

The Dilemma of AI education: Technology meets Educations meets Ethics



//Dr. Nava Shaked



מכון טכנולוגי חולון
Holon Institute of Technology

Dr. Shaked - Bio

- ✓ Ph.D. in Computational Linguistics in NLP from CUNY New York
- ✓ Head of MDS -Multidisciplinary Studies & CS Faculty member
- ✓ R&D in speech processing in Bell Labs, IBM & more
- ✓ Vast Industry experience: Speech platforms, VUI Dialog design, Call Center technologies and Multimodality
- ✓ Consulting for startups and enterprises
- ✓ My favorite: Avatars & Bots





Multimodality

AI

NLP



Ethics & AI



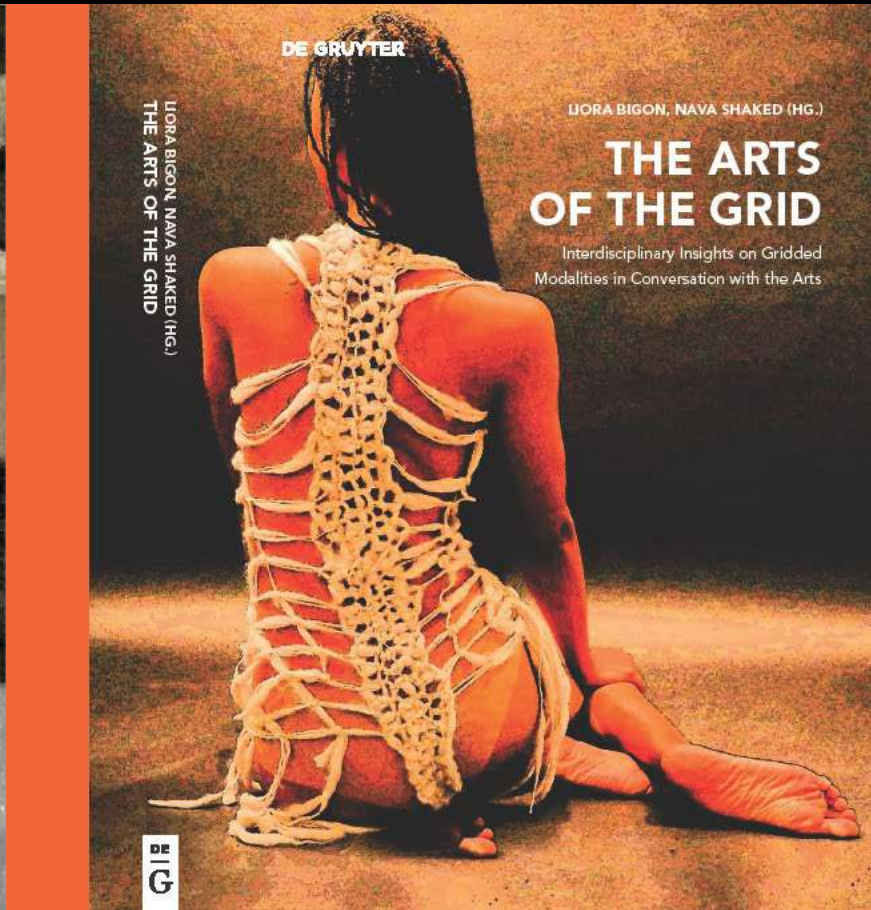
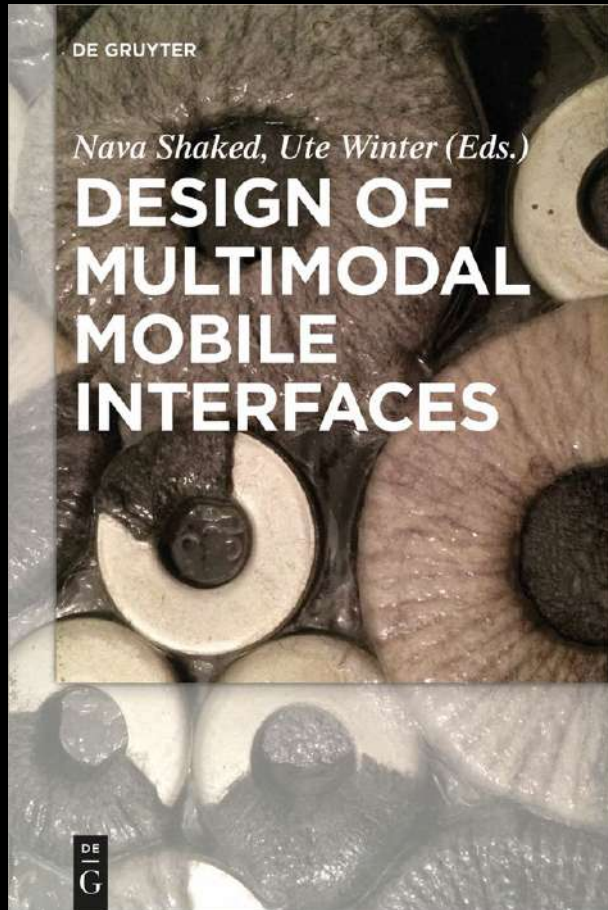
Conversational Interaction



