A Legal Right to an Explanation of Algorithmic Decision Making?

Lilian Edwards

Professor of Law, Innovation and Society

Newcastle Law School

EU citizens might get a 'right to explanation' about the decisions algorithms make

Getty Images

RELATED VIDEOS

Residente: Embra
Your Difference

See Goodman and Flaxman, 'EU Regulations on Algorithmic Decision-Making and a "right to Explanation"' [2016] arXiv:160
Wachter et al , "Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation" International Data Privacy Law, 2017.
Edwards and Veale "Slave to the Algorithm? Why a Right to an Explanation is Not the Remedy You Are Looking For" 16 Duke Law & Technology Review 18 (2017)

GDPR - route to a right to an explanation?

Art 22 (1) : 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.

A right to an explanation how exactly??
(a) Where art 22(1) rt to object is excluded, "safeguards" must be put in place
AND
(b) **Explicit** reference to "safeguards" in art 22(2)(b) (public task)
and
22(4)(special category data, only allowed with explicit consent/substantial public interest)

Safeguards.. Art 22..

(3). In the cases referred to in points (**a) and (c)** of paragraph 2, the data controller <u>shall</u> implement **suitable measures** to safeguard the data subject's rights and freedoms and legitimate interests, at least the right to obtain <u>human intervention</u> on the part of the controller, to <u>express his or her point of view</u> and to <u>contest</u> the decision.

(4) Decisions referred to in paragraph 2 shall not be based on special categories of personal data referred to in Article 9(1), unless point (a) or (g) of Article 9(2) applies and suitable measures to safeguard the data subject's rights and freedoms and legitimate interests are in place.

Safeguards ..-> Explanations!

- **Recital 71**

- [...] In any case, [ie, if basis is contract, explicit consent or public task – art 22(2)]

- ...such processing should be subject to **suitable safeguards**, which should include specific information to the data subject and the right to obtain human intervention, to express his or her point of view, to obtain an explanation of the decision reached after such assessment and to challenge the decision.

Issues with art 22 route to a right to an explanation

How often does art 22 apply anyway?

- How many automated decisions are "solely automated"? Nominal human involvement?
- What's a "decision"? (= a "measure"). Does there have to be an outcome or is a categorisation/optimisation in itself a decision?
- Legal or "significant" effect? Targeted adverts? Targeted *political* adverts? Effect on **group** rather than individual ? ("representation" rather than "allocation" harms)

- Only applies where decision made on basis of explicit consent, contract or public task, or re special category data!

- Recitals are only interpretative not mandatory?

"I wouldn't start from here.."

Another possible route to a "RTE" is via **information rights for data subjects**

- Art 15(h) – DS shall have right to obtain .. The following info:
  - "*meaningful information about the logic involved*" "at least in" cases of "automated decision making, including profiling"
- Issues?
  - Q. Is art 15(h) restricted as per art 22? Ie to only "solely automated" decisions of "legal or significant" effect? (No – see art 4(4)).
  - Is it a right to general (*ex ante*) info or specific details on how an individual decision was reached? (*ex post*)?
  - Possibly limited by IP/trade secrets? See recital 63. Result "should not be a refusal to provide all info to DS"
  - What *meaningful information* can operators provide?
    - Source code; decompositional vs pedagogical explanations; "counterfactual" explanations

Beyond the GDPR

EU

Platform Regulation 2019/1150 (B2B only) – requires transparency as to algorithmic rankings of search engines, for businesses only

Consumer Rights Directive, revised by 2019/2161 (B2C only) – info requirements re online marketplaces now require info as what "parameters" used to generate rankings, and how

National

UK Equality Act – duty to show you were non-discriminatory in hiring, based on protected characteristics, in a format claimant can "meaningfully engage with"

Judicial review – challenging the legality of decisions made by public sector agencies using algorithmic systems

See ICO/ATI draft Guidance on "Explaining Decisions Made With AI", 2019