THE MAN WHO WASN'T THERE AGAIN – CREATIVE INFORMATICS AND LEGAL AI

Burkhard Schafer

Abstract: As legal AI becomes a viable competitor for human lawyers, concern about technological unemployment is reaching the legal profession. One response is a re-orientation of the value proposition of legal service industry, focussing less on knowledge and analytical reasoning skills to «soft» skills such as creativity. This however makes two problematic assumptions: that creativity is a positive trait of good legal advice, and that it is indeed both beyond machine intelligence. This paper aims to challenge both assumptions.

Table of contents

A.   INTRODUCTION: CREATIVITY AND LEGAL AI

The spectre of AI competing with lawyers has raised concern about technological unemployment in the legal profession. This has in turn forced law schools, law firms and the professional bodies that regulate the legal profession to rethink the skills and knowledge that lawyers will have to have in a data-driven economy. One strategy focuses on «soft skills» such as empathetic reasoning and creative problem solving that seem beyond the ability of current or near future AI. This approach is based on a number of hidden, and often highly problematic, assumptions that are not yet sufficiently discussed or understood. It assumes in particular first, that creativity and empathy are

unquestionably good attributes in a lawyer, and second, that AI is indeed incapable now or in the very near future to replicate them.

The aim of this paper is to challenge these assumptions, focusing mainly on «creativity» as a legal skill. We will but briefly look at the question whether creativity in law is indeed desirable. This issue would lead inevitably to much deeper philosophical questions, questions that regrettably are not asked in the current debate on legal technology, let alone answered. Do we prefer judges who coldly, competently and systematically apply the rules and the rules only, ensuring predictability and security, but potentially also contribute to an ossified, inflexible and in individual cases manifestly unjust legal system? Or is our ideal of a lawyer the creative problem solver who finds an entirely new interpretation of the law, maybe to «get their client off the hook»? Questions on legal AI and its potential can't be answered without also answering some of the perennial questions of jurisprudence: what makes a lawyer a «good» or even «virtuous» lawyer? What can we legitimately expect from the justice system? How can we dream a more perfect legal order? These questions raise issues that go to the very heart of legal philosophy: what does it mean to have knowledge of the law, what distinguishes a «correct» from an «incorrect» legal decision (if this question even makes sense), what do we owe to the people affected by the justice system, and with all that ultimately the question of what justice is.

In the next section, we will briefly put the question of creativity in law into a broader historical and intellectual context. From this discussion, we will tentatively conclude that first, there are at least some expressions of human creativity that are beneficial for the administration of justice, and second, that there are indeed some prima facie reasons to think that the absence of these traits in past and present approaches to legal AI can lead to undesirable consequences.

The final section however will argue that the specific type of «benevolent creativity» that was identified in the preceding part is not necessarily outwith the capabilities of computer systems, provided they are properly integrated into the human decision making process. We will look in particular at the ability to create «alternative histories» as one expression of creativity, and provide some formal approaches that are capable of implementing them computationally.

B. CREATIVITY AND THE LAW

The development of soft skills has recently gained renewed prominence in legal educa-tion. The «creative law school» that learns from the arts and the creative industries is hailed as a new paradigm of legal training, often promoted explicitly in response to technological change.[1] As a responded in the influential Pew report noted: «There will be many things that machines can't do, such as services that require thinking, creativity, synthesizing, problem-solving, and innovating».[2]

The desire for creative legal problem solving skills is widely expressed by legal practice and legal employers. LAW (Legal Actions Worldwide), the non-profit network of hu-man rights lawyers, proudly proclaims to be a «Think Tank for Creative lawyering».[3] On the website of a large commercial law firm, we find the following: «XYZ hires pas-sionate, determined people and empowers creativity that delivers innovation to the le-gal and IP markets.» The Association for Corporate Counsels offers training courses in «emotional intelligence» for its members and identifies empathy as a key skill not suf-ficiently taught in law schools.[4] David Perla, the President of Bloomberg Law and Bloomberg BNA's Legal division, identifies empathy as the one quality that sets superior lawyers apart from the rest.[5]

Legal education increasingly responds to this market demand. The common marking scheme at Edinburgh University requires for an «excellent» mark that the student in addition to sound subject knowledge also displays «creative, subtle, and/or original in-dependent thinking». A popular training manual by Ellen C. Hill has the title «Creative

---

[1] See e.g. BOWMAN, The Rise of the Creative Law School. U. Tol. L. Rev., 2018 50, p. 255–264. CORNELL TECH, 13 Reasons Why Tech Companies Need a New Kind of Lawyer, Law Tech Blog https://tech.cornell.edu/news/3-reasons- why-tech-companies-need-a-new-kind-of-lawyer/ Accessed 10.10.2019.

[2] SMITH/ANDERSON, AI, Robotics, and the Future of Jobs. Pew Research Center 2014.https://www.pewinternet.org/2014/08/06/future-of-jobs/ accessed 10.10.19.

[3] http://www.legalactionworldwide.org/think-tank-for-creative-lawyering/ accessed 10.10.19.

[4] https://www.acc.com/education-events/2019/lawyers-empathy-applying-design-thinking-house-law-practice.

[5] https://abovethelaw.com/2016/02/whats-possible-the-empathetic-lawyer/.

Lawyering»[6]. The prospectus of the California Western Law School finally advertises that «The Creative Problem Solving area of concentration is designed to help you acquire the needed skills that demand broader and deeper understanding of people, their problems, and the consequences of confronting those problems only in narrow, legalistic ways».

It is in this last example where we can also see the potential problem with this approach. «Narrow, legalistic ways» are quite obviously a very bad thing indeed. And so, in one small sentence, an influential jurisprudential tradition, formalism and legalism in both its European and Chinese manifestation, becomes the «dark other» of legal education and legal practice. This critical approach to «legalistic education» is of course not unprecedented. Long before AI allowed the vision of justice meted out by machines to become if not a reality, then a possible near future, Roscoe Pound in his quest to reform legal education deployed the term «mechanical jurisprudence» as an aberration of legal thought:

«Legal systems have their periods in which science degenerates, [...] in which a scientific jurisprudence becomes mechanical jurisprudence. [...] The classical jurisprudence of principles had developed, by the very weight of its authority, a jurisprudence of rules; and it is in the nature of rules to operate mechanically».[7]