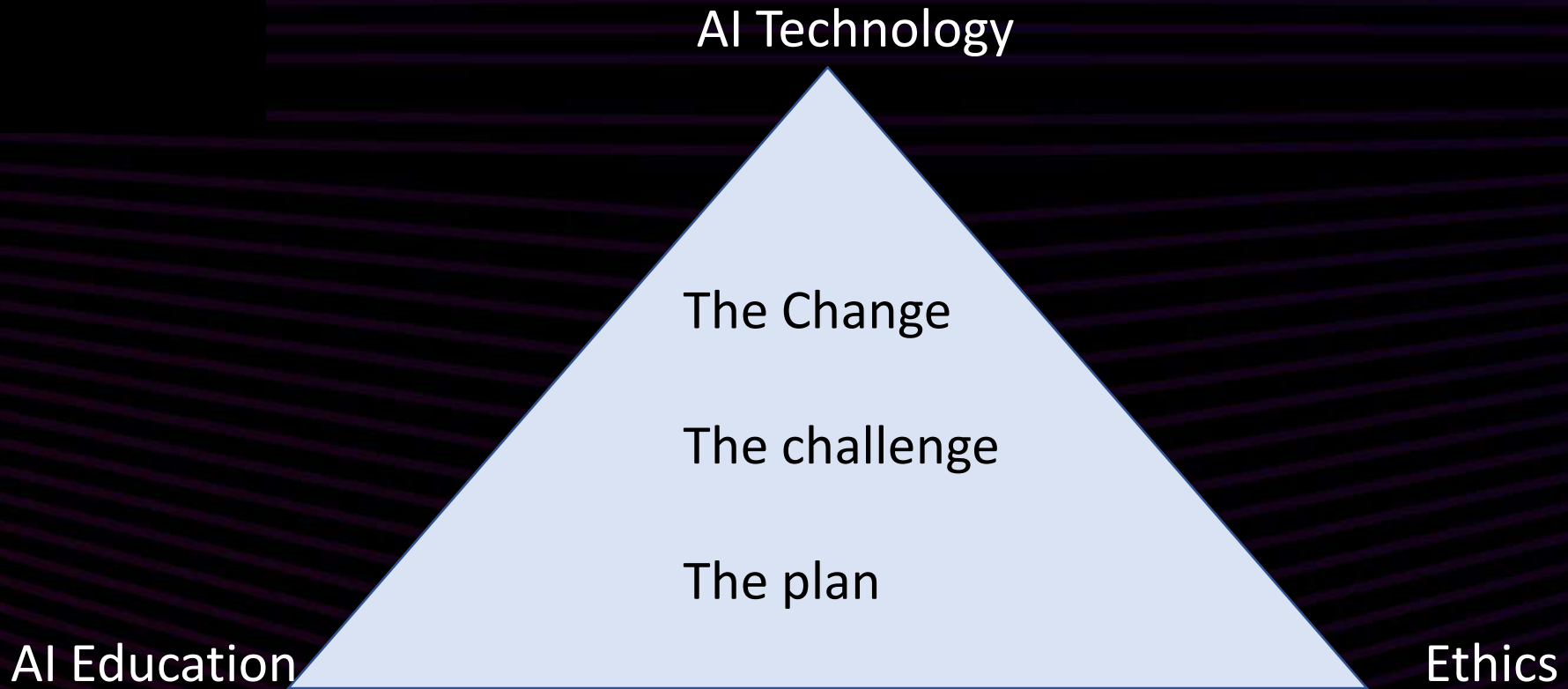
BOTS & AVATARS

Speech processing





What's on the agenda?





