COULD ROBOT JUDGES BELIEVE? EPISTEMIC AMBITIONS OF THE CRIMINAL TRIAL AS WE APPROACH THE DIGITAL AGE. A COMMENT ON SARAH SUMMERS «EPISTEMIC AMBITIONS OF THE CRIMINAL TRIAL: TRUTH, PROOF, AND RIGHTS»[*]

Prof. Dr. Sabine Gless
sabine.gless@unibas.ch

ABSTRACT: Criminal proof is unique, in that it must be able to account for the justification of both: accurate fact-finding and a fair trial. This is Sarah Summers' main message in her article on the epistemic ambitions of the criminal trial, which focusses on belief as a sort of proxy for societal acceptance of truth as a set of facts established by compliance to procedural rules. This commentary tests her finding by scrutinizing whether it is conceivable that robots, complying to all rules, assist in fact-finding with a specific form of legal belief based on a sophisticated probability weighting opaque to humans. The result is in accordance with Sarah Summers: as long as robots cannot explain their beliefs, any criminal proof based on them flounders as it can neither be part of a fair trial nor ensure acceptance in the existing institutional framework.

KEYWORDS: criminal proof, robot judges, legal belief, participation rights in criminal trials, evidence law, electronic monk.

SUMMARY: 1. *TATORT*—. 2. LEGAL BELIEFS: 2.1. Belief Bridging Evidence to Proof. 2.2. Fact-Finding Tailored to Humans. 2.3. Legal Belief as an Epistemic Category—. 3. FUTURE BELIEFS: 3.1. Robots Assisting in Criminal Proof. 3.2. Can Robots Believe? 3.3. Could Robots

1. *TATORT*

«*Glauben können sie in der Kirche. Was wir brauchen sind Beweise*»[1]. Sarah Summers strikes a chord with German speaking legal scholars using the prosecutor's curt reply to a detective's plea for an arrest warrant from an episode of *Tatort* as a thematic anchor. The district attorney's reaction raises the question: «Does belief only have a place in Church?». Over its 50-year run, the iconic detective series *Tatort* has seemed to suggest that the answer is «No». Sarah Summers is a mediator between inquisitorial jurisdictions, which engage less with the very important discussion surrounding evidence, proof, and acceptance of fact-finding, and adversarial jurisdictions, which produced a rich body of evidentiary laws and excellent scholarship on its significance. Therefore, she is excellently equipped to discuss the theoretical underpinnings and practical impacts of the «Epistemic Ambitions of the Criminal Trial: Truth, Knowledge, and the Rights of the Accused», offering a chance to better understand our normative legal world by using the theoretical approach of epistemology[2]. Commenting on her article[3], this paper aims to: a) focus on belief as a crucial element that bridges evidence proffered and fact-finding, and b) change perspectives as the digital age approaches by asking: could a robot believe? And, thereby, assist a human judge with the assessment of evidence?

My search for an answer starts with an analysis of Sarah Summers point of departure on the epistemic ambitions of the criminal trial and her claim that legal fact-finding is primarily aiming for veracity, but that legal adjudication needs not only a process aiming at «accurate outcomes», but also «fair procedures» to achieve «truth». Against this backdrop, this paper scrutinizes whether robots could help human judges in bridging the gap from evidence to proof (as the promoters of digitization of the judiciary might hope).

2. LEGAL BELIEFS

Situating herself in the vast domain of legal epistemology, Sarah Summers explains in her paper (elaborating on her previous work) that the *legal concept* of *believing* is distinguishable from other concepts of believing, for instance from a spiritual

---

[1] *Staatsanwältin Jasmin Winterstein, Tatort Mainz*: *In seinen Augen*, Das Erste, Deutschland, 26.6.22.

[2] For more details, see Ross (2022).

[3] This article also covers work of Sarah Summers' for her panel *Evidence and the European Court of Human Rights* at Michele Taruffo Girona Evidence Week, May 23-27, 2022.

concept shared among churchgoers. This idea is based on there being a legal concept of evidence that can be distinguished from the everyday concept of evidence[4]. Consequently, she illustrates that legal adjudication differs from other forms of decision making (negotiations, democratic elections) with regard to the procedure followed and the extent and nature of the participation of those involved.

