

# Initiatives for AI Ethics: Formulation of Fujitsu Group AI Commitment

● Tsuneo Nakata ● Tatsuki Araki ● Satoshi Tsuchiya ● Yuri Nakao  
● Aisha Naseer ● Junichi Arahori ● Takahiko Yamamoto

While AI technology provides significant benefits, downsides of AI have also been reported, including AI that uses biased data for learning makes unfair decisions. To prevent further advancements in the technology from causing serious side effects, AI ethics are being discussed and many organizations and companies have recently announced new guidelines regarding AI. When Fujitsu released FUJITSU Human Centric AI Zinrai in 2015, we proposed “collaborative, human centric AI” and have paid attention to the ethical aspects of AI. In March 2019, we announced the Fujitsu Group AI Commitment that expresses our approach to AI ethics in a concrete and easy-to-understand manner based on the results of over 30 years of R&D and social implementation of AI. The promises of the Fujitsu Group AI Commitment were formulated in cooperation with AI4People, Europe’s expert forum, and are in accordance with the AI ethics principles that the forum advocates. They are Fujitsu’s message to its major stakeholders, including customers, people, society, shareholders, and employees. This paper presents the social trends related to AI ethics and describes the basic concept of Fujitsu’s AI Commitment.

## 1. Introduction

Progress in AI technologies has led to the development of new methods for prediction and optimization based on the use of massive amounts of data and has brought about major benefits to human society. Along with these benefits, however, are also reports of the negative aspects of AI, such as generating fake news or crime-prediction discriminating against particular ethnic and racial groups—problems that are seen to only increasingly worsen and affect us more directly. It is imperative, therefore, to come up with measures to prevent the adverse effects of AI while ensuring the conveniences it brings. The approach to formulating such preventive measures emphasizes the ethical aspects of AI and involves discussing relevant issues from various perspectives. Many organizations and companies have recently announced their own AI ethics guidelines that summarize their ethical standards when carrying out the research, development, provision, and operation of AI technologies.

For its part, Fujitsu announced the FUJITSU Human Centric AI Zinrai in 2015 under the concept of “collaborative, human centric AI” by paying proper attention to

the ethical aspects of AI. Furthermore, in March 2019, we announced the “Fujitsu Group AI Commitment,”<sup>(1)</sup> which summarizes Fujitsu’s AI ethics promises based on its more than 30 years of experience in R&D and social implementation of AI technologies and expresses its approach to AI ethics in a concrete and easy-to-understand manner.

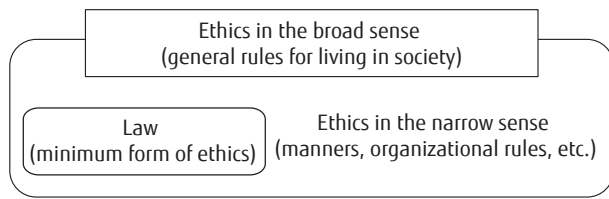
This paper describes the background and social trends surrounding the increasing focus on AI ethics and explains the basic promises of the Fujitsu Group AI Commitment.

## 2. AI ethics

This section defines “AI ethics” and describes the social trends of initiatives related to them.

### 2.1 What is AI ethics?

The field of medicine has led the discussion in defining ethics in science and technology. Modern medical ethics have been discussed and practiced in medicine for more than 40 years. The Japanese Nursing Association (JNA), from its position of being a provider of health care services to patients, has defined ethics



**Figure 1**  
Relationship between law and ethics.

as the “general rules for living in society”<sup>2)</sup> and has laid down a code of ethics for the nursing profession. Laws, which are formal “rules for living,” are referred to as the “minimum form of ethics” that carry legal force, and are therefore included in its broad meaning. Manners and organizational rules are also widely thought of as part of the meaning of ethics. **Figure 1** illustrates the relationship of law and ethics, wherein laws are included as part of the broad meaning of ethics, and manners and organizational rules constitute the narrow definition of ethics.<sup>3)</sup>

AI ethics are the norms that AI must conform with in offering services to humans and society. In this context, AI can be likened to nurses providing health care services. Therefore, in this paper, we define AI ethics as the “rules that AI must follow in collaborating with humans,” in reference to the nursing ethics definition. This definition corresponds to the broad definition of ethics, which also includes future rules and regulations.