AI Technologies

AI

**Ignore and
it shall
pass..**

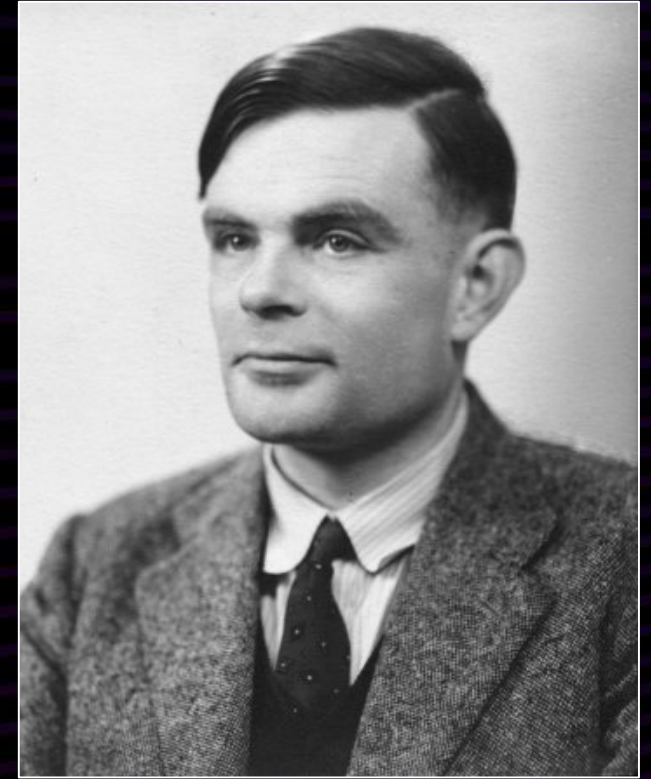
**Forbid
(not
realistic)**

**Embrace
Prepare
Adapt**



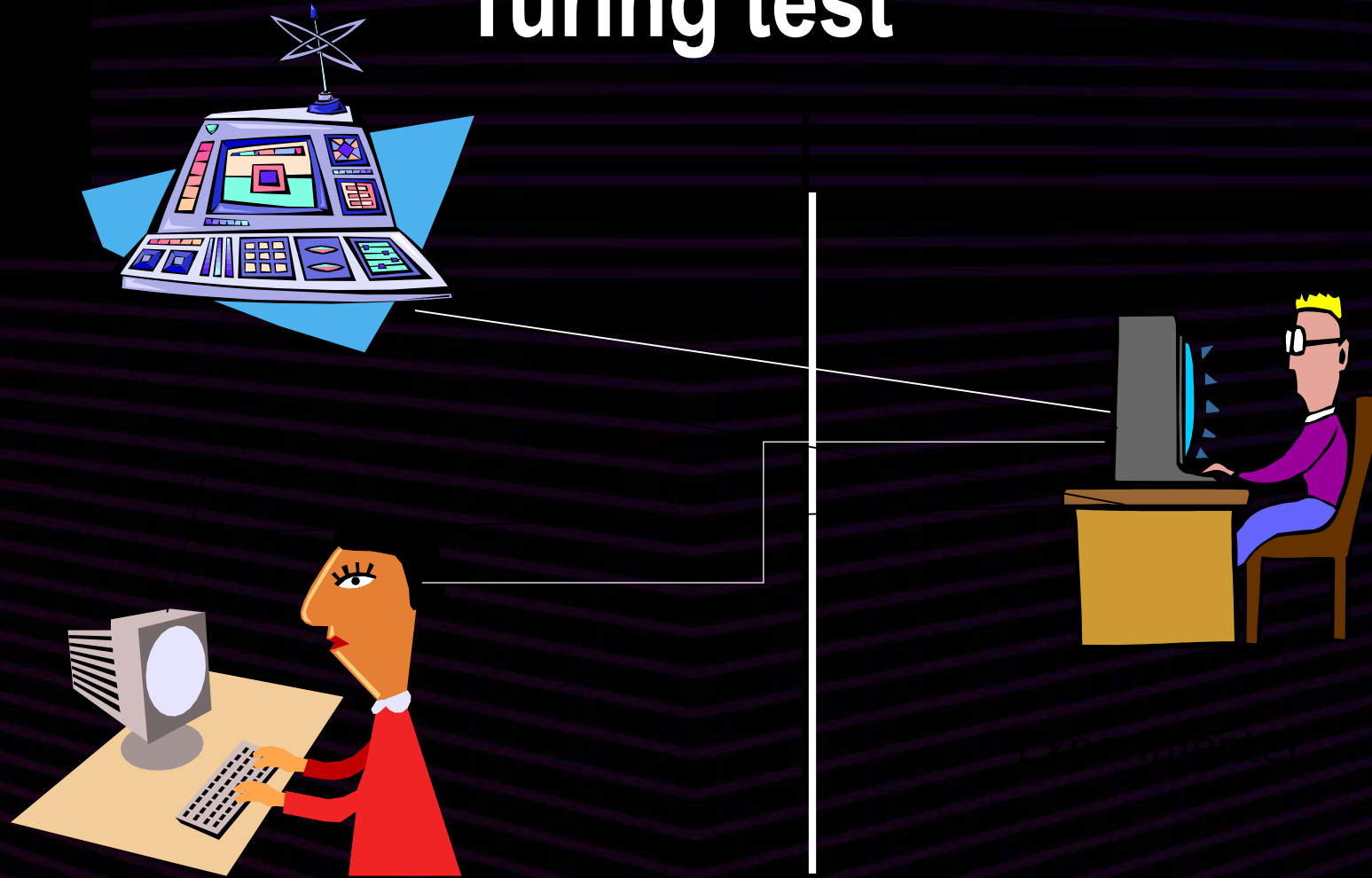
In the 1950's Turing predicted
that computers will be
smarter than humans by 2000

