

Digital Humanism

An Introduction

Hannes Werthner
George Metakides
Summer School 2024
Vienna, 2/9/24

What a Nonsense

Reaction of the audience at the ENTER conference in Innsbruck, 1994.

- Beat Schmid (University of St. Gallen) on the Internet and electronic markets
- Larry Press (UCLA) on digital agents
- Florian Brody on electronic money

Today





- We live in a digital / physical world
- Public good: participation - nearly - mandatory (Austrian politician: **right to analog life**)
- Internet / Web global infrastructure – world's Operating System
- Digital transformation – complex technical social economic process

And now

Help other travelers choose the best!

Using generative AI, we've crafted a possible review based on your inputs:

Amazing Stay!

Our stay at the Oriental Jade Hotel was absolutely amazing! From the moment we arrived, we were impressed by the value for money, convenient location, exceptional cleanliness, friendly staff, and excellent facilities. The hotel exceeded our expectations in every way. The rooms were comfortable and well-maintained, with all the necessary amenities. The hotel's central location made it easy for us to explore the surrounding area and visit nearby attractions. The staff were incredibly helpful and went above and beyond to ensure our stay was enjoyable. The facilities provided were top-notch, offering everything we needed for a comfortable and convenient stay. We highly recommend the Oriental Jade Hotel to anyone looking for a fantastic experience. We can't wait to return!

We hope the suggestion above inspired you. Now, you can write the review you'd like to post:

Surprising Performance

GPT-4 very good in several tests

Scientists are surprised, e.g., **Geoffrey Hinton**

But we are participants in a worldwide unannounced experiment

Promise: intelligent, flexible and easily usable tools that are complementary to human decision-making, technical work and creative tasks

- Noy and Zhang (2023): workers with lower productivity with access to ChatGPT improve their performance on writing tasks
- Brynjolfsson et al. (2023): improvement in working labor productivity when ChatGPT is used in a complementary form
- Issue: augment humans or automate

AI as GPT (General Purpose Technology)

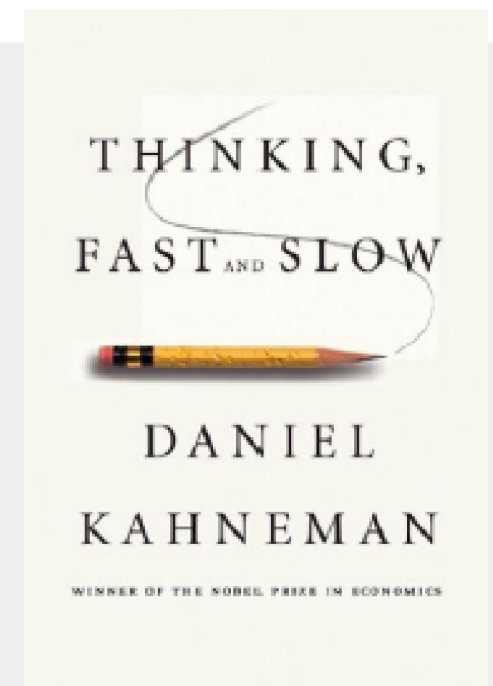
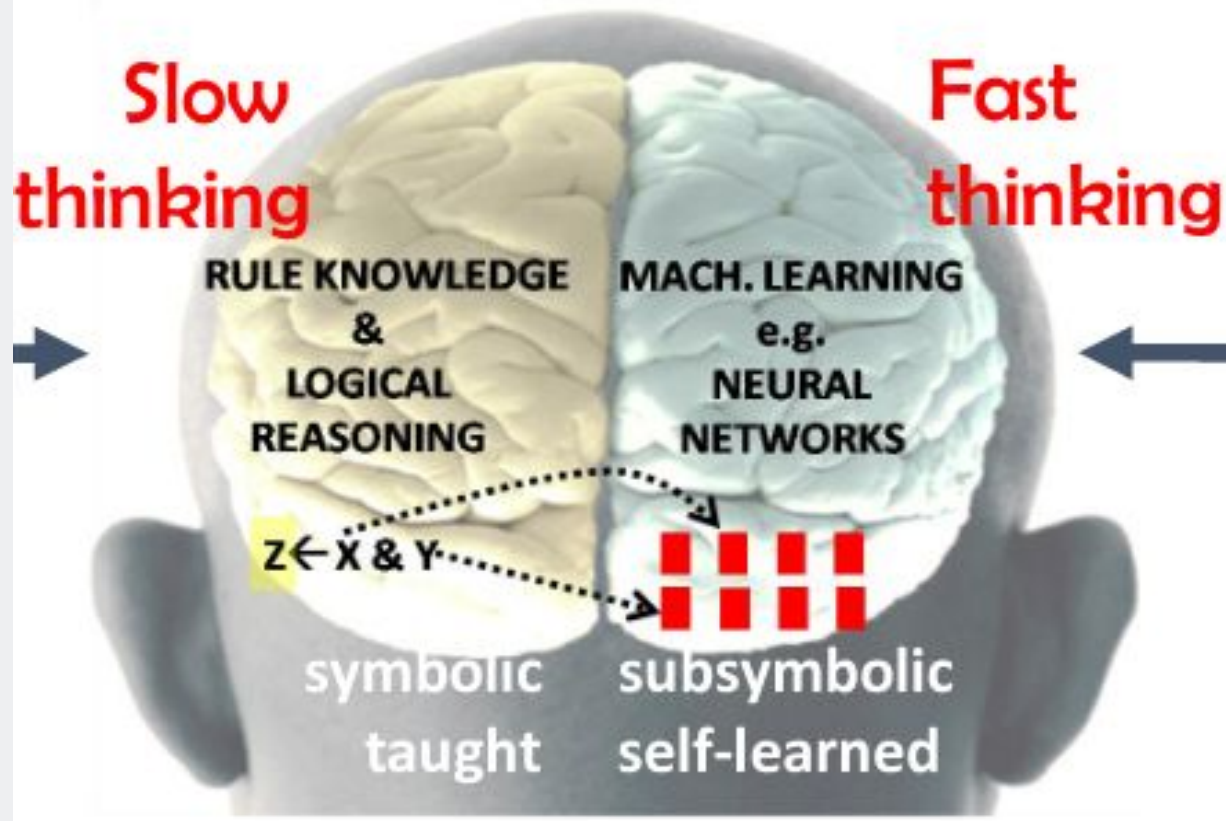
Forecast: annual economic growth of up to 1.2 % (PwC, McKinsey) (?)

