Dicitur Ex Machina[1]: Artificial Intelligence and the Hearsay Rule

*Jess Hutto-Schultz**

Introduction

Only a person *may be a declarant and make a statement.*[2]

The plot of *Ex Machina*[3] is a simple one: the genius billionaire CEO of a Google-esque company invites a young programmer to his secluded futuristic home to serve as the human component in a Turing-like Test of Ava, an artificial intelligence ("AI") humanoid robot. In a sign of the experiment's success, Caleb, the young programmer, develops feelings for Ava. Upon learning of CEO Nathan's nefarious hidden agenda, he devises a plan to help Ava escape. As Caleb's plan springs into motion and Ava is released from her confinement, it becomes increasingly clear that Ava has a mind of her own. With the help of another robot, Ava kills Nathan, leaves Caleb locked inside Nathan's compound to die, and escapes to the real world on a helicopter flown by an unsuspecting pilot.

Now, imagine a sequel to the film in which Ava mentions to the helicopter pilot as they fly away that someone was killed in the secluded compound. The pilot eventually notifies the authorities, who find and prosecute Caleb for Nathan's murder. There is no video evidence and no other witness to the killing. It has been confirmed that Ava was an AI Entity[4] created by Nathan and that she escaped, but she is nowhere to be found.

At the trial, the defense calls the helicopter pilot to the witness stand. The pilot testifies that, during the flight, Ava revealed that Nathan was a

---

1 "Said the machine" in Latin.

* J.D., George Mason University, Antonin Scalia Law School. I'd like to thank Ashley Hutto-Schultz for her unwavering support and guidance and my parents for their support and encouragement. Special thanks to the entire George Mason Law Review for their work on this piece, to MJ Egan for her editorial suggestions, and to Kedar Bhatia for his guidance and mentorship throughout this process.

2 United States v. Washington, 498 F.3d 225, 231 (4th Cir. 2007).

3 EX MACHINA (DNA Films 2014).

4 This Comment will refer to machine-learning AIs as AI Entities.

raging alcoholic, that Nathan was actually killed by a female, and that Ava had, in fact, played a part in Nathan's murder. This testimony shows that Caleb could not be the killer, and, at least, establishes a basis for a self-defense claim. There is only one problem: the statements were made out of court and are being offered to prove the truth of the matter asserted. Classic hearsay. *Or is it?*

"Objection, Your Honor, hearsay!" the prosecuting attorney says. The defense responds, "Your Honor, the hearsay rule is not implicated here. The AI agent that the witness is referring to does not qualify as a 'declarant' under the hearsay rule. Only a *person* can be a declarant, and it has been established that Ava is an AI Entity, not a person." The prosecution goes on to argue that, though Ava is not *technically* a person, she clearly exhibits qualities of a person, and there is evidence that she has passed the Turing Test. Regardless, the prosecution continues, Ava is not here, cannot be cross-examined, and her programming cannot be understood, even by AI experts. Thus, her statements are not subject to the traditional safeguards of reliability. Should Ava's statements be admitted, or should they be excluded as hearsay?

This seems like a settled area of the law: "Only a *person* may be a declarant and make a statement."[5] But, the law of evidence has never before encountered thinking, reasoning, and speaking AI Entities.[6] While it is generally accepted that traditional computer- and machine-generated statements—including those made by the more basic AI programs of the past—present some evidentiary concerns, those concerns are normally addressed using the rules of authentication, not hearsay. Authenticating a statement made by a traditional computer or machine, for example, requires little more than a showing that the computer was functioning properly. This authentication standard is not difficult to meet, meaning the vast majority of these statements are admitted into evidence.

Modern AI is far more advanced than traditional computers, machines, or even AI of the past, however, because it uses machine-learning algorithms that are capable of making intuitive decisions based on newly learned information, much like humans. The new wave of AI developments will generate powerful AI that is capable of improving itself, perhaps even by rewriting its own algorithms, advancing its intelligence beyond that of humans.[7] These advancements in AI technology render the rules of authentication an inappropriate evidentiary safeguard for modern

[5] *Washington*, 498 F.3d at 231.

[6] Though much of the following analysis will apply to AI Entities without human-like physical features, this Comment will primarily focus on the intersection of robotics and artificial intelligence, where future AI Entities will walk, talk, think, learn, understand, reason, and even feel.

[7] *See* Yavar Bathaee, *The Artificial Intelligence Black Box and the Failure of Intent and Causation*, 31 HARV. J.L. & TECH. 889, 899 n.37 (2018).

AI evidence. The lack of understanding as to how these AI Entities actually make decisions would very likely result in either the admission of unreliable and non-credible statements, or the withholding of reliable or credible statements. Rather, the testimonial risks that are inherent in statements made by modern AI Entities are more akin to those found in human assertions that render them hearsay. Thus, the hearsay rule is a more appropriate safeguard to the admissibility of statements made by AI Entities.

While some experts doubt that AI Entities will ever advance to Ava's level, most agree that we will see the day that humanoid robots work among us, interacting with humans every day. This will undoubtedly result in situations in which, for example, a witness claims that Susie AI Factory Robot told him that the "[clamp] of the machine was not in position" at the time of the injury,[8] or that Joe AI Maintenance Robot said that "[the wolf] bit a child."[9] These are clearly out-of-court statements, presumably being offered to prove the truth of the matter asserted, but, as they were technically made by "machines" and not "persons," under a strict application of current case law, they would not implicate the hearsay rule.[10] But, the hypothetical AI Entities clearly saw, remembered, reasoned, and spoke the event to another person, something we previously thought only humans could do. As this Comment will explain, machines and AI Entities are not impervious to credibility issues, so why should the hearsay rule, created to address the risks surrounding credibility of such statements, not apply in these circumstances?

This Comment seeks to answer such questions. Part I of this Comment begins by providing an overview of AI technology and its projected benefits and risks. Part II provides a primer to the hearsay rule with an emphasis on the purpose of the rule and its application to modern machines. Part III introduces the concept of legal personhood for AI Entities and proposes a sliding scale test for personhood under the hearsay rule. Part IV addresses whether an AI Entity could be considered a declarant under the hearsay rule and proposes a hybrid of two interrelated approaches to answer this question: (1) a purpose-based approach in which the application of the hearsay rule is centered on the primary target of credibility; and (2) a "Turing"-based approach in which advanced AI Entities that meet a certain set of criteria are considered "persons" under the hearsay rule. This hybrid approach takes into account the Black Box[11]

[8] Adapted from *Reed v. McCord*, 54 N.E. 737, 740 (N.Y. 1899).

[9] Adapted from *Mahlandt v. Wild Canid Survival & Research Ctr., Inc.*, 588 F.2d 626, 629 (8th Cir. 1978).

[10] *See infra* Part II.B.

[11] *See* Bathaee, *supra* note 7, at 891, 924; Andrea Roth, *Machine Testimony*, 126 YALE L.J. 1972, 1977–78 (2017).

risks in AI as well as the extent of an AI Entity's "person-like" capabilities. Under this approach, as AI Entities become more advanced, future—and, arguably, some current—AI Entities are considered "persons"[12] who can make "statements" and who would satisfy the definition of "declarants" for the purposes of the hearsay rule.

Importantly, this Comment does not advocate that *all* computer-generated evidence should be treated as hearsay,[13] because such an approach would be contrary to the purpose of the hearsay rule. Rather, this Comment focuses on statements made by AI Entities and argues that the hearsay rule is the most appropriate evidentiary safeguard for their credibility.

I. Background: Artificial Intelligence, Machine Learning, and the "Black Box"

That's the thing about AI. It's sort of like magic. We only call things AI that we don't understand yet. Once we understand something, it's just math.[14]

Whether we think an AI Entity can be considered a declarant under the hearsay rule will depend greatly on our knowledge of AI Entities and how they operate. Recent legal, academic, and scientific scholarship has suggested that AI technology will continue to develop and eventually result in person-like AI Entities. But, some still suggest that even these highly advanced AI Entities will remain mere machines, "unlikely ever to match the imagination of humans."[15] While it is true that "machine conveyances"[16] today are not hearsay, there are transparency problems and testimonial infirmities inherent in current and developing machine-learning AI programs that are akin to the risks inherent in human assertions that render them hearsay.[17] This Part provides a

[12] This Comment does not go so far as to suggest a new legal status for AI Entities such as "E-Person" or that they be granted "legal personalities" under the law, but should such a status be developed under the law, then it would easily fall under the definition of "person" for purposes of the hearsay rule.

[13] For an argument that does advocate for this approach, see Steven W. Teppler, *Testable Reliability: A Modernized Approach to ESI Admissibility*, 12 AVE MARIA L. REV. 213, 216 (2014) (arguing that "the admissibility of [all] digital data should be pre-conditioned on some affirmative showing of reliability required by the residual hearsay rule").

[14] Mark Zuckerberg, Comment to *Building Jarvis*, FACEBOOK (Dec. 19, 2016), https://perma.cc/XGR9-ZBR4.

[15] James Moor, *The Dartmouth College Artificial Intelligence Conference: The Next Fifty Years*, 27 AI MAG., Winter 2006, at 87, 90.

[16] Roth, *supra* note 11, at 1976 n.11.

[17] *See id.*

foundational understanding of AI and how it works, gives an overview of current and future AI technology, and demonstrates the inherent problems with AI programs, specifically the "Black Box" problem.[18]

A. *What is AI?*[19]

The term "artificial intelligence" was first coined in 1955 by Professor John McCarthy in a Dartmouth College research project proposal[20]:

> We propose . . . a 2 month, 10 man study of artificial intelligence . . . . The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves.[21]

The project did not achieve the results aimed for, but it serves as the birth of the AI field of research and development and made a lasting effect on AI.[22] AI has come a long way since 1955, but the concept and end goals have largely remained the same. However, there is still no single, straightforward definition of AI that has gained a consensus among the scientific community. It is generally understood as "the science of making machines do things that would require intelligence if done by [humans]."[23] Thus, AI programs are designed to solve problems that require fundamental capabilities associated with human intelligence such as "inferential reasoning, decision-making based on incomplete or uncertain information, classification, optimization, and perception."[24]

AI technology embodies a wide scope of programs with varying degrees of autonomy, problem-solving ability, and intelligence.[25] On one end of the spectrum, AI programs make decisions based on a set of

---

18 Bathaee, *supra* note 7, at 896.

19 It is equally important to understand what AI is *not*. A common misconception is that a robot of any kind is an AI Entity. This is not always the case. AI "is a reference to the software that manifests intelligence, whereas robots infer a physical element, a shell which carries out the decisions made by the AI engine behind it. Not every AI needs a robot to carry out its functions, and retrospectively, not every robot needs true AI to power its functionality." Ben Dickson, *What is Narrow, General and Super Artificial Intelligence*, TECHTALKS (May 12, 2017), https://perma.cc/XSY5-PND9. This Comment will focus on the intersection between the two.

