

# Executive Summary

## Introduction

As the use and impact of autonomous and intelligent systems (A/IS) become pervasive, we need to establish societal and policy guidelines in order for such systems to remain human-centric, serving humanity's values and ethical principles. These systems have to behave in a way that is beneficial to people beyond reaching functional goals and addressing technical problems. This will allow for an elevated level of trust between people and technology that is needed for its fruitful, pervasive use in our daily lives.

To be able to contribute in a positive, non-dogmatic way, we, the techno-scientific communities, need to enhance our self-reflection, we need to have an open and honest debate around our imaginary, our sets of explicit or implicit values, our institutions, symbols and representations.

Eudaimonia, as elucidated by Aristotle, is a practice that defines human well-being as the highest virtue for a society. Translated roughly as "flourishing," the benefits of eudaimonia begin by conscious contemplation, where ethical considerations help us define how we wish to live.

Whether our ethical practices are Western (Aristotelian, Kantian), Eastern (Shinto, Confucian), African (Ubuntu), or from a different tradition, by creating autonomous and intelligent systems that explicitly honor inalienable human rights and the beneficial values of their users, we can prioritize the increase of human well-being as our metric for progress in the algorithmic age. Measuring and honoring the potential of holistic economic prosperity should become more important than pursuing one-dimensional goals like productivity increase or GDP growth.

## Executive Summary

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

**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.**

By “*stakeholder*” we mean anyone involved in the research, design, manufacture, or messaging around intelligent and autonomous systems, including universities, organizations, governments, and corporations making these technologies a reality for society.

Our goal is that *Ethically Aligned Design* will provide insights and recommendations that provide a key reference for the work of technologists in the related fields of science and technology in the coming years. To achieve this goal, in the current version of *Ethically Aligned Design* (EAD2v2), we identify pertinent “Issues” and “Candidate Recommendations” we hope will facilitate the emergence of national and global policies that align with these principles.

The IEEE Global Initiative brings together [several hundred participants](#) from six continents, who are thought leaders from academia, industry, civil society, policy and government in the related technical and humanistic disciplines to identify and find consensus on timely issues.

A second goal of The IEEE Global Initiative is to provide recommendations for IEEE Standards based on *Ethically Aligned Design*. *Ethically Aligned Design* (v1 and v2) and members of The IEEE Global Initiative are the inspiration behind the suite of IEEE P7000™ Standards Working Groups that are free and open for anyone to join.

## Executive Summary

**For more information or to join any Working Group, please click on the links below:**

- 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 for Transparent Employer Data Governance](#)
- IEEE P7006™ - [Standard for Personal Data Artificial Intelligence \(AI\) Agent](#)
- IEEE P7007™ - [Ontological Standard for Ethically Driven Robotics and Automation Systems](#)
- IEEE P7008™ - [Standard for Ethically Driven Nudging for Robotic, Intelligent, and Automation 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](#)

**Disclaimer:** While we have provided recommendations in this document, it should be understood these do not represent a position or the views of IEEE but the informed opinions of Committee members providing insights designed to provide expert directional guidance regarding A/IS. In no event shall IEEE or IEEE-SA Industry Connections Activity Members be liable for any errors or omissions, direct or otherwise, however caused, arising in any way out of the use of this work, regardless of whether such damage was foreseeable.

## Executive Summary

### Who We Are

The [IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems](#) (“The IEEE Global Initiative”) is a program of The Institute of Electrical and Electronics Engineers (“IEEE”), the world’s largest technical professional organization dedicated to advancing technology for the benefit of humanity with over 420,000 members in more than 160 countries.

The IEEE Global Initiative provides the opportunity to bring together [multiple voices in the related technological and scientific communities](#) to identify and find consensus on timely issues.

IEEE will make all versions of *Ethically Aligned Design* (EAD) available under the [Creative Commons Attribution-Non-Commercial 3.0 United States License](#).

Subject to the terms of that license, organizations or individuals can adopt aspects of this work at their discretion at any time. It is also expected that EAD content and subject matter will be selected for submission into formal IEEE processes, including for standards development.

The IEEE Global Initiative and EAD contribute to a broader effort at IEEE to foster open, broad, and inclusive conversation about ethics in technology, known as [the IEEE TechEthics™](#) program.

