13.45pm – 15.00pm

SSH, Technological Developments & Society

Brent Barron (Chair) - CIFAR
Carolyn Watters – National Research Council Canada
Emile Aarts – Tilburg University

Responsible AI
Nexus of development and social impact

Dr. Carolyn Watters
Chief Digital Research Officer
National Research Council Canada

14 October 2020

Overview

1. Responsible AI frameworks
2. Stakeholder engagement
3. Next steps

But first, some definitions. . .

Artificial Intelligence

Simulation of human intelligence using algorithms, including use of machine learning, reasoning through rules, and self-correction through feedback.

Augmented Intelligence

Complementing human intelligence in comprehending and deriving solutions to situations that would otherwise be too complex to resolve in the time available.

Machine Learning

Application of AI to data to improve algorithm outcomes without being explicitly programmed.

What do we mean by "responsible AI"?

Responsible AI is grounded in the human-centric principle of accountability.

Organizations & people developing, deploying, or using AI systems are accountable for harm caused by AI.

Governments are responsible for protecting the best interests of citizens.

Responsible AI in the Canadian context

Canada's Digital Charter
Principles

1. Universal Access
2. Safety & Security
3. Control & Consent
4. Transparency, Portability & Interoperability
5. Open & Modern Digital Government
6. A Level Playing Field
7. Data & Digital for Good
8. Strong Democracy
9. Free from Hate & Violent Extremism
10. Strong Enforcement & Real Accountability

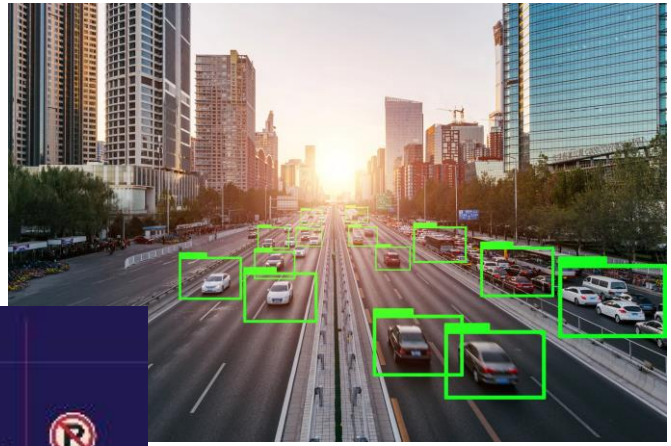
https://www.ic.gc.ca/eic/site/062.nsf/eng/h_00109.html

Responsible AI: A Global Policy Framework
Responsible AI Principles

1. Ethical purpose & societal benefit
2. Accountability
3. Transparency & explainability
4. Fairness & non-discrimination
5. Safety & reliability
6. Open data & fair competition
7. Privacy
8. AI & Intellectual Property

International Technology Law Association. https://www.itechlaw.org/ResponsibleAI

Why do we need a responsible AI framework?



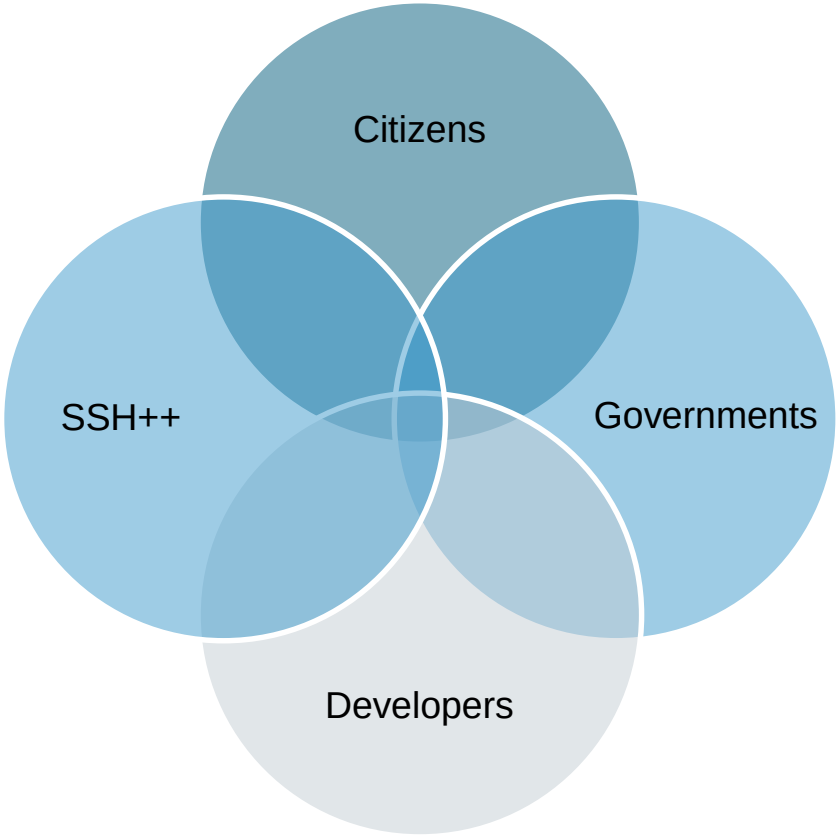
Are self-driving cars a benefit?