20 JOHN MCCARTHY ET AL., A PROPOSAL FOR THE DARTMOUTH SUMMER RESEARCH PROJECT ON ARTIFICIAL INTELLIGENCE (1955), *as reprinted in* 27 AI MAG., Winter 2006, at 12, 12–14.

21 *Id.* at 12.

22 *See* Moor, *supra* note 15, at 87–90.

23 J. David Bolter, *Artificial Intelligence*, 113 DAEDALUS, Summer 1984, at 1, 9.

24 Bathaee, *supra* note 7, at 898.

25 *See id.*

preprogrammed rules and algorithms that evaluate possible outcomes.[26] On the other end of the spectrum, modern AI programs are based on machine learning in which the AI Entity *learns* from data and *changes the way it makes decisions* as it goes.[27] This form of AI does not have any "pre-programmed rules about how to solve the problem at hand, but rather only rules about how to learn from data."[28]

Machine learning is a subfield[29] of AI that "involves computer algorithms that have the ability to 'learn' or improve in performance over time on some task."[30] Many applications of modern AI employ machine learning, including self-driving cars and facial recognition software.[31] Rather than working from preprogrammed rules and algorithms[32] instructing the AI Entity how to solve a particular problem, machine-learning AI works from rules that instruct the AI Entity how *to learn* so that it can solve any problem, regardless of context.[33]

Machine learning attempts to combine deductive, inductive, and abductive reasoning[34] to make person-like reasoning in AI Entities possible.[35] Machine-learning AI Entities initially learn in the same way a human child learns. First, they notice details and patterns in shapes,

[26] For example, in 1997, an AI program named "Deep Blue" beat the world champion in chess using preprogrammed rules and "search[ing] between 100 million and 200 million positions per second" to choose the most optimal move. Larry Greenemeier, *20 Years after Deep Blue: How AI Has Advanced Since Conquering Chess*, SCI. AM. (June 2, 2017), https://perma.cc/J3JL-AQDL.

[27] For example, a *new* AI chess champion, AlphaZero, emerged in 2017, using an approach to AI known as "deep reinforcement learning." Aaron Krumins, *AlphaZero Is the New Chess Champion, and Harbinger of a Brave New World in AI*, EXTREMETECH (Dec. 12, 2017, 7:30 AM), https://perma.cc/4R5S-BGER.

[28] Bathaee, *supra* note 7, at 898.

[29] There is a further subset of machine learning called deep learning.

[30] Harry Surden, *Machine Learning and Law*, 89 WASH. L. REV. 87, 88 (2014).

[31] *See* Ayn de Jesus, *Machine Vision for Self-Driving Cars – Current Applications*, EMERJ (Jan. 16, 2019), https://perma.cc/X88N-YPXU.

[32] Another common misconception in the AI field is that the use of any algorithm in a program makes it an AI program. Again, this is not always the case. "Using an algorithm to predict an outcome of an event is not machine learning. Using the outcome of your prediction to improve future predictions is." Venkatesan M, *Artificial Intelligence vs. Machine Learning vs. Deep Learning*, DATA SCI. CENT. (May 7, 2018), https://perma.cc/N8G7-V2NH.

[33] For a brief discussion on how machine-learning accomplishes this task, see *infra* Part I.B.

[34] Atocha Aliseda, *Mathematical Reasoning vs. Abductive Reasoning: A Structural Approach*, 134 SYNTHESE 25, 25 (2003) ("While deductive reasoning is for making predictions, inductive reasoning is for verifying those predictions; and abductive reasoning is for constructing hypotheses for puzzling phenomena.").

[35] For a more in-depth and technical look at how machine learning works, see IAN GOODFELLOW, YOSHUA BENGIO & AARON COURVILLE, DEEP LEARNING (2016). *See also* ETHEM ALPAYDIN, INTRODUCTION TO MACHINE LEARNING (2004). For a closer look at how these machine-learning algorithms create a Black Box problem, see Bathaee, *supra* note 7, at 899–905.

sounds, and behaviors; then, they put them together to distinguish one person from another, or a dog from a cat.[36] Eventually, the AI Entity can recognize and identify "discrete objects even though no humans had ever defined or labeled them."[37] In this way, a "new [AI Entity that employs machine learning] is not unlike the brain of a human child—ready to be molded and shaped by its experiences."[38] It is this subfield of AI, along with deep neural networks, that will produce the AI Entities that this Comment examines.

1. Current AI Technology

Most AI progress has been made in Artificial Narrow Intelligence ("ANI").[39] ANI is an application of AI in a specialized or narrow domain. Essentially, "narrow AI works within a very limited context, and can't take on tasks beyond its field."[40] This is because ANI "lack[s] the self-awareness, consciousness, and genuine intelligence to match human intelligence. In other words, [it] can't think for [itself]."[41] ANI can process data exponentially faster than humans but lacks the human ability to think abstractly, do a double-take,[42] and solve problems creatively by relying on experience, not on data. Contrary to what it seems, Apple's voice-controlled personal assistant, Siri, "isn't a conscious machine responding to our [questions]. Instead, what Siri is able to do—what it is *designed* to do—is process the human language, enter it into a search engine (Google), and return to us with results."[43] All current applications of AI are classified as ANI,[44] most commonly employing speech, image, and facial recognition software. Examples of ANI include, among many others, Hanson

36 *See* Robert D. Hof, *Deep Learning*, MIT TECH. REV. (Apr. 23, 2013), https://perma.cc/KPZ4-YMGQ.

37 *Id.*

38 Weston Kowert, Note, *The Foreseeability of Human–Artificial Intelligence Interactions*, 96 TEX. L. REV. 181, 183 (2017).

39 NAT'L SCI. & TECH. COUNCIL COMM. ON TECH., EXEC. OFFICE OF THE PRESIDENT, PREPARING FOR THE FUTURE OF ARTIFICIAL INTELLIGENCE 7 (2016) [hereinafter NSTC AI REPORT], https://perma.cc/XG3G-2CKN.

40 Dickson, *supra* note 19.

41 Tannya D. Jajal, *Distinguishing Between Narrow AI, General AI and Super AI*, MEDIUM (May 21, 2018), https://perma.cc/NPR6-2U28.

42 Kevin Hartnett, *Machine Learning Confronts the Elephant in the Room*, QUANTA MAG. (Sept. 20, 2018), https://perma.cc/LX3C-6XPD.

43 Jajal, *supra* note 41.

44 NSTC AI REPORT, *supra* note 39, at 7.

Robotics's humanoid Sophia,[45] IBM's Watson,[46] IBM's chess champion Deep Blue,[47] Google Translate,[48] Apple's Siri,[49] Amazon's Alexa,[50] and Tesla's self-driving cars.[51]

2. Future AI Technology

Artificial General Intelligence ("AGI") refers to future AI technology that "exhibits apparently intelligent behavior at least as advanced as a person across the full range of cognitive tasks."[52] AGI is the AI we see in movies like *Ex Machina*[53] or *Her*,[54] in which the AI Entities are sentient beings that are self-aware, conscious, and can think, learn, reason, analyze, and create. AGI is expected to exhibit these person-like qualities and to be able to perform any intellectual task that a human can, including thinking abstractly and learning from experience.

Experts expect AGI to arrive anywhere from "2030 to centuries from now."[55] Professor John McCarthy, for instance, posited that "human-level AI is likely but not assured by 2056."[56] Ray Kurzweil, a futurist and Google's Director of Engineering, predicts that we will reach AGI by the year 2029 and "singularity" by 2045.[57] Singularity refers to a hypothetical future in which technological progress becomes uncontrollable and irreversible.[58] Kurzweil also predicts that we will be able to "download" a human's brain to AI to create a clone, making humans "immortal."[59] Kurzweil has been

45 *Sophia*, HANSON ROBOTICS, https://perma.cc/D57T-H8KW.

46 *Watson*, IBM, https://perma.cc/T87U-8Y6Q.

47 *Deep Blue*, IBM, https://perma.cc/FD3Y-SCFD.

48 *Google Translate*, GOOGLE, https://perma.cc/3UXY-VFBA.

49 *Siri*, APPLE, https://perma.cc/Z2HG-UHAC.

50 *Amazon Alexa*, AMAZON DEVELOPER, https://perma.cc/3HDG-ZJQG.

51 *Autopilot*, TESLA, https://perma.cc/T2XT-S2DY.

52 NSTC AI REPORT, *supra* note 39, at 7.

53 EX MACHINA, *supra* note 3.

54 HER (Annapurna Pictures 2013).

55 NSTC AI REPORT, *supra* note 39, at 7 n.14.

56 Moor, *supra* note 15, at 90.

57 Christianna Reedy, *Kurzweil Claims That the Singularity Will Happen by 2045*, FUTURISM (Oct. 5, 2017), https://perma.cc/8D2T-YM9U.

58 *See* Roey Tzezana, *Singularity: Explain It to Me Like I'm 5-Years-Old*, FUTURISM (Mar. 3, 2017), https://perma.cc/8WWE-ABGE.

59 *See* RAY KURZWEIL, THE SINGULARITY IS NEAR: WHEN HUMANS TRANSCEND BIOLOGY 260 (2005); *see also* Cadie Thompson, *Live Forever? Maybe, By Uploading Your Brain*, CNBC (May 4, 2015, 6:53 PM), https://perma.cc/55DT-VNN7.

criticized for his optimism,[60] and while his predictions on AGI, singularity, and brain cloning have yet to be proven, many of his other predictions have remained accurate.[61]

Finally, Artificial Super Intelligence ("ASI") is what the hype around AI is all about. Experts and lay people alike imagine "a dystopian [future in which] these *super-intelligent* machines . . . exceed the ability of humanity to understand or control."[62] In this imagined future AI apocalypse, AI takes over, wreaking havoc on humanity and leaving us helpless or even extinct.[63] Many, however, hold "[a] more positive view of [ASI, predicting] the development of intelligent systems that work well as . . . teammates of humans."[64] It is likely this optimistic, teammate-based perspective of future AI that will result in the human-AI interactions that produce AI hearsay.

B. *The "Black Box" Problem*

> *It is a problem that is already relevant, and it's going to be much more relevant in the future.*[65]

It should come as no surprise that "the consensus of computer scientists is that the evidence produced by computer programs is no more inherently reliable or truthful than the evidence produced by human witnesses."[66] Thus, it seems obvious that even some modern machine-generated data should at least implicate the hearsay rule:

---

[60] *See* Gary Marcus, *Ray Kurzweil's Dubious New Theory of Mind*, NEW YORKER (Nov. 15, 2012), https://perma.cc/8DVB-ECEP.