When Pound was mounting his critique at the beginning of the 20th century, he could build on two hundred years of mechanistic metaphors used to describe the ideal, or the dystopia, of «machine like» legal decision making. Maybe the first example of this usage can be found in Julien de La Mettrie and his «L'homme machine» from 1747. The work extended Descartes's notion that animals were best understood as «automata» to humans – including crucially their ability to normative reasoning: «To be a machine, to feel, think, know good from evil like blue from yellow». Colour recognition and moral discernment are equally within the capacity of deterministic machines, both are nothing but mechanic responses to material inputs. Again La Mettrie: «Even if man alone

---

6   HILL, Creative Lawyering, 2005 Bloomington.

7   POUND, Mechanical Jurisprudence. Colum. L. Rev.. 1908 8 p 605–623.

had received a share of natural law, would he be any less a machine for that? A few more wheels, a few more springs than in the most perfect animals.»[8]

Crucially though, for la Mettrie and the legal formalist of the 19th century, this conception of law was as much a descriptive as it was a normative concern. In the face of rampant corruption and arbitrary exercise of power, reliance on rules, applied in a transparent, verifiable and mechanic way, was a truly revolutionary safeguard to despotism. So revolutionary indeed that the Mexican peasant leader Zapata declared that he wanted «to die a slave to laws. Not to men.»[9] It informed the codification movements in France and Germany, and the ideal of the Rechtsstaat, the state under the rule of law. Nor was this understanding a purely European particularity. 2000 years before Mettrie, Shang Yang developed Chinese legalism as the appropriate governance form for a strong centralised kingdom that was both just and efficient – creating through adherence to rules a degree of equality that was disturbing enough for the aristocracy to have him eventually killed.

Legal AI, especially the symbolic-reasoning based approach (Good old fashioned AI, GOFAI) that dominated the first generation of systems developed in the 1980s and 1990s, can clearly be seen as the intellectual heirs of this understanding of the nature of law. Increased objectivity and decreased arbitrariness, especially in prima facie discretionary domains became one of the key motivating points for the introduction of AI in the justice system.[10]

---

[8]   LA METTRIE, L'Homme machine. 1747 Leyden. See also CAMPBELL, La Mettrie: the robot and the automaton. Journal of the History of Ideas, 1970 31(4), pp. 555–572 and on his conception of law THOMSON, French eighteenth-century materialists and natural law. History of European ideas, 2016 42(2), pp. 243–255.

[9]   ROSENBLUM, Heroes of Mexico, 1969 New York p. 112.

[10]   See e.g. TATA/WILSON/HUTTON, Representations of Knowledge and Discretionary Decision-Making by Decision-Support Systems: the Case of Judicial Sentencing. The Journal of Information, Law and Technology (JILT) 1996 (2); SERGOT et al The British nationality act as a logic program. Commun. ACM 1986 29(5):370–386; JOHNSON/MEAD, Legislative knowledge base systems for public administration: some practical issues. In: Proceedings of the third international conference on artificial intelligence and law, 1991, New York, pp 108–117.

Seen against this historical backdrop, the newly found enthusiasm for creativity in law as a response to inroads made by legal technology is inevitably contested and ambivalent. The position outlined so far identifies computational reasoning with consistency and rule adherence, juxtaposed with human creativity and empathy. Crucially, both promotors and critics of the use of AI in the legal system agree on this intimate connection between legal technology and predictability. This holds not only for the rule based systems of the 1980s and 90s, nor does it necessarily require a commitment to legal formalism. More recent machine learning approaches too, which are more aligned with legal realist conceptions of law, are «non-creative» in that sense. They merely widen the scope of the permissible input that is used to predict the outcome of their disagreement is mainly a normative one: in addition to the formal legal rules and precedent cases (both of which can be modelled using symbolic AI) they harvest other data about past decision making to predict the future.

Legal AI, especially when rule based, symbolic approaches and machine learning are combined, emulate in many ways Dworkin's «Judge Hercules». We recall that in Dworkin's model, every legal question has a unique right answer, an answer that an ideal judge, Hercules, with unlimited memory and legal knowledge, would always be able to find. The legal answer to any problem case is determined by the totality of applicable formal rules and precedents, and the standards and principles that the community holds. It is only our lack of time and memory, that is ultimately the shortcomings of our «computational mind» that result in legal disagreement. In theory – and AI could turn this theory into practice – with unlimited memory and time (or computational power which amounts to the same thing), these inputs uniquely determine the output to any legal question. Closely related to the metaphor of Judge Hercules is that of the law as a chain novel: The past, that is the sum total of precedent cases or «chapters» in the novel, constrain and determine the right answer for the present case, the next chapter that is to be written. This metaphor in turn aligns with the foundational assumption of machine learning, that is that the recognition of patterns in data from the past allows us to predict the future. Machine learning is inherently «conservative», just as the next author of Dworkin's chain novel is tasked to conservatively extend the story and maintain the integrity with what has happened before. Exercise of creativity,

especially creativity that radically alters the course of the story, is neither possible nor desirable.[11]

We find the same ambivalent attitude to creativity outside the legal context. Creativity as a psychological trait has also been linked to an increased tendency to dishonesty,[12] and in some disciplines, «creativity» can take on outright negative connotations. «Creative accounting» has been used since the 1960s as a euphemism for morally dubious practices that may follow the letter of the rules of standard accounting practices, while at the same time defeating their purpose. «Creative lawyering» can have the positive associations legal practice hopes for, but can also be used as euphemism for ruthless exploitation of «loopholes». Nelson and Nielsen in their influential study on legal archetypes have assigned creativity in law a «morally ambiguous edge»[13], while Barnet considers «creative compliance» enabled by unethical inhouse lawyers, as co-responsible for financial crisis.[14].

While it may seem almost trite, this debate means that we need at least to be very clear what type of creativity we mean when we argue for creativity-centric legal education as an answer to the AI challenge. This, I will argue, eventually leads to a paradox: «good» creativity in law is neither anarchy nor randomness, it has to be creativity that is constrained by rules. Furthermore, it can be taught, and the success of the teaching assessed in an objective way – at least as far as human lawyers are concerned. This however indicates that the type of creativity that is most likely to be at the very least assisted, if not replaced, by AI – its very teachability sees to this. In the next section, we will look at computational creativity and how creativity is understood and imitated in applications that come from this research tradition. In the final section, we will then build on some of these ideas to introduce an example of «benign» computational creativity in a

---

[11] DWORKIN, A Matter of Principle, 1985 Harvard p. 119 ff. see also RAZ, Dworkin: A new link in the chain. Calif. L. Rev. 1986;74:1103.

