MEMORY SLICES

ANNA STRASSER BERLIN 2019

WORKSHOP 1
SOCIAL DYNAMICS & CULTURE OF AI

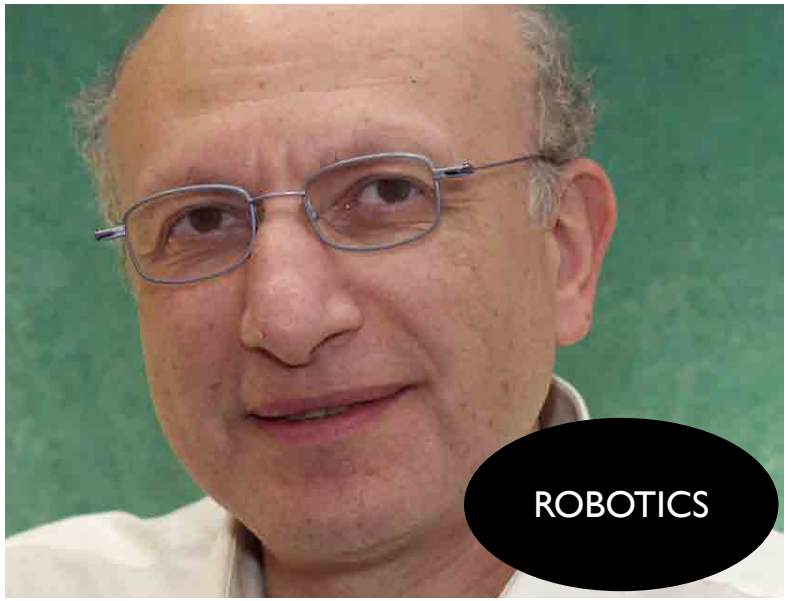
organized by

Marie-Hélène Parizeau
(Laval University, Québec, Chair of COMEST/UNESCO)

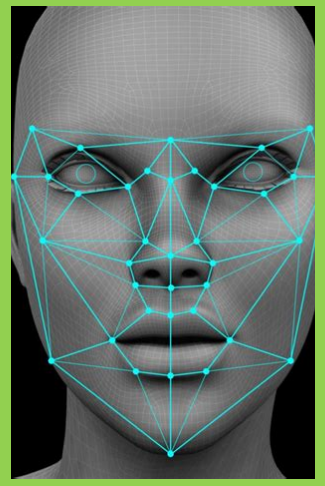
Véronique Guèvremont
(Laval University, Québec)

Vanessa Nurock
(Paris 8 University, France)

Raja Chatila
(Pierre et Marie Curie University, France)



RETHINKING CULTURAL AND ETHICAL ISSUES IN AI



surprising lack of cultural questions

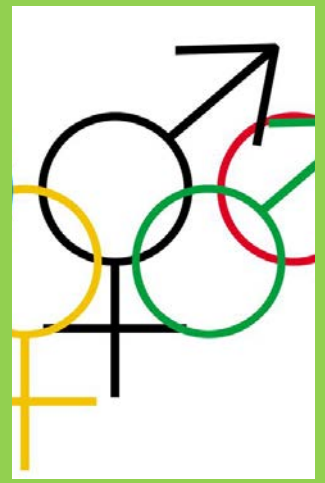
→ ethical framework for AI development & practices (+legislative initiatives)

→ social opportunity for large-scale creativity

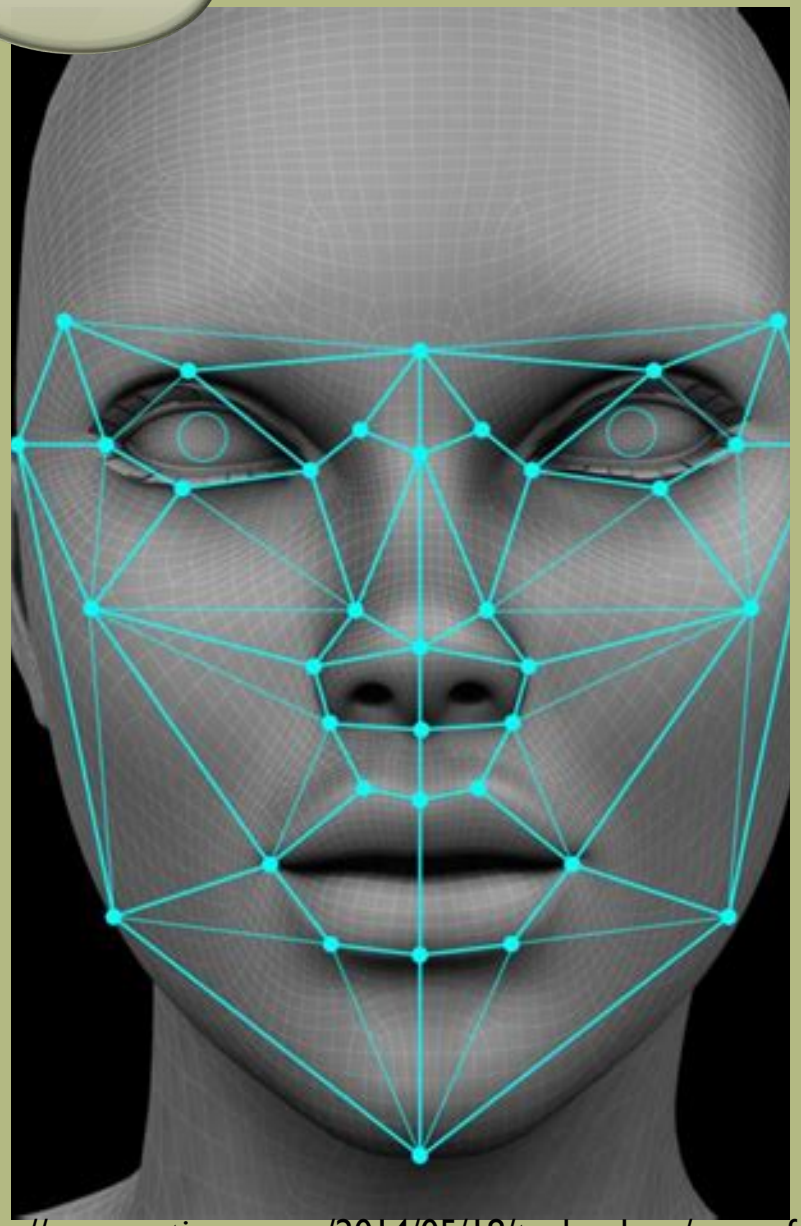
– MOMENT OF NORMATIVE OPENNESS –

1. CONVERGENCE OF SURVEILLANCE AND BIOMETRIC IDENTIFICATION AI + digital technologies + transformation of privacy

2. PROGRAMMING ETHICS INTO "AUTONOMOUS" MACHINES (ETHICS BY DESIGN) + impact on the concept of human freedom
3. IMPACT OF AI AND DIGITAL DEVELOPMENT ON CULTURAL CREATION + standardization to the detriment of cultural diversity & plurality of cultural expressions, as well as the dilution of copyright
4. FIXATION OF CERTAIN WORLDVIEWS THROUGH ALGORITHMIC PROGRAMMING + their ethical & political significance in terms of reports of gender discrimination & domination



SESSION 1



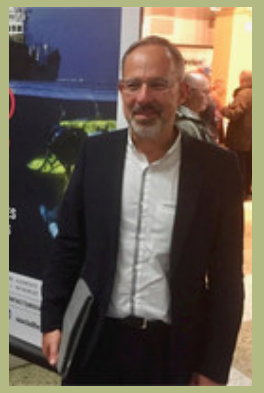
https://www.nytimes.com/2014/05/18/technology/never-forgetting-a-face.html

BETWEEN SURVEILLANCE & IDENTIFICATION: WHAT PRIVACY PROTECTION?

THE EXAMPLES OF FACIAL RECOGNITION AND BIOMETRICS

Chair: Tomislav Bracanović (Institute of Philosophy, COMEST) / Moderator: Francesca Musiani (CNRS))

1. Marie-Hélène Parizeau: **"Digital biometric and assigning identity: ethical and political analysis"**
2. Clare Gravie: **"Face recognition and the right to anonymity"**
3. James Katz: **"Public attitudes toward AI-enabled surveillance and facial/biometric identification technology: Views from a US opinion surveys"**
4. François Berger: **"Digital biometric of the self and the amplification of IA: a new questionable medicine"**



PHILO-SOPHER

MARIE-HÉLÈNE PARIZEAU, LAVAL UNIVERSITY, COMEST/UNESCO - CANADA

DIGITAL BIOMETRIC & ASSIGNING IDENTITY: ETHICAL AND POLITICAL ANALYSIS

- started in the 1980s / used in many applications / data are not protected

- *social logic of working through surveillance and digital identification of the Other*

➢ new parameters of assigning individual digital identity

Q: What could elude this movement and establish ethical limits?

- critical of what it is combined with!

CLARE GRAVIE, CENTER ON PRIVACY AND TECHNOLOGY, GEORGETOWN UNIVERSITY

FACE RECOGNITION AND THE RIGHT TO ANONYMITY

- fundamentally changing the nature of public spaces:
  - expectations of privacy and anonymity (identity, associations, daily routines)
  - loosing identity via face recognition without content

Q: What are the consequences w.r.t. people's willingness to engage in political and social dialogue?

