Automated Decision-Making: A Comparative Perspective

Professor Margot Kaminski

University of Colorado Law

@margotkaminski

University of Colorado Boulder

Automated Decision-Making

- **(1) Legal background for comparative perspective**
  - How the GDPR regulates automated decision-making
- **(2) Normative comparisons**
  - How other efforts compare to the GDPR
  - A comparative framing of impact assessments (risk regulation)
- **(3) Takeaways**

The GDPR in a Nutshell

- The GDPR in a nutshell:
  - Individual rights
  - Data controller (company) obligations

The GDPR in a Nutshell

- Governance style:
  - often vague text
  - delegated interpretation
  - “collaborative” compliance tools
- Against the backdrop of:
  - Potentially huge fines
  - A human right, interpreted by a human rights court
  - Private and public enforcement
  - Regulators who provide guidance
  - Pre-existing law (the Directive) & regulatory infrastructure

Automated Decision-Making in the GDPR

- The GDPR on the whole regulates “the processing of personal data wholly or partly by automated means”
  - *And* contains **specific provisions on automated decision-making** with significant effects.
- Start with ADM-specific provisions, then briefly point to several potentially relevant generally applicable provisions.

GDPR ADM Notice Rights

- Individual Notice Rights
  - Arts. 13 & 14
  - Affirmative obligation for data controller to disclose
  - "the **existence of automated decision-making**, including profiling, referred to in Art. 22…"
  - "and, at least in those cases, **meaningful information about the logic involved**,"
  - "as well as the **significance and the envisaged consequences** of such processing for the data subject."
  - When?
    - when data are collected from the individual (Art. 13) or within a reasonable time of gathering data from a third party (Art. 14)

GDPR ADM Access Right

- Individual Access Right
  - Art. 15
  - Same language.
    - “meaningful information about the logic involved…”
    - “significance and envisaged consequences”
  - May be accessed “easily and at reasonable intervals” (Rec. 63)

Article 22:
ADM “Due Process”

- “The data subject shall have the right **not to be subject to** a decision **based solely on automated processing**, including profiling, which produces **legal effects** concerning him or her or **similarly significantly affects** him or her…”
  - Not a new right/rights
  - Based on language in the Directive, with changes
    - Arguably making Article 22’s right(s) broader, deeper, and stronger.
    - Guidance suggests: “based solely on automated processing” covers more than it may seem; “significant effects” could include manipulative advertising
    - And: this is a ban, not opt-in
    - Except…

Article 22: “Suitable Safeguards”

- If an individual has given **explicit consent** or the ADM is **necessary for a contract** or a Member State has **authorized particular ADM through law**
  - Then ADM may be deployed
  - But only if the data controller “implement[s] **suitable measures to safeguard** the data subject’s rights and freedoms and legitimate interests”
  - What are these “**suitable measures**” or “suitable safeguards”?

Article 22: “Suitable Safeguards”

- “Suitable safeguards” listed in Article 22 largely focus on the **individual**
  - Individual “**due process**” in the face of ADM:
  - the right to **obtain human intervention**
  - The right to **express his or her point of view**
  - and to **contest (challenge)** the decision.

Article 22: “Suitable Safeguards”

- It’s a mistake to think these three rights are all the GDPR does.
  - “Suitable safeguards” in Recital clearly includes a **right to explanation**
    - **Not necessarily the same as “meaningful info about the logic involved**”
    - Enables contestation of a particular decision by the affected individual

Article 22: “Suitable Safeguards”

- Art. 22 also clearly (though not on its face) aims at **systemic (risk) regulation** with substantive goals:
  - Recital:
  - implement **technical and organisational measures** appropriate to ensure, in particular, that factors which result in **inaccuracies** in personal data are corrected
  - **Prevent**… **discriminatory effects** on natural persons on the basis of racial or ethnic origin, political opinion, religion or beliefs, trade union membership, genetic or health status or sexual orientation, or processing that results in measures having such an effect…
- Guidelines on ADM also mention **audits**, **third-party expert oversight**, **DPIAs**

Due Process and… DPIAs

- The Data Protection Impact Assessment (DPIA) (Art. 35) applies to high-risk data processing in general
  - Automated decision-making with significant effects is identified as **one type of high-risk data processing** that requires a DPIA.
  - Guidelines on ADM emphasize the centrality of the DPIA.
    - crucial aspect of the **suitable safeguards** aimed at mitigating harms, on a systemic and ongoing basis, to individual rights and freedoms
  - Thus the DPIA (ideally) links company obligations (risk management) to individual rights
    - Proposed that information produced during DPIA could/should feed into information disclosed to individuals.