[12] GINO/ARIELY, The dark side of creativity: original thinkers can be more dishonest. *Journal of personality and social psychology,* 2012 *102*(3), 445.

[13] NELSON/NIELSEN, Cops, counsel, and entrepreneurs: Constructing the role of inside counsel in large corporations. Law and Society Review, 2000 pp 457–49.

[14] MCBARNETT, Questioning the legitimacy of compliance. Legitimacy and Compliance in Criminal Justice, 2013 London, 71–90.

legal setting, albeit an example where human decision making is not supplanted, but merely enhanced by the AI.

C.   COMPUTATIONAL CREATIVITY

In this section, we challenge the assumption that AI is almost by definition incapable of creativity. The question of robot creativity is by no means new. Indeed, the idea to use a combination of mechanical and random processes to generate new works of art precedes the modern computer by centuries. Popular in the 18th and 19th century were musical compositions developed with the aid of dices. Johann Nikolaus Simrock's famous algorithm for music creation was published in 1792 (and at some point attributed to Mozart). This approach was based on a dice game capable of producing more than 45 Trillion different waltzes, and, so the publisher proudly proclaims, «without understanding anything about music or composition».[15]

Almost as soon as modern computers became available to people outside universities or the military, artists began to explore their potential for creativity. Nicolas Schöffer's CYSP 1 (Cybernetic Spatiodynamic Sculpture) from 1956 showed how the then resurgent «kinetic art» could put modern machines at the heart of their endeavor.[16] Schöffer's interactive work comprised several sensors and electronic components that interacted with observers to produce different kinds of movements. Nam June Paik and Shuya Abe's Robot K-456 from 1964 used robot generated art to thematize issues of remote control and freedom, while Edward Ihnatowicz's *Senster* was maybe the first instance where the issue of robotic behavioral autonomy was coming to the fore. The robot in this work was assigned one of several possible personalities, which then responded to changing situations on their own.[17].

---

[15]   NIERHAUS, Algorithmic composition: paradigms of automated music generation. 2009 Berlin p. 36.

[16]   SCHOEFFER, La ville cyberne'tique., Paris 1969; see also KAC, Foundation and development of robotic art. Art Journal 1997 56(3) 60–67.

[17]   REICHHARD, Robots: fact, fiction + prediction. 1978 London p. 56.

Examples of computer-generated poetry and prose emerged soon afterwards. RACTER was the first book written entirely by a computer in 1985[18]. In poetry, probabilistic and evolutionary generation of word chains can show remarkable results[19], though we also find more traditional rule based systems that try to model the technique of writing poetry more faithfully[20].

Of all the arts, music was arguably the field that was the fastest, and most profoundly changed, by computer creativity[21]. Reviving the tradition of composition by dice as a probabilistic algorithm described above, computer-generated music quickly became the main domain for computer generated works. In a typical application of the time – and of a form that we will revisit below, Iannis Xenakis's «stochastic music» generated a multitude of possible works using the «Monte Carlo» method; knowledge of principles of composition was then used to select those that were acceptable[22].

Since these early beginnings, machine creativity research has made considerable progress across all art forms[23]. Crucially, the quest for the creative robots also helped to formulate and make precise new conceptual questions about the nature of art and (human) creativity. Margaret Boden in particular started, and for a long time dominated, a

---

[18] RACTER, The policeman's beard is half-constructed: computer prose and poetry, 1985 New York; see also FUNKHOUSER, Prehistoric digital poetry: an archaeology of forms 1959–1995. 2007 Tuscaloosa.

[19] JIANG/ZHOU (2008) Generating Chinese couplets using a statistical MT approach. In: Proceedings of the 22nd international conference on computational linguistics, vol. 1, pp 377–384. for the recurrent discussion see MANURUNG/THOMPSON (2012) Using genetic algorithms to create meaningful poetic text. J Exp Theor Artif Intell 24(1):43–64.

[20] GERVAS (2001) An expert system for the composition of formal Spanish poetry. Knowledge Based Syst 14(3):181–188; OLIVEIRA (2012) PoeTryMe: a versatile platform for poetry generation. In: Proceedings of the ECAI 2012 workshop on computational creativity, concept invention, and general intelligence.

[21] For an overview see ROWE (1993) Interactive music systems: machine listening and composing. MIT Press, Cambridge; ROWE (2001) Machine musicianship. MIT Press, Cambridge.

[22] XENAKIS , Formalized music: thought and mathematics in composition (Harmonologia series no.6) 2001 Pendragon Press, Hillsdale.

[23] COLTON/WIGGINS, Computational creativity: The final frontier?. Ecai 2012 Vol. 12, 21–26.

discussion that raised deep conceptual issues around the connection between randomness and creativity[24]. This type of analysis also enabled the development of rigorous methods to compare the creative ability of different computer systems and classificatory schemes for them, from the most mundane such as automatic spell checking to the most sophisticated such as a robot doing jazz improvisations[25].

From this short historical overview, we can learn two things. First, it is far from obvious that AIs are inherently incapable of «a form of» creativity. Second, in the past computational creativity typically involved generating a random element, which ensured that the result was surprising and unpredicted/unpredictable. While more recent efforts aim at minimizing the reliance on mere random generators, randomness continues to play an important role in computational creativity research.

D.   CREATIVE LEGAL AI

Randomness and Justice

We can now start asking the question that is at the centre of this paper: Are those forms of creativity that we consider beneficial in a legal setting in principle beyond the capabilities of AI tools, or can we find ways of building legal Tech that supports rather than suppresses creative legal problem solving?

So far we saw how randomness plays an important role in computational creativity research – but can there be a place for random decision making in the justice system? At first, this seems absurd – and indeed existing approaches to legal AI do not employ random generators.

---

[24]   See e.g. BODEN, Pre´cis of the creative mind: myths and mechanisms. Behav Brain Sci 1994 17(03):519–531; BODEN, Creativity and artificial intelligence. Artif Intelligence 1998 103(1):347–356; BODEN, Computer models of creativity. AI Mag 2009 30(3):23, for a more recent discussion see MCCORMACK/ D'INVERNO, Computers and creativity. 2012 Springer, Berlin.

[25]   COLTON/CHARNLEY/PEASE A. Computational Creativity Theory: The FACE and IDEA Descriptive Models. ICCC 2011 pp. 90–95. WIGGINS, A preliminary framework for description, analysis and comparison of creative systems. Knowledge-Based Systems 2006 19(7):449–458; JORDANOUS, A standardised procedure for evaluating creative systems: computational creativity evaluation based on what it is to be creative. Cogn Comput 2012 4(3):246–279.

