

BEYOND THE GDPR?

**Mireille Hildebrandt, FBA
PI COHUBICOL ERC ADG project**

Beyond the GDPR?

- Like the FTC: destroy not only the data but everything built on it
- For me, the only problem with the GDPR is its enforcement
- And, compliance (because that often introduces compliance software ...)

Beyond the GDPR

- Is the GDPR enough? Does it solve the problem of AI causing havoc?
- No, we are in the era of 'never enough data'
- Ever more applications are built with data-driven AI: health, safety and FR

Beyond the GDPR

- So now we have the upcoming EU data law framework
- Why does the GDPR matter? First bullwork against the growing data hunger
- Data minimisation & purpose limitation may prevent some of the havoc

Beyond the GDPR

- Havoc caused by a remarkable belief in the meaning-driven nature of text-data
- LLMs believed to be trained 'on the entire internet' and thus 'on the entire world'
- In that light I would propose that all data is syntactical, because intra-linguistic

Beyond the GDPR

- Syntactic refers to the relationships between signs: intra-linguistic reference
- Semantic refers to both the intra-linguistic and the extra-linguistic reference
- LLMs cannot ever 'get' anything but intra-linguistic reference (next wordtoken guessing)

Beyond the GDPR

In that sense the problem of data-driven AI is that all **harms** will be driven by meaning-agnostic data and models

Some references

■ On information, performative effect and the law:

- 2016. 'Law as Information in the Era of Data-Driven Agency'. The Modern Law Review 79 (1): 1–30. <https://doi.org/10.1111/1468-2230.12165>
- 2018. 'Law as Computation in the Era of Artificial Legal Intelligence: Speaking Law to the Power of Statistics'. University of Toronto Law Journal, March. <https://doi.org/10.3138/utlj.2017-0044>

■ On regulatory theory and law:

- 2018. 'Algorithmic Regulation and the Rule of Law'. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 376 (2128). <https://doi.org/10.1098/rsta.2017.0355>

■ On data-driven research:

- 2023. 'Ground-Truthing in the European Health Data Space'. In Proceedings of the 16th International Joint Conference on Biomedical Engineering Systems and Technologies - Volume 5 HEALTHINF: BIOSTEC, 15–22. Lisbon, Portugal: Scitepress. <https://www.scitepress.org/Link.aspx?doi=10.5220/0011955900003414>

Some references

■ On the ML pipeline and the political economy of recommender systems:

- 2022. 'The Issue of Proxies and Choice Architectures. Why EU Law Matters for Recommender Systems'. *Frontiers in Artificial Intelligence* 5. <https://www.frontiersin.org/article/10.3389/frai.2022.789076>.

■ On the Rule of Law:

- 2015. 'Radbruch's Rechtsstaat and Schmitt's Legal Order: Legalism, Legality, and the Institution of Law'. *Critical Analysis of Law* 2 (1). <http://cal.library.utoronto.ca/index.php/cal/article/view/22514>
- Laurence Diver, Tatiana Duarte, Gianmarco Gori, Emilie van den Hoven and Mireille Hildebrandt, *Research Study on Text-Driven Law* (Brussels 2023), funded by the ERC Advanced Grant 'Counting as a Human Being in the Era of Computational Law' (COHUBICOL) by the European Research Council (ERC) under the HORIZON2020 Excellence of Science program ERC-2017-ADG No 788734 (2019-2024). <https://publications.cohubicol.com/assets/uploads/cohubicol-research-study-on-text-driven-law-final.pdf>

Some references

- On information, performative effect and the law:
 - 2016. 'Law as Information in the Era of Data-Driven Agency'. *The Modern Law Review* 79 (1): 1–30. <https://doi.org/10.1111/1468-2230.12165>
 - 2018. 'Law as Computation in the Era of Artificial Legal Intelligence: Speaking Law to the Power of Statistics'. *University of Toronto Law Journal*, March. <https://doi.org/10.3138/utlj.2017-0044>
- On regulatory theory and law:
 - 2018. 'Algorithmic Regulation and the Rule of Law'. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 376 (2128). <https://doi.org/10.1098/rsta.2017.0355>
- On data-driven research:
 - 2023. 'Ground-Truthing in the European Health Data Space'. In *Proceedings of the 16th International Joint Conference on Biomedical Engineering Systems and Technologies - Volume 5 HEALTHINF: BIOSTEC*, 15–22. Lisbon, Portugal: Scitepress. <https://www.scitepress.org/Link.aspx?doi=10.5220/0011955900003414>

FUNDAMENTAL ISSUES WITH THE HARMS-BASED APPROACH

The harms approach

- **In law:**
 - tort law
 - to claim damages
 - identifiable individual harm caused by wrongful behaviour of the tortfeasor
- **In quantitative policy science and in utilitarian political philosophy:**
 - positive freedom is limited by the harm principle
 - the state should not interfere unless this is warranted by potential harm
- **Key question:**
 - Can we measure potential future harm?
 - Can we avoid potential future harm by taking evidence-based measures?

