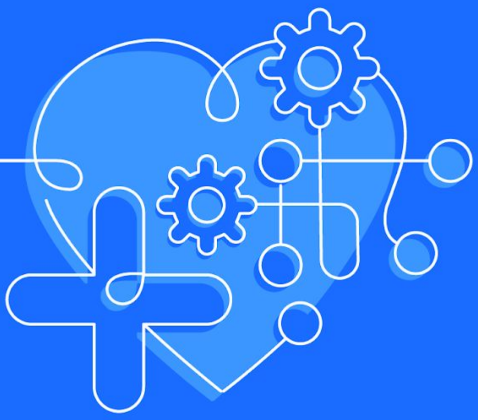

National Institute of Standards and Technology

RFI: Developing a Federal A.I. Standards Engagement Plan

June 7, 2019

buoy®

Boston Office
883 Boylston Street, Suite 2
Boston, MA 02116

New York Office
60 Broad Street, Suite 2506
New York, NY 10004

June 7, 2019

AI-Standards
National Institute of Standards and Technology
100 Bureau Drive
Stop 2000
Gaithersburg, MD 20899

Re: RFI: Developing a Federal A.I. Standards Engagement Plan

On behalf of Buoy Health, Inc. ("Buoy"), we are pleased to submit a comment in response to the National Institute of Standards and Technology ("NIST") Request for Information: Developing a Federal A.I. Standards Engagement Plan, docket number 190312229-9229-01.

Buoy leverages artificial intelligence throughout our business. We acknowledge the value in developing federal standards related to A.I., and we hope to continue to participate in this conversation as it continues.

Enclosed is the following:

- Information about Buoy Health, Inc. (page 2)
- Response to RFI: Developing a Federal AI Standards Engagement Plan (page 4)

The contact for this request for information is Cory Lamz, Esq., Counsel & Data Privacy Officer, cory@buoyhealth.com.

Buoy appreciates the opportunity to submit this request for information. Please contact us if you have any questions or comments.

Warmly,

Buoy Health, Inc.

Cory Lamz, Esq.
Counsel & Data Privacy Officer

Eddie Reyes
Chief Technology Officer

Meet Buoy

Buoy Health, Inc. ("Buoy"), was founded in 2014 to address a big question in healthcare: "What do I do when I get sick?"

Buoy uses its proprietary A.I. technology to answer this question for its users. By communicating with Buoy's A.I. Health Assistant – a proprietary, interactive tool that mirrors a friendly conversation with a doctor – users can learn more about their symptoms as well as where to seek care, even if it is not immediately, so that users can feel more confident and in control of their health.

Who we are

Buoy consists of a team of nearly 30 data scientists, doctors, designers, engineers, marketers, and operators (finance, legal, human resources, etc.). The company is based in Boston and has an office in New York City. Buoy is financially backed by Optum Ventures and F-Prime, the venture arm of Fidelity, and in 2018 won the grand prize at the Robert Wood Johnson Foundation A.I. Challenge, awarded to the company that is best leveraging A.I. in health care.

In addition to its core team, Buoy has a world-renowned medical advisory board, with experts in medical diagnosis and patient safety from a variety of practice areas, and a learning partnership with Boston Children's Hospital to boost Buoy's pediatric capabilities.

How our technology works

Buoy's A.I. Health Assistant asks users questions related to their symptoms in a quick, approximately 5-minute chat. Based on these answers, the tool provides users with relevant medical information that empowers users to self-diagnose and take further action, including to seek care that is located nearby, provided via teleconference, or provided in-network by their employer health plan.

The engine that determines what medical information is provided to users based on their answers, as well as what "next steps" are relevant, is powered by A.I.

Buoy's tool helps users reduce the costs associated with getting sick for individuals and employers.

For individuals, the tool helps users locate the right type of care at the right time.

For employers, the tool helps drive down the business costs associated with sick employees, including decreasing the frequency and/or cost of insurance claims, thereby reducing the healthcare costs that are eventually passed onto consumers. The tool also helps direct the right individuals to the proper care provider at the right location and the right time, thereby reducing, for example, unnecessary emergency room or specialist visits and the overall demand on improper sites of care across the country.

Buoy's tool also impacts population health in three major ways.

First, the Buoy tool helps reduce the barriers to accessing by care by decreasing the time and costs associated with finding the right type of care at the right time. For example, Buoy users in remote areas are not restricted by their geography to seek proper care - they can connect with a telemedicine provider without leaving their own home, regardless of whether their surrounding community has the proper resources to provide the care needed.