## 2.2 Growing interest in AI ethics

One of the earliest references to AI ethics is in “The Three Laws of Robotics” proposed in a novel written in 1950.<sup>4)</sup> AI ethics became subject to wide attention again when the “trolley problem”<sup>5)</sup> came up in the field of autonomous driving. Since then, starting around 2015, a succession of incidents related to ethical problems caused by AI has been reported. Some examples are a chatbot that posted inflammatory and offensive tweets,<sup>6)</sup> facial recognition software that tagged people as gorillas,<sup>7)</sup> an AI recruiting tool that showed bias against women,<sup>8)</sup> and an algorithm used to predict relapse into criminal behavior that showed racial discrimination.<sup>9)</sup> These incidents have exposed the dangers of AI to undermine human dignity and fairness and have pointed out the need for creating ethical AI systems.

## 2.3 Trends in the formulation of AI ethics principles

Despite the expectations from society for AI systems to behave ethically, technologies that enable “ethical AI” have yet to be established. The continued progress in the development and application of AI technologies, however, is leading to heightened risks of serious AI-related ethical problems.

To address this issue, more and more companies and organizations are formulating their own AI ethics principles for the fair conduct of the research, development, provision, and operation of AI technologies. **Table 1** lists the major AI principles pertaining to ethics published thus far.

These guidelines are not legally enforceable and have been formulated primarily to express the particular organization’s position. In Europe, in particular, regulations based on these guidelines are being drafted and are foreseen to become enacted within the next few years.

## 2.4 AI ethics and business

Compliance with AI ethics goes beyond technologies and the statement of principles; it can also serve as a condition for conducting business. In September 2018, the Canadian government began stipulating compliance with a set of criteria based on AI ethics as a requirement in the procurement of AI services, solutions, and products.<sup>10)</sup> The stipulation requires suppliers to explicitly state how they implement ethical AI practices in their frameworks, guidelines, assessment tools, test environments, etc. This move is foreseen to become a global trend, pointing out the need for AI developers and providers to fast-track their efforts in complying with AI ethics principles.

## 3. Fujitsu’s initiatives in AI ethics

In 2009, Fujitsu officially defined its vision for how ICT should bring about a safer, more prosperous, and sustainable world as the “Human Centric Intelligent Society.” In 2015, when the company consolidated the AI technologies it has nurtured for many years into the “FUJITSU Human Centric AI Zinrai” brand,<sup>11)</sup> Fujitsu emphasized “Human Centric AI” as the key feature of the brand. Human Centric AI means that the use of Fujitsu’s AI technologies is premised on respect for human dignity, on providing benefits that contribute to human

**Table 1**  
Major AI principles pertaining to ethics.

Organization/Company	Name	Year	Number of principles included
Partnership on AI	Tenets <sup>15)</sup>	2016	8
Future of Life Institute	Asilomar AI Principles <sup>16)</sup>	2017	23
IEEE (Institute of Electrical and Electronics Engineers)	Ethically Aligned Design <sup>17)</sup>	2017	5
An Initiative of Université de Montréal (Canada)	Montréal Declaration Responsible AI <sup>18)</sup>	2018	10
Microsoft	Future Computed	2018	6
European Group on Ethics in Science and New Technologies	Statement on Artificial Intelligence, Robotics and 'Autonomous' Systems <sup>19)</sup>	2018	9
Select Committee on Artificial Intelligence (UK)	AI in the UK: ready, willing and able <sup>20)</sup>	2018	5
IBM	Everyday Ethics for Planning	2018	5
Conference toward AI Network Society (Japan)	Draft AI Utilization Principles	2018	10
Sony	Sony Group AI Ethics Guidelines	2018	7
AI4People <sup>21)</sup>	An Ethical Framework for a Good AI Society	2018	5
Cabinet Office (Japan)	Principles of Human-centric AI Society	2019	7
High-Level Expert Group on AI (EU)	Ethics Guidelines for Trustworthy AI	2019	4

