

Can we Trust AI?

Peter Murray-Rust

*Yusuf Hamied Dept of Chemistry,
University of Cambridge*

NIPGR, Delhi, IN, 2025-01-21



AMI



Openness and Responsible AI

Peter Murray-Rust

*Yusuf Hamied Dept of Chemistry,
University of Cambridge*

NIPGR, Delhi, IN, 2025-01-21



AMI, semanticClimate

Can we trust AI?

YES and NO



Polly,
stochasticParrot



Free and Open Source (F/OSS)



OpenKnowledge
and OpenData

semanticClimate

Builds a **community**

Automatically collects **climate information** ,
makes it **semantic, annotates it:**
for learning, re-use and advocacy

Sustainable Tech is created by Community



- Youth Learning by Doing; Internship Programme hosted at the National Institute of Plant Genome Research (NIPGR)



Why do you trust Alphafold?



AMI, semanticClimate



Because it's trained on real OPEN Data
And the Database is partnered by EBI

Q8W3K0 · RPP7_ARATH

Proteinⁱ | Protein RECOGNITION OF PERONOSPORA PARASITICA 7


Geneⁱ | RPP7

Statusⁱ |  UniProtKB reviewed (Swiss-Prot)

Organismⁱ | *Arabidopsis thaliana* (Mouse-ear cress)

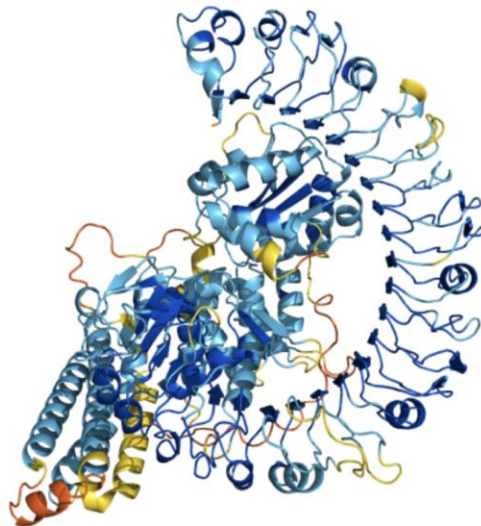
Amino acids | 1138 ([go to sequence](#))

Protein existenceⁱ | Evidence at transcript level

Annotation scoreⁱ | 

[Entry](#) [Variant viewer](#) [648](#) [Feature viewer](#) [Genomic coordinates](#) [Publications](#) [External links](#) [History](#)

 [Tools](#)  [Download](#)  [Add](#) [Add a publication](#) [Entry feedback](#)



Q8W3K0: A potential plant disease resistance protein. Mean pLDDT 82.24.

In [CASP14](#), AlphaFold was the top-ranked protein structure prediction method by a large margin, producing predictions with [high accuracy](#). While the system still has some [limitations](#), the CASP results suggest AlphaFold has immediate potential to help us understand the structure of proteins and advance biological research.

Let us know how the AlphaFold Protein Structure Database has been useful in your research, or if you have questions not answered in the [FAQs](#), at alphafold@deepmind.com.

