

Robots in the Boardroom? Artificial Intelligence & Corporate Governance

2nd December 2020, 13:00 GMT
Webinar, University of Glasgow



Corporate and
Financial Law
Research Group

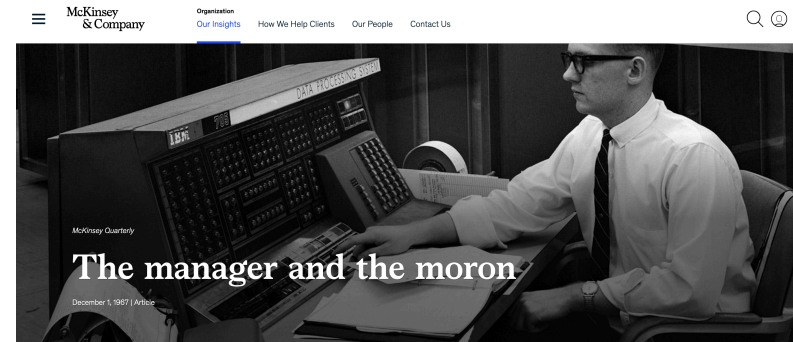


Florian Möslein



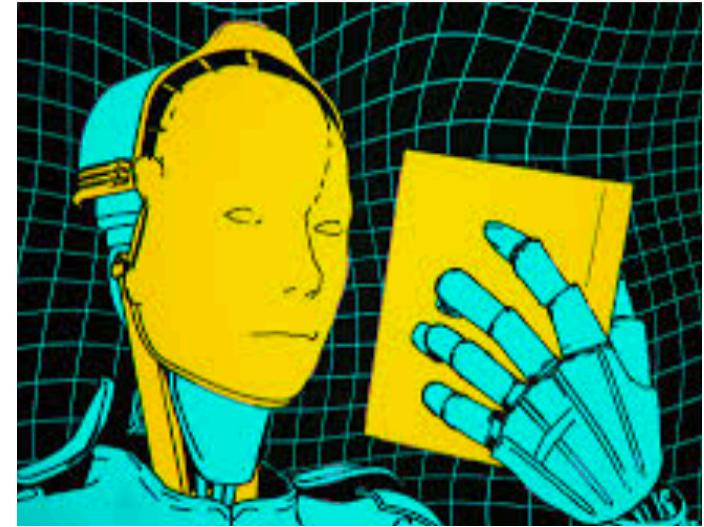
INSTITUT FÜR DAS RECHT
DER DIGITALISIERUNG

AI & CG: State of Debate



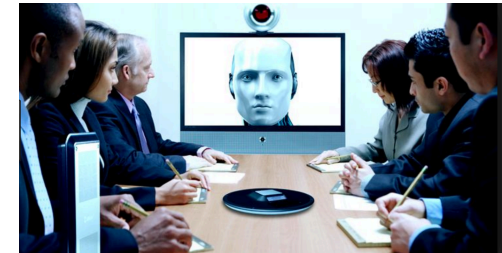
- **Peter F. Drucker (1967). The manager and the moron. McKinsey Quarterly, 3(4), 42–52:** “The computer does not make decisions, it only executes commands. It’s a total moron.”
- **Management literature today: At the top of the agenda for business leaders**
 - Davenport, T. H., & Ronanki, R. (2018). *Artificial intelligence for the real world*. Massachusetts: Harvard Business Review.
 - Libert, B., Beck, M., & Bonchek, M. (2017). *AI in the boardroom: the next realm of corporate governance*. MIT Sloan Management Review Blog.
 - Hilb, M. (2020). *Toward artificial governance? The role of artificial intelligence in shaping the future of corporate governance*. Journal of Management and Governance.
- **Legal Debate:**
 - Möslein, F. (2018). *Robots in the boardroom: artificial intelligence and corporate law*, in: Barfield, W. & Pagallo, U. (eds), *Research Handbook on the Law of Artificial Intelligence*, pp. 649-670.
 - Petrin, M. (2019). *Corporate Management in the Age of AI*, Colum. Bus. L. Rev. 965.
 - Armour, J. & Eidenmüller, H. (2020). *Self-Driving Corporations*, 10 Harv. Bus. L. Rev. 87 (2020).
- **European Commission, DG JUST:** Study on the relevance and impact of Artificial Intelligence for Company Law & Corporate Governance (ongoing, consortium led by EY): „The questions of how AI is being used or is likely to being used in future by company law and corporate governance actors, the consequences of such use, and whether/how they should be addressed in EU company law and corporate governance rules, have not yet been analysed“.