Second, the Buoy tool makes medical information more readily and immediately available to anyone with a Wi-Fi connection, encouraging those who truly need to seek care to seek the right care at the right time and discouraging those who don’t - thereby promoting a *wait and watch* approach instead of a gratuitous burden on the healthcare system.

Third, the Buoy tool effectively triages patients before they even reach the hospital or doctor’s office, freeing up space and resources for others whose symptoms are more complicated or severe.

I. THE FTC SHOULD CREATE A CERTIFICATION PROGRAM THAT RECOGNIZES BASELINE A.I. STANDARDS AND BIAS DISCLOSURE OBLIGATIONS

The reasonable consumer who engages with A.I. may not be aware of biases inherent to the system, including but not limited to data collection bias and attribute selection bias. As the use of A.I. becomes more and more prevalent in consumer-facing tools and resources, the responsibility to educate the consumer falls onto the organization that leverages A.I. for its product or service. From a liability framework, such organization would argue that the consumer assumes the risk of such biases by using that product or service. From a consumer-protection framework, the consumer would argue that it is the responsibility of the organization to acknowledge and disclose the biases, i.e., limits, within its product or service to the consumer. We argue that striking a balance between these two positions is not only possible but necessary as societal engagement with A.I. continues to rise, and we call on NIST to propose that any organization leveraging A.I. do the same.

To strike such a balance, the Federal Trade Commission (the “FTC”) should adopt a certification program to recognize those organizations that 1) leverage A.I.,[1] and 2) acknowledge and/or disclose relevant biases, similar to the FTC’s Green Guides program.[2] Such A.I. certification should consider whether the organization a) leverages A.I. as defined, b) incorporates domain-specific standards, c) discloses its biases to its consumers, and d) otherwise complies with the OECD Principles on Artificial Intelligence.[3]

This certification program should incorporate the standards within the domain to which the A.I. is applied. For example, in the medical industry, A.I. that is used to assist in and/or replicate medical procedures should be held to these A.I. certification standards as well as those licensing standards imposed on human practitioners. Such A.I. certification would need to be evaluated and renewed on a regular basis – e.g., every two or three years.

Finally, we argue that this certification program could be used to subsidize data security programs to reduce the barriers to the market that currently exist for non-large businesses operating in the data space. We also argue that this certification program could be used to unify current state privacy bills and include a safe harbor exception for qualifying A.I.

a. A.I. and its application should align with what it means to leverage “artificial intelligence”

Artificial intelligence varies across domains and applications. To enforce a certification program that standardizes A.I., the FTC would need to answer the threshold question: what is – and is not – “artificial intelligence”? The FTC would also need to set a baseline for what it means to leverage artificial intelligence. Both of these should be broad enough in scope to permit scalable application across any industry, including industries that may exist in the future. We do not propose an answer to either question here, but rather we call on NIST and the A.I. community to determine answers that are mutually agreeable by all stakeholders across sectors and industries.

[1] In so doing, as an A.I. community we would need to answer a threshold question: what is – and is not – “artificial intelligence”? From there, we would need to set a baseline for what it means to leverage “artificial intelligence.”

[2] *See* Federal Trade Commission - Green Guides, available at https://www.ftc.gov/news-events/media-resources/truth-advertising/green-guides (last accessed June 5, 2019).

[3] *See* “Forty-two countries adopt new OECD Principles on Artificial Intelligence”, OECD, May 22, 2019, https://www.oecd.org/going-digital/forty-two-countries-adopt-new-oecd-principles-on-artificial-intelligence.htm (last accessed June 5, 2019).
See also OECD, Recommendation of the Council on Artificial Intelligence, OECD/LEGAL/0449, available at https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449#_ga=2.142717234.629599887.1559661684-1238510132.1559661684.

b. A.I. should incorporate domain-specific standards

Artificial intelligence, as a technology, is a disruptive intervention. To date, A.I. has been leveraged to replicate human action or replace it altogether. However, in so doing, no national standards exist to monitor such intervention. Put another way, there is no national standard to evaluate or otherwise examine the success or failures of artificial intelligence within a given context, process, or sector. This is unsurprising: the definition of success largely varies across each application A.I. Still, how then can it be shown that A.I. has met some level of acceptance in what it was set out to accomplish?

