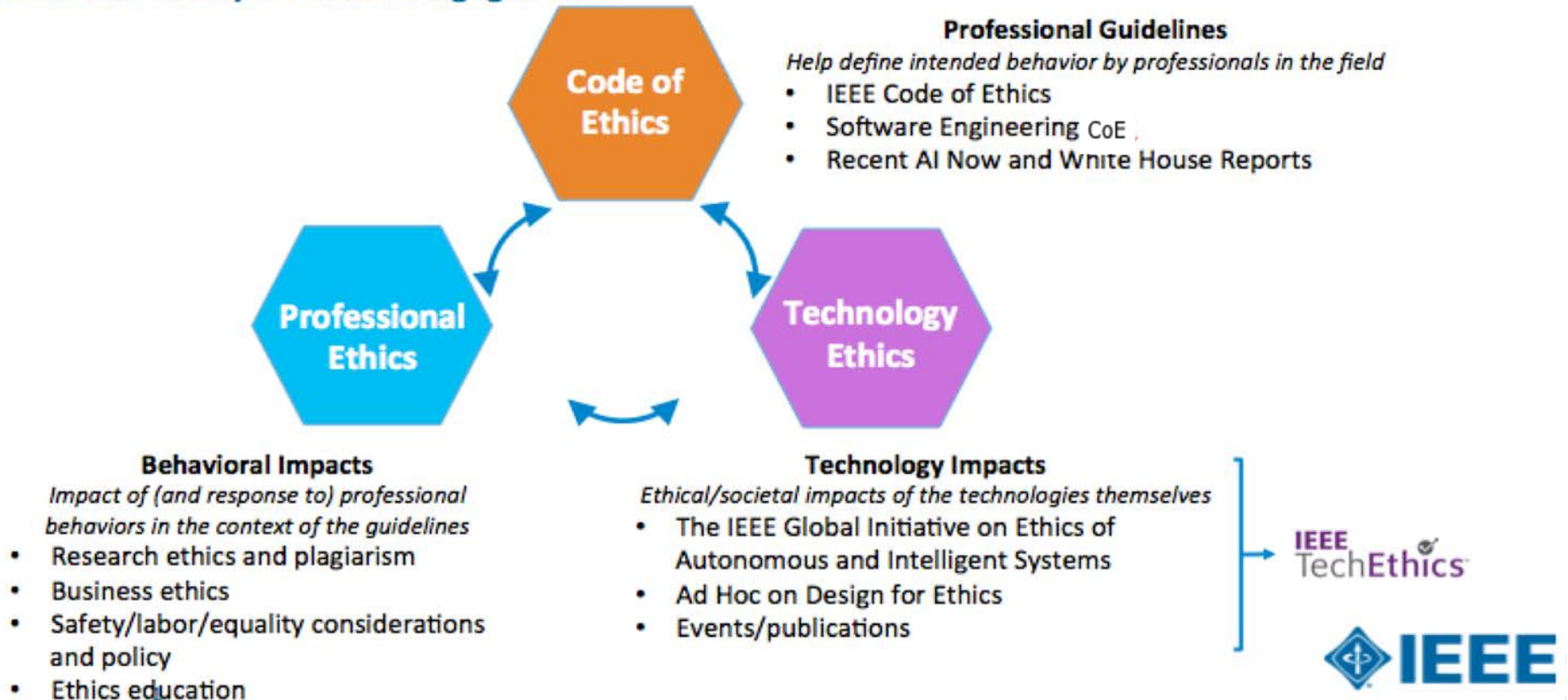# The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems

FairWare2018, 29 May 2018

Ethics Landscape
Overview and Examples of IEEE Engagement

Code of Ethics

Professional Ethics

Technology Ethics

Professional Guidelines
Help define intended behavior by professionals in the field
- IEEE Code of Ethics
- Software Engineering CoE
- Recent AI Now and White House Reports

Behavioral Impacts
Impact of (and response to) professional behaviors in the context of the guidelines
- Research ethics and plagiarism
- Business ethics
- Safety/labor/equality considerations and policy
- Ethics education

Technology Impacts
Ethical/societal impacts of the technologies themselves
- The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems
- Ad Hoc on Design for Ethics
- Events/publications

IEEE TechEthics

◆IEEE

The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems

Overview of The IEEE Global Initiative

Raja Chatila, Initiative Chair with Konstantinos Karachalios, IEEE-SA Managing Director

Kay-Firth Butterfield, Initiative Vice-Chair with Executive Committee Member Francesca Rossi

- Public launch in April 2016

- Currently more than 850 A/IS Ethics professionals involved from the US, EU, Australia, India, Africa, China, Korea, and Japan

- 13 Committees creating content plus groups supporting outreach, visibility, etc.