Responsible AI

- Which values are considered? Whose priorities?
- How do we deal with dilemmas?
- How do we identify and measure effects, both intended and unintended?

- Who participates in these discussions?
- Who is responsible for leading these discussions?
- Who is responsible for regulation?

Who are the stakeholders?



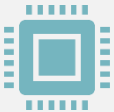
How do they influence responsible AI?

- Inform technology development and deployment
- Provide human perspective of technology impact
- Provide regulatory frameworks & policy
- Develop systems that have accountability

Role of SSH

Multidisciplinary Perspectives

Stewardship role on human-centric impact and potential



Co-partners in AI design and oversight functions

Questions we can address collectively

1. How do we distribute the wealth created by algorithms fairly?

2. How do machines affect our behaviour and personal interaction?

3. Technical nudging? How does intentional direction of human attention to trigger certain actions change society and, possibly, values?

4. How do we eliminate AI bias in direct interactions with algorithms?

5. Can we identify classes of artificial intelligence that have potential for positive change for citizens.

THANK YOU

Carolyn Watters • Chief Digital Research Officer •
Carolyn.Watters@nrc-cnrc.gc.ca

NL AI Coalitie

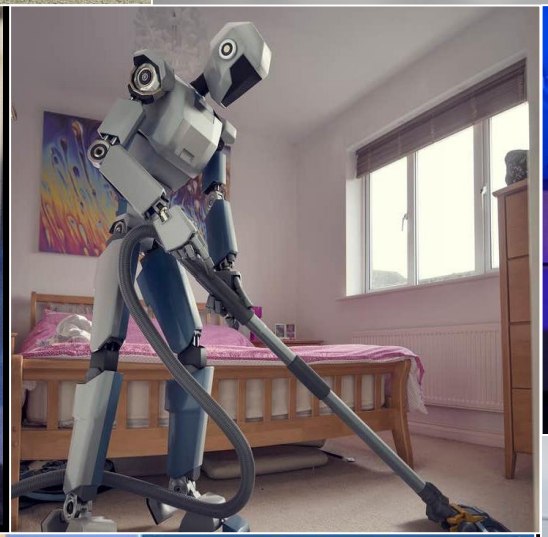
ELSA Labs

A Team Science-Based Approach to Human-Centric AI Innovation

Emile Aarts

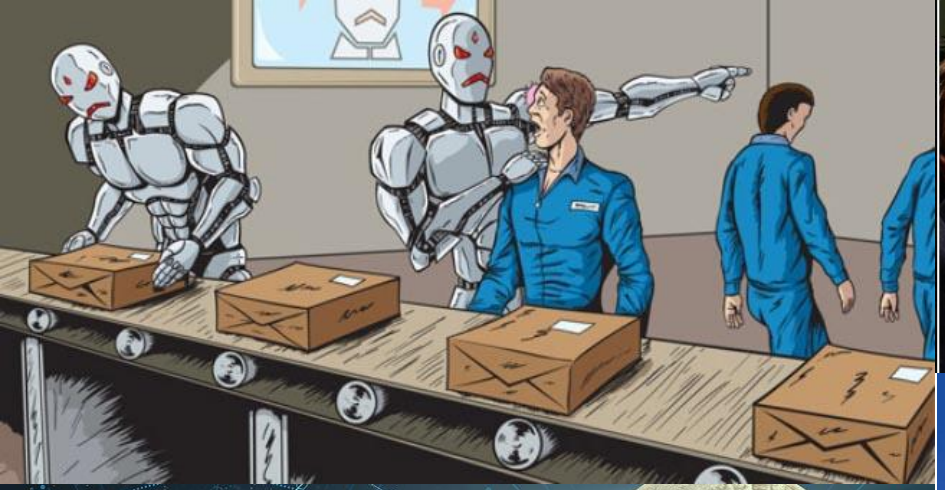
October 14 2020

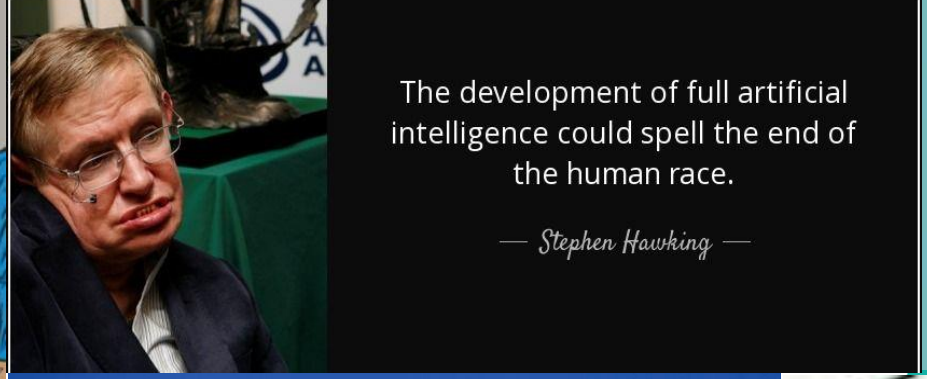

PAINT

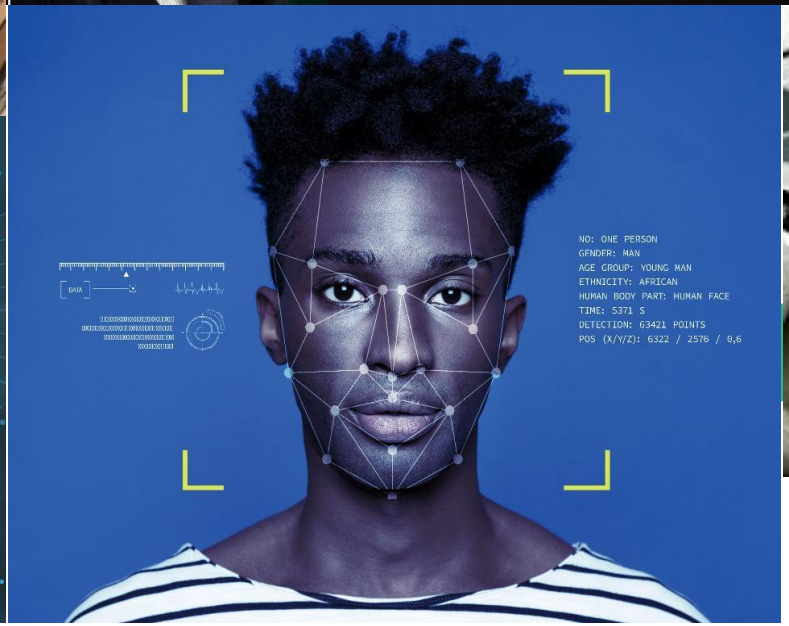


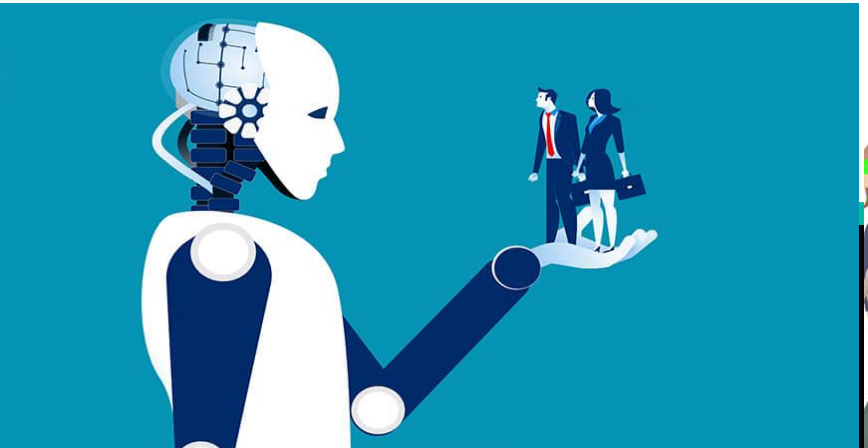

THINK
$24,000
Who is Stoker?
$1,000
$77,147
Who is Bram Stoker?
$ 17,973
$21,600
WHO IS BRAM STOKER?
$5600