AI & CG: Key issues



- **Is it Science Fiction?**
 - AI as autonomous corporate actor?
 - Legal personhood for AI?
- **Current Relevance of AI:**
 - Ability to translate „big data“ into comprehensible information
 - Reducing uncertainty by making reliable predictions
 - AI as assistants to corporate decision-makers
- **Current Role of AI in Corporate Governance:**
 - Assisting shareholders and investors: AI-assisted investment and voting decisions
 - Assisting corporate control: AI-assisted auditing & supervision
 - Assisting directors: AI-assisted management decisions (rather than „robo directors“)

AI & CG: A case in point



- VITAL („Validating Investment Tool for Advancing Life Sciences“)
 - company decisions on investment opportunities
 - „the world’s first artificial intelligence company director“
 - no real voting right, but observer status
- Importance of this technological development:
 - supremacy of AI with respect to (certain) company decisions
 - need for reform in company law?
- Legal question not so much: Can robos replace human directors? But:
 - Can human directors delegate decisions to AI - or are they even required to do so?
 - Directors’ duties in case of delegation to AI? In particular: Duty to monitor
 - Which standards of care and loyalty apply? (P) no case law yet

Variety of Emerging Rules on AI

Mai 2019



OECD-Grundsätze für künstliche Intelligenz

Worum geht es?

Die OECD-Grundsätze für künstliche Intelligenz fördern eine künstliche Intelligenz (KI), die innovativ und vertrauenswürdig ist und die Menschenrechte und demokratische Werte achtet. Sie wurden am 22. Mai 2019 von den OECD-Mitgliedsländern durch die Verabschiedung der *OECD-Richtlinien zur künstlichen Intelligenz* angenommen. Die KI-Grundsätze der OECD sind die ersten von Regierungen unterzeichneten Grundsätze dieser Art. Neben den OECD-Mitgliedern haben weitere Länder, darunter Argentinien, Brasilien, Kolumbien und Costa Rica, den KI-Grundsätzen bereits zugestimmt und die Zustimmung weiterer Länder wäre zu begrüßen.

Die KI-Grundsätze der OECD setzen Standards für künstliche Intelligenz, die hinreichend praktisch und flexibel sind, um dem raschen Wandel auf diesem Gebiet standzuhalten. Sie ergänzen die bestehenden OECD-Standards in Bereichen wie Schutz der Privatsphäre, Risikomanagement im Bereich der digitalen Sicherheit und verantwortungsvolles unternehmerisches Handeln. Der umfassende Anwendungsbereich der Grundsätze stellt sicher, dass sie für KI-Entwicklungen weltweit angewandt werden können.

Wer hat sie entwickelt?

Die OECD hat eine aus mehr als 50 Mitgliedern bestehende Sachverständigengruppe für KI eingesetzt, um eine Reihe von Grundsätzen zu erarbeiten. Die Gruppe bestand aus Vertreterinnen und Vertretern von 20 Regierungen sowie führenden Persönlichkeiten aus der Wirtschaft, der Arbeitswelt, der Zivilgesellschaft, der Hochschulen und der Wissenschaft. Die Vorschläge der Sachverständigen wurden von der OECD aufgegriffen und zu den KI-Grundsätzen der OECD weiterentwickelt.