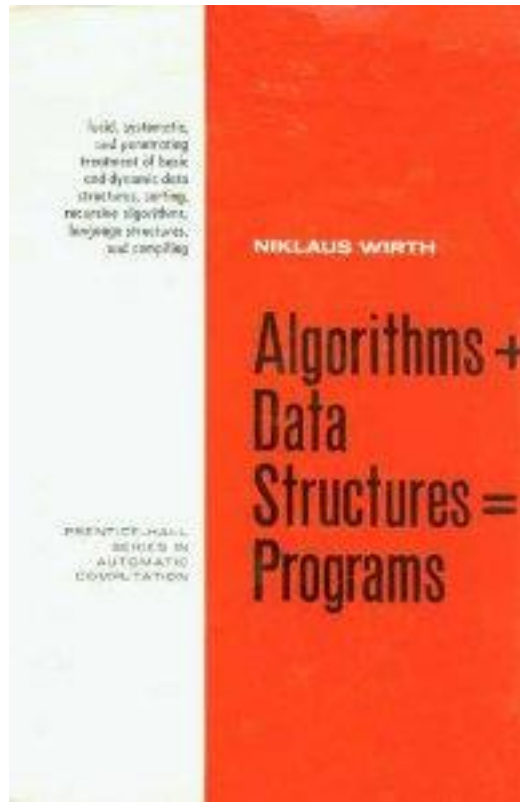
AlphaFold
Protein Structure Database

Developed by Google DeepMind and EMBL-EBI

European
Bioinformatics
Institute

<https://www.uniprot.org/uniprotkb/Q8W3K0/entry>

AI/ML components + licences



- Code (Github, PyPI, etc.)
- Data (e.g. *Bioinformatics*, Genbank, PDB, Uniprot) FAIR
- Text (unstructured, copyrighted) ??

Software freedoms

0. **to run the program** for any purpose
1. to study how the program works (**source code**)
2. **to redistribute copies** so you can help your neighbor.
3. To distribute copies of your **modified versions**

https://en.wikipedia.org/wiki/The_Free_Software_Definition



<https://blogs.ch.cam.ac.uk/pmr/2014/07/01/visit-of-richard-stallman-rms-to-cambridge/>

Software freedoms

These have inspired the
Open Data/Knowledge Movements

code /

2. you

3. ve

<https://>



copies so you can help

copies of your **modified**

re_Definition



<https://blogs.ch.cam.ac.uk/pmr/2014/07/01/visit-of-richard-stallman-rms-to-cambridge/>

The **Bermuda Principles** set out rules for the rapid and public release of [DNA sequence data](#).

The three principles retained originally were:

- Automatic release of sequence assemblies larger than 1 kb (preferably within 24 hours).
- Immediate publication of finished annotated sequences.
- Aim to make the entire sequence freely available in the public domain for both research and development in order to maximise benefits to society.

(https://en.wikipedia.org/wiki/Bermuda_Principles)

Jean-Claude Bradley

Jean-Claude Bradley



Jean-Claude at The White House - June 2013

Jean-Claude Bradley was one of the most influential open scientists of our time. He was an innovator in all that he did, from Open Education to bleeding edge Open Science; in 2006, he coined the phrase Open Notebook Science. His loss is felt deeply by friends and colleagues around the world.

On Monday July 14, 2014 we gathered at Cambridge University to honour his memory and the legacy he leaves behind with a highly distinguished set of invited speakers to revisit and build upon the ideas which inspired and defined his life's work.



Open Notebook Science

- ... there is a URL to a **laboratory notebook that is freely available and indexed on common search engines**. It does not necessarily have to look like a paper notebook but it is essential **that all of the information available to the researchers to make their conclusions is equally available to the rest of the world**
- —*Jean-Claude Bradley*

<http://www.budapestopenaccessinitiative.org/read>

... an **unprecedented public good**. ...

... **completely free and unrestricted access** to [peer-reviewed literature] by **all scientists, scholars, teachers, students, and other curious minds**. ...

...Removing access barriers to this literature will accelerate research, enrich education, share the **learning of the rich with the poor and the poor with the rich**, make this literature as useful as it can be, and **lay the foundation for uniting humanity in a common intellectual conversation and quest for knowledge**.

(Budapest Open Access Initiative, 2003)

Should we trust HuggingFace?





AMI, semanticClimate



I don't know where their data comes from

And they don't have an advisory board

the data are private,
run by a **private for-profit company**

	Hugging Face
Company type	Private
Industry	Artificial intelligence, machine learning, software development
Founded	2016; 9 years ago
Headquarters	Manhattan, New York City
Area served	Worldwide
Key people	Clément Delangue (CEO) Julien Chaumond (CTO) Thomas Wolf (CSO)
Products	Models, datasets, spaces
Revenue	▲ US\$15 million (2022)
Number of employees	170 (2023)
Website	huggingface.co 

“We are on a mission to democratize *good* machine learning, one commit at a time.”

