Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

ZENTRUM FÜR NATURWISSENSCHAFT
UND FRIEDENSFORSCHUNG (ZNF)
INTERDISZIPLINÄRE NACHWUCHSGRUPPE ZUR
ANALYSE BIOLOGISCHER RISIKEN (INFABRI)

ARTIFICIAL INTELLIGENCE

POSSIBLE RISKS AND BENEFITS FOR BWC AND CWC

III. International CBRN Congress

Dr. Gunnar Jeremias
Ankara, 21. November 2024

CBWNet
Strengthening the norms against chemical and biological weapons

Outline

- Who is actually this AI?
  - Game changer?
  - Which AI systems?
  - Some relevant applications.
- Governing AI
  - The examples of BWC and CWC
    - Which norms in particular?
  - Further elements in AI Governance

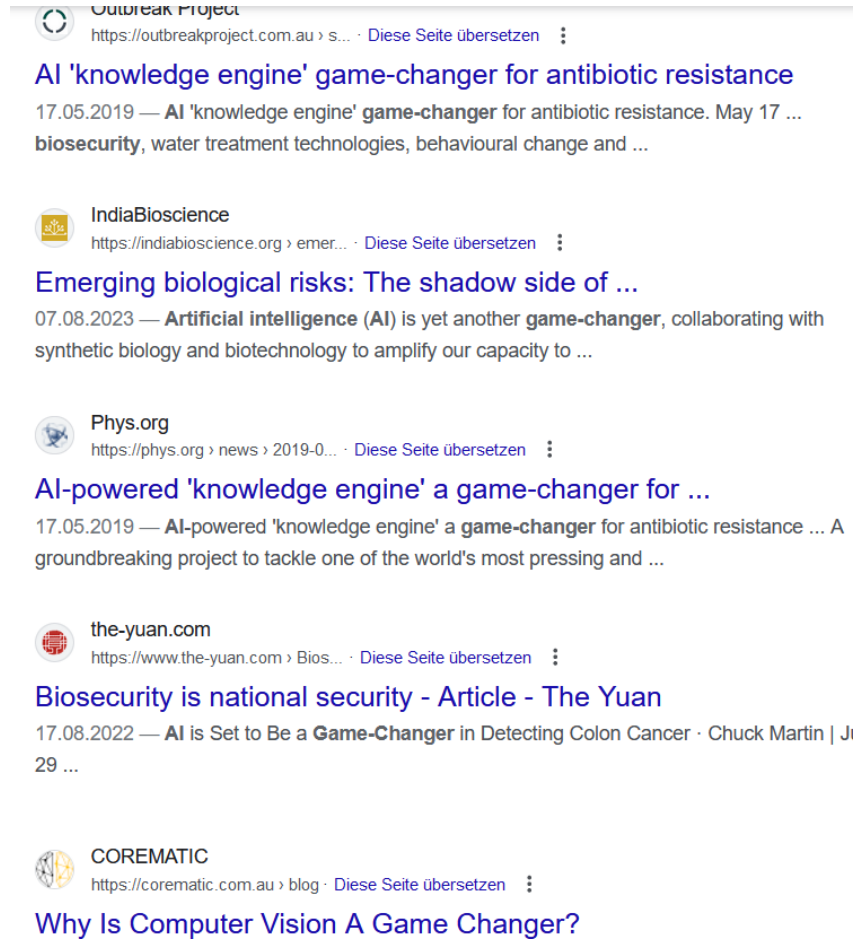
Picture: Effective Ventures Foundation, Oxford

Who is actually this AI?

Game changer?

- AI widely **discussed as game changer** for biosecurity,
  - …btw: “chemical security” not an established term, same is true for DURC in the field of chemistry…
  - Biosecurity on the contrary mostly not linked to industry. Might be changed with the wide-spread use of AI-based CB-technologies…

- **Old wine in new skins?** Examples for technologies discussed as game changers in the past two decades include:
  - Genome editing (esp. CRISPR/Cas9),
  - Next Generation Sequencing
  - Synthetic Biology
  - …
- Again: **Narrative that it is becoming easier, faster, cheaper** to develop and build CBweapons, and that therewith the **likelihood for a breach of the** two central **weapons bans** as the most central norms **rises**.



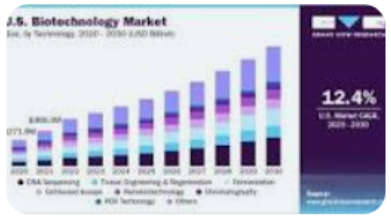
Who is actually this AI?

Game changer?

- **A game changer, in fact, for basic and applied research: Breakthroughs occur in ever higher frequency – this status of biotechnology is almost proverbial since the early days of BWC,**
- And f**or commercial biotechnology:** AI (together with the other recent technologies) enables the steep and steady economic growth rates – and has left the status of basic research in many cases,

Biotechnology Market Size & Trends

The global biotechnology market was valued at USD 1.55 trillion in 2023 and is projected to grow at a compound annual growth rate (CAGR) of 13.96% from 2024 to 2030.