2.1.   Belief Bridging Evidence to Proof

In the first part of her article Sarah Summers adds to the many concepts of believing we share in ordinary life (like belief in God or in the love of someone), the notion of a legal belief: A judge «legally believes» if she trusts the result of a process that complies with all the relevant procedural rules. The rationality of this step of legal adjudication is guaranteed in large part by the way information is gathered, presented in court and processed later than through the subjective belief of the factfinder[5].

In convincing detail, Summers explains that what might at first blush appear a rather formal requirement of strict adherence to a procedure that ensures a high degree of veracity, actually requires more: a value-based search for truth. This is because, in the criminal trial context of a proposition of facts established by evidence, *belief* goes beyond individual acceptance of certain evidence as proof, towards a sort of proxy for societal *acceptance* of truth as a set of facts. This is established by presenting evidence in a fair trial and convincing the triers of facts to believe them.

It is consistent with this line of argument that Sarah Summers criticises the dichotomy when «legal adjudication is understood as capable of being divided into a process involving "accurate outcomes" as distinct from "fair procedure"». She rightly explains that to legally believe also requires compliance with the concept of a «fair trial». Even the much-used example of exclusionary rules limiting the search for truth (Gless and Richter, 2019, p. 5) shows that a dichotomy of «true accuracy» and «fair evidentiary proceedings» is blurry, as exclusionary rules also uphold accuracy of and trust in fact-finding. For instance, when blocking confessions forced from humans, which are of highly suspicious veracity[6].

2.2.   Fact-Finding Tailored to Humans

Summers, like Duff *et al.* (2007) in their work on the narratives of criminal trials that she refers to, depicts the criminal trial process as a human process of calling to account another human. One important element of socially acceptable fact-finding

---

[4]   She therefore refers to Ho (2021).
[5]   *Cf.* for German criminal evidence: Gless (2001, p. 306-307).
[6]   The most extreme case are confessions forced through torture, *cf.* Summers (2022, p. 85).

is meaningful deliberations among all parties of such trials[7], that respect the human rights of all. This is of paramount importance, because—as Sarah Summers explains—the interest in finding the truth has led to procedural rules that expose suspects and witnesses to coercive measures, which frequently interfere with individual rights. Individuals deserve respect throughout the criminal trial, also in order to ensure public acceptance of the verdict.

The way the criminal trial process is tailored for humans is an important part of the institutional framework in which it operates. The process involves a specific setting for presenting and evaluating evidence with particular acknowledgment of the legal position of the defendant who has an «institutionally protected opportunity to present proofs and arguments for a decision in his favor» (Fuller, 1960, p. 2). Building on her previous work, Sarah Summers emphasises the specific institutional setting of the criminal trial and its focus on the human and human rights protection using case law of the ECtHR. In particular, she points to the right to information on the charge, to sufficient time to prepare, to the assistance of counsel, to question witnesses (Article 6[3] ECHR) and to the overall right to be heard (Article 6[1] ECHR). These rights are all inherent to the concept of «fair hearing», and together work towards the right of those charged with a crime to receive a meaningful explanation of the verdict[8].

2.3.   Legal Belief as an Epistemic Category

Sarah Summers' article depicts the criminal trial as a process primarily aimed at a formal legal truth that is nevertheless saturated with human values. The article, however, remains slightly vague on how judges are to decide between conflicting interests when following evidentiary rules that are designed to maximise the likelihood of accurate outcomes[9]. The (apparent) dichotomy between veracity and a fair trial that Sarah Summers shows is shaped by the ECHR case law itself, naturally taking the specific angle of defending individual rights. However, as she (and John Jackson) rightly pointed out, the ECtHR, even when prohibiting the use of evidence obtained in violation of certain human rights, has never made clear how «veracity» and «fairness» relate to each other (Jackson and Summers, 2012, p. 151-152).

Notwithstanding unclarified details, this fundamental consideration appears clearly in Sarah Summers' article: for her, legal truth and proof are epistemic categories that are concerned with the establishment of the existence or nonexistence of

---

[7] Luhmann offers a sociological analysis for this, for instance in «Legitimation durch Verfahren» (1983, p. 28-37).

