# CRi

# Computer Law Review International

A Journal of Information Law and Technology

**cr-international.com**

**otto**schmidt

# CRi Computer Law Review International

A Journal of Information Law and Technology

# Articles

*Ramak Molavi Vasse'i*

# The Ethical Guidelines for Trustworthy AI – A Procrastination of Effective Law Enforcement

## Weaknesses of ethical principles in general and the EU's approach in particular

*In the august issue of the CRi, Nathalie Smuha, the coordinator of the work of the High-Level Expert Group on AI, outlined the approach and considerations leading to the "The EU Approach to Ethics Guidelines for Trustworthy Artificial Intelligence".This paper provides a critical assessment of the Ethical Guidelines of the EU Commission and points out why a law enforcement focused approach must be the essential next step towards a beneficial and humane development of AI. Questioning the diversity of the Commssion's High Level Expert Group on Artificial Intelligence, the dangers of ethics shopping are exposed as well as the UN Universal Declaration of Human Rights explored as already well established alternative reference framework for AI. Having exposed the need for effective red lines, not only the hidden social and ecological cost are assessed, but also the risk of "buying-out" research and other ethical issues neglected in the Ethics Guidelines for Trustworthy Artificial Intelligence. Finally, three key weaknesses concerning the crucial translation of ethical principles into practice (enforcement) are highlighted.*

## I. Ethical Guidelines for Trustworthy AI

1 The AI HLG[1] was tasked to draft two deliverables: AI Ethics Guidelines, published in April 2019 and Policy and Investment Recommendations. During the first AI Assembly on the 26 June 2019 the second deliverable was published: the Policy and Investment Recommendations for Trustworthy Artificial Intelligence.

2 The Ethics Guidelines for Trustworthy AI provide an assessment list that operationalizes the key requirements and offers guidance to implement them in practice. As a next step, the list is currently undergoing a piloting process. The stakeholders are invited to test the assessment list and provide practical feedback on how it can be improved. This phase began with the publication end of June and will run until December 2019. Based on the feedback, the High-Level Expert Group on AI will propose a revised version of the assessment list to the European Commission. This step is planned for early 2020. Over 450 stakeholders have already registered to this special piloting process.

## 1. The Group Structure of the AI High Level Group

In addition to internal know how and assessment, the input of all relevant stakeholders is key for a good regulation. The set up of the High Level Expert Group on Artificial Intelligence (AI HLG) is considered to be very diverse in terms of disciplines and interests.[2] 3

Of the 52 experts, 23 are representatives of companies such as Nokia, Google, IBM or Airbus. In addition Google, Facebook, Amazon, SAP, Microsoft and Apple are represented by an organization called DigitalEurope. SAP is also represented via two additional members: by the AI expert Markus Noga and by the chairman of the expert group Pekka Ala-Pietilä, who is 4

---

1 High Level Expert Group on AI.
2 *Smuha*, CRi 2019, 97 at para. 16.

130    **Articles**
*Molavi Vasse'i* – The Ethical Guidelines for Trustworthy AI – A Procrastination of Effective Law Enforcement

CRi 5/2019

member of the supervisory board of SAP. All together, half of the group is indeed talking for the industry, some companies even represented twice or three times. Not even 10 of the experts are representing consumers or human rights. No academic expert on data protection is part of the group.

5    Diversity is not only about the presence but also about the right balance of different stakeholder's voices. This disproportion and underrepresentation of the civil society can unfortunately be observed in many expert hearings and other government adviser bodies in terms of AI policies[3].

## 2. Another Set of Ethical Guidelines

6    By this date, more than 80 Ethical Guidelines[4] have been published on national and international level. Most of them are industry-led, followed by civil society, government (incl EU), academic sector, industry and other associations. The approaches are not new.

7    Already late 2010, experts from the worlds of technology, industry, arts, law and social sciences met at the joint EPSRC and AHRC Robotics Retreat to discuss robotics and AI, its applications in the real world as well as the possible benefits to society.

8    They came up with five core principles for robots that could easily be applied accordingly to algorithmic decision making systems[5]:

1.    "Robots should not be designed as weapons, except for national security reasons.

2.    Robots should be designed and operated to comply with existing law, including privacy.

3.    Robots are products: as with other products, they should be designed to be safe and secure.

4.    Robots are manufactured artefacts: the illusion of emotions and intent should not be used to exploit vulnerable users.

5.    It should be possible to find out who is responsible for any robot."

9    The first 3 principles were meant as an evolution of Asimov's 3 laws of robotics[6]. The purpose of the principles was to provide consumer and citizen confidence in robotics as a trustworthy technology fit to become pervasive in our society[7] – reflecting the purpose of the Ethical Guideline for Trustworthy AI.

