IEEE SA STANDARDS ASSOCIATION

September 13, 2021 - Submitted electronically via regulations.gov

Mr. Mark Przybocki
U.S. National Institute of Standards and Technology
100 Bureau Drive
Gaithersburg, Maryland 20899

Subject: Comments of the IEEE Standards Association to the National Institute of Standards and Technology on an Artificial Intelligence Risk Management Framework [Docket Number: 210726-0151]

The IEEE Standards Association appreciates the opportunity to submit comments to the National Institute of Standards and Technology in response to the Request for Information to help inform, refine, and guide the development of NIST's AI Risk Management Framework.

We acknowledge NIST for its efforts to gain input on the Framework and appreciate the opportunity to contribute. Further, as the world's largest technical professional organization dedicated to advancing technology for the benefit of humanity and a globally recognized standards developing organization (SDO) grounded in an open, inclusive, transparent, and consensus-building process, we appreciate that NIST is conducting this open call for input.

Our recommendations have been developed by a broad group of global experts from industry, academia, and elsewhere, involved in the risk management and ethical dimensions of AI.

Please find below IEEE SA's input to the questions posed in the Request for Input. As NIST moves forward on the development of its Risk Management Framework, IEEE SA stands ready to share our expertise on this matter. For questions on this submission, or to discuss the recommendations further, please do not hesitate to contact us.

Sincerely,
Kristin Little
Senior Manager, Public Affairs
IEEE Standards Association
k.little@ieee.org

About the IEEE SA
The IEEE Standards Association (IEEE SA) is a globally recognized standards-setting body within IEEE. We develop consensus standards through an open process that engages industry and brings together a broad stakeholder community. IEEE standards set specifications and best practices based on current scientific and technological knowledge. IEEE SA has a portfolio of over 1,500 active standards and over 650 standards under development. As a collaborative body, we liaise and coordinate with many standards organizations from around the world, including international, regional, and national standards bodies, as well as with industry organizations.

About IEEE
IEEE is the world's largest technical professional organisation dedicated to advancing technology for the benefit of humanity. IEEE and its members inspire a global community to innovate for a better tomorrow through highly-cited publications, conferences, technology standards, and professional and educational activities. IEEE is the trusted "voice" for engineering, computing, and technology information around the globe.

***

IEEE offers the following comments in response to the 11 of the 12 topics suggested:

1. The greatest challenges in improving how AI actors manage AI-related risks—where "manage" means identify, assess, prioritize, respond to, or communicate those risks;

Comment
One of the challenges is to reach the widest audience with clear information. To that end, communication of the risks should be conducted in the most accessible manner for any implementer and end user impacted by the outcome of the algorithm. An approach to indicate clearly and consistently as to what this communication should incorporate at a base level could be useful. **IEEE SA sugggests the development of a clear and simple base level communication of risks.**
Contrary to the traditional approaches to RM, the AI risks are largely ethical, hence societal, and one of **the greatest challenges is to identify the appropriate/related key stakeholders for consultation.**
Another challenge is adopting a multidisciplinary approach to understanding and capturing the PESTEL (political, economic, social, technological, environmental, and legal) dimensions of AI-related risks. More specifically, it is difficult to understand how to frame AI and its scope. It is also difficult to involve multidisciplinary AI actors at different stages of AI development (e.g., programmers, systems engineers, lawyers, ethicists, social scientists).
Many actors are involved through time, contributing to making assessments over time difficult. Many AI Actors are involved in the design, deployment, and use of AI. However even more actors are involved in the assessment of risk from their interaction with AI systems. Therefore **a framework of risk evaluation and a common language is needed** to make such an assessment and be able to measure risk change or evolution over time.
Some of the additional challenges in AI-related risks is that **risk analysis is rooted in the assessment of explainable and reproducible situations, whereas in AI explainability might**

