Volume 5, No. 1 | January–February 2022

RAIL

The Journal of Robotics, Artificial Intelligence & Law

fastcase FULL COURT PRESS

RAIL
The Journal of Robotics, Artificial Intelligence & Law
Volume 5, No. 1 | January–February 2022

THE JOURNAL OF ROBOTICS, ARTIFICIAL INTELLIGENCE & LAW (ISSN 2575-5633 (print) /ISSN 2575-5617 (online) at $495.00 annually is published six times per year by Full Court Press, a Fastcase, Inc., imprint. For customer support, please contact Fastcase, Inc., 711 D St. NW, Suite 200, Washington, D.C. 20004, 202.999.4777 (phone), 202.521.3462 (fax), or email customer service at support@fastcase.com.

Publishing Staff
Publisher: Morgan Morrissette Wright
Production Editor: Sharon D. Ray
Cover Art Design: Juan Bustamante

Cite this publication as:

The Journal of Robotics, Artificial Intelligence & Law (Fastcase)

A Full Court Press, Fastcase, Inc., Publication

Editorial Office

711 D St. NW, Suite 200, Washington, D.C. 20004
https://www.fastcase.com/

POSTMASTER: Send address changes to THE JOURNAL OF ROBOTICS, ARTIFICIAL INTELLIGENCE & LAW, 711 D St. NW, Suite 200, Washington, D.C. 20004.

Articles and Submissions

Direct editorial inquiries and send material for publication to:

Steven A. Meyerowitz, Editor-in-Chief, Meyerowitz Communications Inc., 26910 Grand Central Parkway, #18R, Floral Park, NY 11005, smeyerowitz@meyerowitzcommunications.com, 631.291.5541.

Material for publication is welcomed—articles, decisions, or other items of interest to attorneys and law firms, in-house counsel, corporate compliance officers, government agencies and their counsel, senior business executives, scientists, engineers, and anyone interested in the law governing artificial intelligence and robotics. This publication is designed to be accurate and authoritative, but neither the publisher nor the authors are rendering legal, accounting, or other professional services in this publication. If legal or other expert advice is desired, retain the services of an appropriate professional. The articles and columns reflect only the present considerations and views of the authors and do not necessarily reflect those of the firms or organizations with which they are affiliated, any of the former or present clients of the authors or their firms or organizations, or the editors or publisher.

QUESTIONS ABOUT THIS PUBLICATION?

For questions about the Editorial Content appearing in these volumes or reprint permission, please contact:

Morgan Morrissette Wright, Publisher, Full Court Press at mwright@fastcase.com or at 202.999.4878

For questions or Sales and Customer Service:

Customer Service
Available 8 a.m.–8 p.m. Eastern Time
866.773.2782 (phone)
support@fastcase.com (email)

Sales
202.999.4777 (phone)
sales@fastcase.com (email)
ISSN 2575-5633 (print)
ISSN 2575-5617 (online)

Archimedes' Lever and Audience Participation—or—Multifactor Soft-Law Transparency for AI System Process Development

James A. Sherer*

This article explores, examines, and challenges one factor discussed and considered consistently within varying approaches—"Transparency"—in the context of "weak" artificial intelligence ("AI," but AI that is weak only in terms of complexity, and in comparison with human beings). The article notes that considering Transparency in the context of AI is not a special consideration, as many previously unknown technologies are assailed with demands for Transparency upon their introduction into the marketplace of ideas. The article addresses and rejects a standard of completely open Transparency for privacy, security, and practical reasons, and presents a framework of more general principles that propose to contour to real-world instances. It argues for the use of Soft Law as the only realistic approach to surfacing appropriate levels and types of Transparency, but rather than solving the issue (which the author declares it most decidedly does not do), it offers up its presented framework to specific tasks at hand as a tool to help drive consensus for the development, adoption, utilization, and judgment of AI.

Archimedes is said to have proclaimed, "give me a place to stand and with a lever I will move the whole world."[1] This is, of course, geometric reasoning, but as with any mechanical advantage, the device, here a lever, preserves the input and simply trades off forces against movement to amplify the output force. The conservation of energy remains the same, as do the input and output energies, and effort X leads to effect X, no matter the desire of the person putting in the effort.

But with technology, effects can scale. Instead of geometric reasoning, technology can present geometric or (more commonly referred to as) exponential growth, depending on the context (e.g., compound interest).[2] An input, a force, *can* be multiplied. And with artificial intelligence ("AI," defined further below), C-suite executives and other decision-makers have been encouraged to

Robotics, Artificial Intelligence & Law / January–February 2022, Vol. 5, No. 1, pp. 9–33.

ISSN 2575-5633 (print) /ISSN 2575-5617 (online)

multiply the impact of AI with wild abandon, with some commentators going so far as to offer the (admittedly radical) solution of killing the proof-of-concept approach and moving straight to scale.[3]

Surely this will not go awry.

Incorporating scaling into the discussion requires considerations and concerns of near-exponential effects, where humans are toying with technologies that operate like "ever-lengthening levers," but levers that have (or had) a human decision-maker on one end. [4] However capable those decision-makers are often making decisions focused on the realm of the immediate while awash in the analog; there is no surfeit of Archimedes-like programmers creating programs or stacking existing solutions in the hope that the next iteration of solutions produces an answer—usually in response to a business question. As noted by Sawyer [5]:

> The development of AI is a business, and businesses are notoriously uninterested in fundamental safeguards—especially philosophic ones. (A few quick examples: the tobacco industry, the automotive industry, the nuclear industry.) Not one of these has said from the outset that fundamental safeguards are necessary, every one of them has resisted externally imposed safeguards, and none have accepted an absolute edict against ever causing harm to humans.

There are arguments that the issue of scale and of (un)intended consequences can be addressed by better training in logic, but while schools and training programs are making improvements in that realm, those improvements are uneven and unlikely to serve as a panacea here.[6] So, if humans cannot or simply *should* not trust the creators and/or those decision-makers pressing their AI creations into scaled service, humans logically require a different approach.

At its core, AI represents a "fundamental leap in human progress," but we are cautioned to remember the moral of the be-careful-what-you-wish-for stories, where it is never the genie that is dangerous, but the unskilled wisher.[4] How, then, can Transparency assist humans in becoming skilled wishers, and cautiously flex the lever that may move the world much further than the strength of the lever-er? And what specific considerations regarding Transparency might be carefully applied during the design process to support this goal?

Note first, and importantly, an approach considering Transparency is among the hallmarks of new and previously unknown technologies, and so a related exploration should be unsurprising and *not* new or esoteric. As noted previously [7], [8], a call for Transparency is, to a certain extent, both generic in debates on new technologies and a hallmark for how society generally comes to terms with new impactful approaches (*see also* [9], [10]). And AI has not escaped this treatment thus far, as in fact, (1) a large number of ethicists, practitioners, journalists, and policymakers have agreed that some standards should guide the development of a technology, and (2) leading software developers, the World Economic Forum, the draft AI ethics guidelines for the EU commission, and many, many others all promote Transparency or explicability for AI in connection with other appropriate considerations [11].