Note: The shaded guidelines were used as bases for formulating the five principles by AI4People.

happiness and freedom, and on serving the interests of the public. In March 2019, the Japanese government officially adopted the "Social Principles of Human-Centric AI"<sup>12)</sup> through a Cabinet Office decision. In April of the same year, the European Commission presented the "Draft Ethics Guidelines for Trustworthy AI,"<sup>13),14)</sup> emphasizing that AI systems must be "human-centric." The concept of "human-centeredness," therefore, has become an important principle in AI ethics.

More and more stakeholders are announcing their own AI ethics guidelines, but Fujitsu had recognized the importance of AI ethics since 2015. Fujitsu formulated the "Fujitsu Group AI Commitment" to concretely explain its initiatives related to AI ethics.<sup>1)</sup> The Fujitsu Group AI Commitment is an expression of the company's determination to exercise self-discipline as a developer and provider of AI solutions and fulfill its role in widely promoting the uptake of the values of AI in society.

#### 4. Measures aimed at ensuring objectivity and comprehensiveness of AI ethics principles

Fujitsu believed that in formulating its "AI commitment," it was essential to clarify the basis for the necessity of each AI ethics principle (objectivity) and to cover all the necessary considerations (comprehensiveness). To guarantee that these two conditions are met, Fujitsu

partnered with AI4People, a forum of experts on AI ethics.

Fujitsu is one of the founding partners of AI4People, which was established in November 2017 as the first global forum in Europe to discuss the social impacts of AI. Its activities for its first year, in 2018, were aimed at establishing AI ethics principles and laying down the foundations for a "Good AI Society."<sup>22)</sup>

AI4People, taking advantage of it being a relative latecomer in the formulation of AI ethics principles, endeavored to come up with a universal set of principles by comparing the six most important set of AI ethics principles (shaded items in Table 1). This approach enabled them to narrow down these sets of ethics principles into the five most important ones below, which include all four principles of biomedical ethics<sup>23)</sup> and one specific to AI ethics. The first four principles correspond to those of biomedical ethics.

##### Principle 1: Beneficence

AI systems should be designed and developed for the well-being and benefit of humanity; this includes contributing to the establishment of a sustainable and fair society that enables equal access to all, not only generating value and wealth.

##### Principle 2: Non-maleficence

AI systems should not cause harm to humans; this includes not only avoiding physical harm, but also preserving human dignity, freedom, privacy, safety and security.

### Principle 3: Autonomy

Human judgment and decisions should take precedence over those of AI systems. Humans maintain the right to decide over, to know the reason for, and to not obey any judgment, decision, instruction, or command given by AI. The right to entrust decision-making power to AI solely lies in the person concerned, and that power can be revoked anytime.

### Principle 4: Justice

The benefits of AI should be distributed widely to all people, without bias and discrimination against a particular person or group. Compensation for damages including provisions for recovery and relief should be stipulated.

### Principle 5: Explicability

In the principles of biomedical ethics, the requirement of explicability is included under the principle of "autonomy;" wherein patients have the right to make decisions for themselves after receiving the proper explanation about their treatments. Unlike the medical world, however, the provision of AI services cannot be premised on humans and AI sharing the same sense of values. Trust is gained only after an AI's criteria for decisions are explained to humans in a way that the latter, as the recipients of the services, can fully understand. The principle of explicability are as follows.

