

# AI: Law, ethics and responsibility

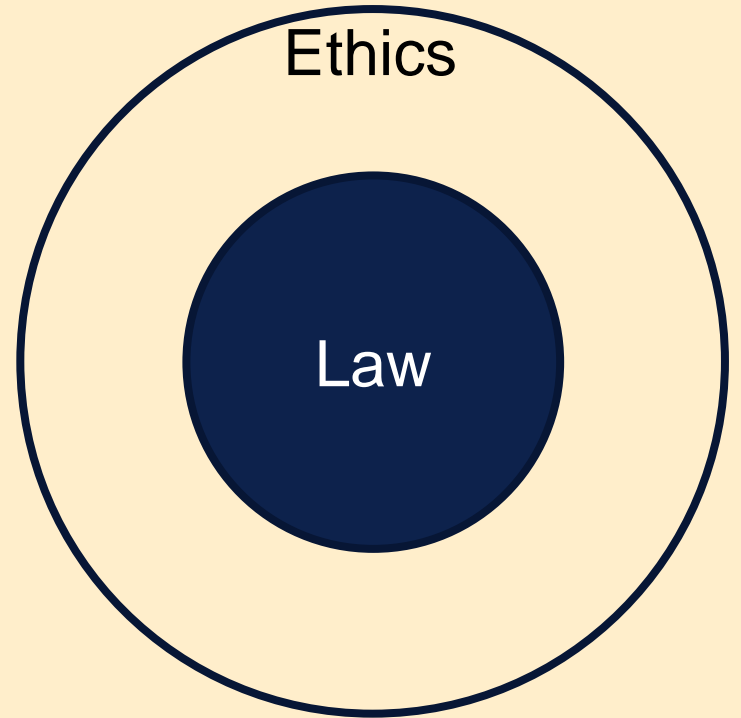
Andrew Cormack, Chief Regulatory Adviser, Jisc (@Janet\_LegReg)

Ari Korhonen, Senior Lecturer, Aalto University

## Traditional view

Law: what you *\*can\** do

Ethics: what you *\*should\** do



# But...

## EU HLEG ethics principles

Respect for autonomy

Prevention of harm

Fairness

Explicability

# But...

EU HLEG ethics principles	GDPR (legal) requirements	GDPR tools
Respect for autonomy	<i>?Automated decisions</i>	<i>See later</i>
Prevention of harm	Balancing test	DPIA (also respects autonomy, through consultation)
Fairness	Fairness/legal basis/purpose limitation	Fairness/legal basis/purpose limitation
Explicability	<i>?Right to explanation</i>	<i>See later</i>

# But...

## EU HLEG ethics requirements

Human agency & oversight

Technical robustness & safety

Privacy & data governance

Transparency

Diversity/non-discrimination

Societal/environmental  
wellbeing

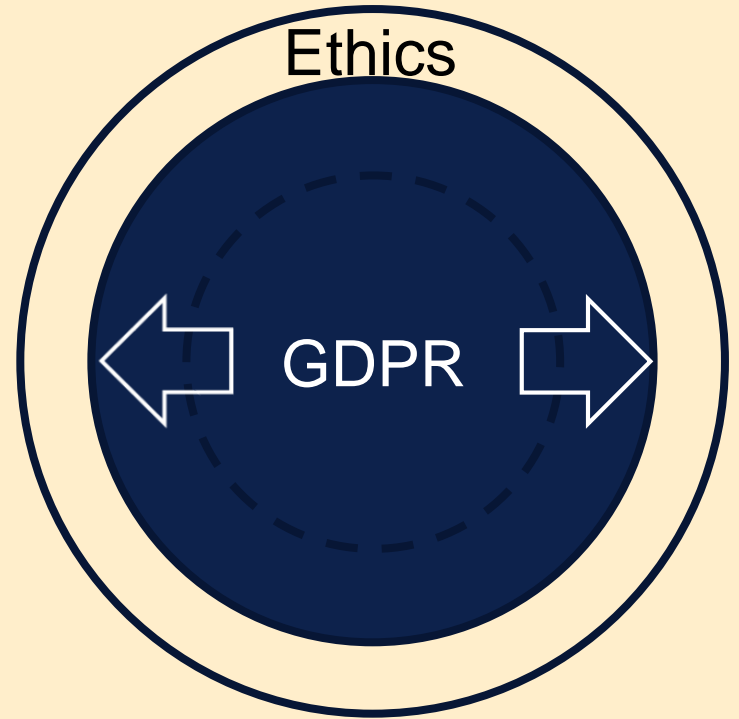
Accountability

# But...

EU HLEG ethics requirements	GDPR (legal) requirements	GDPR tools
Human agency & oversight	?Automated decisions	See later
Technical robustness & safety	Rec.71 on fair AI	Rec.71
Privacy & data governance	GDPR!	GDPR
Transparency	Transparency	Transparency
Diversity/non-discrimination	Rec.71 and Special Category Data	Rec.71 and SCD
Societal/environmental wellbeing		See later
Accountability	Accountability	Accountability

