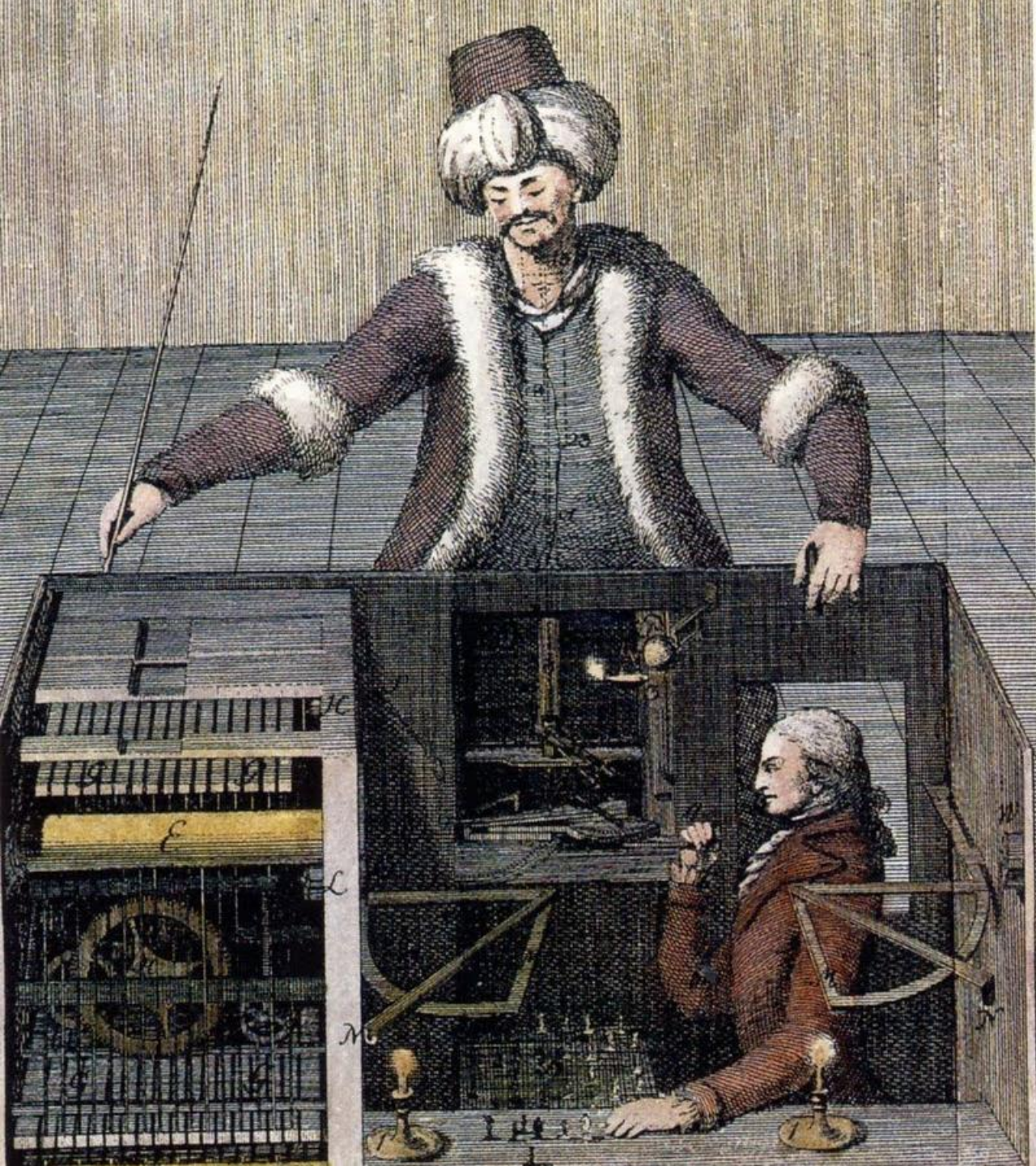


Emberi etika – gépi logika

2019. május 14.