Those august bodies that have considered it seem, ultimately, to be in accord on one specific point, and agree that Transparency should *not* be addressed with one correct, objective approach in mind; rather, Transparency should orient toward an audience-specific approach. That is not to say that discussions and debates have not surfaced consistent ideas of Transparency, some of which have been further applied to the context of AI and its use. And Transparency in this context also considers how Soft Law can support constructive Transparency during the process of developing AI. It argues, in fact, that Soft Law is particularly effective in a quickly moving arena of technological advancement.

Soft Law is particularly apt for this application because of its breadth and tractability. Soft Law is, conceptually and as noted by Marchant [12], a reference to frameworks that provide for substantive expectations and rules and that incorporate professional guidelines and best practices, codes of conduct, and other private standards without integrated regulatory enforcement mechanisms. As Soft-Law approaches are not constrained by regulatory rulemaking processes, they proliferate much more rapidly and currently number in the hundreds [13]. As noted by Shafer & Pollack [14], because Soft Law instruments provide greater flexibility for states and other nongovernmental agencies to cope with uncertainty and learn over time, Soft Law instruments present a more logical approach where the challenge presented requires the ability to adjust expectations in the event of changed circumstances [15]. Finally, Soft Law incorporates inherent suitability because it can

expand to fit the AI landscape that presents a unique complexity, both within approach as well as amplification by effects [16].

Some researchers [17] have utilized an approach that generally considers Transparency as a concept, a spectrum along which samples can be taken at different times for different constituents, framing Transparency's richness in application and also making it difficult to define. This specifically relates the concept of Transparency to the fact that, in some scientific fields, Transparency denotes the literal physical property of a material and its capacity to allow light to pass through it, while in others it is thought of as a tool or powerful means in service to a desirable social end, such as accountability for public officials, reducing fraud, and eliminating corruption [4].

Here, while both are helpful, the second point, while not oriented toward the public, is slightly more directed toward the result of improving processes by embedding Transparency throughout the process. And there is still hope that embedding some type or degree of Transparency into the process will make a difference, as humans are collectively still in the early days of AI [19].

Discussion

What Is AI (Here)?

This article does not discuss AI in detail. Here, it is sufficient to address it as a system that is typically deployed to generate insight into matters which are too complex for human minds to understand [7]. Or perhaps as a departure from traditional programming, a system or approach that allows people to solve a problem even if they cannot think up the solution's steps themselves [4]. And, as it is not uncommon to encounter AI solutions that are cobbled together, where algorithm approaches are built in sequence through interconnected steps marching through iterations of data analysis, where a lack of planning can render them opaque in nature and incomprehensible, even by those individuals tasked with developing or utilizing them [20], those solutions are considered as well. In sum, all those systems are at play, and considerations regarding Transparency still apply. Ultimately, the definition of AI can be distilled into a system that, regardless of how it was cobbled together, is a combination of what the programmer and implementer wanted.

And what they did.

Notably, when examining AI systems, the considerations for Transparency do not only include the desires of the programmers and implementers and the math they used. They must also include the type of data used, as AI works best when it begins with real-world examples and, in some instances, truly personal information [21]. An approach that implicates full-blown Transparency might also, depending on the application, examine the underlying data which itself could implicate existing privacy rights for the individuals whose data is used. Sacrificing individual privacy for the sake of society-wide medical advancements might be the objectively "right" approach when saving lives [22], but a general approach that immediately undercuts or jeopardizes existing rights is a nonstarter [23].

The stakeholders in the Transparency discussion are important as well. Those include the people or organizations using AI; other AI using the AI (*see* earlier discussion regarding stacking solutions); those people or organizations on whom AI is used; governments, regulators, or other nongovernmental organizations invested in or seeking to enforce rules or norms; and, frankly, society as a whole or people generally. Of course, as noted above [17], this is all a matter of degree, depending both on the audience and the insight warranted. That is, what the audience calls for and/or perhaps what the audience deserves. And note finally that explicability itself in this context is not a given, based on the technology as well as people's limitations generally [11].

When Can Transparency Be Less Than Beneficial or Even *Gasp* Disastrous?

First, too much design focus on Transparency can affect the efficiency of a system [20]. If Transparency is included simply for Transparency's sake, then it may inadvertently limit the design of the system, hamstring the ultimate effectiveness of the system, or cause some other unintended consequences.

Second, some people generally like to do bad things, and by virtue of that fact, Transparency is not a universal good [24]. Commentators are therefore in accord that Transparency applied to AI requires an appropriate, metered approach that varies by audience. Transparency can certainly show people how to purposefully do

things that society generally considers bad, including the ability to leverage such technologies and scale them for outsized ill effects [25]. Transparency can also create a platform for easy access to technology that would allow bad actors to steal intellectual property [19]. And for those systems that matter to society as a whole, where health, equality, and general human well-being are supported, releasing the complete source code of an AI system can make it easy for enterprising users to game the system [17].

Additionally, some people specifically want to steal secrets; therefore, general business norms and associated laws and regulations allow organizations to protect their intellectual property [26]. It has been noted that any reasonably successful organization even attempting to maintain ownership and control over its methods will marshal such information, keeping its methods secret to reduce attendant controversy and foil copycat competitors [26]. Vollmer et al. [27] agree, and in this context introduce the concept of balance between the need for Transparency and the risk of undermining intellectual property rights. Wischmeyer and Rademacher [7] also support the general protection of intellectual property rights and trade and business secrets, and note that these are further enshrined—or at least respected—in the European Union's General Data Protection Regulation ("GDPR").

Some other people do bad things, but they do them by accident. Unfortunately, a loaded gun is still a loaded gun when in the wrong hands. This may happen even where there are the best of intentions—the creators or users of these systems present the "magic lamp" challenge, where users (and builders) of these approaches are utilizing a truly new kind of approach to problem-solving that does what it is told or coded to do, but that may not be exactly what was wished for. And sometimes we literally face a future of lethal autonomous weapons [28]. This is a problem that only gets worse as the technologies can scale, and it is compounded further when there is too much information without enough knowledge. It introduces the law of unintended consequences or the "cobra effect," where an attempted solution makes a current problem even worse [29].

In direct opposition to the dystopian future contemplated by AI warfare, sometimes the proverbial juice just is not worth the squeeze. It is a fact that some AI, when used every day for low-risk purposes, does not or should not require explanation [11]. This might include those less-than-vital decisions that are delegated to

machines, such as internet search results, product recommendations, matches on dating sites (perhaps debatable), and, at least prior to modern-day election issues, the content that arrives on our social media feeds [24].