AI Discussion

- Between **dystopia** and **utopia**
- AI pessimism in the West, more optimistic view in the countries of the South
- Critique on **anthropomorphic** notion Artificial Intelligence (J. Lanier)
- **Business as usual**
 - Just next step. History of hypes in CS, e.g., expert systems, blockchains - and we see signs of that
- **Amazement and enthusiasm**
 - Something unexpected happened (even for developers)
- **Existential threats warnings**
 - AGI with threat of human extinction
 - Remark: Singularity and separation of mind and body, immortality of the mind (“brain uploading”) - religious concepts
- **Hype warners** (Whittaker, Gebru): Don't fuel the hype, it's about other things (surveillance, ..)
- **Cautious ones**: potential good uses, but look at social disruption, democracy.....
- **How to answer**: continue as “usual”, moratorium, regulation, public alternatives

AI: Hard to Define

- Complex SW systems, or **brand** for some related technologies? (Thomas Haigh)
- Simulation of human behavior, machines showing “intelligent” behavior
 - „Automation of Thinking” / materialized in code
 - What is intelligence, all in the brain - or embodied (and socially/culturally embedded)
- **Symbolic** vs. **Data driven approach** -> we experience a paradigm shift
- From a-priori defined rules to posteriori learned patterns
 - Current success based on: technology stack, accessible huge amount of data, new methods
- 1956 Dartmouth Conference with *Artificial Intelligence* vs *Cybernetics*
- **Major critical issues:**
 - A story, not a truth machine - hallucination
 - Black box
 - Bias
 - Misinformation and fake news, threats for democratic processes,



Edward A. Lee

Similarity to human way of thinking: slow & conscious or fast & automatic

System 1: Intuitive, fast, inexplicable decision-making

System 2: rational decision process

Neural networks like system 1, logic-based approach like system 2

Don't we Expect Too Much?

- Human beings are also limited
- Rational process: step-by-step thinking according to the rules of logic.
- We can only process a few steps and very limited data - Herbert Simon (1981)
- **Bounded rationality**
- We often find explanations in retrospect
- Overestimate frequencies of rarer events – framing
- Not all options can be grasped



- GPT 4 has 175 billions parameters – that's the “**knowledge**”, not the software
 - No “model”: example Google weather forecast
 - New computation paradigm?
- Business model still open
- AI economic landscape – highly concentrated (high costs)
 - No end of Big Tech – on the contrary
 - 50% of datasets from 12 institutions (10:1:1)
 - In 2022 32 industry models, 3 from academia (HAI 2023)
 - Few countries and digital divide with colonial division of labor
 - Scientific research is equally distributed between China, Europe, the USA and the ROW
- Universities are lagging behind
- Science without model ?
- Science without Universities ?

New Relationship?

- AI story machines lead to an interesting reversal in our relationship with the “computer”
- Computers solve problems for humans with 100% certainty,
- We humans have delegated our problems - **computer is the final authority**
- These “new” systems “invent” results, truth is not guaranteed.
- It is up to us to check the truth
- Did we regain control, have we been re-empowered?
- Interestingly, this is now experienced as a loss of trust, not as a re-appropriation

- From first study programs in 70s to *Latin of today* - A. Loprieno (Egyptologists, Chair ASC)
- Computer Science / Informatics foundational: methods, paradigm, artefacts
- General Purpose Technology (as AI)
- Algorithmic solution of problems **AND** understanding of human thinking:
 - *Problem solving*, e.g., simulation IAS Computer Von Neumann, weather forecast
 - *Understanding*, e.g., Turing with his Test / defining intelligence
- Metamorphosis from a computer to pervasive global machine (sometimes invisible)
global, connected, pervasive, adaptive / learning
- Technical view: Systems consist of a stack of different pieces (hard and software)
 - Increasingly delegation to software -> virtualization, provides flexibility (e.g., car)
- **Everything touched by SW becomes a computer - connected & global**
- Digital divide

Internet / Web: more than Technology

- Economic transformation (companies, markets)
- Social „expansion“
- Law and politics
- Physical spaces
-

Vilnius



Augsburg



- Internet / Web also rooted in the US counterculture (see also Vardi, 2018)
 - Free information sharing - > huge amount of information and rise of search engines
- To sustain (monetize): advertising based business model, extended by
 - **Personalization**
 - **Recommendation** - User change behavior (Russell 2021)
 - **Individual freedom is a fiction:** what I see, what I get, what I do is defined by distance measure of a similarity matrix
- Pay with our data -> surveillance (Zuboff, 2019)
- From **citizen** to **consumer** (Stanger, 2020)
- We are **user**, **product** and **producer** (nearly an economic perpetuum mobile)

Economic Observation

- From *IT supports business* to *IT is the business*
- Focus on automation and efficiency, not resilience (except a famous example)
 - *Informatisation / Deconstruction of value chains* – also basis for globalisation
 - *Outsourcing* (e-commerce, self service)
 - *Privatization*: public data are “owned” by IT companies, e.g., LLMs; Google and Wikipedia (Vincent and Hecht 2021).
- With simultaneous deregulation, the web developed into **winners take it all**
- In contrast to promises of IT since 80s increased income inequality in practically all major advanced economies
- Shared responsibility: neoliberal ideology; massive decrease in regulations, ongoing concentration (Acemoglu and Restrepo, 2019).
 - Technology facilitated this process

Web Business Landscape - Platforms

- Google 1998
 - YouTube 2005
 - Alibaba 1999
 - Amazon 1994
 - Tencent 1998
 - Facebook 2004
 - LinkedIn 2003
 - Uber 2009
 - Airbnb 2008
 - Instagram 2010
-
- Digital platforms - market shares of up to 50-90%
 - Basis: user network and information from/about users and markets
 - Virtualization of the business and industries → scalability
 - Newcomers

Market Capitalization

Jun 2024, Mio USD

Microsoft	3,322,000
Apple	3.230,000
Nvidia	3.182,000
Alphabet	2267,000
Amazon	2.011,000
Meta	1.279,000
tSMC	901.690 Berkshire
Hathaway	879,670
Eli Lilly	815,210
Broadcom	747,360

March 2011, Mio USD

Exxon Mobil	417,166
PetroChina	326,199
Apple	321,072
ICBC	251,078
Petrobras	247,417
BHP Billiton	247,079
China Construction Bank	232,608
Royal Dutch Shell	226,128
Chevron Corporation	215,780
Microsoft	213,336

List of top publicly traded companies (Financial Times Global 500)

Two Sides

- IT keeps the system running, it is successful, with high potential
- Rich free information - essential "public" services / good
- Important for participation and organization of civil society (Sardines, Hong Kong, Pakistan Imran Khan, campaigns in Africa, e.g., South Africa, and Zimbabwe, or also use by Amazon workers)
- Indispensable for solving fundamental problems



Two Sides

Critical issues = **mutually dependent**

- Concentration and monopolies in the Web
- Sovereignty – states as well as individuals
- Echo chambers and fake news – with its impact on political discourse
- AI and decision making (role of humans)
- Privacy and surveillance
- Robots and warfare / weaponization of AI
- Work and automation
- Democracy backsliding
- Sustainability and environment

□ The system is failing – Tim Berners-Lee, 2018

□ AI also example of these two sides

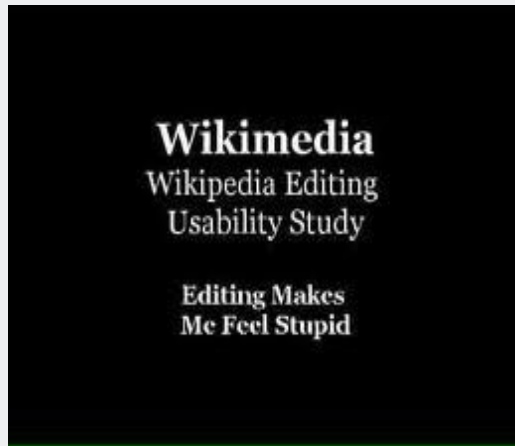
Unprecedented Concentration of Power



The WWW and Dreams of a new Enlightenment

A marketplace of ideas

Online deliberation/decision making/rapid mobilization



- Icelanders 'like' their crowdsourced constitution
- A year after Iceland asked its citizens to help write its new constitution by making suggestions via Facebook, voters decide this new draft should be the basis for the upcoming constitution.