## Executive Summary

# Ethically Aligned Design v2 – Overview

## I. Purpose

Intelligent and autonomous technical systems are specifically designed to reduce human intervention in our day-to-day lives. In so doing, these new fields are raising concerns about their impact on individuals and societies. Current discussions include advocacy for the positive impact, as well as warnings, based on the potential harm to privacy, discrimination, loss of skills, economic impacts, security of critical infrastructure, and the long-term effects on social well-being. Because of their nature, the full benefit of these technologies will be attained only if they are aligned with our defined values and ethical principles. We must therefore establish frameworks to guide and inform dialogue and debate around the non-technical implications of these technologies.

## II. Goals

The ethical design, development, and implementation of these technologies should be guided by the following General Principles:

- **Human Rights:** Ensure they do not infringe on internationally recognized human rights
- **Well-being:** Prioritize metrics of well-being in their design and use
- **Accountability:** Ensure that their designers and operators are responsible and accountable

- **Transparency:** Ensure they operate in a transparent manner
- **Awareness of misuse:** Minimize the risks of their misuse

## III. Objectives

### Personal Data Rights and Individual Access Control

A fundamental need is that people have the right to define access and provide informed consent with respect to the use of their personal digital data. Individuals require mechanisms to help curate their unique identity and personal data in conjunction with policies and practices that make them explicitly aware of consequences resulting from the bundling or resale of their personal information.

### Well-being Promoted by Economic Effects

Through affordable and universal access to communications networks and the Internet, intelligent and autonomous technical systems can be made available to and benefit populations anywhere. They can significantly alter institutions and institutional relationships toward more human-centric structures and they can benefit humanitarian and development issues resulting in increased individual and societal well-being.

## Executive Summary

### Legal Frameworks for Accountability

The convergence of intelligent systems and robotics technologies has led to the development of systems with attributes that simulate those of human beings in terms of partial autonomy, ability to perform specific intellectual tasks, and may even have a human physical appearance. The issue of the legal status of complex intelligent and autonomous technical systems thus intertwines with broader legal questions regarding how to ensure accountability and allocate liability when such systems cause harm. Some examples of general frameworks to consider include the following:

- Intelligent and autonomous technical systems should be subject to the applicable regimes of property law
- Government and industry stakeholders should identify the types of decisions and operations that should never be delegated to such systems and adopt rules and standards that ensure effective human control over those decisions and how to allocate legal responsibility for harm caused by them

### Transparency and Individual Rights

Although self-improving algorithms and data analytics can enable the automation of decision-making impacting citizens, legal requirements mandate transparency, participation, and accuracy, including the following objectives:

- Parties, their lawyers, and courts must have reasonable access to all data and information generated and used by such systems employed by governments and other state authorities

- The logic and rules embedded in the system must be available to overseers thereof, if possible, and subject to risk assessments and rigorous testing
- The systems should generate audit trails recording the facts and law supporting decisions and they should be amenable to third-party verification
- The general public should know who is making or supporting ethical decisions of such systems through investment

### Policies for Education and Awareness

Effective policy addresses the protection and promotion of safety, privacy, intellectual property rights, human rights, and cybersecurity, as well as the public understanding of the potential impact of intelligent and autonomous technical systems on society. To ensure that they best serve the public interest, policies should:

- Support, promote, and enable internationally recognized legal norms
- Develop workforce expertise in related technologies
- Attain research and development leadership
- Regulate to ensure public safety and responsibility
- Educate the public on societal impacts of related technologies

## Executive Summary

### IV. Foundations

#### Classical Ethics

By drawing from over two thousand years' worth of classical ethics traditions, The IEEE Global Initiative explores established ethics systems, addressing both scientific and religious approaches, including secular philosophical traditions, to address human morality in the digital age. Through reviewing the philosophical foundations that define autonomy and ontology, The IEEE Global Initiative addresses the alleged potential for autonomous capacity of intelligent technical systems, morality in amoral systems, and asks whether decisions made by amoral systems can have moral consequences.

#### Well-being Metrics

For extended intelligence and automation based thereupon to provably advance a specific benefit for humanity, there needs to be clear indicators of that benefit. Common metrics of success include profit, occupational safety, and fiscal health. While important, these metrics fail to encompass the full spectrum of well-being for individuals or society. Psychological, social, and environmental factors matter. Well-being metrics capture such factors, allowing the benefits arising from technological progress to be more comprehensively evaluated, providing opportunities to test for unintended negative consequences that could diminish human well-being. Conversely, these metrics could help identify where intelligent technical systems would increase human well-being as well, providing new routes to societal and technological innovation.