10    Was there a need for the creation of yet another new set of ethical guidelines, tasked by the EU Commission? The principles are a blueprint of the existing guidance that can be condensed to 5 to 10 joint Principles: *Inclusive growth, Sustainable development and well-being; Human-centred values and fairness; Transparency and explainability; Robustness, Security and safety as well as Accountability.*

11    As also stated by the OECD Recommendation of the Council on Artificial Intelligence[8]: these should be principles for responsible stewardship of trustworthy AI[9].

12    Even the Beijing AI Principles from China – developed in a dystopian surveillance state, as a place of censorship and suppression of the diversity of opinions – surprisingly share the same list of ethical principles[10].

An alternative option might have been to opt in core principles    13
commonly used in bioethics: *beneficence*, *non-maleficence*, *autonomy*, and *justice* and adding missed aspects like explainability to it[11].

## a) The EU Commission's Approach & Ethics Shopping

Nevertheless, the EU Commission's approach was to come up    14
with a new set of Ethical Guidelines for Trustworthy AI, divided in 3 chapters.

In the introduction, Trustworthy AI is defined to be lawful,    15
ethical and robust. Those attributes are meant to be at the same logical level, cumulatively and not alternatively.

–   In the first chapter, the "Foundation", consisting of the 4 Principles "Respect for human autonomy", "Prevention of harm", "Fairness" and "Explicability" are set.

–   The second chapter (the "Realisation") concentrates on seven key requirements, that are in part redundant with the Principles: "Human agency and oversight", "Technical robustness and safety"," Privacy and data governance", "Transparency", "Diversity", "non-discrimination and fairness" and "societal and environmental wellbeing" as well as "accountability".

–   The third and most interesting chapter provides an assessment list that will undergo a piloting phase by stakeholders to gather practical feedback.

Within contemporary discussions of 'AI ethics', there is no    16
agreed set of ethical standards. At least 42 states opted in the aforementioned OECD principles.

A large number of ethical principles competing with each other    17
can lead to " ethics shopping": In this case, the main, unethical risk is that all this hyperactivity creates a "market of principles and values", where private and public actors may shop for the kind of ethics that is best retrofitted to justify their current behaviours, rather than revising their structures and actions to make them consistent with a socially accepted ethical framework[12].

---

3    See as example the "KI Enquete Kommission" in Germany.

4    AI Ethics Guidelines Global Inventory, assessed on 14.092019 on https://algorithmwatch.org/en/project/ai-ethics-guidelines-global-inventory/.

5    *Margaret Boden*, *Joanna Bryson*, *Darwin Caldwell*, *Kerstin Dautenhahn*, *Lilian Edwards*, *Sarah Kember*, *Paul Newman*, *Vivienne Parry*, *Geoff Pegman*, *Tom Rodden*, *Tom Sorrell*, *Mick Wallis*, *Blay Whitby* & *Alan Winfield* (2017) Principles of robotics: regulating robots in the real world, Connection Science, 29:2, 125-127.

6    *Asimov, Isaac.* 2004. I, Robot.New York: Random House LLC; see also for a good listing of the content of the most known Principles: Alan Winfield's Web Logalso for "An Updated Round Up of Ethical Principles of Robotics and AI" http://alanwinfield.blogspot.com/2019/04/an-updated-round-up-of-ethical.html.

7    *Joanna J. Bryson* "The meaning of the EPSRC principles of robotics" https://www.tandfonline.com/doi/full/10.1080/09540091.2017.1313817.

8    Adopted May 2019.

9    OECD Recommendation of the Council on Artificial Intelligence https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449.

10   Beijing AI Principles assessed 15.9.2019 https://www.baai.ac.cn/blog/beijing-ai-principles.

11   *Luciano Floridi* and *Josh Cowls* 'A Unified Framework of Five Principles for AI in Society'.

12   *Floridi L.* (2019) Translating Principles into Practices of Digital Ethics: Five Risks of Being Unethical.

### b) Alternative Global Reference Framework

18 To avoid this, the EU Commission could have referred directly to human rights as a global framework providing already binding guidance instead of drafting new guidelines.

19 In a footnote, the AI HLG refers to the human rights:

20 *'Fundamental rights lie at the foundation of both international and EU human rights law and underpin the legally enforceable rights guaranteed by the EU Treaties and the EU Charter. Being legally binding, compliance with fundamental rights hence falls under trustworthy AI's first component (lawful AI). Fundamental rights can however also be understood as reflecting special moral entitlements of all individuals arising by virtue of their humanity, regardless of their legally binding status. In that sense, they hence also form part of the second component of trustworthy AI (ethical AI).'*[13]

21 In liberal democracies, human rights are widely recognized as essentially 'constitutional' in status to provide effective guarantees that individual freedom will be cherished and respected. In addition, a commitment to effective human rights protection is part of democratic constitutional orders.