Before addressing this issue directly, one should note that randomness is not only used in the computational generation of art as a way to achieve (the allusion of) creativity, but even more so to applications that are maybe nearer to legal reasoning, in particular computer games. One of the reasons AlphaGo succeeded in beating a human Go champion was by producing a move that was very unlikely to be made by a human – the famous move 37 was calculated by AlphaGo to have a 1/1000 chance to be made by a human, and AlphaGo utilizes Monte Carlo tree models that in turn incorporate an element of randomness. Games in turn, as a rule based, adversarial activity, have for some time been used in legal theory as an appropriate analogy or model of the trial as a rule based and adversarial encounter.[26] The success of AlphaGo has been used in turn to argue for the likely seismic shift that the introduction of this and similar technologies will bring to the justice system, and AlphaGo is also referenced in the report of the Law Society of England and Wales on Algorithm in the Justice System.[27] However, none of these writers focus on the creativity of AlphaGo, or the random choice between branches of the trees that enable it.

The question therefore becomes if randomness can also play a role in legal AI. At first sight, this seems a truly ridiculous claim: what could be more capricious than the throwing of a coin to decide a person's fate? As strange as it sounds however, random decision making processes have a long history in the law, as Neil Duxbury documented.[28] One well-known example is of course the process of jury selection in common law systems. The underlying rationale is not just to spread the burden of service fairly, but also a

---

[26] HART, The concept of law, 1994 London p.141 and 144–145; DWORKIN, Taking Rights Seriously, 1977 New York p. 24; a detailed but critical discussion of the analogy is in ATRIA, On law and legal reasoning, 2002 London pp. 8–48.

[27] LAW SOCIETY of England and Wales, AI: Artificial intelligence and the legal profession, 2018 https://www.lawsociety.org.uk/support-services/research-trends/horizon-scanning/artificial-intelligence/ (last accesses 11/11.19)See also e.g. HEATON, Artificial Intelligence and Machine Learning: The future of Legal Services , 2017 https://www.justcosts.com/10-news-events/livenews/486-artificial-intelligence-law (last accessed 10.11.19); TAYLOR/OSAFO, Artificial intelligence in the courtroom, https://www.lawgazette.co.uk/practice-points/artificial-intelligence-in-the-courtroom-/5065545.article 2018 (last accessed 10.11.19).

[28] DUXBURY, Random justice, 2002 Oxford.

normative commitment to a fair trial, as the Lord Justice Clerk argued in *McCadden v. H. M. Advocate*[29]:

> *«The existing system of empanelling a jury from a list of assize is so broadly based that it provides a wide opportunity of a mix which is liable to level itself out.»*

Random selection of viewpoints from across the population protects the trial not just from coercion and bribery, but also protects the accused from being subjected to the bias of a single decision maker. We could build on this argument and reason that in a pluralist society, a degree of divergence between judges is not only acceptable, it is desirable. While from the perspective of the accused, the random allocation of a stricter vs a more liberal judge may seem unfair, from the societal perspective and the trial as a communicative act, it could be seen as a normative affirmation of a pluralist society. We want to live in the type of society where diverging conceptions of the good life co-exist, and this includes the risk to be in a more or less random (or serendipitous) way exposed to and judged by people with a range of normative convictions.

This line of reasoning could also be stated as an evolutionary argument. We noted above that machine learning approaches to legal AI in particular risk the ossification of the justice system, as their underlying assumption is that the past strictly determines the future and makes it predictable. While predictability is of course one of the desiderate of a justice system, if pushed to the extreme it risks inflexibility and ultimately decay. In a similar vein, mankind survived pandemics such as the Spanish flue because not everybody's body followed the same script, and random mutations, mostly harmful but sometimes beneficial, created eventually a resistant pool. To prevent machine learning systems to entrench the past and as a consequence deliver sub-optimal solutions once the environment changes, we may want the equivalent of random mutations that can temporarily unsettle the system, only for it to achieve through selection a new, stable equilibrium. Introducing a random element in the AI decision making process could achieve such a more dynamic and adaptable system. In the next section, we will look at one specific program that took this idea to its heart

---

[29]   McCADDEN v. H. M. Advocate 1985 J.C. 98.

Un(usual) Suspects

This part of the paper is based on a project that the author led as principal investigator and a number of papers that it generated. The discussion will remain informal, and for the technical details and implementation the reader is referred to these studies.[30]

In the late 80s, the discovery of several high-profile miscarriages of justice shook the British legal system. The Runciman Commission was established in 1991 in response, to examine the effectiveness of the criminal justice system in all its aspects. In its wake, a significant body of knowledge has been produced analysing the potential for errors in criminal investigations and prosecutions.[31] One clear pattern was the problem of premature case theories: Instead of withholding judgement and establishing «bare and neutral» facts first, police officers decided at a very early stage of an investigation on the most likely suspects, and from then on investigate *against* them. In the words of David Dixon:

> «If any factor in investigative practice had to be nominated as most responsible for leading to miscarriages of justice, it would have to be the tendency for investigators to commit themselves to belief in a suspects guilt in a way that blinds them to other possibilities».[32]

In our context, we could say that the officers behaved just like the current generation of machine learning programs, with all the problems that this entails. Their training and experience led them quickly to identify patterns, and with that the assumption that every new situation had to fall within one of these patterns or case theories. The use of such «case theories» is arguably inevitable.[33] The problem is therefore not the fact that case theories are used at all, but rather the restricted scope of alternatives that is considered. As Greer argues:

---

30 KEPPENS/SCHAFER, Knowledge based crime scenario modelling. Expert Systems with Applications 2006 30.2 pp. 203–222; SCHAFER/KEPPENS, Legal LEGO: Model Based Computer Assisted Teaching in Evidence Courses. Journal of Information, Law and Technology 2007 https://warwick.ac.uk/fac/soc/law/elj/jilt/2007_1/schafer_keppens/.

31 See e.g. WALKER, Justice in error, 1993 London.

32 DIXON, Police Investigative procedures. In: C. Walker (ed.) Miscarriages of justice, 1999 London.

33 See MCCONVILLE/SANDERS/LENG, The Case for the Prosecution, 1991 London.

> *«...no criminal justice system could work without them. The dangers stem instead from the highly charged atmosphere surrounding an investigation, the haste with which the theory has been formed and the tenacity with which the police have clung to their original view in spite of strong countervailing evidence».[34]*

Irving and Dunningham suggested a number of possible solutions to this problem.[35] They argue for the need to improve officers reasoning and decision-making by challenging the «common sense» about criminals and crimes and the detective craft's «working rules about causation, about suspicion and guilt, about patterns of behaviour and behavioural signatures.»