not always be feasible, and the in-the-field continuous evolution and optimization of automated learning does not guarantee the reproducibility of the results. Yet these situations can pose risks.
A **lack of understanding of AI can result in undue suspicion in nonexistent risks, or resistance to its adoption** at the population level. This reassurance and non-rejection needs to be managed through communication as well.
Anticipating future contexts of application, and articulating and envisioning AI uses that do not yet exist in order to identify potential risks presents a particular challenge in improving how AI actors manage AI-related risks.
Establishing protective and preventive obligations to recognize that AI systems have limitations and therefore may not be appropriate to deploy in particular contexts. These limitations include susceptibility to system errors, 'black box' scenarios, unpredictability of machine-learning systems, and an inability to engage in 'pure reasoning'.
A bias to recognize in any AI oriented risk is **understanding and taking into account the cultural context that forms the basis for legal norms followed**. AI principles regarding an overall design approach and an approach to ethics is critical. For additional perspective, please see the Classical Ethics chapter of the IEEE publication, "Ethically Aligned Design: Prioritizing Human Wellbeing with Autonomous and Intelligent Systems."

2. How organizations currently define and manage characteristics of AI trustworthiness and whether there are important characteristics which should be considered in the Framework besides: Accuracy, explainability and interpretability, reliability, privacy, robustness, safety, security (resilience), and mitigation of harmful bias, or harmful outcomes from misuse of the AI;

Comment
There are several levels of risk to consider in relation to the approach. **At the meta level is the need to ensure that human-centricity is integrated within the systems overview with clear expectations of assessment evaluations** to indicate levels of measured risk inherent to the system's environment, the development of the system, and relative assessment in relation to risk association with the AI implementation. **Adopting a 'human-centric and lifecycle approach' to identifying and managing risks.** This approach would focus on the human user and their rights. It would involve prioritizing human wants, needs, and values through user awareness, protecting human rights, recognizing non-deterministic influences on decision-making, and respecting AI limitations, and applying

these throughout the AI system lifecycle (i.e., design, development, manufacture, deployment, and post-deployment phases).
Risk management approaches will have their role in relation to this work. **What needs to be clearly understood as a part of this effort is how the risk efforts are categorized and how the grey areas are managed in relation to this framework. Developments that initially instantiate at a minimal risk but over time engage with higher risk environments need to be considered.** Providing explicit clarity would be necessary.
The narrative on trust has become confused. Some references consider "trust" in the machine-to-machine sense, of maintaining a score, trusting those devices that meet the score, not trusting (isolating) those that do not. Others suggest "trust" to be a desirable state of confidence in AI technology when it appears warranted, in order to overcome a fear of the intention of machinery. A third category simply wants us to "trust" AI and anywhere it will take us. These three approaches are intermingled in the trust literature. For example an AI may display all of the characteristics in the question, and still not deserve our trust. **Much more multi-disciplinary examination of "trust" is needed.**
There are a number of emergent ethical properties of products, services, and systems. These are *contributory* to "trustworthiness," however, and no single property can ensure societal and stakeholder acceptability. **The ethical values at stake are highly community-, context-, and culture-driven and require a consultative and comprehensive approach to identification, ranking, fostering, and harm mitigation.**
Trustworthiness, through its many properties, is often narrowly-defined within the scope of the tasks being executed by a given AI module implementation. It has been demonstrated that **AI systems can develop correlations independent of the task at hand** (e.g., identify race constructs from x-rays), and **it is important to look into the unintended consequences of these, both at the level of the task itself or in conjunction with other (possibly AI) implementations leveraging it.** One could call this *information leakage*.
The additional element with respect to **Trust** is that it **is heavily contextualized**. **What one might see as trustworthy may not be seen the same by another.** In order to deliver on the characteristics noted, **it is important to consider trust also from the perspective of the end user *and* the operator of the system as that is where the relationship unfolds between expectation and realization.**
One element to consider for trustworthiness is volume. **Is the AI system behaving similarly in handling large numbers of cases or situations, or is it narrowly tested for accuracy and explainability**, **etc.**?
Awareness of human dignity is a characteristic that IEEE SA suggests should be considered in the Framework. A comprehensive Framework **should consider the design, development, and**