22 The UN Universal Declaration of Human Rights (UNHR) is the best-known international human rights charter. Despite the number of, and variation between, regional human rights instruments in the Americas, Africa and Europe, they are all grounded on a shared commitment to uphold the inherent human dignity of each and every person, in which each individual is regarded of equal dignity and worth, wherever situated.

23 These shared foundations reflect the status of human rights standards as basic moral entitlements of every individual in virtue of their humanity, whether or not those entitlements are given explicit legal protection and therefore would be a worthy reference for the development of Trustworthy AI[14].

24 As a framework, human rights are the most promising standard to ensure ethical development of technology.

### c) Potential Conflicts Between Ethical Norms

25 Another shortcoming of the existing codes of ethical conduct is that they fail to acknowledge potential *conflicts* between ethical norms.

26 The HLG notes itself that *'The law provides both positive and negative obligations, which means that it should not only be interpreted with reference to what cannot be done, but also with reference to what should be done and what may be done. The law not only prohibits certain actions but also enables others. In this regard, it can be noted that the EU Charter contains articles on the 'freedom to conduct a business' and the 'freedom of the arts and sciences', alongside articles addressing areas that we are more familiar with when looking to ensure AI's trustworthiness, such as for instance data protection and non-discrimination.'*[15]

27 At the same time, no guidance is provided to solve the conflicts, being one of the major issues of the current development and implementation of AI. The same is true for most other ethical codes of conduct. The existing conflict in the technological development of recent years due to extreme power asymmetry between industry and wider society naturally always comes at the expense of citizens and human rights and this development is reinforced by the use of AI. And there is no guidance on how to deal with these situations.

28 Another strength of human rights as a framework is that they do provide a solution for conflicts between competing rights and diverging interests. This structured framework for reasoned resolution of conflict arising between competing rights and collective interests in specific cases is widely understood by human rights lawyers and practitioners, forming an essential part of a 'human rights approach'[16].

29 Even if the need for another set of ethical guidelines as guidance can be doubted, the Ethical Guidelines for Trustworthy AI are a detailed piece of guidance towards the operationalizing of existing moral rights. They can be used by companies to train and develop internal standards and to develop regulatable AI. The more important question to tackle, however, is how to solve conflicts and to restore the level playing field between competing rights again. As the power mismatch persists, the more powerfully represented interests prevail in conflicts. This is traditionally to the detriment of citizens.

## 3. Red Lines

30 With a deterministic view on technology development, the question of red lines seems to be a taboo. "What is technically realizable should also be allowed to be done", "Technology development can only be positive for everyone if you only let it go." are common assumptions towards AI.

31 Non negotiable red lines were initially demanded, later cancelled. Thomas Metzinger, a Professor of Theoretical Philosophy and member of the AI HLG and his colleague Urs Bergmann, a machine-learning expert from Zalando, were initially given the task of defining red lines for values that are not negotiable. In the final draft, however, no red lines are formulated.

32 What Metzinger reports to a newspaper[17] sheds an unfortunate light on the process. Metzinger and Bergmann had organized several workshops with many stakeholders on the subject and defined as a deliverable a number of such ethical restricted areas:

− the research on autonomous weapon systems and citizen scoring, as it practiced in China;

− the automated identification of persons with the help of face recognition; or

---

13  Footnote 12 Ethical Guidelines for Trustworthy AI.

14  *Karen Yeung*, *Andrew Howes*, and *Ganna Pogrebna* "An End to Ethics Washing"To be published in M Dubber and F Pasquale (eds.) The Oxford Handbook of AI Ethics, Oxford University Press (2019) forthcoming.

15  S.6 Ethical Guidelines for Trustworthy AI.

16  S.4 *Karen Yeung*, *Andrew Howes*, and *Ganna Pogrebna* "An End to Ethics Washing"To be published in M Dubber and F Pasquale (eds.) The Oxford Handbook of AI Ethics, Oxford University Press (2019) forthcoming.

17  *Thomas Metzinger* 'Ethik-Waschmaschinen made in Europe' (2019) assessed on 15.9.2019 https://background.tagesspiegel.de/ethik-waschmaschinen-made-in-europe.

132 **Articles**
*Molavi Vasse'i* – The Ethical Guidelines for Trustworthy AI – A Procrastination of Effective Law Enforcement

**CRi** 5/2019

– AI systems that work in disguise has been detected by the experts as red lines during the workshops.

33 What then happened is a lesson about the political processes in Brussels and the lobbying of industry. The numerous industry representatives within the AI HLG pushed for the term "red line" to be erased from the entire document. The word "non-negotiable" and the expression "red lines" no longer appear in the final version. Red lines turned to "critical concerns"[18].