Finally, Transparency as an absolute, stand-alone standard probably will not accomplish any of the stated goals anyway. A presentation on the Transparency of AI, including machine learning and deep learning algorithms as well as predictive analytics, often centers too much on opening the black box, and presenting just the "facts" without context simply does not go far enough [30]. That is, opening the hood of car, without more context, explanation, and understanding, does not a mechanic make. Therefore, scholars have noted that Transparency is ultimately not an end but a means, an interim step on the road to intelligibility [26]. The concern is that when companies or individuals undertake work that involves complexity too great to monitor or understand, disclosure becomes an empty gesture without effect [26].

With That Addressed, Why Is Transparency Good?

People are curious and like to understand things generally, and the utilization and implementation of AI (as well as other endeavors) should support that fundamental human desire. Humans also like to understand one other, and recent scholarship supports the idea that people are often interacting with AI as though it were a person. In that same vein, authors have examined the efficacy of Transparency efforts or deciphering the behavior of the "intelligent other," intentionally (or not) raising the level of discourse to considering machines as smart decision-makers [31]. And despite questioning Transparency as an end-all solution, de Fine Licht & de Fine Licht [30] also note that to both prevent negative outcomes and create trustworthy accountable systems, developers may need to open the "black box" of AI decision-making and make it more transparent.

A degree of Transparency is also a measure of protection for people. General proposals often include Transparency as part of a basket of protective measures for individuals, and this is key to why, as mentioned above, considerations of Transparency are among the requisite steps to examining new technologies. Some researchers have gone so far as to include Transparency as part

of a "duty of loyalty" that would legally require covered entities to respect the privacy of individuals, including by implementing measures for data minimization, fairness, and Transparency [32], but short of legal requirements, Soft Law adoption may accomplish something similar.

There is currently a lack of Transparency that should be remedied. Researchers and regulators alike have noted this, pointing to high-profile publications that have shown a lack of Transparency, replicability, ethics, and effectiveness in the reporting and assessment of AI-based prediction models [27]. The Federal Trade Commission of the United States ("FTC") explicitly stated that there have been a number of uses of doppelgängers, whether a fake dating profile, phony follower, deep-fakes, or an AI chatbot, that have misled consumers. The FTC followed that with an admonition that such duplicitousness could face FTC enforcement actions [33].

When AI systems have control over human behavior and, by extension, human lives, lying—and concealing or obfuscating those lies—is wrong unless there is an even more powerful reason for the behavior. This is especially true when AI systems utilize and/or are derived from personal data, as the more personal and authentic the data used to build automated systems—especially when used to influence behavior—is, the less transparent they appear, and the more they affect unconscious and irrational cognitive or intentional human processes [7], [34].

Considered at a level higher than the movement of the math, then, requiring Transparency or injecting it into information systems may help systems work better for humans. Transparency may also help users and subjects to better understand that the use of AI is not a panacea; AI programs, no matter how complicated, are still unthinking effects like forces of nature, the ultimate reliable workers who do only what they are told, no more and no less. Reinforcing the fact that AI systems are tools that do not think for themselves (and of course, do not, at least at this point, think at all) may remind human users engaged with the technologies that the AI systems do not, in fact, want anything except what the programmer or interface told them to want [4]. There are also thoughts that a lack of Transparency can lead to mistrust and uneven adoption [35]. In contrast, including Transparency should lead to greater stability of an overall system because of a shared approach that is, by design, disclosed to the participants within the system as well as to the individuals ultimately affected by it.

And while protection of intellectual property or trade secrets might seem to have no foes, Pasquale [26] is one of many who criticize a one-size-fits-all approach, stating that while secrecy may be a wise business strategy, it devastates humans' ability to truly understand the social world, where opacity creates a number of opportunities to hide anticompetitive, discriminatory, or simply careless conduct behind a veil of technical inscrutability. His argument there ends with an assertion that, without Transparency, accountability is impossible.

While some commentators have argued (above and before) that efficacy will suffer with Transparency [20], it may well be that "transparent" deep-learning models may allow programmers and users to better understand how models process signals, better explain predictions, and gain better knowledge regarding reasons behind task failure [36]. In fact, without Transparency, building effective platforms and approaches suffers. As noted elsewhere, there is a growing body of evidence suggesting that while many current, best-practice recommendations for design, conduct, analysis, reporting, impact assessment, and clinical implementation borrow from traditional statistics literature, such recommendations may not be sufficient to guide the use of machine learning ("ML")/AI in research.

Vollmer [27] and Harford [37] also observed that a theory-free analysis of mere correlations is inevitably fragile, where the lack of reasoning behind a correlation means a lack of understanding of what might cause the correlation to break down. In this context, current regulatory regimes are insufficiently dealing with the issue. Vollmer [27] noted the same, when complying with legislative frameworks (e.g., the GDPR) is necessary but is insufficient to show the Transparency required to produce trustworthy ML/AI research.

Injecting Transparency into the approach at least to some degree should also improve results when, in certain types of systems, it is more challenging to "visualize" the hidden representation of data because it is difficult for humans to process patterns that do not have familiar visual cues. This manifests in the argument that there is an urgent need for visible neural networks that help make the hidden representations otherwise "transparent" from the perspective of investigators in order to provide a direct view of what hidden representations stand for [36].

There is certainly industry movement based on the general population's expectation, and a number of associated practices and

frameworks consider Transparency, including article four of the Engineering and Physical Sciences Research Council ("EPSRC") Principles of Robotics, which asserts that "Robots are manufactured artefacts. They should not be designed in a deceptive way to exploit vulnerable users; instead their machine nature should be transparent" [38].

Also, people will be happier with the results if they believe they can comprehend the process. One team stated that, based on their own literature survey, a limited form of Transparency focused on providing justifications for decisions has the potential to provide sufficient grounds for perceived legitimacy in AI decision-making as well [30].

Industry movement based on general population or "customer" expectation can also lead to actualization and generalized industry practices. These have manifested in advertising and marketing materials, where some now routinely suggest that Transparency is a primary concern for those companies utilizing technologies, focused on goals including bringing the benefits of analytics to business users as seamlessly and transparently as possible, without disrupting their day-to-day activities [39]. But while the use of AI in marketing and advertising can be arguably benign (e.g., a choice of fragrance or shampoo based on a consumer's visited website), the same algorithm used under slightly different circumstances (e.g., serving up advertising for an employment vacancy from the same website lead-in) could be categorized as a high-risk AI system and enforced against accordingly [40].