[8] See *e. g. Berhani v Albania* (2010, §12); *Khamidov v Russia* (2007, §107); *Ajdarić v Croatia* (2011, §47-52); *Anđelković v Serbia* (2013, §26-29).

[9] See also Duff *et al.* (2007, p. 89).

facts «to the satisfaction of a legal tribunal» that decides on the verdict[10]. This conviction is based on scholarship covering legal epistemology, the criminal trial process and doxastic logic concerned with reasoning about beliefs when handing down a verdict[11]. The question is: does it have a promising future?

3.   FUTURE BELIEFS

Our world is being transformed by new technology. Thus, the question arises: what form of legal belief will be acceptable in a digitized future? Or rather: what will belief as an epistemic category mean when robots assist human judges? Theoretically, all procedural rules can be coded into such smart systems. However, a robot would not base its inferences on a form of human faith that something is true, but rather on data crunching and sophisticated statistical methods. Could a robot judge bridge the gap from evidence to proof, transcending «human fallibility», with a specific robot belief?

3.1.   Robots Assisting in Criminal Proof

Smart tools, driven by artificial intelligence (AI), already assist in the process of adjudication. As things stand today, however, such robots cannot replace the humans on the bench.

Nevertheless, they are already enhancing the judicial process in certain situations[12]. For instance, robots assist with fact-finding in criminal procedure through their use in digitized breathalyzers that measure and record breath alcohol concentration or when smart systems are used to help with DNA-analysis[13]. Robots that could assist human judges promise a scientific based assessment, accompanied by seemingly unlimited resources for information processing and learning capacity. Yet, as we all know, robots do have flaws. They cannot explain their reasoning nor their results (Roth, 2017, p. 1988-1990). Complex systems, which are trained with machine-learning, in particular cannot provide meaningful explanations of their inferences. As a result, they have been depicted as «mindless agents», even when they produce a decision (Hildebrandt, 2015, p. 22).

Weighing pros and cons, the technological development nevertheless suggests that the computing power of robots will make them an interesting option for proofing crimes in the future. A case reported in 2016 by Swiss news media[14] provides

---

[10]   She here references Twining (2006, p. 193).
[11]   For more details, see Ross (2020).
[12]   For a more detailed overview, see Ashley (2017, p. 3-37).
[13]   *Cf.* Gless (2020, p. 198, 207-215).
[14]   For details of the press coverage, see https://bit.ly/3wsPqVM.

a glimpse of potential proofs in a future in which robots monitor human behavior in real time and offer an «evidence-based» assessment of an allegedly criminal act. Our example involves a drowsiness detection system embedded in a modern car which ran over a scooter driver, injuring her badly. The drowsiness detection system had repeatedly alerted the driver prior to the accident of driving errors made due to perceived fatigue, alerts that were ignored by the driver. The system collects data on the driver's steering movements, body tension, seating position, respiratory rate and/or eyelid movements, evaluates these indicators for «signs of drowsiness» and, on the basis of complex algorithms and elements of machine learning, chooses whether or not to issue an alert to the driver (Ramzan *et al.*, 2019, p. 61908). As the events antecedent to the accident were unreproducible, only the recorded alert could be proffered as a proof of the unfitness of the driver to steer a vehicle when running into the motor scooter driver. Thus, the judge had to either believe what the robot believed, that the driver was indeed unfit, or to discard the proof.

3.2.   Can Robots Believe?

Robots can be coded to: a) assess human behavior that is ambiguous and requires an individual appraisal, and b) accept at a certain threshold, based on their evaluation (derived from information processed from their sensors), that a human behaves in that state. Drowsiness detection systems in particular are an interesting example as drowsiness is a vague concept. And yet, these robots are still capable of monitoring human drivers and determining, at a predetermined threshold, that a human's behavior is indicative of fatigue, based on the data they process and the use of statistical methods. Drowsiness detection systems don't have to operate with 100% certainty, as the threshold for issuing a warning can be set lower, at 80% certainty for instance. If a drowsiness detection system assesses a human driver's steering movements, muscle convulsion, squinting eyes and issues a warning for fatigue as a result, can we say that this warning was the result of the system believing that the driver was fatigued in an epistemic sense?

