coupling



- 1 Topics
- 2 Overlaps,
- 3 Connections in astro and other particles

Thank you!


sulcantonin.github.io

Contact

Deutsches Elektronen-
Synchrotron DESY

www.desy.de

A. Sulc, A. Eichler, G. Hartmann, T. Wilksen

 0000-0001-7767-778X

MCS

antonin.sulc@desy.de

