

AI and society: governing the game changer

Dr. Clara Neppel

Vienna, 8 May 2019



“AI” to address societal challenges



AI in the media



["Tech Support," by R. Kikuo Johnson.](https://www.newyorker.com/magazine/2017/10/23)

<https://www.newyorker.com/magazine/2017/10/23>

“Success in creating effective AI, could be the biggest event in the history of our civilization. Or the worst. We just don't know. So we cannot know if we will be infinitely helped by AI, or ignored by it and side-lined, or conceivably destroyed by it”

Stephen Hawking, Nov 2017

“Doomsday AI will cause a nuclear war by 2040 that could DESTROY humanity, and there may be no way to prevent it”

<https://www.dailymail.co.uk/sciencetech/article-5651327/Doomsday-AI-cause-nuclear-war-2040-DESTROY-humanity.html>

AI for the media

AI generating content – images of society?



<https://www.nytimes.com/2019/02/05/business/media/artificial-intelligence-journalism-robots.html>

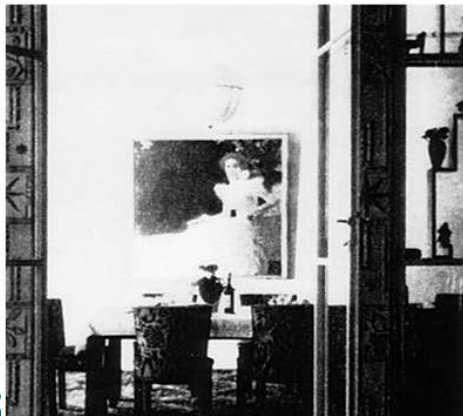


Sonja Knips, 1898, 24 years old



Sonja Knips, 1908, 34 years old

The influence of images



What images will AI offer us?

Exploring creative frontiers



AI self-portrait. Credit: IBM

What is truth?

(Deep) Fakes



Photo: University of Washington.



Illustration: iStockphoto

Algorithmic decision making

Transparency?



<https://hrtechweekly.com/2018/10/19/how-recommendation-engine-use-artificial-intelligence-for-e-commerce-industry/>

Digital Trust

Intelligent Digital Mesh – Privacy and Ethics

- Intelligent:
 - How AI is in virtually every existing technology, and creating entirely new categories.
- Digital:
 - Blending the digital and physical worlds to create an immersive world.
- Mesh:
 - Exploiting connections between expanding sets of people, businesses, devices, content and services.

Source: <https://www.gartner.com/smarterwithgartner/gartner-top-10-strategic-technology-trends-for-2019/>

Top 10 Strategic Technology Trends for 2019



[gartner.com/SmarterWithGartner](https://www.gartner.com/SmarterWithGartner)

Source: Gartner
© 2018 Gartner, Inc. and/or its affiliates. All rights reserved.
Gartner is a registered trademark of Gartner, Inc. or its affiliates. FR_499528

Gartner.

IEEE at a Glance

Global Reach

423,000+ Members



190+ Countries



117,000 Student Members



46 Technical Societies and Councils



Technical Breadth

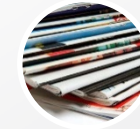
1,800+ Annual Conferences



4,000,000+ Technical Documents



190+ Top-cited Periodicals



1,300+ Active Standards



Social Impact

Global Public Policy



Global Humanitarian Efforts



Continuing Education & Certification



Ethics in Technology



The Fuel of IEEE

Advancing Technology for Humanity

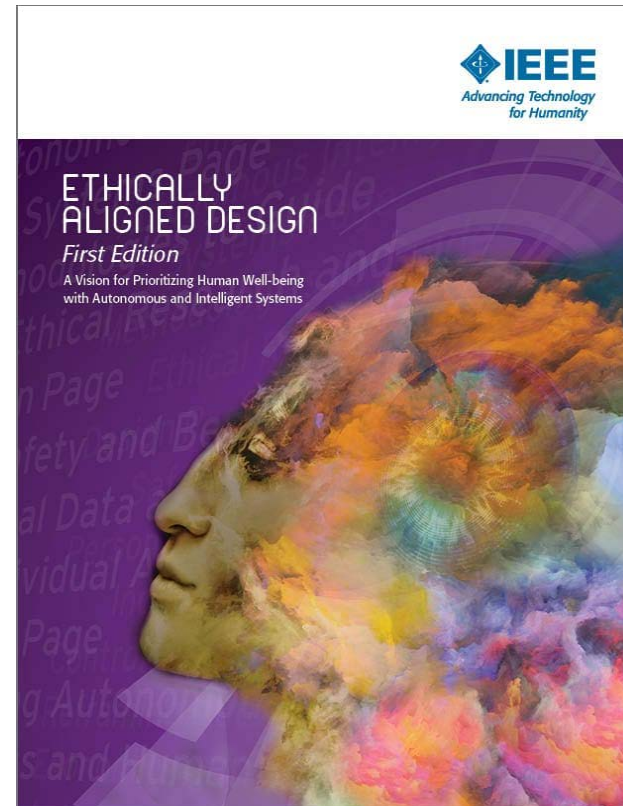
- ▶ Forward-thinking technology professionals coming together ...
 - to discover the next technological innovation,
 - to develop international standards,
 - to form communities,
 - to share research and educate,
 - in the spirit of collaboration.