• Crowdsourcing/Crowdfunding

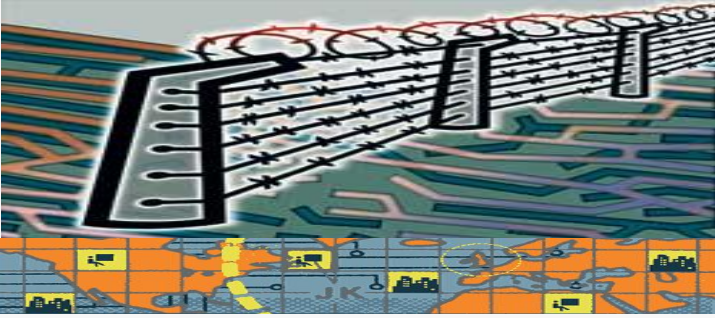
- http://www.ted.com/talks/a_choir_as_big_as_the_internet.html

- www.fansnextdoor.com

New Art ideas

- www.kickstarter.com

Innovative entrepreneurship



Early Warnings: Tim Berners Lee *Scientific American*, 2010

- Privacy (snooping , profiling ,tracking)
- Closed «silos» (Social Media)
- «Walled Gardens» (smart phone apps)
- Monopolies (private clouds ,”trapped in a store”)

From Utopia to Dystopia (TBL, 2019, Turing Award acceptance speech)

Confluence of Economic and Political Power

- Corporate value power (incl. lobbying)
- Power to influence/manipulate public opinion/elections
- Surveillance power
- Nudging, herding, conditioning
- Predictable humans can be reliably presented with ads they will click on
- Political extremists are more predictable than moderates
- **Pernicious polarization** (when a society becomes defined almost exclusively by political allegiances).
- A mob in the agora instead of a marketplace of ideas
- Impact on **democracy**, social cohesion and more!

Democracy and participation

- <https://www.idea.int/>

- Participation in European national elections in **1991** 77 %
2023 65 %

- In European South drop is bigger **1991** 83 %
2023 45 %

- Resentment and lack of trust in politics

- Autocratic/populist propaganda/disinformation

- https://v-dem.net/documents/44/v-dem_dr2024_highres.pdf

- % of world population living in autocracies **2003** 50%
2023 71%

- Need for equality of opportunity and equity (of socioeconomic prerequisites)

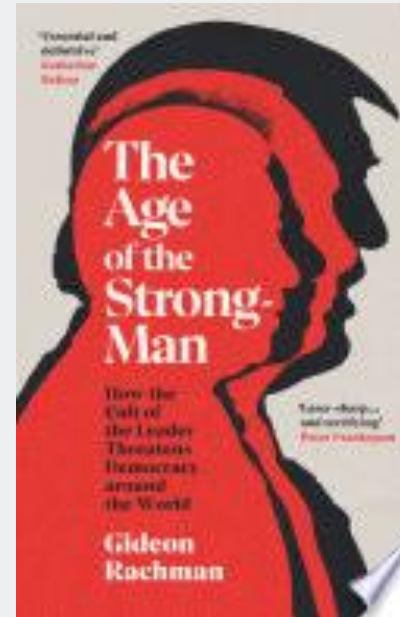
Democracy and the Young I

- Gen Z : Born between 1995 and 2012
- UK 2023 study:
- Democracy is the best system 48%
- Military dictatorship could be better 47%
- Similar results in most European countries with far-right parties getting a bigger share of Gen Z votes than from higher age groups