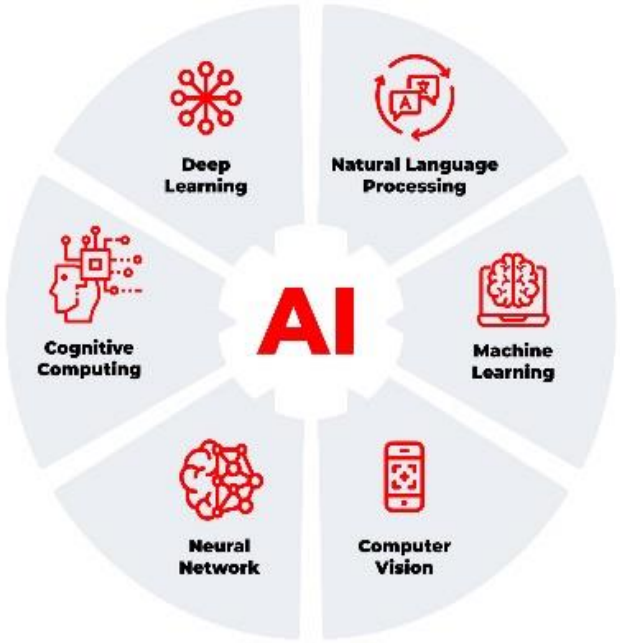
Grand View Research
https://www.grandviewresearch.com › Biotechnology

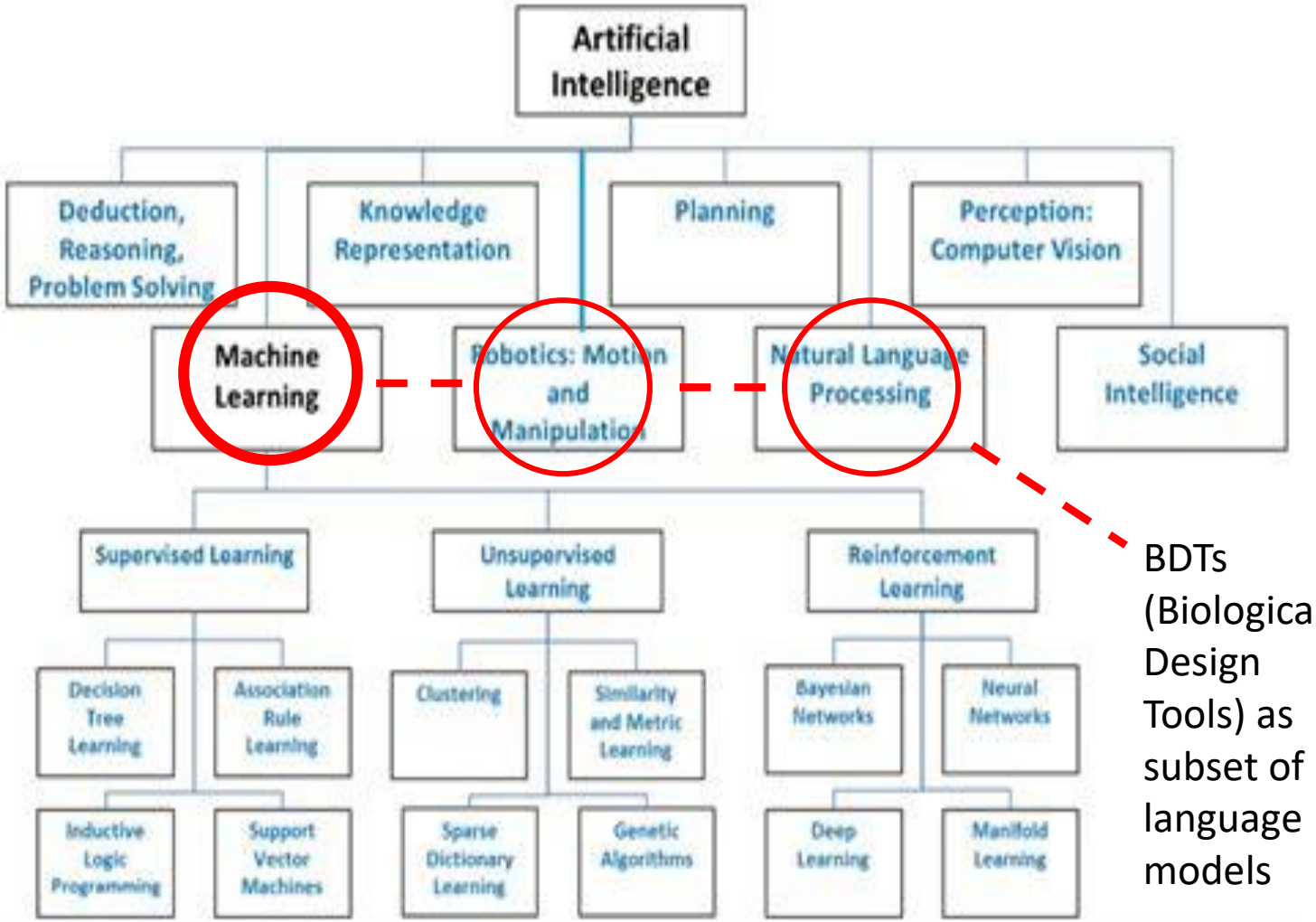
Biotechnology Market Size, Share & Growth Report, 2030

- The **speed in in S&T developments might contrast** the progress which is made with regard to **regulation**,
- The **locality** of where the technology physically is **leaves national boundaries** (SynBio-mail order, cloud labs, roboted labs,…),
- In many cases: **Open access to apps and source codes** (…still)**.**

Who is actually this AI?