- Events featuring *Ethically Aligned Design* in India, Japan, China, World Economic Forum, EU Parliament, The Vatican and UN/ITU in Geneva

- *Ethically Aligned Design*, version 2 (EAD) launched December, 2017

- Thirteen approved Standards Working Groups inspired by EAD so far

- Final version of EAD to be launched in February, 2019

Ethically Aligned Design
A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems
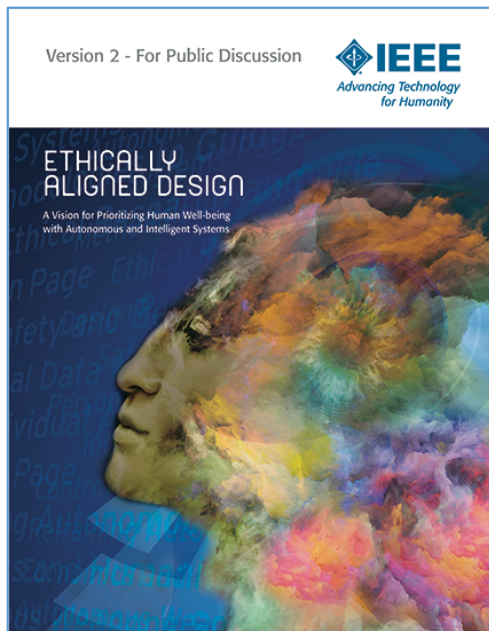
Version 2

- Launched December 2017 as a Request for Input

- Created by over 250 Global A/IS & Ethics professionals, in a bottom up, transparent, open and increasingly globally inclusive process

- Incorporates over 200 pages of feedback from public RFI and new Working Groups from China, Japan, Korea and more

- Thirteen Committees / Sections

- Contains **over one hundred twenty** key Issues and Candidate Recommendations

https://ethicsinaction.ieee.org/

Inspiration for IEEE P7000 Working Groups

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

**IEEE P7003™,
Standard for Algorithmic
Bias Considerations
Working Group**

IEEE Computer Society/Software & Systems
Engineering Standards Committee (C/S2ESC)

**Open invitation to join the P7003 working group
http://sites.ieee.org/sagroups-7003/**

Key questions when developing or deploying an algorithmic system

▸ Who will be affected?

▸ What are the decision/optimization criteria?

▸ How are these criteria justified?

▸ Are these justifications acceptable in the context where the system is used?

◆IEEE

P7003 foundational sections

▶ Taxonomy of Algorithmic Bias

▶ Person categorization and identifying affected population groups

▶ Legal frameworks related to Bias

▶ Psychology of Bias

P7003 algorithm development sections

▶ Algorithmic system design stages

▶ Assurance of representativeness of testing/training/validation data

▶ Evaluation of system outcomes

▶ Evaluation of algorithmic processing

▶ Assessment of resilience against external manipulation to Bias

▶ Documentation of criteria, scope and justifications of choices

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

◆IEEE

Global Influence *(Reports Mentioning EAD)*

- [AI in the UK: ready, willing and able?](#)

- [The Age of Artificial Intelligence: Towards a European Strategy for Human-Centric Machines](#)

- [Challenge Ahead: Integrating Robotics, Artificial Intelligence and 3D Printing Technologies into Canada's Healthcare Systems](#)

- [Artificial Intelligence Standardization Whitepaper (2018 Edition)](#)

HOUSE OF LORDS
Select Committee on Artificial Intelligence
Report of Session 2017–19

AI in the UK: ready, willing and able?

EPSC Strategic Notes

Issue 29
27 March 2018

European Political Strategy Centre

The Age of Artificial Intelligence
Towards a European Strategy for Human-Centric Machines

人工智能标准化白皮书
（2018 版）

ESI 中国电子技术标准化研究院
China Electronics Standardization Institute

CHALLENGE AHEAD
Integrating Robotics, Artificial Intelligence and 3D Printing Technologies Into Canada's Healthcare Systems

Standing Senate Committee on Social Affairs, Science and Technology
The Honourable Kelvin Kenneth Ogilvie, Chair
The Honourable Art Eggleton, P.C., Deputy Chair

October 2017

IEEE

Thank you!

Ansgar Koene
Chair of IEEE P7003 Standard for Algorithmic Bias Considerations
ansgar.koene@ieee.org

http://sites.ieee.org/sagroups-7003/
http://unbias.wp.horizon.ac.uk/

◆IEEE