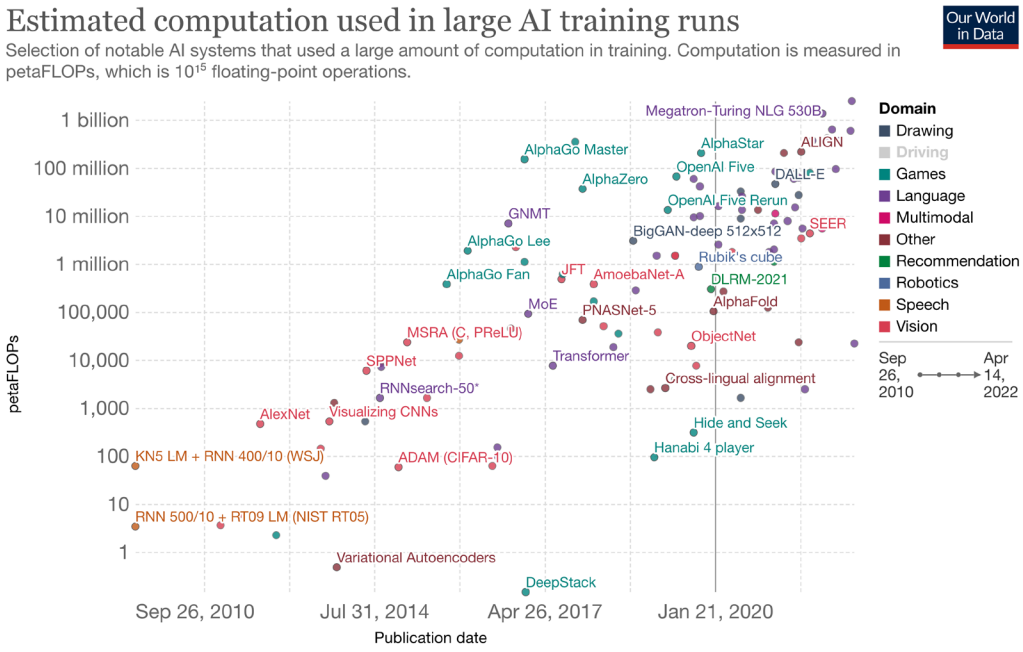
May 2022

General Purpose AI and the AI Act

Contact
risto@futureoflife.org
+372 5843 4759



Source: Sevilla et al. (2022)
Note: The estimates have some uncertainty but are expected to be correct within a factor of ~2.

WHY REGULATE GENERAL PURPOSE AI?

ALLAI, an organisation focused on responsible AI, recently published a policy paper, warning that excluding general purpose AI systems from the AI Act runs the risk of stifling innovation.[10] They argue that the exclusion of these systems from the AI Act means that the burden of making these systems compliant with the regulation falls entirely on the users of the AI systems instead of the developers. Users would have to make the systems meet the requirements for high-risk AI. This could be too much of a burden, especially for small businesses and startups, and might even be impossible to do. In this case, the AI Act may cut off SME access to cutting-edge general purpose AI systems and place Europe at a competitive disadvantage compared to the US or China.

ALLAI further explains that even if the general purpose AI developer helped downstream users with the technicalities of complying with the AI Act, it would put the latter in a position of full responsibility for any damage the system may cause, without providing the appropriate means for the user to seek redress. This could limit the uptake of general purpose AI systems and cause AI innovation to further concentrate with the developers. The latter could thus gain a strong competitive advantage over downstream users.

Furthermore, general purpose AI systems can be complex and often exhibit behaviors that surprise even their own developers. For example, GPT-3, which is trained to process natural language, unexpectedly acquired the ability to write rudimentary programs in a programming language.[11] Importantly, the complexity and opacity of these systems poses risks to fundamental rights, health and safety. Some of these systems have already caused alarm by propagating extremist content,[12] exhibiting anti-Muslim bias[13] or inadvertently revealing personal data.[14] A chatbot built using GPT-3 told a person to commit suicide.[15] We will likely see more incidents such as these if regulation fails to promote the safe use of this important new area within AI development.