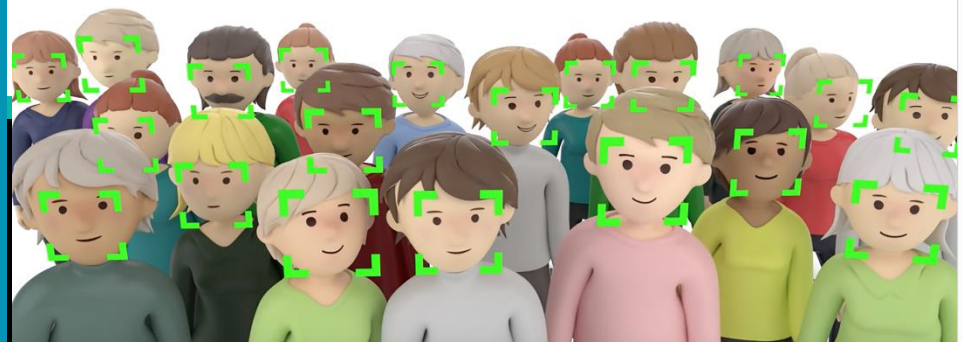

ALPHAGO
01:51:38



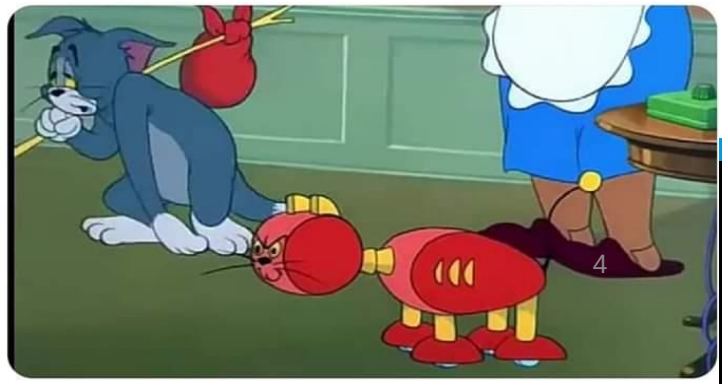






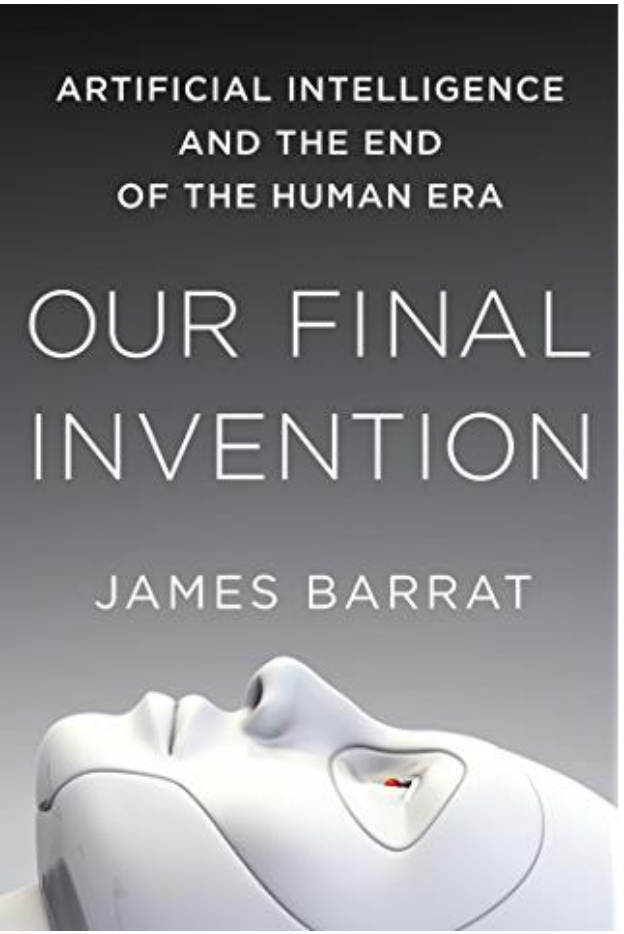


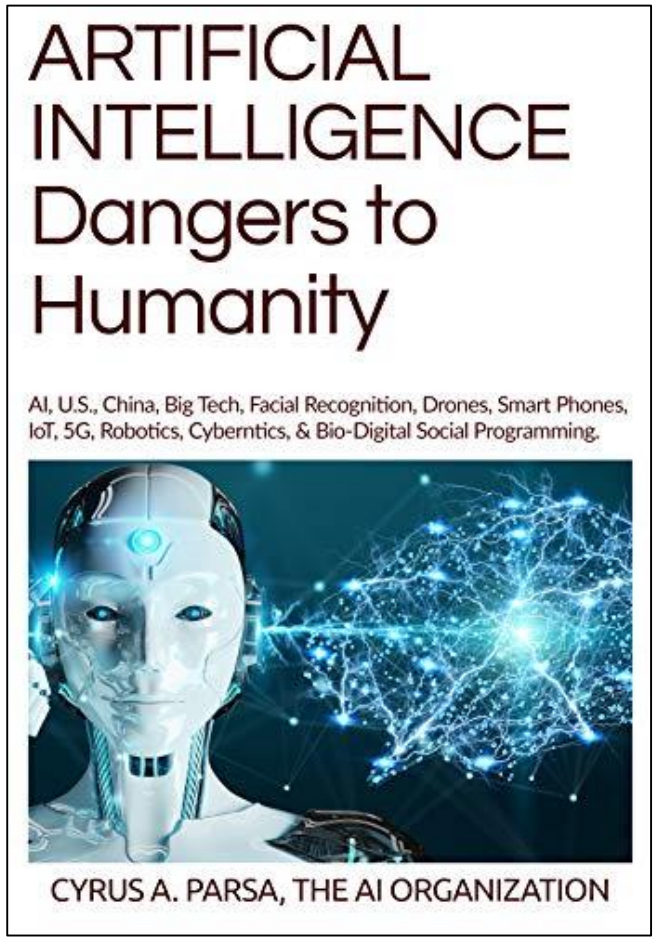
Tom was the first guy losing his job because of Artificial intelligence