The harms approach

- The utilitarian assumptions of the harms-based approach
 - John Stuart Mill:
 - liberalism and rule utilitarianism
 - the harm principle
- Anglo-American common sense is *marinated in* utilitarianism
 - Harms-based approach as a panacea (naïve)
 - CBA as a panacea (naïve)

The harms approach

- Regulatory theory is directly linked with utilitarianism
- “legal regulatory theory, legal regulation should target phenomena that
 - **cause** the problems that regulation aims to solve or
 - the phenomena that are **instrumental** for the desired regulatory outcomes”

[difference between instrumentality and instrumentalism]

The harms approach

- Legal theory is not equivalent with regulatory theory, it is *both more and less*
 - Legal norms are directly related to *legal certainty, instrumentality and justice*
 - Acknowledging the antinomian relation between them (Radbruch)
 - While rejecting a final ranking (Radbruch)
 - See also Waldron, Dworkin, Toope, Sen, Sandel, Brownsword

The harms approach

- *Regulatory theory prioritises the instrumentality of legal norms*
 - Thereby 'reading' them as - only - means to policy ends
 - This easily results in weighing means in terms of their efficiency

The harms approach

- Economists have 'discovered' the Goodhart effect, elegantly worded by Strathern:
 - *'if you use a measure as a target, it ceases to be a good measure'*
 - check complexity theory about predicting complex systems
 - human society is a complex system, if anything
- *Law is defined by its performative effect (= legal effect)*
 - *This is not causal or logical but constitutive of our institutional reality*
- Legal effect is far more effective in the long run
 - It is related to Hart's 'internal aspect' of legal norms

Tort law

Fundamental Right

- Tort law = private law
 - Aims for the *compensation of harm caused by wrongdoing*
 - Need to *identify* individual harm
 - At societal level that would be an aggregate (collective action)
- Fundamental rights law = constitutional law
 - Aims to counter *violation, for which harm is NOT a condition*
 - Violation of rights whose substance is not computable
 - Violation of norms, whether that violation results in identifiable harm, or not
 - *The consequences of norm violation are potential anomia, not harm caused*

EU data law

- *Replacing a legal approach with one of policy science* is not a good idea:
 - It will invite social engineering
 - Nudging (behavioural economics, combined with machine learning)
 - Game theoretical manipulation (rational choice theory, combined with MAS)
 - Psychometrics (combined with sentiment analysis and the above)
- The legal approach is not only more effective in the long run
 - It also *takes human agency seriously* in a way that causal influencing cannot
 - Regulatory theory is a *cynical approach to human agency*
 - The legality principle supports a qualitative, normative understanding of effectiveness

EU data law

- The AI Act aims to protect health, safety and fundamental rights
 - It hopes to prevent risks to health, safety and fundamental rights
- It takes a *risk* approach which is not the same as a *harm* approach
 - *A risk to a right is not a risk of harm but a risk of violation*
 - A health risk is a risk of trauma or disease
 - A safety risk is a risk of physical injury or death
- The risk approach in relation to fundamental rights is *a precautionary approach*
- The AI Liability Act takes a harm approach as it concerns tort law

Meaning-driven law

Meaning-agnostic regulation

- **Meaning-agnostic systems:**
 - Symbolic computing systems (knowledge-based, logic processing, rules as code)
 - Sub-symbolic computing (ML, DL, NLP, GPT)
- **In the context of a computing system all data is meaning-agnostic**
 - Symbolic computing systems operate on the syntactic relations
 - Ontology, semantic web (intra-linguistic references)
 - Logical operators define syntactice relations
- **Meaning is created on the cusp of intra- and extra-linguistic reference**
 - this is not in the remit of digital computers

Meaning-driven law

Meaning-agnostic regulation

- **Meaning-agnostic systems:**
 - Symbolic computing systems (knowledge-based, logic processing, rules as code)
 - Sub-symbolic computing (ML, DL, NLP, GPT)
- **In the context of a computing system all data is meaning-agnostic**
 - Symbolic computing systems operate on the syntactic relations
 - Ontology, semantic web (intra-linguistic references)
 - Logical operators define syntactice relations
- **Meaning is created on the cusp of intra- and extra-linguistic reference**
 - this is not in the remit of digital computers

Meaning-driven law

Meaning-agnostic regulation

- **Meaning-agnostic influencing (regulation):**
 - perlocutionary instead of performative effects
 - nudging 'humans as puppets' instead of addressing 'humans as agents'
- **Meaning-driven interaction (legislation, case law, doctrine, fundamental principles, custom):**
 - Governing through written and unwritten legal speech acts
 - Taking human agency seriously
 - Also in the sense of not trying to see legislation as regulation-only



m.e. mcNair