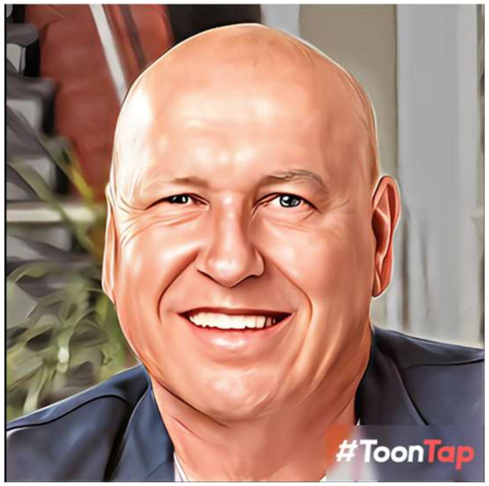
In Vitro Lecture and Luncheon for Students

AI Will Only Replace Toxicologists Who Do Not Use It!

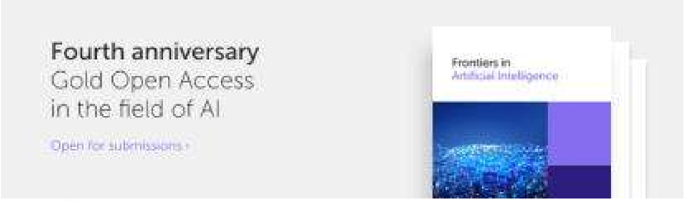
Thomas Hartung

SOT SALT LAKE CITY MARCH 10–14

JOHNS HOPKINS
BLOOMBERG SCHOOL of PUBLIC HEALTH



CAAT
Johns Hopkins University
Center for Alternatives to Animal Testing

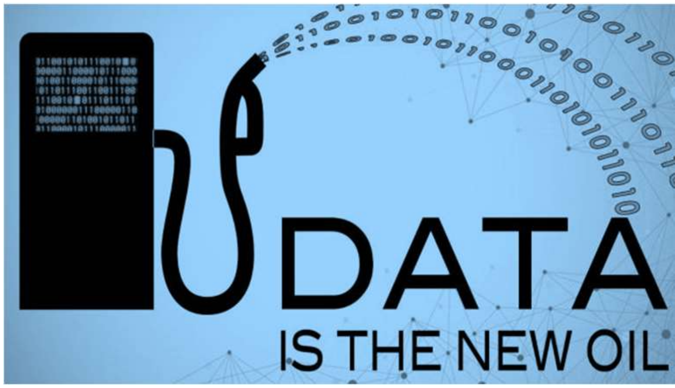


frontiers
in Artificial Intelligence

Thomas Hartung & team

AI will only replace toxicologists, who don't use it

A.I. = Making big sense of



If data is the new oil (Clive Humby, 2006), AI is the new combustion engine!

https://theamericangenius.com/editorials/big-data-is-watching-you-some-will-panic-others-will-rejoice/

Data: double every 18month
= 90% in last three years
Computer: double every 24 months (Moore's law)
AI: double every 3 months since 2010

Together increase >1 billion-fold (since we engaged in AI ten years ago)

Plagiarism?

