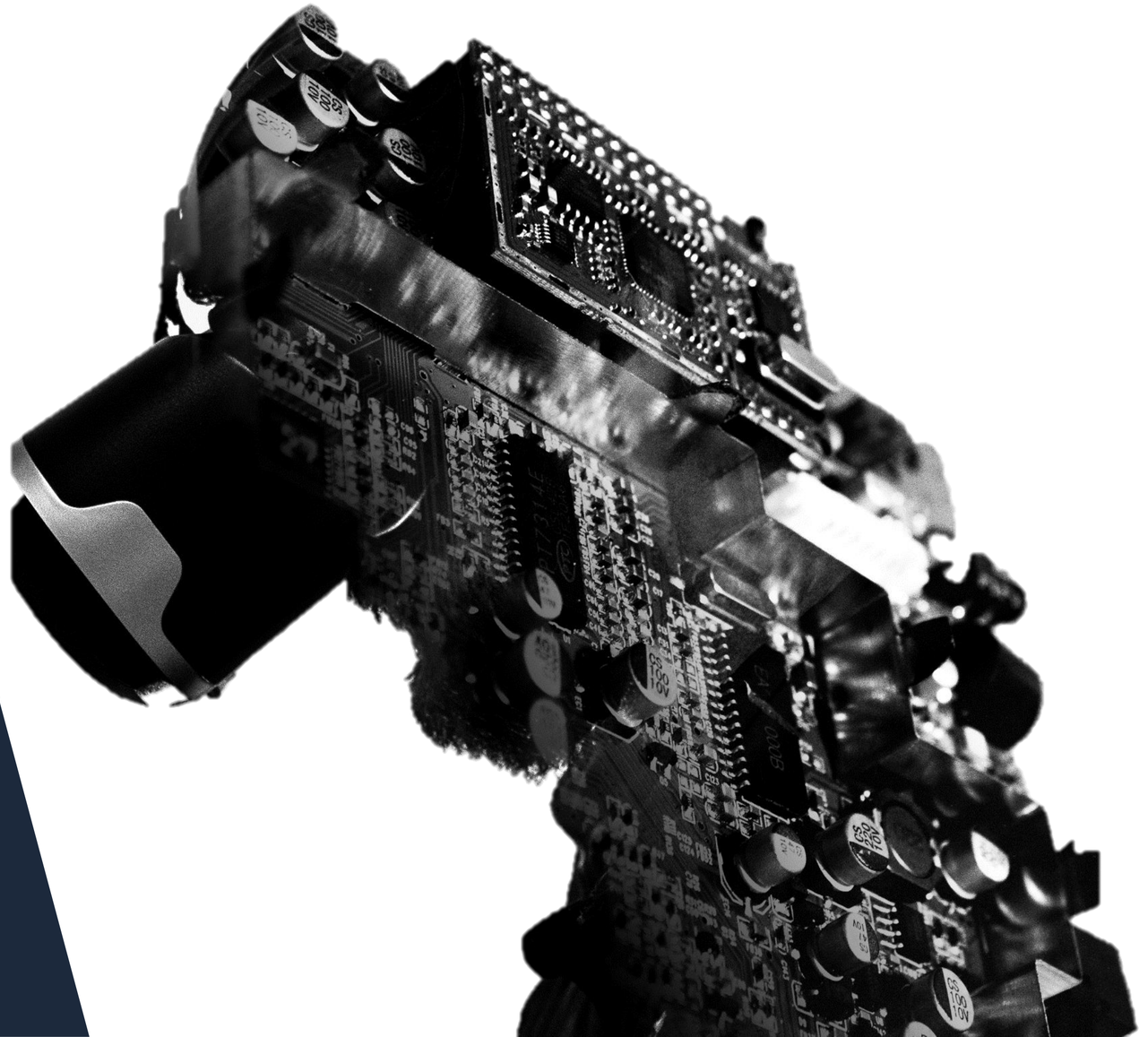
DEMYSTIFYING ARTIFICIAL INTELLIGENCE

Tim Miller

Professor of Artificial Intelligence
School of Electrical Engineering and Computer Science
The University of Queensland, Australia
E: timothy.miller@uq.edu.au
T: @tmiller_uq
W: https://uqtmiller.github.io/