- → Joseph J. Atick

JAMES KATZ, COLLEGE OF COMMUNICATION, BOSTON UNIVERSITY

PUBLIC ATTITUDES TOWARD AI-ENABLED SURVEILLANCE AND FACIAL/BIOMETRIC IDENTIFICATION TECHNOLOGY: VIEWS FROM A US OPINION SURVEYS

leaders of security & civil rights communities control the agenda for AI-enabled surveillance & personal identification

- public's attitudes towards surveillance & personal identification are not considered → decreases interventions & trade-offs that people might wish to make

information to guide selection & deployment of technologies + creation of appropriate laws & ethical guidelines

- studies (US public) concerning AI in people's lives
  - attitudes towards roles that AI might take / appropriateness of AI-supported surveillance and facial + biometric identification systems in the public's daily lives

range of value and ethical issues worthy of consideration in policy development

- → apparatgeist

FRANÇOIS BERGER, INSTITUT NATIONAL DE LA SANTÉ ET DE LA RECHERCHE MÉDICALE (INSERM)

DIGITAL BIOMETRIC OF THE SELF AND THE AMPLIFICATION OF IA: A NEW QUESTIONABLE MEDICINE

- major progress in digital technologies
  - algorithms processing of large amounts of data
  - miniaturization of micro-nanoelectronics → integrate perfectly into the environment and daily practices of individuals

→ connected flow of multimodal data that AI algorithms can decrypt in a predictive way to prevent, diagnose and treat

- modalities must be questioned on the basis of their epistemology, feasibility, and societal and ethical impact
  - preserving humans is progressive

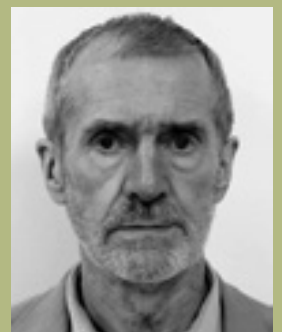
SESSION II

CAN WE PROGRAM ETHICS INTO AUTONOMOUS MACHINES?
THE EXAMPLE OF AUTONOMOUS CARS AND WEAPONS

Chair: Peter-Paul Verbeek (University of Twente, COMEST) / Moderator: Mélanie Dulong de Rosnay (CNRS)

1. Raja Chatila: **"AI as omniscience: a threat to human dignity and autonomy?"**
2. Patrick Lin: "**AI as omniscience"**
3. John Finney: **"Can autonomous weapons obey the 'laws' of war?"**
4. Jennifer Ang: **"Are autonomous weapons truly autonomous?"**






RAJA CHATILA, SORBONNE UNIVERSITY

AI AS OMNISCIENCE: A THREAT TO HUMAN DIGNITY AND AUTONOMY?

- AI is increasingly giving us the power of omniscience
  - better understand the world & better make certain decisions and interventions
- BUT it will disrupt existing norms and values, similar (expectations of privacy)
  - Can we measure to protect important ethical values in the age of AI?
  - *ethical deliberation =?= algorithm NO*
  - *machines do not know reasons, they have no moral judgment*
- responsibility remains to the human

PHILO-SOPHER

PATRICK LIN, CALIFORNIA POLYTECHNIC STATE UNIVERSITY

AI AND OMNISCIENCES. UNCOVERING …

- AI concerns many subjects *(40 issues in 4 areas listed)*

What is the nature of AI? → technology as a superpower

- superpower ethics / information encoded everywhere → info-sphere (like biosphere) – information = power
- decoding the universe is difficult as world is designed for unperfect perception of humans
- restore autonomy
  - NEW SOLUTIONS: transparency / consent …
  - explainability
  - ability to opt out

JOHN FINNEY, UNIVERSITY COLLEGE LONDON

CAN AUTONOMOUS WEAPONS OBEY THE 'LAWS' OF WAR?

- Geneva Conventions & Additional Protocols: weapons must satisfy requirements of distinction and proportionality
  - International Humanitarian Law: requires human decision making
  - Hague convention: only humans can command to kill other humans

'ethical governor's programmed to follow the relevant 'rules' cannot work because

- → critical decisions (during conflict) require human judgment
- → effective human control over advanced weaponry must be retained
  - by nature only humans can take such decisions

JENNIFER ANG, CENTRE FOR UNIVERSITY CORE, SINGAPORE UNIVERSITY OF SOCIAL SCIENCES

ARE AUTONOMOUS WEAPONS TRULY AUTONOMOUS?

- IF humans are taken fully out of the loop in decisions related to life and death

  → new ethical and legal concerns

- unresolved ethical challenges
  - examine the idea of 'autonomy' in terms of the capacity of self-rule

→ because 'autonomous weapons' are not truly autonomous ethical challenges can never be fully addressed

SESSION
III

CULTURAL POLITICS IN THE ERA OF AI: CREATIVITY AND DISCOVERABILITY

Chair: Luka Omladič (University of Ljubjana, COMEST) / Moderator: Yves Citton (Eur ArTeC, Paris 8 University)

- Véronique Guèvremont: **"The new measures in favour of discoverability"**
- Pierre-Luc Déziel: **"Use personal data in a context of valuing cultural diversity"**
- Alexandra Bensamoun: **"The protectability of AI creations: a legal and ethical issue"**
- Octavio Kulesz: **"Artificial intelligence and the cultural sector: opportunities and challenges"**

VÉRONIQUE GUÈVREMONT,
LAVAL UNIVERSITY

THE NEW MEASURES IN FAVOR OF DISCOVERABILITY

discoverability

- digital growth of cultural & creative industries → cultural policies
  - methods of intervention for local content do not translate well to digital platforms
  - → discoverability of this content requires the implementation of particularly innovative public policy

cultural identity / fundamental rights

- → UNESCO Convention on the Protection and Promotion of the Diversity of Cultural Expressions

GOVERNANCE FOR CULTURE

National policies and measures promote creation, production, distribution and access with regard to diverse cultural goods and services and contribute to informed, transparent and participatory systems of governance for culture.

FLOWS AND MOBILITY

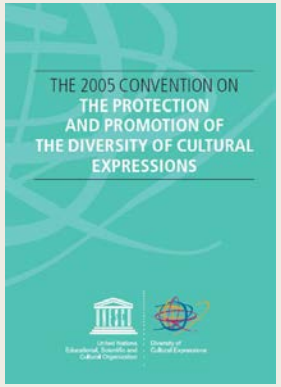
Preferential treatment measures facilitate a balanced flow of cultural goods and services and promote the mobility of artists and cultural professionals around the world.

SUSTAINABLE DEVELOPMENT

Sustainable development policies and international assistance programmes integrate culture as a strategic dimension.

HUMAN RIGHTS

International and national legislation related to human rights and fundamental freedoms promote both artistic freedom and the social and economic rights of artists.



PIERRE-LUC DÉZIEL, FACULTY OF LAW, LAVAL UNIVERSITY

USING PERSONAL DATA IN THE CONTEXT OF PROMOTING CULTURAL DIVERSITY

critique

- collaborative filtering and personalized recommendation algorithms → reinforce users' consumption habits and freeze their preferences in time
- → normalization of behaviors that threaten the dissemination of diverse cultural expressions.
- → alternate methods of using artificial intelligence that encourage the diversity of cultural expressions

ALEXANDRA BENSAMOUN, CERDI, UNIVERSITY OF RENNES I

THE PROTECTION OF CREATIONS GENERATED BY AI: A LEGAL AND ETHICAL ISSUE

- AI is capable of generating artistic creations that are comparable to those made by man
  - copyright protection?
  - new rights?

Q: Is the public domain the most reasonable way forward?

OCTAVIO KULESZ, TESEO/UNESCO

ARTIFICIAL INTELLIGENCE AND THE CULTURAL SECTOR: OPPORTUNITIES AND CHALLENGES

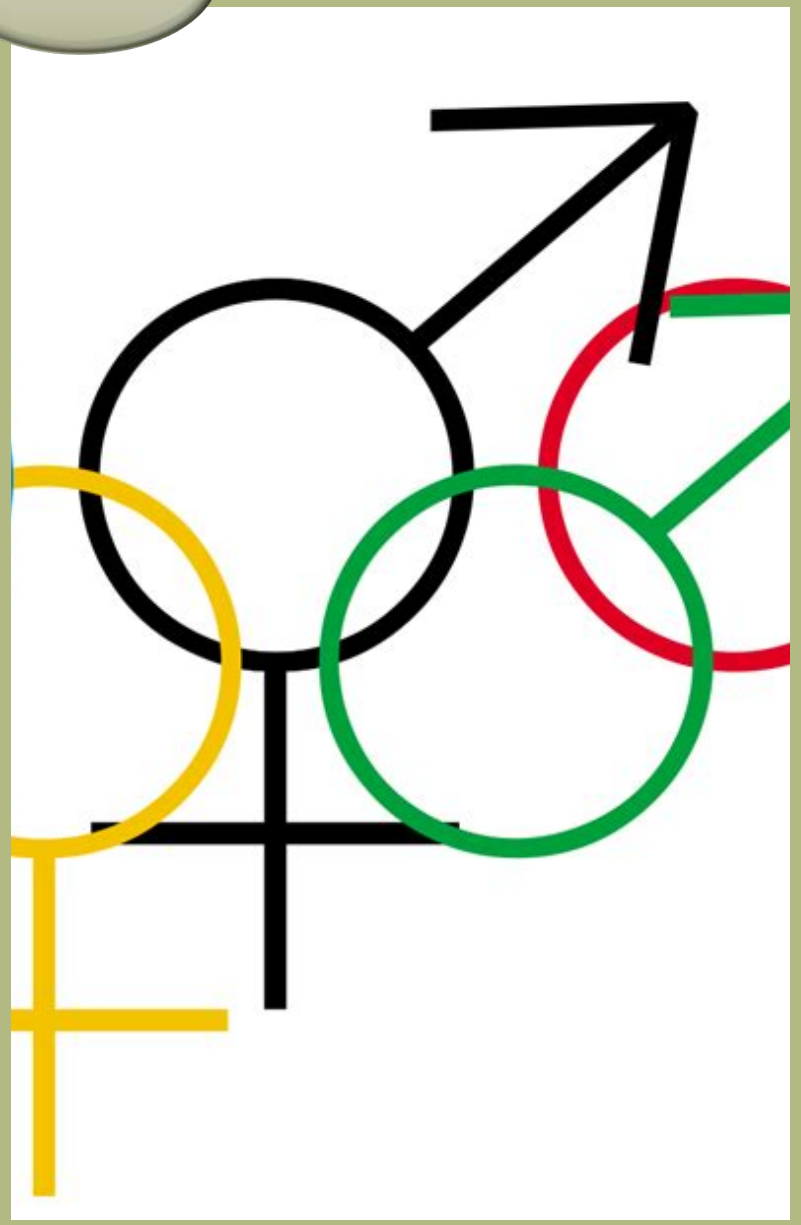
- empowers creators, makes the cultural industries more efficient, increases the number of artworks,
- BUT few artists know how to use tools such as machine learning