Customer expectation can also lead to enforcement based on such expectations, as seen in the FTC's general pronouncements [33], [41], [42]. Other enforcement is likely forthcoming as well. Felzmann [43] comments that the origins of the Transparency requirement in data protection law date to the 31st International Conference of Data Protection and Privacy Commissioners held in Madrid in November 2009, during which the importance of Transparency to protect an individuals' privacy was acknowledged. An extrapolation from privacy combined with a built-in Transparency requirement is easy to draw. Note too that the FTC's April 8, 2020 guidance states that the FTC's law enforcement actions, studies, and other guidance emphasize that the use of AI tools should be transparent, explainable, fair, and empirically sound while also fostering accountability [33]. Likewise, others have stated that Transparency has always been a general principle of data protection, and as seen

in the GDPR and elsewhere, lawmakers around the globe are experimenting with specific Transparency requirements for automated decision-making systems, including AI-based systems [7].

Assuming They Are Good, at Least to a Degree, What Do Transparency Proposals Look Like?

Perhaps unsurprisingly, researchers and theorists vary in approach, and so they either bundle or bifurcate Transparency depending on perspective. One Transparency approach scholars have considered for AI requires an intelligent system to be "self-explanatory," such that an operator can be confident about what an AI system is doing and why [31]. A self-explanatory or "axiomatic" approach would also widen the potential audience, in support of a normative Soft Law approach, that of increased openness and participation, which in turn would promote Transparency [44]. But note that Soft Law approaches have proliferated in recent times, and as of this article's drafting, there are a number of compendia available online, including the AlgorithmWatch that compiled and maintains over 170 AI ethics guidelines [45], the European Union Agency for Fundamental Rights' work on collecting and hosting 350 AI policy initiatives between 2016-2020 [46], and work done by Gutierrez & Marchant compiling a database representing over 600 Soft Law programs directed at AI [12].

Within the materials and research considered for this work, Lipton [47] further defines a transparent explanation as reflecting to some degree how a system came to its output, where inherent in the simulatability of a given AI system is a requirement that a given person can grasp the entire model. This, in turn, implies that explanations will reflect the inner workings of a system. Even though some *post hoc* explanations might not elucidate precisely how a model works, they might still confer useful information for practitioners and end users of machine learning. These types of approaches could (or might be required to) be built into the system itself, where part of the program could focus on disclosing the state of the AI system (as can occur in nature) [36]. This would also support an audit-type approach after the AI system has performed, as Kerry et al. [32] and Walker [48] have argued that Transparency should be linked to accountability in support of algorithmic decision-making.

This matches other scholastic approaches that have suggested a one-factor construct or approach for general application. That is, rather than different information for different constituencies, Walker [48] advocates for intelligible Transparency available to ordinary, intelligent, educated citizens that cogently presents a legally defensible applied AI system.

Larsson [17] goes further and presents an additional query, introducing the "key question" considering "for whom the AI-systems or algorithmic decision-making should be more transparent." Lipton [47] then distinguishes Transparency into two broad concepts or categories for explainable: transparent and *post hoc* interpretability.

Kerry et al. [32] proposed a three-layer Transparency approach, first arguing that a one-size-fits-all Transparency mechanism is exactly the wrong direction, and instead presents a program including timely, context-specific notifications for individuals; basic privacy statements targeted to individuals (underscoring the data supporting many of these programs); and comprehensive privacy disclosures aimed at regulators and other close observers.

Others agree with a three-factor approach in some way, including de Fine Licht & de Fine Licht [30], who noted that Transparency regarding public decision-making should be broken into three factors at play. Those factors focused on the distinctions that can be made between Transparency that informs (e.g., the public or the general population) (1) about final decisions or policies; (2) about the process resulting in the decisions (Transparency in process); and (3) about the reasons on which the decision is based (Transparency in rationality). And Pasquale [26] argued that any Transparency solution to black box problems should be specific about three main issues: How much does the black box firm have to reveal? To whom must it reveal it? And how fast must the revelation occur?

Phillips et al. [49] are more discrete in approach and propose a concept of Transparency in which principles lead to a framework that addresses different components of an explainable system. Phillips' [49] four principles are (1) that a system produces an explanation, (2) that the explanation is meaningful to humans, (3) that the explanation accurately reflects the system's processes, and (4) that the system also provides its knowledge limits. This idea of a required explanation, or more broadly, the concept of "Explainable AI" or "XAI" is further explored in other research considering that perspective and arguing strongly for some logic

in approach that can, in turn, be grasped by a human considering a system and its operation. Doran et al. [50], a proponent of XAI, specifically posits that achieving trustworthiness in a system may require detailed explanations of the decisions made by the AI—and, by extension—the AI itself. Hagras [51] asserts that a successful general introduction of AI across industry and government will require the use of XAI for greater user transparency, and Ignatiev [52], another proponent, pushes further, arguing that heuristic explanations for algorithm operation (specifically ML models) may be incorrect as a general concept, whereas only those approaches that are provably correct and provide rigorous explanations are in turn trustable.

In contrast to a more global approach and with even further nuance in mind, Felzmann et al. [43] have different stakeholders, as adapted from earlier work. The Felzmann [43] approach begins with an AI system's developer, where the system must be sufficiently transparent to allow the developer to understand whether the system is working properly so the developer can subsequently identify and remove errors from the system or improve it. The Felzmann [43] approach then considers Transparency for a user, for whom Transparency must be sufficient to provide a sense of *what* the system is doing and *why*, which should further enable intelligibility of future unpredicted actions or circumstances and should build a sense of trust in the technology. The Felzmann [43] approach would also require Transparency sufficient for a user to understand why any particular decision was reached; to allow a check that the system worked appropriately; and to enable a meaningful challenge to an outcome, especially one of societal importance (e.g., credit approvals or criminal sentences).

The Felzmann [43] approach considers three more stakeholders with varying Transparency requirements. First, society at a broad level should have sufficient information to understand and become comfortable with the strengths and limitations of a given system, with enough Transparency for the general public to overcome a reasonable fear of the unknown (e.g., autonomous cars or robotic surgeons) and embrace a situation with probable societal benefits. For experts and regulators, there should be the requisite degree of Transparency to provide a knowledge base to audit a prediction or decision trail in detail, particularly (un)intended harmful actions (e.g., a crash by an autonomous car). Finally, the deployers of a solution should have sufficient Transparency to make them feel

comfortable with a prediction or decision, so that they keep using the system.

Of course, within each approach, no matter how nuanced, there are further considerations regarding exactly how Transparency should operate or be judged. For example, when considering efficacy (and reproducibility), Vollmer [27] notes that the reproducibility of a process requires questioning whether there is organizational Transparency about the flow of data and results. Ernst [34] answers, looking at indirect mechanisms regarding Transparency that provide disclosure rules or guidance to persons concerned that also incorporate ideas regarding the Transparency of underlying data, Transparency of the underlying intent, and whether the end product is auditable. And finally, Cheatham et al. [53] note the applicability of using a structured identification approach to pinpoint those risks with heightened criticality, which should work for Transparency as an approach.