Körmendi György
Clementine





Áttekintés

Gépi gondolkodás hibái

3 fejlődési szint

Az MI működés monitorozása

2018 – az őszinte aggodalom éve



Hírek » Mérleg rovat

A Microsoft elnöke aggódik az arcfelismerő szoftverek miatt



TECH

CYBERSECURITY | ENTERPRISE | INTERNET | MEDIA | MOBILE | SOCIAL MEDIA

Tim Cook: Personal data collection is being 'weaponized against us with military efficiency'



nek megszólalását, mely cég erősen függ a

asonlóan még mindig vezető jogtanácsosa –, Brad Leiber fejte ki, hogy a ma már mindennaposnak számító telen, mivel olyan eszközről van szó, mely az állami lehetőségeket biztosíthat. Smith ezért azt tartaná különben „2024-ben egyszer csak 1984-ben könyvére.

l ismerés egy olyan technológia, mely túlhatalmat a felett: a bekamerázott világban roppant mód szempontból is aggályos szerinte: az üzleti világ is gy kereskedő az arcfelismerés alapján azonosít egy

EDITION: EU

ZDNet



CENTRAL EUROPE MIDDLE EAST SCANDINAVIA AFRICA UK ITALY SPAIN MORE NEWSLETTERS

MUST READ: Marijuana means business: Everything you need to know about the cannabis industry

AI 'more dangerous than nukes': Elon Musk still firm on regulatory oversight

The man building a spaceship to send people to Mars has used his South by Southwest appearance to reaffirm his belief that the danger of artificial intelligence is much greater than the danger of nuclear warheads.

By Asha McLean | March 12, 2018 -- 18:50 GMT (18:50 GMT) | Topic: Innovation

2018 - az MI politikai téma

- Oroszok - propaganda, trollok - indirekt
- Kínaiak – megfigyelőrendszer, pontrendszer
- USA – személyi szabadság – nagy multik

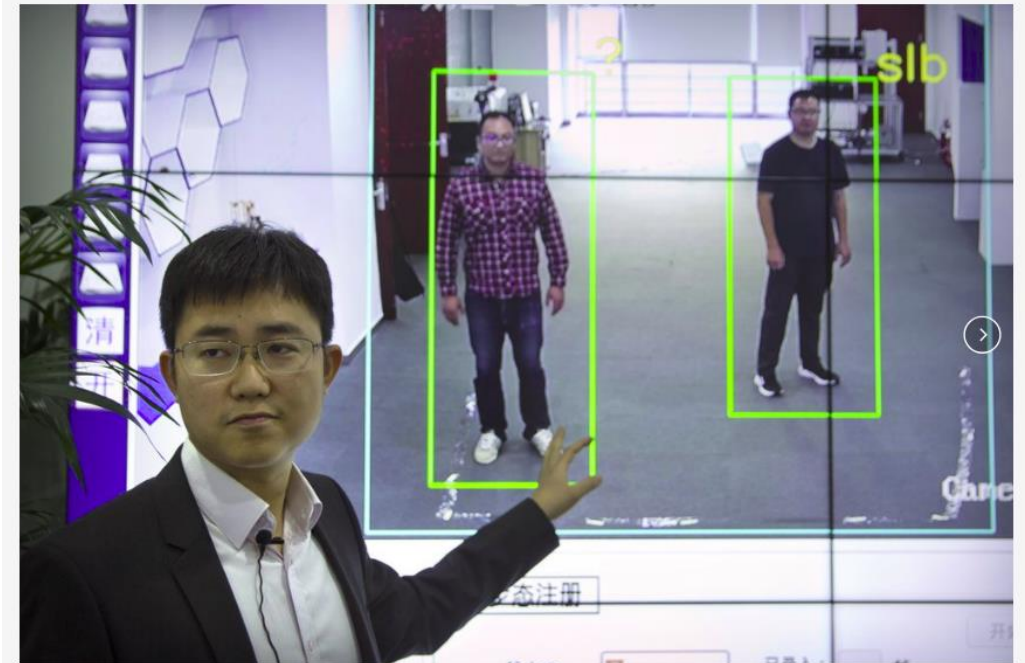
The New York Times

U.S. Begins First Cyberoperation Against Russia Aimed at Protecting Elections



Chinese 'gait recognition' tech IDs people by how they walk

By DAKE KANG November 6, 2018



Click to copy

RELATED TOPICS

Beijing

BEIJING (AP) — Chinese authorities have begun deploying a new surveillance tool: “gait recognition” software that uses people’s body shapes and how they walk to identify them, even when their faces are hidden from cameras.

Already used by police on the streets of Beijing and Shanghai, “gait recognition” is part of a push across China to develop artificial-intelligence and data-driven surveillance that is raising concern about how far the technology will go.