# GDPR is more than a “privacy law”

Also covers lots of “ethics”



# Remaining issues 1 – Right to Explanation

## May be in Art.15(1)(h)...

- Academics & regulators disagree 😞
- But educators need it anyway
  - If AI predicts student could do better, we need to know how
- This requires *retrospective, per decision* explanation
  - “Why did system think that about me?”



# Remaining issues 2 – Automated Processing

## Art.22 “right not to be subject...”

- Regulators disagree with themselves (contrast Art29WP views in 2013 and 2018)
  - Is this a right to review (see above) or a prospective ban on full automation?
- If it’s a ban, some decisions require human input
  - Using more knowledge than the AI has, e.g. context
  - Can be AI-informed
  - Probably requires a *prospective understanding* of what AI does
    - “What conclusions can it reach, in what circumstances?”
- Which (educational) decisions “produce legal effect ... or similarly significantly affect”?
  - Acceptance by institution? Acceptance onto course?...

# Remaining issues 3 – Societal/Environmental Wellbeing

- Should we do this at all?
  - See, UN HR (Art.26(2)) on purpose of education: “full development of human personality”
  - And UK Parliament report: “must be employed by the institution for improved educational outcomes as well as for ‘administrative’ purposes”
  - Is “improved educational outcome” better marks, better career, better fulfilment...?
- Should this be done by computer, or reserved for humans like us?
  - E.g. soldier, judge, physician,...
  - Does education raise any of these?

# Examples to think about

# Sorting Applicants/Automated Processing

How to create a balance between human oversight of AI and automation.

- Should AI choose which applicants are accepted onto an over-subscribed course?
- Raise the green card for yes, and the red for no

# Personalised Learning/Social Wellbeing

How to ensure AI enhances social and learning skills, rather than contributing to their deterioration.

- Should AI be used to tell students which texts they need to read to pass their exam?
- Raise the green card for yes, and the red for no

# Understanding Decisions/Explainability

How much understanding do we need when using algorithms.  
Should this limit our choice of techniques?

- Should a teacher know full details of an algorithm used to predict student grades, or just what data led to its conclusion for an individual student?
- Raise the green card for full details, and the red for significant data

# Discussion

# Before you start, think...

- If this works, what will I do with the results?
  - Is that something I should do?
- What's the minimum data I need to achieve it?
  - And is that still too intrusive/risky?
- What's the effect on students (from Slade & Prinsloo) and staff?
  - Should be collaborators, not recipients
  - Should change during education process
  - Success is wider than can be measured by data alone



# References

- GDPR
  - Text <https://ji.sc/gdpr-text>
  - DPIA [https://ec.europa.eu/newsroom/article29/item-detail.cfm?item\\_id=611236](https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=611236) (Art29WP adopted by EDPB)
  - Purpose [https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2013/wp203\\_en.pdf](https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2013/wp203_en.pdf)
  - Accountability [https://edps.europa.eu/data-protection/our-work/subjects/accountability\\_en](https://edps.europa.eu/data-protection/our-work/subjects/accountability_en) (EDPS)
  - Accountability <https://ico.org.uk/about-the-ico/news-and-events/news-and-blogs/2019/10/blog-embedding-accountability-we-want-to-hear-from-you/> (UK ICO)
- Ethics
  - EU HLEG <https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines>
  - UK Parliament <https://www.policyconnect.org.uk/appgda/research/trust-transparency-and-technology-building-data-policies-public-good>
  - UN Human Rights <https://www.un.org/en/universal-declaration-human-rights/index.html>
  - Slade S and Prinsloo P., Learning analytics: ethical issues and dilemmas, American Behavioural Scientist, Vol. 57 (10) 2013, pp 12-14.
  - Explainability <https://community.jisc.ac.uk/blogs/regulatory-developments/article/explaining-ai-algorithms>