commercial logic of the large platforms

- increasing supply, data and income & impoverishment of cultural expressions in the long term
- all stakeholders should act in a coordinated way to take ownership of this powerful technology

essential to implement policies to take advantage of the opportunities and mitigate the challenges that AI poses for the cultural sector

SESSION IV



GENDER IN ARTIFICIAL INTELLIGENCE: ETHICAL, LEGAL AND POLITICAL ISSUES

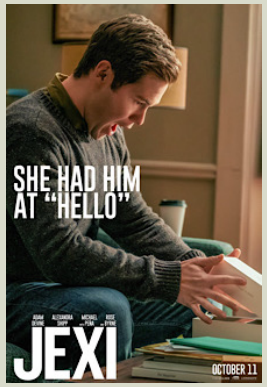
Chair: Sang Wok YI (Hanyang University) / Moderator: Marie Lechner (Gaîté Lyrique)

- Heather **Woods:** "The Politics of Gendering 'AI' Assistants: Learning from Popular Culture Contexts"
- Susan **Leavy:** "Gender Proof: Preventing AI from Learning and Perpetuating Gender"

HEATHER WOODS, COMMUNICATION STUDIES, KANSAS STATE UNIVERSITY

THE POLITICS OF GENDERING 'AI' ASSISTANTS: LEARNING FROM POPULAR CULTURE CONTEXTS

- pop culture helps us to make sense of technology
  - "smart" devices: ubiquitous, widespread distribution of voice-enabled virtual assistants by technology hegemons, including Amazon, Apple, Google …
  - **AI**-enabled **V**irtual **A**ssistants: gender stereotypes are reproduced
- film examples
  - JEXI (comedy)
  - AMI (horror film)



SUSAN LEAVY, *COMPUTER SCIENCE, UNIVERSITY COLLEGE DUBLIN*

GENDER PROOF: PREVENTING AI FROM LEARNING AND PERPETUATING GENDER BIAS

- widespread use of AI → potential to set back decades of advances in gender equality in society
- bias embedded in training data → reinforce bias in society
  - bias particularly problematic with language
  - gender theory, feminist linguistics & sociolinguistics deconstruct how gender ideology is embedded in language

Q: How to bridge AI and research in gender and language?

PHILO-SOPHER

VANESSA NUROCK (PARIS 8 UNIVERSITY, FRANCE)

THROUGH THE LOOKING GLASS OF ARTIFICIAL INTELLIGENCE

- lens to analyze
- justification through artification
- question of care
- rethink ethical & political issues in AI

SECOND DAY

Opening speech by **Cédric Villani** (Sorbonne University, Parliament of France)

1. **Trustworthy AI:** cybersecurity, safety, privacy, intelligibility of AI techniques
2. **Data governance:** ethical, political & legal challenges
3. **Future of work and economic impacts of AI**
4. **Delegation of decisions to machines, human agency and oversight**

TRUSTWORTHY AI: CYBERSECURITY, SAFETY, PRIVACY, INTELLIGIBILITY OF AI TECHNIQUES

Chairs: Raja Chatila (Sorbonne University), Virginia Dignum (Umeå University)

1. David **Sadek**: "AI and Cybersecurity"
2. Fosca **Giannotti**: "Intelligibility of AI"
3. Michael **Fisher**: "Verification of AI"
4. Stuart **Russell**: "AI and Safety"
5. Karen **Yeung**: "AI and Human Rights"





DAVID SADEK, THALES

AI AND CYBERSECURITY

perspective from industrial AI

necessary conditions for trust:

- safety: verification, qualification, certification, compliance, validation → continuous-learning systems
- explainability: a posteriori + explained by the system in real time
- cybersecurity: victim + actor (causing & preventing cyber-attacks)
- responsibility: fairness, questioning the usefulness of big data /compliance to ethic principles (laws, regulation, rules)

FOSCA GIANNOTTI (CNR ITALY)

INTELLIGIBILITY OF AI

5 principles

beneficence - non-maleficence – autonomy – justice - explicability

explicability:

1. intelligibility for experts and non-experts
2. right to explanation
3. how to explain: blackbox explanation / via conversation (why–questions) /give reasons

MICHAEL FISHER (LIVERPOOL UNIVERSITY)

VERIFICATION OF AI

AUTONOMY:= make decisions and take actions

- Can we anticipate them and see reasons why they behave as they behave?
- **trustworthiness**
  - reliable / intention

strong verification is needed (testing samples is not enough)

- avoid black boxes

STUART RUSSELL (UNIVERSITY OF CALIFORNIA AT BERKELEY)

AI AND SAFETY

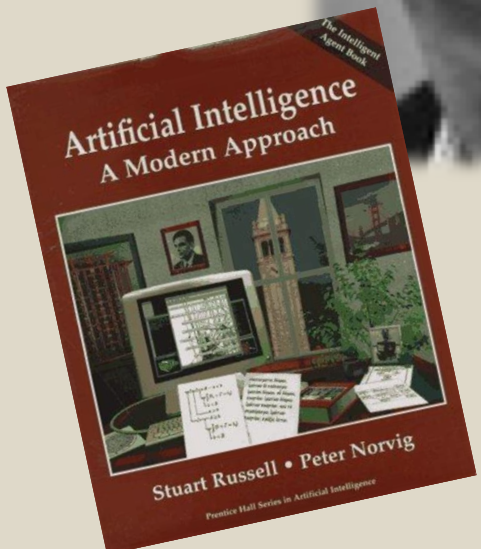
Asimov's 1st Law of Robotics is ignoring uncertainty

- specification is the wrong place to look for SAFETY

replacing the standard model for AI

MACHINE ARE BENEFICIAL TO THE EXTENT THAT THEIR ACTIONS CAN BE EXPECTED TO ACHIEVE OUR OBJECTIVES

- **machines do NOT KNOW what our objectives are!!!!**
  - interpret behavior reverse engineering…
- OPTIMAL SOLUTION:
  defer to human, ask permission, allow to be switched off



REWRITE THE BOOK → OBJECTIVES ARE UNCERTAIN!!

KAREN YEUNG (UNIVERSITY OF BIRMINGHAM)

AI AND HUMAN RIGHTS

CLAIM: only ethical AI can be trustworthy

- diversity of the human race → core of ethical norms?
- human rights

DATA GOVERNANCE: ETHICAL, POLITICAL & LEGAL CHALLENGES

Chairs: Sylvie Delacroix (University of Birmingham & Alan Turing Institute) and Joelle Pineau (McGill University & Facebook)

1. Paul **Nemitz**: "Iterative or linear moves between Ethics and law?"
2. Neil **Lawrence:** "Bottom-up data Trusts"
3. Lise **Getoor:** "Optimization, Quantification and Data Dignity"
4. Nigel **Shadbolt:** "Building an equitable data ecosystem in the age of AI"

PAUL NEMITZ

EUROPEAN COMMISSION, GERMAN DATA ETHICS COMMISSION

ITERATIVE OR LINEAR MOVES BETWEEN ETHICS AND LAW?

- GDPR / EU AI High Level Group → legislation on AI

Q: How far is law and technology in a substitutive relation?

Maintain democracy!

Literature:

- Constitutional democracy and technology in the age of artificial intelligence. Phil. Trans. R. Soc. A 376: 20180089, http://dx.doi.org/10.1098/rsta.2018.0089.
- The Report of the **German Data Ethics Commission** https://www.bmjv.de/DE/Themen/FokusThemen/Datenethikkommission/Datenethikkommission_EN_node.html

NEIL LAWRENCE, CAMBRIDGE UNIVERSITY

BOTTOM-UP DATA TRUSTS

- humanity → diversity! → human rights interpreted by humans
- AI is not anthromorphic intelligence – rather statistic
- ethics is about vulnerabilities

complementary to top-down regulation: bottom-up data aim to 'give a voice' to data subjects

- shape possible futures (natural resources, medical care etc.) is unlikely to be as impactful when leveraged individually

The power that stems from aggregated data should be returned to groups of individuals through the legal mechanism of Trusts.

LISE GETOOR, UNIVERSITY OF CALIFORNIA AT SANTA CRUZ

OPTIMIZATION, QUANTIFICATION AND DATA DIGNITY

socio-technical-systems are broken
because of outmoded notions of identity, property, power

DATA DIGNITY

- ability to understand & decide how your data is used

critical questions:

- group ownership / who owns the network
  - e.g. California Data Freedom Act
- new data governance strategies + deeper understanding and examination of AI algorithms.