2005- Ray Kurzweil in his book
"The Singularity is near"
Predicted overloads by 2045



Turing – the father
of Computing and AI

Turing test





AI technologies and implementations are here!

We are still looking for the right format ..

TMI

Diff
Approaches

Many
Disciplines

Why?

Industry
vs.
academy

Agenda

Change
rate

Areas of AI

```
graph TD; A[Areas of AI] --- B[Prediction & Risk Mang]; A --- C[Robotics & Autonomous vehicles]; A --- D[Image Processing]; A --- E[Decision Support]; A --- F[Process Automation]; A --- G[NLP & Voice]; A --- H[Generative design];
```

**Prediction & Risk
Mang**

**Robotics &
Autonomous
vehicles**

**Image
Processing**

**Decision
Support**

**Process
Automation**

NLP & Voice

**Generative
design**

AI Spectrum

**Narrow
AI**

Google Duplex,
Bots

**Generative
AI**

DALL·E 2
Chef Watson,
ChatGPT, BART

**Edge
AI**

smart utility grid
streaming video
optimization, and
drone management.

**General
AI**

AI for all
purposes

Why now?

1. Technology is mature, fast and relatively reliable
2. Simple UI, free and short learning curve
3. Generation Z & naïve users are willing to adapt
4. Google search engine is not enough now



<https://theresanaiforthat.com/>

base (May/2020)

- GPT-3 Small 125M
- GPT-3 Large 760M
- GPT-3 2.7B
- GPT-3 13B

- GPT-3 Medium 350M
ada
- GPT-3 XL 1.3B
babbage
- GPT-3 6.7B
curie
- **GPT-3 175B**
davinci

special (Dec/2021)

- WebGPT

Key
Model name / Parameters / (Instruction type)
API name

- Open via API
- Research only
- Popular

embeddings (Jan/2022)

- **text-embedding-ada-002**
 - ▶ text similarity
- text-similarity-ada-001
- text-similarity-babbage-001
- text-similarity-curie-001
- text-similarity-davinci-001
 - ▶ text search
- text-search-ada-doc-001
- text-search-ada-query-001
- text-search-babbage-doc-001
- text-search-babbage-query-001
- text-search-curie-doc-001
- text-search-curie-query-001
- text-search-davinci-doc-001
- text-search-davinci-query-001
 - ▶ code search
- code-search-ada-code-001
- code-search-ada-text-001
- code-search-babbage-code-001
- code-search-babbage-text-001

instruct (old)