- 1) Providing a straightforward explanation of the reasons for AI decisions
- 2) Ensuring the ability to simulate the process that leads to a problem and identify its causes, in case humans feel that something is wrong with a decision that AI makes
- 3) Allowing third parties to carry out audits and inspections.

These five principles are in line with the long-accepted principles of biomedical ethics and therefore not only comprehensively cover all areas of AI ethics, but are also guaranteed of objectivity. We, therefore, believe that they will continue to remain valid even as AI technologies continue to advance and become widely applied in society.

## 5. Fujitsu Group AI Commitment

As explained above, the Fujitsu Group AI Commitment is an expression of Fujitsu's commitment to all its stakeholders, including its customers and society as a whole. Its goal goes beyond the mere delivery

of AI solutions and services that satisfy ethical principles stipulating "the rules that AI must comply with." In accordance with the FUJITSU Way,<sup>24)</sup> Fujitsu endeavors to offer "customer-centric AI," "trustworthy AI," and "responsible AI" by putting a premium on its relationship with its customers and society, as part of its obligations as an organization engaged in the research, development, provision, and operation of AI technologies. To this end, we have defined promises of conduct to be followed by all officers and employees engaged in AI in the Fujitsu Group.

The Fujitsu Group AI Commitment, which is made up of five core promises, is a declaration to stakeholders (customers, people, society, shareholders, and employees) of Fujitsu's intention on how it provides or will provide AI solutions that conform to the ethics principles. Fujitsu came up with ethics promises that take stakeholders into account by creating a matrix of the five AI ethics principles AI4People developed, against all the stakeholders of Fujitsu. **Table 2** shows the relationship between the five AI4People AI ethics principles and the five core promises of the Fujitsu Group AI Commitment.

The five core promises of the Fujitsu Group AI Commitment are explained below.

#### 1) Provide value to customers and society by AI

This promise is a major guiding promise in Fujitsu's efforts to provide AI solutions and services. Our role does not end in creating and providing AI systems to customers, but we are committed to generating innovations through co-creation with our customers from the early stages of the process, and to continually offer value to our customers through "continuously evolving AI." Because our business is providing ICT infrastructures to society, adhering to this promise will also enhance our ability to contribute to society.

#### 2) Strive for Human-Centric AI

This promise corresponds with the AI4People principles of beneficence, non-maleficence, and justice towards humans. Fujitsu is committed to building AI solutions and services that are meant for the benefit of humans, and to making sure that AI supports people with diverse values and capabilities to exhibit their potential according to their aspirations and circumstances. At the same time, to minimize the possible harmful effects of AI, Fujitsu will aim to create AI solutions and services that people can use with peace of

**Table 2**  
**Relationship between the Fujitsu Group AI Commitment promises and the five AI ethics principles adopted by AI4 People.**

		Five AI ethics principles by AI4People				
		Benevolence	Non-maleficence	Autonomy	Justice	Explicability
Stakeholders	Customers	✓	✓	✓	✓	✓ 1)
	People	✓	✓ 2)	✓ 4)	✓ 2)	✓ 5)
	Society	✓	✓	3)	✓	✓
	Shareholders	Contribution through the issuance of appropriate AI ethics guidelines				
	Employees	Covered through specific rules and regulations on AI ethics				

Note: Numbers 1) to 5) in the table correspond to the Fujitsu Group AI Commitment promises and indicate the particular areas they cover.

mind, by focusing our efforts on establishing a high level of quality for our AI solutions and services. This includes reducing or eliminating discrimination, ensuring security, and protecting the privacy of users.

3) Strive for a sustainable society with AI

This promise mainly corresponds to the principle of benevolence, as well as of non-maleficence and justice, towards society. It also expresses Fujitsu’s commitment to leverage AI in ensuring the sustainability of the earth’s environment. Our vision for a Human Centric Intelligent Society embodies the Fujitsu Group’s goal of making sustainable contributions to society through co-creation. This vision is consistent with the Sustainable Development Goals (SDGs)<sup>25)</sup> adopted by the international community.