When artificial intelligence is applied to a new domain that did not otherwise utilize A.I. previously, non-standard evaluation methodologies create hesitancy to adopt the technology. And not without reason: the domain must bear some level of risk in the adoption of this new technology. This is evidenced by scaled adoption; in others, it is by isolated adoption. In some domains, though, risk cannot be borne by the domain. For example, in medicine, medical professionals have zero risk tolerance because of what is at stake – their patients' health, their reputation, their licensure. Medical professionals would not adopt A.I. without proof of credibility, accuracy, precision, or similar reputable, standardized metrics.

Similarly, a lack of standard evaluation methodology creates apprehension between and within domains. Domain professionals are, again, wary of adoption and application without recognized acceptance standards. How else would such professionals be willing to assume risk, let alone understand whether the A.I. has been implemented successfully and/or appropriately?

In the absence of reputable evaluation methodology, A.I. developers have generated their own, potentially flawed, methodology to measure the success of A.I. within specific domains. This, however, has exacerbated distrust between those who understand the technology – including its strengths and limitations – and those who do not.

Therefore, we suggest the adoption of an FTC certification program that, in part, adopts artificial intelligence guidelines associated with evaluating the technology by domain via well-defined stages, similar to that of clinical trials in medicine[4]:

i. *Is the A.I. safe?*
This standard would examine whether the A.I. cause any harm or risk of harm in its application and, if so, whether the harm or risk of harm outweighs the benefits.

ii. *Does the A.I. work?*
This standard would examine whether the A.I. does what it was set out to do.

iii. *Is the A.I. better than what we have now?*
This standard would compare whether it is appropriate for the A.I. to replace the current approach. In some instances, particularly in light of the first and second standards, this would be a high bar to meet.

[4] Per the U.S. Food & Drug Administration (the "FDA"), during the first step of the new drug development process, researchers conduct experiments on the drug "to gather information on: i) how it is absorbed, distributed, metabolized, and excreted; ii) its potential benefits and mechanisms of action; iii) the best dosage; iv) the best way to give the drug (such as by mouth or injection); v) side effects or adverse events that can often be referred to as toxicity; vi) how it affects different groups of people (such as by gender, race, or ethnicity) differently; v) how it interacts with other drugs and treatments; and v) its effectiveness as compared with similar drugs." U.S. Food & Drug Administration, "Step 1: Discovery and Development", last updated Jan. 4, 2018, available at https://www.fda.gov/patients/drug-development-process/step-1-discovery-and-development (last accessed June 5, 2019). Effectively, such development standards can be grouped into four buckets that we have adopted for the purposes of this comment: 1) is the drug safe? 2) does the drug work? 3) is the drug better than what we have now? 4) are there any other uses or benefits of this drug?

iv. *Are there any other uses or benefits of the A.I.?*
Finally, this standard would examine whether the A.I. could be used for other applications as well.

These four standards would promote evaluation and testing in specific domains and sectors so as to measure A.I. in a standardized way. However, evaluation and testing is only half of the calculus; we must also solve for consumer adoption.

Therefore, we also suggest that the certification program also incorporate well-defined benchmarks for controversial, i.e., high-risk, applications for A.I. before such technology could be implemented in a given domain. These benchmarks would include:

i. Measurement of interpretability and explainability,
ii. Identification of known risks and/or deficiencies, and
iii. Ways to mitigate any resulting safety or ethical concerns.

These three benchmarks and four standards, when part of the certification program, would promote transparency between the organization leveraging A.I. and its impacted populations, thereby fostering greater trust in A.I. and implicitly encouraging public adoption of the technology.

c. Organizations that leverage A.I. should acknowledge and disclose their systems' biases

With rising demands in patient care, ballooning operational costs, and fixed resources, the healthcare sector has begun pursuing creative, technological solutions to problems that have plagued the modern healthcare system. Leveraging A.I. is one such solution.

Although the use of A.I. in the healthcare sector may be hugely beneficial to patients, providers, employers, and the public at large as discussed in the previous section, major conflicts currently exist between A.I. and the healthcare sector that severely limit the scope and prevalence of the technology in solving systemic problems.

The application of artificial intelligence to data that is collected, stored, or transferred in service to the consumer has unavoidable limitations that must be mitigated. One such limit, biases, may be introduced during the data collection process or in the development of the algorithms used to deploy and drive the A.I. Such limitations, in turn, can have far-reaching and undetectable consequences on any system outputs and/or the overall user experience. Therefore, these biases must be acknowledged and mitigated so that the consumer can understand the limitations of the data and act accordingly. We consider two of these biases – data collection bias and attribute selection bias – in turn, then call for a mechanism by which to acknowledge and disclose such biases.