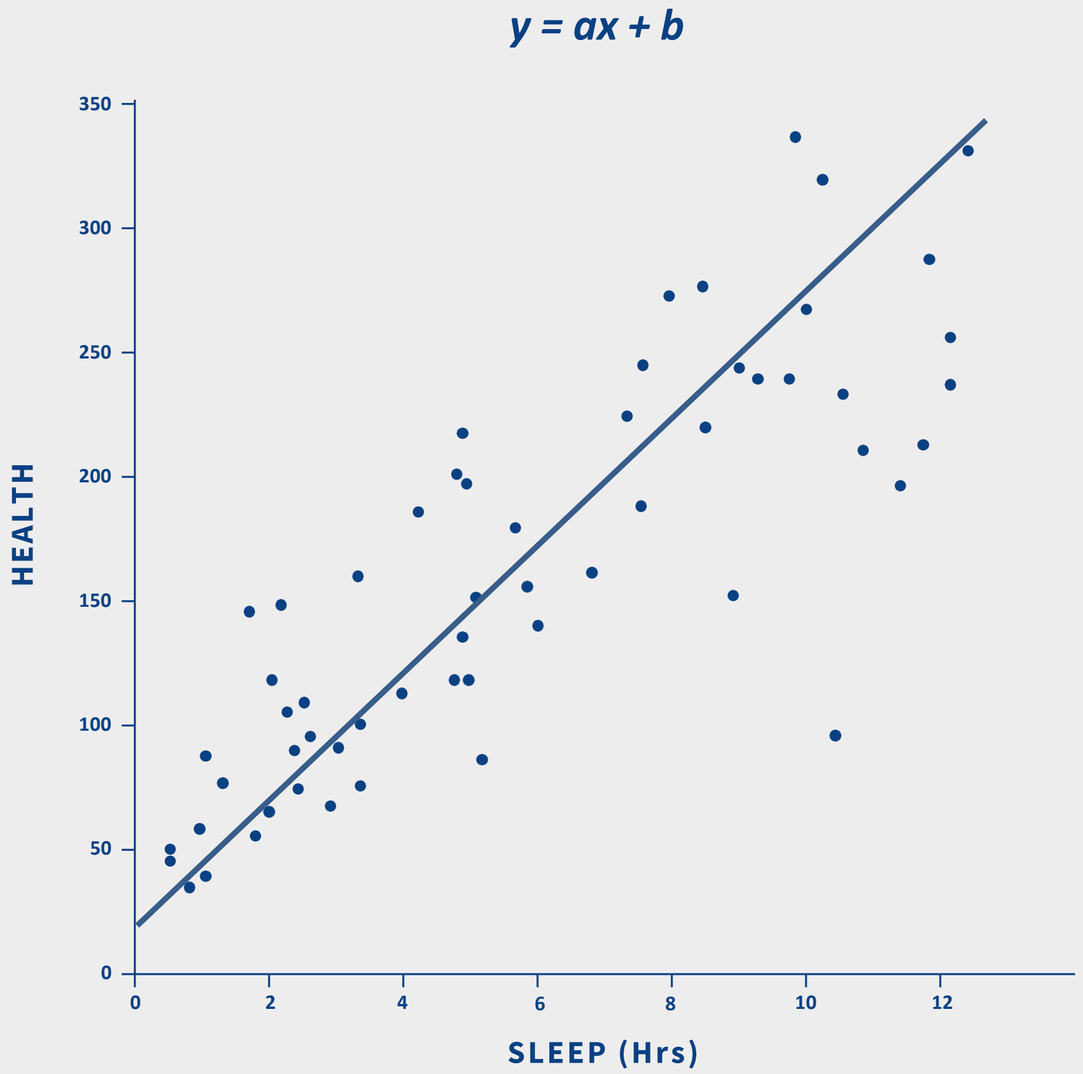


TALK OUTLINE

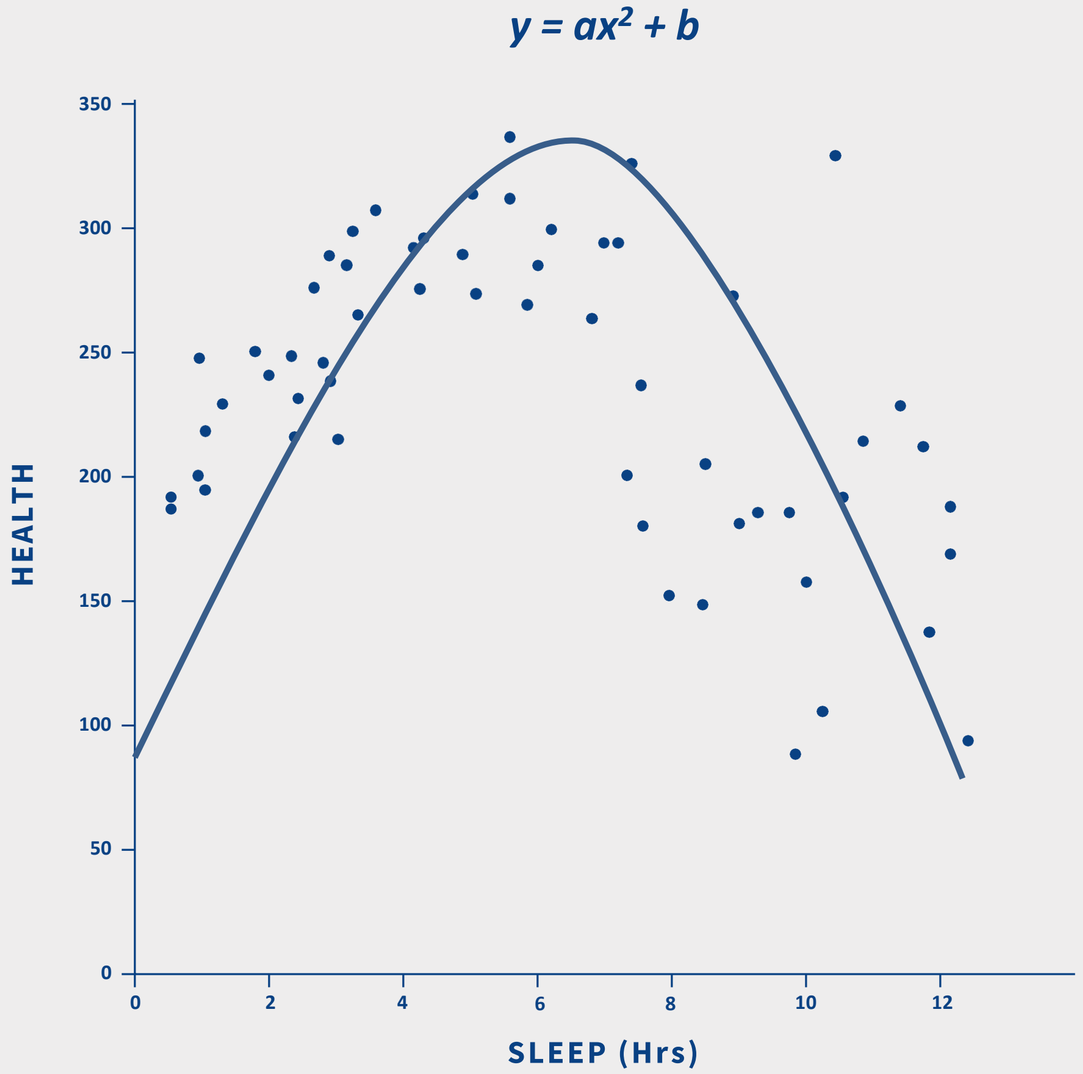
What is (modern) artificial intelligence?

How does it work?