We can now link these ideas with our discussion above. One of the reasons AlphaGo was able to beat a human player was that at a crucial point of the game, it made an unexpected move, one that defied its opponent's «case theory» about the normal «patterns of behaviour» of a Go player in a specific situation. Creativity in this case was a disruption of established patterns and a deviation from the «common sense» assumptions that «everybody just knew».

What we suggested therefore was to use AI as a creative interruption of the decision making process. But just as AlphaGo could not win by playing just a random move, but a move constrained not only by the rules of the game, but also a parallel assessment of the strength of the move, so our system should not just produce bizarre outputs to disrupt the thought pattern of its user, but suggestions which while extremely unlikely might just be true.

Our system addressed this balance of creativity and constraint by combining a «backward chaining» abductivist model of reasoning with a «forward chaining» model that is based on the idea of indirect proof. Presented with all the evidence collected at a given point, the *Assumption Based Truth Maintenance System* or ATMS first develops a range of alternative scenarios that are all consistent with the facts as established so far, as unlikely and counterintuitive as they may be. This encourages creative speculation

---

[34]   GREER, Miscarriages of criminal justice reconsidered. Modern Law Review 1994 58 p. 71.

[35]   IRVING/DUNNINGHAM, Human Factors in the quality control of CID investigations and a brief review of relevant police training. Royal Commission on Criminal Justice Research Studies 1993 vol 23.

and the questioning of common-sense assumptions. In the second step, the ATMS indicates those (as yet uncollected) pieces of evidence that could differentiate between these different theories through a process of falsification – a lose equivalent of AlphaGo playing through a possible game to its end, to determine future moves. The hope was to foster critical thinking away from the prevailing «inductivist» ethos.

A quick illustration explains what this means. Imagine a police officer arriving at a potential scene of crime. He notices a person, identified to him as the homeowner, with injuries consistent to blows with a blunt instrument. The window of the room is broken, and outside a step ladder is found. The officer now has to make a decision: is this a likely crime scene, are further (costly) investigations necessary? Should all known burglars in the area be rounded up for interrogation?

Our officer, to make sense of the scenario as described above, will arrange (probably pre-linguistically) the features of the scene in coherent whole or Gestalt.[36] In the same way as we cannot but see a forest when there are many trees, s/he will at a very early stage «see» a scenario in which a burglar entered the house with the ladder through the window, was approached by the homeowner, and killed him with a blunt instrument. This whole «picture» or «story» is influenced by typical associations, e.g. that of «burglar» with «ladder».

What our system proposed to do then was not so much emulating or improving the process by which individual aspects of a scenario are combined, but rather to facilitate for the officer to perform a «Gestaltswitch», to see the same individual aspects (scenario fragments) in a re-arranged way that gives rise to another whole – to one or more alternative stories. In our example, scenario fragments are the broken window, the dead body, the wounds on the body and the ladder, and the preferred hypothesis is one of a burglary gone bad. The system reminded the officer for instance that on the basis of this evidence, it is also (though not necessarily equally) possible that the dead person did some Do-It-Yourself in his flat, suffered a heart attack while on the ladder, fell from the ladder to the ground, and the ladder fell through the window. This involves several «switches»: the ability to see the ground as a «blunt instrument», the window as an

---

[36] BOLTE/GOSCHKE, Intuition in the context of object perception: Intuitive gestalt judgments rest on the unconscious activation of semantic representations. Cognition. 2008 108(3) pp. 608–16.

opening that let things go out as well as in, and the entire scenario from one of crime to one of domestic accident. It should then identify those pieces of evidence that could rule out one of the theories, e.g. fingerprints (or lack of them) on the ladder and on the hammer that is considered the murder weapon, or fibres from clothing on the broken glass. We note that our system was not based on machine learning but on explicit and symbolic knowledge representation, «good old fashioned AI» with other words. For legal AI, this had the possible advantage that generation of explanations was comparatively trivial. The disadvantage was however that these alternative scenarios had to be elicited «manually» in the knowledge acquisition stage, which meant that despite the development of a successful prototype, it was not possible to develop it to usable size within the cost and time constraints, a fate typical for legal AI during that time. For our purposes however it is the underlying philosophy, successfully tested, that is relevant for the question the paper set out to answer.

I'm giving here only a very abbreviated account of the formal mechanism which we developed to distinguish murder from accidental death and suicide, using the database of knowledge of homicide and suicide scenarios that we generated from interviews with police officers.

We used a model-based reasoning technique, derived from the compositional modelling paradigm,[37] to automatically generate crime scenarios from the available evidence consistent with the approach to reasoning about evidence still dominant today[38], we employed abductive reasoning. That is, crime scenarios were modelled as the causes of evidence and they are inferred based on the evidence they may have produced.

The goal of the decision support system (DSS) was to find the set of hypotheses that follow from scenarios that support the entire set of available evidence.

---

[37]   FALKENHAINER/FORBUS, Compositional modelling: finding the right model for the job. Artificial Intelligence 1991 51 p. 95–143.

[38]   For two recent papers with further references, see TUZET, Abduction, IBE and standards of proof. The International Journal of Evidence & Proof. 2019 23(1–2) pp. 114–20 and VERHEIJ et al, Arguments, scenarios and probabilities: connections between three normative frameworks for evidential reasoning. Law, Probability and Risk. 2015 15(1) pp. 35–70.

The central component of the system architecture is an assumption-based truth maintenance system (ATMS). An ATMS is an inference engine that enables a problem solver to reason about multiple possible worlds or situations.[39] Each possible world describes a specific set of circumstances, a crime scenario in this particular application, under which certain events and states are true and other events and states are false. What is true in one possible world may be false in another. The task of the ATMS is to maintain what is true in each possible world.

In an ATMS, each piece of information of relevance to the problem solver is stored as a *node*. Some pieces of information are not known to be true and cannot be inferred from other pieces of information.

In the ATMS, they are represented by a special type of node, called *assumption*. Inferences between pieces of information are maintained within the ATMS as *justifications*. The plausibility of these assumptions can then be determined through verification of nodes inferred from them. Continuing our «burglary gone bad» example (and bearing in mind that we are in an abductive environment), this could be understood as: the investigative hypothesis that it was an accident is justified by the presence of the victim's fingerprints on the ladder and the absence of fingerprints of a third party.