deployment of AI systems that do not undermine or lead to loss of human dignity through coercion, manipulation, deception, or loss of autonomy. In this context, human dignity means innate human worthiness that relates to the status of human beings as agents with rational capacity to exercise reasoning, judgement, and choice, and which entitles them to respectful treatment.
IEEE SA suggests that the Framework should consider **the issue of legal responsibility for AI - designing, developing, and deploying AI systems so that there is legal responsibility present throughout the system life cycle, and attributable to a broad spectrum of human agents** (e.g., designers, programmers, engineers, manufacturers, operators, and system owners).
Trust must involve genuine agency from a user, wherein agency is derived via explicit and meaningful consent, where "meaningful" is not only about a personal understanding of what a technology *does*, and with whom the information is shared, for example. Building trust also needs to happen through disclosure of information concerning possible effects of the technology. Forthright sharing of this type of information represents an opportunity to build trust. **To engender trust, communication must be bidirectional** and consider how a person interacts / communicates with technology. Agency and trust do not occur where trust is assumed or the communication designed to start trust creation only comes from designers who have not set up models of participatory design.
AI-related risks need to be incorporated into an enterprise's risk management framework. For example, it should be a category for consideration when thinking about program risk, project risk, technology risk, etc.

3. How organizations currently define and manage principles of AI trustworthiness and whether there are important principles which should be considered in the Framework besides: Transparency, fairness, and accountability;

Comment
Privacy and ethical governance Apart from the current focus on Transparency, Accountability, and Fairness (lack of unacceptable bias), there are other societal concerns in so far as **respect for privacy, ethical governance** (capability and maturity of enterprises in this context), and potentially other emergent properties of AI solutions. Matters of **technical dependability, safety, security, reliability,** etc., are also relevant but subsidiary to the principal ethical concerns.

Power and empowerment **Power and empowerment** is currently not generally found in the AI trustworthiness literature. An AI could be transparent, fair, and accountable, yet create a circumstance in which its possession significantly disempowers a population. For example, a facial recognition system may treat everyone fairly (accurately identifying people of every race using transparent algorithms), yet the outcome may be vast social control by a small minority. In that case, the safety and trustworthiness of the technology is questionable, even though it may be "transparent" and "accountable."
Trustworthiness of the organizations controlling the AI AI is currently adopted as an optimization technology (cost, performance, etc.) and trustworthiness is mostly considered only along the lines of the task it performs. Yet its reliance on big data in many of its implementations induces a position of control with implications at the level of the *trustworthiness of the organizations controlling the AI* (and the data) both in terms of the announced intent of the AI application, and in terms of its possible evolution, which might call for *a principle of precaution*.
Human dignity **An important element to consider would be the design, development, and deployment of AI systems that do not undermine or lead to loss of human dignity through coercion, manipulation, deception, or loss of autonomy.** In this context, human dignity means *innate human worthiness* that relates to the status of human beings as agents with rational capacity to exercise reasoning, judgement, and choice, and which entitles them to respectful treatment.
Legal responsibility for AI It would be important to consider designing, developing, and deploying AI systems so that there is **legal responsibility present throughout the system life cycle, and attributable to a broad spectrum of human agents** (e.g., designers, programmers, engineers, manufacturers, operators, and system owners).