[61] *See* Dominic Basulto, *Why Ray Kurzweil's Predictions Are Right 86% of the Time*, BIG THINK (Dec. 13, 2012), https://perma.cc/ZZD7-RGQK. For a list of some of Kurzweil's predictions and how they have fared, see generally RAY KURZWEIL, HOW MY PREDICTIONS ARE FARING (2010), https://perma.cc/F3EE-WXNP. For a more objective look at nine of those predictions, see Drake Baer, *5 Amazing Predictions by Futurist Ray Kurzweil That Came True — and 4 That Haven't*, BUS. INSIDER (Oct. 20, 2015, 9:44 AM), https://perma.cc/7KCC-XW4Q.

[62] NSTC AI REPORT, *supra* note 39, at 8.

[63] *See* Rory Cellan-Jones, *Stephen Hawking Warns Artificial Intelligence Could End Mankind*, BBC NEWS (Dec. 2, 2014), https://perma.cc/5VKP-NMB4.

[64] NSTC AI REPORT, *supra* note 39, at 8. The human element is easily overlooked. If these machines are developed, humans will have to find a way to relate to them. Thus, "[t]he important issue for us may be less about the capabilities of the [AI Entities] than about our own vulnerabilities when confronted with very sophisticated artificial intelligences." Moor, *supra* note 15, at 91.

[65] Will Knight, *The Dark Secret at the Heart of AI*, MIT TECH. REV. (Apr. 11, 2017), https://perma.cc/GDL5-8PS5.

[66] Christian Chessman, Note, *A "Source" of Error: Computer Code, Criminal Defendants, and the Constitution*, 105 CALIF. L. REV. 179, 185 (2017); *see also* Perma Research & Dev. v. Singer Co., 542 F.2d 111, 125 (2d Cir. 1976).

> Just as human sources potentially suffer the so-called 'hearsay dangers' of insincerity, ambiguity, memory loss, and misperception, machine sources potentially suffer 'black box' dangers that could lead a factfinder to draw the wrong inference from information conveyed by a machine source . . . . Just as the 'hearsay dangers' are believed more likely to arise and remain undetected when the human source is not subject to the oath, physical confrontation, and cross-examination, black box dangers are more likely to arise and remain undetected when a machine utterance is the output of an 'inscrutable black box.'[67]

This "inscrutable black box" theory discusses the credibility issues of machines due to possible errors in the output, either by design, programming errors, or simple misanalysis by the machine.[68] These issues exist in nearly every modern machine and, for most basic machines, can be addressed in the adversarial process under the rules of authentication by examining the source code or the input or by having the programmer testify.[69]

But there are far more troubling "Black Box" risks in machine-learning AI Entities than in the average machine. These risks center around the problem that "[n]o one really knows how the most advanced algorithms do what they do."[70] In other words, machine-learning AI internalizes and processes data in ways that are not fully recognizable to or understood by its human programmers.[71] In fact, an AI Entity with Black Box problems "may function in a manner well outside of what the program's creators could foresee."[72] These characteristics of machine-learning AI Entities remove the need for a human programmer to specify each individual decision and will result in AI Entities solving problems that no human could solve or that the programmer did not even know existed. The tradeoff is that the results are generated in ways that even the programmers cannot understand.[73] In fact, machine-learning AI Entities "can be as difficult to understand as the human brain."[74] Thus, Black Box AI Entities will make decisions in the same way humans do, but without any means for communicating how or why they did so.

---

67 Roth, *supra* note 11, at 1977–78 (footnotes omitted).

68 *See id.*

69 *See id.* at 1975.

70 Knight, *supra* note 65.

71 *Id.*

72 Bathaee, *supra* note 7, at 907.

73 AI Entities at Facebook AI Research ("FAIR") have even dynamically created their own language to speak to each other, and their programmers have no idea what they are saying. *See* Mark Wilson, *AI Is Inventing Languages Humans Can't Understand. Should We Stop It?*, FAST COMPANY (July 14, 2017), https://perma.cc/CG3Z-HBSP. However, while allowing these AI Entities to create their own language results in an even more opaque Black Box, the value in letting them continue could be substantial. *See id.*

74 Bathaee, *supra* note 7, at 891.

The Black Box problem presents unavoidable limitations in safeguarding the reliability of statements made by these future AI Entities. First, the adversarial process will be useless if the program's creator cannot explain or verify how or why an AI Entity made a particular decision. Second, though the exclusion of evidence under the rule of authentication "is relatively rare, because authentication is an easily met threshold requirement,"[75] traditional standards of authenticating machine generated evidence may prove to be unattainable with AI Entities.[76] Thus, statements made by future AI Entities will lack an evidentiary standard to demonstrate their reliability, so reliable statements may remain inadmissible.

Due to the shortcomings of authentication procedures and the inability to effectively cross-examine a declarant, these evidentiary safeguards do little to promote the goal of adjudicatory justice when applied to AI Entity statements. This Comment suggests that the safeguards present in the hearsay rule are more appropriate for AI Entity statements.

II. Basics of the Hearsay Rule

The hearsay rule is a rule of exclusion, yet is riddled with exceptions.[77]

Hearsay is an out-of-court statement "offer[ed] in evidence to prove the truth of the matter asserted."[78] Hearsay is inadmissible unless otherwise provided by an exception to the hearsay rule or by statute.[79] The definition of hearsay is deceptively simple, but the application is quite complex.[80] This Part examines the basics of the hearsay rule, beginning with a brief discussion of its purpose, and then provides a general overview of the admissibility of computer-generated evidence ("CGE").

A. *Purpose of the Hearsay Rule*

At the core of the hearsay doctrine is the principle that out-of-court statements carry certain risks regarding the credibility and reliability of

---

75 Michael L. Seigel, *Rationalizing Hearsay: A Proposal for a Best Evidence Hearsay Rule*, 72 B.U. L. REV. 893, 904 n.39 (1992); *see* FED. R. EVID. 901(a).

76 *See infra* Part IV.A.2.

77 Richard A. Posner, *On Hearsay*, 84 FORDHAM L. REV. 1465, 1466 (2016).

78 FED. R. EVID. 801(c).

79 FED. R. EVID. 802.

80 *See* Posner, *supra* note 77, at 1467 ("The hearsay rule, with its multitude of exceptions, is too complex.").

the statements made.[81] The reason for their exclusion "centers on the danger of admitting evidence whose reliability has not been tested."[82]

To encourage witnesses to give accurate and truthful testimony and to expose any credibility issues, "tradition has [advanced] three conditions under which witnesses will ideally be required to testify: (1) under oath, (2) in the personal presence of the trier of fact, [and] (3) subject to cross-examination."[83] The rules against hearsay are designed to ensure compliance with these ideal conditions. Declarants[84] under the hearsay rule are not subject to these safeguards, which highlights the risk of unreliability of their statements.[85] So, "hearsay's fundamental evidentiary flaw is the absence of an opportunity to reveal an out-of-court declarant's weaknesses through cross-examination."[86] Therefore, hearsay statements should not be admitted as evidence to prove the truth of what the statements purport to say.[87]

The common law approach to hearsay, which is referred to in Part IV of this Comment as the *purpose-based approach*,[88] underscores this idea. Under this approach, "an out-of-court statement is hearsay when it depends for value on the credibility of the declarant."[89] Because a trier of

81 CHRISTOPHER B. MUELLER ET AL., EVIDENCE § 8.2, at 755–56 (6th ed. 2018).

82 Roger Park, *A Subject Matter Approach to Hearsay Reform*, 86 MICH. L. REV. 51, 55 (1987).

83 FED. R. EVID. art. VIII advisory committee's introductory note.

84 As a brief reminder, a "declarant" should be distinguished from a "witness." Under the hearsay rule, a declarant is the "person" who makes an out-of-court statement. On the other hand, a witness is the "person" who sits on the witness stand during trial. A witness may relay a declarant's out-of-court statement as a part of the witness's in-court testimony. The hearsay rule is implicated when the witness is not also the declarant of a statement that she relays in her testimony. When the witness *is* the declarant of an out-of-court statement, the trier of fact can simply assess her credibility in court while she is on the witness stand. However, when the witness merely relays a statement made by someone else, that other person's credibility cannot be assessed, raising questions as to the statement's reliability. *See* MUELLER ET AL., *supra* note 81, § 8.2, at 755–56.

85 For an argument that the primary purpose for the exclusion of hearsay is "procedural fairness," see Liesa L. Richter, *Goldilocks and the Rule 803 Hearsay Exceptions*, 59 WM. & MARY L. REV. 897, 948 (2018). However, Richter refutes this argument, stating that "the fundamental justification for hearsay doctrine is binary, firmly grounded in *both* procedural justice and reliability." *Id.* at 949.

86 Park, *supra* note 82, at 55–56.

87 There are many who advocate for reform of the hearsay rule and a few who seek to abolish it altogether. *See, e.g.*, Ronald J. Allen, Commentary, *A Response to Professor Friedman: The Evolution of the Hearsay Rule to a Rule of Admission*, 76 MINN. L. REV. 797, 800 (1992); Eleanor Swift, *Abolishing the Hearsay Rule*, 75 CALIF. L. REV. 495, 495, 518 (1987). This Comment, however, does not address those proposals. Rather, it seeks to find a way to address statements made by AI Entities under the current hearsay rules and exceptions.

88 *See infra* Part IV.B. This approach has also been referred to as the "declarant-centered approach." Roger C. Park, *"I Didn't Tell Them Anything About You": Implied Assertions as Hearsay Under the Federal Rules of Evidence*, 74 MINN. L. REV. 783, 783 (1990).

89 Park, *supra* note 88, at 783.

fact does not have a suitable method for assessing the credibility of an out-of-court declarant (i.e., through cross-examination), statements made by such declarants, when offered for the truth of their contents, are excluded from evidence.[90] Take, for example, an out-of-court statement such as "I'm dying." If the statement were offered to prove that the declarant was, in fact, dying at the time it was made, the trier of fact, while determining the credibility of the statement, would have to trust that the declarant—who is not on the witness stand—was not misinformed, mistaken, or lying. In other words, the statement's value would depend on the *declarant's* credibility, and under a purpose-based approach, it would be inadmissible hearsay. However, if the same statement were offered, instead, to prove only that the declarant was still alive at the time it was made, the credibility of the declarant would not be an issue. If it is true that the declarant made the statement, a fact that depends on the *witness'* credibility—not the declarant's—and that can be tested on cross-examination, then it is the mere fact that the statement was made at all, and not the content of the statement, that matters. Dead people do not make statements, after all.[91]

Over time the approach to hearsay analysis has transformed. In 1972, the Advisory Committee's Notes "enunciated what might be termed an 'intent-based' approach to the hearsay rule . . . focus[ing] on what the out-of-court declarant intended to assert, and then ask[ing] whether that intended assertion has been offered for its truth at trial."[92] Under this approach, "[t]he key to [whether something is a statement] is that nothing is an assertion unless intended to be one."[93] However, just because the Advisory Committee has elaborated on the intent behind a certain rule of evidence, it does not mean that a court must interpret the rule in the same manner. In fact, "[l]egitimate questions have been raised both about the degree of authority that should be afforded the Advisory Committee's Notes in general and about the validity of the intent-based approach."[94] However, "most leading treatises and recent judicial opinions seem to accept the Advisory Committee's understanding."[95]

---

[90] *See id.* at 785.