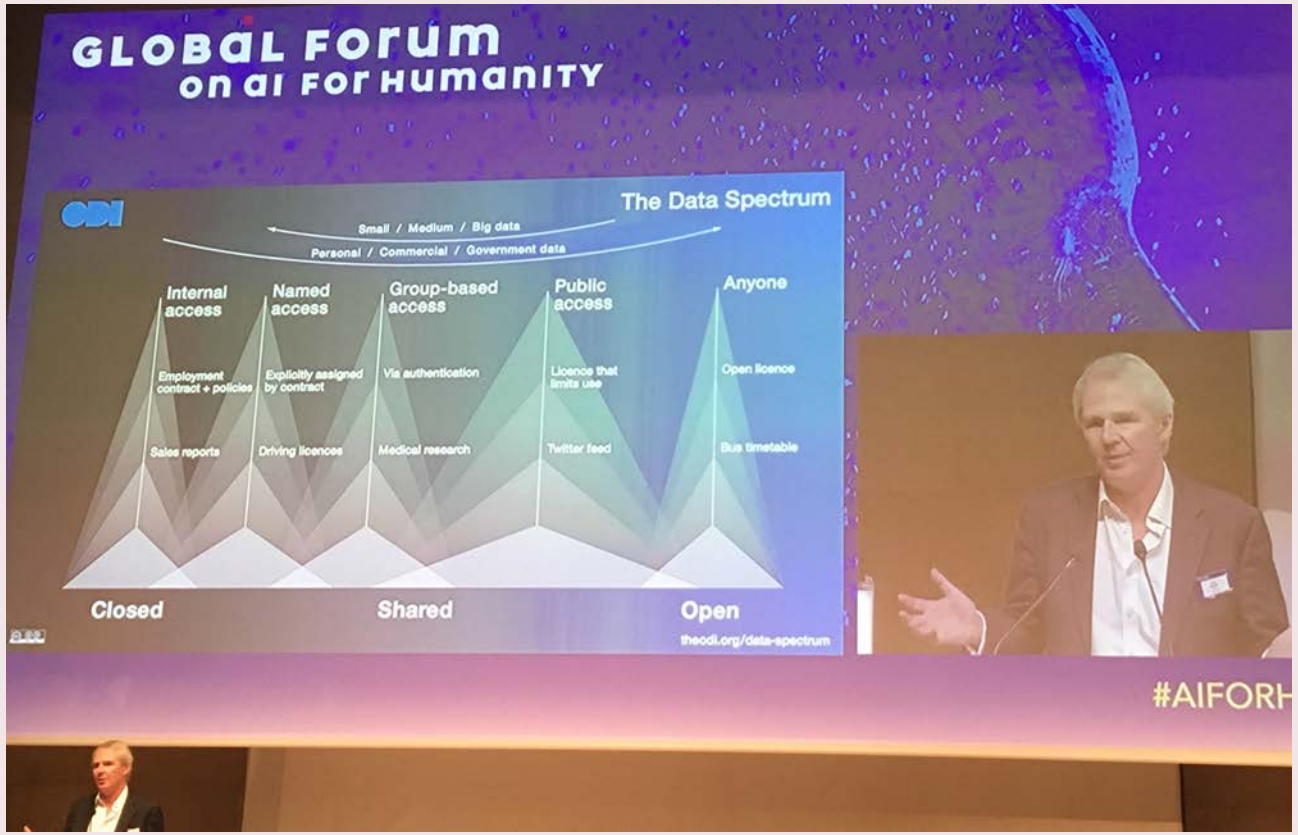
Understanding core assumptions that are inherent in AI systems are key to empowering individuals and social groups to make informed choices about both data sharing and platform participation.

NIGEL SHADBOLT, UNIVERSITY OF OXFORD AND OPEN DATA INSTITUTE

BUILDING AN EQUITABLE DATA ECOSYSTEM IN THE AGE OF AI

- components & characteristics of a data infrastructure sufficient to enable FAIR AI
  - e.g. flow of data of several combinations of app
  - e.g. used for medicine / marketing ….
- regulation needed

→ Kantian



FUTURE OF WORK AND ECONOMIC IMPACTS OF AI

Chairs: Yuko Harayama (Tohoku University) and Michela Milano (University of Bologna)

1. **Céline Antonin**: "Machines and Employment in France: Evidence from French Plants (2019)"
2. **Janine Berg**: "ILO's Survey of workers on digital labour platform (2018)"
3. **Anousheh Karvar**: "Proposal for new approach to formulate public policies related to labour market"
4. **Andrew Wyckoff:** "OECD Going Digital Project (2019)"
5. **Richard Baldwin** "The Globotics Upheaval (2019)"





STUDIES

Céline Antonin, OFCE and Collège de France:

Paris Institute of Political Studies

“Machines and Employment in France: Evidence from French Plants (2019)”

- towards a jobless world?
- new step in a continuum of automation
-

- Janine Berg, International Labor Organization: “ILO’s Survey of workers on digital labour platform (2018)”
- Geneva

POLITICS

- **Anousheh Karvar,** French Ministry of Labour/French Ministry of Europe & Foreign Affairs

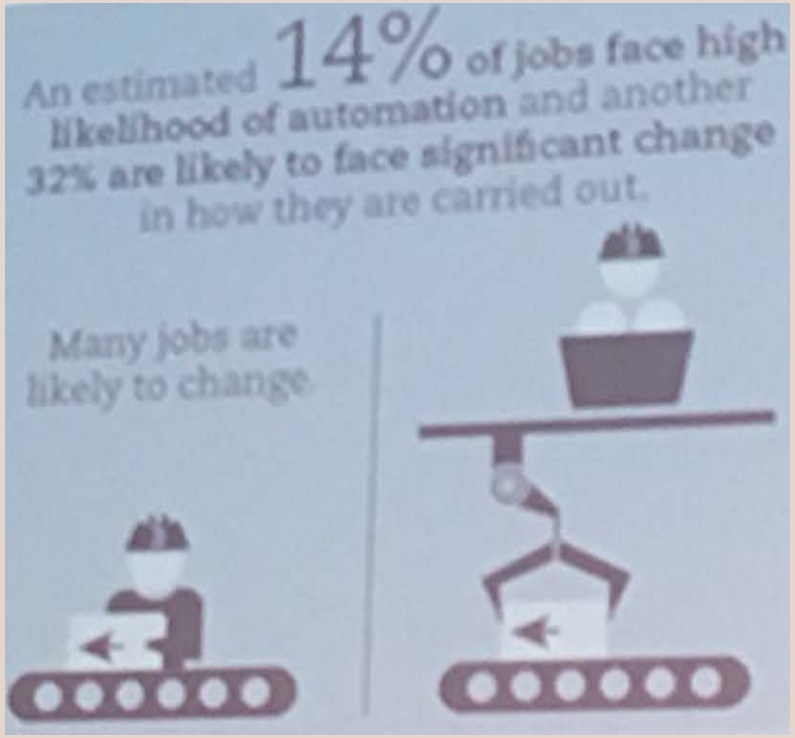
PROPOSAL FOR NEW APPROACH TO FORMULATE PUBLIC POLICIES RELATED TO LABOUR MARKET

- good human–machine-interaction
- AI seems good for everything?

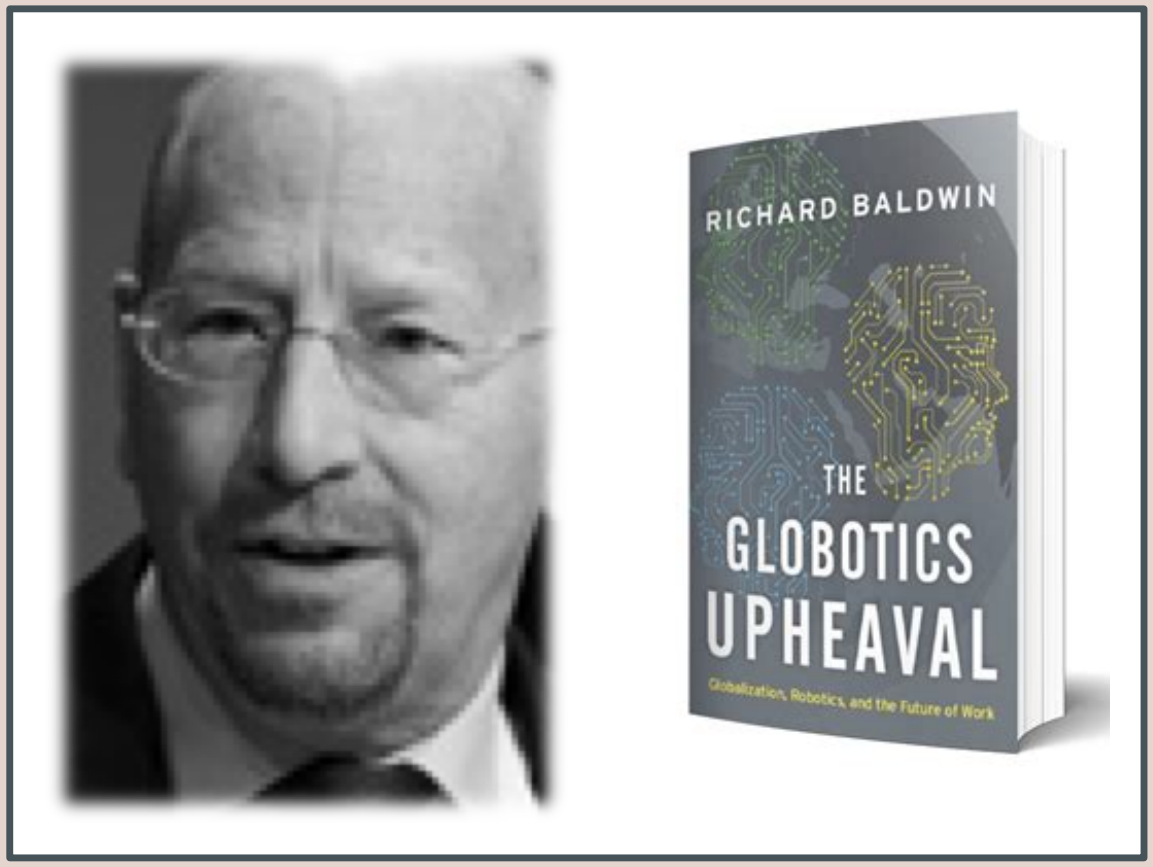
- **Andrew Wyckoff,** OECD Directorate for Science, Technology and Innovation:

OECD GOING DIGITAL PROJECT

- http://www.oecd.org/going-digital/project/
- *anticipation / inclusion directionality*
  → process governance in inovation

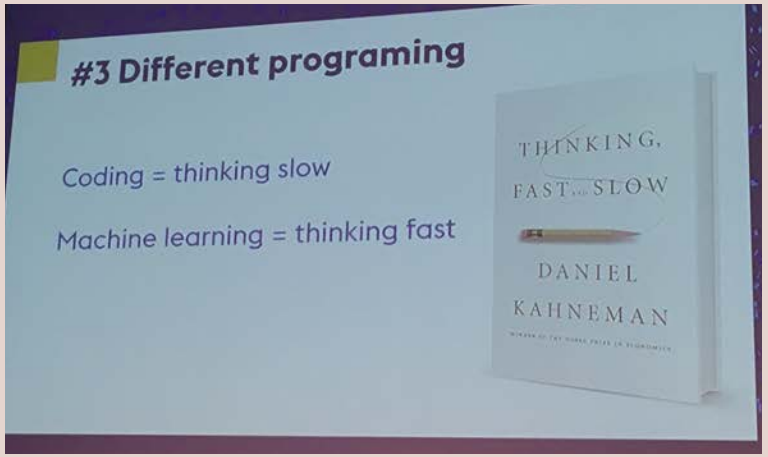


RICHARD BALDWIN, INTERNATIONAL ECONOMICS, GENEVA
"THE GLOBOTICS UPHEAVAL (2019)"



THE GLOBOTICS UPHEAVAL (2019)

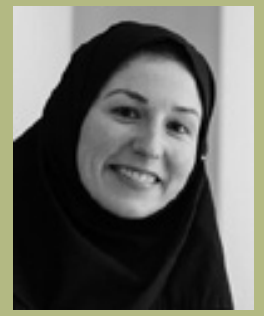
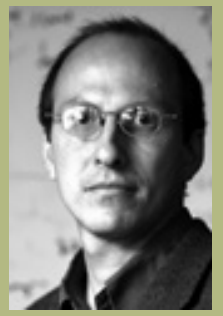
- think telemigration (RI) / think automation (AI)
- several myth: false physics --- faster …
- AI can't do human / RI can't be in the room



DELEGATION OF DECISIONS TO MACHINES, HUMAN AGENCY AND OVERSIGHT

- Jeroen van den **Hoven:** AI Ethics: design for responsibility and Human Agnecy
- Virginia **Dignum:** We are responsible
- Marzyeh **Ghassemi:** Learning Healthy Models for Healthcare
- Illah **Nourbakhsh:** AI Fluency in Education
- Carme **Torras:** Assistive AI: Ethics Education Initiatives based on Science Fiction




JEROEN VAN DEN HOVEN, DELFT UNIVERSITY OF TECHNOLOGY

AI ETHICS: DESIGN FOR RESPONSIBILITY AND HUMAN AGENCY

- …is about Humans!
  - safe for society → fair distribution of risk
- we have to design for responsibility knowledge / control / freedom & choice

VIRGINIA DIGNUM, UMEÅ UNIVERSITY

AI: WE ARE RESPONSIBLE

not just AI - AI applications are socio-technical systems

Which decisions AI should make (what are *better* decisions?)

- responsible AI systems:

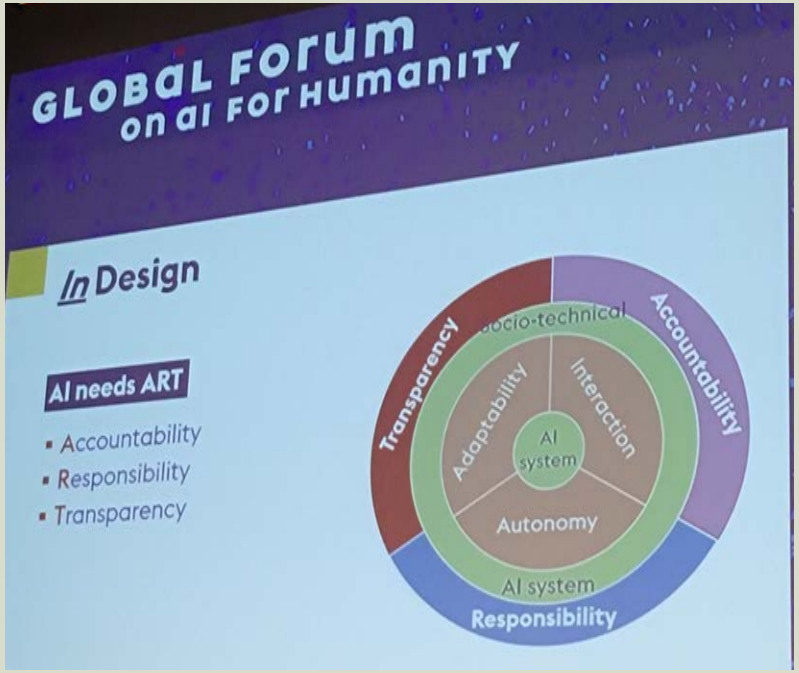
  requires more than designing systems whose result is aligned with ethical values.

→ **design, develop and use them, why, and who is involved**

1) regulatory & engineering processes → support the design and evaluation of AI systems

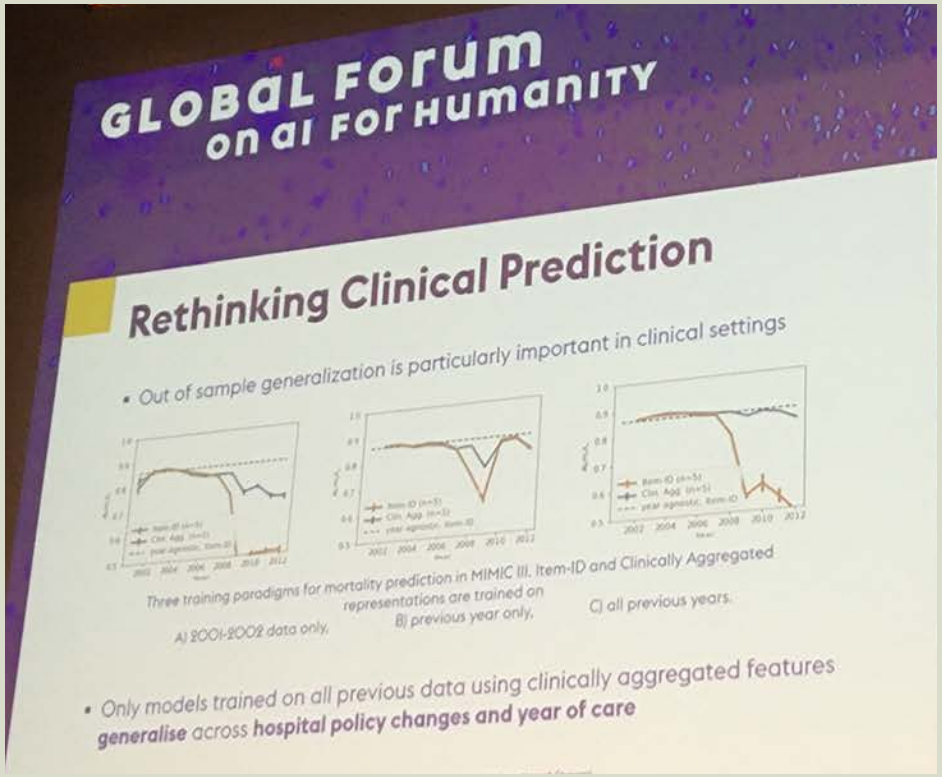
2) ensuring that behavior of AI systems is beneficial and ethical and follow a human-in-command stance, and

3) to the codes of conduct, the standards and certification processes that ensure the integrity of all actors as they research, design, construct, employ and manage artificial intelligent systems

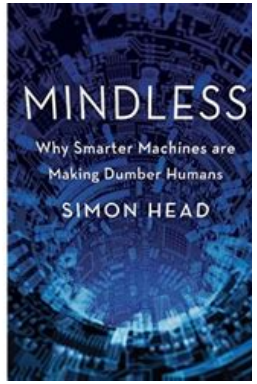


MARZYEH GHASSEMI, VECTOR INSTITUTE, UNIVERSITY OF TORONTO

Learning Healthy Models for Healthcare



ILLAH NOURBAKHSH, CARNEGIE MELLON UNIVERSITY:



A PENGUIN SPECIAL

Drone Theory

Gregoire Chamayou

AI FLUENCY IN EDUCATION

- **learn machine learning in a context**
- **we have to prepare future decision makes – education!!**

CARME TORRAS, SPANISH SCIENTIFIC RESEARCH COUNCIL

ASSISTIVE AI: ETHICS EDUCATION INITIATIVES BASED ON SCIENCE FICTION

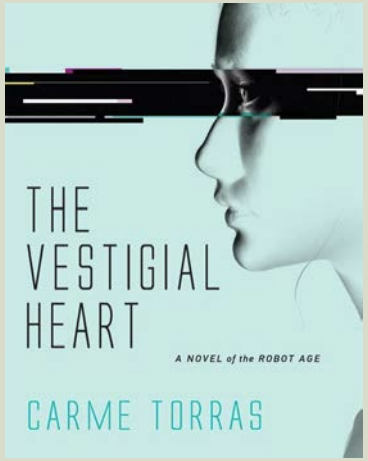
assistive AI → great expectations for social wellbeing + automatic decision-making → ethical concerns regarding human freedom and dignity

- Regulation and education initiatives

BOOK: *The vestigial heart*:

- provocative ethical issues:
- What kind of robots do we want when robot companions become as common as personal computers are now?
- Is it the responsibility of researchers to design robots that make the human mind evolve in a certain way?