2018 – EU és a GDPR

- Oroszok - propaganda, trollok
- Kínaiak – megfigyelőrendszer, pontrendszer
- USA – személyi szabadság – nagy multik
- EU – GDPR



TECH

[CYBERSECURITY](#) | [ENTERPRISE](#) | [INTERNET](#) | [MEDIA](#) | [MOBILE](#) | [SOCIAL MEDIA](#)

Tim Cook: Personal data collection is being 'weaponized against us with military efficiency'



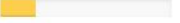
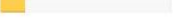


GDPR

Don't feed the AI

Az előítéletes MI

- Arcfelismerés problémája
nők <-> férfiak
fehérek <-> feketék
- AI alapú komoly döntések

Gender Classifier	Female Subjects Accuracy	Male Subjects Accuracy	Error Rate Diff.
 Microsoft	89.3% 	97.4% 	8.1% 
 FACE++	78.7% 	99.3% 	20.6% 
 IBM	79.7% 	94.4% 	14.7% 



<https://www.aclu.org/blog/privacy-technology/surveillance-technologies/amazons-face-recognition-falsely-matched-28>

<https://www.media.mit.edu/projects/gender-shades/>

Az előítéletes MI

- Arcfelismerés problémája
nők <-> férfiak
fehérek <-> feketék
- AI alapú komoly döntések
- Amazon Rekognition
kongresszusi képviselők vs
bűnözők (99.9997% accuracy!)
- Google Photos témafelismerő hibája



<https://www.aclu.org/blog/privacy-technology/surveillance-technologies/amazons-face-recognition-falsely-matched-28>

<https://www.media.mit.edu/projects/gender-shades/>

Az előítéletes MI

- Arcfelismerés problémája
nők <-> férfiak
fehérek <-> feketék
- AI alapú komoly döntések
- Amazon Rekognition
kongresszusi képviselők vs
bűnözők (99.9997% accuracy!)
- Google Photos témafelismerő hibája
- Lehet-e Clark Kent bűnöző?
- Az utálkozó chatbot esete
- Sokszor black box a
működési elv ->
jó lenne látni, hogy mikor és
miért torzít a model