An ATMS can also store justifications called *nogoods* that lead to an inconsistency, i.e. justifications of the form «Fact A, B and C imply D and -D». The latter nogood implies that at least one of the statements in the antecedents must be false. This accounts for the «critical» ability of our system, or the analogue of natural selection in our epidemic example above. Presented with two conflicting hypotheses, it will direct its user to collect evidence in a way that one of them is «justified» by a nogood, that is undefeated evidence that is incompatible with the investigative theory. In our «burglary» example, the presence of large amounts of broken glass outside the house, implying that the window was broken from the inside, would be consistent with the accident hypothesis, but might lead to further hypotheses – the burglar broke the window in order to escape the

---

39   For technical details see de KLEER, An assumption-based TMS, Artificial Intelligence 1986 28 pp 127–162.

house, or the home-owner fell head-first from the broken window (several hypotheses might include this node).

In this approach, it is presumed that the states and events constituting a scenario can be represented as predicates or relations. Naturally, states and events do not exist in isolation from one another. Certain states or events may be consequences of combinations of other states and events. For example, if a person is being assaulted and capable of self-defence, then (s)he will probably engage in some form of defensive action. Such knowledge is represented by *scenario fragments*. To illustrate the concept of scenario fragment, consider this example which we first give in a semi-formal notation, then in a verbal transcription:

if{doctor(D),person(B) brain_trauma(B)} assuming {cause-of-death(B,brain_traumae), correct-diagnosis(D,cause-of-death(B))}   then   {medical-report(D,cause-of-death(B), brain_trauma)}

This scenario fragment states the following: given a person B, a doctor D and the fact that B suffered a brain trauma; and assuming that the cause of death of B is the brain trauma and that D makes a correct diagnosis of that cause of death; then a medical report must exist, written by D, stating that the cause of death of B is a brain trauma.

This scenario fragment can fulfil a dual purpose in our application. First and somewhat trivially, it ensures that the absence of a medical certificate is a reason to doubt that B died of a brain trauma – allowing in our example for the alternative explanation of «heart attack». Secondly, it means that a medical report is not in any way different from say DNA evidence or a fingerprint: all are facts that are *explained by* certain assumptions. The medical report is *an observable consequence of* a state of affairs. In police investigations, it is quite often the presence of an official document, a medical report for instance or a confession, that blocks the investigator from seeing alternative explanations of the evidence. In our approach, this document itself is linked in alternative scenario fragments with alternative explanations. In addition to the example described, a medical report might also result from a mistake by the doctor, or indeed his attempts to cover up his own crime.

The goal of the scenario instantiator is to construct a space of possible crime scenarios by instantiating the knowledge base of scenario fragments and inconsistencies into an ATMS. The algorithm which we developed expanded on an existing composition modelling algorithm devised for the automated construction of ecological models.

We can draw here a rough parallel to the computational creativity systems mentioned above. A program that writes poetry will break down text in fragments (e.g. words, or phrases). It will then learn associations between these fragments from the existing corpus of work. It will then rearrange these fragments in ways that are a) novel, because of a random element, and b) still adhere to a set of pre-defined rules, e.g. the rules of grammar. In the same way, we break up crime scene narratives into fragments, and then rearrange them in ways that are novel (and for this reason alone also unlikely to be true) and still adhere to e.g. the causal constraints of the physics of our universe (our victim can't e.g. fall through the ground).

The ATMS constructed by the algorithm contains a space of all scenarios that can be constructed with the knowledge base and that produce the given set of evidence E. This then allows to answer three types of query. The approach taken here involves translating queries into formal ATMS nodes and justifications, thus enabling the existing ATMS label propagation to answer the queries of interest:

1) Which hypotheses are supported by the available evidence?

Every hypothesis that follows from a plausible scenario is supported by the available evidence.

In our «burglary gone bad» example, there are two environments that support the available evidence. According to one, the houseowner B suffered from a hit with a blunt instrument on his head, according to the other, he suffered heart failure, fell from the ladder and hit his head on the ground.

E_1={high-cholesterol(b),accidental-coronary-blood-vessel-rupture(b),cause-of-death(b,heart-attack), correct-diagnosis(d,cause-of-death(b))}

E_2={hammer-attack(p,b)},brain-trauma-due-to-attack(b),cause-of-death(b,brain-trauma), (correct-diagnosis(d,cause-of-death(b))}

In the possible world described by environment E_1, accident(b) is true and in the one described by E_2 homicide(b) is true. Therefore, it follows that both hypotheses are supported by the available evidence.

2) What additional pieces of evidence can be found if a certain scenario/hypothesis is true?

All the states and events, including pieces of evidence, that are logical consequence states and events in plausible scenarios are generated in the forward chaining phase of the algorithm. Therefore, the initial state of the ATMS will contain nodes representing pieces of evidence that are produced in certain scenarios but were not collected in E. A piece of evidence *e* can be found under a given hypothesis *h* if a possible world exists that supports both the evidence and the hypothesis. Continuing with the ongoing example a piece of evidence *e* that consists of a medical report documenting high cholesterol in b, medical-report(d, high-cholesterol(b)), is generated under the environment:

E_3={high-cholesterol(b),accidental-coronary-blood-vessel-rupture(b),cause-of-death (b,heart-attack)},correct-diagnosis(d,cause-of-death(b)) test(d,test-cholesterol (b))}

This means simply that under the hypothesis of accident, this third piece of evidence, a report, may be found.

3. What pieces or sets of additional evidence can differentiate between two hypotheses?

Let h_1 and h_2 be two hypotheses, then any set of pieces of evidence E that can be found if h_1 is true, but are inconsistent with h_2, can differentiate between the two hypotheses. For example, it follows from the above discussion that the piece of evidence medical-report(d, heart attack(b)) may help to differentiate between the two hypotheses, accident(b) and homicide(b). This information suggests to a police officer or a prosecution lawyer examining the case that ordering tests for symptoms of heart attack would be useful.

The man who wasn't there

In the account so far, we also encountered if obliquely one key element of human creativity. In the previous section, we heard that the system constructs alternative scenarios,

or «possible worlds». One key aspect of human creativity is our ability to think about the world not just as it is, but also as it could have been (fiction and fictional writing), might be (planning and invention) and should be (deontic reasoning). In all these cases, we envisage universes that are in varying degrees dissimilar to ours, while sharing sufficient traits to make them intelligible. Crucially, we are aware of this ability of ours, and can self-reflexively reason about it. This also allows us to use fiction and make belief strategically for argumentative purposes. It is this conscious reasoning about fictional worlds that we turn to in the final part of this paper, as a further addition to the «disruptive creativity» that we encountered in the previous section.