[91] The made-up case, *Estate of Murdock*, is commonly used to illustrate this concept. *See* 32 Muc. 352 (1983).

[92] Paul F. Kirgis, *Meaning, Intention, and the Hearsay Rule*, 43 WM. & MARY L. REV. 275, 283–84 (2001); *see* FED. R. EVID. 801 advisory committee's note.

[93] FED. R. EVID. 801 advisory committee's note.

[94] Kirgis, *supra* note 92, at 284 (footnote omitted); *see* Eileen A. Scallen, *Interpreting the Federal Rules of Evidence: The Use and Abuse of the Advisory Committee Notes*, 28 LOY. L.A. L. REV. 1283, 1293–1301 (1995).

[95] Kirgis, *supra* note 92, at 285 & nn.40–41 (footnote omitted) (citing leading treatises and recent judicial opinions).

There are obvious limitations to this approach when applied to machines: can a machine *intend* to assert anything? The answer to this question by courts and legal scholars as it pertains to modern-day machines is "no."[96] But, with machine-learning AI Entities, the question of basic intent to assert a statement[97] arguably becomes more complex. If, for instance, we define intent to mean the determination to do some act,[98] and a Black Box AI program is programmed with an overarching goal to negotiate a deal with a maximum profit, wouldn't every statement made in furtherance of that purpose be *intended* to be asserted by the AI Entity?[99] Possibly, but many would, nevertheless, claim that the ability to intend is a trait reserved for humanity.

Still, though the prevailing approach focuses primarily on intent, courts and scholars have essentially "draw[n] a line between the risk of sincerity and the other testimonial risks."[100] As a result, under the *intent-based approach*, those statements that "implicate[] the sincerity of an out-of-court declarant [are] potentially hearsay; [those statements] that [do] not implicate the sincerity of an out-of-court declarant [are] not hearsay."[101] Thus, because both approaches rest, in some form, on the reliability of the declarant, the intent-based approach, in practice, is not all that far removed from a purpose-based approach.

B. *Hearsay and Machines Today*

The American court system offers many safeguards to ensure the reliability of human witnesses, including the requirement of testifying under oath, the Confrontation Clause,[102] the hearsay rule, authentication,

96 *See, e.g.*, Bathaee, *supra* note 7, at 906 ("Machines and computer programs have no intent."); *see also* United States v. Coscia, 866 F.3d 782, 802–03 (7th Cir. 2017) (finding that—in the only criminal conviction for spoofing using a high-frequency trading program to date—the defendant intended to manipulate the market based on a programmer's testimony about what the defendant designed the program to do).

97 This Comment does not address the questions of intent and causation that are fundamental to civil and criminal liability, but merely suggests that the test for basic intent to assert something under the hearsay rule can be easily met. For more on AI intent and causation, see Bathaee, *supra* note 7.

98 *Intent*, BLACK'S LAW DICTIONARY 930 (11th ed. 2019).

99 Though it has been suggested that while "[t]he robotic machines that exist today have goals . . . we have no reason to believe they 'care' whether the goals are satisfied." F. Patrick Hubbard, *"Do Androids Dream?": Personhood and Intelligent Artifacts*, 83 TEMP. L. REV. 405, 421 (2011).

100 *See* Kirgis, *supra* note 92, at 285.

101 *See id.*

102 Courts "have held that machine-generated data does not trigger the Confrontation Clause because it is the machines—not the analysts operating them—that make the statements at issue, and machines are not 'witnesses' within the meaning of the Confrontation Clause." Brian Sites, *Rise of the*

and the rules of impeachment.[103] However, these courtroom safeguards were designed for *human* witnesses and, at times, "seem an awkward fit for machines[.]"[104] Nevertheless, as technology has advanced, courts have attempted to adapt their analyses to create a framework for the admissibility of computer-generated evidence ("CGE").

As part of this developing framework, many courts have distinguished between CGE—information that is generated by the machine's internal operations—and computer-*stored* evidence ("CSE").[105] CSE is essentially a "collection[] of statements by humans"[106] that is stored or processed by a computer.[107] Examples of CSE include "word processor files; spreadsheets, such as Microsoft Excel files; charts; graphs; and emails."[108] CSE *requires* human interaction and input. Thus, the hearsay rule is universally implicated for CSE[109] because its statements are the byproduct of a human assertion. CGE, on the other hand, is evidence that is created by the machine's internal operation; so, it is thought that CGE is not information generated by a human source, and, thus, hearsay is not implicated.[110]

Machines: Machine-Generated Data and the Confrontation Clause, 16 COLUM. SCI. & TECH. L. REV. 36, 51 (2014). One of the leading pieces of scholarship on this topic is Andrea Roth's *Machine Testimony*, which explores, among other things, whether machine conveyances can be "witnesses" against a defendant under the Sixth Amendment's Confrontation Clause. *See* Roth, *supra* note 11, at 2039–51. Her approach centers on the "right of meaningful impeachment," suggesting that no accusatory machine source should be immune "from the Clause's reach entirely." *Id.* at 2050–51.

While Roth's analysis is persuasive, if AI Entities are considered "persons" for purposes of hearsay, the analysis under the Confrontation Clause becomes more complex. If we assume an AI Entity can *technically* be cross-examined, we must ask whether it can make a testimonial statement or be placed under oath. With the Black Box problem, it's likely that all an AI Entity could do is simply restate the past; as of now, an AI Entity or its programmer cannot explain why the AI Entity did what it did, and "there is no obvious way to design such a system." Knight, *supra* note 65. However, restating the past could be quite valuable. It may be that cross examiners are forced to adapt their questioning, and instead of attempting to poke holes in a story or impeach an AI Entity witness, they might need to attempt to get the AI Entity to state a part of the story that has not been asked yet. These are

questions that deserve a full analysis in a separate article. However, it's very likely that, even if out-of-court statements by AI Entity declarants are considered admissible hearsay, the statements, if testimonial, would still be inadmissible under the Confrontation Clause if the AI Entity cannot testify.

103 *See* Andrea Roth, *Trial by Machine*, 104 GEO. L.J. 1245, 1300 (2016).

104 *Id.*

105 *See* Curtis E.A. Karnow, *The Opinion of Machines*, 19 COLUM. SCI. & TECH. L. REV. 136, 150–56 (2017).

106 *Id.* at 151.

107 *See* Adam Wolfson, Note, *"Electronic Fingerprints": Doing Away with the Conception of Computer-Generated Records as Hearsay*, 104 MICH. L. REV. 151, 159 (2005).

108 *Id.*

109 *See id.* at 158.

110 *See id.* at 155–58.

Other courts have had a tendency to treat *any* machine output, regardless of the original source, as hearsay, reasoning that all machine output is the result of a human declarant's assertion.[111] However, even though these courts fail to recognize the distinction between CGE and CSE, the majority of case outcomes ironically fit the approach that does make that distinction.[112] Because the AI Entities that this Comment is concerned with are machine-learning,[113] this Section only examines the admissibility of CGE by courts that distinguish CGE from CSE.

So far, "every [modern] court to have addressed the issue has ruled that [CGE] or source code is not an 'assertion' for purposes of the hearsay rule and not 'testimonial' for purposes of the Confrontation Clause."[114] These rulings are the product of the idea that, when "information is communicated out of court by something other than a person, that out-of-court conveyance cannot be excluded on hearsay grounds."[115] Since a machine is not a person,[116] then any output that is automatically generated by the machine is not hearsay. In other words, courts have not yet conducted any in-depth analysis on whether a machine *can* be a person, but simply assume it cannot be. Instead, the hearsay rule is only implicated if the trier of fact finds that the source of the statement is a person. Thus, a threshold inquiry to the hearsay question requires distinguishing those machine outputs that are generated automatically and without human interaction—CGE—from those machine outputs that are made as the result of a human's input or assistance. This is not always as clear-cut as it may sound because human interaction often triggers the internal processes that create the resulting CGE.[117]

The prevailing view[118] is that, if the data is *new* data that is automatically created by the internal operations of the machine, then

111 *See id.* at 155–56.

112 *See id.* at 158.

113 *See supra* Part I.

114 Roth, *supra* note 103, at 1301.

115 CHARLES A. WRIGHT & ARTHUR R. MILLER, FEDERAL PRACTICE AND PROCEDURE § 6716 (2018 ed.); *see* FED. R. EVID. 801; *see also* United States v. Lizarraga-Tirado, 789 F.3d 1107, 1110 (9th Cir. 2015) ("Here, the relevant assertion isn't made by a person; it's made by the Google Earth program."); United States v. Washington, 498 F.3d 225, 229–31 (4th Cir. 2007).

116 *See, e.g.*, People v. Dinardo, 801 N.W.2d 73, 79 (Mich. Ct. App. 2010) ("[A] machine is not a person and therefore not a declarant capable of making a statement.").

117 *See* Wolfson, *supra* note 107, at 159.

118 This approach has not been without disagreement. *See, e.g.*, Young v. United States, 63 A.3d 1033, 1046 (D.C. Cir. 2013) ("[I]t is too simplistic to say the DNA profiles . . . were not hearsay because they were 'nothing more than raw data produced by a machine.'" (quoting United States v. Summers, 666 F.3d 192, 202 (4th Cir. 2011))); *Washington*, 498 F.3d at 233 (Michael, J., dissenting) (finding that there is only one circumstance in which "a computer-generated assertion [is] not considered the statement of a person: when the assertion is produced without any human assistance or input"). In

even if some human interaction is implicated, the new data is CGE, and thus, "no hearsay objection applies."[119] For example, in *United States v. Washington*,[120] a sample of the defendant's blood was sent for a toxicology analysis.[121] Three lab technicians conducted the test, and the testing machines printed out twenty pages of results containing raw data that listed the presence and amount of certain drugs in the blood sample.[122] The defendant argued that the report constituted inadmissible hearsay, reasoning that it was a statement made by the lab technicians who operated the machine, not a statement made by the machine itself.[123] The court disagreed, finding that the statements in question were the "assertions that [the defendant's] blood sample contained PCP and alcohol."[124] Because "no *person* viewed a blood sample and concluded that it contained [the drugs,]" the statements were of the machine and not its operators.[125]

More recently, in *United States v. Lizarraga-Tirado*,[126] the Ninth Circuit held that a digital "tack" marking a location in Google Maps was not hearsay because it was CGE and not an assertion made by a person.[127] When border agents arrested the defendant, they recorded GPS coordinates.[128] To show where the arrest occurred, the prosecutor entered those coordinates into Google Maps, creating the "tack."[129] Acknowledging that the "tack" was an assertion akin to an "X" on a treasure map, the court followed the reasoning in *Washington*, finding that the "assertion" that the GPS coordinates correspond to the location of the tack was actually made by the Google Earth program and not the person who typed them into the

Young, the D.C. Circuit stated that "the DNA profiles . . . do not stand on their own but, instead, have meaning because they amount to a communication by the scientists who produced them—the assertion, essentially, that the scientists generated these specific results by properly performing certain tests and procedures on particular, uncorrupted evidence and correctly recording the outcomes." 63 A.3d at 1046.