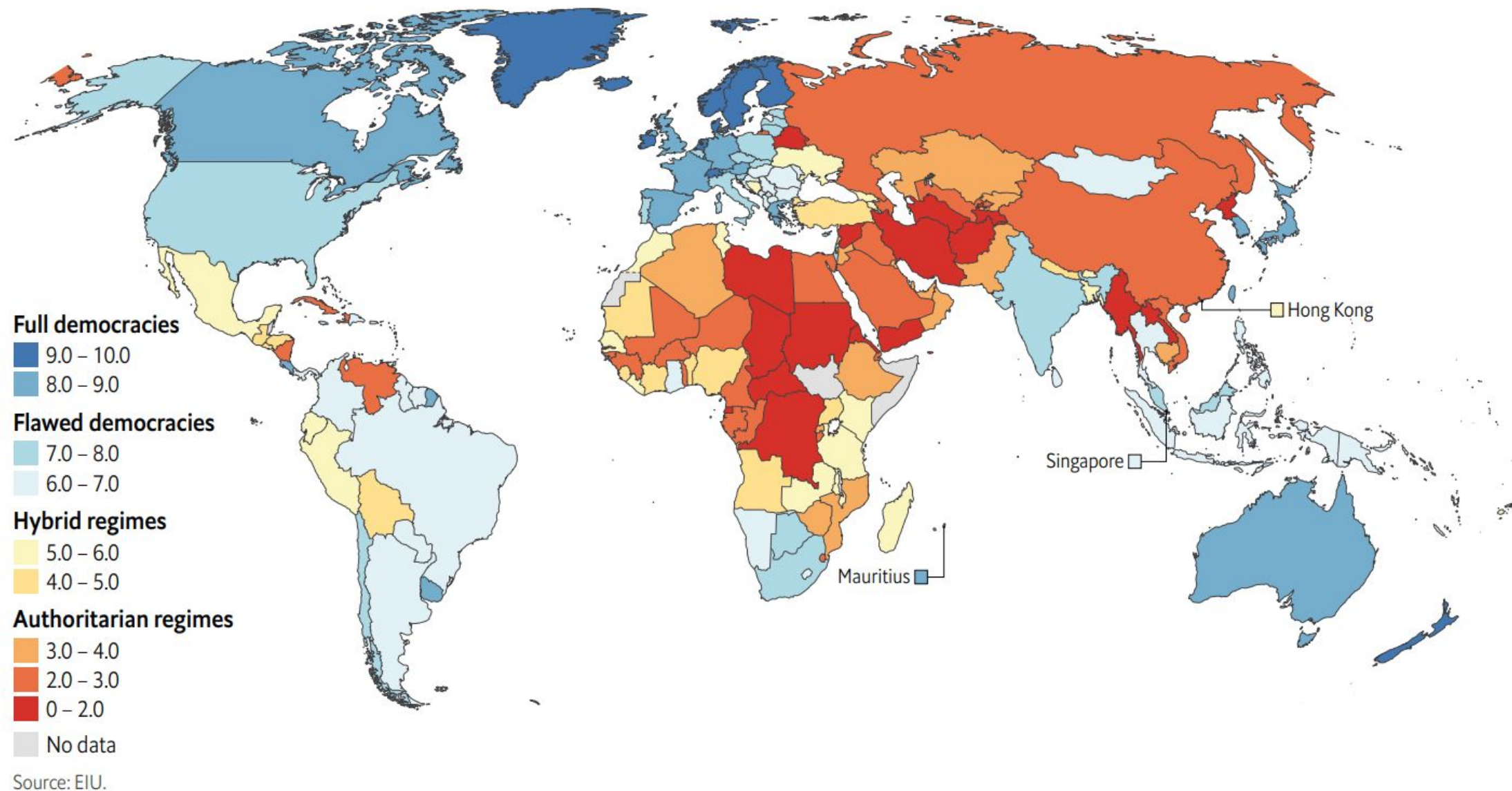
Democracy and the Young II

- Increasing percentages among the young agree with:
- “In a democracy nothing gets done, we need perhaps less democracy and more effectiveness”
- “a strong leader who does not have to deal with parliaments and elections”

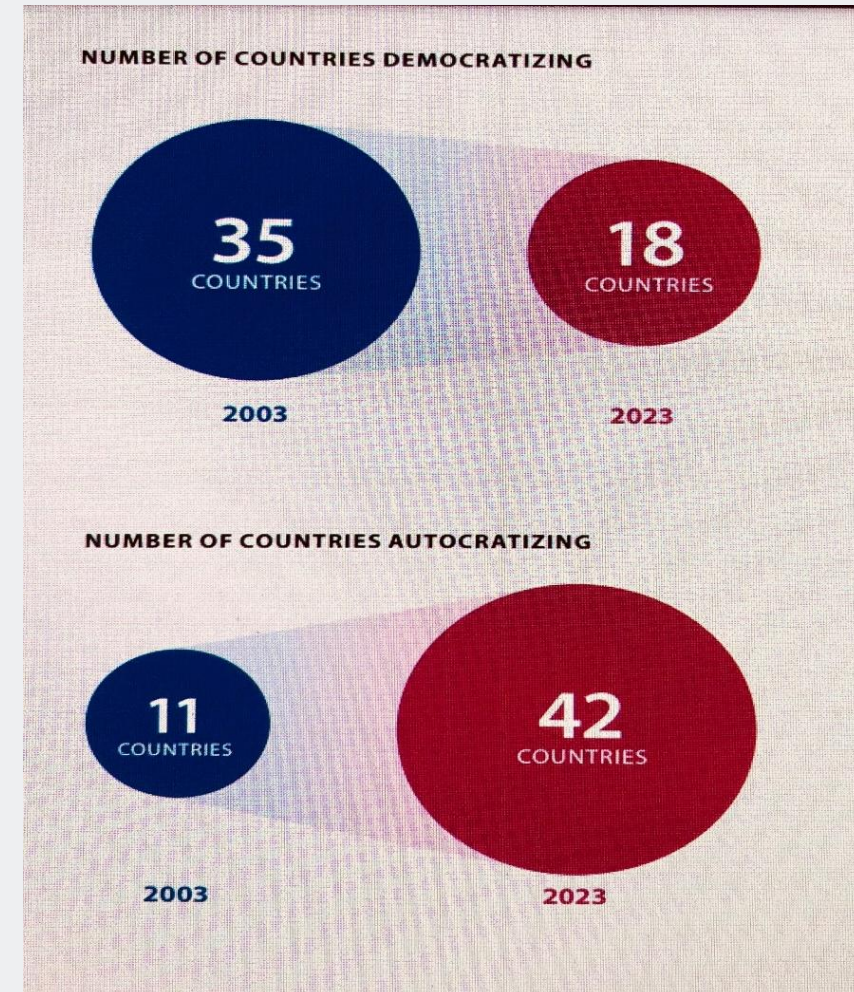
Worldwide polarization, extremism, attraction to “strong” leaders



Democracy Index 2023, global map by regime type



Democracy backsliding worldwide



Democracy will not live or die by Digital alone

Most cited (and intertwined) threats:

- Economic inequalities (especially after 2008 crisis)
- Globalization effects (winners and losers)
- Immigration (especially after 2015 crisis)
- Internal degeneration (constitutional defects, plutocracy/ kleptocracy, Orbanism)
- The fallacy of inevitability (that somehow globalization , capitalism and trade will preserve and spread democracy)
- Big Tech and Social media. Concentration of Power
- Cyberthreats to democratic processes (internal and external) AI-powered Misinformation/disinformation
- As such content proliferates people may be manipulated or simply end up not believing anything (mistrusting ALL sources, not able to agree on ANY facts)
- This would erode a core tenet of a democratic system which is trust
- “Disinformation is the 5th horseman of the Apocalypse”
- **Two sides again**
- Most see digital connectivity/social media making people easier to manipulate – but also more informed (Pew 2023)
- Is the “bad side” inevitable? (short answer: NO)

2024 elections

- Euroelections
- Russians used cyberattacks (e.g. to Von der Leyen's campaign page [ursula2024.eu](https://www.ursula2024.eu)) and misinformation but relied more heavily on old-fashioned spying and buying /bribing (e.g. MPs)
- Examples: Attacks by APT 28 which is linked to the GRU Russian Secret Services
- Russia bypassed bans on Russia Today and Sputnik via proxies like (recently) Voice of Europe (in Chechia)
- EU responded with EUvsDisinfo as DSA gets results very slowly
- France established since 2017 Viginum against Cyber-interventions
- Uncovered pro-Russian trolls and 193 web-pages spreading disinfo on Ukraine
- In Germany 1 mio tweets from Russian false accounts in a month criticising policy vs Ukraine
- China does same with disinfo on Taiwan and Uigurs.APT31 is controlled by China. Called Glucksman a CIA Trojan horse in Europe
- Role of Gen AI in all this: Hard to measure ! Effort needed