34 This is regrettable as the underlying topics are serious and need strict rules and enforcement.

### a) Lethal Autonomous Weapons

35 The military use of AI in cyber warfare, of weaponized unmanned vehicles or of drones[19] is one the most destructive and fatal trends and a big threat for security of humans around the world.

36 According to media reports, the US government alone intends to invest two billion dollars in military AI projects over the next five years[20]. The development and use of machine-driven warfare is spreading[21].

37 The use of autonomous weapons for warfare leads to a lowering of the inhibition threshold for decisions on military interventions with unforeseen chain reactions and dangerous deterioration of relations between states. The lowering of the "marginal costs" of an attack due to the lack of traceability and the dramatic reduction of the probability of own endangerment carries a high risk for peace and freedom of the citizens of the world. From 2015 until today, 4,502 AI/Robotics researchers and 26,215 other supporters have signed an open letter[22] against the development and use of lethal autonomous weapons (LAW's).

38 It is difficult to think of any reasonable explanation for not banning autonomous weapons. Security and Freedom within and outside Europe are at disposal. This is in particular incomprehensible in view of the narrative of the EU itself as a peace project.

### b) Profiling, Nudging and Manipulation, Social Scoring

39 By means of technology equipped with sensors of all kinds in digital space, in public space, at home and even on and in the body, detailed profiling of citizens has become possible. The evaluation of movement patterns, the analysis of user behavior allows more or less accurate but also false predictions of future behavior.

40 Most citizens do not want their behavior to be measured and clustered. As an example, according to a study by PWC, the majority of German citizens reject the analysis of their online behavior[23]. Most participants of the same study perceive social scoring as a risk and not as an opportunity.

41 Private data from social networks, which would otherwise not be publicly visible, would only be revealed by one in four of the participants even for a reward like a cheaper loan. Even the collection and analysis of all data already publicly available for the assessment of creditworthiness is only acceptable to about one in five Germans.

42 More than 70 percent of the people interviewed are afraid that their user data will lead to false conclusions about their creditworthiness.

43 Not even the rejected and much-criticized social scoring and mass surveillance with its heavy impact on the autonomy of individuals and groups was accepted as a red line. In regard of the massive and invasive spreading of surveillance technology, the scoring and measurement of the human behavior and emotions, related manipulation and disinformation campaigns as well as social control, it is worrying that the HLG AI could not agree to ban such harmful practices.

44 The AI HLG could have taken the opportunity and addressed the aspect of democratic control, governance and political deliberation of AI systems and determine principles against the application of AI systems for the reduction of social cohesion, for example by isolating people in echo chambers or by increasingly anthropomorphized technical devices. Further Guidance would have been needed against the abuse of AI systems for political reasons in the context of automated propaganda, bots, fake news, micro targeting, election fraud or Deepfakes. This would have been a clear added value to the existing guidelines and a reaction to the status quo.

45 Civil society is completely uninvolved in the decision on the excessive use of surveillance technology online, in the public sphere and even at home with IoT. For such online surveillance

46 often a parallel secondary use is employed, since state actors resort to technology of large private platforms, such as the image recognition of Amazon or the storage of highly sensitive data on AWS servers owned by Amazon. This explains the lack of willingness on the part of the states to take action against violations of the law: they are the beneficiaries of the development.

47 This unresolved conflict of interests at the expense of the citizens therefore remains.

## II. Neglected Ethical Issues

48 The HLG AI missed the chance to reflect major upcoming issues as well as wanted and unintended harms and downsides of the Tech and AI development as a push for a course correction.

49 Decisions within the criminal justice process may threaten several human rights, including the right to a fair trial, the presumption of innocence and the right to liberty and security.

---

18 *Thomas Metzinger* 'Ethik-Waschmaschinen made in Europe' (2019) assessed on 15.9.2019 https://background.tagesspiegel.de/ethik-waschmaschinen-made-in-europe.

19 E.g. recent drone strikes on Saudi Arabian oil processing centers, see: https://www.nytimes.com/2019/09/14/world/middleeast/saudi-arabia-refineries-drone-attack.html.

20 *Zachary Fryer-Biggs*, (2018)THE PENTAGON PLANS TO SPEND $2 BILLION TO HELP INJECT MORE ARTIFICIAL INTELLIGENCE INTO ITS WEAPONRY https://publicintegrity.org/national-security/the-pentagon-plans-to-spend-2-billion-to-help-inject-more-artificial-intelligence-into-its-weaponry/.