4. The extent to which AI risks are incorporated into different organizations' overarching enterprise risk management—including, but not limited to, the management of risks related to cybersecurity, privacy, and safety;

Comment
The ethical risks are primarily context sensitive and cannot be broadly managed unilaterally by the enterprise without suitable and sufficient consultation with the principal stakeholders. Cybersecurity and safety have other standards and business drivers and

invariably exist in the enterprise RMFs but ethics is unlikely to be treated systematically.
Currently, large organisations do not consistently incorporate "AI adoption risk" into their risk registers. This type of risk is likely closest in character to "outsourcing" or "offshoring" risk, in that it has a significant impact on employment and employment conditions. Many organizations lack incentives to declare such risks.
Understanding of AI potential, mechanisms, and limitations, have not become pervasive enough to replace the association of AI with an "intelligent entity" in the minds of the non-experts, and evaluate its intrinsic risks. Its perceived infallibility, combined with human behavioral traits of influence from and delegation to machines, can make resulting situations riskier compared to a human-only mode.
IEEE SA suggests that the extent to which AI risks are incorporated into different organizations' overarching enterprise risk management could be determined by the existence of: 1) **minimum assessment requirements** comprising a) sector risks, including web-based global operation risks, b) potential harms/adverse impacts from AI systems, c) end-user needs, and d) supply chain awareness 2) **overall legal compliance** taking account of cross-jurisdictional reach and sector-specific AI system operations 3) **early warning systems** or messages for dynamic or learning systems 4) **'black box' scenario protocols** 5) **user pre-use information and opt out mechanism** 6) **emergency response mechanism for random and systematic errors**

5. Standards, frameworks, models, methodologies, tools, guidelines and best practices, and principles to identify, assess, prioritize, mitigate, or communicate AI risk and whether any currently meet the minimum attributes described above;

Comment
IEEE SA would like to bring to NIST's attention for consultation the following IEEE resources, as they are prominent in this space and provide rules, standards, and processes for ethical assurance:

IEEE's portfolio of initiatives under the Global Initiative for Ethics of AIS comprised of

- The publication *Ethically Aligned Design*. The chapter on wellbeing might be of particular interest.
- The P70xx suite of standards, especially IEEE 7010-2020 - IEEE Recommended Practice for Assessing the Impact of Autonomous and Intelligent Systems on Human Well-Being to provide a framework of KPIs for both "risk" and "success" for AI in design and in use, which means AI can't be considered "free of risk" where "success" is factored largely by economic or financial metrics or metrics of exponential growth in isolation.
- The Ethics Certification Program for Autonomous and Intelligent Systems (ECPAIS),

6. How current regulatory or regulatory reporting requirements (e.g., local, state, national, international) relate to the use of AI standards, frameworks, models, methodologies, tools, guidelines and best practices, and principles;

Comment

IEEE SA observes that current regulatory controls and reporting requirements fall short of protecting from the potential impact of AI on social values and human rights. That said, ongoing efforts continue. To name a few:

- EU GDPR;
- the Council of Europe Modernised Convention for the Protection of Individuals with Regard to Automatic Processing of Personal Data (Convention 108+))
- bilateral and multilateral treaties;
- 2019 OECD Recommendation on AI;
- 2019 G20 Human-Centred AI Principles;
- 2019 EU Ethics Guidelines for Trustworthy AI;
- 2019 Recommendations of UN SG's HLP on Digital Cooperation;
- 2019 IEEE EAD; 2015 UN Sustainable Development Goals,
- the European draft legislation AI Act
- Voluntary ethical certification under IEEE ECPAIS
- The California Consumer Privacy Act (CCPA) (https://oag.ca.gov/privacy/ccpa).
- Estonia's Data standards / practices frame actual user / citizen agency (https://www.dataguidance.com/notes/estonia-data-protection-overview).