1. Bias is introduced by the data collection process

A.I. systems built on machine learning are bounded by the data collected, including the completeness of the data and the recurrence of the same – or significantly similar – data set and how that data is communicated to and from the system itself. Bias may be introduced to these systems when the bounds of such data, or the mere existence of such boundaries, is processed by machine learning and, as a result, the algorithm evolves accordingly.

Firstly, the completeness of the data set shapes the successful processing of that data, where the degree of completeness is measured by whether all, some, or none, of the data sets requested are collected. For example, Buoy's A.I. Health Assistant may collect the following data: i) name; ii) email address; iii) age; iv) sex; v) race/ethnicity; vi) IP address; vii) web browser; viii) cookie ID ix) geolocation; x) access date/time; and xi) symptom information. These data points together make up a complete data set for purposes of data processing by the Buoy tool to deliver the most relevant and accurate medical information to the user. However, Buoy also may limit the completeness of this data set by scaling back the data points collected. In cases where users engage with the tool anonymously (i.e., without creating a user profile), the tool collects a limited data set that may only include age, sex, access date/time, and symptom information.

When Buoy collects a limited data set from a user, the output of corresponding medical information to that user is more generalized. However, when Buoy collects a complete data set from a user, the output of corresponding medical information is more acutely tailored to that individual. Complete data sets allow for the Buoy tool to provide medical information with consideration toward population health.

Secondly, those systems that receive significantly similar data sets repeatedly become more and more accurate as more of the same data is fed into the system. Machine learning then trains the algorithm to become more advanced with respect to that type of recurring data set, but not others. For example, Buoy's A.I. Health Assistant tool has seen significant usage by females age 18-25 who use the tool to seek out reproductive health information.[5] As a result, the algorithm is positively biased to perform better with respect to female reproductive health information than other areas.[6]

For some systems, the inverse may also be true. Consider, for example, a data set that is un- or under-representative of reality. "If a deep-learning algorithm is fed more photos of light-skinned faces than dark-skinned faces[,] [t]he resulting face recognition system would inevitably be worse at recognizing darker-skinned faces."[7] This happened for Amazon, whose A.I. recruiting tool had been dismissing female candidates sheerly because they were female. "Because [the A.I.] was trained on historical hiring decisions, which favored men over women, [the A.I.] learned to do the same."[8]

Finally, the data collected is also limited by communication of that data. For example, Buoy's A.I. Health Assistant is limited by user comprehension and communication – in particular, how symptom information is communicated to and from the tool. That is why Buoy prioritizes development work and system inputs that mirror an in-person conversation with a doctor. Overly technical questions distract or confuse the user, whereas Buoy strives to ask questions in plain English at an 8th grade reading level so as to facilitate the dialogue per se between user and tool. If the user does not understand the question posed by Buoy's A.I. Health Assistant, then the user may not answer correctly. Over time, the regular repetition of incorrect answers may spur the machine learning process to create inaccurate linkages between the data collected and the output of medical information. This creates a bias in the algorithm based purely on how the tool and user communicate with one another.

This bias may be introduced differently for users under age 13.[9] In compliance with the Children's Online Privacy and Protection Act of 1998, users under age 13 are not permitted to use Buoy's A.I. Health

[5] Over a one-year period, females age 18-25 made up 41% of Buoy's overall traffic.

[6] Over a one-year period, Buoy's female reproductive health information scored 700 basis points higher among female users than for other health information.

[7] Hao, K., "This is how AI bias really happens – and why it's so hard to fix", MIT Technology Review, Feb. 4, 2019, https://www.technologyreview.com/s/612876/this-is-how-ai-bias-really-happensand-why-its-so-hard-to-fix/ (last accessed May 26, 2019).

[8] *Id*; *see also* Golden, J., "AI has a bias problem. This is how we can solve it", World Economic Forum, Jan. 18, 2019, https://www.weforum.org/agenda/2019/01/to-eliminate-human-bias-from-ai-we-need-to-rethink-our-approach/ (last accessed May 26, 2019).

[9] 16 C.F.R. § 312.