21 *Zachary Fryer-Biggs*, (2019) COMING SOON TO THE BATTLEFIELD: ROBOTS THAT CAN KILL https://publicintegrity.org/national-security/future-of-warfare/scary-fast/ai-warfare/, assessed on 15.9.2019.

22 https://futureoflife.org/open-letter-autonomous-weapons/.

23 P.10 "Is Germany ready for social scoring?", PWC Study April 2018.

## 1. Hidden Social and Ecological Costs

50  Almost no guideline talks about AI in contexts of nurture, help and care, welfare, social responsibility or ecological networks. In AI ethics, technical artefacts are primarily seen as isolated entities that can be optimized by experts so as to find technical solutions for technical problems[24].

51  In connection to this, no guideline deals in detail with the question where systems of algorithmic decision making are superior or inferior, respectively, to human decision routines. None of the "hidden" social and ecological costs of AI systems are addressed[25].

52  Issues of unsustainable technology development such as lithium mining, the exploitation of rare earth minerals, e-waste, or energy consumption are mostly ignored.

53  AI implementations might lead to massive job losses. The digital transformation and automation tendencies has to be monitored with care and caution and shaped to avoid to leave individuals or groups behind.

54  Education, innovation of work culture, decrease of working hours and better distribution of profit and work must be implemented.

55  Precarious working conditions of clickworkers are also not addressed. The tax avoidance of big Tech, the psychological situation of content moderators remain unaddressed.

56  Although "clickwork" is a necessary prerequisite for the application of methods of supervised machine learning, it is associated with numerous social problems[26], such as low wages, awful work conditions and psychological consequences, which tend to be ignored by the AI community. The situation of content moderators who observe and rate thousands of often deeply disturbing images and videos every day, leading to lasting psychological impacts, is perfectly traced in the the documentary "the cleaners[27] ". These are the externalized costs of the AI development which need to be addressed as well.

## 2. Artificial General Intelligence (AGI)

57  One striking aspect is that the danger of a malevolent AGI[28] or existential threats is not discussed in any of the ethics guidelines, even though this is a thoroughly discussed topic in other contexts[29].

58  There is a vivid debate about the danger and, depending on the discipline, the opinions and estimations about the emergence of superintelligence and possible threats are very controversial.

59  While some (Kurzweil) see the breakthrough to general AI as imminent and solely beneficial, others perceive AGI as science fiction narrative and reject the possibility of a consciousness of machines.

60  A third group of experts expects a breakthrough in AGI development and warns strongly of the dangers.

61  At least, the debate could have been mentioned. It is for society to decide whether further research should be done in this direction in order to speed up the findings and the application or not.

## 3. "Buy-out" of Research

62  Large parts of university AI research are financed by corporate partners.

63  In January 2019, Facebook invested US\$7.5 million in a Centre on Ethics and AI at the Technical University of Munich, Germany[30].

64  Two months later, Amazon announced it would partner with the US National Science Foundation (NSF) "to commit up to \$10m each in research grants over the next three years focused on fairness in AI". On 10 May, letters of intent are due to the US National Science Foundation (NSF) for a new funding program for projects on Fairness in Artificial Intelligence, in collaboration with Amazon[31].

65  Google has spent millions of pounds funding research at British universities over the last five years. Oxford University alone has received at least £17m from Google.

66  In light of this, it remains questionable to what extent the ideal of freedom of research can be upheld – or whether there will be a gradual "buyout" of research institutes.

67  Companies' input in shaping the future of AI is essential, but they cannot retain the power they have gained to frame research on how their systems impact society or on how we evaluate the effect morally. Governments and publicly accountable entities must support independent research, and insist that industry shares enough data for it to be kept accountable[32].

68  A university abdicates its central role when it accepts funding from a firm to study the moral, political and legal implications of practices that are core to the business model of that firm[33].

69  And society must demand increased public investment in independent research rather than hoping that industry funding will fill the gap without corrupting the process.

## 4. Problematic Public-Private Partnerships

70  None of the existing guidelines raises the issue of public-private partnerships and industry-funded research in the field of AI.

---

24  Dr. *Thilo Hagendorff* (2019) The Ethics of AI Ethics An Evaluation of Guidelines S.4.

25  Dr. *Thilo Hagendorff* (2019) The Ethics of AI Ethics An Evaluation of Guidelines S.5.

26  By *M. Six Silberman*, *Lilly Irani*, and *Joel Ross* 'Ethics and Tactics of Professional Crowdwork' (2010).

27  "The Cleaners" A film by *Hans Block* and *Moritz Riesewieck* (2018).

28  See *Zeng/Lu/Huangfu*, Linking AI Principles, p. 4, available at: https://arxiv.org/ftp/arxiv/papers/1812/1812.04814.pdf.