#### Embedding Values into Autonomous Systems

If machines engage in human communities as quasi-autonomous agents, then those agents will be expected to follow the community's social and moral norms. Embedding norms in such systems requires a clear delineation of the community in which they are to be deployed. Further, even within a particular community, different types of technical embodiments will demand different sets of norms. The first step is to identify the norms of the specific community in which the systems are to be deployed and, in particular, norms relevant to the kinds of tasks that they are designed to perform.

#### Methodologies to Guide Ethical Research and Design

To create intelligent technical systems that enhance and extend human well-being and freedom, value-based design methodologies put human advancement at the core of development of technical systems, in concert with the recognition that machines should serve humans and not the other way around. System developers should employ value-based design methodologies in order to create sustainable systems that can be evaluated in terms of both social costs and also advantages that may increase economic value for organizations.



## Executive Summary

### V. Future Technology Concerns

#### Reframing Autonomous Weapons

Autonomous systems designed to cause physical harm have additional ethical dimensions as compared to both traditional weapons and/or autonomous systems not designed to cause harm. These ethical dimensions include, at least, the following:

- Ensuring meaningful human control of weapons systems
- Designing automated weapons with audit trails to help guarantee accountability and control
- Including adaptive and learning systems that can explain their reasoning and decisions to human operators in a transparent and understandable way
- Training responsible human operators of autonomous systems who are clearly identifiable
- Achieving behavior of autonomous functions that is predictable to their operators
- Ensuring that the creators of these technologies understanding the implications of their work
- Developing professional ethical codes to appropriately address the development of autonomous systems intended to cause harm

#### Safety and Beneficence of Alleged Artificial General Intelligence (AGI) and Artificial Superintelligence (ASI)

Similar to other powerful technologies, the development and use of intelligent and

potentially self-improving technical systems involves considerable risk, either because of misuse or poor design. However, according to some theories, as systems approach and surpass AGI, unanticipated or unintended system behavior will become increasingly dangerous and difficult to correct. It is likely that not all AGI-level architectures can be aligned with human interests, and as such, care should be taken to determine how different architectures will perform as they become more capable.

#### Affective Computing

Affect is a core aspect of intelligence. Drives and emotions such as anger, fear, and joy are often the foundations of actions throughout our life. To ensure that intelligent technical systems will be used to help humanity to the greatest extent possible in all contexts, artifacts participating in or facilitating human society should not cause harm either by amplifying or damping human emotional experience. Even the rudimentary versions of synthetic emotions already deployed in some systems impact how they are perceived by policy makers and the general public.

#### Mixed Reality

Mixed reality could alter our concepts of identity and reality as these technologies become more common in our work, education, social lives, and commercial transactions. The ability for real-time personalization of this mixed-reality world raises ethical questions concerning the rights of the individual and control over one's multifaceted identity, especially as the technology moves from headsets to more subtle and integrated sensory enhancements.



## Executive Summary

# Our Process

To ensure greatest cultural relevance and intellectual rigor in our work, The IEEE Global Initiative has been globally crowdsourcing feedback for Versions 1 and 2 of *Ethically Aligned Design*.

We released [Ethically Aligned Design Version 1](#) as a Request for Input on December of 2016 and received [over two hundred pages](#) of in-depth feedback about the draft. As a way to highlight insights inspired by the feedback we received, Sara Mattingly-Jordan of The IEEE Global Initiative also wrote the report, [Becoming a Leader in Global Ethics](#).

We are releasing *Ethically Aligned Design Version 2 (EADv2)* as a [Request for Input](#) once again to gain further insights about the eight original sections from EADv1, along with unique/new feedback for the five new sections included in EADv2.

### Next Steps

The IEEE Global Initiative is currently creating an organizational committee composed of representatives of all our Committees and IEEE P7000™ Working Groups to do the following in order to prepare the final version of *Ethically Aligned Design* to be released in 2019:

- Create criteria for Committees to vote on all “Candidate Recommendations” becoming “Recommendations” based on the General Principles of *Ethically Aligned Design* that are in accordance with the Mission Statement of The IEEE Global Initiative. This voting process will be based on the consensus-based protocols provided by IEEE-SA.
- Create a rigorous methodology to best incorporate feedback received from EADv1 and EADv2, working to holistically consider global and diversity-based considerations for content inclusion.
- Use the [glossary](#) we have produced as a key tool for synthesizing content for final version of EAD, unifying terms as much as possible.