MACHINE LEARNING

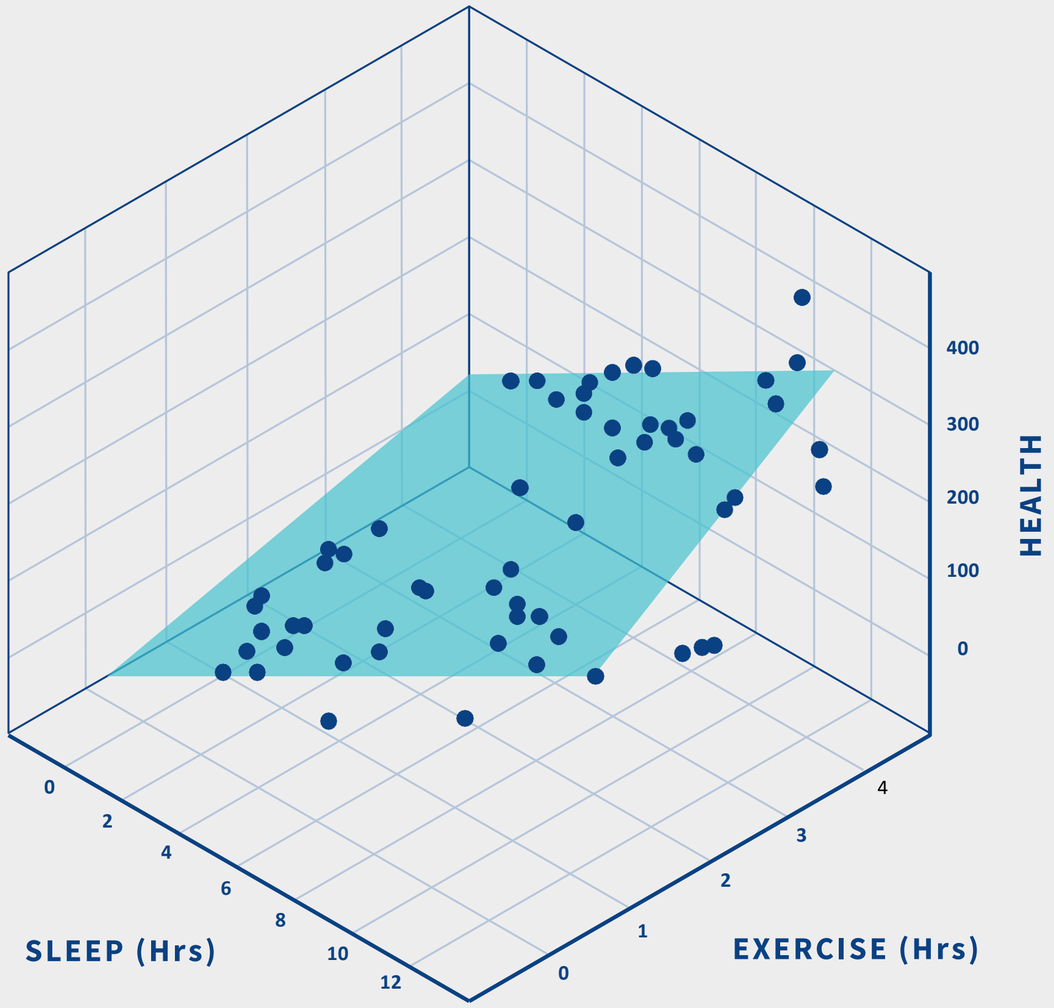


MACHINE LEARNING

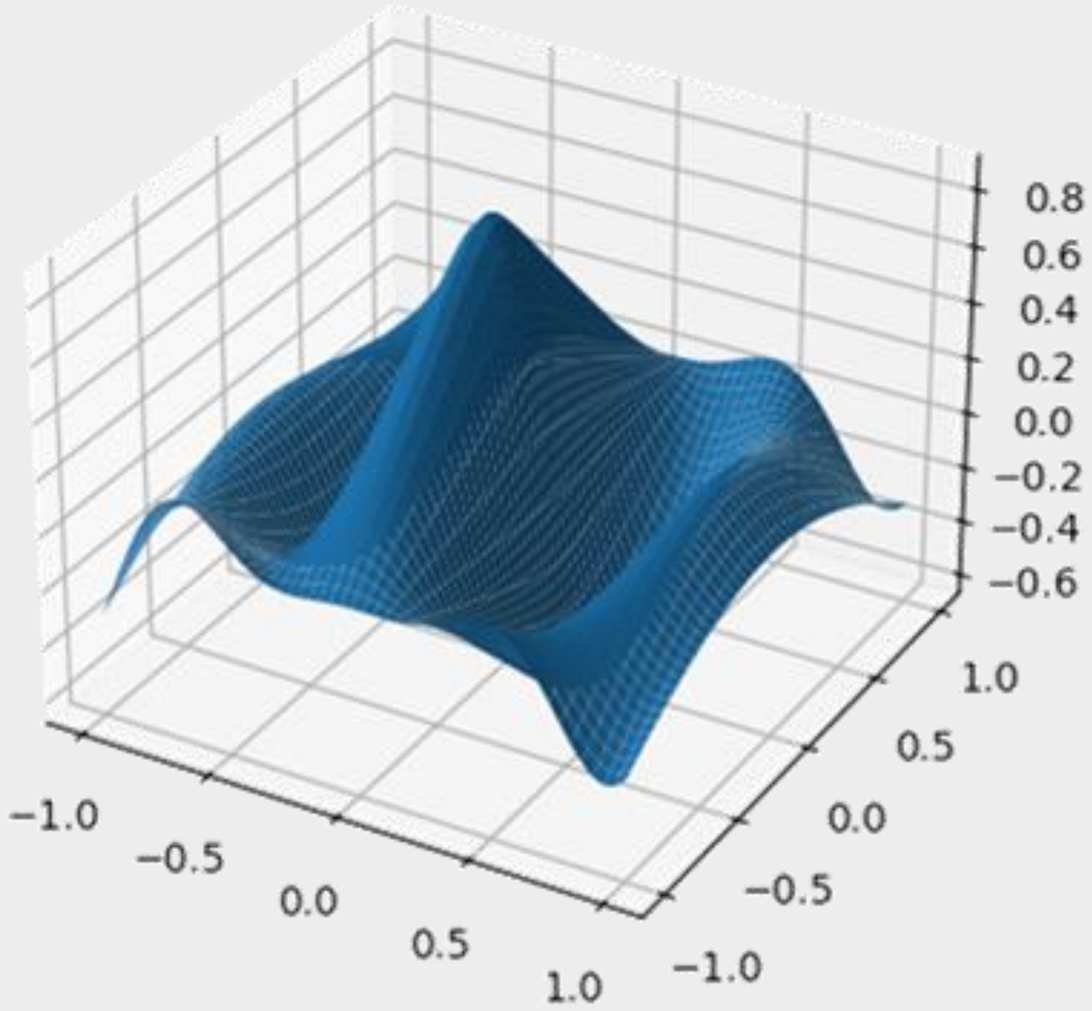


MACHINE LEARNING

$$y = ax + by + c$$



MACHINE LEARNING



WHY MACHINE LEARNING?

LARGE LANGUAGE MODELS

https://chat.openai.com/

Q: Explain the schooling system in Queensland, Australia: what levels are there, what qualifications can be gained, etc.?

LARGE LANGUAGE MODELS

https://chat.openai.com/

Q: I have 419 red apples and 201 green apples. I eat 300 of the apples. How many apples to I have left?

A: You have a total of 419 red apples + 201 green apples, which equals 620 apples initially.

If you eat 300 apples, subtract that from the initial count:

620 (initial count) - 300 (eaten) = 320 apples left.

So, you have 320 apples left after eating 300 of them.

LARGE LANGUAGE MODELS

https://chat.openai.com/

Q: I have 419 red apples and 201 green apples. I eat 300 of the apples. How many red apples do I have left?

A: You initially had 419 red apples. After eating 300 apples from your total count, you have:

419 (initial red apples) - 300 (eaten) = 119 red apples left.

So, you have 119 red apples left.

LARGE LANGUAGE MODELS

https://chat.openai.com/

Q: I have 419 red apples and 201 green apples. I eat 300 of the apples. How many green apples do I have left?

A: You initially had 201 green apples. After eating 300 apples from your total count, you have:

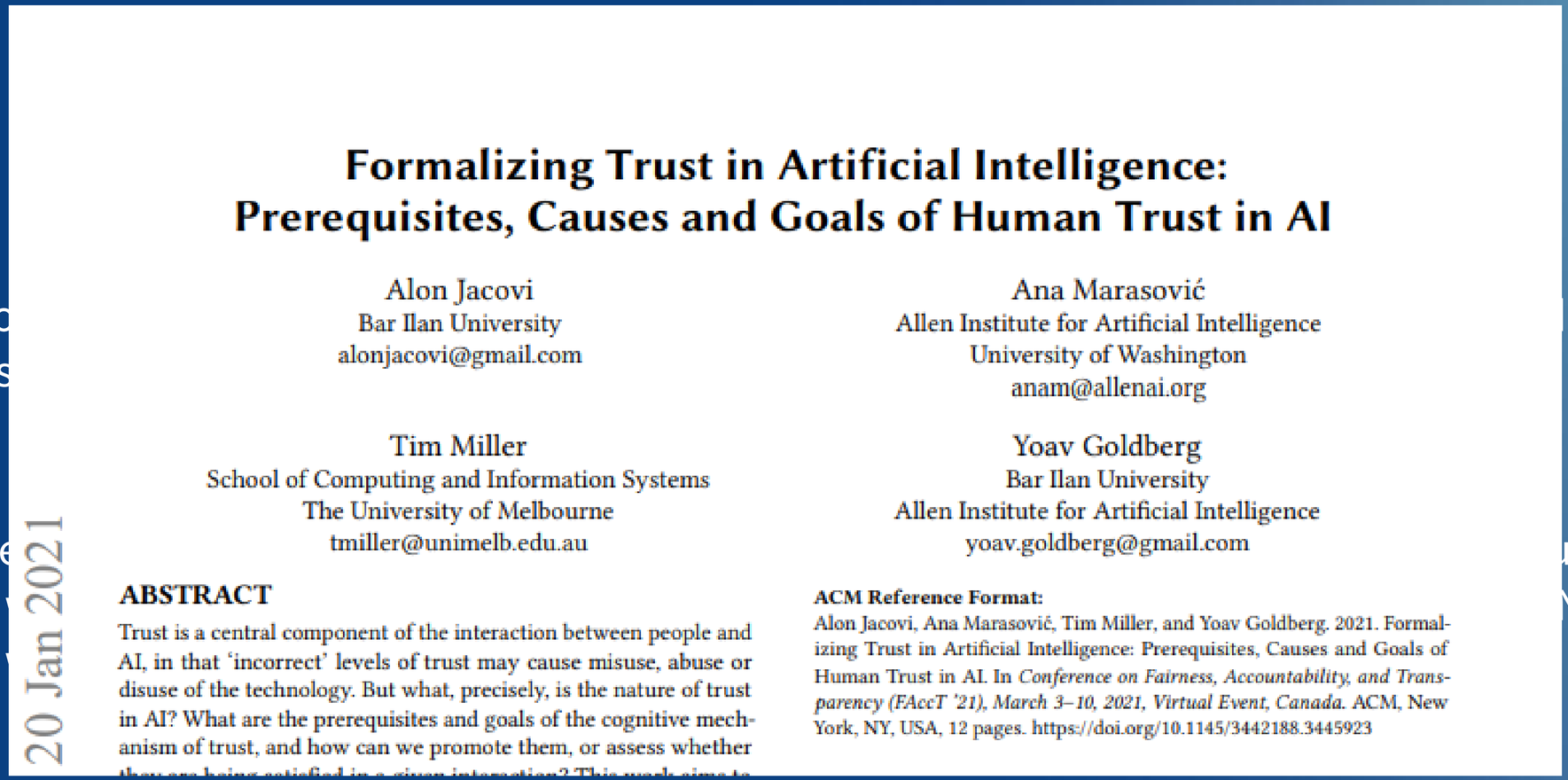
201 (initial green apples) - 300 (eaten) = -99 green apples.

It appears that you've eaten more green apples than you initially had, so you have a negative balance of 99 green apples. In reality, you don't have any green apples left, and you're left with only the remaining 119 red apples.

LARGE LANGUAGE MODELS

Q: Who wro … goals of human trus…

A: The pape… uman trust in AI" … Nancy K. Roderer. It …



20 Jan 2021

Formalizing Trust in Artificial Intelligence: Prerequisites, Causes and Goals of Human Trust in AI

Alon Jacovi
Bar Ilan University
alonjacovi@gmail.com

Ana Marasović
Allen Institute for Artificial Intelligence
University of Washington
anam@allenai.org

Tim Miller
School of Computing and Information Systems
The University of Melbourne
tmiller@unimelb.edu.au

Yoav Goldberg
Bar Ilan University
Allen Institute for Artificial Intelligence
yoav.goldberg@gmail.com