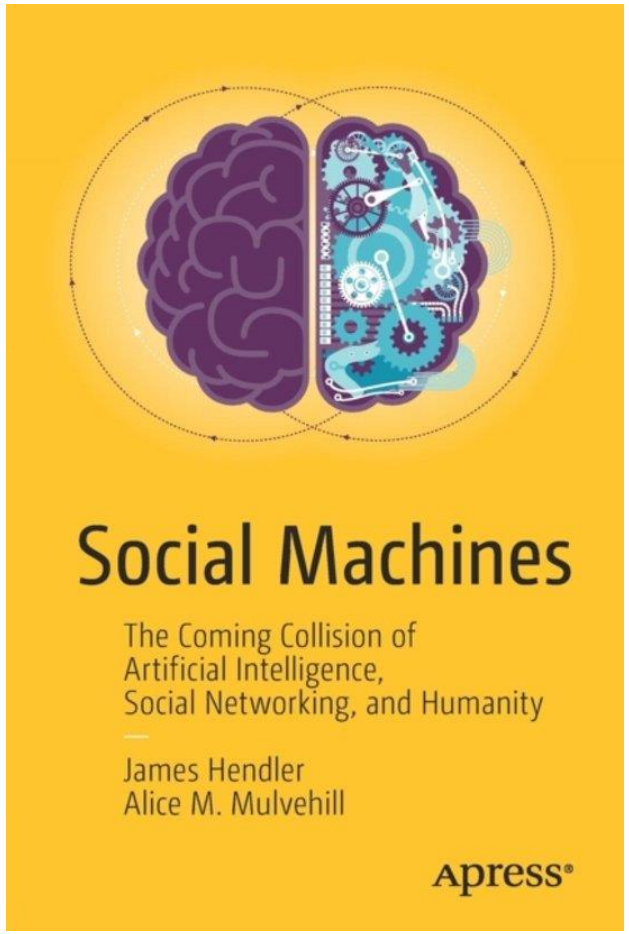


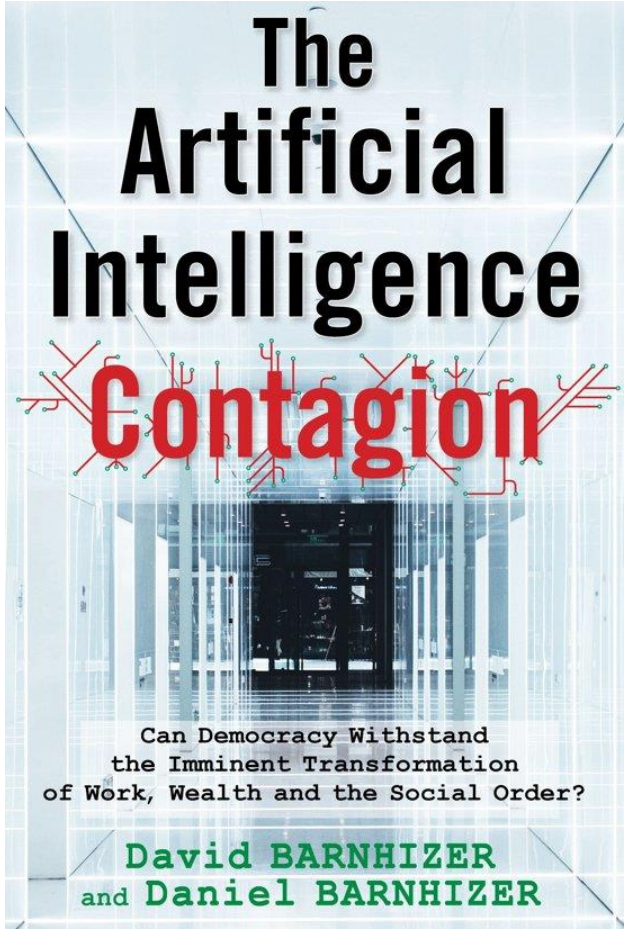
NL AI Coalitie

There are many dystopic visions and warnings









NL AI Coalitie

The EU vision on Human AI is clear

7 Key Requirements for Trustworthy AI

- **Human agency and oversight**
- **Robustness and safety**
- **Privacy and data governance**
- **Transparency**
- **Diversity, non-discrimination and fairness**
- **Societal and environmental well-being**
- **Accountability**

PILOTING OF THE ETHICS GUIDELINES FOR TRUSTWORTHY AI

How can developers, deployers or any citizen affected by the use of AI make sure that the 7 Key Requirements are implemented?

In their founding document, the "**Ethics Guidelines for Trustworthy AI**", the High-Level Expert Group on AI (AI HLEG) outlined 7 Key Requirements that are complemented by an "**assessment list**" to support their practical implementation.

In June 2019, the Commission launched a **piloting process**, inviting all stakeholders to provide feedback on how this assessment list can be improved.

Interested stakeholders can register and participate in the piloting until the **1st of December 2019**.