29  *Müller* and *Bostrom* 2016 assessed on 15.9.2019 https://aiimpacts.org/muller-and-bostrom-ai-progress-poll/.

30  https://www.newstatesman.com/science-tech/technology/2019/06/how-big-tech-funds-debate-ai-ethics.

31  Blogpost Amazon Developer, 25. March 2019, assessed on 15.9.2019 https://developer.amazon.com/blogs/alexa/post/1786ea03-2e55-4a93-9029-5df88c200ac1/amazon-and-nsf-collaborate-to-accelerate-fairness-in-ai-research.

32  *Yochai Benkler* "Don't let industry write the rules for AI" (2019) Nature.com.

33  *Yochai Benkler* "Don't let industry write the rules for AI" (2019) Nature.com.

71 On the contrary, the HLG AI encourages the enforcement of Public-Private Partnerships (PPP) by recommending the set-up of a new PPP to foster sectoral AI ecosystems[34].

72 Today's leading technology companies – that have grown to control monopolies over the last 10 years – were founded at a time of high faith in market-based mechanisms. In the 1990s, regulation was limited, and public services such as railways and utilities were privatized.

73 Governments and publicly accountable entities must support independent research, and insist that industry shares enough data for it to be kept accountable.

74 The big tech companies use their immense wealth and power to influence policymakers at every level. Governments cannot delegate any more to them, neither the construction of a digital infrastructure nor the drafting of a policy frameworks to the private sector.

75 Citizens and businesses need a neutral and public digital basic infrastructure. This is a clear government task and must not conflict with economic interests.

76 Sensitive areas such as e-health and e-government cannot be uncritically operationalized by PPP. Own solutions for basic digital public utilities for health and state services must be deployed. In this area, economic interests and basic public utilities should be decoupled.

77 Instead of more dependencies on private partners, this requires a massive build-up of government in-house expertise and personnel and is the right next step to gain the needed independence and freedom of action.

78 Investments are a powerful tool of steering and governance. The EU Commission should seize the opportunity to only invest in beneficial tech development, in independent and neutral digital infrastructure, and especially in education.

79 The realization of projects by private partners with a track record of disinformation, massive data protection breaches and human unfriendly business models is alarming. The set-up of new PPP would reward and perpetuate the history of malpractice by the industry.

80 The partnerships also lead to aggravate the monopolization and therefore the contortion of competition. Often instead of encouraging small businesses, or to use of open source solutions, big tech companies are chosen to implement the e-government infrastructure[35] or even health services.

## 5. Protection and Enabling of Whistleblowers

81 The AI development is purely driven by an economic logic. The companies strive solely for a profitable use of machine learning systems. The developers are not focused on values or ethics.

82 This comes as little surprise as there is no systematic education of developers and engineers about ethical issues, nor are they empowered to raise ethical concerns. Especially as the engagement in ethical issues can slow down the development and the time to market.

83 Many of the major scandals of the last years would have been unthinkable without the use of AI. Echo chamber effects, the spread of fake News and propaganda bots, AI always played a key role to the effect of diminishing social cohesion, fostering instead radicalization, the decline of reason in public discourse and social divides[36].

84 Countless companies strive for diminishing and overcoming human autonomy, employing more and more subtle techniques for manipulating user behavior via micro targeting[37], nudging, dark patterns in UX-design.

85 It is elementarily necessary to create an environment and infrastructure, where developers and insiders have the opportunity to inform the public about unethical and unlawful business practices to fight the disinformation und to help creating a human friendly economy. Projects like "#notokgoogle" or "#techwontbuildit" should be listened to.

## III. Enforcement – From the WHAT to the HOW

86 The existence of mere guidelines by itself has no effect. The most important step forward, therefore, is the operationalization of the guidance and the installation of governance and oversight mechanisms by the executives.

## 1. Weaknesses of the Regulatory Framework

87 When even the Beijing AI Principles are largely aligned with the Ethical Guidelines created by the HLG AI, the difference must lie in the interpretation of the principles.

88 This leads to the weakness of ethical guidelines in general: they are mostly too vague and elastic and secondly, they are non-binding and therefore without any effect as long as they are not adopted and truly implemented by the industry.

89 In several parts of the Ethical Guidelines for Trustworthy AI document, the HLG emphasizes that the algorithmic systems must be ethical, robust AND lawful. The three necessary and cumulative requirements are meant to be at an equal level. But at the same time the ethical guidelines – that in part (only) reflect existing laws like human rights, non-discriminations laws, data protection laws, IT security and integrity of systems or liability laws – are explicitly described as non-binding.

90 The fact that the same requirement can at the same time be a binding norm and a non-binding ethical principle, is confusing and harmful, as the industry might interpret it as weakening of the existing binding laws.