Ethically Aligned Design, First Edition

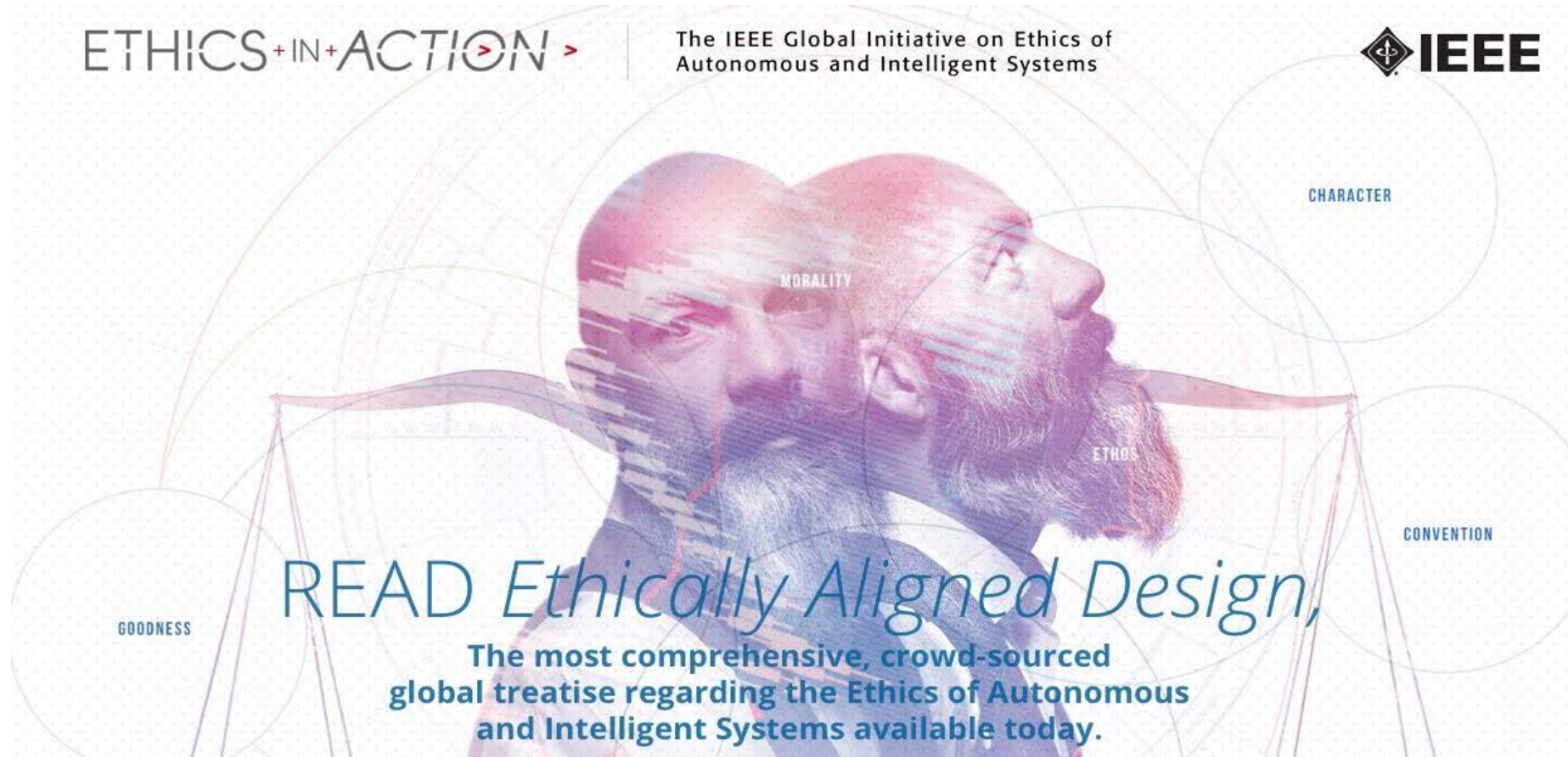
A key component of our work with The IEEE Global Initiative is the creation of our work, ***Ethically Aligned Design***.