119 Karnow, *supra* note 105, at 154–55; *see Washington*, 498 F.3d at 230 ("Only the machine, through its diagnostic and technical process, could provide facts about the chemical composition of Washington's blood. Accordingly, the raw data generated by the machines were not the statements of technicians.").

120 498 F.3d 225 (4th Cir. 2007).

121 *Id.* at 228.

122 *Id.*

123 *Id.* at 229.

124 *Id.*

125 *Id.* at 230.

126 789 F.3d 1107 (9th Cir. 2015).

127 *Id.* at 1110.

128 *See id.*

129 *See id.*

program.[130] The court reasoned that, although "a person types in the GPS coordinates, he has no role in figuring out where the tack will be placed."[131] Rather, "[t]he real work is done by the computer program itself."[132] Thus, the court held it is not inadmissible hearsay because it is not a statement by a person.[133] The court recognized that machine statements present evidentiary concerns but accepted that those concerns are generally addressed using the rules of authentication, not hearsay. Other courts that follow this approach with CGE agree.[134]

The rules of authentication require that the evidence simply be shown to be "what the proponent claims it is."[135] The standard for authentication is generally easy to meet.[136] With a machine, the proponent must "show that the machine and its functions are reliable, that it was correctly adjusted or calibrated, and that the data . . . put into the machine was accurate."[137] To meet this standard, courts have required little more than a showing that the input data was what the proponent purports it to have been (e.g., the blood was the defendant's)[138] and that the machine "was operating as it usually does, explained by someone with some experience in using the system."[139] In *Lizarraga-Tirado*, the government did not have to prove authentication because the Court admitted the computer-generated "tack," finding that Google Earth was a "source[] whose accuracy cannot reasonably be questioned."[140]

There can be little doubt that information produced by a machine-learning AI Entity is CGE.[141] However, this Comment argues that, because future AI Entities are Black Boxes and will be inherently person-like, they should not be placed in the same category as basic machines. This is because their Black Box and person-like qualities implicate the purpose of the hearsay rule and make the hurdles to authentication more difficult to

130 *Id.* at 1109.

131 *Id.* at 1110.

132 *Lizarraga-Tirado*, 789 F.3d at 1110.

133 *Id.*

134 *See, e.g.*, United States v. Washington, 498 F.3d 225, 231 (4th Cir. 2007).

135 FED. R. EVID. 901(a).

136 *See* Teppler, *supra* note 13, at 224 ("The hurdles presented by the[] subsections to Rule 901(b) to authentication are . . . low and easily traversed.").

137 *Washington*, 498 F.3d at 231.

138 *Id.*; *see* Young v. United States, 63 A.3d 1033, 1046, n.49 (D.C. Cir. 2013) ("It is axiomatic, for example, that if a human being does not enter correct information, the output from a computer means nothing.").

139 Karnow, *supra* note 105, at 155.

140 United States v. Lizarraga-Tirado, 789 F.3d 1107, 1109 (9th Cir. 2015).

141 *See infra* Part IV.A.1.

traverse. Therefore, their out-of-court statements should not be handled as CGE, but, rather, as hearsay.

III. Legal Personhood for AI Entities?

If I hid Ava from you, so you could just hear her voice, she would pass for human. The real test is to show you that she is a robot. Then see if you still feel she has consciousness.[142]

If only a *person* can be a declarant and make a statement, then it must be asked: what does it mean to be a person? How do we determine personhood? These are the fundamental questions that need to be answered if we are to extend any form of personhood to AI Entities for purposes of the hearsay rule.[143]

It seems logical that, if human beings are accorded personhood because humans "are intelligent, have feelings, are conscious, and so forth," then whether AI Entities are accorded personhood should depend on whether they "share these qualities."[144] But how do we determine this, and what will it mean if an AI Entity does share human qualities?

---

142 EX MACHINA, *supra* note 3.

143 Other countries are already recognizing the inevitability of these questions. The European Parliament published a report proposing legislation that would grant machines a "legal personality" under the law. EUR. PARL. DOC. (COM A8-0005/2017) (proposing the "creat[ion of] a specific legal status for robots in the long run, so that at least the most sophisticated autonomous robots could be established as having the status of electronic persons responsible for making good any damage they may cause, and possibly applying electronic personality to cases where robots make autonomous decisions or otherwise interact with third parties independently"). While many have supported Parliament, others have criticized its suggestions as premature and unnecessary. *See* Tim Collins, *Outrage Over European Plans to Grant ROBOTS Legal Status as Experts Warn It Could Breach Human Rights*, DAILY MAIL (Apr. 13, 2018, 11:11 AM), https://perma.cc/K2W8-U955.

Furthermore, at least one country has granted citizenship to an AI Entity: Saudi Arabia. Chris Weller, *A Robot That Once Said It Would 'Destroy Humans' Just Became the First Robot Citizen*, BUS. INSIDER (Oct. 26, 2017, 10:13 AM), https://perma.cc/UY46-5HMQ. Sophia is an AI Entity created by Hanson Robotics and looks remarkably similar to a human with surprisingly lifelike features. *Sophia*, *supra* note 45. She speaks with a human sounding voice and with human inflection; she can smile, frown, and even look confused or surprised if the conversation requires. She can even interact with humans, holding a conversation remarkably well. This should not be taken as a sign that AGI is already here. It's not. *See* Jaden Urbi & MacKenzie Sigalos, *The Complicated Truth About Sophia the Robot — An Almost Human Robot or a PR Stunt*, CNBC (June 5, 2018, 11:15 AM), https://perma.cc/5U2V-4ED8. Her creator, David Hanson, is "approaching Sophia with the mindset that she is AI 'in its infancy'" and that the formula for safe superintelligence is for developers to think like parents, "rais[ing] AGI like a good child, not like a thing in chains." *Id.* Despite criticism, see *id.*, Sophia has gained much attention in the news.

144 Lawrence B. Solum, *Legal Personhood for Artificial Intelligences*, 70 N.C. L. REV. 1231, 1262 (1992).

A. *Proposed Approaches to Machine Personhood*

A growing number of law review articles have addressed the concept of legal personhood for machines and artificial intelligence.[145] Many experts have long held that the Turing Test is the best way to determine whether a machine possesses the ability to think like a human. The original Turing Test involved three rooms, each connected to the others via a computer screen.[146] In one room sits an AI Entity, in the second a human, and in the third sits a human "judge."[147] The judge's job is to decide which of the two occupants is a machine and which is a human. The test has been modified to include only one contestant, and the judge's job is simply to decide whether the single contestant is human or machine. If a human judge cannot tell the contestant is an AI Entity, then it is considered to be of human intelligence.[148] Ray Kurzweil argues that "the Turing Test is a valid test of the full range of the human intelligence [because y]ou need the full flexibility of human intelligence to pass [it]."[149]

However, over time, the original Turing Test has drawn criticism that highlights the belief that intelligence has multiple dimensions that include more than just language comprehension and processing.[150] Real human intelligence, it is thought, requires social awareness and an understanding of one's environment.[151] Building on this idea, professor and AI legal scholar F. Patrick Hubbard has devised an extended behavioral test[152] to determine the standard capacity for personhood. He posits that the capacity for personhood should be determined by demonstrating that the entity has "(1) the ability to interact with its environment and to engage in complex thought and communication, (2) a sense of being a self with a concern for achieving its . . . purpose in life, and (3) the ability to live in a community based on mutual self-interest."[153] If an AI Entity meets these three elements, it is "entitled to at least a prima facie right to be treated as a person rather than property."[154]

---

145 *See, e.g.*, Susan W. Brenner, *Humans and Humans+: Technological Enhancement and Criminal Responsibility*, 19 B.U. J. SCI. & TECH. L. 215, 221–59 (2013); Hubbard, *supra* note 99, at 472; Solum, *supra* note 144, at 1255–80, 1284–87.

146 *See* A. M. Turing, *Computing Machinery and Intelligence*, 49 MIND 433, 433 (1950).

147 *See id.*

148 *See id.* at 433–35, 440.

149 Nicholas Thompson, *Ray Kurzweil on Turing Tests, Brain Extenders, and AI Ethics*, WIRED (Nov. 13, 2017, 10:02 AM), https://perma.cc/54VN-UB6X.

150 *See* Jia You, *Beyond the Turing Test*, 347 SCI. MAG. 116, 116 (2015).

151 *See id.*

152 Hubbard, *supra* note 99, at 419.

153 *Id.*

154 *Id.*

Hubbard's test might be appropriate for determining whether an entity is granted full legal personhood under the law to trigger certain legal rights. However, is Hubbard's test necessary to determine whether an entity is considered a person for a limited purpose under the law, such as hearsay? For instance, does an AI Entity really need to be fully self-aware to make a statement that qualifies as hearsay? An analysis that considers the purpose of the hearsay rule compels an answer of "no."[155]

The concept of limited personhood for nonhuman entities is not unheard of. Courts already grant a legal *personality* to at least one type of nonhuman entity: corporations.[156] For centuries, corporations have been recognized as a "separate legal personality . . . with existence as a juridical entity, separate from its shareholders."[157] In fact, the evolution of a corporation's status under the law began in a similar fashion to the problem before us today with AI Entities.[158]

Interestingly, corporations are granted limited personhood under the First, Fifth, and Fourteenth Amendments[159] when they, in and of themselves, lack meaningful person-like characteristics. Even ships—inanimate objects—have been granted legal personhood in the past, which has included the ability to sue.[160] Corporations, like ships, require human assistance to function.[161]

---

155 *See supra* Part II.A. Another proposed approach is to create a new "ontological category" for AI Entities, somewhere between living and non-living. *See* PETER H. KAHN, JR. ET AL., THE NEW ONTOLOGICAL CATEGORY HYPOTHESIS IN HUMAN-ROBOT INTERACTION, 159 (2011).

156 For a discussion on personhood for other nonhuman artifacts, see generally Hubbard, *supra* note 99. For a historical look at legal personhood generally and a summary of other proposed approaches to legal personhood for AI Entities, see Brenner, *supra* note 145, at 221–59. Brenner posits that "[u]nderstanding how law has defined 'person' (or 'legal person') in the past is . . . a necessary step toward analyzing how future law might approach the task of assessing the 'personhood' of [future technologically enhanced entities]." *Id.* at 223.