Bias

Data gaps

Black box

Hallucination

Autonomous AI

Productivity

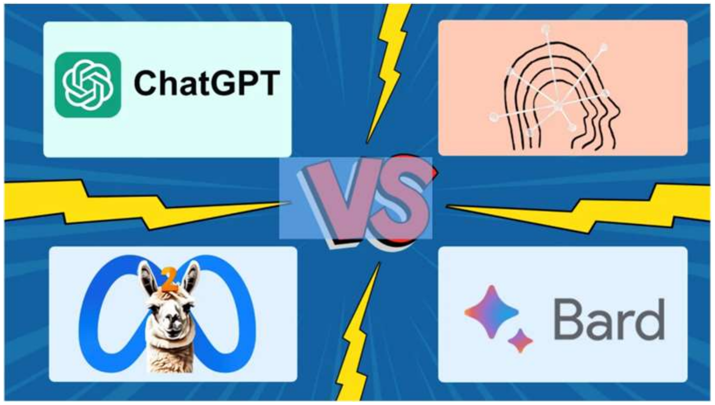
Information retrieval

Evidence integration of Big Data

Multi-modal

Toward xAI

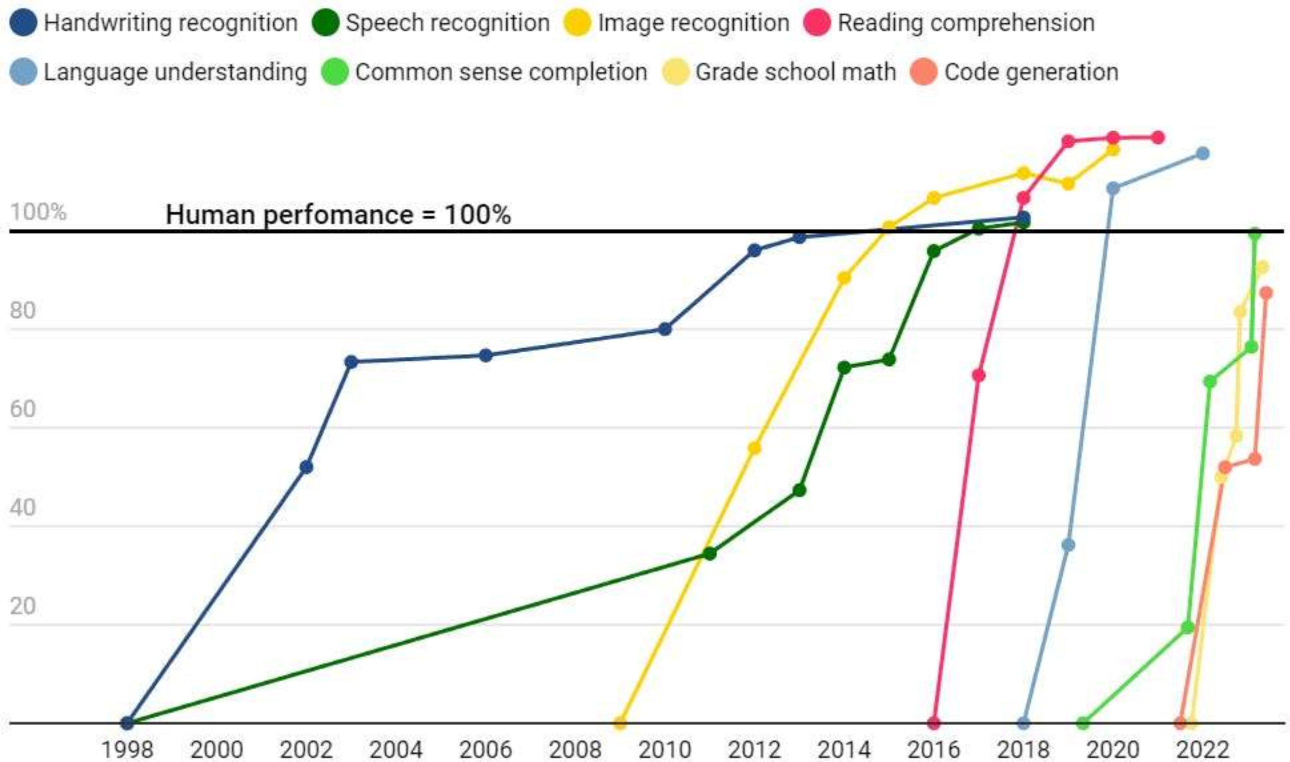
Human-in-loop



GPT-4 performed at the 90th percentile on a simulated bar exam, the 93rd percentile on an SAT reading exam, and the 89th percentile on the SAT Math exam, OpenAI claimed.

AI has surpassed humans at a number of tasks and the rate at which humans are being surpassed at new tasks is increasing

State-of-the-art AI performance on benchmarks, relative to human performance



For each benchmark, the maximally performing baseline reported in the benchmark paper is taken as the "starting point", which is set at 0%. Human performance number is set at 100%. Handwriting recognition = MNIST, Language understanding = GLUE, Image recognition = ImageNet, Reading comprehension = SQuAD 1.1, Reading comprehension = SQuAD 2.0, Speech recognition = Switchboard, Grade school math = GSK8k, Common sense completion = HellaSwag, Code generation = HumanEval.