Ethically Aligned Design, First Edition (EAD1e) represents the culmination of a three-year process guided from the bottom up since 2015 by the rigor and standards of the engineering profession and by a globally open and iterative methodology involving hundreds of global experts.



ETHICS+IN+ACTION >

The IEEE Global Initiative on Ethics of
Autonomous and Intelligent Systems



Main Website

<https://standards.ieee.org/industry-connections/ec/autonomous-systems.html>

Ethics in Action

<https://ethicsinaction.ieee.org>



Our Mission

To ensure every stakeholder involved in the design and development of autonomous and intelligent systems is educated, trained, and empowered to prioritize ethical considerations so that these technologies are advanced for the benefit of humanity.



Three Pillars

The Pillars of the *Ethically Aligned Design* Conceptual Framework fall broadly into three areas, reflecting anthropological, political, and technical aspects:

- **Universal Human Values**
- **Political Self-Determination and Data Agency**
- **Technical Dependability**

General Principles

- **Human Rights** - A/IS shall be created and operated to respect, promote, and protect internationally recognized human rights.
- **Well-being** - A/IS creators shall adopt increased human well-being as a primary success criterion for development.
- **Data Agency** - A/IS creators shall empower individuals with the ability to access and securely share their data, to maintain people's capacity to have control over their identity.
- **Effectiveness** - A/IS creators and operators shall provide evidence of the effectiveness and fitness for purpose of A/IS.
- ▶ **Transparency** - The basis of a particular A/IS decision should always be discoverable.
- ▶ **Accountability** - A/IS shall be created and operated to provide an unambiguous rationale for all decisions made.
- ▶ **Awareness of Misuse** - A/IS creators shall guard against all potential misuses and risks of A/IS in operation.
- ▶ **Competence** - A/IS creators shall specify and operators shall adhere to the knowledge and skill required for safe and effective operation.

IEEE P7000™ Standardization Projects

The IEEE P7000™ series of standards projects under development represents a unique addition to the collection of over 1,900 global IEEE standards and projects. Whereas more traditional standards have a focus on technology interoperability, functionality, safety, and trade facilitation, the IEEE P7000 series addresses specific issues at the intersection of technological and ethical considerations. Like its technical standards counterparts, the IEEE P7000 series empowers innovation across borders and enables societal benefit.

For more information: <https://ethicsinaction.ieee.org/#set-the-standard>

Set the Global Standards

Join an IEEE P7000™ Working Group
and Set the Global Standards for the future of Ethical A/IS.



IEEE P7000™ Standardization Projects

IEEE P7000™ - Model Process for Addressing Ethical Concerns During System Design

IEEE P7001™ - Transparency of Autonomous Systems

IEEE P7002™ - Data Privacy Process

IEEE P7003™ - Algorithmic Bias Considerations

IEEE P7004™ - Standard on Child and Student Data Governance

IEEE P7005™ - Standard on Employer Data Governance

IEEE P7006™ - Standard on Personal Data AI Agent Working Group

IEEE P7007™ - Ontological Standard for Ethically driven Robotics and Automation Systems

IEEE P7008™ - Standard for Ethically Driven Nudging for Robotic, Intelligent and Autonomous Systems

IEEE P7009™ - Standard for Fail-Safe Design of Autonomous and Semi-Autonomous Systems

IEEE P7010™ - Wellbeing Metrics Standard for Ethical Artificial Intelligence and Autonomous Systems

IEEE P7011™ - Standard for the Process of Identifying and Rating the Trustworthiness of News Sources

IEEE P7012™ - Standard for Machine Readable Personal Privacy Terms

IEEE P7013™ - Inclusion and Application Standards for Automated Facial Analysis Technology.



The Ethics Certification Program for Autonomous and Intelligent Systems (ECPAIS)

Developing metrics and processes towards the implementation of a certification methodology addressing transparency, accountability and algorithmic bias

The Ethics Certification Program for Autonomous and Intelligent Systems (ECPAIS) has the goal to create specifications for certification and marking processes that advance transparency, accountability, and reduction in algorithmic bias in autonomous and intelligent systems. ECPAIS intends to offer a process and define a series of marks by which organizations can seek certifications for their processes around the A/IS products, systems, and services they provide.