The Dutch AiNed Investment Plan

- **Economic**
  increasing the number of companies active as well as investment levels in development and application of AI in impactful sectors, missions, platforms
- **Economic**
  contributing to economic growth (recovery) for Netherlands (potential 1.6% GDP)
- **Societal**
  human centric AI and autonomy (in EU context)

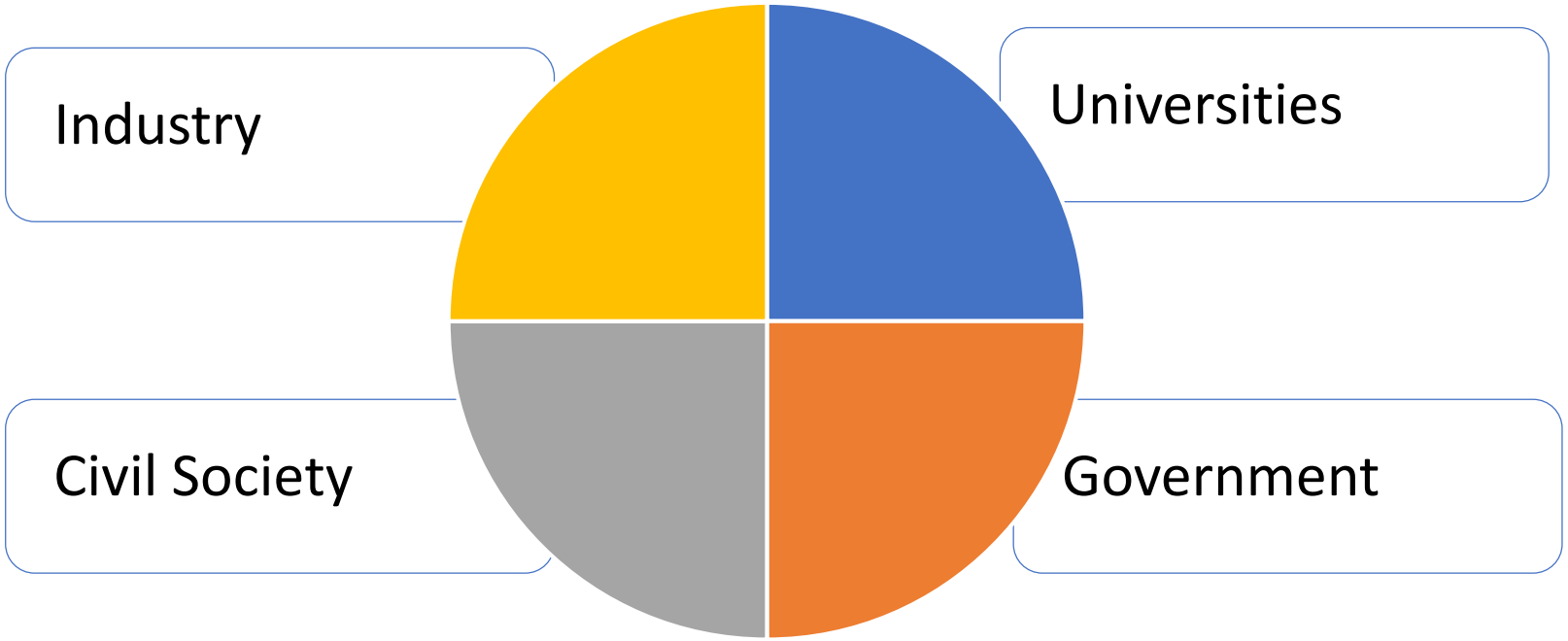
The ELSA Concept

1. **Participation.** Involvement in large scale (scientific) research into novel technological innovations.
2. **Anticipation.** Early identification of societal impact and signaling of potential controversies.
3. **Integration.** Stimulation of active involvement of citizens.
4. **Interdisciplinary.** Bridging the controversies between disciplines.

Zwart H., Nelis A. (2009), What is ELSA genomics? Science and Society Series on Convergence Research, EMBO Reports 10 (6), 1-5.