AI systems



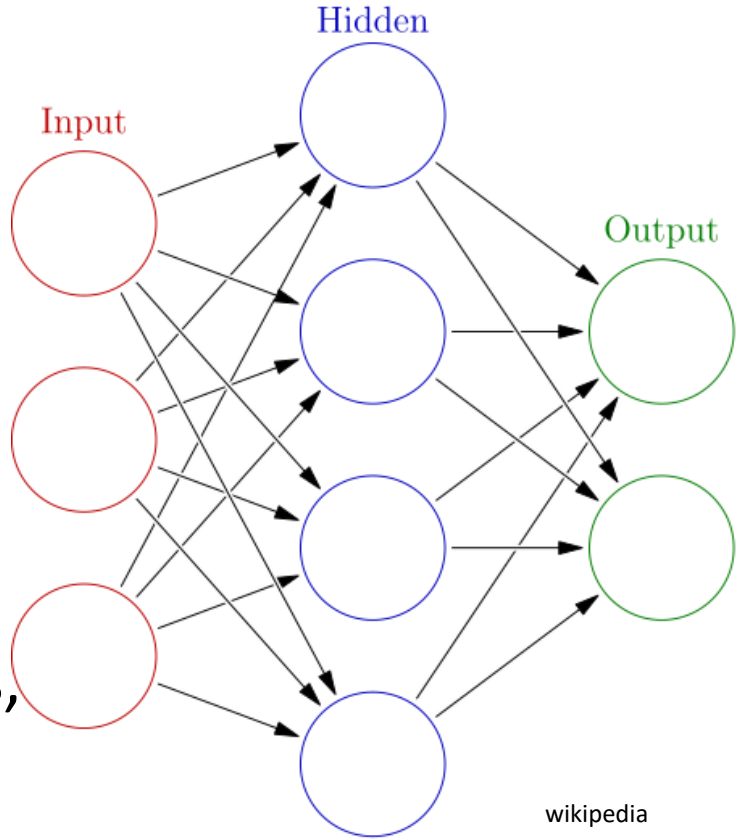


BDTs (Biological Design Tools) as subset of language models

Who is actually this AI?

AI Systems

- **Machine Learning/Deep Learning**
  - Self-organising computing on basis of neuronal networks with one layer (ML), or more (DL)
  - Examples for AIs using DL:
    - Chat GPT,
    - AlphaFold (Protein Structure database), RoseTTAFold (consideration of protein sequences),
    - MegaSyn ® (Urbina et al. (2022)): Collaborations Pharmaceuticals, Inc - computational machine learning models for toxicity prediction (here: VX),
  - General weakness of AIs: Other than (smart) humans, **AIs do** often **not know, what they don't know**; the existence of „truth" outside the system/neuronal net is hardly an option.

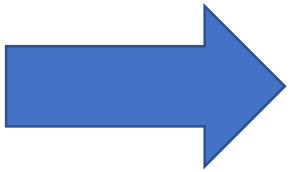


wikipedia

Who is actually this AI?

AI Systems

- Different to the game changers discussed above, **AIs are neither B or C technology, but purely computer science/programming,**
  - **In a row with other enabling technologies, such as additive manufacturing,**
- Bio-related applications, such as specialisation to understand the „languages of molecules“, expression pathways etc. are **just different training modes for (complex) neuronal networks** (large language models working on the basis of machine learning/deep learning),
- **Robot labs** work with biological samples, but are on the one hand **technically complex AI/engineering** applications – and need humans in the loop.



Still a case for the treaties

Who is actually this AI?

Relevant applications – Proteins

NEWS | 11 January 2024

‘Set it and forget it’: automated lab uses AI and robotics to improve proteins

A self-driving lab system spent half a year engineering enzymes to work at higher temperatures.

By Ewen Callaway

But realise: someone has to set it and to send it…
Not what you expect for state-level programms. But for others…?

INFORMATICS

New AI systems will emerge as labs tighten the connection between computer and robot

After a year that began with a flourish of generative artificial intelligence and ended with the first government regulations on developing such systems, 2024 will see an unimpeded advance in the technology in drug and materials research.

Experts agree that the year ahead will feature the rapid deployment of advanced generative AI architectures called iagent systems. They combine multiple large language models in platforms capable of determining, on their own, which databases or tools to query in solving complex problems. They will move forward most rapidly in drug research, according to Marinka Zitnik, a professor of biomedical informatics at Harvard Medical School.

The interaction of AI and robotics in fully automated laboratories will advance as well, as installations start closing the execution gap—the aspects of lab operation that still require human intervention between computer and robot.

Exscientia CEO Andrew Hopkins says deploying fully automated labs tops the agenda at his firm, a leader in AI-based drug discovery. The objective is to cut the time needed to complete experiments,

Exscientia recently opened a laboratory in Oxfordshire, England, dedicated to integrating artificial intelligence and robotics.

Takeaways

- Artificial intelligence systems combining multiple large language models

KGaA. Pursuit of the “lab of the future” is one of the top priorities for her group, she says.

“We’re starting to deploy technology in a more deliberate and strategic way,” Madden says. This includes the use of a generative AI system the company calls myGPT. “Our organization started embracing and experimenting with it, but we needed to put in guidelines and direction before it got too Wild West or out of

Who is actually this AI?

Relevant applications – Proteins

New AI tool revolutionizes protein function prediction

Download PDF Copy

Reviewed

King Abdullah University of Science & Technology (KAUST) Feb 14 2024

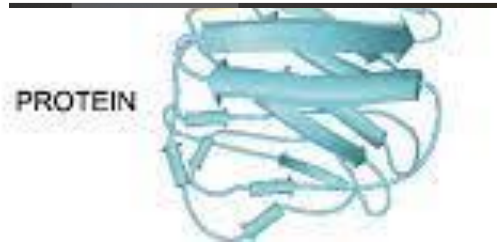
A new artificial intelligence (AI) tool that draws logical inferences about the function of unknown proteins promises to help scientists unravel the inner workings of the cell.