Other Relevant Aspects of the GDPR

- Crucial to understand that the ADM Rights do not exist in a vacuum.
  - E.g. DPIAs for other forms of processing, including systematic large-scale surveillance of public spaces
  - Fundamental principles: fairness & transparency, purpose limitation/data minimization, accountability
  - Data protection by design and by default
  - Other individual rights, including for your purposes: a **right to object** (Art. 21) to processing, that includes and goes beyond ADM
    - Includes an **absolute right** to object to direct marketing.

(2) Normative Comparisons

- A few observations:
- The GDPR approach to Impact Assessments is very much situated in both the GDPR's two-prong approach and its governance style.
  - **Systemic risk regulation** but targeted at **protecting core individual rights**
  - Stylistically "**Meta regulation**" or collaborative regulation:
    - Meant to affect the **internal infrastructure, norms, heuristics** of a company
    - In conversation with a regulator/enforcer

(2) Normative Comparisons

- By contrast, proposals in U.S.:
- Notable absence of individual rights (of explanation, contestation, “fair” decisions, data protection writ large)
  - Makes us an outlier, internationally
- Impact Assessments are more of an **enterprise risk-management** tool.
  - Self-assessments aimed at internal risk mitigation.

(2) Normative Comparisons

- There's a third model for impact assessments that neither jurisdiction appears to be using: NEPA
- Impact Assessments as **public accountability/iterative policymaking**
  - Wyden/Booker/Clarke Algorithmic Accountability Act attempts to thread needle between enterprise risk management and NEPA model, with various forms of public disclosure (reports, partial database)
  - Where a proposed WA law (SB 5116) requires public comment on algo accountability reports (albeit for state actors)

(2) Normative Comparisons

- Brings me to two major weaknesses for the GDPR regime: public/third-party accountability and stakeholder participation
  - Very regulator-centric
- Lacks voices in, voices out.
  - E.g. DPIA: controllers “shall” consult data subjects… but only “where appropriate”
  - DPIAs are not made public (it’s recommended, at least a summary)
- Voices in/out are crucial for
  - (a) defining the harms to be mitigated
  - (b) accountable “meta regulation”/collaborative regulation.

(2) Normative Comparisons

- By contrast, a number of proposed U.S. laws take stakeholder participation seriously– even where there isn't public disclosure of AIAs.
  - Regulator consults with affected communities during rulemaking
  - Company consults with affected communities/representatives during impact assessment
- Again, necessary for:
  - Oversight over the algorithm/data set itself
  - But also, oversight over the company/agency deciding how it's going to mitigate the risks of the algorithm.

(3) Takeaways

- The GDPR (and for that matter, the draft EU AI Act) doesn't raise big concerns about having a broad definition of automated decision-making
  - Because the regulations are really triggered by "significant effects" of the decision or "high risks" from processing.

(3) Takeaways

- Placing a “human in the loop” is the least sophisticated– and probably most problematic– mode for governing ADM.
  - Because “hybrid” systems create plenty of challenges of their own
  - Instead, systemic oversight (including ex ante risk mitigation, ideally ongoing and iterative) coupled with robust individual rights is the emerging model.

(3) Takeaways

- The developers and users of ADM should share responsibility for its harms (comparative weakness of the draft EU AI Act)
  - Like the driver of a car and the car manufacturer
  - “crashworthy” AI…

(3) Takeaways

- An impact assessment can be a very different tool in different regimes, towards different goals.
  - Collaborative governance, versus enterprise risk-management, versus public accountability and policy iteration
  - Ex ante, versus ongoing, versus iterative and including post-market measures

(3) Takeaways

- Voices in, voices out are crucial to effective governance.
  - Transparency matters not just because algorithms are opaque, but for governance to be effective/accountable.
    - The difference between a self-assessment and governance.
  - AIAs/DPIAs can be linked to individual disclosure rights: the information produced during a DPIA can/should be used to constitute the "meaningful info" disclosed to an individual
  - Impacted stakeholders matter and should be meaningfully consulted, to afford accountability and participate in constituting harms, whether it's at a rulemaking stage or during an impact assessment, or both