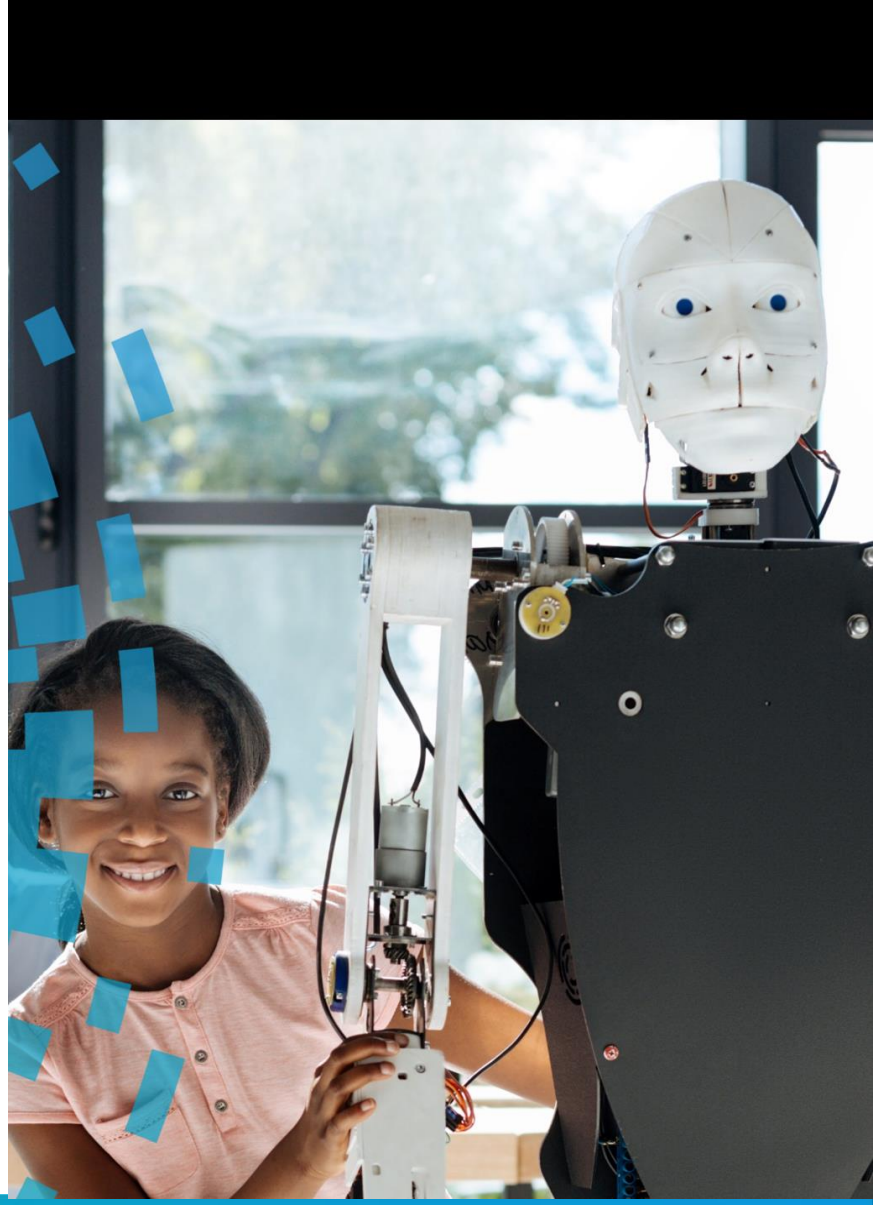
The Quadruple Helix Innovation Model

An innovation model that exploits the interaction between universities, industry, government, and civil society



Carayannis, Elias G.; Campbell, David F.J. (2009), "'Mode 3' and 'Quadruple Helix': toward a 21st century fractal innovation ecosystem", International Journal of Technology Management **46** (3/4): 201.

ELSALabs are a novel approach to the development of human centered AI solutions to hard societal problems