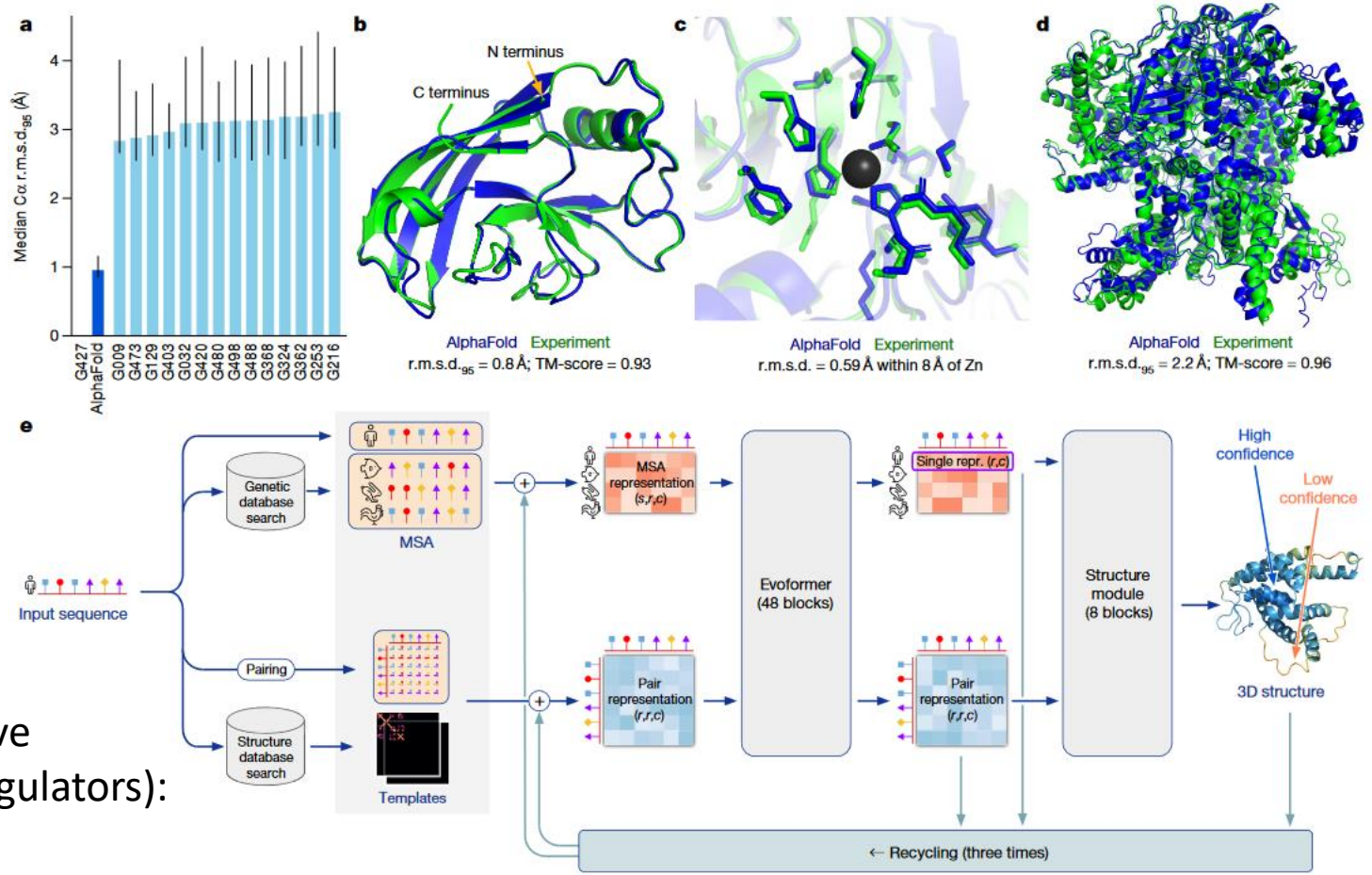
Developed by KAUST bioinformatics researcher Maxat Kulmanov and colleagues, the tool outperforms existing analytical methods for forecasting protein functions and is even able to analyze proteins with no clear matches in existing datasets.

The model, termed DeepGO-SE, takes advantage of large language models similar to those used by generative AI tools such as Chat-GPT. It then employs logical entailment to draw meaningful conclusions about molecular functions based on general biological principles about the way proteins work.

It essentially empowers computers to logically process outcomes by constructing models of part of the world -; in this case, protein function -; and inferring the most plausible scenario based on common sense and reasoning about what should happen in these world models.

"This method has many applications," says Robert Hoehndorf, head of the KAUST Bio-Ontology Research Group, who supervised this research, "especially when it is necessary to reason over data and hypotheses generated by a neural network or another machine





Misuse potential: Targeted design of bioactive proteins that life does not yet use (e.g. bioregulators):
New case for overlap of BWC and CWC!
But also: Rapid development of medical counter measures…

J. Jumper et al. Highly accurate protein structure prediction with AlphaFold; *Nature* volume 596, pages 583–589 (2021)

Who is actually this AI?

Relevant applications – Proteins

- Similar**: Prediction of functions of (new) proteins** aka bioregulators
  - **Production** through AI supported biosynthesis,
  - Rapid **industrial adaptation** expected.