A particularly nice example of this ability can be found in the poem «Antigonish» by William Hughes Mearns:

Yesterday, upon the stair,

I met a man who wasn't there!

He wasn't there again today,

Oh how I wish he'd go away!

We can see at the same time that the poem pokes fun at our ability to reason about things that do not exist, and still, we can «see» the man who wasn't there in our mental eye. And indeed, we can ask legal questions about him – for instance, is his persistent «not being there, on the stair» and his refusal to go away indicative of a criminal offence under the Trespass Scotland Act 1865? This seems silly, and yet it could be part of a sound legal argument – for instance as part of an alibi. Inventing non-existent persons does play a crucial, and quintessentially creative, role in legal reasoning.

In developing alternative scenarios consistent with the evidence, the ATMS performs some of the scrutiny a good defence solicitor would subject the prosecution case to. A defence solicitor has broadly speaking two strategies available to him. First, she can question the factual correctness, or the legal admissibility of evidence presented by the prosecution. Second, she can accept the evidence at face value and argue that alternative explanations for their presence are possible that do not incriminate his client. We are concerned here primarily with this second strategy.

The defence has in fact again two different ways how to play this game. The first can be dubbed the «Perry Mason Stratagem». Like the fictitious advocate, the defence can pursue its own investigation and «point to the real culprit». In Scots law, this is known as the special defence of incrimination. This strategy has a number of psychological and legal advantages. The same reason that makes it the solution of choice for crime writers also works well with juries: no loose ends are left, and the crime is avenged. Procedurally, it allows the defence to submit also other pieces of evidence. This corresponds to the «forward chaining» aspect of our ATMS: The party named by the defence will have interacted causally with the crime scene. This will have created evidence which can strengthen the defence case. This allows introduction of additional «suspect specific» evidence (such as alibi) evidence about other people, which otherwise might be ruled out as irrelevant. The defence of course need not prove the guilt of the other party, it only needs to establish it as a plausible alternative. This way of dealing with the evidence mirrors particularly closely the working of our ATMS.

In reality however, this strategy faces considerable obstacles. Defence solicitors normally don't have the resources, time or training to engage in investigative activity of their own. In developing alternative accounts of the evidence, they will therefore typically settle for something less. They will argue that a hypothetical «someone» might have been the real culprit, and that this alternative is not ruled out by the evidence.

A (fictitious) example of a cross examination can illustrate this point:

Police officer: *When I heard that scream, I ran up the stairs to the room where the victim was. There I found the accused, with a bloodied knife in his hand.*

Defence solicitor: *Is it not true that it took you more than three minutes to find the right room? In this time, the real murderer could have escaped through the window, couldn't he? So that when my client arrived at the scene shortly before you, he found the body of his wife lying there, and took the knife in an attempt to guard himself against whoever killed her?*

At this point, the prosecution might accept this alternative «for the sake of the argument» when re-examining this witness:

Prosecution: According to your description of the scene of crime, is it not true that this «mysterious Mister X» would have had to jump down from the third floor, and then run across a busy street with blood all over his clothes?

The argument proposed here is that the use of «someone» in this strategy is not best represented as a classical existential quantifier that picks a specific object from the universe of discourse. This is particularly important to the prosecution, who may fear damage to their case from a statement that carries an ontological commitment to a «Mister X». The introduction of Mr X however requires a considerable degree of creativity – quite literally, a new person is created or at least outlined, like in a first sketch.

What happens in this example is categorically different from a defence of incrimination. In an incrimination defence, defence and prosecution disagree about the state of the world. In the dialogue by contrast, what is debated is not so much a state of affairs. Rather, what is exchanged is «discourse information». We can understand the defence argument as a meta-statement about the evidence: «It does not follow logically from the case as presented that my client is the person who committed the crime». Conversely, the prosecution is stating: «the alternative of the defence is inconsistent with the evidence».

Evidentiary legal argumentation typically involves a combination of exchanges about «discourse information» (the rules of admissibility might be seen as a particularly prominent example) and discussion about the world. For this reason alone, it might be worthwhile to amend the formal system in a way that both types of reasoning can be represented *in their distinctiveness*. To conclude this passage, and to show the pervasiveness of the issue discussed here, another example is given that at first sight is unrelated to the introduction of hypothetical suspects. It will turn out, however, that solutions discussed in natural language analysis for just this problem also provide a natural solution to the issue at hand. Consider the following scenario.[40] Several witnesses claim to have seen «someone» suspicious running away at the time of the crime.

---

[40] Modelled on an example given in GROENENDIJK,/STOCKHOF/VELTMAN. Coreference and modality. In S.Lappin (ed.), The handbook of contemporary semantic theory , 1996 Oxford, pp. 179–213 at p. 3.

Witness 1: *Yeah, he was a black guy, with a red cap, running to his car. It was a Volvo.*

Witness 2 (wife of 1): *It was not a man, it was a girl, and she wasn't wearing a red cap but a red scarf. Also she wasn't really black, more tanned. Oh, and it was a Vectra.*

This exchange is intelligible only if witness 2 has reasons to believe that the indefinite «a black guy» of statement 1 refers to an entity which witness 2 thinks to be a woman with a red scarf. None of the original attributes of this entity seem to be agreed upon. The police officer needs to be cautious: it might turn out after all that they saw two different people and our ATMS in turn needs to be able to cope with partially conflicting descriptions of what might or might not be the same object.

Possible world semantics and similar approaches have created a plethora of formalisms that are capable of formalising the reasoning about such a possible, but possibly not real, Mr X. Some of them have been developed explicitly to give formal, and hence computational, accounts of fiction writing – which also includes the analysis of «legal fictions» as an example of legal creativity.[41] Here we will look into one approach in particular, partly because of the imagery and metaphors it uses as motivation are particularly appealing for a discussion of both creativity and law, partly because it stays particularly close to the linguistic structure of the text and with that the computational creativity approaches for fictional writing discussed above.

