Data Ethics & AI: Best Practices for Actuaries

Cathal Leahy, Laura Rossi

Deloitte Ireland

EAA e-Conference on
Data Science & Data Ethics

14 May 2024

USES OF DATA

01

DATA ETHICS PRINCIPLES INFORM DATA USAGE

- Data ethics is based on core ethical principles.

- Principles enable development of holistic ethical framework to data processing.

LIMITATIONS

02

Stated purpose at collection

Internal congruence

Open source data

Owner consent & protection

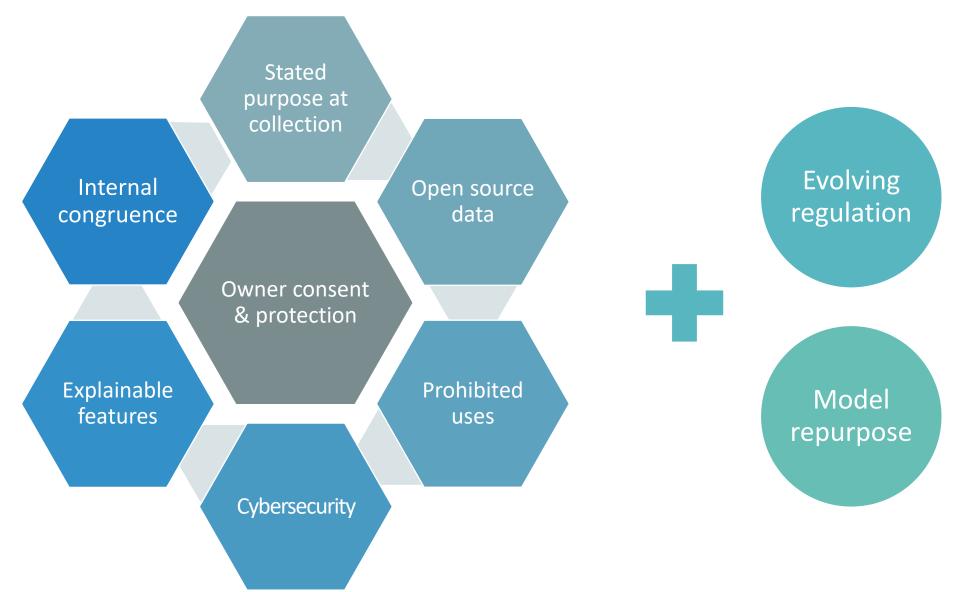Explainable features

Prohibited uses

Cybersecurity

+

Evolving regulation

Model repurpose

UnitedHealth uses faulty AI to deny elderly patients medically necessary coverage, lawsuit claims CBS News, November 20, 2023

Google hit with class-action lawsuit over AI data scraping
Reuters, July 12, 2023

What Air Canada Lost In 'Remarkable' Lying AI Chatbot Case Forbes, February 19, 2024

EMBEDDED DATA IN THIRD PARTY AI MODELS

03

CLASSIFICATION
Exact definition and designation criteria to be set out in final legal text or implementing acts

KEY REQUIREMENTS FOR GPAI PROVIDERS
Non-exhaustive list

ALL GPAI MODELS AND SYSTEMS

Examples[a]:
- Large language models and systems (GPT-3, Chat GPT-3, LaMda, Google Bard)
- Diffusion models and systems (DALL-E 2, Midjourney)

- GPAI models, including large Generative AI models, that can be used for a variety of tasks
- May be integrated into a large number of AI systems

- Transparency obligations **before market placement**, including:
  - Drawing up technical documentation for downstream providers
  - Complying with EU copyright law and disseminating detailed summaries about the content used in training
  - Watermarking AI generated or manipulated content

"HIGH-IMPACT" GPAI MODELS

Examples[c]:
- OpenAI's GPT-4
- (likely) Google DeepMind's Gemini

- GPAI models that could pose **systemic risks**
- **Designation criteria:**
  - FLOPs[b] > $10{\sim}25$ for computing used in training (to be kept under review)
  - Other qualitative / quantitative criteria (e.g. number of business users)

- <u>**All**</u> requirements applicable to all GPAI models and systems
- **Additional requirements**, including:
  - Conducting model evaluations
  - Assessing and mitigating systemic risks
  - Conducting adversarial testing
  - Ensuring sufficient cybersecurity protection
  - Reporting on energy efficiency
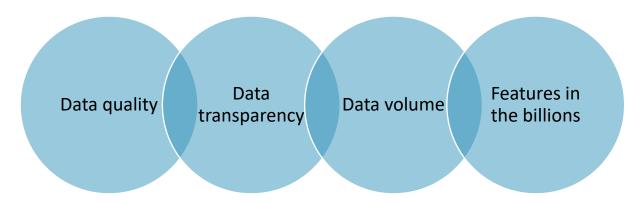 - Reporting of serious incidents to the EU Commissions

