Defining "AI"

PROFESSOR KAMINSKI

MARGOT.KAMINSKI@COLORADO.EDU

OFFICE 417

Terminology: Technical definitions

- There is a lot of technical terminology that comes up when we discuss AI systems.
- Briefly go over a few:
- ***Machine learning***
- ***A model***
- ***Generative AI***
- ***AI***

Terminology: Machine learning

- *Machine learning*
  - "a category of methods aimed at *generalizing rules from data* to produce a desired outcome or prediction."
  - "The focus on learning from data is what sets machine learning apart from other forms of computer programming."
  - **BUT** importantly, humans are often still involved: supervised learning, reinforcement learning.

Terminology: A model

- *A model*
  - “the thing that machine learning produces”
  - "a collection of code and data that contains the general rules discovered during machine learning, and it can be used to make predictions or otherwise analyze using new data”
  - Understanding the difference between machine learning and a model lets us realize: machine learning is actually a *process*, not an *object*:
    - Two basic steps: (1) playing with the data, (2) the running model

Terminology: Generative AI

- *Generative AI*
  - Large Language Models (LLM) are an example.
  - A model (the product of machine learning) trained on a data set, to produce *statistically likely* outputs on the basis of that data set.
  - What Generative AI produces depends on what it's been trained on:
    - Computer code
    - Images (image diffusion models)
    - video
    - translations

Terminology: AI

- *Artificial Intelligence*
- What is Artificial Intelligence?
- Could mean any of:
  - Algorithms
  - machine learning models
  - deep learning and neural networks
  - large language models
  - other generative AI.
  - More broadly, AI is the *field of making computers perform tasks by approximating the way humans think*—or, generating human-like outputs.

Critiques of “AI”

- *Artificial Intelligence* comes with baggage.
- CS joke:
- You call it AI when you want to impress executives
- You call it machine learning when you talk to other computer scientists.

Legal Definitions

- Do lawmakers just adopt these or other technical definitions when they pass laws?
- Absolutely not.

Summary of how Legal AI definitions have evolved

- Legal definitions of AI at first tried to target particular programming methods, **or** the autonomy/the "intelligence" of a system
  - E.g. earlier EU AI Act definition (focus on programming approaches)
  - The Singapore definition (focus on approximating human intelligence)
  - "levels of autonomy" from NHTSA
- Then definitions started harmonizing around the 2024 OECD definition (which emphasizes *output*, especially *predictions as output*)
  - *Then* started making room for "general purpose AI" definitions, often shoehorned in.

Summary of how Legal AI definitions have evolved

- The paradigmatic case *was* an automated decisionmaker
  - Generative AI and foundation models in particular threw policymakers, especially in the EU, for a loop.

TAKEAWAYS

- (1) How you define AI depends on what legal harms you're trying to address (your paradigm case)
- (2) The definition is often broad, and the gatekeeping happens in other parts of the law
  - E.g. "significant effects"
- (3) Legal definitions don't track, necessarily, technical definitions– and that's an intentional thing
  - "futureproofing"
- (4) Harmonization is a policy choice.

OECD AI Principles

- The Organization for Economic Co-operation and Development (OECD) is an intergovernmental organization.
  - Economic orientation
  - Rather than human rights
  - Produces “soft law” principles to harmonize, encourage trade
    - Two versions of the definition: 2019 and 2024.

OECD Revised AI Principles (2024):

- *AI system:* An AI system is a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate **outputs** such as **predictions, content, recommendations, or decisions** that can influence physical or virtual environments. Different AI systems vary in their levels of **autonomy** and **adaptiveness** after deployment.

- *AI system lifecycle:* An **AI system lifecycle** typically involves several phases that include to: plan and design; collect and process data; build model(s) and/or adapt existing model(s) to specific tasks; test, evaluate, verify and validate; make available for use/deploy; operate and monitor; and retire/decommission. These phases often take place in an iterative manner and are not necessarily sequential…

EU AI Act (2024 Version, final)

- ‘AI system’ means a machine-based system that is designed to operate with **varying levels of autonomy** and that may exhibit **adaptiveness**s after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as **predictions, content, recommendations, or decisions** that can influence physical or virtual environments;

Colorado AI Act

- "Artificial Intelligence System" means any machine-based system that, for any explicit or implicit objective, **infers from the inputs the system receives how to generate outputs**, including **content, decisions, predictions, or recommendations**, that can influence physical or virtual environments.

Executive Order on AI (EO 14110)(Oct. 30, 2023)

- Issued in October, 2023
  - Definition looks awfully familiar by now…
  - But also, watch for slight wrinkles that emphasize human actors slightly more…
  - And for how trying to define technology in the law runs into a "pacing problem"

Executive Order on AI (2023)

- (b) The term “artificial intelligence” or “AI” has the meaning set forth in 15 U.S.C. 9401(3): a **machine-based system** that can, for a given set of **human-defined objectives**, make **predictions, recommendations, or decision**s influencing real or virtual environments. Artificial intelligence systems use machine- and **human-based inputs** to perceive real and virtual environments; abstract such perceptions into models through analysis in an automated manner; and use model inference to **formulate options for** information or action.

Executive Order on AI (2023)

- Another section of the Executive Order sets forth special reporting requirements for:

- (i) any model that was trained **using a quantity of computing power greater than 1026 integer or floating-point operations**, or using primarily biological sequence data and using a quantity of computing power greater than 1023 integer or floating-point operations; and

- (ii) any computing cluster that has a set of machines physically co-located in a single datacenter, transitively connected by data center networking of over 100 Gbit/s, and having a theoretical **maximum computing capacity of 1020 integer or floating-point operations** per second for training AI.

Harmonization and the definition of “AI”

- Why use vague(r) definitions?
  - What are the benefits of harmonizing a definition across different laws?
  - What are potential problems?
  - Should we focus on the definition of AI– or, on defining “bad AI acts rather than bad AI actors”?

The Problem/Puzzle of “general-purpose AI” models

- Both the EU AI Act and the Executive Order have provisions on generative AI and foundation models or “general-purpose AI models”
  - Really regulating a different set of uses and envisioned harms, especially for the Executive Order (national security concerns)
  - AI Act also addresses systemic threats (threats to democracy, to climate)

Executive Order on AI (2023)

- (k) The term "**dual-use foundation model**" means an AI model that is trained on broad data; generally uses self-supervision; contains **at least tens of billions of parameters**; is applicable across a wide range of contexts; and that exhibits, or could be easily modified to exhibit, high levels of performance **at tasks that pose a serious risk to security, national economic security, national public health or safety, or any combination of those matters…**

EU AI Act (2024 Version, final)

- ‘General-purpose AI model’ means an AI model, including where such an AI model is trained with a large amount of data using self-supervision at scale, that displays **significant generality and is capable of competently performing a wide range of distinct tasks** regardless of the way the model is placed on the market and that can **be integrated into a variety of downstream systems or applications**, except AI models that are used for research, development or prototyping activities before they are placed on the market…