

United States House of Representatives  
Committee on Financial Services  
2129 Rayburn House Office Building  
Washington, D.C. 20515

October 7, 2021

## Memorandum

**To:** Members, Committee on Financial Services  
**From:** FSC Majority Staff  
**Subject:** October 13, 2021, Task Force on Artificial Intelligence Hearing entitled “Beyond I, Robot: Ethics, Artificial Intelligence, and the Digital Age.”

---

The Task Force on Artificial Intelligence will hold a hearing entitled, “Beyond I, Robot: Ethics, Artificial Intelligence, and the Digital Age” on October 13, 2021, at 12 p.m. ET, on the virtual meeting platform Cisco Webex. The single-panel hearing will have the following witnesses:

- **Meredith Broussard**, Associate Professor, Arthur L. Carter Journalism Institute of New York University
- **Meg King**, Director, Science and Technology Innovation Program, The Wilson Center
- **Miriam Vogel**, President and CEO, EqualAI
- **Jeffery Yong**, Principal Advisor, Financial Stability Institute, Bank for International Settlements
- **Aaron Cooper**, Vice President for Global Policy, BSA – The Software Alliance

### Overview

Isaac Asimov’s three laws of robotics, popularized in his short story collection “I Robot” outlines the ethical guidelines intelligent machines are bound to follow – 1) to serve and protect their creators (i.e., humankind); 2) to not harm humans either by their action or inaction, and 3) to protect themselves as long as this does not interfere with the other two laws. However, these laws may be a poor guide for our current age of Artificial Intelligence (AI) technologies.<sup>1</sup> AI can broadly be thought of as computerized systems that work and react in ways commonly thought to require intelligence, such as learning, solving problems, and achieving goals under uncertain and varying conditions. The field encompasses a range of methodologies and application areas, including machine learning (ML), predictive modeling, and natural language processing (NLP). These technologies are often defined to optimize a particular objective function but can be less aware of any unintended harm. In financial services, these systems often tackle complex problems in real-world situations and offer new tools, products, and services, from businesses to consumers, including credit underwriting, customer service, and cybersecurity.<sup>2</sup> Using AI irresponsibly could result in financial market instability or unintended discrimination against protected groups when they do not perform as intended.<sup>3</sup> Consequently, there is growing concern over the extent to which AI’s effects are fully known or understood when a system is implemented.<sup>4</sup>

---

<sup>1</sup> See e.g. The Conversation, [After 75 years, Isaac Asimov’s Three Laws of Robotics need updating](#) (Mar. 17, 2017).

<sup>2</sup> See e.g. Finextra, [How AI and ML Will Be Transforming Banking and Finance](#) (Jan. 19, 2021).

<sup>3</sup> See McKinsey & Company, [Confronting the risks of artificial intelligence](#) (Apr. 26, 2019).

<sup>4</sup> See The Harvard Gazette, [Great promise but potential for peril: Ethical concerns mount as AI takes bigger decision-making role in more industries](#) (Oct 26, 2020).

AI has the potential to offer a broad range of potential benefits, including the speed of data analysis and the ability to synthesize large datasets at a rate that could not possibly be achieved through human analysis.<sup>5</sup> However, complex AI calculations can produce undesirable consequences as well. The Task Force on Artificial Intelligence has examined concerns regarding AI usage, including how human-centered AI can address systemic racism<sup>6</sup> and how AI can affect jobs in financial services.<sup>7</sup> Building off these prior hearings, this hearing will consider ethical frameworks to assess the potential benefits and harms that AI technologies could cause in society, including in highly consequential decision areas such as financial services and housing. Additionally, this hearing will examine how domestic and international frameworks on AI ethics are being considered, and the ethical implications of new technologies, such as predictive modeling.