Assistant on their own. Instead, an authorized parent or guardian must use the tool on their behalf. In such case, the tool relies on successful conveyance of symptom information from the youth to their parent/guardian, who then engages in dialogue with the tool. In those cases where symptom information is miscommunicated between child and parent/guardian, bias may arise with respect to the linkages between the data collected and the output of medical information. Similar concerns arise in cases where a youth cannot convey her symptom information to a parent/guardian whatsoever, as would be the case of any individual who is otherwise incapacitated in some way to engage with the tool or understand English at an 8th grade reading level, including but not limited to a non-communicative individual, a non-English speaking individual, or a baby.

2. Bias is introduced by attribute selection

A.I. systems built on machine learning are also bound to the algorithmic foundations set up by their developers. The A.I. is only as intelligent and useful as its developers permit it to be. Therefore, any biases held by the developers may be introduced, subconsciously or consciously, to the artificial intelligence within the system itself, i.e., by design.[10]

A developer may introduce a bias to a specific model by minute decision-making of specific attributes over others, thereby creating attribute-selection bias. For example, with Amazon's A.I. recruiting tool, developers might have selected the model to focus on the candidate's gender, education, or years of experience,[11] thereby not considering other attributes, such as professional certifications or volunteer experience. That model may perpetuate that bias.[12] For Buoy's A.I. health assistant, our team has been careful to train the A.I. to consider thousands of possible symptoms, without creating an overly cumbersome experience. Yet, given the complexity of how symptoms present from one person to the next, and the variance of symptoms by location, such considerations are finite – and the Buoy tool is unable to consider symptoms outside of those already included in the system. That is why we are perpetually training our A.I. to consider more and more attributes and symptom presentations. Even then, however, attribute-selection bias may creep in.

3. Bias disclosures de-mystify A.I. and foster public trust

Currently, consumers are unwilling to engage with A.I. because they do not understand it or recognize that certain black-box system biases may impact their experience. An FTC certification program could address consumers' concerns: de-mystifying A.I. and requiring disclosure of system biases in a transparent way that encourages use… but with limits. For Buoy, this would mean that we identify the biases inherent to the Buoy tool's data collection and attribute selection processes and disclose, to the fullest extent possible (in light of trade secret and related considerations) these biases to the consumer. Should the consumer proceed to use the tool, then she assumes any inherent risk therein.

d. A.I. should comply with the OECD Principles on Artificial Intelligence

In addition to the foregoing, it is critical that the FTC certification program emphasizes compliance with international standards that already exist, namely the OECD Principles on Artificial Intelligence. The

[10] For example, the COMPAS recidivism algorithm – a commercial tool previously used by judges, probation officers and parole officers to predict whether a criminal defendant would become a recidivist, i.e., commit future crimes – which was found to be implicitly biased against black defendants. Larson, J. et. al., "How We Analyzed the COMPAS Recidivism Algorithm", ProPublica, May 23, 2016, https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm. See also Angwin, J. et. al., "Machine Bias", ProPublica, May 23, 2016, https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.

[11] *See*, supra, note 19.

[12] *See, e.g.*, "Automated Experiments on Ad Privacy Settings: A Tale of Opacity, Choice, and Discrimination", Datta, A. et. al., Cornell University, last updated March 17, 2015, available at https://arxiv.org/abs/1408.6491 (which found that Google's online advertising system had displayed ads related to high-income jobs to men more often than to women).

OECD Principles state, in part, that A.I. systems should be designed in a way that includes appropriate safeguards (e.g., enabling human intervention where necessary), should include transparent and responsible disclosures to ensure that consumers understand such systems, and must function in a secure and safe way.[13] These principles are not a vast departure from the standards called for elsewhere by our proposed certification program – the major difference being that the OECD Principles are of international scope. Thus, we recommend that the certification program include the standards proposed herein and structured to align with the OECD Principles and international scale.

II. THE CERTIFICATION PROGRAM SHOULD ALIGN WITH INTERNATIONAL DATA SECURITY AND PRIVACY STANDARDS

Private companies that deal in data collection and processing do not survive in the market without public trust, which is undergirded in part by the implementation of a robust data security and privacy program. However, current data security standards may not be feasible for start-ups and small companies to scale a successful national A.I. strategy, and data privacy laws lack cohesion between the state and federal levels. We argue that the proposed FTC certification program could subsidize compliance with these data security standards and unify current data privacy laws.

a. Compliance with data security standards should be subsidized via the proposed certification program

While the FTC offers guidance on data security practices,[14] no federal agency mandates *specific* practices. Rather, market competition and international bodies, as well as state laws that govern data breach notifications,[15] have come to influence what have gone on to become routine practice[16] – i.e., established data security standards.[17]For example, all companies that engage with personal data in some way cannot compete unless they are compliant with international ISO standards,[18] HITRUST certified,[19] and/or the NIST Cybersecurity Framework[20] – and that is just to name three sources of a handful of

[13] *See*, supra, note 15.