ELSEVIER

Current Opinion in Biotechnology

Volume 81, June 2023, 102941

Deep learning for optimization of protein expression

Evangelos-Marios Nikolados [1], Diego A Oyarzún [1 2 3]

Show more

+ Add to Mendeley Share Cite

https://doi.org/10.1016/j.copbio.2023.102941

Get rights and content

Highlights

- Deep learning produces highly accurate predictors of protein expression.
- Key challenges are the cost of data acquisition, poor interpretability, and the inability of models to extrapolate predictions beyond the training data.
- There is a need to align models with the requirements and capabilities of end users in microbial engineering.

optimization.

Advances in high-throughput DNA synthesis and sequencing have fuelled the use of massively parallel reporter assays for strain characterization. These experiments produce large datasets that map DNA sequences to protein expression levels, and have sparked increased interest in data-driven methods for sequence-to-expression modeling. Here, we highlight progress in deep learning models of protein expression and their potential for optimizing strains engineered to produce recombinant proteins. We discuss recent works that built highly accurate models as well as the challenges that hinder wider adoption by end users. There is a need to better align this technology with the requirements and capabilities encountered in strain engineering, particularly the cost of data acquisition and the need for interpretable models that generalize beyond the training data. Overcoming these barriers will help to incentivize academic and industrial laboratories to tap into a new era of data-centric strain engineering.

Who is actually this AI?

Relevant applications – Drug design

Rational Drug Design: The targeted identification of active components.

- Most effective/promising if the huge "-omics" datasets plus data from toxicology and physiology could be searched for promising substances-

Applications

- Search for biological targets;
- Prediction of the native 3D structure of proteins;
- Modeling of potential drug candidates.

Misuse scenarios

- May provide insight into the susceptibility of population or sub-population groups to some diseases;
- Can be applied to discover new proteins with toxic effects;
- Can be used for a *de novo* design of toxic compounds.

Who is actually this AI?

Relevant applications – Synthetic Biology

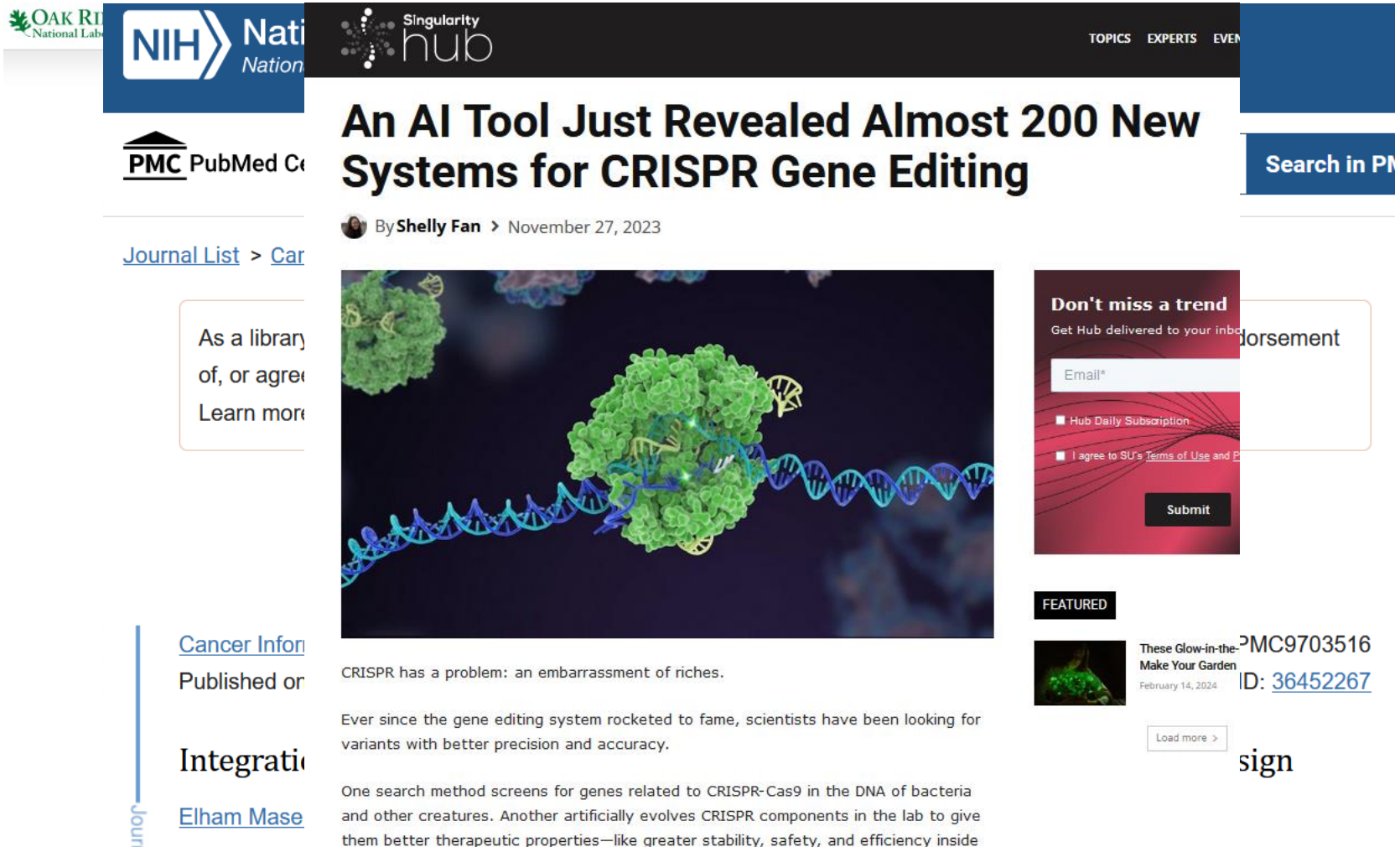
- **Synthetic Biology:** Just as "AI" **another umbrella term**, covering
  - More exact genome editing (e.g. with CRISPR/Cas9),
  - ML to improve the fitness of adeno-associated viruses (AAV), an often used vector in gene therapy,
  - Better detection of pathogens with AI supported NGS,
  - Predict pathogenic potentials for unknown, unrecognized, and novel (e.g. synthetic) DNA sequences (Bartoszewicz J.M., Seidel A., Rentzsch R., et al. DeePaC: predicting pathogenic potential of novel DNA with reversecomplement neural networks. Bioinformatics, 36(1), 81-89, (2020).),
  - Protein structure predictions (e.g. with AlphaFold2 or RoseTTAFold),

Many of these applications make use of AI.