Die KI-Grundsätze der OECD

Die *Empfehlung* enthält fünf komplementäre wertebasierte Grundsätze für eine verantwortungsvolle Steuerung vertrauenswürdiger KI:

- KI sollte für die Menschen und den Planeten Nutzen bringen, indem sie ein inklusives Wachstum, eine nachhaltige Entwicklung und die Lebensqualität fördert.
- KI-Systeme sollten so konzipiert werden, dass sie das Prinzip der Rechtmäßigkeit, die Menschenrechte, demokratische Werte und die Vielfalt achten, und sie sollen angemessene Schutzmaßnahmen – z.B. bei Bedarf die Möglichkeit menschlichen Eingreifens – vorsehen. Das Ziel ist dabei eine faire und gerechte Gesellschaft.
- KI-Systeme sollten Transparenz und verantwortungsvolle Offenlegung gewährleisten um sicherzustellen, dass die Menschen KI-basierte Ergebnisse verstehen und hinterfragen können.
- KI-Systeme müssen über ihren gesamten Lebenszyklus robust und sicher funktionieren, und potenzielle Risiken sollten kontinuierlich beurteilt und kontrolliert werden.
- Die Organisationen und Personen, die KI-Systeme entwickeln, einführen oder betreiben, sollten für den einwandfreien Betrieb gemäß den oben aufgeführten Grundsätzen rechenschaftspflichtig sein.

© OECD 2019

goInnovat



Variety of Emerging Rules: International level

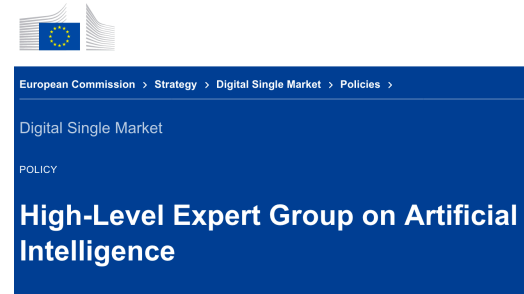
- G7: Declaration of March 2018, Multi-stakeholder conferences – cf. also G20 summit June 2019 in Osaka
- UNICRI Centre for Artificial Intelligence and Robotics
- OECD Recommendation of the Council on Artificial Intelligence, 22 May 2019 (<https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>)
 - Signed by 36 Member States and six additional States, supported by G20
 - Principles and policy recommendations
 - Developed by commission of 50+ experts
 - Not legally binding, but promises to form international standard (cf. Corporate Governance principles of 1999)

Variety of Emerging Rules: European level (I)



- High-level expert group on Artificial Intelligence (June 2018): 52 members with different background
- Ethic guidelines on Artificial Intelligence (December 2018, final version April 2019)
 - Requirements for trustworthy AI
 - „These guidelines do not explicitly deal with [...] lawful AI, but instead aim to offer guidance on fostering and securing [...] ethical and robust AI“ => no legal standard (?)

Variety of Emerging Rules: European level (II)



- High-level expert group on Artificial Intelligence (June 2018): 52 members with different background
- Ethic guidelines on Artificial Intelligence (December 2018, final version April 2019)
- Communication of the Commission on „Building Trust in Human-Centric Artificial Intelligence“ (COM 2019 168 final)
 - builds on ethic guidelines, but shortened and reformulated
 - does not claim binding character, but important step „towards international AI ethics guidelines“

Variety of Emerging Rules: National level – and self-regulation



Strategie Künstliche Intelligenz
der Bundesregierung

Stand: November 2018

- Example: Germany
 - no cross-cutting legal (or ethical) rules on artificial intelligence
 - dispersed specific rules, for instance on algorithmic trading (§ 80 para. 2 WpHG)
 - National strategy on Artificial Intelligence (November 2018), with promise to test and, if necessary, to modify the rules applicable to AI
 - Focus on implementation of ethical standards (Ethics by, in and for Design)
- Self-regulation
 - Rules implemented by companies (German Telekom) or by industry associations (Global Policy Framework of the International Technology Law Association)

Variety of Emerging Rules: Mapping exercise



16(c) Berkman Klein Center - <https://cyber.harvard.edu/story/2019-06/introducing-principled-artificial-intelligence-project>