4) Strive for AI that respects and supports people’s decision making

This promise corresponds to the principle of respecting human autonomy. It expresses Fujitsu’s belief that human prerogative takes precedence over AI proposals and decisions, i.e., humans always have the right to make the final decisions.

5) As corporate social responsibility, emphasize transparency and accountability for AI

This promise is based on the principle of explicability, but more strongly expresses it from Fujitsu’s position of being a provider of ICT infrastructures that underpin society. It expresses Fujitsu’s commitment to ensure the transparency of AI and to provide sufficient information regarding decisions made by AI, in order to gain stakeholders’ trust towards Fujitsu’s AI solutions

and services. Recognizing that AI technologies are not omnipotent, Fujitsu continually endeavors to reduce failures and to create systems that are capable of minimizing the effects of failures and immediately implementing measures to prevent their recurrence.

Not all the goals pertaining to the above five promises are immediately achievable at the moment. As AI technologies continue to progress and be applied in society at a rapid pace, it has become imperative to discuss the technologies and frameworks for creating “trusted AI.” Other than being the first step towards creating such frameworks, the Fujitsu Group AI Commitment also serves as the foundation for formulating the strategies for the development of such technologies. These commitments, therefore, are not set in stone, but rather should be revised as necessary on the basis of the state of social acceptance of AI technologies and the directions of discussions on AI ethics.

## 6. Conclusion

This paper described the promises of the Fujitsu Group AI Commitment announced by Fujitsu in March 2019 and the background for its formulation.

The Fujitsu Group AI Commitment expresses Fujitsu’s current set of promises—not concrete solutions for problems that arise in day-to-day activities—pertaining to the research, development, provision, and operation of AI technologies. Going forward, Fujitsu will formulate specific rules and regulations to be used for making on-site decisions based on this set of promises. At the same time, Fujitsu will leverage its experience

in applying these promises to promote the appropriate enactment of AI ethics regulations around the world.

-----  
All company and product names mentioned herein are trademarks or registered trademarks of their respective owners.