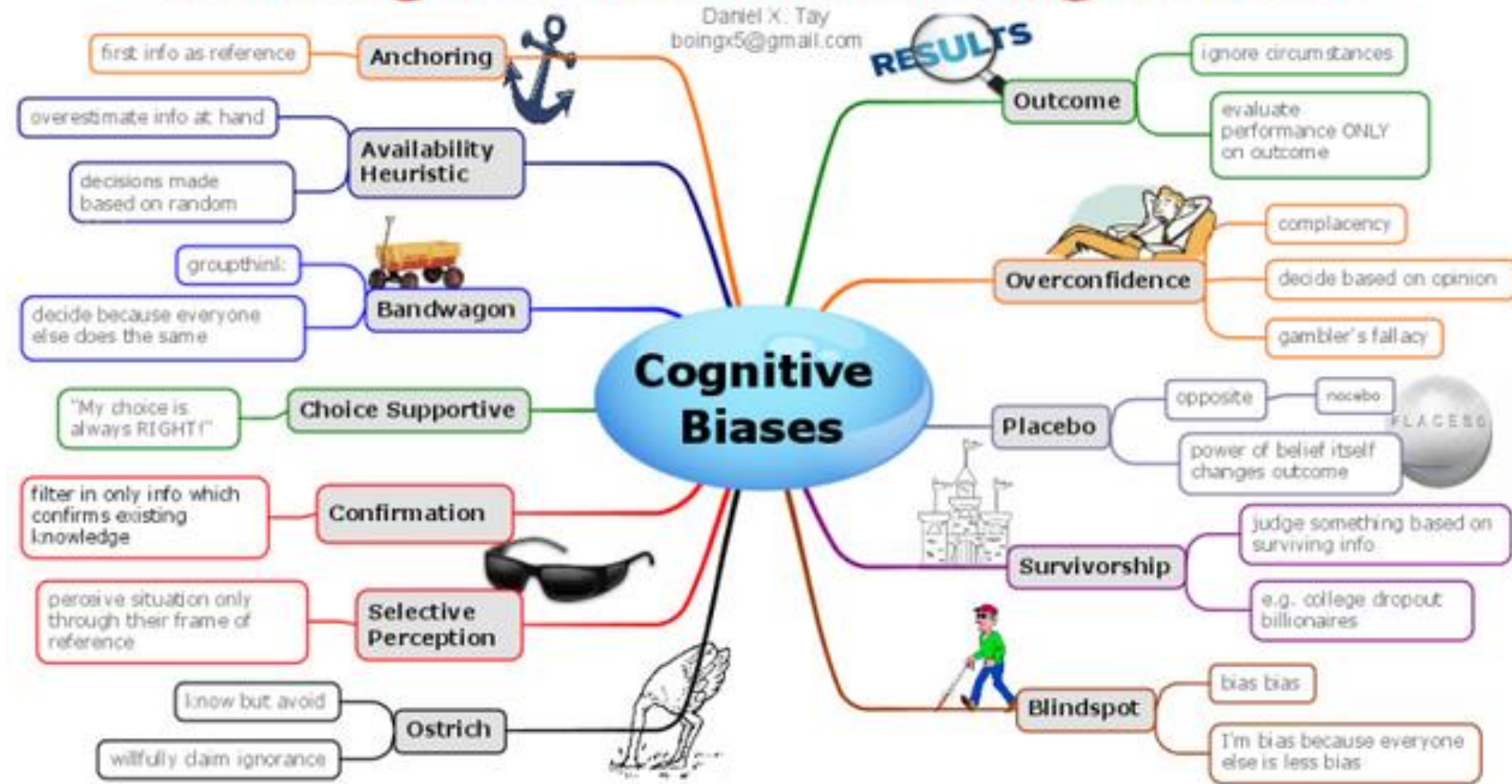


<https://www.aclu.org/blog/privacy-technology/surveillance-technologies/amazons-face-recognition-falsely-matched-28>

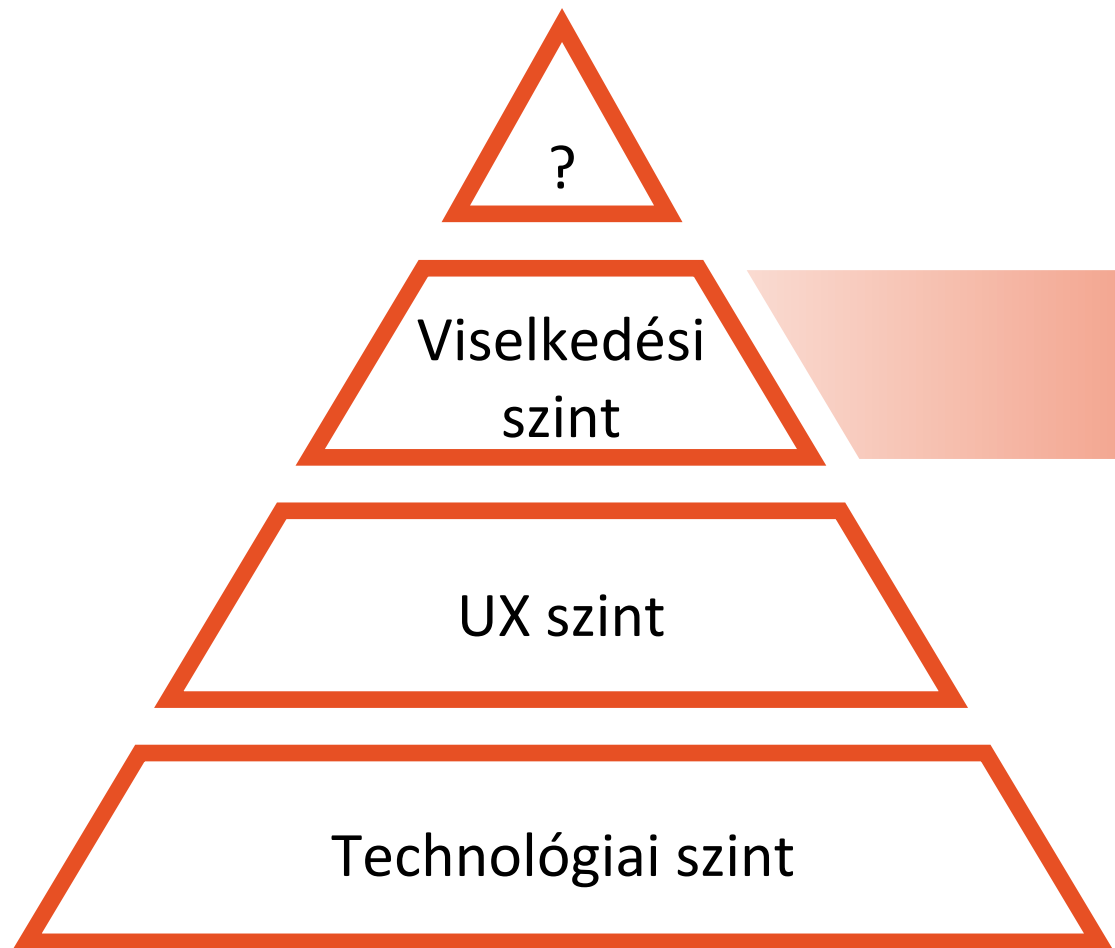
<https://www.media.mit.edu/projects/gender-shades/>