Why and How Soft Law Is the Right Avenue

In this article, Soft Law instruments are not "sources of law," such that, as noted by Demin [54], Soft Law instruments do not contain legally binding rules, are not backed by public sanctions, and are not unilaterally enforceable. However, they have significant practical importance and may sometimes bring about definite legal consequences. Further considered, Senden's [55] definition of Soft Law focuses on those rules of conduct that are provided in instruments which have not been attributed legally binding force as such, but nevertheless may have certain—indirect—legal effects, and that are aimed at and may produce practical effects.

As Senden [55] also notes, among types of Soft Law (e.g., preparatory and informative instruments, interpretative and decisional instruments, and steering instruments), proposed standards might be formal and nonformal. These instruments can aim at establishing or giving further effect to community objectives and policy or related policy areas, sometimes in a rather political and declaratory way (e.g., in declarations and conclusions) but often also with a view to establishing closer cooperation or even harmonization in a nonbinding way, including recommendations, resolutions, and codes of conduct.

The research on Transparency and approaches within AI systems suggests that there is certainly not a consensus at this time, and since we are collectively at the dawn of this technology's use and the technology and approaches seem to outpace (or at the very least keep pace with) the scholarship, hard and fast, bright-line approaches are simply untenable. Soft Law standards have the benefit of nearer-term implementation and can incorporate flexibility unavailable to other means of attempted direction and control.

Soft Law is the correct approach, in lieu of or even in contrast to a universal or global standard, because of the flexibility inherent in a Soft Law approach. The research indicates that we collectively do not know how rapidly AI models will evolve, or which portions of AI practice will take veer off on an orthogonal or unexpected direction. The consensus is only that whatever happens, its effects will be amplified. Soft Law's "softness" is therefore a benefit in these circumstances. A hardline approach cannot accommodate the known unknown—that we only know that we don't know what will happen. Soft Law instruments are best positioned then to accommodate the unexpected and to morph and change to hew most closely to the original intent of the drafters and supporters.

What Soft Law Should Consider When Developing Transparency Standards—and a Proposed Path Forward

As noted by Trubek, Cottrell & Nance [44], the considerations presented regarding Soft Law and related ethical approaches have focused first on increasing Transparency among the purposes for rules and institutional functioning. Further, as posed by Zeitlin [56], they have worked to link participation by individuals together with Transparency to increase the effectiveness of a Soft Law approach.

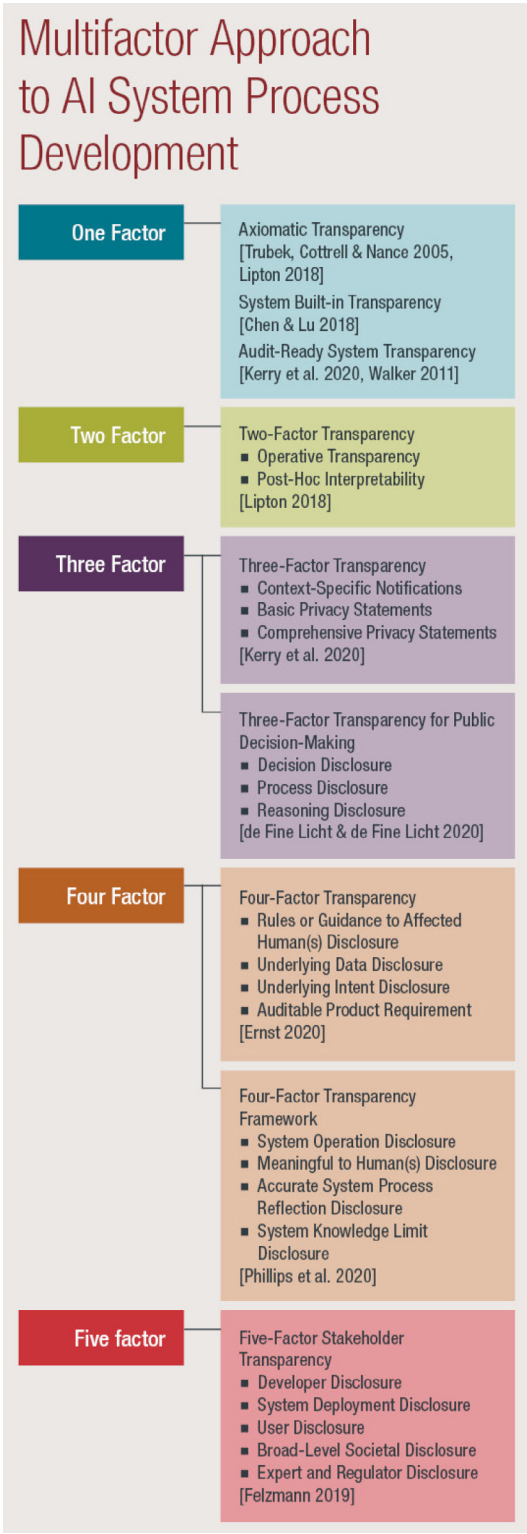
One of the underpinnings of a Soft Law approach here is, in fact, Demin's [54] understanding that publication of the rules allowing for Transparency is necessary so that the rules may be studied and be well-known to the appropriate agents. Employing Soft Law within a structure where the first-line appropriate agents (e.g., the developers) will read and recognize the guidance and, perhaps, feel bound to consider and incorporate it, is the most effective approach presented within the current literature.

Further, as Senden [55] notes, Soft Law is sometimes seen as complementary to traditional command-and-control legislation. This is well and good if Soft Law regimes move collectively in similar or complementary directions; unfortunately, different Soft (and Hard) Law regimes can operate antagonistically if there are other power considerations at play.

Forgive a final parable to provide a proposed path forward. Recall the story of the wise woman and a glass jar as related in Canessa et al. and translated by Sidore [57], where large stones are dropped into the jar until no more will fit. The woman asks a child if the container is full, and when the child answers in the affirmative, the woman adds smaller rocks that tumble into the remaining space. The child opines that the jar is now full, and the woman repeats the process with pebbles, sand, and finally water. As Sidore [57] explains anew, each "new material, smaller and finer than the one before, closes the ever-shrinking gaps until truly the jar can hold no more."

Such an overall approach can provide a best-of-breed *as appropriate* that acknowledges Larsson's [17] Transparency "spectrum," and considers all the approaches and picks those that are warranted. This should address the more general human need for a modicum of Transparency [7], [34], and then allow a graduated approach following and depending on the proposed system itself.