## References

- 1) Fujitsu: Fujitsu Formulates "Fujitsu Group AI Commitment," Demonstrating Dedication to Safe and Secure Use of AI. <https://www.fujitsu.com/global/about/resources/news/press-releases/2019/0313-01.html>
- 2) Japanese Nursing Association: Nursing Practice Information (in Japanese). [https://www.nurse.or.jp/nursing/practice/rinri/text/basic/what\\_is/index.html](https://www.nurse.or.jp/nursing/practice/rinri/text/basic/what_is/index.html)
- 3) M. Onohara: All About Ethics (in Japanese). <https://blog.goo.ne.jp/masaoonohara/e/8cf1e287a4c420ec976acd38e893f87>
- 4) I. Asimov: I, Robot, Gnome Press, 1950.
- 5) J. F. Bonnefon: "Autonomous Vehicles Need Experimental Ethics: Are We Ready for Utilitarian Cars?" arXiv:1510.03346v1, 2015. <https://pdfs.semanticscholar.org/13d4/56d4c53d7b03b90ba59845a8f61b23b9f6e8.pdf>
- 6) Wikipedia: Tay. [https://en.wikipedia.org/wiki/Tay\\_\(bot\)](https://en.wikipedia.org/wiki/Tay_(bot))
- 7) Wikipedia: Criticism of Google. [https://en.wikipedia.org/wiki/Criticism\\_of\\_Google#Portrayals\\_of\\_race\\_and\\_gender](https://en.wikipedia.org/wiki/Criticism_of_Google#Portrayals_of_race_and_gender)
- 8) Reuters: Amazon scraps secret AI recruiting tool that showed bias against women. <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G>
- 9) A. Chouldechova: "Fair prediction with disparate impact: A study of bias in recidivism prediction instruments." Big Data, Vol. 5, pp. 153–163, 2017. <https://arxiv.org/pdf/1703.00056.pdf>
- 10) Public Works and Government Services Canada: Invitation to qualify for opportunities in Artificial Intelligence. <https://buyandsell.gc.ca/invitation-to-qualify-for-opportunities-in-artificial-intelligence>
- 11) Fujitsu: AI technology developed by Fujitsu systematized as the "Human Centric AI Zinrai." (in Japanese). <https://pr.fujitsu.com/jp/news/2015/11/2.html>
- 12) Japanese Cabinet Office: Council for Social Principles of Human-Centric AI (in Japanese). <https://www8.cao.go.jp/cstp/tyousakai/humanai/index.html>
- 13) European Commission: Draft Ethics guidelines for trustworthy AI. <https://ec.europa.eu/digital-single-market/en/news/draft-ethics-guidelines-trustworthy-ai>
- 14) European Commission: Ethics Guidelines for trustworthy AI. <https://ec.europa.eu/futurium/en/ai-alliance-consultation>
- 15) Partnership on AI: Tenets. <https://www.partnershiponai.org/tenets>
- 16) Future of Life Institute: Asilomar AI Principles. <https://futureoflife.org/ai-principles/>
- 17) IEEE: Ethically Aligned Design. <https://ethicsinaction.ieee.org/>
- 18) An Initiative of Université de Montréal: Montréal Declaration Responsible AI. <https://www.montrealdeclaration-responsibleai.com/>
- 19) Publications Office of the EU: "Statement on Artificial Intelligence, Robotics and 'Autonomous' Systems." <https://op.europa.eu/en/publication-detail/-/publication/dfebe62e-4ce9-11e8-be1d-01aa75ed71a1/language-en/format-PDF/source-78120382>
- 20) House of Lords–Select Committee on Artificial Intelligence: "AI in the UK: ready, willing and able?" <https://publications.parliament.uk/pa/ld201719/ldselect/ldai/100/100.pdf>
- 21) AI4People: AI4People. <http://www.eismd.eu/ai4people/>
- 22) L. Floridi et al.: "AI4People–An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations." Minds and Machines, Vol. 28, Issue 4, pp. 689–707, 2018. <https://link.springer.com/content/pdf/10.1007%2Fs11023-018-9482-5.pdf>
- 23) T. L. Beauchamp et al.: Principles of Biomedical Ethics, Oxford University Press, 1983.
- 24) Fujitsu: Our Corporate Philosophy "FUJITSU Way." <https://www.fujitsu.com/global/about/philosophy/>
- 25) Ministry of Foreign Affairs of Japan: What is the SDGs? <https://www.mofa.go.jp/policy/oda/sdgs/index.html>



**Tsuneo Nakata**

Fujitsu Laboratories Ltd.

Dr. Nakata is currently engaged in R&D on AI ethics.



**Tatsuki Araki**

*Fujitsu Ltd.*

Mr. Araki is currently engaged in creation of new business and development of new contract models in the field of Data and XAI.



**Satoshi Tsuchiya**

*Fujitsu Ltd.*

Mr. Tsuchiya is currently engaged in development of AI service platforms.



**Yuri Nakao**

*Fujitsu Laboratories Ltd.*

Mr. Nakao is currently engaged in research on recommendation systems, fairness-aware machine learning, and accessibility user observation.



**Aisha Naseer**

*Fujitsu Laboratories of Europe Ltd.*

Dr. Naseer is currently engaged in research on AI ethics, particularly in collaboration with standardization groups in Europe.



**Junichi Arahori**

*Fujitsu Ltd.*

Mr. Arahori is currently engaged in policy-making related to cutting-edge technologies, such as AI, digital annealers, and data utilization.



**Takahiko Yamamoto**

Mr. Yamamoto was engaged in formulation of strategies for promoting standards in AI ethics and other issues related to the social implementation of new technologies at Fujitsu up to September 2019 prior to his retirement.