### Final Version of *Ethically Aligned Design* – Format and Goals

The final version of *Ethically Aligned Design* will be made available in the following formats:

- **Handbook.** While specific formatting is still under consideration, the final version of *Ethically Aligned Design* will feature “Recommendations” (versus “Candidate Recommendations”) for all existing and future “Issues” voted on by Members of The IEEE Global Initiative. It is very likely the final version of EAD will not be broken into sections according to Committees (as with EADv1 and EADv2) but according to themes or principles to be decided on by the organizational committee mentioned above. While not an official IEEE position statement, “Recommendations” will be

## Executive Summary

created to be easily utilized by technologists and policy makers focusing on autonomous and intelligent systems design, usage, and governance.

- **Educational materials.** The IEEE Global Initiative would like to convert the handbook version of *Ethically Aligned Design* into an academically oriented book/educational materials. Evergreen in nature, these would be targeted to academics, engineers, and technologists looking for global guidance to be used in university, post-grad, or other educational settings where ethics in technology or the issues EAD comprises would be taught.

### Incorporating Feedback

While it was our intention to directly accept or review all feedback we received for EADv1, we were (happily) overwhelmed with the fantastic response we received. However, to most holistically include feedback from EADv1 and EADv2 into our overall process we have created a Glossary and are working to increase more global representation and diversity in our work. Specifically:

### Glossary

We received a great deal of feedback on the need for aligned recommendations for key terms in *Ethically Aligned Design*. To that end, we created a Glossary Committee and launched the first draft of our [Glossary](#) at the same time we released EADv2. Our goal is to refine our Glossary so that by mid-2018,

based on aggregated feedback to all sections of EAD (Versions 1 and 2), we can standardize definitions that reflect a global and holistic set of definitions to be implemented by all Committees in the final version of EAD.

### More Global Representation/Diversity

We received a great deal of feedback noting that EADv1 was fairly “Western” in its cultural orientation. This makes sense, as the initial 100 members working on EADv1 were largely from North America and the European Union. Since the release of EADv1, however, we have:

- Added members from China, Korea, Japan, Brazil, Mexico, the Russian Federation, Iran, Thailand, and Israel along with new people from the United States and the European Union. In addition to the 250 members of the Initiative, there are also now more than 400 global members in the IEEE P7000™ Working Groups that EAD inspired.
- Supported the members translating the Executive Summary of EADv1 into multiple languages.
- Added our new “Classical Ethics in A/IS” Committee.
- Created the [Becoming a Leader in Global Ethics](#) report.
- Commissioned a report from our newer global members about [the state of A/IS Ethics in their regions](#).
- Created an Outreach Committee to help identify and incorporate work being done

## Executive Summary

in A/IS ethics by women, people of color, students, and other groups representing the full spectrum of society that we are hoping to positively influence with our work. We are currently working with members of [The Reboot Retreat](#), [AI4ALL](#), and other leaders within IEEE to help us ensure that The IEEE Global Initiative and the final version of *Ethically Aligned Design* are as holistically representative and relevant as possible.

### Terminology Update

There is no need to use the term *artificial intelligence* in order to conceptualize and speak of technologies and systems that are meant to extend our human intelligence or be used in robotics applications. For this reason, we use the term, *autonomous and intelligent systems* (or *A/IS*) in the course of our work. We chose to use this phrase encapsulating multiple fields (machine learning, intelligent systems

engineering, robotics, etc.) throughout *Ethically Aligned Design*, Version 2 to ensure the broadest application of ethical considerations in the design of these technologies as possible.

### How the Document Was Prepared

This document was prepared using an open, collaborative, and consensus building approach, following the processes of the [Industry Connections program](#), a program of the IEEE Standards Association.

Industry Connections facilitates collaboration among organizations and individuals as they hone and refine their thinking on emerging technology issues, helping to incubate potential new standards activities and standards-related products and services.