Common Cognitive Biases when Making Decisions and Judgements

Szerencsére
mi
bölcsebbek
vagyunk!



A gépi alkalmazás fejlődési piramisa



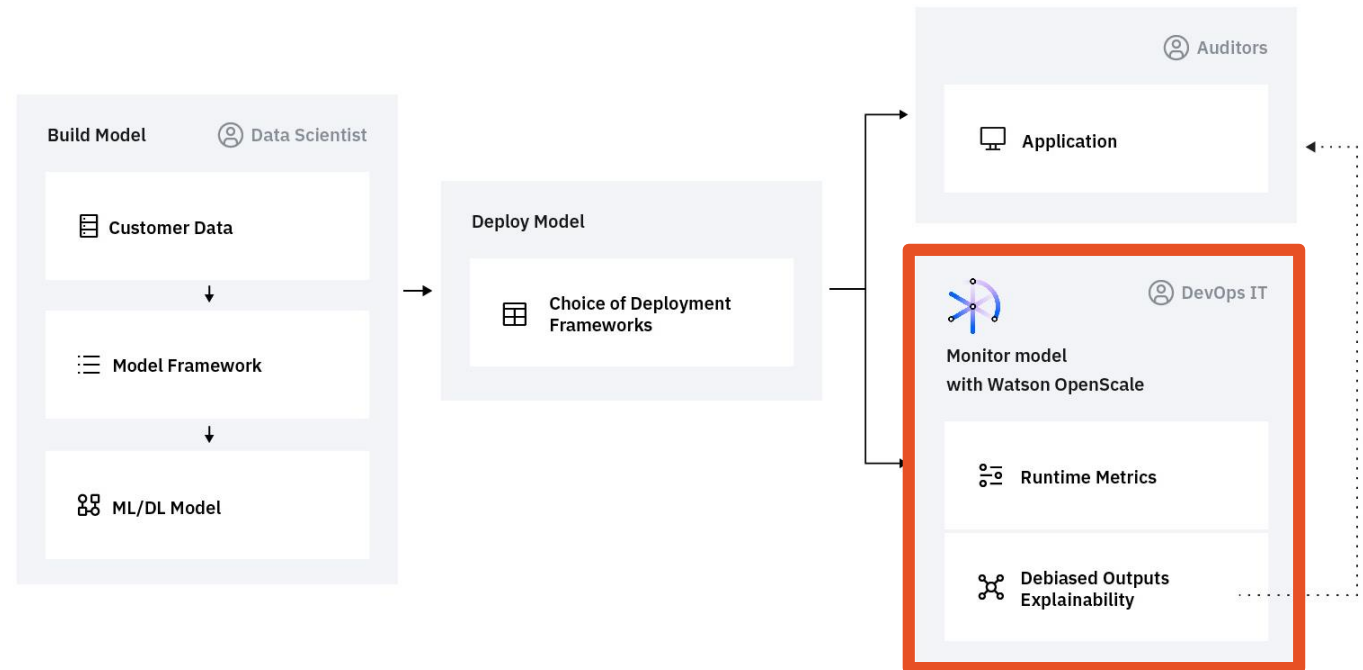
Legyen etikus!

A gépet használjuk, legyen használható!

Összerakjuk, működik.

Kövessük nyomon az MI döntéseit!