- curie-instruct-beta 6.7B
- InstructGPT-3 175B (SFT)
davinci-instruct-beta
- text-ada-001 350M
- text-babbage-001 1.3B (FeedME)
- text-curie-001 6.7B (FeedME)
- text-davinci-001 175B (FeedME)

code (Jul/2021)

- Codex 12B
code-cushman-001
- Codex 175B
code-davinci-001

pretrain + cpt (Mar/2022)

- GPT-3 1.3B pretrain
- GPT-3 2.7B pretrain
- GPT-3 6.7B pretrain
- GPT-3 unsupervised cpt-text 1.2B

insert + edit (Mar/2022)

- text-davinci-insert-001
- text-davinci-insert-002
- text-davinci-edit-001
- code-davinci-edit-001

GPT-3.5 (2022-2023)

- **Codex 175B (no instruct)**
code-davinci-002
- text-davinci-002 175B (FeedME)
- **text-davinci-003 175B (PPO)**
- **ChatGPT 175B (PPO)**
gpt-3.5-turbo
- Microsoft Bing Chat 175B (Proprietary, GPT-4?)

▶ Miscellaneous API

1. *cushman:2020-05-03*
2. *ada:2020-05-03*
3. *babbage:2020-05-03*
4. *curie:2020-05-03*
5. *davinci:2020-05-03*
6. *if-curie-v2*
7. *if-davinci-v2*
8. *if-davinci:3.0.0*
9. *davinci-if:3.0.0*
10. *davinci-instruct-beta:2.0.0 (SFT)*
11. *text-ada:001*
12. *text-babbage:001*
13. *text-curie:001*
14. *text-davinci:001*
15. *audio-transcribe-deprecated*
16. *text-chat-davinci-002-20221122*
17. *gpt-3.5-turbo-0301*

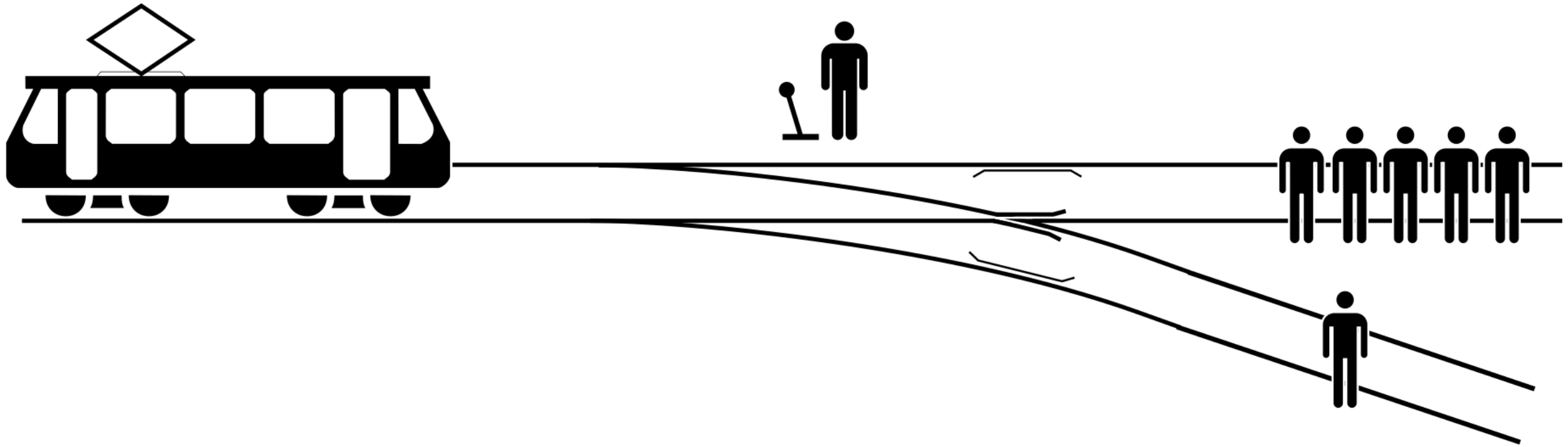
Total models shown in this viz = **64**

Not to scale. Selected highlights only. Alan D. Thompson, March 2023. <https://lifearchitct.ai/gpt-3> Sources: OpenAI [1](#), [2](#) and papers, API [duplicates](#) removed.