## Executive Summary

# How to Cite Ethically Aligned Design

Please cite Version 2 of *Ethically Aligned Design* in the following manner:

The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems. *Ethically*

*Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems*, Version 2. IEEE, 2017. [http://standards.ieee.org/develop/indconn/ec/autonomous\\_systems.html](http://standards.ieee.org/develop/indconn/ec/autonomous_systems.html).

## Executive Summary

# Our Appreciation

We wish to thank our Executive Committee and Chair of The IEEE Global Initiative:

### Executive Committee Officers

Raja Chatila, *Chair*

Kay Firth-Butterfield, *Vice-Chair*

John C. Havens, *Executive Director*

### Executive Committee Members

Dr. Greg Adamson, Ronald C. Arkin, Virginia Dignum, Danit Gal, Philip Hall, Malavika Jayaram, Sven Koenig, Raj Madhavan, Richard Mallah, Hagit Messer Yaron, AJung Moon, Monique Morrow, Francesca Rossi, Alan Winfield

### Committee Chairs

- **General Principles:** Alan Winfield and Mark Halverson
- **Embedding Values into Autonomous Intelligent Systems:** Francesca Rossi and Bertram F. Malle
- **Methodologies to Guide Ethical Research and Design:** Raja Chatila and Corinne J.N. Cath
- **Safety and Beneficence of Artificial General Intelligence (AGI) and Artificial Superintelligence (ASI):** Malo Bourgon and Richard Mallah

- **Personal Data and Individual Access Control:** Katryna Dow and John C. Havens
- **Reframing Autonomous Weapons Systems:** Peter Asaro
- **Economics/Humanitarian Issues:** Kay Firth-Butterfield and Raj Madhavan
- **Law:** Kay Firth-Butterfield and Derek Jinks
- **Affective Computing:** Ronald C. Arkin and Joanna J. Bryson
- **Classical Ethics in A/IS:** Jared Bielby
- **Policy:** Kay Firth-Butterfield and Philip Hall
- **Mixed Reality:** Monique Morrow and Jay Iorio
- **Well-being:** Laura Musikanski and John C. Havens
- **Drafting:** Kay Firth-Butterfield and Deven Desai
- **Industry:** Virginia Dignum and Malavika Jayaram
- **Communications:** Leanne Seeto and Mark Halverson
- **Glossary:** Sara M. Jordan
- **Outreach:** Danit Gal

## Executive Summary

We wish to express our appreciation for the reports, organizations, and individuals that have contributed research and insights helping to increase awareness around ethical issues in the realm of intelligent and autonomous systems, including (*but not limited to, and in no particular order*):

### Reports

[The Future of Life Institute's Asilomar AI Principles](#), [The AI Now 2017 Report](#), [Human Rights in the Robot Age Report](#) from [The Rathenau Instituut](#), [Report of COMEST on Robotics Ethics](#) from [UNESCO](#), [The European Parliament's Recommendations to the Commission on Civil Law Rules on Robotics, Artificial intelligence – The Consequences of Artificial Intelligence on the \(Digital\) Single Market, Production, Consumption, Employment and Society](#) report from the *European Economic and Social Committee* (Rapporteur: Catelijne MULLER), OECD's report, [Going Digital: Making the Transformation Work for Growth and Well-Being](#), [USACM's Statement on Algorithmic Transparency and Accountability](#), [Guide to the Ethical Design and Application of Robots and Robotic Systems](#) (British Standards Institute),

[Japan's Basic Rules for AI Research](#), [Éthique de la Recherche en Robotique](#) (CERNA), [Charta der Digitalen Grundrechte der Europäischen Union \(Charter of the Digital Fundamental Rights of the European Union\)](#), Telecommunications Research Laboratory, "AI Network Kent kai Kaigi H kokusho 2016: AI Network no Eiky to Risk – Chiren Shakai (WINS) no Jitsugen ni Muketa Kadai" (AIネットワーク化検討会議 報告書2016 公表 - 「AIネットワーク化の影響とリスク - 智連社会 (WINS(ウインズ)) の実現に向けた課題 -」) [[The Conference on Networking among AIs Report \(2016\): Impacts and Risks of AI Networking Issues for the Realization of Wisdom Network Society](#), (WINS)], Japanese Ministry of Internal Affairs and Communications, The Information Technology Industry Council's [AI Policy Principles](#), Intel's [Artificial Intelligence – The Public Policy Opportunity](#), IEEE European Public Policy Initiative's position statement, [Artificial Intelligence: Calling on Policy Makers to Take a Leading Role in Setting a Long Term AI Strategy](#), IEEE-USA's position statement on [Artificial Intelligence Research, Development and Regulation](#), The IEEE Global Initiative's [Prioritizing Human Well-being in the Age of Artificial Intelligence](#).