### **Domestic and International Ethical AI Proposals from Government Entities**

The use of AI has become widespread worldwide, and the resulting ethical concerns have spurred conversation by policymakers on the necessity for the development of ethical AI frameworks which could provide guardrails to ensure that new AI technologies benefit humankind.<sup>8</sup> AI's ethical principles vary between frameworks, but may include principles of fairness, accountability, transparency, justice, security, and privacy.<sup>9</sup> Often, these principles are summarized as good AI governance. Still, the interpretations or achievements of these principles, such as the allowance of third-party assessments and audits of AI systems are subject to debate.<sup>10</sup> In many cases, AI development firms will employ a cost/benefit calculation in an attempt to fully understand the implications of a particular AI system before deployment. These analyses consider the extent to which the AI system may produce unintended or tangential consequences that may end up harming end-users in the future.<sup>11</sup>

Ethics and good governance are widely viewed as essential elements of trustworthy AI systems.<sup>12</sup> In the U.S., a variety of federal entities are working on standards and frameworks for AI, including considerations of ethics, bias, and explainability. For example, in March 2021, the Office of the Comptroller of the Currency, the Federal Reserve System, the Federal Deposit Insurance Corporation, the Consumer Financial Protection Bureau, and the National Credit Union Administration issued a Request for Information and Comment on Financial Institutions' Use of AI, Including ML, to examine issues such as bias, fairness, and transparency.<sup>13</sup>

---

<sup>5</sup> See e.g. Brookings, [How artificial intelligence is transforming the world](#) (Apr. 24, 2018).

<sup>6</sup> House Financial Services Committee, [Equitable Algorithms: How Human-Centered AI Can Address Systemic Racism and Racial Justice in Housing and Financial Services](#) (May 7, 2021).

<sup>7</sup> House Financial Services Committee, [Robots on Wall Street: The Impact of AI on Capital Markets and Jobs in the Financial Services Industry](#) (Dec. 6, 2019).

<sup>8</sup> See Congressional Research Services, [Artificial Intelligence: Background, Selected Issues, and Policy Considerations](#) (May 19, 2021).

<sup>9</sup> See e.g. Jobin, A., Ienca, M. & Vayena, E. [Artificial Intelligence: the global landscape of AI ethics guidelines](#), Natural Machine Intelligence (2019).

<sup>10</sup> *Id.*

<sup>11</sup> See e.g. Leslie, D., [Understanding artificial intelligence ethics and safety: A guide for the responsible design and implementation of AI systems in the public sector](#), The Alan Turing Institute, (2019).

<sup>12</sup> See National Institute of Standards and Technology, [U.S. Leadership in AI: A Plan for Federal Engagement in Developing Technical Standards and Related Tools](#) (Aug. 9, 2019).

<sup>13</sup> Federal Reserve, FDIC, CFPB, NCUA, OCC, [Request for Information and Comment on Financial Institutions' Use of Artificial Intelligence, Including Machine Learning](#) (Mar. 31, 2021).

As part of its 2019 plan for federal engagement in the development of AI technical standards and related tools, the National Institute of Standards and Technology (NIST) stated that the U.S. federal government should “facilitate research and collaboration across scientific disciplines to increase understanding of how societal and ethical considerations relate to and can advance the development and use of standards for reliable, robust, and trustworthy AI technologies.”<sup>14</sup> The National Science and Technology Council (NSTC) has discussed the challenges and potential approaches to designing and building ethical AI, noting that “[e]thics is inherently a philosophical question while AI technology depends on, and is limited by, engineering.... However, acceptable ethics reference frameworks can be developed to guide AI system reasoning and decision-making to explain and justify its conclusions and actions.”<sup>15</sup> To reach those goals, NSTC’s Strategic Plan stated a need for multidisciplinary research in designing AI architectures that incorporate ethical reasoning by design.<sup>16</sup>

The General Services Administration’s (GSA) AI Center of Excellence has created a “Guide to AI Ethics” that lists questions to be “continually revisited through the design-develop-deploy cycle and extensively tested to mitigate any unintended consequences from the application of AI technology.”<sup>17</sup> Examples include: “Have you trained your business and technical teams about AI ethics and bias? . . . Have you established a process to identify and address ways that the algorithm could discriminate against protected classes, either by direct inclusion of these attributes or by proxy? . . . Have you determined quantitative evaluation criteria to measure the impact on users, stakeholders affected by the decision made or informed by the AI solution?”<sup>18</sup> As part of the Federal Data Strategy 2020 Action Plan, GSA further released a data ethics framework in December 2020, including an AI and bias use case.<sup>19</sup> In June 2021, the Government Accountability Office (GAO) released an AI accountability framework for federal agencies and other entities involved in the design, development, deployment, and continuous monitoring of AI systems, centered on the principles of governance, data, performance, and monitoring.<sup>20</sup>