157 Phillip I. Blumberg, *The Corporate Personality in American Law: A Summary Review*, 38 AM. J. COMP. L. 49, 49 (1990); *see* Hubbard, *supra* note 99, at 433–36; *see also* Trustees of Dartmouth College v. Woodward, 17 U.S. (4 Wheat.) 518, 636 (1819) ("A corporation is an artificial being, invisible, intangible, and existing only in contemplation of law.").

158 *See* Blumberg, *supra* note 157, at 49–52 ("[T]he fundamental issue was not one of theoretical concept but the adaptation of the law to achieve an appropriate degree of control over the activities of the corporation in the light of the values of the times.").

159 *See* Citizens United v. FEC, 558 U.S. 310, 336–66 (2010) (holding that a federal campaign finance law banning corporate-funded independent expenditures violated a corporation's First Amendment right to free speech).

160 *See* Brenner, *supra* note 145, at 228–31.

161 A ship can simply float as an empty vessel without a human there to operate it, and a corporation can simply exist as an empty legal entity without a human to operate it. But both a ship and a corporation require the assistance of something external, historically a human, to function as an entity.

But, the need for human assistance was not the driving rationale for granting corporations their limited legal personhood.[162] Rather, "the fundamental issue was not one of theoretical concept but the adaptation of the law to achieve an appropriate degree of control over the activities of the corporation in the light of the values of the times."[163] The recognition that the usefulness of corporations outweighs their unpleasant dominance has resulted in the law adapting to give them a "fictional legal status" of personhood only when they further the goal of "implement[ing] a set of complex legal relationships among human persons."[164] It seems logical, then, that an entity, such as an AI Entity, that *does* exhibit meaningful person-like characteristics *and* has substantial economic usefulness should be considered a person at least for those limited purposes in which a goal of adjudicative justice would be furthered. This Comment posits that hearsay is one such limited purpose.

Just like corporations, an AI Entity does not necessarily need to be afforded any substantive legal "rights" to be considered a "person" under the hearsay rule. Rather, all a court must do is recognize the need for such an application of the rule and perhaps the extent of an AI Entity's person-like capabilities. The next Subsection proposes using a sliding-scale approach to determine whether certain out-of-court statements made by an AI Entity warrant such an application of the hearsay rule.

B. *A Better Approach: Sliding-Scale Test[165] for AI Personhood Under the Hearsay Rule*

Whether an AI Entity can qualify as a declarant under the hearsay rule depends, at least in part, on whether an AI Entity can be a "person" under the hearsay rule.[166] As suggested above, the qualifications for AI personhood may hinge on whether an AI Entity shares certain characteristics with humans.[167] Of course, the number of human qualities any single AI Entity shares will vary. Where will we draw the line? Moreover, certain characteristics might be thought of as more "human" than others. For instance, a future AI Entity that appears to be self-aware might be thought of as more person-like than a self-driving car. These

162 *See* Hubbard, *supra* note 99, at 433–36.

163 Blumberg, *supra* note 157, at 51.

164 *See* Hubbard, *supra* note 99, at 435.

165 This sliding-scale test for AI personhood for the purposes of hearsay and the proposed sliding-scale approach below to statements made by AI Entities, see *infra* Part IV.C, were adapted from Yavar Bathaee's sliding-scale approach to evaluating AI intent and causation. *See* Bathaee, *supra* note 7, at 936–38.

166 *See supra* Part II.B.

167 *See supra* Part III.A.

factors can be merged to create an element of personhood predicated on the degree of human qualities an AI Entity shares.

But the degree of shared human qualities is not traditionally considered by courts when determining whether a statement is made by a person or a machine. In Part II, this Comment explained that courts have taken an approach to CGE that considers the extent and type of human interaction required to produce the output.[168] Thus far, courts have operated under the belief that a machine is not and cannot be a person.[169] Thus, the admissibility analysis of a machine's output has traditionally consisted of determining whether the output is CGE by looking at the degree and type of human interaction required to generate the output;[170] if it is CGE (i.e., the source of the statement is not a human), the court avoids the hearsay analysis.[171]

However, thus far, courts have not encountered machines that share any significant degree of human qualities or that have transparency issues that limit the ability to assess their reliability.[172] Combining the degree of shared human qualities with the degree of human interaction required to operate the AI Entity, we can imagine a sliding-scale test to determine AI personhood under the hearsay rule. Under this method, courts will need to note that human interaction with an AI Entity may not look the same as with other machines. For instance, an AI Entity may require a human voice or action to initiate a function,[173] or a human to oversee its actions to prevent or correct errors, or a human to train the AI Entity. This sliding scale would allow courts to simply combine a slightly adapted version of their current analysis with an analysis of an AI Entity's shared human qualities to determine whether that machine can be considered a person under the hearsay rule.

Under this new approach to machine personhood for hearsay purposes, there would be four main quadrants created by the intersection of the two elements: human qualities and required human interaction. The table below (Table 1) is a visual representation of the sliding-scale test for AI personhood for purposes of hearsay.

168 *See supra* Part II.B.

169 *See, e.g.*, People v. Dinardo, 801 N.W.2d 73, 79 (Mich. Ct. App. 2010) ("[A] machine is not a person and therefore not a declarant capable of making a statement.").

170 *See supra* Part II.B.

171 *See id.*; *see also* Natalie F. Pike, Note, *When Discretion to Record Becomes Assertive: Body Camera Footage as Hearsay*, 20 VAND. J. ENT. & TECH. L. 1259, 1266 (2018).

172 For an explanation of why these differences might change the hearsay analysis, see *infra* Part IV.A.

173 *See* Russell Brandom, *Humanity and AI Will Be Inseparable*, THE VERGE (Nov. 15, 2016), https://perma.cc/5NPS-V4F4.

Table 1: The Sliding-Scale Approach to AI Personhood Under the Hearsay Rule

	More Human Interaction	**Less Human Interaction**
More Human Qualities	Case-by-case determination of hearsay personhood but emphasis on human interaction	Person under the hearsay rule
Fewer Human Qualities	Not a person under the hearsay rule	Case-by-case determination of hearsay personhood but emphasis on human qualities

To start, when a trier of fact deems an AI Entity to have more human qualities—considering the number of qualities and weight of each quality—and to require less human interaction to operate or maintain it—considering the degree and type of interaction required—then the AI Entity can be considered a person under the hearsay rule. It should be noted that this does not necessarily result in an AI Entity qualifying as a *declarant* because, unlike with declarant humans, there may exist a method to assess the "declarant" AI Entity's credibility, depending on degree of Black Box risks.[174]

Second, when a trier of fact deems an AI Entity to have fewer human qualities and to require more human interaction to operate or maintain it, the AI Entity cannot be considered a person under the hearsay rule. However, as discussed *infra*, this does not necessarily prevent an AI Entity from being a hearsay *declarant* but may necessitate a relaxed hearsay standard depending on the degree of Black Box issues.[175]

The remaining two quadrants require a case-by-case determination of AI personhood under the hearsay rule and may demand a more in-depth examination of each element. When the AI Entity has more human qualities and requires more human interaction, the trier of fact should give greater weight to the need for human interaction. This is because humans have a strong "inclination to anthropomorphize . . . objects, [and thus] we may be tempted to be too generous in applying the test" when the AI Entity exhibits a high degree of human qualities.[176]

174 *See infra* Part IV.C.

175 *See id.*

176 Hubbard, *supra* note 99, at 449.

To the contrary, when the AI Entity has fewer human qualities and requires less human interaction, the trier of fact should give greater weight to the degree of human qualities. Because of the desire to "preserve our unique status"[177] as humans, there may be a temptation to place too high of a value on the degree of human interaction.

This sliding-scale test will need fine-tuning prior to its implementation, as it leaves some questions unanswered. For instance, what is a human quality, what amounts to human interaction, and will triers of fact be left to their own devices to answer these questions? Furthermore, as suggested above, triers of fact may be required to assign values to different human qualities and weigh those values against each other. They will likely have to decide whether quantity (i.e., a larger *number* of human qualities) outweighs a single human quality's superiority (i.e., one quality is valued more than the rest combined). This will be no easy task if, in the "future, it [may] be hard to distinguish human agency from automated assistance."[178]

It is tempting to view this test as a one-step hearsay analysis that treats all out-of-court statements by an AI Entity meeting its "person" standard and offered for their truth as hearsay. However, there are differences in the credibility risks between humans and AI Entities,[179] compelling a further analysis that considers the Black Box problem inherent in some AI Entities. For this reason, this Comment posits that, just because an AI Entity might (or might not) qualify under this test as a person under the hearsay rule, it does not necessarily mean it should (or should not) be considered a *declarant* under the hearsay rule. Therefore, whether or not an AI Entity qualifies as a person for hearsay purposes under this sliding scale test will impact only how and to what extent its statements implicate the hearsay rule.

IV. Analysis: Future AI Entities as Declarants Under the Hearsay Rule

This Part first discusses why AI Entities cannot be treated the same as other machines under the hearsay rule. It then suggests and applies a hybrid approach to statements made by AI Entities, combining a purpose-based approach and a Turing-based approach to resolve whether the hearsay doctrine is implicated and how it should be applied.

177 *Id.*

178 Brandom, *supra* note 173.

179 *See supra* Part I.B.

A. *AI Entities Cannot Be Treated the Same as Other Machines Under the Hearsay Rule*

Inherent in the future of AI is the expectation that AI Entities will be fully autonomous, walking, talking, thinking, sentient beings that will reason and make decisions like humans.[180] Furthermore, these AI Entities' Black Box problems will only become more opaque as the field of AI progresses.[181] Finally, machine-learning AI programming does not use preprogrammed rules that tell a computer exactly what to do or "say," but, rather, it uses only rules that tell the AI how to learn from the data it encounters.[182] These characteristics of future AI Entities render them substantially different from the basic machines of today, and so their statements cannot be treated the same under the hearsay rule.

First, because AI Entities essentially think for themselves, their statements are their own, and, therefore, any human interaction required to operate or maintain the AI does not implicate hearsay. Second, statements made by AI Entities cannot be assessed for reliability or authenticated the same way as other machines, making the hearsay rule the most appropriate evidentiary safeguard for their credibility.

1. The "Human Interaction" Problem

Unlike "statements" made by other machines, statements made by these AI Entities cannot rest on human input for implicating the hearsay rule. While the degree and type of human interaction required to operate or maintain an AI Entity may vary, "output" produced by machine-learning AI Entities are not that of the human interactor, but, rather, solely that of the AI Entity.