91 This impression is supported by the public communication on the Ethical Guidelines. The HLG KI broadcasts mixed signals with a discrepancy between disclaimer and framing:

92 The Guidelines state that "We however recall that it is the duty of any natural or legal person to comply with laws – whether

---

34 Policy and Investment Recommendations for Trustworthy AI, High-Level Expert Group on Artificial Intelligence S.17.

35 Article in Tagesspiegel "The state in the Microsoft corset?" assessed 15.9.2019 https://background.tagesspiegel.de/der-staat-im-microsoft-korsett.

36 S. 8 *Hagendorff* (s.a.).

37 *Matz, S.* et al. 2017 'Psychological targeting as an effective approach to digital mass persuasion'.

applicable today or adopted in the future according to the development of AI. These Guidelines proceed on the assumption that all legal rights and obligations that apply to the processes and activities involved in developing, deploying and using AI systems remain mandatory and must be duly observed[38] ". However, during the first AI Assembly in Brussels, self-regulation was promoted as a favorable way to govern AI in general.

93 The name chosen for the *Regulation* Panel was "50 shades of self-regulation" and in the final summary by the moderator of the panel – Andrea Renda, Chairholder of the Google Chair in Digital Innovation and member of the HLG AI – which was streamed in contrast to the actual panel itself, the dominant critical voices of the majority of the panelists against self-regulation and soft law were not mentioned with a single word.

## 2. Systematic Weaknesses

94 Besides the problematic framing of the Ethical Guidelines, systematically it would have been favorable to classify the ethical design of AI and robustness as necessary prerequisites for the lawfulness of algorithmic systems to avoid confusion about the relation to existing laws.

95 Transparency and explainability are not the goal, but the condition for the verifiability and accurate use of AI. Algorithmic systems that are neither explainable and nor controllable should not be used in critical areas that impact humans or society.

96 AI ethics formulated in countless guidelines lack reinforcement mechanisms. There are very few initiatives with governance or oversight mechanisms that ensure and enforce the compliance of voluntary commitments. Most documents are recommendations, presentation of principles, or guidelines. Of the 21 examples[39] that can be labelled as voluntary commitments, quality seal or similar, only three mention some sort of oversight mechanism or quality control. 3 out of 83 analyzed Guidelines.

97 According to a recent study, the effectiveness of guidelines or ethical codes is *almost zero* as they do not change the behavior of professionals from the tech community. In the survey, 63 software engineering students and 105 professional software developers were scrutinized. They were presented with software-related ethical decision scenarios, testing whether the influence of the ethics guideline of the Association for Computing Machinery (ACM) influences ethical decision-making.

98 The results are disillusioning but foreseeable: *"No statistically significant difference in the responses for any vignette were found across individuals who did and did not see the code of ethics, either for students or for professionals."*[40]

99 The wish for self-regulation by the industry might be understandable, but illusionary.

## 3. Similar Weaknesses in EU Legislation

100 The trend of inappropriate regulatory and enforcement hesitation can also be seen in the current example of digital platform regulation.

101 On the 20 June 2019, a long awaited and relevant but barely mentioned legislation came into force for more fairness of business practices and greater transparency in online platforms like

Google, Amazon, Apple (regarding the App store etc.), Ebay or Booking.com. While the rules affect B2B[41] relationships and are meant to protect small and medium sized businesses, the positive effects would also affect consumers.

The Regulation (EU) 2019/1150 on promoting fairness and    102
transparency for business users of online intermediation services[42] will apply from 12 July 2020.

This Regulation substantiates the need for (also algorithmic)    103
transparency/explainability and fairness – two of the core ethical AI principles of all existing guidelines.

As an example the law states in Art. 5 ("Ranking"), that Provi-    104
ders shall set out the main parameters determining ranking and the reasons for the relative importance of those main parameters as opposed to other parameters.

Art. 7 ("Differentiated treatment") states that the Platform pro-    105
viders shall include in their terms and conditions a description of any differentiated treatment which they give. This description shall refer to the main economic, commercial or legal considerations for such differentiated treatment.

The EU Commission as the regulatory body is tasked[43] to con-    106
tinuously monitor the application of this Regulation and to periodically evaluate this Regulation and closely monitor its effects on the online platform economy. Codes of conduct are also encouraged[44].

But unfortunately no direct sanctions or fines are included in    107
the regulation, only mediation (Art. 14) and the possibility of judicial proceedings by representative organizations or associations and by public bodies. Evaluations seem to replace sanctions.

## 4. Weaknesses of "ePerson" Concept

In other respects, it can be positively emphasized that the HLG    108
AI resists against efforts to deregulate and limit accountability and to create harmful liability silos by pushing for the introduction of an "e-person":

Recommendation 29.7: "In addition, we urge policy-makers to    109
refrain from establishing legal personality for AI systems or robots. We believe this to be fundamentally inconsistent with the principle of human agency, accountability and responsibility, and to pose a significant moral hazard[45]".