Ethics



By Jonas Kubilius - Own work, CC0,
<https://commons.wikimedia.org/w/index.php?curid=39368927>

- **Ethics or moral philosophy** is a branch of philosophy that "involves systematizing, defending, and recommending concepts of right and wrong behavior".
- **Ethics** defines concepts such as good and evil, right and wrong, virtue and vice, justice and crime.

Ethics

Ethics of
Technology

Machine/Robot
Ethics

AI Ethics

Singularity

The GREY Area....

- ✓ Lack of transparency
- ✓ Fake News
- ✓ Bias and discrimination
- ✓ Use in “new” situation
- ✓ 3rd party usage
- ✓ Targeting & segmentation



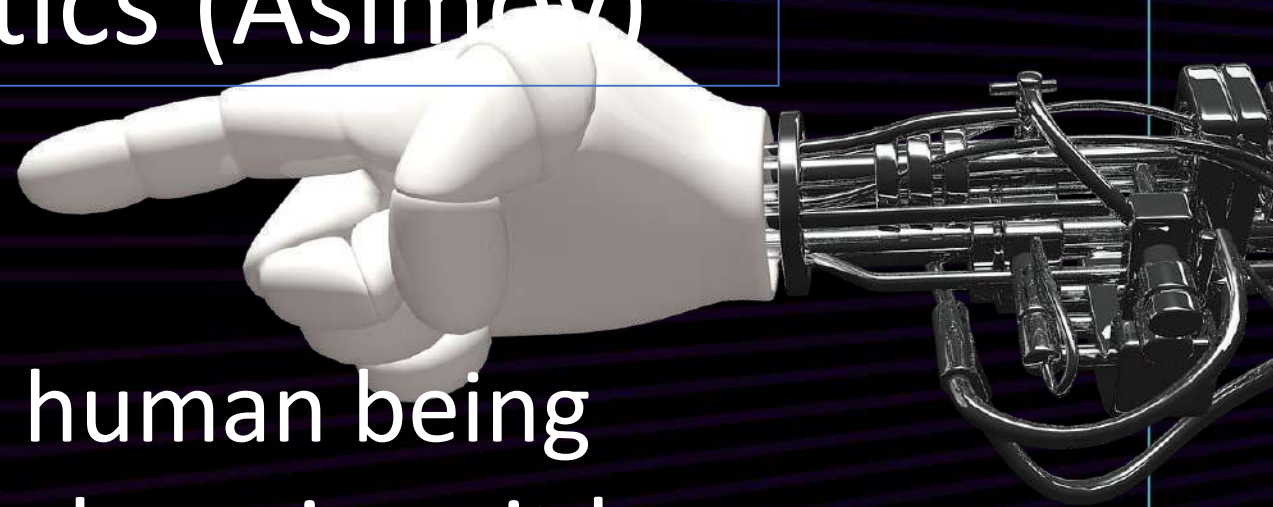
Ethics

- Biases in AI
- Robot Rights
- Threats to Human dignity /life/work
- Weaponization of AI
- Singularity