Transparency requirements

Generative AI, like ChatGPT, will not be classified as high-risk, but will have to comply with transparency requirements and EU copyright law:

- Disclosing that the content was generated by AI
- Designing the model to prevent it from generating illegal content
- Publishing summaries of copyrighted data used for training

From: EU AI Act: first regulation on artificial intelligence | Topics | European Parliament (europa.eu)
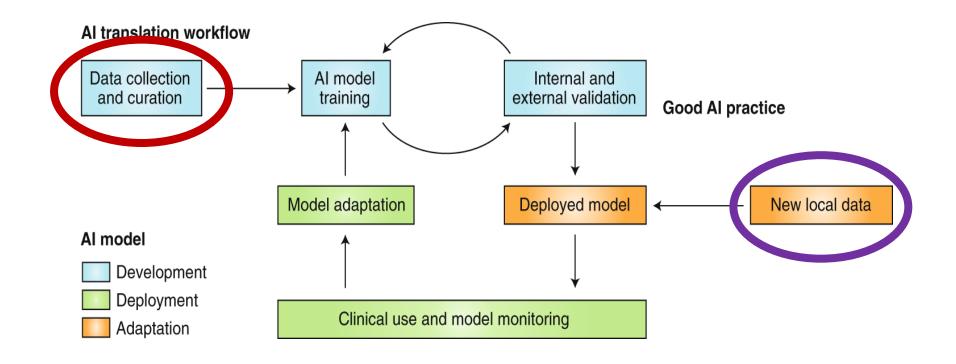
Areas to consider:

Data quality

Data transparency

Data volume

Features in the billions

➤ It's just my data, isn't it?

From: Hu, Y., Jacob, J., Parker, G.J.M. *et al.* The challenges of deploying artificial intelligence models in a rapidly evolving pandemic. *Nat Mach Intell* **2**, 298–300 (2020).https://doi.org/10.1038/s42256-020-0185-2

GOVERNANCE STRUCTURES

04

Cross-Functional Data Ethics Framework

Chief Risk Officer

Data Protection Officer and Privacy Team

Chief Information and Security Office

CEO and the Board of Directors

Chief Compliance Officer

Operations Business Functions

Heads of Departments

- Important role played by everyone in the organisation

- The CEO, the Board and other key individuals are integral to defining data ethics values

- Collaboration amongst stakeholders to integrate the defined practices into the organisation's culture.

What should organisations do?

Strategy and framework

Data ethics strategy design

Data ethics roadmap

Data ethics policies and procedures

Description of how data ethics is considered within the organisation.

Demonstrating accountability

Data Ethics Board

Data ethics by design framework

Data ethics impact assessment

How to embed Data ethics into your organisation

Culture, training and awareness

Employee awareness training

Data ethics leadership programme

Operation and Sustenance

Data ethics health check scan

Data ethics framework implementation

Third party data ethics due diligence assessment

Steps to a successful ethical data-processing organisation

START

Create Data Ethics Strategy and Framework

Policies and Procedures | Accountability

Culture ,Training and Awareness

Operations

Sustenance and established lines of defence

ABOUT ME

Cathal is a Senior Manager in the Artificial Intelligence and Data team in Dublin with over ten years of management consulting experience.

Cathal has practical experience of AI and Data programmes across public and private sector clients delivering projects across a wide range of areas such as regulatory compliance, cyber-security, fraud detection, Gen AI and many more.

Cathal also delivers a module on data ethics on behalf of the Analytics Institute in Ireland.

Cathal Leahy

Deloitte IE

ABOUT ME

Laura is a Consultant in the Actuarial Modelling Centre, and an Associate of both the Society of Actuaries in Ireland and the Institute and Faculty of Actuaries.

Laura has worked at Deloitte Ireland for 2.5 years, serving insurance clients across the life, non-life and reinsurance markets. She specialises in software and technology solutions.

Prior to Deloitte, Laura worked in the American software consulting industry and led courses in data ethics and data communication at Purdue University.

Laura Rossi

Deloitte IE

Thank you very much for your attention

Contact

Contact Cathal or Laura at:

caleahy@deloitte.ie

laurossi@deloitte.ie

Deloitte Ireland

EAA e-Conference on
Data Science & Data Ethics

14 May 2024