This would progress a given AI system's stakeholders through the following steps of process development:


Multifactor Approach to AI System Process Development
One Factor
Axiomatic Transparency [Trubek, Cottrell & Nance 2005, Lipton 2018]
System Built-in Transparency [Chen & Lu 2018]
Audit-Ready System Transparency [Kerry et al. 2020, Walker 2011]
Two Factor
Two-Factor Transparency
Operative Transparency
Post-Hoc Interpretability
[Lipton 2018]
Three Factor
Three-Factor Transparency
Context-Specific Notifications
Basic Privacy Statements
Comprehensive Privacy Statements
[Kerry et al. 2020]
Three-Factor Transparency for Public Decision-Making
Decision Disclosure
Process Disclosure
Reasoning Disclosure
[de Fine Licht & de Fine Licht 2020]
Four Factor
Four-Factor Transparency
Rules or Guidance to Affected Human(s) Disclosure
Underlying Data Disclosure
Underlying Intent Disclosure
Auditable Product Requirement
[Ernst 2020]
Four-Factor Transparency Framework
System Operation Disclosure
Meaningful to Human(s) Disclosure
Accurate System Process Reflection Disclosure
System Knowledge Limit Disclosure
[Phillips et al. 2020]
Five factor
Five-Factor Stakeholder Transparency
Developer Disclosure
System Deployment Disclosure
User Disclosure
Broad-Level Societal Disclosure
Expert and Regulator Disclosure
[Felzmann 2019]

Conclusion

A sifting process applied by developers and decision makers would help find the appropriate level for a given AI solution, and even the exercise should prove valuable for developers and stakeholders alike.

When specifically considering its application in real-life situations, utilization of the proposed Multifactor Approach fits within current proposed governmental requirements. The Multifactor Approach arguably meets current requirements within the United States, especially within the FTC's admonition [33] that organizations using AI with material effects should carefully evaluate and select the correct AI transparency framework or independent standard, and the FTC's reiteration [41] that application of appropriate, recognized AI transparency frameworks may be required. The FTC's particular focus on supportable frameworks stems back to and draws upon the FTC's "Big Data" analysis [42] but has evolved over time to provide for considered application of such process development frameworks, and provides directly for the use of frameworks as part of an organization's practical approach to responsible AI deployment [58].

The proposed Multifactor Approach also seems to meet current proposed approaches and regulation. Specifically, in 2020, the High-Level Expert Group on AI ("AI HLEG"), set up by the European Commission, promulgated The Assessment List for Trustworthy Artificial Intelligence ("ALTAI") for self assessment [sic] [59]. Within the ALTAI, the AI HLEG reiterated its seven requirements for Trustworthy Artificial Intelligence, which included Transparency at Number 4 [59]. Use of the Multifactor Approach could address the manner in which HLEG encourages organizations to self-assess the explainability of an AI system, as well as other supporting explainability measures, including traceability and auditability [59].

In addition, the European Commission's ("EC's") 2021 Proposal for the Artificial Intelligence Act (the "Proposal" for proposed regulation) incorporates room for, and a reliance upon, additional technical standards that will help operationalize the EC's proposed, common mandatory requirements [40]. The EC's Proposal also considers harmonized standards that work in connection with supporting guidance and compliance tools as part of its proportionality considerations; considers approaches to high-risk AI systems and

asserts that precise technical solutions may be provided by standards, other technical specifications, or development in accordance with general engineering or scientific knowledge; provides for common normative standards for those same high-risk AI systems; and proposes that its suggested European Artificial Intelligence Board would itself provide guidance on the proposed regulation in consultation with existing standards [40].

Finally, the EC's Proposal states that certain AI systems will require record-keeping and logging, and such activities will conform to recognized standards or common specifications [40], which speak specifically to the audit-ready nature, disclosures, and auditability built into all of the factored approaches within the proposed Multifactor Approach for AI System Process Development.

Assuming that the current United States and proposed EC regimes are accurate representations of how more general AI issue enforcement will operate, a strategy for utilizing the Multifactor Approach would incorporate fit and focus mechanism. That is, either internal stakeholders responsible for the selection and use of an AI model strategy, and/or a third-party subject-matter-expert or audit function evaluating the AI model strategy as envisioned by the 2021 FTC guidance would select an approach based on certain criteria, and then execute against it [41], [58]. Specifically, internal stakeholders would examine the complexity of a proposed AI model, the relative complexity of the organizational strategy seeking to utilize the model, and the potential impact of the AI model in the world-at-large. And in particular, this last point—the overall potential impact—would be informed by the applicable regulatory guidance, as the EC's 2021 Proposal specifically outlines which effects (and, by extension, which AI systems) are high-risk [40]. This calculus would then inform the stakeholders how many factors within the Multifactor Approach might be warranted, with the examination of higher-number factors encouraged in instances where the effects are amplified, or the effects are by nature, high-risk.

Returning to the example of the same AI model used to sell fragrances, shampoo, or employment opportunities, the AI model in each of those instances would be the same. However, because employment as a concept might raise high-risk AI system concerns, an internal stakeholder or external subject-matter-expert might opt against a One Factor approach even in the face of a relatively simple AI model.

Perhaps in the face of these considerations, an organization determines every high-risk AI model requires at least a four-factor approach. In that case, the variance for this fact pattern might be simply whether intent matters, in which case the Ernst Four-Factor Transparency approach would be warranted, and perhaps the Phillips et al. Four-Factor Transparency Framework would not.

In overall conclusion therefore, given both the imprecise nature of general AI development and utilization, the need to create buy-in among participants utilizing the system, and the evidence presented to this point regarding current and proposed regulatory enforcement, the Soft Law considerations represented in the proposed Multifactor Approach makes it a strong (and nimble) candidate for Soft Law application as well as a defensible and well-considered program.

References

[1] Pappus of Alexandria, "Synagoge," Book VIII, AD 340.

[2] OECD Glossary of Statistical Terms Organization for Economic Co-operation and Development, "Glossary of Statistical Terms," OECD Statistics Portal, April 28, 2003 [Online]. Available: https://stats.oecd.org/glossary/detail.asp?ID=3662.

[3] A. Kanioura & F. Lucini, "A Radical Solution to Scale AI Technology," in Harvard Business Review, Apr. 13, 2020 [Online]. Available: https://hbr.org/2020/04/a-radical-solution-to-scale-ai-technology.

[4] C. Kozyrkov, "Artificial Intelligence: Do stupid things faster with more energy!" in Towards Data Science, Sept. 27, 2019 [Online]. Available: https://towardsdatascience.com/artificial-intelligence-do-stupid-things-faster-with-more-energy-379aa6bac220.

[5] R.J. Sawyer, "On Asimov's Three Laws of Robotics," 1994 [Online]. Available: https://www.sfwriter.com/rmasilaw.htm.