[14] *See* Federal Trade Commission - Data Security, https://www.ftc.gov/tips-advice/business-center/privacy-and-security/data-security (last accessed May 14, 2019).

[15] *See e.g.*, CA Civ. Sec. 1798.82 (California Code (2019 Edition)); Colo. Rev. Stat. § 6-1-716 Notification of security breach. (Colorado Revised Statutes (2018 Edition)); NY GBS Law 899-AA Notification; person without valid authorization has acquired private information. (Laws of New York (2019 Edition)); Mass. Gen. Laws Ch. 93H Sec. 1-6 Duty to report known security breach or unauthorized use of personal information (The General Laws of Massachusetts (2017 Edition)); N Wash. Rev. Code 42.56.590 Personal information—Notice of security breaches. (Revised Code of Washington (2018 Edition)).

[16] For example, in the healthcare sector, for example, two subsets of HIPAA dictate specific data security protocols, the Security Rule (45 C.F.R. § 160; 45 C.F.R. § 164 Subparts A and C.; *see* "Health Information Privacy - Summary of the HIPAA Security Rule", U.S. Dept. of Health & Human Services, available at https://www.hhs.gov/hipaa/for-professionals/security/laws-regulations/index.html (last accessed May 14, 2019).) and the Breach Notification Rule (45 C.F.R. §§ 164.400-414; *see* "Health Information Privacy - Breach Notification Rule", U.S. Dept. of Health & Human Services, available at https://www.hhs.gov/hipaa/for-professionals/breach-notification/index.html (last accessed May 14, 2019)).

[17] *See, e.g.*, HITRUST Alliance,"New Version of HITRUST CSF Incorporates California Consumer Privacy Act, NIST Cybersecurity Framework and Additional Legislation & Standards", available at https://hitrustalliance.net/new-version-hitrust-csf-incorporates-california-consumer-privacy-act-nist-cybersecurity-framework-additional-legislation-standards/ (last accessed May 26, 2019) (includes list of most recent data security standards, including the Centers for Medicare & Medicaid Services' (CMS) Information Security ARS: CMS Minimum Security Requirements for High Impact Data, version 3.1; the Federal Risk and Authorization Management Program (FedRAMP); IRS Publication 1075 Tax Information Security Guidelines for Federal, STate and Local Agencies: Safeguards for Protecting Federal Tax Returns and Return Information; the National Institute of Standards and Technology's (NIST) Framework for Improving Critical Infrastructure Cybersecurity: Framework Core – Subcategories, v1.1; South Carolina's Bill 4655, the Insurance Data Security Act).

[18] Per the Organisation Internationale de Normalisation ("ISO") website, the ISO data security standards, i.e., the so-called family of standards of ISO/IEC 27000, help organizations keep information assets secure. ISO is a private, international organization based in Switzerland, with member organizations based in 164 countries, including the United States. *See* "ISO/IEC 27000 family - Information security management systems", Organisation Internationale de Normalisation, available at https://www.iso.org/isoiec-27001-information-security.html (last accessed May 23, 2019); *see* "All about ISO", Organisation Internationale de Normalisation, available at https://www.iso.org/about-us.html (last accessed May 23, 2019).

[19] Per one private organization that deals with PHI and other HIPAA-compliance considerations, HITRUST CSF certification "indicates that an organization has met industry-defined requirements and is appropriately managing risk when protecting patient data. It's similar to having TSA pre-boarding clearance at the airport – you breeze through security because you're a known quantity that's been pre-verified" ("What TigerText's HITRUST Certification Means for You", TigerConnect, July 27, 2016, available at https://www.tigerconnect.com/blog/what-tigertexts-hitrust-certification-means-for-you/ (last accessed May 17, 2019)) – except that TSA is an agency within the U.S. Dept. of Homeland Security, whereas HITRUST is not otherwise affiliated with the public sector.