- + free materials to teach a university course on Ethics in Social Robotics and AI based on science fiction



Instructor Resources

Digital Exam/Desk Copy
Print Exam/Desk Copy
Ancillary Material

SESSION 8.1. GLOBAL COOPERATION ISSUES AND THE GPAI ORGANIZATION

- **Panel 1.GPAI**
  Panel chairs:
  - Henri Verdier (French Ministry of European and Foreign Affairs),
  - Jordan Zed (DG of External and Trade Policy Branch, Canada)


Panelists: Moez Chakchouk (UNESCO), Michela Milano (University of Bologna, EurAI), Junichi Tsuji (Artificial Intelligence Research Center), Zee Kin Yeong (PDPC)




AI & HUMAN VALUES: INEQUALITIES, BIASES, FAIRNESS, REPRESENTATION OF MINORITIES, GENDER BALANCE

Chair: Françoise Soulié-Fogelman (Hub France IA)

1. Eric **Salobir:** AI & solidarity: when technology strengthens social bonds
2. Ricardo **Baeza-Yates:** Interaction biases
3. Francesco **Bonchi:** Fairness and privacy
4. Kate **Crawford:** The Politics of Classification
5. Laurence **Devillers:** Nudging with Affective computing system: inequalities and ethical issues





ERIC SALOBIR, OFFICE OF PRIESTS FOR ICT - OPTIC

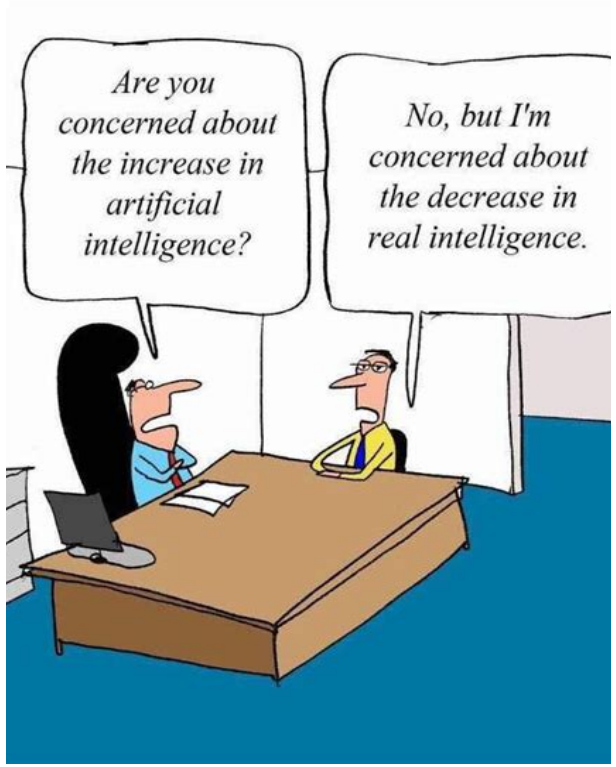
AI & SOLIDARITY: WHEN TECHNOLOGY STRENGTHENS SOCIAL BONDS

personalization versus solidarity: we are unique but we live together

example insurance: how AI can be part of the solution

- increased personalization → implement technologies that forge new ties of solidarity and strengthen the social fabric
- What we need to rebuild the social contract
  - public awareness
  - interdisciplinary work
  - political participation

RICARDO BAEZA-YATES, NTENT AND NORTHEASTERN UNIVERSITY



INTERACTION BIASES

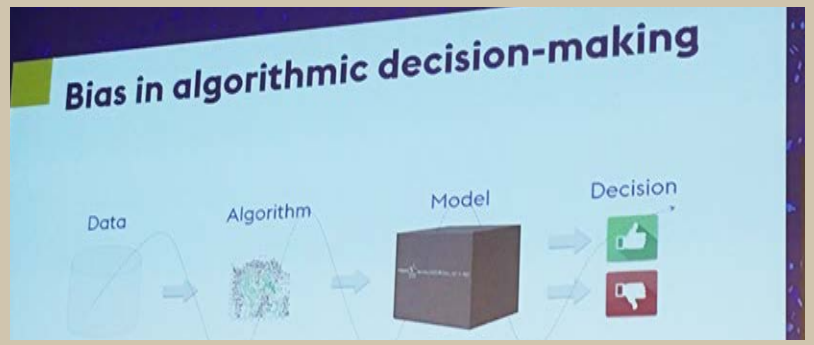
IMPACT IN SEARCH AND RECOMMENDER SYSTEMS

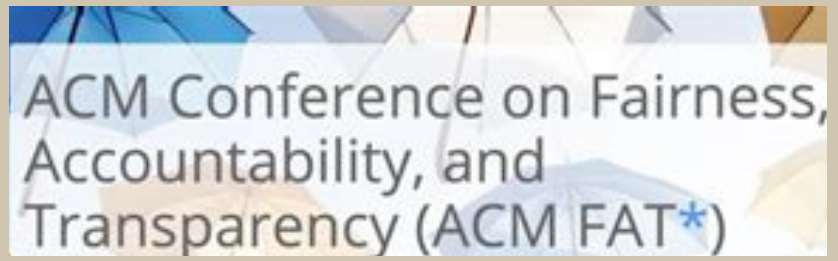
- shaped by their user = implicit user feedback (user data = subtle biased) → reinforce biases...
- position bias / ranking bias / presentation bias / social bias
- second order bias ....

FRANCESCO BONCHI, ISI FOUNDATION

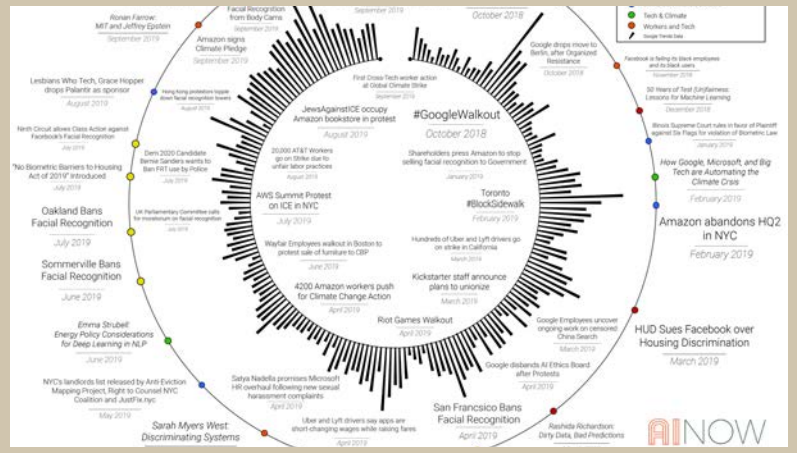
FAIRNESS AND PRIVACY

- → designing machine learning methods which are provably 'fair by-design'
- tradeoff between fairness and utility
- hiding sensitive information → hiding presence of discriminatory patterns





KATE CRAWFORD, AI NOW INSTITUTE / MICROSOFT RESEARCH LAB



THE POLITICS OF CLASSIFICATION

- relevant for social institutions, (education, healthcare, criminal justice)
- need more forensic analysis of the data
- methods of classification and human assumptions entrench bias in software
- socio-technical problem!

LAURENCE DEVILLERS, SORBONNE UNIVERSITY/DEPARTMENT OF HUMAN-MACHINE, INTERACTION OF CNRS-LIMSI

NUDGING WITH AFFECTIVE COMPUTING SYSTEM: INEQUALITIES AND ETHICAL ISSUES

- 80 % male programmers 80% social robots have female names
- How will Human co-learn, co-create and co-adapt with machines?
  - human machine co-evolution: → emotions play a central role

recognize interpret process simulate human language and affect

- nudging:= subtly modifying behavior
- How can vulnerable people be protected against potential threats (or nudges) from the machine?
- Do we need rules in what way nudging is okay...