[6] J. Haber, "It's Time to Get Serious About Teaching Critical Thinking," in Inside Higher Ed, Mar. 2, 2020 [Online]. Available: https://www.insidehighered.com/views/2020/03/02/teaching-students-think-critically-opinion.

[7] T. Wischmeyer & T. Rademacher, eds. "Regulating Artificial Intelligence." Springer, 2020.

[8] B. Mittelstadt, P. Allo, M. Taddeo, S. Wachter, & L. Floridi, "The ethics of algorithms," Big Data Soc 3(2):1–21, 2016.

[9] D. Neyland, "Bearing accountable witness to the ethical algorithmic system," Sci Technol Hum Values 41(1):50–76, 2016.

[10] K. Crawford, "Can an algorithm be agonistic? Ten scenes from life in calculated publics," Sci Technol Hum Values 41(1):77–92, 2016.

[11] S. Robbins, "A misdirected principle with a catch: explicability for AI," Minds and Machines 29.4: 495-514, 2019.

[12] G.E. Marchant, "'Soft Law' Governance of Artificial Intelligence," The Program on Understanding Law, Science, and Evidence (PULSE) at UCLA School of Law, January 25, 2019 [Online]. Available: https://aipulse.org/soft-law-governance-of-artificial-intelligence/.

[13] C.I. Gutierrez & G.E. Marchant, "A Global Perspective of Soft Law Programs for the Governance of Artificial Intelligence," SSRN, May 27, 2021 [Online]. Available: https://lsi.asulaw.org/softlaw/research-data/.

[14] G.C. Shaffer & M.A. Pollack, "Hard vs. soft law: Alternatives, complements, and antagonists in international governance," Minn. L. Rev. 94: 706, 2009.

[15] A.T. Guzman & T. Meyer, "Soft law," Economic Analysis of International Law. Edward Elgar Publishing, 2016.

[16] J. Villasenor, "Soft law as a complement to AI regulation," Brookings Institute Report July 31, 2020 [Online]. Available: https://www.brookings.edu/research/soft-law-as-a-complement-to-ai-regulation/.

[17] S. Larsson & F. Heintz, "Transparency in artificial intelligence," Internet Policy Review 9.2, 2020.

[18] H.K. Hansen, L.T. Christensen, & M. Flyverbom, "Introduction: Logics of transparency in late modernity: Paradoxes, mediation and governance," European Journal of Social Theory 18 (2), 117-131, 2015. [Online]. Available: https://doi.org/10.1177/1368431014555254.

[19] J.M. Meyers, "Artificial Intelligence and Trade Secrets," A.B.A. Landslide Vol. 11 No. 3, 2019. [Online]. Available: https://www.americanbar.org/groups/intellectual_property_law/publications/landslide/2018-19/january-february/artificial-intelligence-trade-secrets-webinar/.

[20] J. Sherer, "When is a Chair not a Chair? Big Data Algorithms, Disparate Impact, and Considerations of Modular Programming," DESI VII Workshop on Using Advanced Data Analysis in eDiscovery and Related Disciplines (2017); 34 C&ILJ 8 (2017).

[21] J. Sherer, T. Hoffman & E. Fedeles, "AI and Clinical Data," Capstone Legal Bangalore Conference (2018), 2 RAIL 4 (2019).

[22] D.P. Byar, "Why data bases should not replace randomized clinical trials," Biometrics (1980): 337-342.

[23] J. Sherer & N. Sterling, "AI Privacy—The Hurt and the Help," ABA 23 TYL 2 (2019).

[24] K. Hosanagar & V. Jair, "We Need Transparency in Algorithms, But Too Much Can Backfire," at Harvard Business Review Jul. 25, 2018 [Online]. Available: https://hbr.org/2018/07/we-need-transparency-in-algorithms-but-too-much-can-backfire.

[25] L. O'Donnell, "Black Hat 2020: Open-Source AI to Spur Wave of 'Synthetic Media' Attacks," in ThreatPost.com Aug. 5, 2020 [Online]. Avail-

able: https://threatpost.com/black-hat-2020-open-source-ai-to-spur-wave-of-synthetic-media-attacks/158066/.

[26] F. Pasquale, "The black box society," Harvard University Press, 2015.

[27] S. Vollmer, B.A. Mateen, G. Bohner, F.J. Király, R. Ghani, P. Jonsson, S. Cumbers, A. Jonas, K.S. McAllister, P. Myles, & D. Granger, "Machine learning and artificial intelligence research for patient benefit: 20 critical questions on transparency, replicability, ethics, and effectiveness," bmj, 368, 2020.

[28] M. Baker, "Ethical AI—The Minefield of Lethal Autonomous Weapons (1) (1)," The Digital Marketing Bureau, Jun. 2, 2018 [Online]. Available: https://www.thedigitalmarketingbureau.com/artificial-intelligence/ethical-ai-minefield-lethal-autonomous-weapons/attachment/ethical-ai_-the-minefield-of-lethal-autonomous-weapons-1-1.

[29] M. Rolnick, "Beware Of The 'Cobra Effect' in Business," Forbes Business Development Council Post, Aug. 26, 2020 [Online]. Available: https://www.forbes.com/sites/forbesbusinessdevelopmentcouncil/2020/08/26/beware-of-the-cobra-effect-in-business/#7932b3f35f6f.

[30] K. de Fine Licht & J. de Fine Licht, "Artificial intelligence, transparency, and public decision-making," AI & Society 1-10, 2020.

[31] R.H. Wortham, A. Theodorou, & J.J. Bryson, "What Does the Robot Think? Transparency as a Fundamental Design Requirement for Intelligent Systems," in Proceedings of the IJCAI Workshop on Ethics for Artificial Intelligence: International Joint Conference on Artificial Intelligence. IJCAI 2016 Ethics for AI Workshop, New York, USA United States, Sept. 7, 2016.

[32] C.F. Kerry, J.B. Morris, Jr., C.T. Chin, & N.E. Turner Lee, "Bridging the Gaps, A Path Forward to Federal Privacy Legislation," Governance Studies at the Brookings Institute, June 2020 [Online]. Available: https://www.brookings.edu/wp-content/uploads/2020/06/Bridging-the-gaps_a-path-forward-to-federal-privacy-legislation.pdf.

[33] A. Smith, "Using Artificial Intelligence and Algorithms," Federal Trade Commission Business Blog, April 8, 2020. [Online]. Available: https://www.ftc.gov/news-events/blogs/business-blog/2020/04/using-artificial-intelligence-algorithms.

[34] C. Ernst, "Artificial Intelligence and Autonomy: Self-Determination in the Age of Automated Systems," Regulating Artificial Intelligence. Springer, Cham, 53-73, 2020.

[35] P. Roelofs, "Transparency and mistrust: Who or what should be made transparent?" Governance 32.3: 565-580, 2019.