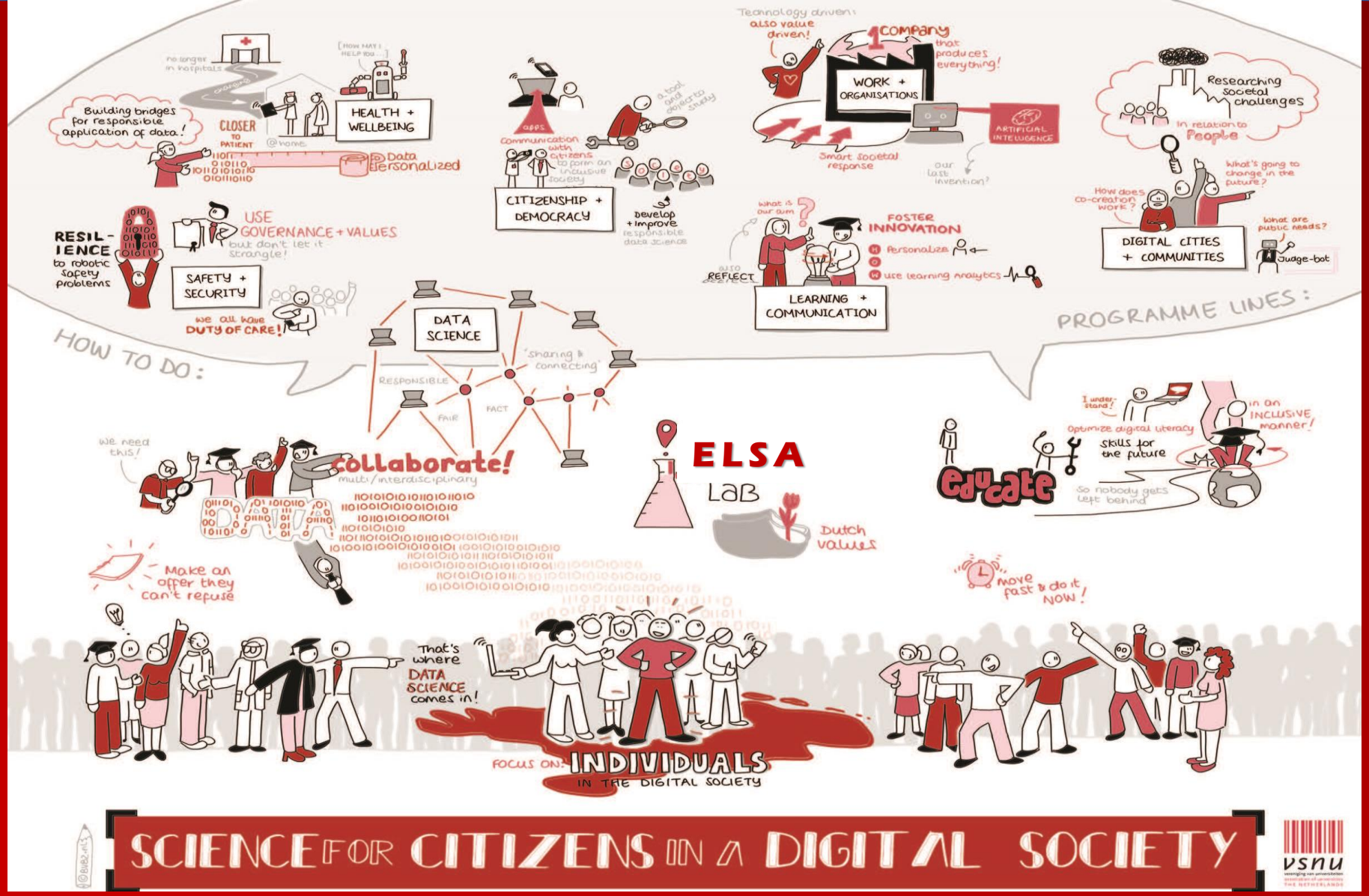


An ELSALabs reference framework

1. ELSALabs **address societally relevant issues** which are aimed at sustainable prosperity in the broadest sense, examples may come from the 17 UN Sustainable Development Goals.
2. ELSALabs **collect validated and documented insights** in a multi-stakeholder context.
3. Solutions are developed **applying design thinking methods** applying improvement cycles in real-life settings.
4. Insights and solutions are generated with **data heavy and algorithm savvy** techniques and methods.
5. In an ELSA Lab **all four innovation helix dimensions assume equal responsibility** for the development and coordination of the portfolio of activities carried out.
6. ELSA Labs apply a communication policy to **share the insights and solutions** with stakeholders and society at large.
7. ELSA Labs take a **responsibility to scale-up** solutions in order to impact society.

The ELSA concept as part of the AINed strategy

ELSA (Ethical Legal Societal Aspects):
Learning approach (use cases, co-creation) to AI ELSA and technology development in mutual coherence
Address needs of society at large, as well as partners in the quadruple helix
Medium to higher TRL level (if TRL is applicable at all)
Focus on multidisciplinary ELSA challenges to incentivize human-centric AI technology development
Lab-structure, 5 years, 7 fte, ~4 Meuro*
(2021): 10 ELSA labs
One call or selection procedure with NWO, SIA, or SMO and add first money stream to knowledge partners
Strengthen collaboration in quadruple helix and between AI-tech and SSH domain.
Governance through ELSA Board


Building bridges for responsible application of data!
CLOSER to PATIENT
HEALTH + WELLBEING
Data personalized
CITIZENSHIP + DEMOCRACY
Develop + improve responsible data science
Technology driven: also value driven!
1 COMPANY that produces everything!
WORK + ORGANISATIONS
ARTIFICIAL INTELLIGENCE
Smart societal response
Researching societal challenges
In relation to People
DIGITAL CITIES + COMMUNITIES
How does co-creation work?
What's going to change in the future?
What are public needs?
Judge-bot
RESILIENCE to robotic safety problems
USE GOVERNANCE + VALUES but don't let it strangle!
SAFETY + SECURITY
we all have DUTY OF CARE!
FOSTER INNOVATION
Personalize
Use learning analytics
LEARNING + COMMUNICATION
HOW TO DO:
PROGRAMME LINES:
DATA SCIENCE
RESPONSIBLE
FAIR
FACT
'sharing & connecting'
We need this!
collaborate!
multi/interdisciplinary
ELSA LAB
Dutch values
Make an offer they can't refuse
educate
Optimize digital literacy
skills for the future
So nobody gets left behind
in an INCLUSIVE manner!
move fast & do it NOW!
That's where DATA SCIENCE comes in!
FOCUS ON: INDIVIDUALS IN THE DIGITAL SOCIETY
SCIENCE FOR CITIZENS IN A DIGITAL SOCIETY
vsnu