If we argue that belief is a broad term, with one being able to believe in many different ways in everyday life, one could think that a robot is also capable of believing, following this understanding of belief. Take our drowsiness detection system, for instance, that issues an alert when the information reported back from its sensors crosses a predetermined threshold based on its machine-learning experience. Under this broader conception of belief, it could be said that the system issued the warning because it believed that the driver was fatigued and therefore incapable of steering the vehicle safely.

However, if one insists, like Sarah Summers might do, that belief is a psychological state aiming at truth, then robots—lacking a psyche—cannot believe. The question, however, arises as to whether what she is really asking for is a human mind, or whether a set of faculties capable of intuitive thinking that allows for the adequate

assessment of evidence is all that is necessary. This can be important as we trust robots with very intuitive assessments based on probability weighting. For example, when we trust drowsiness detection systems to assess whether a human is sleepy or just happens to swerve whilst evading toads crossing the street, or is relaxing after a workout in the gym, or is squinting her eyes because of a trachoma (Gless *et al.*, 2022, p. 290-291).

Before getting lost in a complicated discussion about the epistemic properties of probability-weighted evidence approaches in ordinary life, we must foreshadow the issues crucial for legal belief to be designated as an epistemic category. As Sarah Summers points out, there are obvious parallels between the forming of factual beliefs in everyday life and in the legal context. The factfinder in legal proceedings—like an individual engaged in factual deliberations in everyday life—is concerned with the establishment of true belief, primarily aiming at veracity. The same is true for the robot that aims to reproduce reality when issuing a drowsiness alert. Ultimately, it may just be an issue of semantics, as to whether we can claim that robots «believe» as we conceive in everyday language. For instance, it is likely that many of us have been engaged in a discussion as to what the optimum route to take is whilst driving, where someone insists that the navigation system on their smartphone «believes» another route is better.

3.3.   Could Robots Legally Believe?

Could robots legally believe? At first glance, they are in a much better position to do so in law than they are in ordinary life. Robots can be made to follow all relevant rules without the possibility of deviation, due to the prescriptive nature of coding. Thus, they comply with all the standards necessary to ensure an accurate outcome, *i. e.* compliance with the relevant rules. And—following Sarah Summers line of argument—they meet the specific characteristics of legal belief—as legal belief differs from ordinary belief; just as legal evidence differs from the ordinary concept of evidence (Ho, 2021).

Representatives of legal doxasticism, like Ross (2022, p. 8)—who Sarah Summers refers to in her article—, would reject this because robots can only believe through statistics and probability and thus lack the relevant epistemic status to believe [15]. Sarah Summers seems to raise the same argument in a slightly different way when she claims that criminal proof must be able to account for the justification of both the ruling and the process. Even if robots can produce a high degree of confidence in the accuracy or inaccuracy of the probative facts, they most probably would fail when making a decision on fair trial issues embedded in fact-finding. Furthermore, one

---

[15]   A broader point is made by Buchak (2014), when she argues that blame must be based on belief, and not on mere credence.

can argue that it violates a human's dignity if—charged with a crime—her allegedly criminal behaviour is assessed by a robot. However, this might be more an issue of the inadequate distribution of tasks than a fundamental objection against assigning a robot a capacity for legal belief.

Using a robot to evaluate evidence could, conversely, ground this specific part of adjudication with a rational foundation that is far stronger than some individual's belief. This is because robots can, ideally, be calibrated towards objectivity, thereby providing a stronger basis for rationality. But to go down this road, we must accept statistical or mathematical probabilities as a decisive reason in a legal sense. Sarah Summers seems reluctant to rate accuracy over certain epistemic properties. But robots will strongly pose the question of whether an evaluation based on a statistical assessment should be less worthy than a more or less well-defined belief. If one rates probability-weighted evidence higher than—possibly complicated and vague—epistemic properties, then robots have a strong likelihood of being considered capable of believing in a legal sense. Should this change occur, it would represent a departure from human fallibility and a move in lockstep with the developments in everyday language used to talk about assessments made by robots.