7. AI risk management standards, frameworks, models, methodologies, tools, guidelines and best practices, principles, and practices which NIST should consider to ensure that the AI RMF aligns with and supports other efforts;

Comment
IEEE SA recommends that NIST consult the European Union's AI regulation and the other resources listed in Q6, along with the following IEEE resources (listed in Q5): *IEEE's portfolio of initiatives under the Global Initiative for Ethics of AIS comprised of* ● *The publication Ethically Aligned Design. The chapter on wellbeing might be of particular interest.* ● *The P70xx suite of standards, especially IEEE 7010-2020 - IEEE Recommended Practice for Assessing the Impact of Autonomous and Intelligent Systems on Human Well-Being to provide a framework of KPIs for both "risk" and "success" for AI in design and in use, which means AI can't be considered "free of risk" where "success" is factored largely by economic or financial metrics or metrics of exponential growth in isolation.* ● *The Ethics Certification Program for Autonomous and Intelligent Systems (ECPAIS),*

8. How organizations take into account benefits and issues related to inclusiveness in AI design, development, use and evaluation—and how AI design and development may be carried out in a way that reduces or manages the risk of potential negative impact on individuals, groups, and society.

Comment
The following will be important to AI design and development: ● Adopting a multidisciplinary approach to understanding and capturing the PESTEL (political, economic, social, technological, environmental, and legal) dimensions of AI-related risks. ● Anticipating future contexts of application, and articulating and envisioning AI uses that do not yet exist in order to identify potential risks. ● Establishing protective and preventive obligations to recognize that AI systems have limitations and therefore may not be appropriate to deploy in particular contexts. These limitations include susceptibility to system errors, 'black box' scenarios, unpredictability of machine-learning systems, and an inability to engage in 'pure

reasoning'.

- Adopting a 'human-centric and lifecycle approach' to identifying and managing risks. This approach would focus on the human user and their rights. It would involve prioritizing human wants, needs, and values through user awareness, protecting human rights, recognizing non-deterministic influences on decision-making, and respecting AI limitations, and applying these throughout the AI system lifecycle (i.e., design, development, manufacture, deployment, and post-deployment phases).

- Human dignity-aware AI - designing, developing, and deploying AI systems that do not undermine or lead to loss of human dignity through coercion, manipulation, deception, or loss of autonomy. In this context, human dignity means innate human worthiness that relates to the status of human beings as agents with rational capacity to exercise reasoning, judgement, and choice, and which entitles them to respectful treatment.

- Legal responsibility for AI - designing, developing, and deploying AI systems so that there is legal responsibility present throughout the system life cycle, and attributable to a broad spectrum of human agents (e.g., designers, programmers, engineers, manufacturers, operators, and system owners).

A model process for addressing ethical issues in system design is depicted in IEEE 7000-2021 - IEEE Standard Model Process for Addressing Ethical Concerns during System Design, which will be published soon. This applies to all sizes of enterprises, public and private.

IEEE SA would also like to emphasize that it is critical to have a "societal impact assessment" lens for AI to help avoid the risk of prioritizing economic or financial metrics or growth in isolation. The "risks" encountered when designing for GDP measures are different from those considered when environment and human wellbeing are also taken into account. For precedents here, please see:

- The IEEE standard: IEEE 7010-2020 - IEEE Recommended Practice for Assessing the Impact of Autonomous and Intelligent Systems on Human Well-Being
- The IEEE paper: Measuring What Matters in the Era of Global Warming and Algorithmic Promises
- The work of New Zealand and their "wellbeing economy."

9. The appropriateness of the attributes NIST has developed for the AI Risk Management Framework. (See above, "AI RMF Development and Attributes");

Comment
--

10. Effective ways to structure the Framework to achieve the desired goals, including, but not limited to, integrating AI risk management processes with organizational processes for developing products and services for better outcomes in terms of trustworthiness and management of AI risks. Respondents are asked to identify any current models which would be effective. These could include—but are not limited to—the NIST Cybersecurity Framework or Privacy Framework, which focus on outcomes, functions, categories and subcategories and also offer options for developing profiles reflecting current and desired approaches as well as tiers to describe degree of framework implementation; and