## Executive Summary

### Organizations

[The Association for the Advancement of Artificial Intelligence](#) and their formative work on [AI Ethics](#), [The Future of Life Institute](#), [The Partnership on AI to Benefit People and Society](#), [The Foundation for Responsible Robotics, AI & Society](#), [Machine Intelligence Research Institute](#), [The International Center for Information Ethics](#), [The African Center of Excellence for Information Ethics](#), [The 4TU Center for Ethics and Technology](#), [The Center for the Study of Existential Risk](#), [The Leverhulme Center for the Future of Intelligence](#), [The Future of Humanity Institute](#), [The Japanese Society for Artificial Intelligence](#), [The Association for Computing Machinery](#), [Future Advocacy](#), [ACM Special Interest Group on Artificial Intelligence](#), [The World Economic Forum's Global Future Council of Artificial Intelligence and Robotics](#), [The Digital Asia Hub](#), [The AI Initiative](#), [The Open Roboethics Institute](#), [The Dalai Lama Center for Ethics and Transformative Values at MIT](#), [The Ethics Initiative at MIT Media Lab](#), [The IEEE-USA Government Relations Council Artificial Intelligence Committee](#), [The IEEE Robotics and Automation Society Committee on Robot Ethics](#), [The IEEE Robotics and Automation Society](#), [The IEEE Society on Social Implications of Technology](#), [The IEEE Computer Society](#), [The IEEE Computational Intelligence Society](#), [The IEEE Systems, Man and Cybernetics Society](#), [The IEEE Symbiotic Autonomous Systems Initiative](#).

### People

We would like to warmly recognize the leadership and constant support of The IEEE Global Initiative by Dr. Ing. Konstantinos Karachalios, Managing Director of the IEEE Standards Association and a member of the IEEE Management Council.

We would especially like to thank Eileen M. Lach, the IEEE General Counsel and Chief Compliance Officer, who invested her time and expertise in fully reviewing this entire document, with the heartfelt conviction that there is a pressing need to focus the global community on highlighting ethical considerations in the development of autonomous and intelligent systems.

Special thanks to Dr. Peter S. Brooks for his contributions to the Overview of EADv2.

### Thank You to Our Members and IEEE Team

Our progress and the ongoing positive influence of this work is due to the volunteer experts serving on our Committees and IEEE P7000™ Standards Working Groups, along with the IEEE staff who support our efforts. Thank you for your dedication toward defining, designing, and inspiring the ethical PRINCIPLES and STANDARDS that will ensure that intelligent and autonomous systems and the technologies associated therewith will positively benefit humanity.



## Executive Summary

# Disclaimers

*Ethically Aligned Design* is not a code of conduct or a professional code of ethics. Engineers and technologists have well-established codes, and we wish to respectfully recognize the formative precedents surrounding issues of ethics and safety and the professional values these codes represent. These codes provide the broad framework for the more focused domain addressed in this document, and it is our hope that the inclusive, consensus-building process around its design will contribute unique value to technologists and society as a whole.

This document is also not a position, or policy statement, or formal report of IEEE or any other organization with which is affiliated. It is intended to be a working reference tool created in an inclusive process by those in the relevant scientific and engineering communities prioritizing ethical considerations in their work.

### **A Note on Affiliations Regarding Members of The Initiative**

The language and views expressed in *Ethically Aligned Design* reflect the individuals who created content for each section of this document. The language and views expressed in this document do not necessarily reflect the positions taken by the universities or organizations to which these individuals belong, and should in no way be considered any form of endorsement, implied or otherwise, from these institutions.

This is the second version of *Ethically Aligned Design*. Where individuals are listed in a Committee it indicates only that they are Members of that Committee. Committee Members may not have achieved final concurrence on content in this document because of its versioning format and the concurrence-building process of The IEEE Global Initiative. Content listed by Members in this or future versions is not an endorsement, implied or otherwise, until formally stated as such.