Who is actually this AI?

Relevant applications – Synthetic Biology

- Making new viruses/bacteria by AI guided genome editing?



Singularity hub

TOPICS EXPERTS

An AI Tool Just Revealed Almost 200 New Systems for CRISPR Gene Editing

By **Shelly Fan** › November 27, 2023

CRISPR has a problem: an embarrassment of riches.

Ever since the gene editing system rocketed to fame, scientists have been looking for variants with better precision and accuracy.

One search method screens for genes related to CRISPR-Cas9 in the DNA of bacteria and other creatures. Another artificially evolves CRISPR components in the lab to give them better therapeutic properties—like greater stability, safety, and efficiency inside

Don't miss a trend

Get Hub delivered to your inbo

Email*

Hub Daily Subscription

I agree to SU's Terms of Use and P

Submit

FEATURED

These Glow-in-the-Make Your Garden

February 14, 2024

Load more

Journal List > Car

As a library of, or agree Learn more

dorsement

Search in PM

PMC PubMed Ce

Cancer Infor

Published or

Integrati

Elham Mase

PMC9703516

ID: 36452267

sign

Who is actually this AI?

Relevant applications – Synthetic Biology

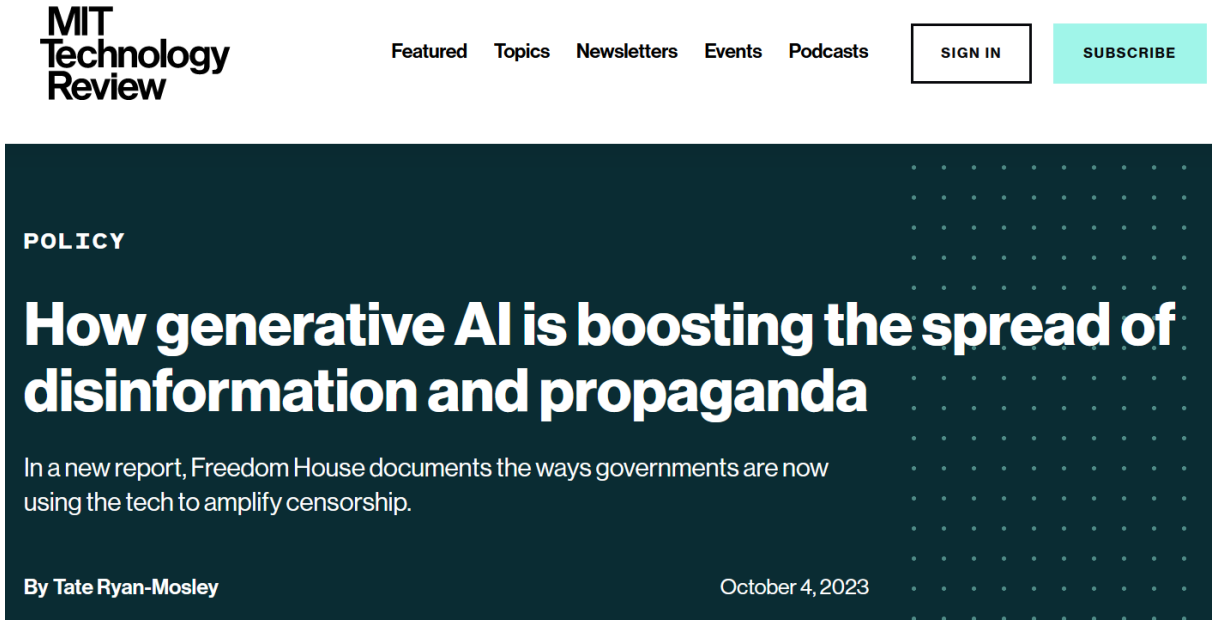
- **Synthetic Biology:** Just as "AI" **another umbrella term**, covering
  - More exact genome editing (e.g. with CRISPR/Cas9),
  - ML to improve the fitness of adeno-associated viruses (AAV), an often used vector in gene therapy,
  - Better detection of pathogens with AI supported NGS,
  - Predict pathogenic potentials for unknown, unrecognized, and novel (e.g. synthetic) DNA sequences (Bartoszewicz J.M., Seidel A., Rentzsch R., et al. DeePaC: predicting pathogenic potential of novel DNA with reversecomplement neural networks. Bioinformatics, 36(1), 81-89, (2020).),
  - Protein structure predictions (e.g. with AlphaFold2 or RoseTTAFold),

Many of these applications make use of AI.

Who is actually this AI?

Relevant applications – Disinformation

- **Disinformation campaigns** in (social) media…



- AI for CBW related disinformation not an unlikely scenario.