Chart: Will Henshall for TIME • Source: ContextualAI

TIME

ToxAIcology

Big Data

- High-content (~omics & imaging)
- High-throughput (Robotized testing, e.g., Tox21 & ToxCast)
- Sensors
- Literature, Internet
- Legacy studies

Big Computer

AI & Machine Learning

- Natural Language Processing

Big Sense

- Data retrieval
- Evidence integration (systematic reviews, risk assessments)
- Predictive toxicology
- Digital pathology
- Reporting

Food for Thought ...

ToxAIcology – The Evolving Role of Artificial Intelligence in Advancing Toxicology and Modernizing Regulatory Science

Thomas Hartung[1,2]

Artificial intelligence as the new frontier in chemical risk assessment

frontiers | Frontiers in Artificial Intelligence

Thomas Hartung[1,2*]

ACCEPTED MANUSCRIPT

Machine learning of toxicological big data enables read-across structure activity relationships (RASAR) outperforming animal test reproducibility

Thomas Luechtefeld, Dan Marsh, Craig Rowlands, Thomas Hartung

Toxicological Sciences, kfy152, https://doi.org/10.1093/toxsci/kfy152

Published: 11 July 2018

Tom Luechtefeld

- Combining read-across with machine-learning
- Very large database
- Nine OECD test predicted
- 87% accuracy for 190,000 chemicals with known classifications
- 81% reproducibility of respective animal tests

Uses chemical similarity (network effect)

Uses transfer learning (74 labels)

AnImal Replacement

2022: Nine most used animal tests predicted by AI

AI predicted 4700+ food chemicals 83% correctly in 1h

= 38,000 animal studies at $250+ million

2018: Nine most used animal tests

AI predicted 190,000 chemicals 87% correctly

Animal reproducibility 81%

2023: Systemic toxicities

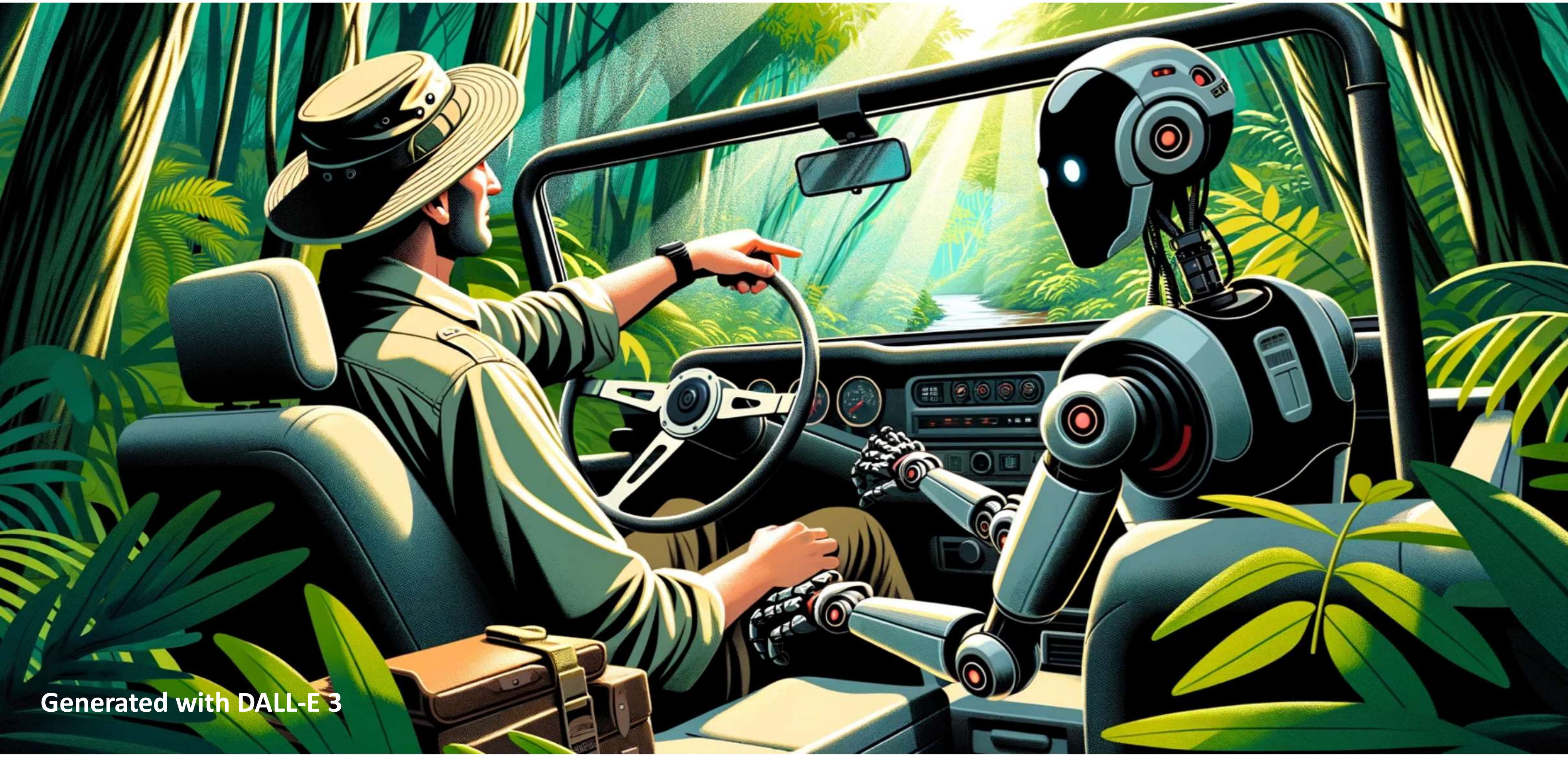
AI predicted 75% cancer risk of 950 chemicals and 82% reproductive tox of 1152 chemicals correctly