Converging Substance of Emerging Rules

- Comparison EU-OECD: Convergence of regulatory content
- Elements (similar in OECD guidelines)
 - Technological basis (computer systems)
 - Purpose-orientation (complex aimed as defined by humans)
 - Impact of technological systems
 - Degree of autonomy?
 - Modes of action

Converging Substance of Emerging Rules:

1. Control & controllability of technology

- Human agency and oversight
 - Human-in-the-loop, human-on-the-loop, human-in-command
 - Which degree of human involvement?
- Technical robustness and safety
 - Human responsibility for well-functioning of technology
 - Protection against manipulation, cyber-security, technological risk-management

Converging Substance of Emerging Rules:

2. Disclosure

- Transparency
 - Protocols and documentation of technological processes
 - Explainable AI (black-box problem, related to autonomous decision-making) – but subject to various restrictions and limitations („as far as possible“)
- Accountability
 - Building on transparency, but related to responsibility
 - Third-party evaluations
 - „accessible mechanisms should be foreseen that ensure adequate redress“ => link to the „legal sphere“

Converging Substance of Emerging Rules:

3. Safeguarding individual rights

- Privacy and Data Governance
 - Going much further than data protection requirements
 - Integrity and control of data
 - Relevance, for instance, for robo advice
- Diversity, non-discrimination and fairness
 - Duty to avoid biases
 - Potential collisions with existing legal rules

Converging Substance of Emerging Rules:

4. Public good requirements

- Contribution to societal and environmental well-being
 - Vaguely formulated
 - Relevance largely unclear (AI services for everybody?)

Legal Relevance?

1. Distinguishing Law & Ethics

- Means of enforcement
 - If pure ethical standards, then no legal, but only social sanctions
 - If legal rules, then either private or public enforcement possible (for instance in case of violations of non-discrimination or disclosure rules)
- Example: Violation of non-discrimination rules by AI
 - Social sanction:
 - Loss of reputation
 - Legal sanctions:
 - Private persons can claim damages
 - Public authorities can impose penalties

Discriminating algorithms: 5 times AI showed prejudice

Artificial intelligence is supposed to make life easier for us all – but it is also prone to amplify sexist and racist biases from the real world

Legal Relevance?

2. Formal Classification

- From formal perspective, classification seems obvious:
 - Non-binding character of OECD rules and of EU recommendations
 - Wording of these instruments (f. ex. EU Commission Communication, p. 3: „non-binding and as such do not create any new legal obligations“)
 - Nonetheless raising the claim to establish international standards (and expectation towards States to take action)

Legal Relevance?

3. Effective impact

- Implicit legal relevance:
 - „Rechtsgeltungs- und Rechtserkenntnisquellen“
 - Hardening of soft law
 - Reasons: Lack of experience; New category of cases; Cross-cutting regulatory question
 - EU communication: „Finally, when unjust adverse impact occurs, accessible mechanisms should be foreseen that ensure adequate redress“.

Legal Relevance? Conclusion

- Crucial distinction: Formally no legal force, but potential implicit legal relevance
- Implementation: Concretization of legal standards, in particular corporate directors' duties of care and loyalty
- Example: Company directors delegating decisions to AI
 - Liability: Standard of care \Leftrightarrow Human responsibility for well-functioning of technology
 - Disclosure: Information rights of shareholders \Leftrightarrow Explainable AI
- Expectation:
 - Hardening of Ethical Standards on AI

Thank you for your attention.



INSTITUT FÜR DAS RECHT
DER DIGITALISIERUNG

PROFESSOR DR.
FLORIAN MÖSLEIN
LL.M. (LONDON)

MARBURG UNIVERSITY SCHOOL OF LAW
UNIVERSITÄTSSTR. 6
D-35032 MARBURG

T: +49 6421 28 – 21704

F: +49 6421 28 – 27046

MOESLEIN@IRDI.INSTITUTE

WWW.IRDI.INSTITUTE