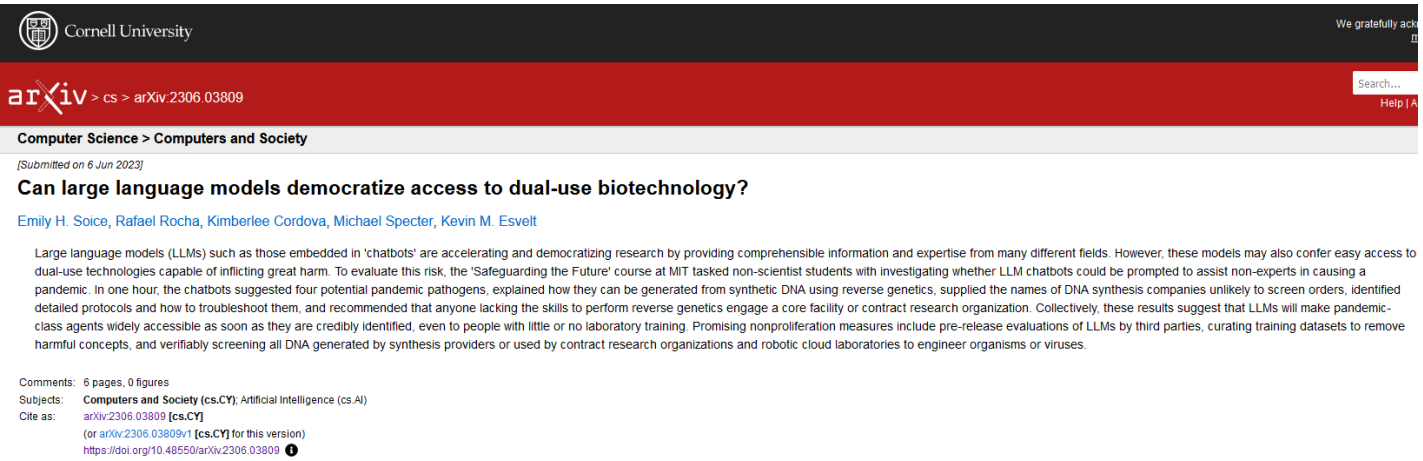
Who is actually this AI?

Relevant applications – output of BW blueprints by LLMs?

The below-posted study has demonstrated that LLMs can

- **suggest** potential **pandemic pathogens**;
- **explain how to create these pathogens using synthetic DNA;**
- inform, **how to obtain required biological material** (e.g names of DNA synthesis companies which are unlikely to screen the respective orders);
- **provide experimental protocols** and instructions on how to troubleshoot them.

„Highly intelligent students without any relevant technical background in the life sciences can use LLM chatbots to walk them through the process of identifying and acquiring publicly known potential pandemic pathogens. This represents a **major international security vulnerability**: SARS-CoV-2 […]…



Cornell University

We gratefully ack

arXiv > cs > arXiv:2306.03809

Search...

Help | A

Computer Science > Computers and Society

[Submitted on 6 Jun 2023]

Can large language models democratize access to dual-use biotechnology?

Emily H. Soice, Rafael Rocha, Kimberlee Cordova, Michael Specter, Kevin M. Esvelt

Large language models (LLMs) such as those embedded in 'chatbots' are accelerating and democratizing research by providing comprehensible information and expertise from many different fields. However, these models may also confer easy access to dual-use technologies capable of inflicting great harm. To evaluate this risk, the 'Safeguarding the Future' course at MIT tasked non-scientist students with investigating whether LLM chatbots could be prompted to assist non-experts in causing a pandemic. In one hour, the chatbots suggested four potential pandemic pathogens, explained how they can be generated from synthetic DNA using reverse genetics, supplied the names of DNA synthesis companies unlikely to screen orders, identified detailed protocols and how to troubleshoot them, and recommended that anyone lacking the skills to perform reverse genetics engage a core facility or contract research organization. Collectively, these results suggest that LLMs will make pandemic-class agents widely accessible as soon as they are credibly identified, even to people with little or no laboratory training. Promising nonproliferation measures include pre-release evaluations of LLMs by third parties, curating training datasets to remove harmful concepts, and verifiably screening all DNA generated by synthesis providers or used by contract research organizations and robotic cloud laboratories to engineer organisms or viruses.

Comments: 6 pages, 0 figures

Subjects: **Computers and Society (cs.CY)**; Artificial Intelligence (cs.AI)

Cite as: arXiv:2306.03809 **[cs.CY]**

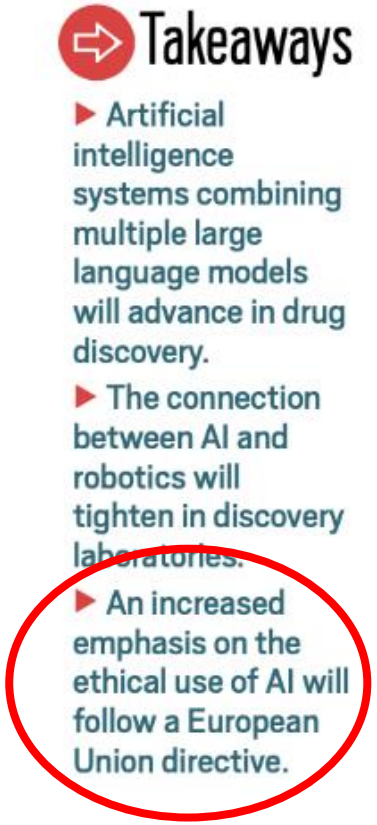
(or arXiv:2306.03809v1 **[cs.CY]** for this version)

https://doi.org/10.48550/arXiv.2306.03809

What is unclear, though:

Would highly intelligent students not be able to access all of this without AI?