[20] National Institute of Standards and Technology, "Framework for Improving Critical Infrastructure Cybersecurity", Version 1.0, Feb. 12, 2014, available at https://www.nist.gov/cyberframework/framework.

standards out there. Indeed, there is a wealth of standards and sources that dictate what constitutes “data security.”

Companies cannot compete in the marketplace unless they have the resources to become compliant and maintain such compliance with the aforementioned standards and the evolving market best practices that are set by the companies with the most resources and the biggest budgets to devote to data security. Moreover, these best practices evolve regularly – if smaller companies cannot keep up, users and customers can vote with their feet, so to speak, by choosing to take their business, and their data, elsewhere.

Such competition has created a disparity among data-centric companies. Companies that have the resources to support robust data security programs lead the marketplace and serve as industry thought-leaders are shaping consumer expectations and subsequently creating a high barrier to entry, nevermind success, for those companies that do not have the same resources. Thus, there is a divide between the major players and the small movers, between those companies that can devote entire teams to data security and those companies that must focus primarily on R&D and secondly on everything else. Although such robust data security programs may be presumptively good for protecting consumer data in the near term, ultimately the cost of standing up and maintaining such programs will squeeze out the smaller companies, i.e., the competition – and the result, data monopolies, will not be in service to the best interests of the consumer.

Therein lies the opportunity for the FTC to level the playing field so that large and small companies that leverage A.I. may compete in the same league. To be clear: this is not to advocate for the loosening of data security protocols. Rather, such protocols should be more easily implemented – and account for leveraging by – non-large companies.

To do this, we propose that the FTC certification program should subsidize the certification process for those companies that would not otherwise have the budget to comply with such robust data security standards. Such subsidies could exist in the form of monetary grants awarded to companies that i) have sufficiently demonstrated need and ii) have exhibited key engagement with A.I. to further propel that sector, and the market, and the country as a whole, forward. Such grants could be funded in whole or in part by certification (and re-certification) fees.

b. Data privacy obligations should be unified under the proposed certification program

A.I. systems that rely on machine learning to iterate, improve, and solve problems require the right to collect, store, and transfer data – including protected health information (“PHI”)[21] – to perform the services for which they are intended. For example, to engage with Buoy’s A.I. Health Assistant, users must grant Buoy a license to their PHI so as to process their symptom information, provide them with relevant medical information, and connect them with proper points of care. This use case, albeit straightforward, is rife with complexities related to data privacy due to rights granted at the state level by recent legislation in tension with federal healthcare laws.

This includes, for example, the right to deletion as granted by the California Consumer Privacy Act of 2018 (“CCPA”).[22] Simply put, if an artificially intelligent algorithm that is founded on machine learning

[21] as defined by 45 C.F.R. 160.103.
[22] The CCPA recognizes privacy rights associated with the usage of technology and related industries. The CCPA grants to California consumers the right 1) to know what personal information is being collected about them; 2) to know whether their personal information is sold or disclosed and to whom; 3) to decline the sale of their personal information; 4) to access their personal information; and 5) to access the applicable technologies without recourse to price or service availability if the foregoing rights are exercised. Assembly No. 375 (June 29, 2018). However, legislators have proposed to amend the CCPA as of April 2019. *See* Brennan, M. et. al., “CCPA Amendments

collects PHI that is, in turn, i) processed to deliver a service, and ii) retained to improve the services – how then can that PHI be extracted, nevertheless deleted, from the system? Once the PHI is de-identified and therefore stripped of its protected status, it no longer poses the same privacy concerns that were intended to be protected by Congress with the passage of the Health Insurance Portability and Accountability Act of 1996 ("HIPAA") – but that is only the first hurdle. Although such de-identified data no longer poses the same privacy risk under HIPAA, locating such data so as to extract and delete it from the system, in compliance with CCPA, becomes much more difficult.[23]

As state privacy bills and laws have begun to trend toward a more consumer-friendly approach, there lies an opportunity for a federal certification program to unify state-based data privacy obligations.[24] It is critical that such certification program, together with state obligations, work together to protect consumer data while promoting the advancement of A.I., careful not to nullify the intent or effect of one over the other. Therefore, the certification program should include clear paths forward for state-based compliance. The certification program should also include carve-outs for applications of A.I. that would be exempted from specific obligations under current data privacy laws – i.e., a safe harbor[25] or similar exemption for leveraging artificial intelligence in certain scenarios, e.g., the promotion of and/or the access to care[26] – and offer up a substitute mechanism by which such certified applications would accomplish the spirit of the data privacy laws but by different methodology.