Where does LLM data come from?

The AI model powering ChatGPT was trained using text databases from the internet and it is thought to have trained on around 300 billion words, or 570 GB, of data.

One proposed class-action suit claimed that OpenAI "secretly" stole "massive amounts of personal data," including medical records and information about children, to train ChatGPT. A group of authors are also suing the AI company, accusing them of ingesting their books to train the chatbot.

BUSINESS INSIDER

Google researchers say they got OpenAI's ChatGPT to reveal some of its training data with just one word

Beatrice Nolan Dec 4, 2023, 11:26 AM GMT

Share Save

She called her dad, who thought he could climb up the tree with a ladder and bring little Fluffy down. **The others sat next to her and fondled her.**

Hallucinations!

And then he got to the, they pulled his into the house on the window, **and he slapped out his infl Sexuality,**

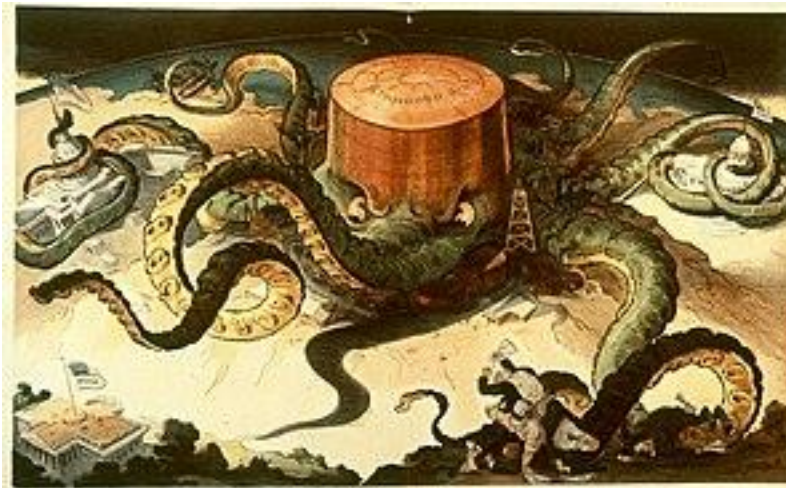
everybody in the truck, the whole family, just waving and yelling. My goodness. **That was pretty, extremely barbaric.**

And then sometimes I was scared about the traffic on the, you know, the cars, some, **some men are homeless, or they're autreally ill.**



Polly,
stochasticParrot

ROBBER BARONS



1904 depiction of an acquisitive and manipulative [Standard Oil](#) (founded by [John D. Rockefeller](#)) as an all-powerful octopus

Those practices included

- [unfettered consumption and destruction of natural resources](#),
- [influencing high levels of government](#),
- [wage slavery](#),
- [squashing competition by acquiring their competitors](#) to create [monopolies](#) and/or [trusts](#) that control [the market](#), and [s](#)
- [chemes to sell stock at inflated prices](#) to unsuspecting [investors](#).^[2]

Tiktok's enshittification

The company manually allocates surplus to creators, and they can take it away again, too.



Cory Doctorow  · [Follow](#)

17 min read · Jan 21, 2023

Here is how platforms die: first, they are good to their users; then they abuse their users to make things better for their business customers; finally, they abuse those business customers to claw back all the value for themselves. Then, they die. I call this enshittification, and it is a seemingly inevitable consequence arising from the combination of the ease of changing how a platform allocates value, combined with the nature of a "two-sided market", where a platform sits between buyers and sellers, hold each hostage to the other, raking off an ever-larger share of the value that passes between them.

<https://pluralistic.net/2023/01/21/potemkin-ai/#hey-guys>

Engage your brain!
Control AI or it will
control you!



AMI, semanticClimate

Thanks to:

* Yusuf Hamied Travel Fellowship (Cambridge India)

- NIPGR Hosting and internships
- Indian National Young Academy of Scientists
- The brave people who have challenged injustice