Comment
Please see the response to Q8, copied below for reference:
The following will be important to AI design and development: ● *Adopting a multidisciplinary approach to understanding and capturing the PESTEL (political, economic, social, technological, environmental, and legal) dimensions of AI-related risks.* ● *Anticipating future contexts of application, and articulating and envisioning AI uses that do not yet exist in order to identify potential risks.* ● *Establishing protective and preventive obligations to recognize that AI systems have limitations and therefore may not be appropriate to deploy in particular contexts. These limitations include susceptibility to system errors, 'black box' scenarios, unpredictability of machine-learning systems, and an inability to engage in 'pure reasoning'.* ● *Adopting a 'human-centric and lifecycle approach' to identifying and managing risks. This approach would focus on the human user and their rights. It would involve prioritizing human wants, needs, and values through user awareness, protecting human rights, recognizing non-deterministic influences on decision-making, and respecting AI limitations, and applying these throughout the AI system lifecycle (i.e., design, development, manufacture, deployment, and post-deployment phases).* ● *Human dignity-aware AI - designing, developing, and deploying AI systems that do not*

undermine or lead to loss of human dignity through coercion, manipulation, deception, or loss of autonomy. In this context, human dignity means innate human worthiness that relates to the status of human beings as agents with rational capacity to exercise reasoning, judgement, and choice, and which entitles them to respectful treatment. ● *Legal responsibility for AI - designing, developing, and deploying AI systems so that there is legal responsibility present throughout the system life cycle, and attributable to a broad spectrum of human agents (e.g., designers, programmers, engineers, manufacturers, operators, and system owners).*
A model process for addressing ethical issues in system design is depicted in IEEE 7000-2021 - IEEE Standard Model Process for Addressing Ethical Concerns during System Design. This applies to all sizes of enterprises, public and private.
IEEE SA would also like to emphasize that it is critical to have a "societal impact assessment" lens for AI to help avoid the risk of prioritizing economic or financial metrics or growth in isolation. The "risks" encountered when designing for GDP measures are different from those considered when environment and human wellbeing are also taken into account. For precedents here, please see: ● *IEEE standard: IEEE 7010-2020 - IEEE Recommended Practice for Assessing the Impact of Autonomous and Intelligent Systems on Human Well-Being* ● *IEEE paper: Measuring What Matters in the Era of Global Warming and Algorithmic Promises* ● *The work of New Zealand and their "wellbeing economy."*

11. How the Framework could be developed to advance the recruitment, hiring, development, and retention of a knowledgeable and skilled workforce necessary to perform AI-related functions within organizations.

Comment
We refer NIST once again to the comments listed in Q8, copied below for convenience.
The following will be important to AI design and development: ● *Adopting a multidisciplinary approach to understanding and capturing the PESTEL (political, economic, social, technological, environmental, and legal) dimensions of AI-related risks.*

- *Anticipating future contexts of application, and articulating and envisioning AI uses that do not yet exist in order to identify potential risks.*

- *Establishing protective and preventive obligations to recognize that AI systems have limitations and therefore may not be appropriate to deploy in particular contexts. These limitations include susceptibility to system errors, 'black box' scenarios, unpredictability of machine-learning systems, and an inability to engage in 'pure reasoning'.*

- *Adopting a 'human-centric and lifecycle approach' to identifying and managing risks. This approach would focus on the human user and their rights. It would involve prioritizing human wants, needs, and values through user awareness, protecting human rights, recognizing non-deterministic influences on decision-making, and respecting AI limitations, and applying these throughout the AI system lifecycle (i.e., design, development, manufacture, deployment, and post-deployment phases).*

- *Human dignity-aware AI - designing, developing, and deploying AI systems that do not undermine or lead to loss of human dignity through coercion, manipulation, deception, or loss of autonomy. In this context, human dignity means innate human worthiness that relates to the status of human beings as agents with rational capacity to exercise reasoning, judgement, and choice, and which entitles them to respectful treatment.*

- *Legal responsibility for AI - designing, developing, and deploying AI systems so that there is legal responsibility present throughout the system life cycle, and attributable to a broad spectrum of human agents (e.g., designers, programmers, engineers, manufacturers, operators, and system owners).*

A model process for addressing ethical issues in system design is depicted in IEEE 7000-2021 - IEEE Standard Model Process for Addressing Ethical Concerns during System Design. This applies to all sizes of enterprises, public and private.