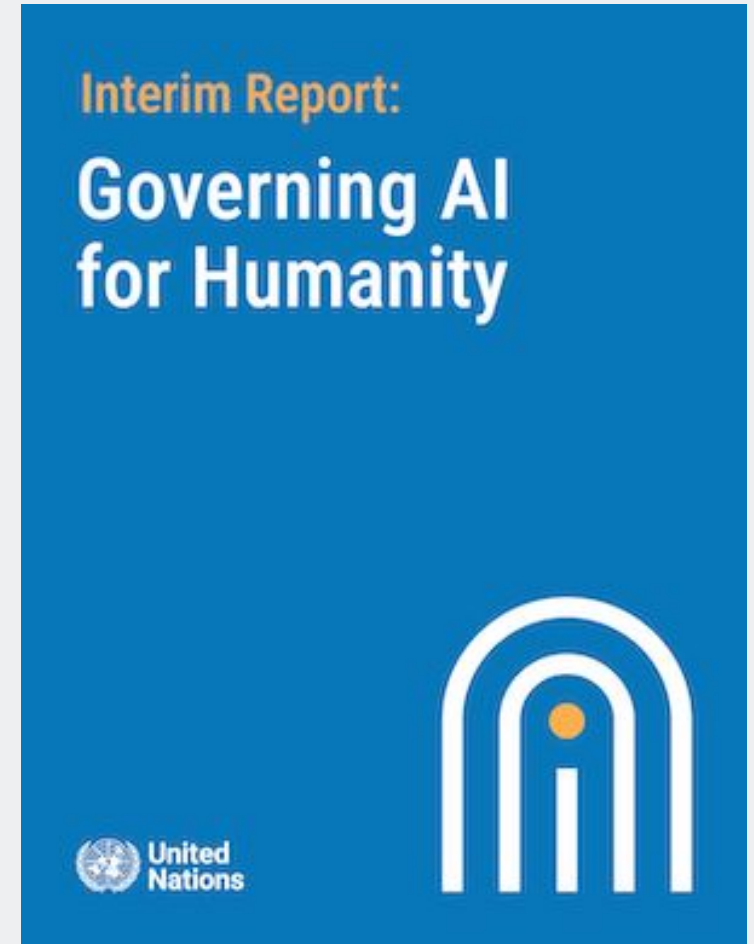
The Need for Regulation and new research directions:

- A new impetus for regulation including anti-trust
 - Current Big Tech platforms were developed in the absence of any form of regulation .
 - EU: DMA/DSA, Data Act and **AI Act** (ex-ante precautionary principle)
 - US: Anti –Trust (Remedies?)
 - New AI R&D directions aimed at the collective good
 - This will not happen if control of the direction of R&D of transformational technologies like generative AI is left in the hands of the few and powerful
 - Today only 10% carried out in Universities
 - Big Tech argues for self-regulation because “bureaucrats” do not understand the technology and they, the experts, do.
 - Two responses:
 - Regulators regulate Aviation Safety, Pharmaceuticals etc. for decades!
 - AI is not just about technology as it implicates human rights and democracy
- Big Tech is not an expert in democracy and their track record in human rights is abominable

A global race for regulation !

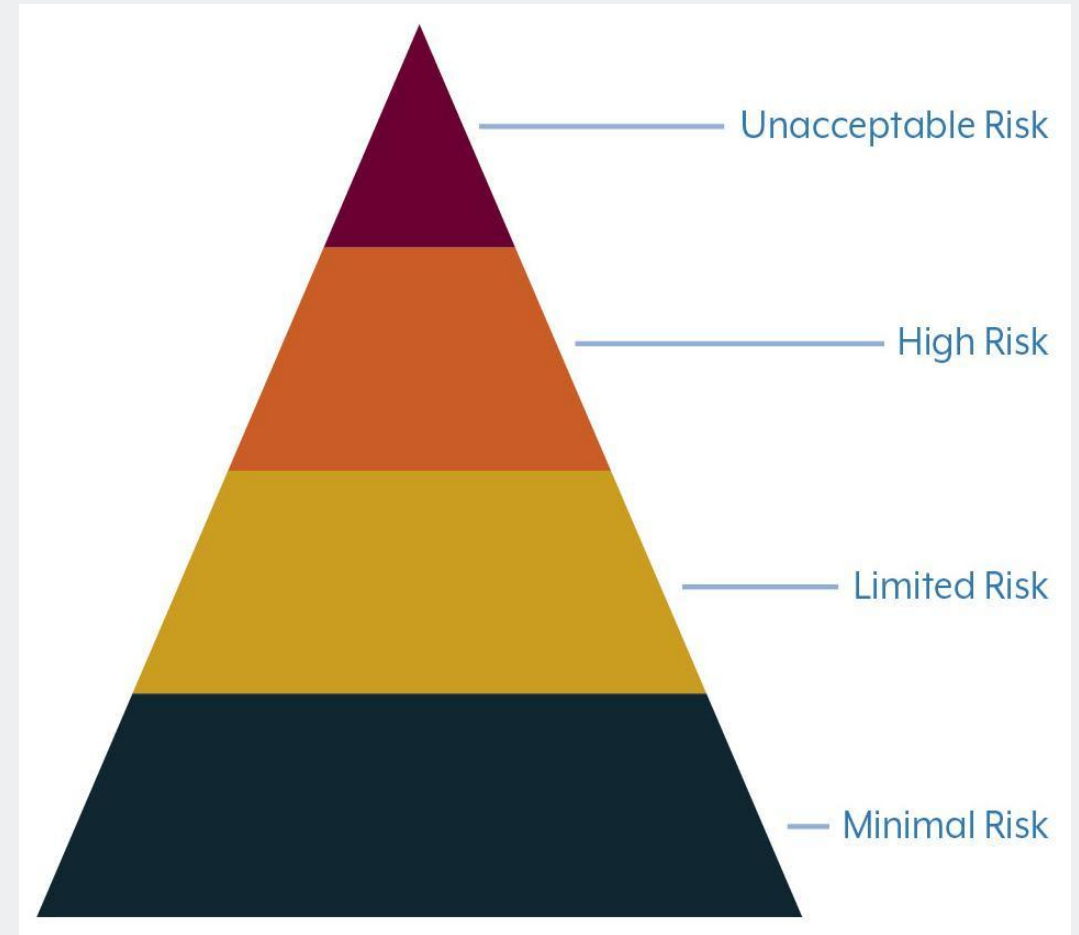
Source: UN Advisory Body for AI (Report 21 /12 /23)

- Biden's executive order
- China's measures for AI services
- UK AI Safety summit
- Global Partnership for AI
- OECD AI principles
- G7 AI principles (Hiroshima process)
-
- **EU AI Act** (The only one with enforcement power besides China)



AI Act: Categories of Risk

- **Not Acceptable** (social scoring, biometric surveillance, emotion detection...)
- **Ban**
- **High Risk** (critical infrastructure, health products/services, hiring, justice, loan approvals ...)
- **Strict conditions of transparency and human control** διαφάνειας. Especially for «large» GPAI
- **Limited Risk** (deep fakes, chatbots...)
- **User must be informed about use of AI**
- **Minimal Risk** (spam filters, search,inventories, games)