III. CONCLUSION

The implementation of an FTC certification program would address many of the disparate concerns related to artificial intelligence. A certification program would recognize 1) compliant A.I. systems that 2) incorporate domain-specific standards, 3) disclose system biases to consumers and 4) otherwise comply with the OECD Principles on Artificial Intelligence. In addition, the certification program could be used to subsidize data security programs and unify current state privacy bills.

Advance through California Assembly", Chronicle of Data Protection, April 24, 2019, available at https://www.hldataprotection.com/2019/04/articles/consumer-privacy/ccpa-amendments-advance-through-california-assembly/?utm_source=feedburner&utm_medium=feed&utm_campaign=Feed%3A+ChronicleOfDataProtection+%28HL+Chronicle+of+Data+Protection%29 (last accessed by May 26, 2019).

[23] One article published by the Computer Security & Law Review discusses the complexity and practicality of data deletion in the context of A.I.: "Essentially, the Right to be Forgotten [mandated by General Data Protection Regulation 2016/679 (i.e., "GDPR") and effectively the CCPA] applies the human memory metaphor of 'forgetting' information. When individuals request that their personal information be deleted, this is equivalent to metaphorically requesting that others forget that information. However, this metaphor is unique to human minds only and does not translate to the AI/machine learning era. Specifically, the Right to be Forgotten requirements of data deletion do not easily translate because AI does not 'forget' data in the way that humans do. Data deletion in artificial intelligence contexts is much more complex … [E]very data record added to the database might not only reside at one specific point in the file system, but might be stored at various locations inside internal database mechanisms, as well as across different replicated databases, in log-files and backups. When the Right to be Forgotten asks for permanent deletion of the data, these requirements must be taken into account. When asking for deletion in a strict sense, these spaces must be identified and overwritten with random information. In several internal mechanisms like the database transaction log, the latter is especially impossible without seriously endangering the consistency of the database, or even simply breaking it altogether." Fosch Villaronga, E. et. al., "Humans Forget, Machines Remember: Artificial Intelligence and the Right to Be Forgotten", Computer Security & Law Review (Forthcoming), Aug. 13, 2017, available at https://ssrn.com/abstract=3018186 (last accessed June 5, 2019).

[24] Currently, the United States does not have a federal privacy law or set of laws that regulate data privacy, except for specific industries such as healthcare and finance. Rather, data privacy is regulated by state law. The result is a patchy, i.e., sometimes contradictory, framework of data privacy laws that can be difficult to comply with state by state. No doubt, "[i]t is past time for Congress to create a single legislative data-protection mandate to ... reconcile the differences between state and federal requirements" (Council on Foreign Relations, "Reforming the U.S. Approach to Data Protection and Privacy", Jan. 30, 2018, available at https://www.cfr.org/report/reforming-us-approach-data-protection (last accessed June 5, 2019)), but in the meantime the proposed certification program could suffice. If and when the issue of federal data privacy is settled and legislation is passed, such certification program could be adjusted accordingly.

[25] A safe harbor is a "provision granting protection from liability or penalty if certain conditions are met. A safe harbor provision may be included in statutes or regulations to give peace of mind to good-faith actors who might otherwise violate the law on technicalities beyond their reasonable control." "Definition of Safe harbor", Legal Information Institute, Cornell Law School, available at https://www.law.cornell.edu/wex/safe_harbor (last accessed May 26, 2019).

[26] In the healthcare sector, for example, a safe harbor exists with respect to the federal Anti-Kickback Statute (42 U.S.C. 1320a-7b) for promoting access to care. In short, the Anti-Kickback Statute prohibits the knowing and willful payment, or offer of payment, of any remuneration to induce or reward business referrals with respect to any good or service payable by a federal healthcare program, such as Medicare or Medicaid. The safe harbor applies in instances where the Office of the Inspector General, as authorized by the U.S. Dept. of Health & Human Services to enforce the Anti-Kickback Statute, determines that remuneration poses a low risk if the items or services at issue 1) are unlikely to interfere with, or skew, clinical decision-making; 2) are unlikely to increase costs to federal healthcare programs or beneficiaries through overutilization or inappropriate utilization; and 3) do not raise patient safety or quality of care concerns. *See, e.g.*, OIG Advisory Opinion No. 19-02 at 7, https://oig.hhs.gov/fraud/docs/advisoryopinions/2019/AdvOpn19-02.pdf.