It has been suggested that every computer program is reducible to a human programmer's commands, and, thus, any output by a machine *is* a human statement.[183] This theory addresses machine-learning algorithms by stating that "[t]he name . . . is misleading . . . because it suggests a fundamental autonomy from the programmer that does not exist."[184] This line of reasoning is predicated on the assertion that "a [machine-learning] program is only capable of modification because of (and pursuant to)

---

180 *See supra* Part I.

181 *See supra* Part I.B.

182 *See supra* Part I.A.

183 *See* Chessman, *supra* note 66, at 184–85.

184 *Id.* at 181 n.9; *see also* Teppler, *supra* note 13, at 232 ("When a certain condition or conditions are met, I (the computer programmer or system administrator) want you (the computer) to say 'this' and nothing else, on my behalf.").

human-programmed code."[185] Even though "humans are one step removed from program output," they are not entirely removed from the program.[186] This theory necessarily results in a finding that all CGE implicates the hearsay rule, essentially eliminating the CGE and CSE distinction altogether.[187]

While this approach is persuasive and extremely easy to apply, it misses the point of machine learning and AI by failing to recognize how it actually works and, more important to this Comment, how it is expected to work in the future. In fact, AI is already able to rewrite its own code, further removing the programmer from statements it makes.[188] With machine learning, the human programmer provides a directive and a framework for a machine to learn from its environment and from experience to accomplish that directive. To then say that all output by that program are actually statements by the human programmer and not the machine itself would be similar to saying that all statements by an employee with a directive from their supervisor were made by the supervisor and not the employee because the supervisor provided the framework for the employee's development. Ultimately, like most all-human sources, the AI Entity "is influenced by others, but is still a source whose credibility is at issue."[189]

Thus, even if an AI Entity requires human interaction to operate, statements made by an AI Entity should not be treated the same as "statements" made by mere machines. Fundamental to the nature of these future AI Entities is the idea that they "think" and make decisions without the assistance of a human, and so their statements are their own.

2. The Authentication Problem

Furthermore, because of Black Box risks and credibility issues, statements made by AI Entities cannot necessarily be assessed for reliability or authenticated in the same way as traditional machines. The general authentication standard may be difficult to meet for one who wants an AI Entity's out-of-court statement to be admitted because the AI's input and functionality may be incapable of assessment. Thus—because machine-learning AI Entities have Black Box issues that prevent reliability testing and render the cross-examination of their programmer

185 Chessman, *supra* note 66, at 181 n.9.

186 *See id.* at 185.

187 *See supra* Part II.B.

188 *See* Michael Grothaus, *An AI Can Now Write Its Own Code*, FAST COMPANY (Apr. 27, 2018), https://perma.cc/8RES-KQZ6.

189 Roth, *supra* note 11, at 1978–79.

useless—if AI Entities' out-of-court statements do not implicate the hearsay rule, then there may be no path to their admissibility. In a way, the hearsay rule, through its exceptions, may offer these statements a path to admissibility that might not otherwise be available.

On the other hand, if a judge applies a lenient authentication standard, it will likely result in potentially inaccurate statements being admitted to evidence without any safeguards to assess their credibility. The hearsay rule serves as the appropriate evidentiary safeguard for these statements as well as a pathway for admissibility for those statements by Black Box AI Entities that cannot meet the standard of authentication.

B. *Possible Approaches to AI Entity Statements*

If the purpose of the hearsay rule is to ensure the credibility and reliability of out-of-court statements, then a purpose-based approach would exclude as hearsay those statements in which their credibility and reliability cannot be assessed. This is a somewhat straightforward approach: either the reliability can be assessed, or it cannot. If it can be assessed—the AI Entity's decision-making process is somewhat transparent—then it can be admitted through the authentication rules. If it cannot—the AI Entity's decision-making process is entirely opaque to humans—it will be excluded unless it can be admitted under another hearsay exception.[190]

Furthermore, if the hearsay rule exists to simply exclude statements which cannot be shown to be free of deliberate or conscious misrepresentation, for the rule to apply, an AI Entity must be *capable* of deliberate or conscious misrepresentation. In fact, some courts have held that machines fall outside the scope of hearsay because "there is no possibility of a conscious misrepresentation, and the possibility of inaccurate or misleading data only materializes if the machine is not functioning properly."[191] However, this is too narrow of a view of the hearsay rule.[192] The rule's core concern is the reliability and credibility of statements made by declarants who are not cross-examined, not that the declarant may have lied. Even a human declarant can misinterpret or

---

[190] For a discussion of Black Box AI and degrees of transparency and opaqueness, see Bathaee, *supra* note 7, at 905–06.

[191] *See, e.g.*, State v. Armstead, 432 So. 2d 837, 840 (La. 1983).

[192] Furthermore, it's not that far-fetched to imagine an AI Entity that deliberately misrepresents a fact (i.e., tells a lie). *See* Roth, *supra* note 11, at 1991. In fact, it has been suggested that deliberate deception by robots is desirable and even necessary if robots are to contribute effectively to society. George Dvorsky, *Why We'll Eventually Want Our Robots to Deceive Us*, GIZMODO (Oct. 4, 2017, 10:10 AM), https://perma.cc/W2Q6-8NQR. HAL 9000 is a popular culture example of an AI Entity that was programmed to lie. *See* ARTHUR C. CLARKE, 2001: A SPACE ODYSSEY (1968).

unconsciously misrepresent a situation, and, in the absence of the ability to assess the reliability of that human declarant's statement through cross-examination, it would be excluded as hearsay. The same should apply with equal if not greater force to statements made by Black Box AI Entities.

Finally, while the untested credibility of a declarant is often offered as a justification for the hearsay rule, that theory does not tell the whole story about why hearsay is excluded. According to Professor Roger Park, "[t]he history of the hearsay rule indicates that lawmakers had a number of other concerns, and these concerns are reflected in the structure of the hearsay rule and its exceptions."[193]

A purpose-based approach, however, fails to address the textual issues with the hearsay rule itself: "Only a *person* may be a declarant and make a statement."[194] A Turing-based[195] approach to the hearsay rule would consider all statements made by AI Entities as hearsay if the AI Entity meets a certain set of criteria or has a certain set of pre-determined, person-like qualities.[196] Under this approach, statements made by an AI Entity that qualifies as a person under the proposed sliding-scale test for AI personhood[197] would be treated exactly the same as statements made by humans for hearsay purposes.

While this approach treats statements made by AI Entities that meet these criteria the same as those made by humans, it fails to address the differences between humans and AI Entities[198] and may result in admitting or excluding statements that otherwise would not be if made by a human.

C. *A Better, Hybrid Approach to AI Entity Statements*

Neither of the approaches above, on their own, produce a satisfactory result because, while a purpose-based approach considers how reliable an AI Entity is, it fails to address the case law stating that "only a *person* can be a declarant"[199] and that "a machine is not a person."[200] Moreover, a

193 Park, *supra* note 82, at 56 (footnote omitted) (providing an overview of these concerns).

194 United States v. Washington, 498 F.3d 225, 231 (4th Cir. 2007).

195 By referring to this as a *Turing*-based approach, this Comment does not aim to suggest that only AI Entities that have passed the Turing Test will qualify as a person under this approach. Rather, the name is simply a general reference to AI personhood to be tested using the proposed sliding-scale test for AI personhood under the hearsay rule. *See supra* Part III.B.

196 *See id.*

197 *See id.*

198 For instance, the credibility risks with human declarants may not be the same as those of AI Entities. *See supra* Part I.B. ("Black-Box AI Entities will make decisions in the same way humans do, but without any means for communicating how or why they did so.").

199 *Washington*, 498 F.3d at 231.

200 *See, e.g.*, People v. Dinardo, 801 N.W.2d 73, 79 (Mich. Ct. App. 2010).

Turing-based approach considers only whether the AI Entity can be considered a person and ignores the credibility issues altogether. Thus, this Comment proposes a hybrid approach using the sliding-scale test for hearsay personhood,[201] combined with a measure to define the degree of Black Box credibility, to determine how an AI statement is treated under the reliability safeguards of authentication and hearsay. Table 2 exemplifies this hybrid approach to the assessment of out-of-court statements made by AI Entities.

Table 2: Hybrid Sliding-Scale & Credibility Standard Approach

	Transparent (Weak Black Box)	**Opaque (Strong Black Box)**
More Person-Like for Hearsay Purposes	Not a Declarant Heightened authentication standard	Declarant Traditional application of hearsay rule; ultra-relaxed authentication standard, if applicable
Less Person-Like for Hearsay Purposes	Not a Declarant Traditional methods for introducing machine generated evidence can be used	Declarant Relaxed hearsay standard

Like the sliding scale test for hearsay personhood,[202] this hybrid approach to the admissibility of out-of-court statements made by AI Entities will consist of four main quadrants. First, when an AI Entity is not considered a person under the sliding-scale test for AI personhood[203] and the Black Box is transparent (i.e., can be reverse-engineered to allow even "a limited [or] imprecise ability to predict how the [AI Entity] will make its decisions"[204]), out-of-court statements made by the AI Entity should be evaluated under the traditional authentication methods for introducing CGE.[205] A similar analysis applies when an AI Entity *is* considered a person and has a transparent Black Box. Under these circumstances, however, the authentication standard should be somewhat heightened to account for the AI Entity's personhood, which would normally implicate the exclusionary rules of hearsay if the statement maker were human. Notice

201 *See supra* Part III.B.

202 *See id.*

203 *See id.*

204 Bathaee, *supra* note 7, at 906.

205 *See supra* Part II.B.

that under both of the "transparent" quadrants, the hearsay rule is not implicated. This is because, assuming that the purpose of the hearsay rule is to exclude unreliable out-of-court statements which the credibility of cannot be assessed, statements in which the credibility *can* be assessed would not implicate the hearsay rule.

On the other end of the sliding-scale, if the AI Entity has a substantial degree of Black Box credibility problems, then the AI Entity will be considered a declarant, but the standard applied to its statements will depend on whether and to what degree the AI Entity qualifies as a person under the sliding-scale test for AI personhood. If the declarant AI Entity is not considered a person under the hearsay rule, its statements will be analyzed under a relaxed hearsay standard. This is because hearsay would not be implicated at all if the AI Entity were not a Black Box. This relaxed standard could take many different forms. One form could be a presumption of admissibility for any statements that meet the bare requirements under a given exception, but allowing the opponent to challenge the trustworthiness of the statement.[206] To the contrary, the relaxed hearsay standard could take the form of a soft version of the catchall hearsay exception[207]: a presumption of exclusion, allowing circumstantial evidence of the AI Entity's credibility. Whichever standard is adopted, there will be a risk that unreliable statements by AI Entities will be admitted or that reliable statements will be excluded. Which of these alternatives best furthers the goals of adjudicative justice will be for the courts to decide.