For More Information:

standards.ieee.org/industry-connections/ecpais.html



Artificial Intelligence and Ethics in Design (*Business Course*)

IEEE Innovation
at work

 **IEEE**

**New Course Program
Now Available: AI and
Ethics in Design**



Led by thought leaders in AI, philosophy and policy management fields, [Artificial Intelligence and Ethics in Design](#) is a five-course training for business professionals offered by IEEE Continuing Professional Education.

Created by thought leaders contributing to *Ethically Aligned Design* from The IEEE Global Initiative, the course will educate and empower professionals to practically implement ethical considerations when developing intelligent and autonomous products and services.

[Artificial Intelligence and Ethics in Design](#) is intended to help professionals leverage the power of AI while minimizing the risk. Participants will explore:

- ▶ Responsible Innovation in the Age of AI
- ▶ The Economic Advantage of Ethical Design for Business
- ▶ Values by Design in the Algorithmic Era
- ▶ The Nature of Nudging
- ▶ Ensuring Data Protection and Data Safety

Ethically Aligned Design University Consortium (EADUC)



Hagit Messer

*EADUC Lead Organizer
IEEE Global Initiative
Executive Committee*

Background:

Global engineering education today is largely focused on scientific and technological courses that are not intrinsically related to the analysis of human values overtly designed to positively increase human and environmental wellbeing.

The Consortium will focus on developing and promoting the issues, recommendations, and themes outlined in *Ethically Aligned Design* along with the A/IS Ethics oriented curriculum already being taught by Member Universities. The initial course, and eventual global curriculum, will be designed so it can be integrated into any program educating future engineers.

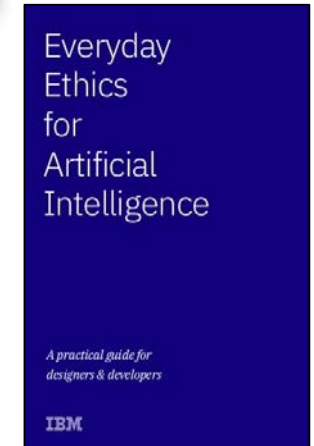
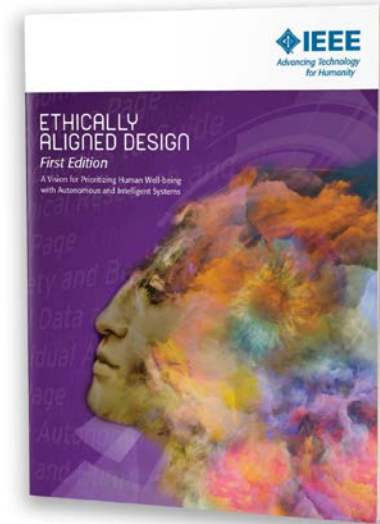
- Currently in development
- Will feature 15-20 Global Partner Universities
- Program to be launched in Fall, 2019
- Program to utilize *Ethically Aligned Design, First Edition*