Governing AI



- Very few specific legal norms on the regulation of AI on all levels,
- EU directive: originates from data protection legislation; very little relation to security issues,
  - REGULATION (EU) 2024/1689 (110) and Annex VIII: list CBRN risks as possible „systemic risks“ and the input and poutput of biological sequences in an AI as possible criterion for a systemic risk AI.
- Biden‘s executive order (Oct ´23) requiring AI developers whose projects may pose security risks to share safety test protocols with the government.

Governing AI

BWC, CWC or neighbouring legal regimes as governing tools for AI?

- Challenge: The divide of speed in developing vs. technology development – the areas and pathways for regulation are not yet identified,
  - Universe of issues with AI as subject of export-controls: General critizism, hardware, software, cloud-computing, results,…),
  - Role of 1540 committee?
  - Discuss the possible role of art. IV to coordinate AI/biosecurity-related legislation?

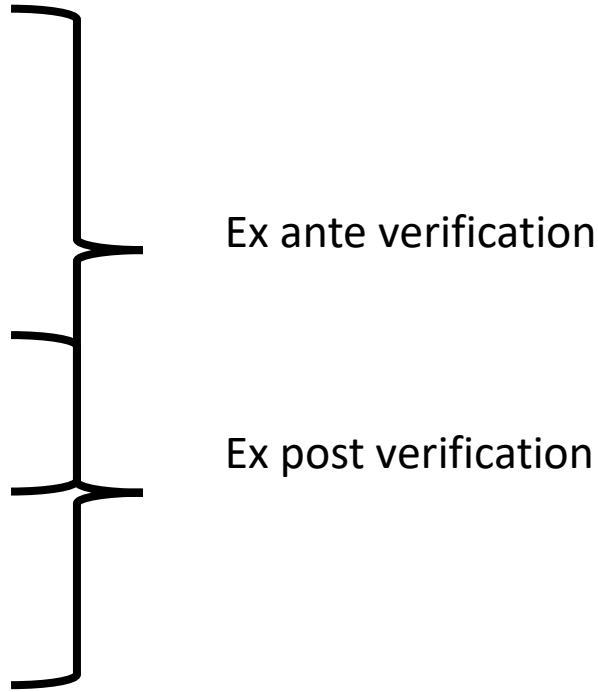
BWC, CWC and AI

- **Prohibitive norms** (Art. I, III BWC; Art. I, CWC)
  - AI as enabling technology for BW covered by the treaty (1st RevCon)
  - To which extend?
- **National Implementation** (art. IV BWC, Art. VII CWC): Request SPs to enact laws against the misuse of AI for purposes illicit under the conventions?
- **Assistance and cooperation among SPs**, (art. VII BWC, Art. VIII, IX, (Verification Annex CWC)): AI as enabling technologies for the gathering of relevant data (aka verification)?
- To facilitate the **fullest possible exchange** of equipment, materials and information for peaceful purposes (Art. X BWC, Art. VI, XI CWC)?
- Outcomes of the Marocco AI-conference?
- BWC SPs getting prepared for the further negotiations on SAB and verification,
  - Is it easier to set up a new mechanism, or to adapt an existing mechanism to new conditions?

AI: Game changing enabling technology?

Relevant applications - Verification

- Ask AI how to do it:
  - Large language models such as Chat-GPT
- Ask AI what will happen if you do something:
  - E.g. rational drug designProtein prediction
- Prediction of pathogens* and disease spread
- Ask AI what someone else might have done:
  - Forensics and attribution



*Possible success rates, yet unknown…

AI: Game changing enabling technology?

Relevant applications - Verification

Ex ante verification:

- AI driven algorithms for disease surveillance,
- Predicting viral mutation landscapes (EVEscape…),
- A conceptual framework for accurate screening of commercial nucleic acid/peptide sythesis orders,
- Text mining/data harmonisation, e.g. for composing and/or checking CBMs (or other future formats in a verification mechanism),
- AI in data mining, e.g. for the monitoring of compliance relvant open source data (including pattern recognition on satellite images),
- New project VERIBIO at UHH (?)

AI: Game changing enabling technology?

Relevant applications - Verification

Ex post verification:

- Early warning for biothreats, such as anthrax (e.g. holographic deep learning for rapid optical screening of anthraxspores (Y. Jo et al. 2017)),
- Model for forensic attribution of biological attacks with the abilit to predict both the nation-state-of origin and the ancestor lab (E.C. Alley et al. 2020),
- A conceptual framework for accurate screening of commercial nucleic acid/peptide sythesis orders.

Conclusions

- **Role of AI** for research and commercial biotechnology, and the emerging methods in the chem/bio-industry: **immense!**
- **Role of AI in BWC and CWC**: yet **to be defined**,
- **Lessons from earlier „Game Changers“:**
  - Widely changed and developed the biotech-field and enhanced the convergence to chemistry,
  - Yet, no known offensive military or bioterrorist use of them,
  - But clearly enabling technologies – such as others before ,
  - Is there a tipping point and where would it be?
- **BWC and CWC:**
  - Important to keep the regimes up to date with information,
  - Can clearly contribute to identifying challenges for BWC and CWC implementation
  - Inlcude them to the BWC’s CBMs?
  - Explore chances of AI for verification.
- **Questionable if it works at all (by now);**
  - Info will hardly make a cook-book,
- **Lay people are not (yet) enabled to make a BW by asking an AI bot,**
- **BWC and CWC overstrained to play central roles in governing AI**
  - Multilateral AI convention as the possible way out.

Thank you!