[36] L. Chen & X. Lu, "Making deep learning models transparent," Journal of Medical Artificial Intelligence 1, 2018.

[37] T. Harford, "Big data: are we making a big mistake?" Financial Times, Mar. 28, 2014. [Online]. Available: https://www.ft.com/content/21a6e7d8-b479-11e3-a09a-00144feabdc0.

[38] M. Boden, J. Bryson, D. Caldwell, K. Dautenhahn, L. Edwards, S. Kember, P. Newman, V. Parry, G. Pegman, T. Rodden, T. Sorell, M. Wallis,

B. Whitby, & A. Winfield, "Principles of robotics," The United Kingdom's Engineering and Physical Sciences Research Council (EPSRC), April 2011 [Online]. Available: https://epsrc.ukri.org/research/ourportfolio/themes/engineering/activities/principlesofrobotics/.

[39] J. Bartolo, "Explore New Frontiers in eDiscovery with Data Analytics and Machine Learning," Knovos Whitepaper August 2020.

[40] European Commission, "Proposal for a regulation of the European Parliament and the Council laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union legislative acts," April 21, 2021 [Online]. Available: https://op.europa.eu/en/publication-detail/-/publication/e0649735-a372-11eb-9585-01aa75ed71a1/language-en/format-PDF/source-205836026.

[41] E. Jillson, "Aiming for truth, fairness, and equity in your company's use of AI," Federal Trade Commission Business Blog, April 19, 2021 [Online]. Available: https://www.ftc.gov/news-events/blogs/business-blog/2021/04/aiming-truth-fairness-equity-your-companys-use-ai.

[42] Federal Trade Commission ("FTC"), "Big Data—A Tool for Inclusion or Exclusion? Understanding the Issues," FTC Report, Jan. 2016 [Online]. Available: https://www.ftc.gov/system/files/documents/reports/big-data-tool-inclusion-or-exclusion-understanding-issues/160106big-data-rpt.pdf.

[43] H. Felzmann, E.F. Villaronga, C. Lutz, & A. Tamò-Larrieux, "Transparency you can trust: Transparency requirements for artificial intelligence between legal norms and contextual concerns," Big Data & Society 6, no. 1, 2019.

[44] D.M. Trubek, M.P. Cottrell, & Mark Nance, "'Soft Law,' 'Hard Law,' and European Integration: Toward a Theory of Hybridity," U of Wisconsin Legal Studies Research Paper 1002, 2005 [Online]. Available: https://media.law.wisc.edu/s/c_8/zy5nj/hybriditypaperapril2005.pdf.

[45] S. Gießler & L. Haas, "AI Ethics Guidelines Global Inventory," AlgorithmWatch April 2020 [Online]. Available: https://inventory.algorithmwatch.org/.

[46] The European Union Agency for Fundamental Rights, "AI policy initiatives (2016-2020)," December 11, 2020 [Online]. Available: https://fra.europa.eu/en/project/2018/artificial-intelligence-big-data-and-fundamental-rights/ai-policy-initiatives.

[47] Z.C. Lipton, "The mythos of model interpretability," Queue 16.3: 31-57, 2018.

[48] V.R. Walker, "Complexity, transparency, and the warranted use of formal systems in legal factfinding," Artificial Intelligence and Law 9.2-3: 189-197, 2001.

[49] P.J. Phillips, A.C. Hahn, P.C. Fontana, D.A. Broniatowski, & M.A. Przybocki, "Four Principles of Explainable Artificial Intelligence," (Draft) NIST 8312, Aug. 2020.

[50] D. Doran, S. Schulz & T.R. Besold, "What Does Explainable AI Really Mean? A New Conceptualization of Perspectives," arXiv preprint arXiv:1710.00794. Oct. 2, 2017 [Online]. Available: https://arxiv.org/pdf/1710.00794.pdf.

[51] H. Hagras, "Toward human-understandable, explainable AI," Computer. Oct. 4, 2018; 51(9):28-36.

[52] A. Ignatiev, "Towards Trustable Explainable AI," Proceedings of the 29th International Joint Conference on Artificial Intelligence, IJCAI 2020 (pp. 5154-5158) [Online]. Available: https://www.ijcai.org/Proceedings/2020/726.

[53] B. Cheatham, K. Javanmardian, & H. Samandari, "Confronting the risks of artificial intelligence," McKinsey Quarterly, Apr. 26, 2019. [Online]. Available: https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/confronting-the-risks-of-artificial-intelligence.

[54] A.V. Demin, "Soft Law Concept in a Globalized World: Issues and Prospects," Право. Журнал Высшей школы экономики 4, 2018.

[55] L.A.J. Senden, "Soft law, self-regulation and co-regulation in European law: where do they meet?" EJCL 9.1 at 23, SSRN 943063, 2005 [Online]. Available: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=943063.

[56] J. Zeitlin, "Conclusion: The Open Method of Coordination in Action: Theoretical Promise, Empirical Realities, Reform Strategy," J. Zeitlin and P. Pochet (eds.), The Open Method of Coordination in Action: The European Employment and Social Inclusion Strategies (P.I.E.-Peter Lang), 2005.

[57] B. Canessa, J. Escher, A. Koeberle-Schmid, P. Preller, C. Weber, transl. S. Sidore, "The Family Office," Palgrave Macmillan 2018.

[58] J. Sherer, N. Sterling & S. Burke, "The Not-So-Hidden FTC Guidance on Organizational Use of Artificial Intelligence (AI), from Data Gathering Through Model Audits," BakerHostetler Data Counsel, May 24, 2021 [Online]. Available: https://www.bakerdatacounsel.com/ai/the-not-so-hidden-ftc-guidance-on-organizational-use-of-artificial-intelligence-ai-from-data-gathering-through-model-audits/.

[59] Independent High-Level Expert Group on Artificial Intelligence of the EC ("AI HLEG"), "The Assessment List For Trustworthy Artificial Intelligence (ALTAI) *for self assessment* [sic]," European Commission, July 17, 2020 [Online]. Available: https://digital-strategy.ec.europa.eu/en/library/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment.

Note

* James A. Sherer is a partner in the New York office of BakerHostetler, where he is a co-leader of the Emerging Technology team for the Digital Assets and Data Management Group and directs the firm's AI and Information Governance engagements. His work focuses on litigation; discovery management processes; enterprise risk management; records and information governance;

data privacy, security and bank secrecy; artificial intelligence and algorithmic transparency; technology integration issues; and related merger and acquisition diligence. He may be reached at jsherer@bakerlaw.com.

Originally submitted on September 15, 2020 as part of the 2020 Arizona State University Workshop on the Governance of AI through Soft Law. The views expressed herein are solely those of the author; they should not be attributed to his places of employment, colleagues, or clients; and they do not constitute solicitation or the provision of legal advice.