Ethical AI frameworks have been considered internationally as well. The European Commission’s independent High-Level Expert Group on Artificial Intelligence (HLAI) developed ethical guidelines on the usage of AI, which puts forth four ethical principles and seven key requirements for developing and implementing trustworthy AI. The ethical principles formulated by HLAI are: respect for human autonomy, prevention of harm, fairness, and explicability.<sup>21</sup> Additionally, HLAI’s key requirements for an AI to be deemed trustworthy are human agency and oversight, technical robustness and safety, privacy and data governance, transparency, diversity, non-discrimination, fairness, societal and environmental well-being, and accountability.<sup>22</sup>

---

<sup>14</sup> National Institute of Standards and Technology, [U.S. Leadership in AI: A Plan for Federal Engagement in Developing Technical Standards and Related Tools](#) (Aug. 9, 2019).

<sup>15</sup> Select Committee on Artificial Intelligence of the National Science and Technology Council, [The National Artificial Intelligence Research and Development Strategic Plan: 2019 Update](#) (Jun. 2019).

<sup>16</sup> *Id.*

<sup>17</sup> General Services Administration, Artificial Intelligence Center of Excellence (CoE), [CoE Guide to AI Ethics](#) (May 2020).

<sup>18</sup> *Id.*

<sup>19</sup> General Services Administration, [Federal Data Strategy: Data Ethics Framework](#) (Dec. 2020).

<sup>20</sup> Government Accountability Office, [Artificial Intelligence: An Accountability Framework for Federal Agencies and Other Entities](#), (Jun. 30, 2021).

<sup>21</sup> High-Level Group on Artificial Intelligence, [Ethics Guidelines for Trustworthy AI](#), (Apr. 8, 2019).

<sup>22</sup> *Id.*

Additionally, the United Nations released a report in September 2021, urging action to assess and address the serious risks that AI technologies may pose to human rights.<sup>23</sup>

### **Ethical AI Proposals from Civil Society Stakeholders**

Researchers, nonprofit entities, and professional organizations have also proposed ethical frameworks. A framework proposed by the American Council for Technology-Industry Advisory Council (ACT-IAC) would describe and quantify five components (bias, fairness, transparency, responsibility, and interpretability) used to create, operate, and improve AI capabilities, delivering “an overall score that should be used to evaluate the ethics of an AI system, which should continually be monitored and checked over time.”<sup>24</sup> Managerial decision-making—the process by which we revise and reconsider AI ethics—will also be an essential part of any conversation on AI ethics frameworks.<sup>25</sup> Additionally, various nonprofit groups and governments have proposed Algorithmic Impact Assessments (AIAs) for automated decision systems, as well as frameworks for evaluating AIAs.<sup>26</sup>

Some groups have also described a need for risk-based, sector-specific approaches, particularly for AI applications that directly affect U.S. citizens’ lives and livelihoods, sometimes referred to as “high risk” or “systems critical” uses, such as many financial services applications like credit decisions. For example, the professional association the Institute of Electrical and Electronics Engineers (IEEE), which includes a standards development unit actively working to help develop consensus-based industry standards across many technologies, released a 2019 report that created a “conceptual framework addressing universal human values, data agency, and technical dependability with a set of principles to guide [autonomous and intelligent systems (AIS)] creators and users through a comprehensive set of recommendations.”<sup>27</sup> The principles, similar to those of the HLAI, are human rights, well-being, data agency, effectiveness, transparency, accountability, awareness of misuse, and competence of creators and operators.<sup>28</sup>

Subsequently, IEEE released an implementation playbook for financial services technologists and institutions “to provide a theoretical framework for implementing responsible data and AIS in financial institutions while outlining the global regulatory landscape with its impacts on the future development of AIS and ethics.”<sup>29</sup> The Finance Playbook includes a roadmap for organizations to “develop, execute, and scale trusted data and AIS” that ties together considerations for three building blocks: (1) people, including leadership buy-in and education and awareness across the organization; (2) process, including principles, data, and AIS impact assessments, as well as measurements and performance indicators; and (3) technology, including

---

<sup>23</sup> See NPR, [The U.N. Warns That AI Can Pose A Threat To Human Rights](#) (Sep. 16, 2021).

<sup>24</sup> Artificial Intelligence Working Group, [Ethical Application of Artificial Intelligence Framework](#), American Council for Technology-Industry Advisory Council (ACT-IAC) (Oct. 8, 2020).