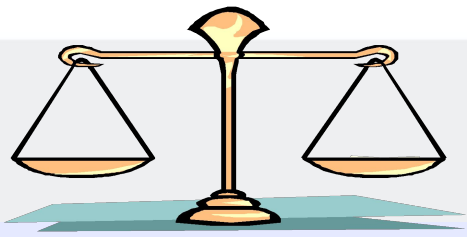


AI Act :The Timetable and more

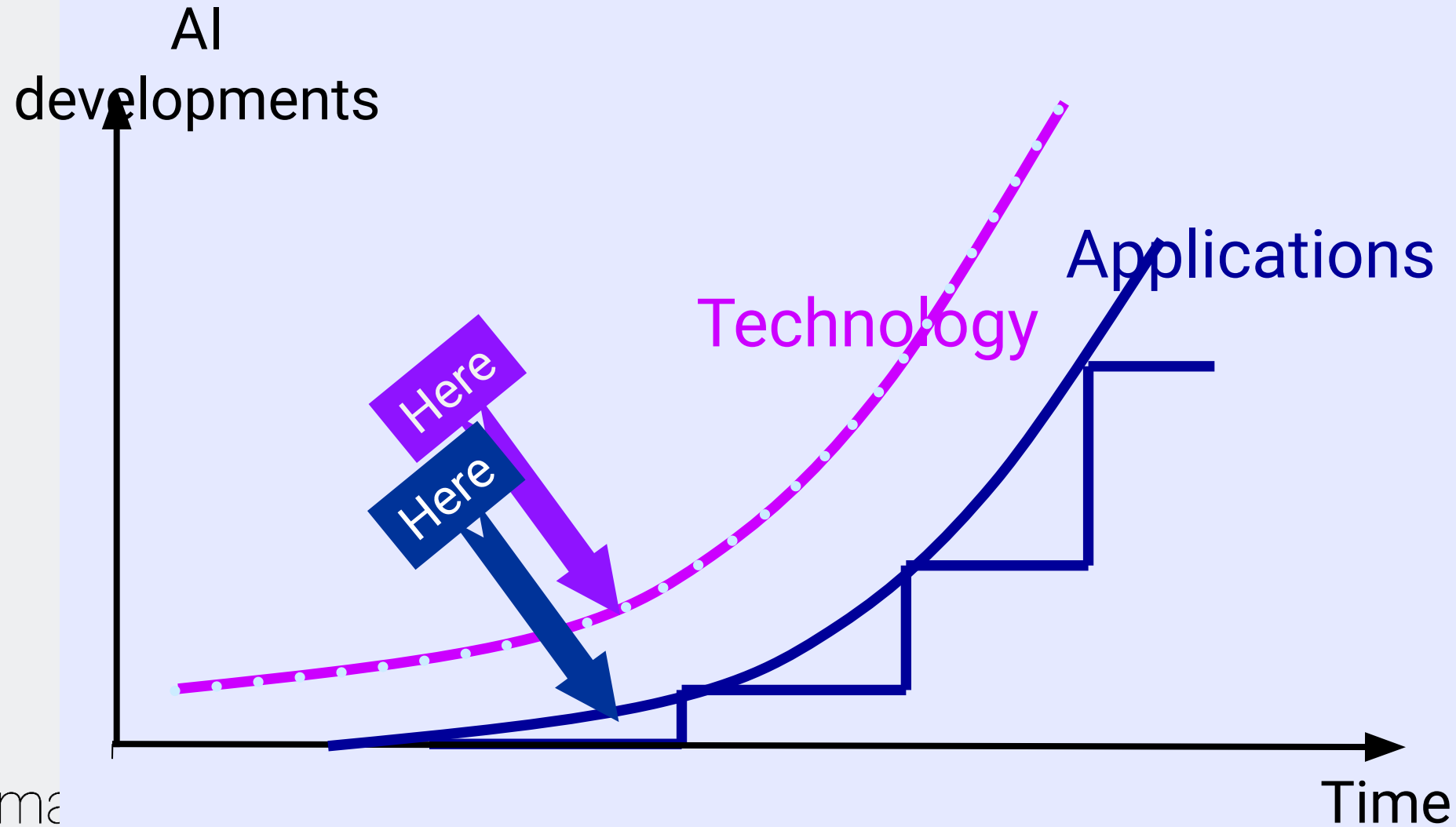
- The final text was published in the Official Journal of the European Union on July 12 2024 and came into effect on August 1 2024
- Full enforcement two years hence (August 2 2026)
- Rules for unacceptable risk : February 2 2025
- Rules for GPAI : August 2 2025
- Civil Society VS corporate lobby (digium open letter and more)
- During next two years : harmonized standards
- “AI Office” has been created in Brussels for monitoring/enforcement
- EU outreach to Canada,India, S.Corea, Japan, Philippines et al aiming at a ... Brussels effect

Digital Humanism : Questions begging for answers

- How can we best deal with the use of AI by malicious actors to manipulate the beliefs and understanding of citizens?
- How can we ensure that the engagement of the public with elected representatives—a cornerstone of democracy—is not drowned out by AI-generated noise?
- How can we help everyone, including our scientific, political, industrial, and educational leaders, develop the **critical thought** needed to identify AI-generated misinformation, impersonation, and manipulation?
- How can we promote the development of a new generation of AI tools designed for the benefit of society and for the support of democratic processes?
- Most importantly : **Do not give up , it is still early days !**



It is still early days!



HOMO FABER

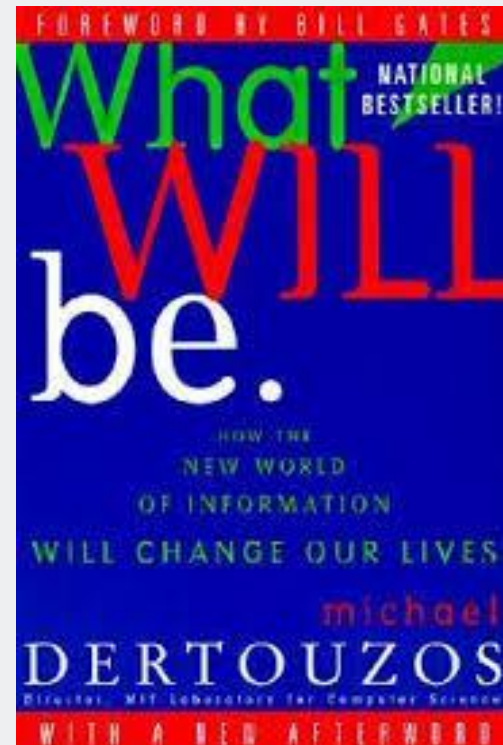
- AI is NOT a force of nature
- It is (like the Web) a human artifact
- Both beneficial and harmful impacts (e.g. Web)
- Beneficial applications to be designed and promoted
- Potential harmful applications NOT to be treated like volcanoes or earthquakes
- “Negative Externalities” of social media (like pollution as a byproduct of profit-seeking businesses)
- Do we have to accept them as the price to pay for living in a digitally connected world?
- New regulatory impetus globally
- Techno-optimism: It depends!