---

38  S. 6 Ethical Guidelines for trustworthy AI.

39  Veronika Thiel,"Ethical AI guidelines": Binding commitment or simply window dressing?"(https://algorithmwatch.org/en/ethical-ai-guidelines-binding-commitment-or-simply-window-dressing/).

40  *McNamara, Andrew*, *Justin Smith*, and *Emerson Murphy-Hill*. 2018. "Does ACM's code of ethics change ethical decision making in software development?" S.4.

41  Business to Business.

42  Assessed 14.9.2019 https://eur-lex.europa.eu/legal-content/EN/TXT/?uri= CELEX:32019R1150.

43  Recital 47 assessed 15.9.2019 https://eur-lex.europa.eu/legal-content/EN/ TXT/?uri=CELEX:32019R1150.

44  Recital 48 assessed 15.9.2019 https://eur-lex.europa.eu/legal-content/EN/ TXT/?uri=CELEX:32019R1150.

45  Policy and investment recommendations for trustworthy AI, assessed on 14.9.2019 https://ec.europa.eu/digital-single-market/en/news/policy-and-investment-recommendations-trustworthy-artificial-intelligence.

## IV. Conclusions

110 Numerous governance proposals are already formulated by researchers and institutions. The establishment of 'an FDA for algorithms', a 'right to reasonable inferences', new roles for consumer protection agencies, proposals based on tort liability in combination with algorithm certification by a regulatory agency, mandatory algorithmic impact assessments on human rights and common standards to name some of them are all viable proposals towards the right direction. Besides this, 28 countries in the United Nations have explicitly endorsed the call for a ban on lethal autonomous weapons systems. Unfortunately all these proposals are still in the ideation phase.

111 The use of AI is capable of expanding existing threats, increase risks, introduce new threats and to modify and worsen the character of threats. The AI race and competition narrative is not helpful. Safeguards are needed to avoid a race to the bottom regarding human rights and autonomy. This takes time and joint efforts. A piloting project to evaluate innovative and tech driven ways to regulate and govern AI would be a valuable new project to overcome the enforcement crisis in the digital area.

### *Ramak Molavi Vasse'i*

Attorney-at-law and Advisor Legal Affairs at iRights Lab and Innovation Law Expert (The Law Technologist). Futher, independant policy advisor as well as member of the EU Commission's AI Alliance and visiting lecturer at the University of Potsdam.

IT-law, artificial intelligence, distributed ledger technologies and data protection

## *John P. Beardwood / Paula Millar**

# Failed ERP Implementation Case Study of MillerCoors v HCL

## Lessons Learned – Again – from a Failed SAP ERP Implementation Project

*The story is eerily familiar. Customer engages service provider to implement an enterprise SAP solution across its various sites. Alas, the implementation is increasingly subject to delays and cost overruns. The relationship between the parties deteriorates. Customer refuses to accept and pay the provider invoices and ultimately terminates, and then sues the provider for more than $100 million. Provider counterclaims. Let the finger pointing begin!*

*The following are the specifics of this particular case. In 2013, MillerCoors LLC ("MillerCoors"), a large brewery company, engaged HCL Technologies Limited and HCL America, Inc. ("HCL"), an SAP consultant and implementer, to implement an enterprise SAP solution across its various breweries. The engagement was complex, multi-year and required significant staffing contributions from both parties. The implementation was subject to delays and cost overruns. The relationship between the parties deteriorated to the point that MillerCoors refused to accept and pay three invoices and ultimately terminated Work Order No. 1-3 related to the implementation. MillerCoors then brought a claim against HCL for damages suffered in an amount in excess of $100 million (the "Claim"). HCL responded with a defence and counterclaim, alleging breach of contract and tortious interference with the agreements by MillerCoors and sought damages for MillerCoors' wrongful termination of the agreement (the "Reply").*

*The parties settled the dispute in 2018. Court records do not detail the terms of the settlement.*

*Based on the pleadings alone, the truth of which were never proved in court, this paper reviews some of the key issues of the lawsuit, and identifies certain lessons learned for practitioners that draft these contracts.*

## I. Background

### 1. The Project

1 In September 2013, MillerCoors issued a request for proposals ("RFP") for a Business Process and System Transformation Realization project (the "Project").

2 The intent of the Project was to adopt a common set of best practice business processes and implement them in a new, a single, uniform enterprise SAP software solution across MillerCoors' various breweries, thereby standardizing business processes across the organization.

3 HCL responded to the RFP. In October 2013, MillerCoors awarded the Project to HCL.