3.4.   Institutional Framework for Robot Beliefs

However, there is no place for robot beliefs in criminal proceedings when they cannot meet the standards of the «institutionally protected opportunity to present proofs and arguments for a decision» that Sarah Summers refers to[16]. Justly, she demands that criminal evidence be gathered and evaluated in a distinct process, designed to ensure fair treatment by guaranteeing individual rights—and that the success of a verdict cannot be based merely on a certain percentage of accuracy of the result (Buchak, 2014). The main deficiency of robots is their inability to explain themselves. A meaningful explanation of fact-finding as a human right is explained above. If we want to use robots in the future, we need an adequate institutional framework to justify their beliefs. As Sarah Summers rightly points out: «The process of adjudication takes on central importance, not just in defining the criteria according to which proceedings are to be understood as appropriate or fair, but also in relation to establishing what constitutes a "successful outcome"»[17].

The first step: establish a way for robots to give reasons for a «belief» in a meaningful way and find possibilities to meaningfully confront it. The benchmark can be the criteria set by the case law of the ECtHR on the obligation to provide reasons for the verdict[18]. The giving of reasons is of fundamental importance to treating in-

_____

[16]   Referencing Fuller (1960, p. 2).
[17]   *Cf.* Duff *et al.* (2007, p. 89).
[18]   See *Moreira Ferreira v Portugal* (2017, §84).

dividuals with dignity and allowing them to determine whether or not an exercise of authority is in fact justified[19]. It would certainly violate human dignity to be judged by a machine without reasons then being given for the judgement.

This is the reason why lie detection has been rejected as violating human dignity. The enigma is powerfully set in scene at the beginning of the movie *Blade Runner* when the protagonist runs the so called Voight-Kampff test, that is designed to distinguish replicants from humans, on a replicant who is tricked into thinking she is a human. Like the drowsiness detection system, the imaginary Voight-Kampff test is intended to measure bodily functions such as respiration, heart rate, blushing and pupillary dilation in response to emotionally provocative questions. Humans submitted to robot assessment must have access to explanations of their functioning and, if they wish so, must be given clarification for the results[20]. It is for good reason, as Sarah Summers points out, that the ECtHR assigns a «substantive» element to the obligation to give reasons. It aims to make sure that not only the individual, but «indeed the public» is able to understand the process as well as the verdict[21].

4.   LOOKING AHEAD

If one would want to answer Sarah Summers' question as to whether it is possible to identify a distinct, normative concept of criminal proof in brief, one could be tempted to draw on the widespread view that such proof foremostly requires *legal belief* in the evidence proffered, with this evidence following all relevant rules of fact-finding in a specific jurisdiction. But this answer would miss her more important message. In her article she convincingly explains that criminal proof is special, because it must be able to account for the justification of both the fact-finding and the process. A minimum requirement for every step on this path is a meaningful explanation of the fact-finding and the process to the humans that are involved in the process, as well as to the public, by the court.

This also answers my question of whether robots could legally believe, thereby assisting human judges in fact-finding. As long as robots cannot explain their beliefs, they cannot assist criminal proofs. Being able to explain one's belief is fundamental and central to eliminating major risks, and not only in criminal trials, as Douglas Adams (1987, p. 4-5) foresaw when envisioning how robots could disburden humans in the future decades ago:

> The Electric Monk was a labour-saving device, like a dishwasher or a video recorder. Dishwashers washed tedious dishes for you, thus saving you the bother of washing them yourself, video recorders watched tedious television for you, thus saving you the bother of looking at it yourself;

---

[19] *Cf.* Heinze (2021, p. 225-229).
[20] It is for a good reason that human dignity is underpinning European data protection rights, namely the GDPR which provides a right to transparency of data processing.
[21] Sarah Summers references *Lhermitte v Belgium* (2016, §66 and 67).

Electric Monks believed things for you, thus saving you what was becoming an increasingly onerous task, that of believing all the things the world expected you to believe.

Unfortunately this Electric Monk had developed a fault, and had started to believe all kinds of things, more or less at random. It was even beginning to believe things they'd have difficulty believing in Salt Lake City. It had never heard of Salt Lake City, of course. Nor had it ever heard of a quingigillion, which was roughly the number of miles between this valley and the Great Salt Lake of Utah.

In both court and church, it is good advice to carefully examine where one's faith lies.