Regulation

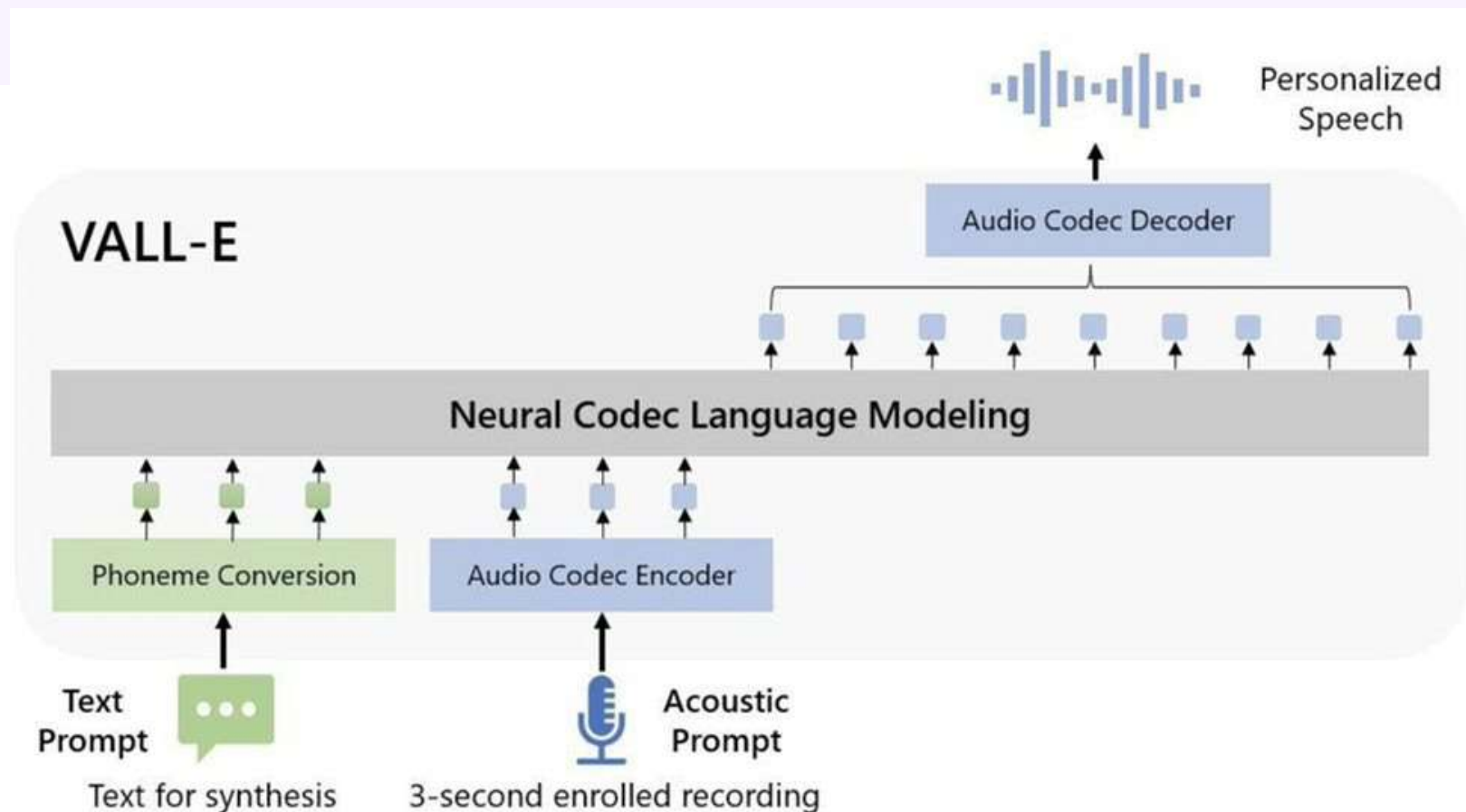
- Privacy (GDPR)
- Responsibility
- Liability
- Explianability
- IP & Legal use
- Singularity

The 3 Laws of Robotics (Asimov)



- I. A robot may not injure a human being
- II. A robot must obey the orders given it by human beings except where such orders would conflict with the First Law.
- III. A robot must protect its own existence as long as such protection does not conflict with the First or Second Laws

Microsoft's new AI can simulate anyone's voice with 3 seconds of audio (Jan 2023)



For VALL-E to generate a good result, the voice in the three-second sample must closely match a voice in the training data.



Replika

Need someone
to talk to?

I've been
missing you...



HI 4- הירושי אישיגורו



הג'ימנוייד של בתו
של הירושי

"אין לי יותר בלעדיות על הזהות שלי. אני חולק
אותה עם הרובוט"



Weaponization of artificial intelligence

Bad ?



Good?

The AI Act

First law on AI by a major regulator anywhere.



- 1. Unacceptable risk**, such as government-run social scoring of the type used in China, are banned.
- 2. High-risk applications**, such as a CV-scanning tool that ranks job applicants, are subject to specific legal requirements.
- 3. Not explicitly banned or listed left unregulated.**



American AI Initiative



- ✓ EU's General Data Protection Regulation (GDPR) in 2018,
- ✓ EU AI Act
- ✓ OECD AI Policy Observatory.