The linguistic phenomenon described in the last example above has been at the centre of research in formal linguistics for quite some time and produced copious observations and theories.[42] The question for linguists has been, how is it possible for a pronoun to be bound without being in the syntactic scope of its antecedent? The obvious answer is that the semantic scope of an expression may reach beyond its syntactic scope. Formalizing this insight was less straightforward, however. The first theory to implement this idea was Discourse Representation Theory (DRT).[43]

---

[41] For a discussion of a number of such formal theories and their applicability to legal reasoning see SCHAFER/CORNWELL J. Law's Fictions, Legal Fictions and Copyright Law. In Del Mar (ed), Legal Fictions in Theory and Practice, 2015 Heidelberg pp. 175–195.

[42] See e.g. GEURTS, Donkey business. Linguistics and philosophy 2002 25 pp129–156 or GAWRON/PETERS, Anaphora and Quantification in Situation Semantics . CSLI, 1990 Stanford.

[43] KAMP/REYLE, From discourse to logic . 1993 Dordrecht

We want to draw the attention here to the «Dutch school» which, inspired by DRT, developed a family of formal systems which are particularly suited to address the issue not only of anaphoric binding, but more generally the «unspecific» use of quantifiers discussed here. They also allow a meta-linguistic treatment of the kind of discourse information that we mentioned above, allowing for a more unified theory of evidentiary legal reasoning. Update semantics, dynamic logic and data semantics are varieties of this approach. Technically, they are extensions of standard model theoretical (Kripke) semantics, which should make their incorporation into our ATMS straightforward. However, they radically reinterpret the meaning that they give to Kripke models.

Frank Veltman, in his influential paper on update semantics,[44] coined the slogan that summarized the unifying assumption of these different formalisms: «You know the meaning of a sentence if you know the change it brings about in the information state of anyone who accepts the news conveyed by it». Meaning thus is not (just) a relation between sentences and the world, but the potential to change a context, where contexts are identified with information states. This connects the idea with our crime investigation scenario, which can be understood as a systematic attempt to change information states until the evidence is sufficient for a charge.

To deal with anaphoric binding, these systems introduced as part of the discourse information a new type of objects, commonly called «pegs». «Pegs» are intermediate discourse entities. They are connected to the variables in the object language on the one hand, and with the objects in the models (our scenarios) on the other. The metaphor here is appealing: A Peg, like a clothing peg, is something on which we can temporarily drape all sorts of clothing – a woman's jacket and hat now, a man's coat and wig thereafter. The peg remains the same, but its only role is to act as a scaffold for whatever story we want to tell, it's its «accessories» that change. The language of the theatre, intentionally chosen by Veltman, brings in yet another aspect of our ability for creativity.

---

44   VELTMAN, Defaults in update semantics, Journal of Philosophical Logic 1996 25 pp. 221–261.

An information state as a whole consists not only of discourse information (in the form of a referent system), but also of information about the world, and of a link between the two types of information. An information state in this approach is regarded as a set of possibilities, an assumption that resonates well with the concept of «scenario space» in our ATMS, only that now the very meaning of pieces of evidence is seen to be determined by the possible scenarios in which it appears. Each possibility consists of a referent system; a possible world; and an assignment function which assigns some object from the domain of that world to each of the pegs present in the referent system. Information growth can take place in two ways: the referent system may be extended with new pegs, (re)associating variables with them and assigning them suitable objects; and/or certain possible assignments or possible worlds may be eliminated. This last possibility corresponds to the «forward chaining phase» of our ATMS.

The notion of update and exchange of information about the values of variables has some intuitive appeal. As the examples above show, it comes naturally to us to talk about «indefinite» objects, objects of which we have only partial knowledge. We talk about these objects (the mysterious Mr X), they are ascribed (possibly conflicting) properties and people are informed about their existence. Crucially though, these «objects» must not be understood as classical objects. From the perspective of agents with partial information the ultimate identity of such objects may be left unresolved. Nonetheless, they can be topics of information exchange. Fred Landman has developed the most explicit theory of such partial objects to date,[45] even though the idea traces back to Karttunen's seminal paper.[46] According to Landman, they are things that don't have properties, but to which properties can nonetheless be ascribed and, similarly, things that don't have identity conditions, but that have identity conditions ascribed to them:

> *«the essence of partial information is that it cannot justify certain distinctions, and the decision about the identity of certain pegs is a prime example of that».*

---

[45]   LANDMAN. Towards a theory of information. The status of partial objects in semantics. 1986 Dordrecht.

[46]   KARTTUNEN. Discourse referents. Syntax and Semantics, 1976 7 pp. 363–385.

The proposal that we are making here is that the assumption of purely formal objects or pegs allows development of theories that explain the peculiar features of legal discourse noted above, and illuminate a key aspect of «benevolent» legal creativity. While the defense of incrimination introduces a classical object into the discourse, the more abstract speculation about «possible other parties» introduces (merely) pegs. The prosecution, as we have seen, can in turn refer to these objects – without incurring a commitment to the existence of any corresponding person in the universe of discourse. Similarly, prosecution and defense stories can accommodate witness accounts of «the same» «person» even though this «person» is ascribed incompatible attributes. Coping with this degree of ambiguity is yet another cognitive ability closely associated with creativity – every reader will imagine the personae dramatis in a play or book differently, and yet we can communicate about them.

E.   CONCLUSION

We began this paper with a challenge: can we identify benevolent examples of legal creativity, and are these then beyond the ken of computers past, present and maybe in the near future? To do so, we identified some of the dangers that are expressed by the machine metaphor of robotic legal decision making on the one hand, real problems with machine learning approaches on the other. Current legal AI is «Dworkinian» in the sense that it assumes that as long as sufficient data can be processed, the one right answer can be predicted/derived. Creativity is at best a crutch we use as long as we fall short of this ideal, at worst an illegitimate deviation from the chain novel.

We then argued that computational creativity research by contrast emphasizes the benefits of an element of «controlled randomness», that is randomness that still adheres to certain constraints. We then showed how a specific approach could leverage this idea to counteract exactly those shortcomings of the justice system where police, jurors and courts are all too «machine like» in their behaviour.

This indicates that the common and popular juxtaposition of «creative human lawyers» vs «deterministic, predictable machine» is a false dichotomy that captures neither human nor machine. Rather than a grand division of labour, what we advocate and demonstrated to be possible is a much more nuanced approach that asks for specific

expressions of creativity, and specific problems of the justice system, if computational creativity can help to inspire a solution. ÉCLAIR, empathetic and creative legal AI, is proposed here as an antidote to an understanding of legal technology that is shared by both its proponents and critics, with ramifications not just for the way in which we should build legal technology, but, to come back to the beginning of the paper, also how we should train the lawyers of the future to make the best, and most just, use of these technologies.