ABSTRACT

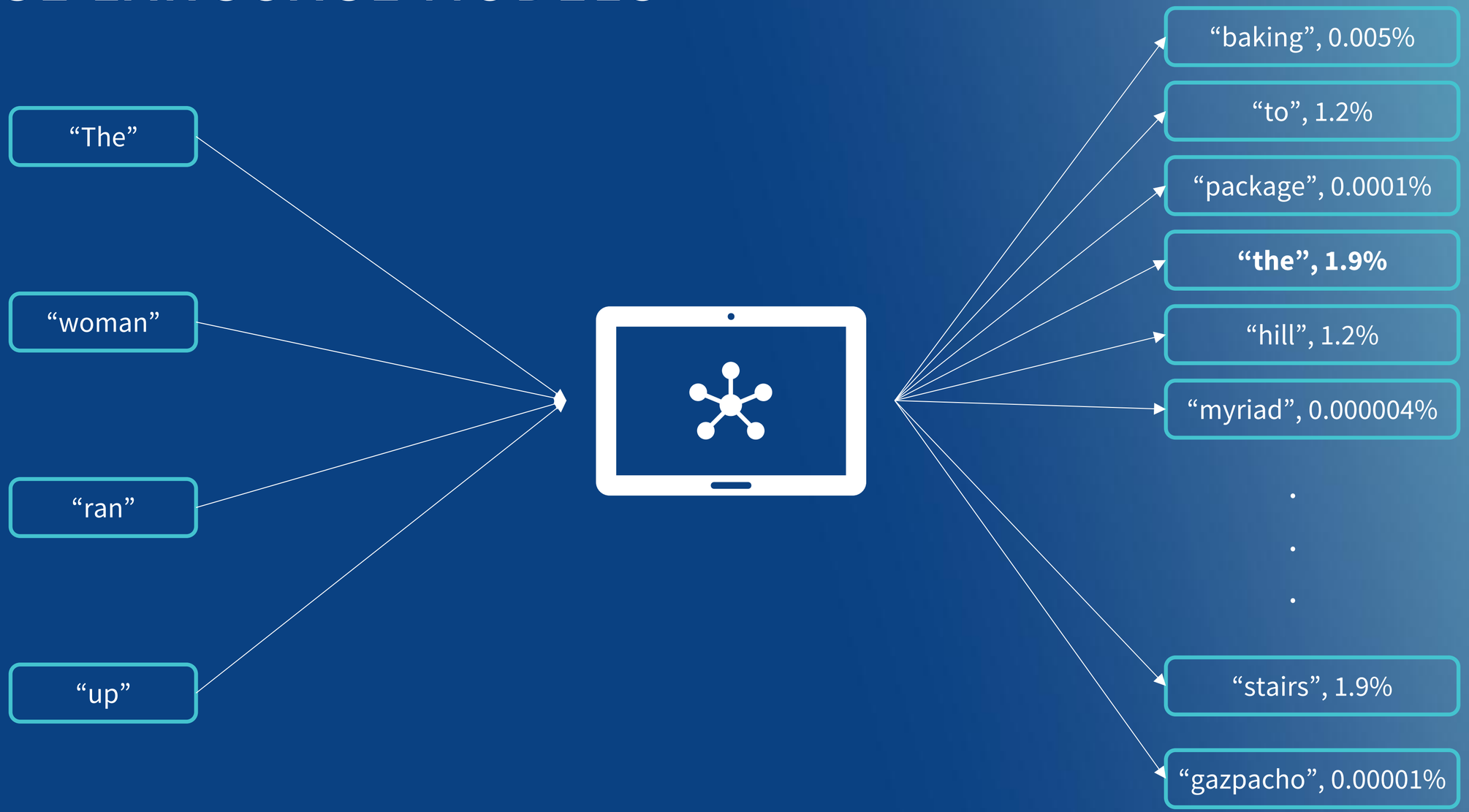
Trust is a central component of the interaction between people and AI, in that 'incorrect' levels of trust may cause misuse, abuse or disuse of the technology. But what, precisely, is the nature of trust in AI? What are the prerequisites and goals of the cognitive mechanism of trust, and how can we promote them, or assess whether

ACM Reference Format:
Alon Jacovi, Ana Marasović, Tim Miller, and Yoav Goldberg. 2021. Formalizing Trust in Artificial Intelligence: Prerequisites, Causes and Goals of Human Trust in AI. In *Conference on Fairness, Accountability, and Transparency (FAccT '21), March 3–10, 2021, Virtual Event, Canada.* ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3442188.3445923

ACTIVITY: HOW DO THESE LANGUAGE MODELS WORK?

“The woman ran up the _____”

LARGE LANGUAGE MODELS



LARGE LANGUAGE MODELS

HOW DO THESE LANGUAGE MODELS WORK?

“The woman ran up the _____”

“What did the woman run up? _____”

RETRIEVAL-AUGMENTED GENERATION (RAG)

https://www.perplexity.ai/

https://copilot.microsoft.com/

Q: Who wrote the paper "Formalizing trust in artificial intelligence: Prerequisites, causes and goals of human trust in AI"?

THANK YOU

Tim Miller

School of Electrical Engineering and Computer Science
The University of Queensland, Australia
E: timothy.miller@uq.edu.au
T: @tmiller_uq
W: https://uqtmiller.github.io/