### **A Note Regarding Candidate Recommendations in This Document**

*Ethically Aligned Design* is being created via multiple versions that are being iterated over the course of two to three years. The IEEE Global Initiative is following a specific concurrence-building process where members contributing content are proposing candidate recommendations so as not to imply these are final recommendations at this time.

### **Our Membership**

The IEEE Global Initiative currently has more than 250 experts from all but one continent involved in our work, and we are eager for new voices and perspectives to join our work.

## Executive Summary

### Copyright, Trademarks, and Disclaimers

IEEE believes in good faith that the information in this publication is accurate as of its publication date; such information is subject to change without notice. IEEE is not responsible for any inadvertent errors.

The Institute of Electrical and Electronics Engineers, Incorporated

3 Park Avenue, New York, NY 10016-5997, USA

Copyright © 2017 by The Institute of Electrical and Electronics Engineers, Incorporated

Published December 2017

Printed in the United States of America.

IEEE is a registered trademark owned by The Institute of Electrical and Electronics Engineers, Incorporated.

PDF: ISBN 978-0-7381-xxxx-x STDVxxxxx

Print: ISBN 978-0-7381-xxxx-x STDPDVxxxxx

IEEE prohibits discrimination, harassment, and bullying. For more information, visit <http://www.ieee.org/web/aboutus/whatis/policies/p9-26.html>

This work is made available under the [Creative Commons Attribution Non-Commercial License](https://creativecommons.org/licenses/by-nc/3.0/).

To order IEEE Press Publications, call 1-800-678-IEEE.

Find IEEE standards and standards-related product listings at: [standards.ieee.org](http://standards.ieee.org)

### Notice and Disclaimer of Liability Concerning the Use of IEEE-SA Industry Connections Documents

This IEEE Standards Association (“IEEE-SA”) Industry Connections publication (“Work”) is not a consensus standard document. Specifically, this document is NOT AN IEEE STANDARD. Information contained in this Work has been created by, or obtained from, sources believed to be reliable, and reviewed by members of the IEEE-SA Industry Connections activity that produced this Work. IEEE and the IEEE-SA Industry Connections activity members expressly disclaim all warranties (express, implied, and statutory) related to this Work, including, but not limited to, the warranties of: merchantability; fitness for a particular purpose; non-infringement; quality, accuracy, effectiveness, currency, or completeness of the Work or content within the Work. In addition, IEEE and the IEEE-SA Industry Connections activity members disclaim any and all conditions relating to: results; and workmanlike effort. This IEEE-SA Industry Connections document is supplied “AS IS” and “WITH ALL FAULTS.”

Although the IEEE-SA Industry Connections activity members who have created this Work believe that the information and guidance given in this Work serve as an enhancement to users, all persons must rely upon their own skill and judgment when making use of it. IN NO EVENT SHALL IEEE OR IEEE-SA INDUSTRY CONNECTIONS ACTIVITY MEMBERS BE LIABLE FOR ANY ERRORS OR OMISSIONS OR DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO: PROCUREMENT OF

## Executive Summary

SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS WORK, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE AND REGARDLESS OF WHETHER SUCH DAMAGE WAS FORESEEABLE.

Further, information contained in this Work may be protected by intellectual property rights held by third parties or organizations, and the use of this information may require the user to negotiate with any such rights holders in order to legally acquire the rights to do so, and such rights holders may refuse to grant such rights. Attention is also called to the possibility that implementation of any or all of this Work may require use of subject matter covered by patent rights. By publication of this Work, no position is taken by IEEE with respect to the existence

or validity of any patent rights in connection therewith. IEEE is not responsible for identifying patent rights for which a license may be required, or for conducting inquiries into the legal validity or scope of patents claims. Users are expressly advised that determination of the validity of any patent rights, and the risk of infringement of such rights, is entirely their own responsibility. No commitment to grant licenses under patent rights on a reasonable or non-discriminatory basis has been sought or received from any rights holder. The policies and procedures under which this document was created can be viewed at [standards.ieee.org/about/sasb/iccom/](https://standards.ieee.org/about/sasb/iccom/).

This Work is published with the understanding that IEEE and the IEEE-SA Industry Connections activity members are supplying information through this Work, not attempting to render engineering or other professional services. If such services are required, the assistance of an appropriate professional should be sought. IEEE is not responsible for the statements and opinions advanced in this Work.