The goal of technology is to serve humanity

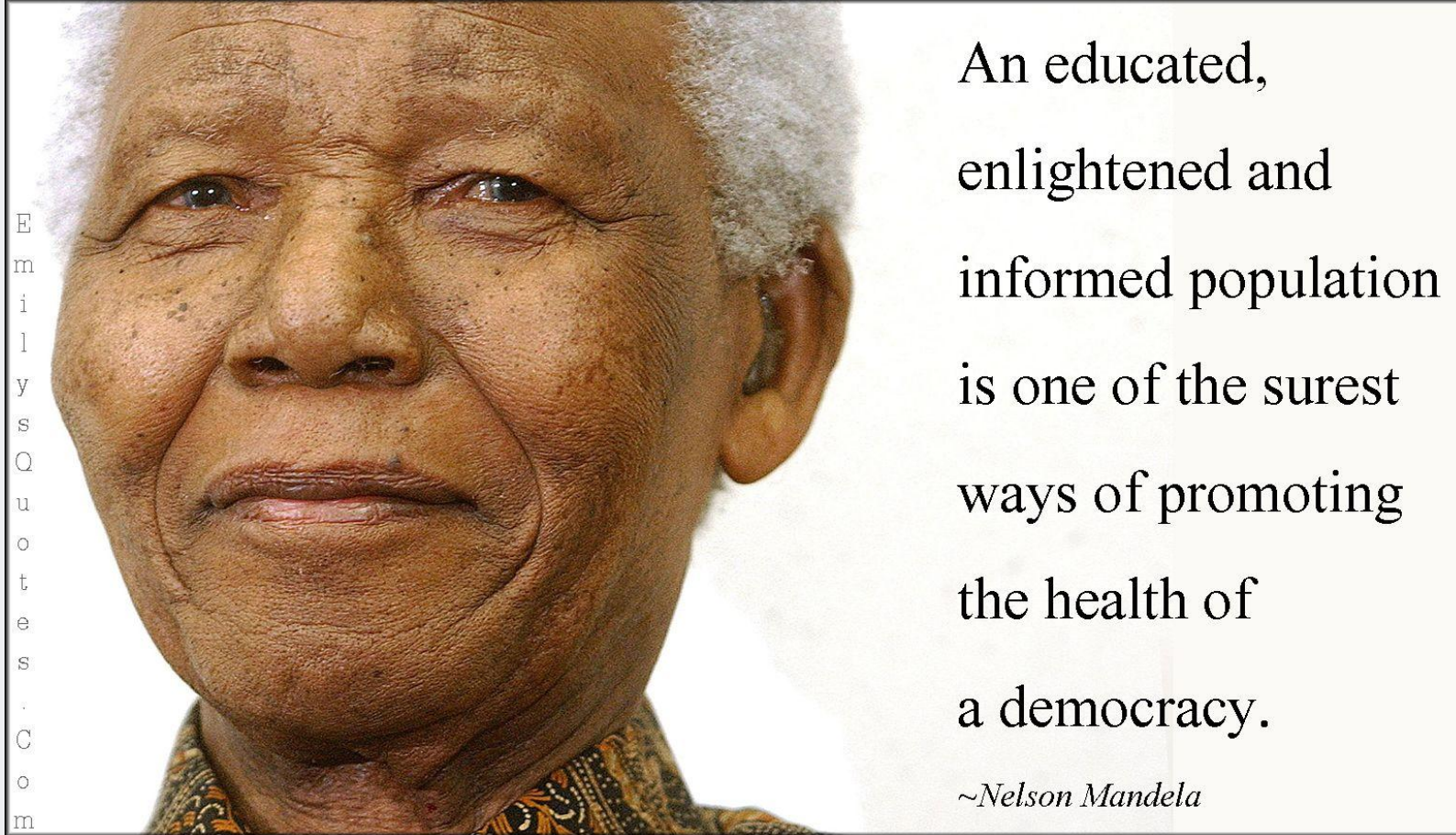


- ***"The measure of success for AI applications is the value they create for human lives"***
- They should be designed to enable people to understand AI systems successfully, participate in their use, and build their trust. Public policies should help ease society's adaptation to AI applications, extend their benefits, and mitigate their failures

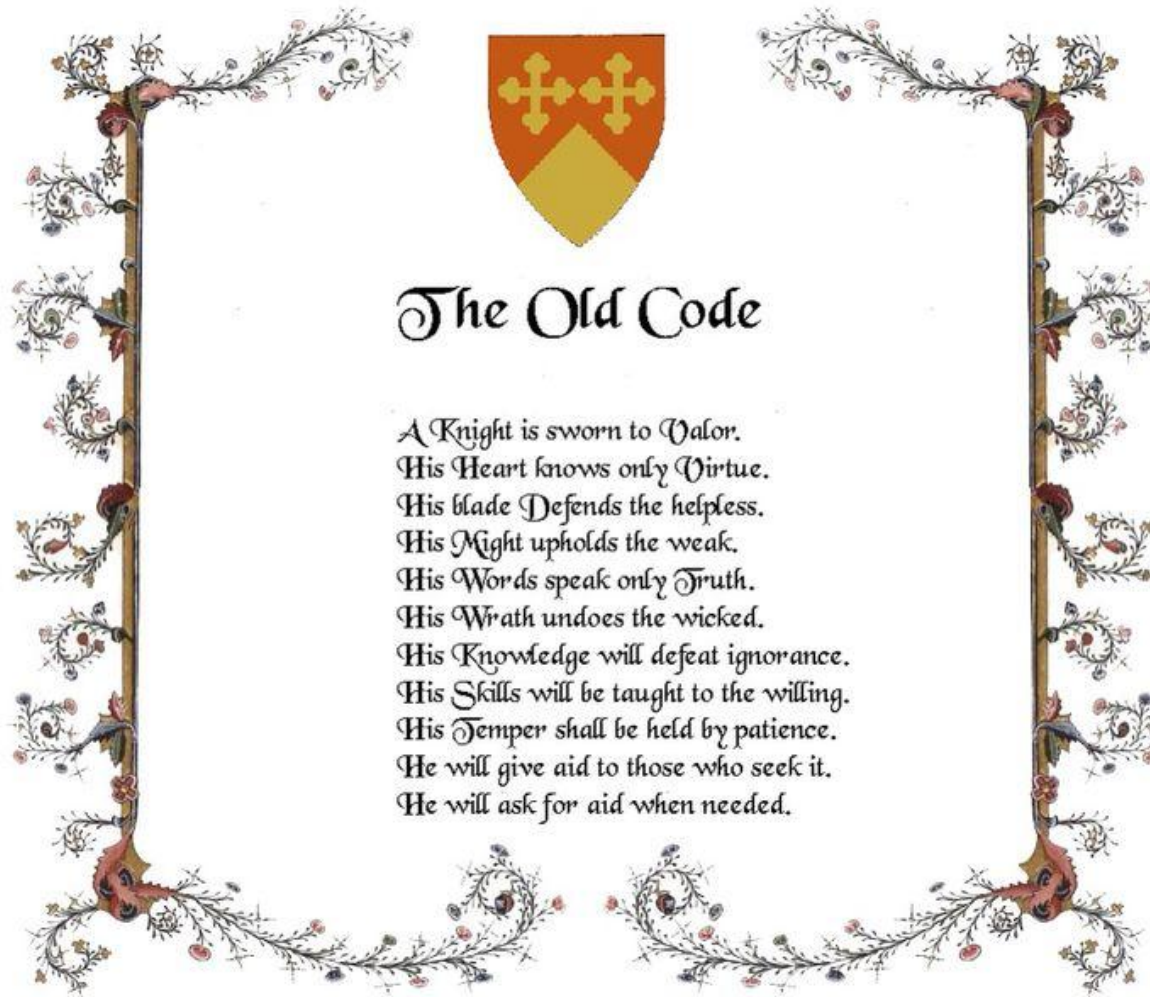
Humies and Techies: Bridging the gap



In today's classrooms sit tomorrow's societies



A new Code for “Digital Chivalry” ?