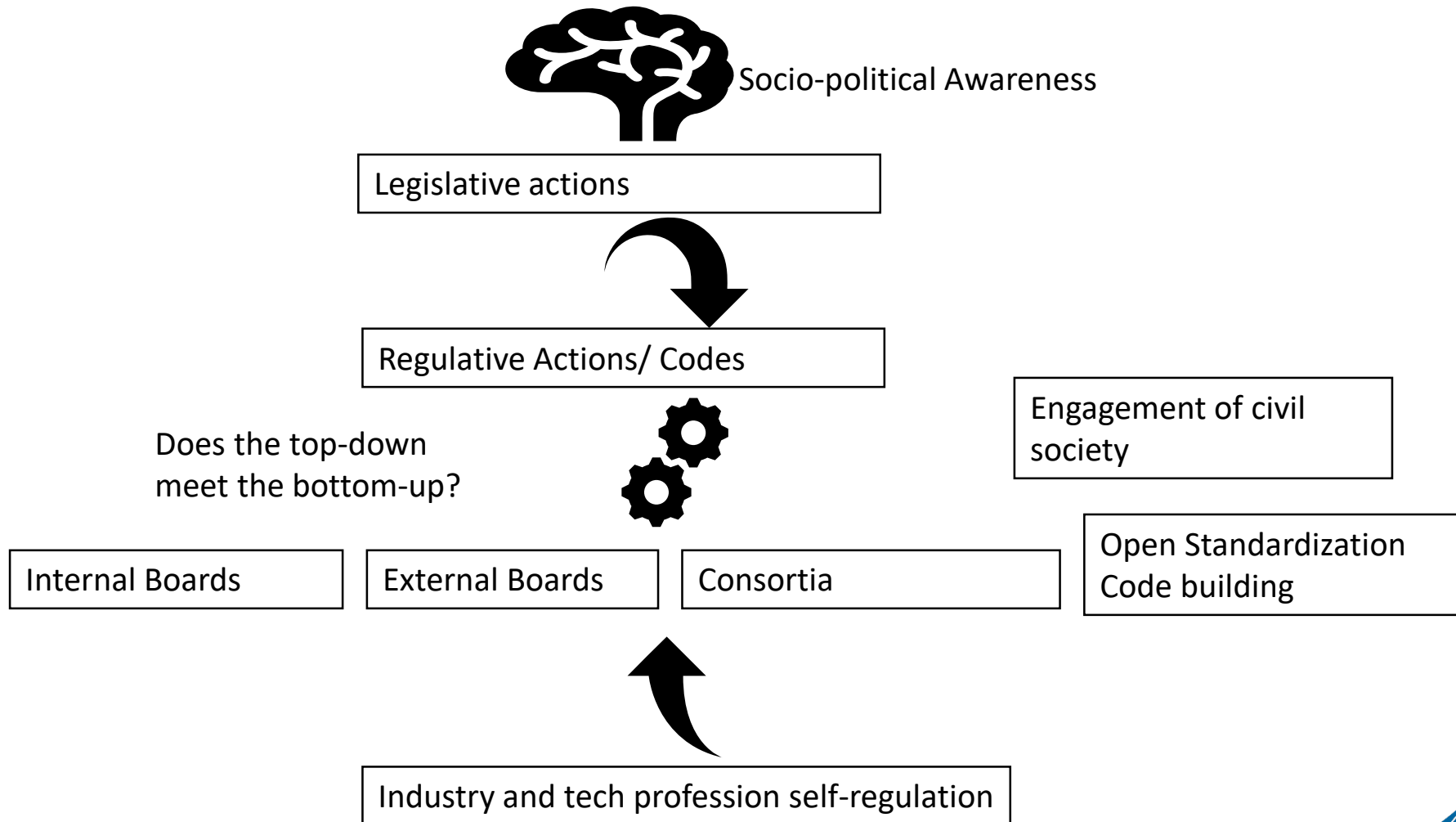
EAD for...Series

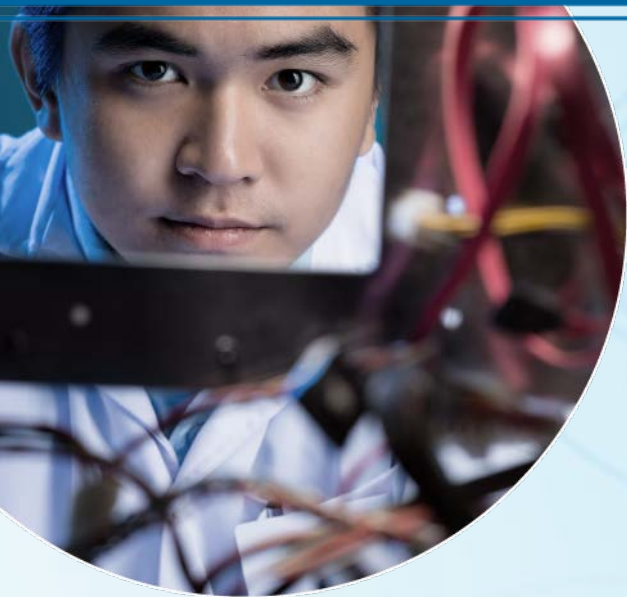
Designed to quickly educate specific business verticals and other key societal stakeholders on the relevant value of *Ethically Aligned Design*, the "EAD for..." Series currently has the following Committees working on short papers (8-12 pages) to be produced by Q3 or Q4 of 2019:

- ▶ **EAD for Industry**
- ▶ **EAD for Policymakers**
- ▶ **EAD for Law**
- ▶ **EAD for Health**



Working together towards Responsible AI





"No kidding? — you broke all
three laws of robotics?"

Dr Clara Neppel,
Senior Director, IEEE European Office, Vienna
c.neppel@ieee.org