BIBLIOGRAPHY

Adams, D. (1987). *Dirk Gently's Holistic Detective Agency*. William Heinemann.

Ashley, K. (2017). *Artificial Intelligence and Legal Analytics: New Tools for Law Practice in the Digital Age*. Cambridge University Press.

Ashworth, A. and Redmayne, M. (2010). *The Criminal Process* (4th ed.). Oxford University Press.

Buchak, L. (2014). Belief, Credence, and Norms. *Philosophical Studies, 169*(2), 1-27. https://philpapers.org/archive/BUCBCA.pdf

Duff, A., Farmer, L., Marshall, S. and Tadros, V. (2007). *The Trial on Trial: Vol. 3. Towards a Normative Theory of the Trial*. Hart Publishing.

European Convention on Human Rights (1950). Convention for the Protection of Human Rights and Fundamental Freedoms. https://echr.coe.int/Documents/Convention_ENG.pdf

Fuller, L. L. (1960). Adjudication in the Rule of Law. *Proceedings of the American Society of International Law, 54*, 1-8.

Gless, S. (2001). Zur «Beweiswürdigungs-Lösung» des BGH. *Neue juristische Wochenschrift 54*, 306-307.

Gless, S. (2020). AI in the Courtroom: A Comparative Analysis of Machine Evidence in Criminal Trials. *Georgetown Journal of International Law, 51*, 195-253.

Gless, S., Di, X. and Silverman, E. (2022). Ca(r)veat Emptor: Crowdsourcing Data to Challenge the Testimony of In-Car Technology. *Jurimetrics, 62*(3), 285-302.

Gless, S. and Richter, T. (2019). *Do Exclusionary Rules Ensure a Fair Trial? A Comparative Perspective on Evidentiary Rules*. Springer Open.

Heinze, A. (2021). Evidence Illegally Obtained by Private Investigators and Its Use Before International Criminal Tribunals. *New Criminal Law Review, 24*(2), 212-253.

Hildebrandt, M. (2015). *Smart Technologies and the End(s) of Law*. Edward Elgar.

Ho, H. L. (2021). The Legal Concept of Evidence. In E. N. Zalta and U. Nodelman (Eds.), *The Stanford Encyclopedia of Philosophy* (Winter 2021 Edition). https://plato.stanford.edu/entries/evidence-legal/#toc

Jackson, J and Summers, S. (2012). *The Internalisation of Criminal Evidence: Beyond the Common Law and Civil Law Traditions*. Cambridge University Press.

Luhmann, N. (1983). *Legitimation durch Verfahren*. Suhrkamp.

Ramzan, M., Khan, H. U., Awan, S. M., Ismail, A., Ilya, M. and Mahmood, A. (2019). A Survey on State-of-the-Art Drowsiness Detection Techniques. *IEEE Access, 7*, 61904-61919.

Ross, L. (2022). The Foundations of Criminal Law Epistemology. *Ergo*. https://www.researchgate.net/publication/361228461_The_Foundations_of_Criminal_Law_Epistemology

Roth, A. (2017). Machine Testimony. *The Yale Law Journal, 126*(7), 1972-2053.

Summers, S. (2022). Evidential Remedies for Procedural Rights Violations Comparative Criminal Evidence Law and Empirical Research. In J. Ferrer and C. Vázquez (Eds.), *Evidential Legal Reasoning: Crossing Civil Law and Common Law Traditions*. Cambridge University Press.

Twining, W. (2006). *Rethinking Evidence: Exploratory Essays* (2nd ed.). Cambridge University Press.

Williams, B. (1973). Deciding to Believe. In B. Williams (Ed.), *Problems of the Self: Philosophical Papers* (pp. 1956-1972). Cambridge University Press.

Cases and legislation

Ajdarić v Croatia, nº 20883/09, 13 December 2011.

Anđelković v Serbia, nº 1401/08, 9 April 2013.

Berhani v Albania, nº 847/05, 27 May 2010.

Khamidov v Russia, nº 72118/01, 15 November 2007.

Lhermitte v Belgium [GC], nº 34238/09, ECHR 2016.

Moreira Ferreira v Portugal (No 2) [GC], 19867/12, 11 July 2017.