SESSION 6. **NEXT BIG CHALLENGES IN CORE AI TECHNOLOGY**

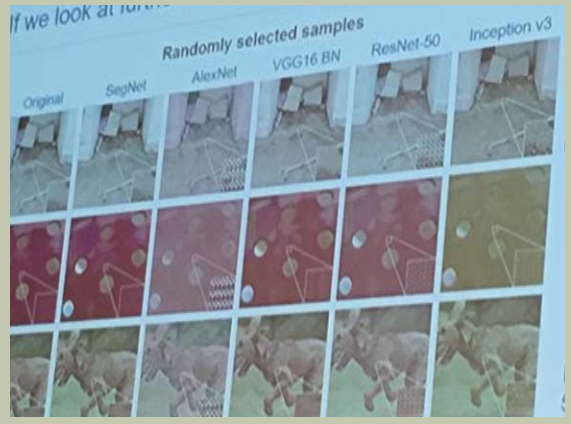
Chairs: Fei-Fei Li (Stanford University) and Junichi Tsujii (Artificial Intelligence Research Center, Japan)

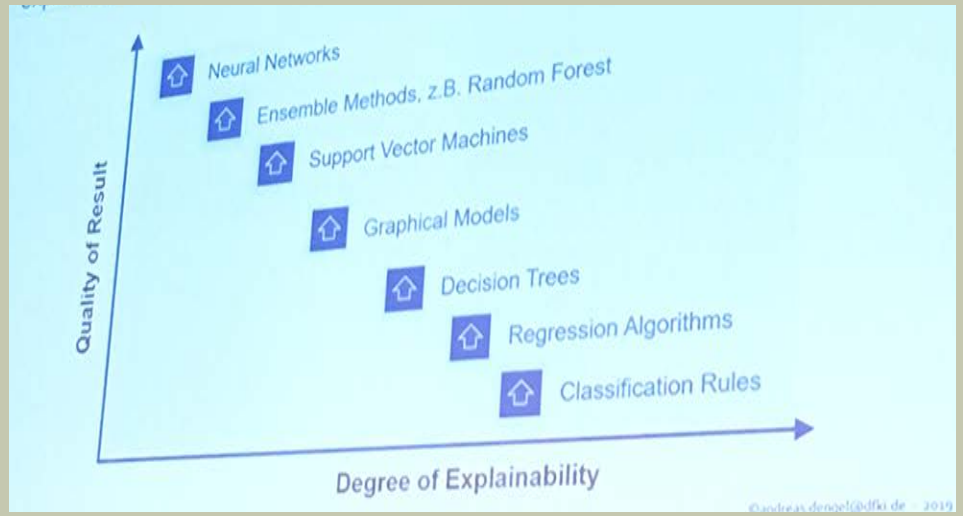
- Andreas Dengel: Endowing Deep Neural Networks to Show and Explain Behavior and Decision Making
- Oren Etzioni: The Challenge of Trustworthy AI
- Martial Hebert: AI and autonomy
- Holger Hoos: Addressing the AI Talent Bottleneck by Automating Artificial Intelligence
- Bernhard Schölkopf: Statistical and causal organizing principles of intelligence

ANDREAS DENGEL, DFKI

ENDOWING DEEP NEURAL NETWORKS TO SHOW AND EXPLAIN BEHAVIOR AND DECISION MAKING

- deep neural networks is extremely critical and risky → 'black box' character
  - difficult to interpret how or why the models come to certain conclusion
- innovative approaches to understand and explain decision-making of deep neural networks
- landscape / understand how information is used
  - use networks to explain network
  - how to engineer deep neural networks
    - how much information is used / which layers apply the most important filters

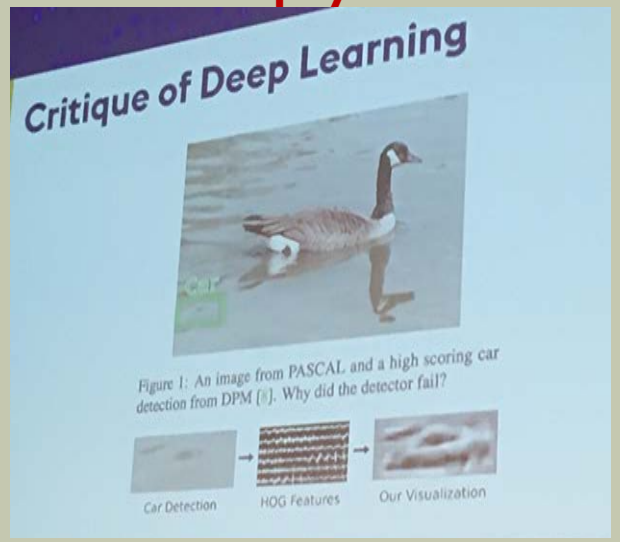


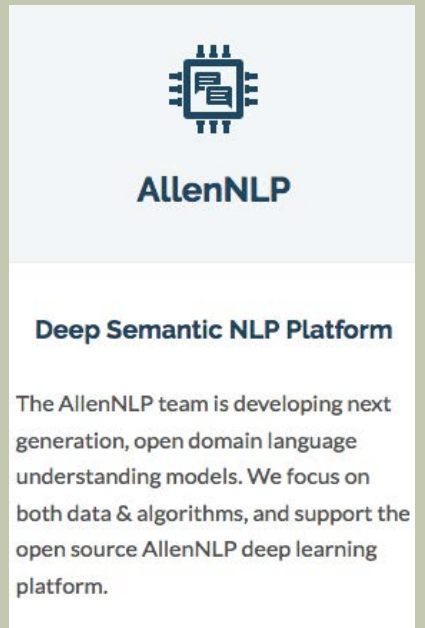


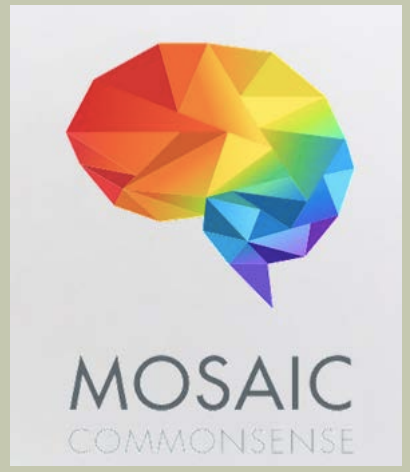
OREN ETZIONI, ALLEN INSTITUTE FOR AI

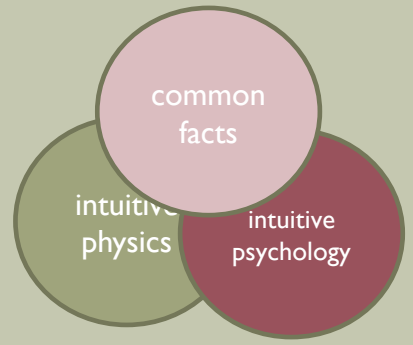
THE CHALLENGE OF TRUSTWORTHY AI

- step 1: training data
- step 2: apply deep learning
- step 3: observe impressive gains
- catastrophic & unexpected errors (← biases / adversarial attacks)
- **→ endowing AI systems with common facts / intuitive psychology / intuitive physics**

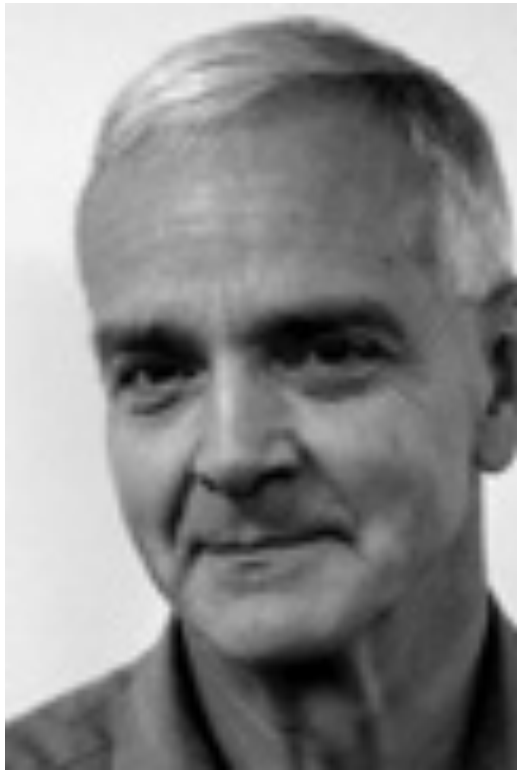








MARTIAL HEBERT, CARNEGIE MELLON UNIVERSITY



AI AND AUTONOMY

- learn a visual model with minimal supervision
  - learning from limited data
- modelling intent
  - read intention of the user
- deep learning techniques → has led to a rapid expansion of applications opportunities.
- BUT operational autonomous systems, → critical failure modes
  - introspection/self-awareness of performance, anytime algorithms for computer vision, multi-hypothesis generation, rapid learning and adaptation, from perception to action

HOLGER HOOS, LEIDEN UNIVERSITY

ADDRESSING THE AI TALENT BOTTLENECK BY AUTOMATING ARTIFICIAL INTELLIGENCE

- biggest risks of AI? → **incompetent use of weak AI**
- increased inequality & wide-spread use of low-quality AI systems
- machine learning automatic construction of algorithms
- opportunities climate change inequitity

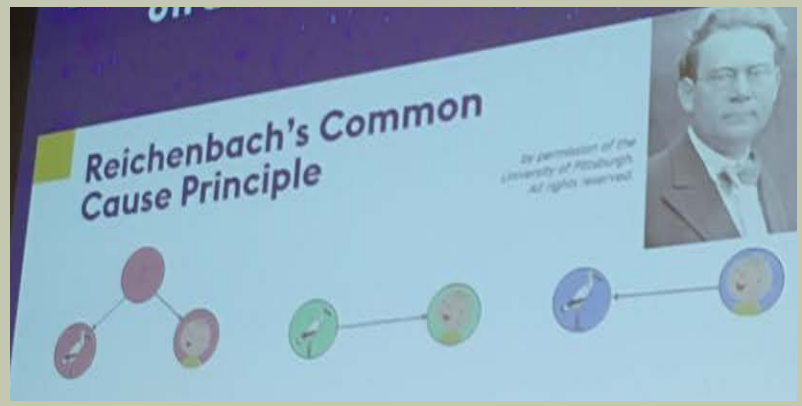
TALENT BOTTLENECK!