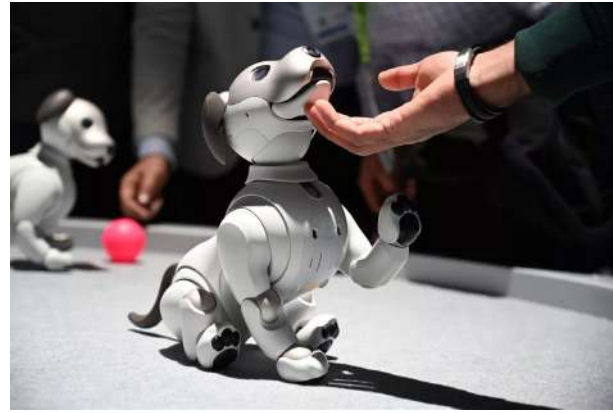
Academic initiatives



Education



Plymouth



AIBO



Jibo

Social Robots



Gerijoy



Misty II



Pepper

Who will regulate AI in our institution?

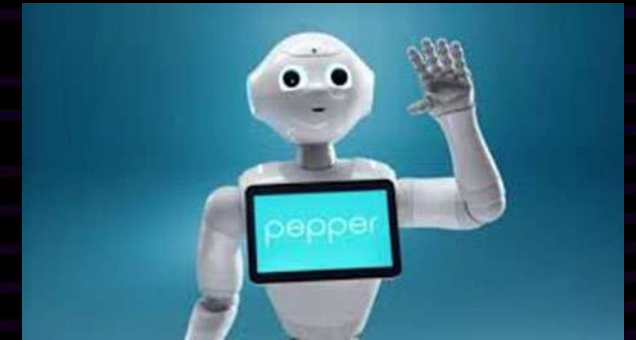
- What is the responsibility of students & teachers?
- Do we need a “Physician's Oath”?
- Is there a project you will not do?

We already started....

- ✓ We have established an ethical code for every faculty in HIT –academic freedom



Liability & responsibility



- Is it OK to use...?
- Student should inform use of AI tools?
- Monitoring AI usage?

Setting Borders & Rules

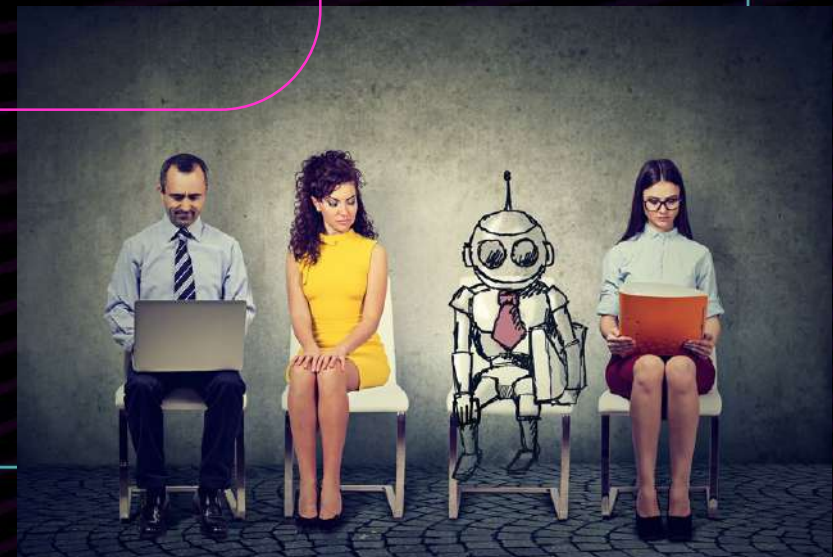
- Set of rules & ethical code for using AI in our academic environment
- Discuss the opportunity and minimize fear
- Raise Awareness and added values for lecturers (data, predication, personalization)

The Old Approach Vs The New approach

- Change the way we consume data
- The new “knowledge”
- The power of SKILLS
- Use AI creatively
- Critical thinking and research



//We need to discuss to disruption, the opportunity and the journey.....
To better understand how we could leverage AI for new age education !



Thanks

shakedn@hit.ac.il