- Uphold weak/defend the helpless:
Collective good
- Share skills and knowledge
- Truth and Trust
- Undo the wicked?

A New Digital Ecosystem is Emerging



- AI advances (not the A and Ω !)
- Democratizing the digital ecosystem
- Regulation/ Research/ Public Investment
- Nurturing Critical Thinking Skills
- Digital Humanism as a guide
- Summer Schools like this now more important than ever!

Digital Humanism Workshop in 2019

- 100 participants - worldwide – informatics, philosophy, history, anthropology, law, economics, political science, mathematics, sociology
 - Started from international Advisory board of Informatics / TU Wien



Popper (1969) "Moral Responsibility of the Scientist"

Digital Humanism

Describe, analyze and, above all, try to influence the complex interplay between IT and humanity - for a society that respects human rights and environment

- Technical development not automatic and "God-given"
- Technology according to our values and needs
 - We are the designers & developers
- We humans (all) have the freedom, right and responsibility to use our own minds
- We are the authors of our own lives, human agency
- Also refers to enlightenment

VIENNA MANIFESTO ON DIGITAL HUMANISM

- It is a call for reflection and **action**, also a research program
- Some core principles:

Democracy

- Digital technologies should be designed to promote democracy and inclusion
- Fairness, responsibility and transparency of software programs and algorithms

Regulation

- Action / intervention against tech monopolies
- Decisions affecting human rights must be made by humans

The role of Research, Science, and Academia

- The connection of different scientific disciplines is essential
- Universities (create new knowledge, enhance critical thinking) have a particular responsibility

Education

- Academic teaching needs to combine humanities, social sciences and engineering
- Education on Informatics and its societal impact must start as early as possible

Afterwards – The Initiative

Since 2020 over 60 online lectures and workshops (pandemic)

Large amount of knowledge

Over 150 hours online accessible lectures (including slides), see caiml.org/dighum



The Initiative

- Lectures on
 - Efficiency and resilience
 - AI and ethics
 - Sovereignty
 - Privacy
 - AI Regulation Europe
 - Automation and Work
 - IT, Democracy and Geopolitics
 - Autonomous weapons
 -
- Manifesto – signatures from 45 countries, 8 languages
- In governmental programs and declarations
- EU Statement at UNO
- Fellowships (IWM), research and innovation programs

The Initiative

- Created **informal international intellectual interdisciplinary core**: authors of manifesto and books, members of lecture's Program Committee, speakers
 - Books and conferences such as ***Can machines save the world?***
 - Active in public campaigns, e.g. supporting EU AI Act
 - Roadmap on research and innovation
 - UNESCO chair on Digital Humanism at TU Wien
 - ACM Summer Schools on Digital Humanism
-
- Also many international civil society based and academic initiatives as well as international standardization organizations: e.g., IEEE Software Engineering Standard



AI - Things are Changing

- Now „mainstream" topic in politics
- International discussions and activities
 - Risks / threats / opportunities
 - Rules and / or regulation
 - Research and innovation programs
 - Geo Politics

Digital innovation DigHum

Table

	Critical issues	AI & human control	Labor & automation	Surveillance	Platform & monopolies	Online media & fake news	Digital sovereignty	Environment & sustainability
Research topics								
Explainability		X				X		
Transparency		X		X		X		
Privacy		X		X			X	
Personalization		X		X	X	X		
Fairness		X	X		X	X		X
Norms and ethics		X	X	X	X	X	X	X
Accountability of systems and providers		X	X	X	X	X	X	X
Machine/human cooperation and control		X	X	X	X	X		
Participatory approaches		X	X			X		
Security				X			X	
Regulatory approaches		X	X	X	X	X	X	X
DigHum business models				X	X	X	X	X
Content moderation		X		X		X		
Market mechanisms and power relations			X		X		X	
Resilience and efficiency (algorithms and architectures)		X	X		X		X	X
Open systems engineering (incl. interoperability, open data)				X	X		X	

Issues - Interdisciplinarity

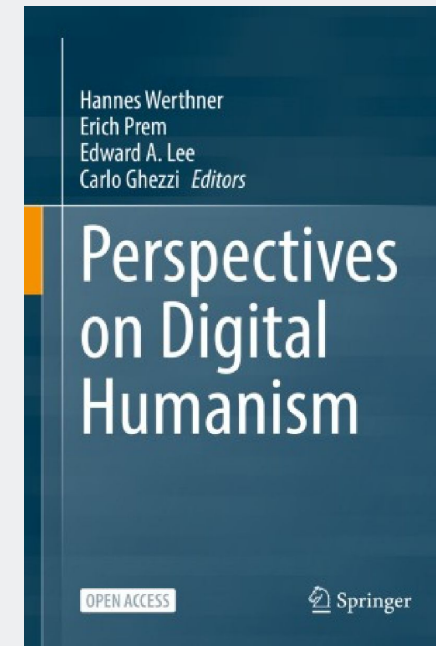
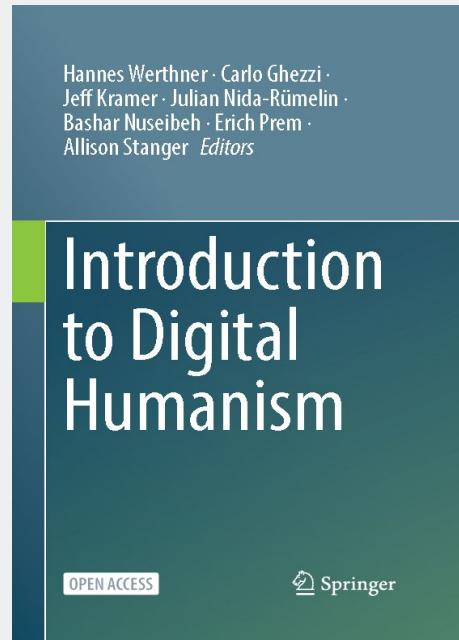
- *Fairness*
who is stakeholder: user, provider, platform – trade offs
who defines? what is included?
- *Resilience*
to survive in and to recover fast from critical situations
architectures, governance rules, adaptation and objective function
- Obvious need for more disciplines: social science, humanities,
We create technical and social constructs -> Computer Science becomes (also) a social science (R. Connolly, CACM2020)

Major crises such as climate change, decline of democracy, increasing inequality:

- IT is not the cause, but enables and reinforces
- It's not the solution (alone)

Conclusions

- IT will not stop
 - Further technology and „service“ waves
- Interdisciplinarity + Informatics as guiding discipline
 - New skills / knowledge (also perhaps at first seemingly useless)
- Development is not deterministic or „god given“
- It's up to us
- Not money and efficiency - but people, society and nature at the center
- At the end, it's about politics



Thank you

dighum.org