- humans tend to overestimate capacities of AI
- underqualified 'experts' -→ low-quality AI
- us AI to construct AI
- AI → **AUTO AI**

BERNHARD SCHÖLKOPF, MPI FOR INTELLIGENT SYSTEMS AND ETH ZURICH

STATISTICAL AND CAUSAL ORGANIZING PRINCIPLES OF INTELLIGENCE

- machine learning: use data to find dependences in the world to predict future observations
- build on statistics + assaying causal structures underlying statistical dependences
- Can causal knowledge help us produce intelligent behavior?
  - central role in addressing some of the hard open problems of machine learning
  - causal models are more robust to changes
  - implications of causal models for machine learning tasks→ 'independent mechanisms'



GLOBAL COOPERATION ISSUES AND THE IPAI ORGANIZATION

Chairs: Jocelyn Maclure (Laval University) & Stuart Russell (University of California at Berkeley)

- Stephen **Muggleton**, AI for Science
- Benjamin **Combes**, AI for Earth:
- Nicholas **Ayache**, AI for Health: hopes and challenges
- Cristina **Conati**, AI for Education: Let's not Miss the Opportunity

STEPHEN MUGGLETON, IMPERIAL COLLEGE LONDON: AI FOR SCIENCE

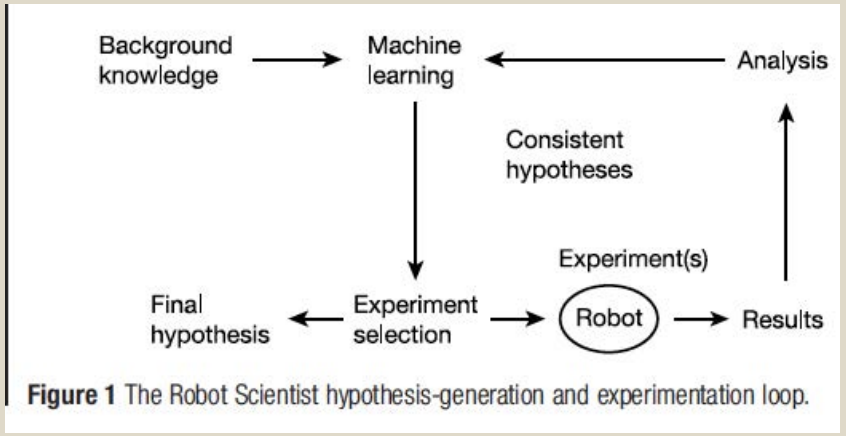


Figure 1 The Robot Scientist hypothesis-generation and experimentation loop.

King, R. D. et al. (2004). Functional genomic hypothesis generation and experimentation by a robot scientist. *Nature* **427**, 247–252.

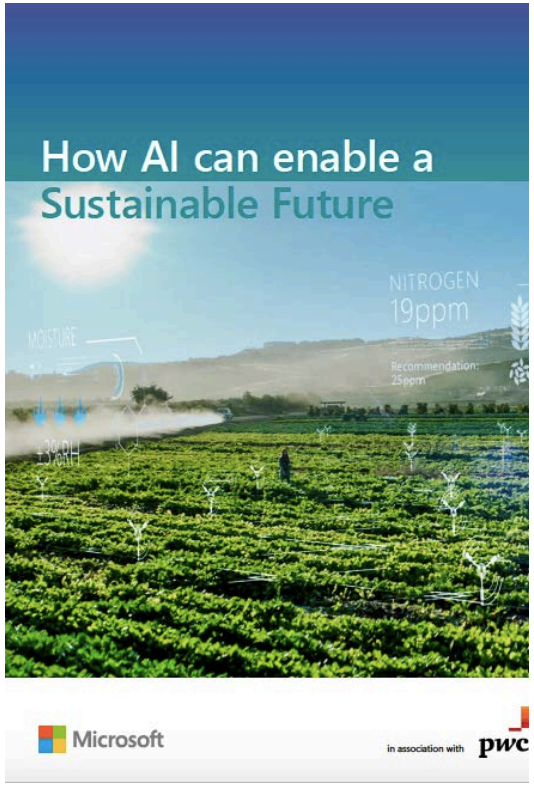
LOGIC-BASED AI TECHNIQUES APPLIED TO SCIENTIFIC DISCOVERY PROBLEMS

- → human-comprehensible hypotheses, allowing comparison against the literature by human scientist
  - results in Drug Design, Protein Structure prediction, Systems Biology and Synthetic Biology

TOPICS: global warming / disease / food security (e.g. proteine structure discovery)

D.A. Bohan, G. Caron-Lormier, S.H. Muggleton, A. Raybould, & A. Tamaddoni-Nezhad. (2011). Automated discovery of food webs from ecological data using logic-based machine learning. *PloS ONE*, 6(12)

BENJAMIN COMBES, PRICE WATERHOUSE COOPERS



AI for Earth: Innovation & Sustainiblity

← download report

- https://www.pwc.co.uk/services/sustainability-climate-change/insights/how-ai-future-can-enable-sustainable-future.html

NICHOLAS AYACHE, INRIA & SCIENTIFIC DIRECTOR OF THE NEW AI INSTITUTE 3IA CÔTE D'AZUR

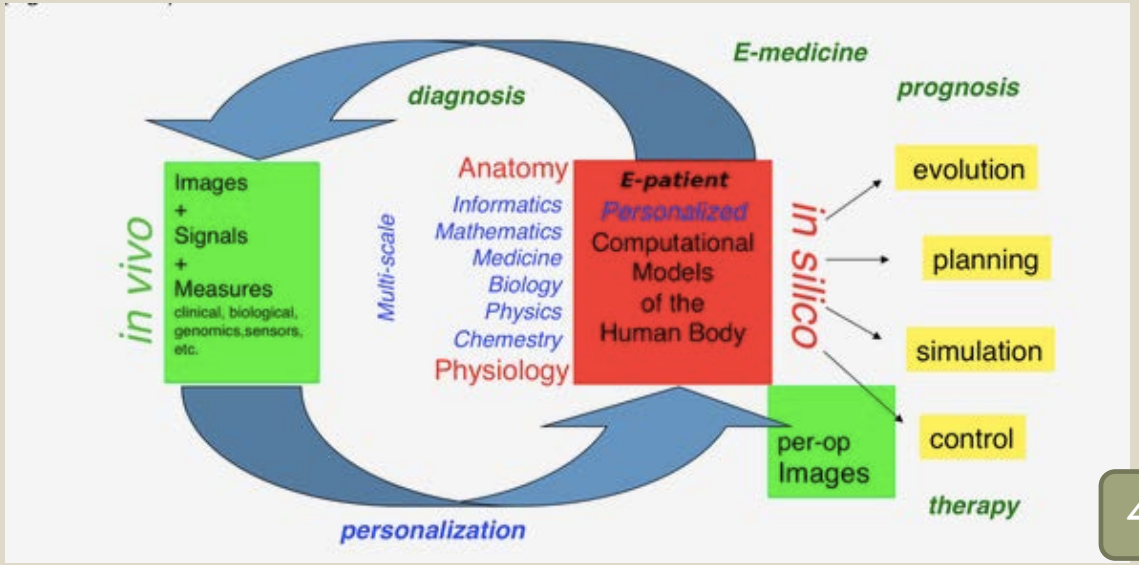
AI FOR HEALTH: HOPES AND CHALLENGES

fields: dermatology / radiology ...

limitations:

- LARGE ANNOTATED DATA
- RISK OF BIAS
- ETHICAL ISSUES
- BLACK BOX EFFECT

→ model-based AI:
e-patient (digital twin) for e-medicine



CRISTINA CONATI, UNIVERSITY OF BRITISH COLUMBIA

AI FOR EDUCATION: LET'S NOT MISS THE OPPORTUNITY

- ideal AIED Tutor
- Conati, C., (To Appear). Commentary on: "Toward Computer-Based Support of Meta-Cognitive Skills: a Computational Framework to Coach Self-Explanation". *International Journal of AI in Education, 25th-anniversary Special Issue: Most Impactful IJAED research papers commented by their authors.*

SESSION 8 - GLOBAL COOPERATION ISSUES AND THE GPAI ORGANISATION

REGULATION & REDLINES – CHAIR PEKKA ALA-PIETILÄ

Konstantinos **Karachalios**, IEEE Standards Association

- expert in standards development & intellectual property / managing director of IEEE Standards Association, member of IEEE Management Council
- WE ARE PART OF THE PROBLEM

David **Sadek**, Thales

- pioneer in human-machine natural language processing technology / Vice President of Research and Technology at Thales Group
- COOPERATION HAS TO BECOME STRONGER +INDUSTRY

Yuko **Harayama**, Tohoku University

- 2010: director of Science, Technology and Industry OECD in France / until 2018 Council of Science, technology and innovation
- Japan: AI strategy" (03, 2017) technological + ethical / social aspects of AI "Artificial Intelligence and Human Society"
- NEED FOR GLOBAL COOPERATION TO DEVELOP RESPONSIBLE AI

Jim **Snabe**, Siemens

- formerly SAP now chairman of Siemens
- EVEN IF AI IS IN ITS INFANCY IT IS TIME TO REGULATE NOW