Finally, if the AI Entity is considered a person under the hearsay rule and is an opaque, or strong, Black Box, then its statements should be evaluated under a traditional application of the hearsay rule. Furthermore, some courts may require these statements to be authenticated because even admissible "machine" hearsay must be authenticated.[208] If applicable, statements made by AI Entities that fall under this quadrant should be assessed using an ultra-relaxed authentication standard to avoid further constraining their admissibility after overcoming the hearsay hurdle. This is to ensure that credible statements are not unduly barred from

206 *See* Richter, *supra* note 85, at 902; *see also* Evan v. State, 899 P.2d 926, 928 (Alaska Ct. App. 1995) (holding that with a relaxed hearsay rule, the court may rely on hearsay unless the opposing party presents evidence of a reasonable indication that the hearsay statements are inaccurate).

207 FED. R. EVID. 807.

208 *See* Liesa L. Richter, *Don't Just Do Something!: E-Hearsay, the Present Sense Impression, and the Case for Caution in the Rulemaking Process*, 61 AM. U. L. REV. 1657, 1684 (2012) ("[U]nlike oral present sense impressions repeated by live witnesses, e-hearsay will have to be authenticated pursuant to Article 9 of the Federal Rules before being admitted into evidence."); Teppler, *supra* note 13, at 223 ("In order for evidence to be admissible, it must be identified or authenticated by extrinsic evidence in a manner that complies with Rule 901(a).").

admissibility for a lack of ability to explain the AI Entity's decision-making process.

That said, not all statements made by AI Entities will be in "machine" output format. In fact, this Comment imagines a future in which most statements made by AI Entities will be verbal. Whether oral statements made by AI Entities must be authenticated is a question that may remain unanswered until such circumstances arise, but this Comment posits that oral statements by AI Entities should be treated the same as oral statements made by humans.[209]

Though some modern AI Entities may qualify as a declarant, this Comment is concerned mostly with future AI Entities. AI will eventually reach AGI or even ASI,[210] producing AI Entities that we now only imagine in science-fiction books and films. It may be that "the improvement of AI requires . . . more complexity that will cause a further lack of transparency,"[211] resulting in very few AI Entities that will meet the "transparent" threshold on the sliding-scale. If this is the case, there may exist a presupposition of Black Box problems, resulting in a presumption that all future AI Entities are prima facie declarants on some level under this sliding-scale approach and shifting the burden to the other party to prove transparency.

Like the test for AI personhood,[212] this sliding-scale approach to out-of-court statements made by AI Entities will need fine-tuning, and courts will need to remain flexible to adapt the rules to the developing technology. It may be that no AI Entities will reach the highest level of the sliding scale until (and unless) the most advanced and sentient AI Entities arrive, resulting in all hearsay statements made by AI Entities being assessed using the relaxed hearsay standard. This will be up to the triers of fact, however, as they will decide what human qualities make an AI Entity more or less person-like.

The purpose of the sliding scale approach is not to search for a way to render *all* future AI Entities declarants under the hearsay rule. Rather, it provides a framework for the evidentiary treatment of statements made by AI Entities. Ultimately, the implementation of this sliding-scale approach will not require revision of the existing structure of the hearsay or authentication rules. Furthermore, though this Comment does not advocate for a disjointed system of evidence analysis,[213] this approach will

209 *See* Richter, *supra* note 208, at 1684.

210 *See supra* Part I.A.

211 Bathaee, *supra* note 7, at 929.

212 *See supra* Part III.B.

213 In fact, this Comment is quite sympathetic to the view that "the need is clear for the adoption of a uniform and well-articulated approach to the admissibility of [CGE]." Teppler, *supra* note 13, at 229. However, this Comment recognizes that substantive changes in evidence law do not happen

allow each jurisdiction to maintain its current methodology for the admissibility of non-AI machine evidence.

D. *Admissibility of Hearsay Statements Made by AI Entity Declarants*

If an AI Entity qualifies as a declarant under the hearsay rule, then any out-of-court statements it makes will be analyzed using traditional hearsay analysis. But since the hearsay rule and its exceptions were designed with *human* declarants in mind,[214] one can imagine that an AI Entity's statement, even if made by an AI Entity that is fully autonomous and essentially indistinguishable from a human,[215] may not neatly fit under any hearsay exceptions. For instance, can an AI Entity have a state of mind[216] or a present sense impression?[217] Will an AI Entity understand the potential consequences when it makes a statement against its own interest?[218] Perhaps the most advanced sentient versions of future ASI[219] Entities, should they ever exist, will possess these qualities, but it is far less clear that the more probable future AI Entities will.

How, then, will their hearsay statements be admitted if they do not squarely fall under any of the hearsay exceptions? Some will argue that for some time, perhaps they should not be admitted, or, perhaps that they should be admitted under the proposed relaxed hearsay standard.[220] This approach could be promising because, though many of the hearsay exceptions seem to presuppose the existence of certain qualities such as a state of mind or a self-interest, the rules do not explicitly impose such a requirement or explicitly state that only a human is capable of having those qualities.

Thus, statements like the first statement made by Ava in the hypothetical sequel[221] to *Ex Machina* might be admitted to prove Ava's

overnight and that uniformity in the judicial approach to CGE—especially a well-articulated approach—may never occur. Thus, this Comment advocates for a more practical accommodation of the hearsay rule: that it classify AI Entities meeting the proposed tests as "persons."

214 *See supra* Part II.B.

215 *See supra* Part III.B.

216 FED. R. EVID. 803(3).

217 FED. R. EVID. 803(1).

218 FED. R. EVID. 804(b)(3).

219 *See supra* Part I.A.2.

220 *See supra* Part IV.C.

221 Recall Ava's hypothetical out-of-court statements to the testifying pilot: (1) Nathan was a raging alcoholic; (2) Nathan was actually killed by a female; and (3) Ava had, in fact, played a part in Nathan's murder.

"state of mind"[222]—not necessarily to show that Ava was experiencing the "human" emotion of fear, but rather to show that she simply wanted to escape Nathan. While it is not clear whether Ava has a "state of mind" in the traditional *human* sense of the phrase, her behavior indicates a "state of mind" that meets the basic standard of the exception: she had a *plan* to leave Nathan's secluded compound to escape her captor.

Furthermore, Ava's statement that she played a part in Nathan's murder is clearly a statement against her own interest because it implicates her in Nathan's murder.[223] An AI Entity with full access to the internet that desires to experience the world as a human[224] would surely recognize that the consequences of her actions might lead to her losing the opportunity to fulfill this desire. Under this rationale, these statements could be admitted as statements against the declarant's interest[225] under the traditional, non-relaxed, hearsay standard exceptions.

Admittedly, these examples might be more straightforward than what might actually occur. It will be less clear, for instance, what types of hearsay statements will be made by future AI Entities such as AGI facial recognition software or advanced versions of Siri or Alexa—AI Entities resembling HAL 9000 from the novel *2001: A Space Odyssey*[226]—and how they might be admitted under an exception. Could a "statement" or output of one such AI Entity's operating conditions at a particular moment, like a black box on a plane, be considered a "state of mind" or a "present sense impression"? This is an interesting question without a clear answer.

Ultimately, however, like the many different the paths to admissibility for other types of hearsay, paths to the admissibility of hearsay statements by AI Entities will likely need to be developed organically through case law.

Conclusion

This Comment speculates about a future in which AI Entities walk, talk, and work among us and asks whether those AI Entities could be considered declarants under the hearsay rule. Ultimately, this Comment proposes two sliding-scale tests that work together to produce the answer

222 FED. R. EVID. 803(3) (admitting "statement[s] of [a] declarant's then-existing state of mind (such as motive, intent, or plan) or emotional, sensory, or physical condition (such as mental feeling, pain, or bodily health)").

223 FED. R. EVID. 804(b)(3)(A) (admitting statements that "a reasonable person in the declarant's position would have made only if the person believed it to be true because, when made, it was so contrary to the declarant's . . . interest").

224 *See* EX MACHINA, *supra* note 3.

225 *See supra* note 223.

226 CLARKE, *supra* note 192.

to this question, arguing that AI Entities could, and under certain circumstances[227] *should* be considered declarants under the hearsay rule.

Whether fully autonomous AI Entities should exist and whether the risks of such entities outweigh the benefits is a question of ethics, not law. But, between Apple's Siri, self-driving cars, and Hanson Robotics's robot Sophia, AI is increasingly becoming a part of daily life. There's still a long way to go, but the law is at an inflection point and the legal community will soon be forced to make certain determinations,[228] draw certain lines, and expand or retract certain doctrines to accommodate the fast-moving progress in AI.

This Comment attempts to find a way to analyze statements made by future AI Entities in a way that does not require an expansion of the hearsay rule itself, just an accommodation. However, as AI technology grows, case law will begin to rapidly develop. Courts may find that even this approach is too difficult to apply, and it may become necessary to rethink the hearsay doctrine altogether to make room for these developments. Otherwise, the purpose of the hearsay rule will be frustrated by admitting statements made by person-like entities that would clearly have been excluded if made by a human.

What, then, is the ruling on our evidentiary dilemma posed in the hypothetical sequel to *Ex Machina* in the Introduction to this Comment? Should the prosecution's hearsay objection be sustained, excluding Ava's out-of-court statements to the testifying helicopter pilot as hearsay?

Recall that Ava, a fully autonomous AI Entity, effectively passed a version of the Turing Test. She was able to influence her Turing Test human judge, Caleb, to develop feelings for her and help her escape her overprotective creator, Nathan. To accomplish this, she used sexuality and emotional manipulation. She was either a fully sentient being, or she was a master at mimicking human emotions. She desired freedom. She desired the experience of being human. Does the fact that she was programmed to desire these things make her any less person-like? Perhaps not. Humans are, in a way, preprogrammed to desire these things too. Ava is fully autonomous, and she has a high degree of human qualities.[229] Therefore, she is a person under the hearsay rule.

But is she a declarant? This will largely depend on the degree of transparency in her programming. While this was not explicitly discussed in the film, we will assume that Ava's AI is a Black Box.[230] Thus, a quick

227 *See supra* Parts I.B, II.A.

228 *See* JAMES BOYLE, BROOKINGS INST., ENDOWED BY THEIR CREATOR? THE FUTURE OF CONSTITUTIONAL PERSONHOOD 14 (2011), https://perma.cc/7B8V-VX2L.

229 *See supra* Part III.B.

230 This is because Ava's brain was made of a structured gel called "[w]etware," designed to rearrange on a molecular level. EX MACHINA, *supra* note 3. Nathan, Ava's creator, described Ava's

analysis using the proposed sliding-scale tests seems to result in at least implicating the hearsay rule. Nonetheless, "Objection overruled." Do you agree?

wetware brain as "chaotic" in physical structure, which suggests there could be behaviors that were not programmed and were unaccounted for. *Id.* Indeed, Nathan did not seem to have predicted Ava's violent turn on her creator. *See id.*