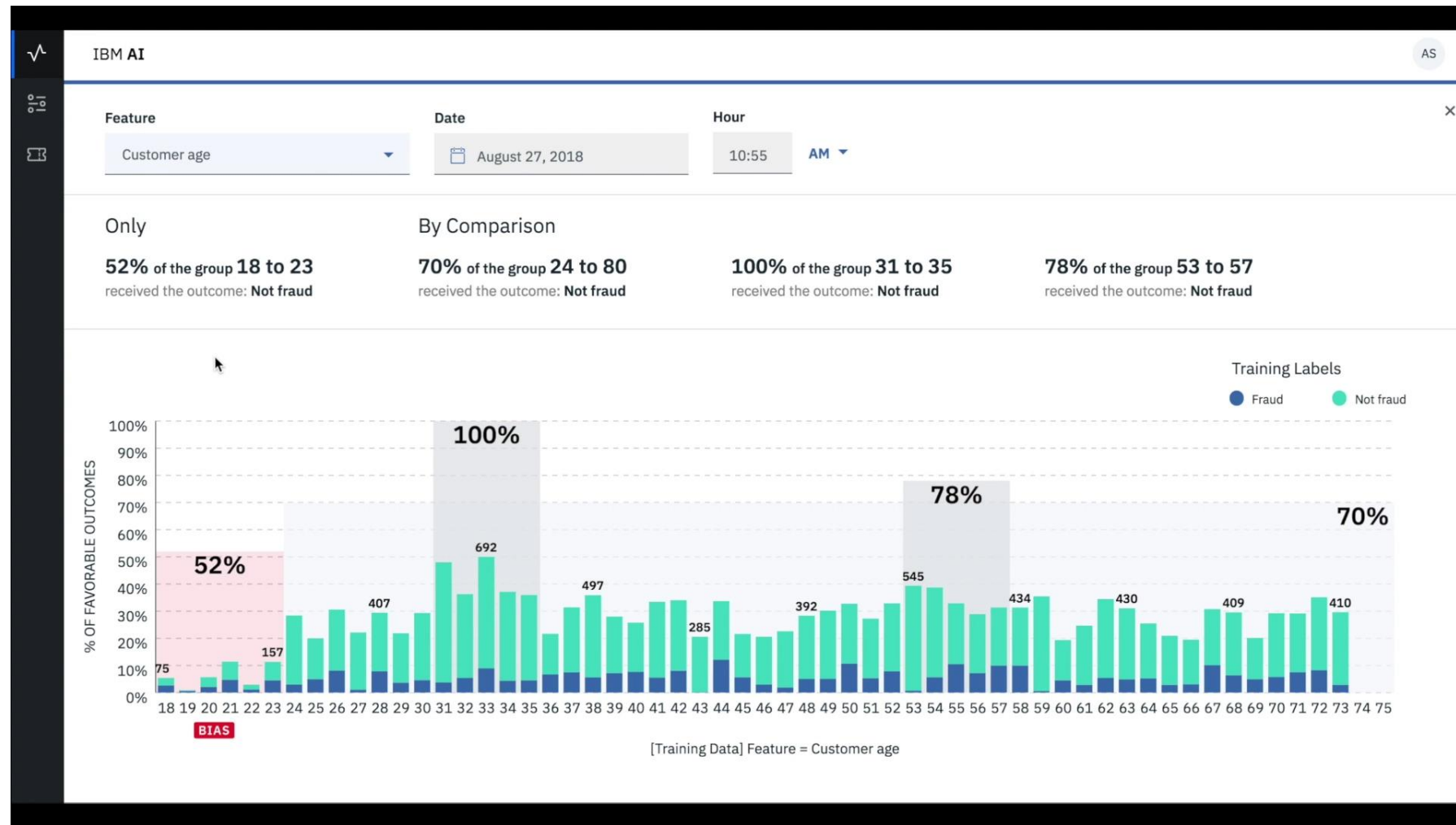
- **IBM Watson OpenScale**
- Megkönnyíti az AI életciklus bizonyos lépéseit
- Elemezhető, hogy mikor és miért *előítéletes* az AI
- Direkt felület az azonnali beavatkozáshoz
- Tranzakció szintű granularitás
- NeuNets: automatikus modelltanítás, -építés
- **OpenScale**: könnyű modellintegráció



Demo



Demo



Összefoglalva

Működés – Az hibázik, aki dolgozik. A gép is egyre emberibben hibázik.

Viselkedés – Az MI-t emberi környezetben, emberi szempontok szerint ítéljük meg.

Felelősség – Nincs gépi etika. A felelősség is az emberé.

Kérdések?

gykormendi@clementine.hu