<sup>25</sup> See The National Artificial Intelligence Act of 2020, [Division E of the William M. \(Mac\) Thornberry National Defense Authorization Act \(NDAA\) for Fiscal Year 2021, P.L. 116-283](#), see also [AI in Government Act of 2020](#), Division U, Title I of the Consolidated Appropriations Act, 2021, P.L. 116-260.

<sup>26</sup> See e.g. Dillon Reisman, Jason Schultz, Kate Crawford, et al., [Algorithmic Impact Assessments: A Practical Framework for Public Agency Accountability](#), AI Now (Apr. 2018).

<sup>27</sup> 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](#) (2019).

<sup>28</sup> *Id.*

<sup>29</sup> Institute of Electrical and Electronics Engineers, [IEEE Finance Playbook Version 1.0: Trusted Data and Artificial Intelligence Systems \(AIS\) for Financial Services](#), IEEE (Apr. 2021).

standardized operations and capabilities such as differential privacy, AIS and analytics models, and testing and optimization. The Finance Playbook further lays out key ethical considerations in what it lists as the top 20 high-value AI use cases in the financial sector.<sup>30</sup>

While these AI ethics frameworks vary, some common themes that emerge include involving human decision-makers throughout the AI life cycle of design, development, and implementation; maintaining a continuous or periodic evaluation of framework and tools (e.g., guiding principles, evaluation questions, auditing tools); and being transparent about the underlying values informing AI applications and their ultimate impacts.

### **AI Predictive Analytics and Ethics in Financial Services**

The proliferation of AI has transformed production processes and created new consumer services across many industries.<sup>31</sup> AI ethicists have raised concerns about AI tools that perform predictive modeling, which broadly involves the use of statistics to predict future outcomes.<sup>32</sup> Predictive modeling allows an AI to forecast future results or activities by analyzing data sets and patterns.<sup>33</sup> Analysts have noted that AI and ML can magnify the power of predictive models because “ML is able to manage vast amounts of data and detect many more complex patterns within them, often attaining superior predictive power.”<sup>34</sup> The use of AI, including ML, in predictive modeling is sometimes referred to as predictive analytics, “a branch of advanced analytics that makes predictions about future outcomes using historical data combined with statistical modeling, data mining techniques and machine learning.”<sup>35</sup>

The use of AI and ML in predictive analytics can have both positive and negative effects for businesses and individuals. On one hand, their use can mitigate risks to businesses, such as through increasing the accuracy of customer credit scoring, and protect consumers, such as by helping financial institutions monitor suspicious behavior patterns to detect fraud.<sup>36</sup> On the other hand, some academics have noted that the application of AI and ML can embed racial or gender-based biases in predictive analytics models, which can be introduced in the “generation, collection, labeling and management of data that the AI and ML algorithm learns from, as well as the design and evaluation of algorithms.”<sup>37</sup> The use of ML models on new data sources, often called alternative data, has also been shown in some instances to result in discriminatory decisions.<sup>38</sup> Therefore, the use of AI and ML in predictive analytics may result in inaccurate or discriminatory predictions for certain subsets of the population and can lead to an unfair allocation of opportunities, resources, or information; infringe on civil liberties; and fail to provide the same quality of service to some people as to others.

---

<sup>30</sup> *Id.*

<sup>31</sup> See CRS, [Artificial Intelligence: Background, Selected Issues, and Policy Considerations](#) (May 19, 2021).

<sup>32</sup> See e.g. Harvard Business Review, [When Does Predictive Technology Become Unethical?](#) (Oct. 23, 2020).

<sup>33</sup> See IBM, [Predictive Analytics](#) (last accessed Oct. 7, 2021).

<sup>34</sup> See McKinsey & Company, [Controlling machine-learning algorithms and their biases](#) (Nov. 10, 2017).

<sup>35</sup> See IBM, [Predictive Analytics](#).

<sup>36</sup> See IBM, [A business guide to modern predictive analytics](#).

<sup>37</sup> See [Mitigating Bias in Artificial Intelligence: An Equity Fluent Leadership Playbook](#) (July 2020).

<sup>38</sup> American Banker, [Proceed with Caution on Credit Scoring with Alternative Data](#) (Jun. 11, 2015).