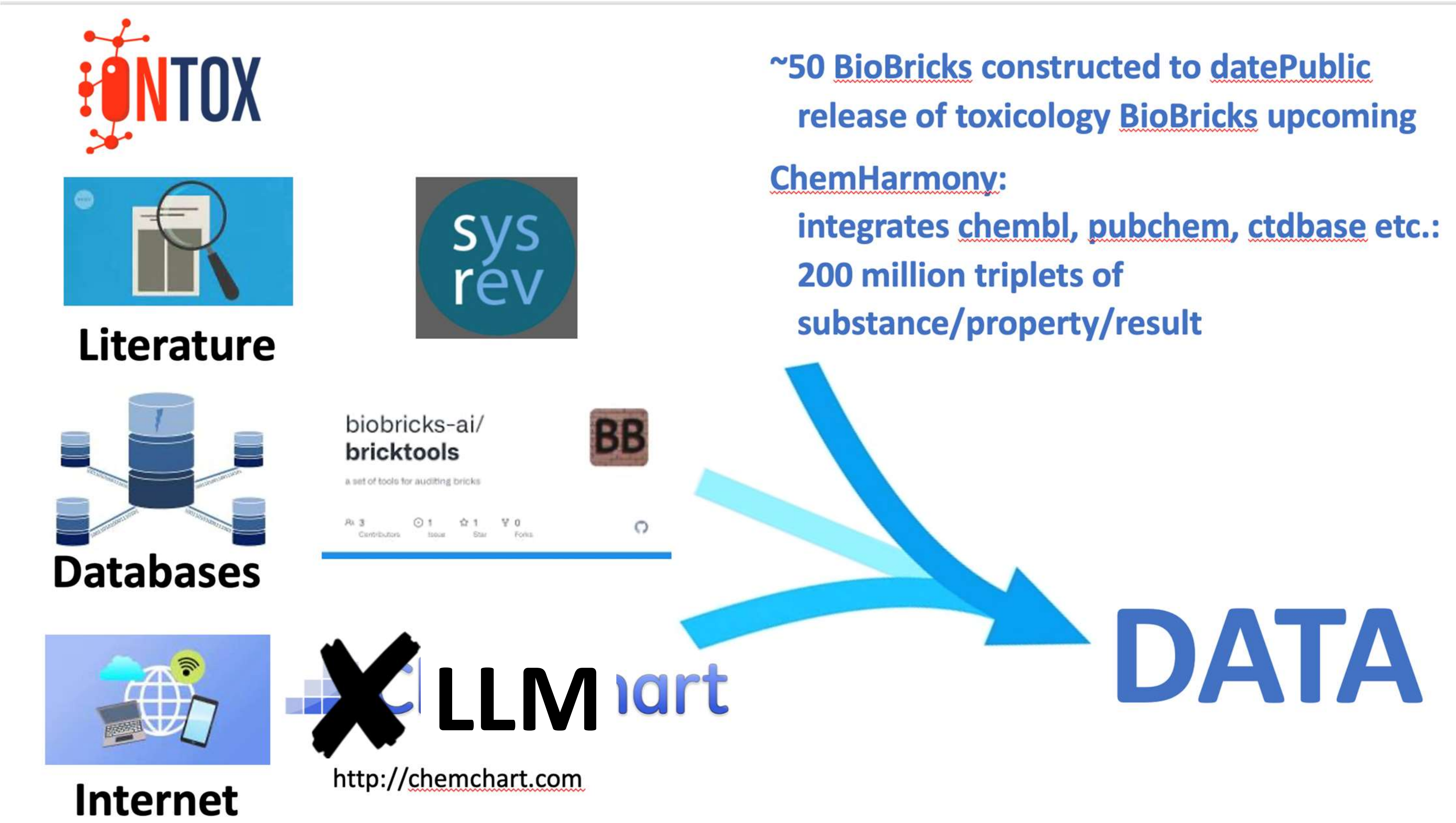
2020: Human Skin Sensitization

AI predicted 506 chemicals 80% correctly

Animal 74% correct

AI as copilot for toxicology




ONTOX
~50 BioBricks constructed to datePublic release of toxicology BioBricks upcoming
ChemHarmony:
integrates chembl, pubchem, ctdbase etc.:
200 million triplets of substance/property/result
sys rev
Literature
biobricks-ai/
bricktools
a set of tools for auditing bricks
BB
Databases
Internet
LLM
http://chemchart.com
DATA

Can we make a better similarity metric?

Structural similarity
(e.g., Morgan fingerprints)

Biological similarity

Test run:

- 1 million data points for 152 properties
- ~70% accuracies with chemical similarity
- ~ 80% accuracy for combined chemical and biological similarity

DATA

AOP networks
Biological & physiological maps

"Probability is the very guide of life."
Cicero (106 – 43 B.C.)

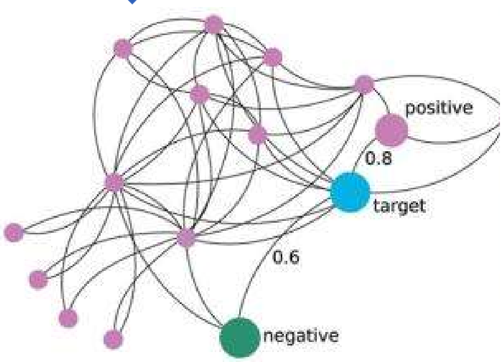
Food for Thought …

Probabilistic Risk Assessment – the Keystone for the Future of Toxicology

Alexandra Maertens[1], Emily Golden[1], Thomas H. Luechtefeld[1,2], Sebastian Hoffmann[1,3], Katya Tsaioun[1] and Thomas Hartung[1,4]

From perturbation of physiology

Probability of hazard



RASAR + QSAR

From chemical structure and properties


ProbRA only becomes
beautiful through AI

Generated with DALL-E and LeiaPix

- **Integrate Disruptive Technologies with Existing Knowledge**
- **Accelerate Drug Development**
- **Optimize Prevention and democratize Healthcare Access**

The Smart Path Forward

- **Open access, machine readable**
- **Identify bias in data, explainable AI**
- **Mechanistic and evidence-based approaches**

Discussion

AI Will Only Replace Toxicologists Who Do Not Use It!

Thomas Hartung

SALT LAKE CITY
MARCH 10–14

SOT SALT LAKE CITY MARCH 10–14

Discussion Questions

1. What was the risk assessment that was possible with these data in 2014? What could be done to support decisions with *in vivo* and *in vitro* tests?
2. How can AI help information retrieval in such cases nowadays? What are the problems?
3. Information is often contradictory and comes from very different sources. How does this impede emergency risk assessments? How can AI help to integrate information?
4. How can AI fill data gaps and assist in predicting toxicity by automated read-across in such cases?
5. Why are big data and transfer learning superior to traditional computational methods?
6. What are hallucinations in AI? How can these be mitigated? What would this mean in an emergency case?
7. What is the black box problem of AI and how can it be mitigated? Is this important in an emergency risk assessment?
8. Can AI support mechanistic and systems toxicology? Is this important for emergency situations?
9. How fast can we expect AI to develop in toxicology? How can it be used already?
10. What are the bottlenecks for AI use? How can we overcome them?