IEEE SA would also like to emphasize that it is critical to have a "societal impact assessment" lens for AI to help avoid the risk of prioritizing economic or financial metrics or growth in isolation. The "risks" encountered when designing for GDP measures are different from those considered when environment and human wellbeing are also taken into account. For precedents here, please see:

- *IEEE standard: IEEE 7010-2020 - IEEE Recommended Practice for Assessing the Impact of Autonomous and Intelligent Systems on Human Well-Being*
- *IEEE paper: Measuring What Matters in the Era of Global Warming and Algorithmic Promises*

• *The work of New Zealand and their "wellbeing economy."*
In addition, IEEE SA recommends that the RMF include a governance dimension in which competence plays a major role in assurance. Mature enterprises include evaluation, assessment, development and management of human competencies.

12. The extent to which the Framework should include governance issues, including but not limited to make up of design and development teams, monitoring and evaluation, and grievance and redress.

Comment
IEEE SA suggests including governance issues in the Framework. For more information, please refer to IEEE P2863 - Recommended Practice for Organizational Governance of Artificial Intelligence.
While design, manufacturing, and deployment teams generally have a stronger role and influence in the context of the design and deployment of products and services, IEEE SA would like to underscore the importance of identifying and managing societal harms and benefits at the whole enterprise level.
All those listed in Q8 comment.
The following will be important to AI design and development: • *Adopting a multidisciplinary approach to understanding and capturing the PESTEL (political, economic, social, technological, environmental, and legal) dimensions of AI-related risks.* • *Anticipating future contexts of application, and articulating and envisioning AI uses that do not yet exist in order to identify potential risks.* • *Establishing protective and preventive obligations to recognize that AI systems have limitations and therefore may not be appropriate to deploy in particular contexts. These limitations include susceptibility to system errors, 'black box' scenarios, unpredictability of machine-learning systems, and an inability to engage in 'pure reasoning'.* • *Adopting a 'human-centric and lifecycle approach' to identifying and managing risks.*

This approach would focus on the human user and their rights. It would involve prioritizing human wants, needs, and values through user awareness, protecting human rights, recognizing non-deterministic influences on decision-making, and respecting AI limitations, and applying these throughout the AI system lifecycle (i.e., design, development, manufacture, deployment, and post-deployment phases).

- *Human dignity-aware AI - designing, developing, and deploying AI systems that do not undermine or lead to loss of human dignity through coercion, manipulation, deception, or loss of autonomy. In this context, human dignity means innate human worthiness that relates to the status of human beings as agents with rational capacity to exercise reasoning, judgement, and choice, and which entitles them to respectful treatment.*

- *Legal responsibility for AI - designing, developing, and deploying AI systems so that there is legal responsibility present throughout the system life cycle, and attributable to a broad spectrum of human agents (e.g., designers, programmers, engineers, manufacturers, operators, and system owners).*

A model process for addressing ethical issues in system design is depicted in IEEE 7000-2021 - IEEE Standard Model Process for Addressing Ethical Concerns during System Design, which will be published soon. This applies to all sizes of enterprises, public and private.

IEEE SA would also like to emphasize that it is important to have a "societal impact assessment" lens for AI to help avoid the risk of prioritizing economic or financial metrics or growth in isolation. The "risks" encountered when designing for GDP measures are different from those considered when environment and human wellbeing are also taken into account. For precedents here, please see:

- *IEEE standard: IEEE 7010-2020 - IEEE Recommended Practice for Assessing the Impact of Autonomous and Intelligent Systems on Human Well-Being*
- *IEEE paper: Measuring What Matters in the Era of Global Warming and Algorithmic Promises*
- *The work of New Zealand and their "wellbeing economy."*