10 AIA in-depth #1: Objective, Scope, Definition: https://allai.nl/wp-content/uploads/2022/03/AIA-in-depth-Objective-Scope-and-Definition.pdf

11 Evaluating Large Language Models Trained on Code: https://arxiv.org/abs/2107.03374

12 The Radicalization Risks Of GPT-3 And Advanced Neural Language Models: https://www.middlebury.edu/institute/sites/www.middlebury.edu.institute/files/2020-09/gpt3-article.pdf

13 'For Some Reason I'm Covered in Blood': GPT-3 Contains Disturbing Bias Against Muslims: https://onezero.medium.com/for-some-reason-im-covered-in-blood-gpt-3-contains-disturbing-bias-against-muslims-693d275552bf

14 Extracting Training Data from Large Language Models: https://arxiv.org/abs/2012.07805

15 Doctor GPT-3: hype or reality?: https://www.nabla.com/blog/gpt-3/

RECOMMENDATIONS

ARTICLE 3

NEW ARTICLE

Article 3 — 'General purpose AI system' means an AI system that is able to perform generally applicable functions such as image/speech recognition, audio/video generation, pattern detection, question answering, translation, etc, and is able to have multiple intended and unintended purposes.

An image recognition system that identifies signs of skin cancer has an intended purpose. Another system that identifies potholes in roads from images also has an intended purpose. A system able to identify skin cancer and potholes, has two (quite different) intended purposes. General purpose AI systems can have many more intended purposes as well as many unintended uses. The definition of general purpose AI systems hinges on its ability to do several different tasks.

General purpose AI systems are software, which means they can very quickly be applied to a wide range of areas - much faster than the EU can adopt new acts. Therefore, the solution is to cover them in this regulation by default, and to ensure the responsibility for their safety is not just on EU companies, but shared with the creators of general purpose AI systems. MEP Voss (EPP, DE) suggested a version of this recommendation in his first draft opinion on the AI Act for the Legal Affairs (JURI) Committee.

ARTICLE 4A

NEW ARTICLE

Article 4a: Obligations of providers of general purpose AI systems

Providers of general purpose AI systems shall:

a) ensure that their general purpose AI systems are compliant with the requirements set out in Article 15 in Chapter 2;
b) comply with the other requirements set out in Chapter 2 to the fullest extent possible;
c) assess the reasonably foreseeable misuse of their systems;
d) provide instructions and information about the safety of these systems to users and other relevant stakeholders in the supply chain;
e) regularly assess whether the AI systems have presented any new risks, including risks discovered when investigating novel use cases;
f) register their systems in the EU database referred to in Article 60.

General purpose AI systems are able to perform generally applicable functions such as image/speech recognition, audio/video generation, pattern detection, question answering, translation, etc. These systems are trained on broad data that can be adapted to a wide range of downstream tasks and applications[16], and have many intended and unintended purposes.

The wide range of applications for which general purpose systems can be used means that any flaw can have overarching effects on many sectors – a single failure could affect hundreds of downstream AI applications. For example, researchers recently showed that one general purpose AI system had an anti-Muslim bias. If left unaddressed, a bias of this type could affect media articles, educational materials, chatbots as well as other uses that might only be discovered when SMEs experiment with these systems. The potential use of general purpose AI systems for many tasks with different levels of risk justifies pre-market and post-market obligations.

Furthermore, general purpose AI systems will be integrated into various services and products. (European) companies integrating (mainly American) general purpose systems will be unable to understand the full scope of risks involved. Product safety regulation commonly uses the concept of 'reason-

16 AIA in-depth #1: Objective, Scope, Definition: https://allai.nl/wp-content/uploads/2022/03/AIA-in-depth-Objective-Scope-and-Definition.pdf

ably foreseeable use.' It is reasonable to ask providers to try to foresee the potential misuses of their general purpose AI systems because of the likely significant economic and societal impacts of these systems. They should address these risks to health, safety and fundamental rights beforehand, while acknowledging that not all such uses can be foreseen. Developers of these systems are best placed to estimate foreseeable uses and misuses.

ARTICLE 4B

NEW ARTICLE
Article 4b: Conformity assessment for general purpose AI systems Providers of general purpose AI systems shall ensure that their systems undergo a conformity assessment prior to their placing on the market or putting into service. The conformity assessment procedure will be based on assessment of the quality management system and assessment of the technical documentation, with the involvement of a notified body, referred to in Annex VII. Where the compliance of the AI systems with the requirements set out in Chapter 2 of this Title has been demonstrated following that conformity assessment, the providers shall draw up an EU declaration of conformity in accordance with Article 48 and affix the CE marking of conformity in accordance with Article 49.

As general purpose AI systems will form the foundation of many other AI applications, requiring providers to undergo a conformity assessment will lower the regulatory burden for the hundreds of potential companies using those systems as the basis of narrower applications.

As ALLAI notes[17], general purpose systems ought to be included in the Act's scope to avoid a situation where the burden of bringing these systems into compliance with the AI Act falls entirely on 'downstream' users of the general purpose AI systems. Downstream users would otherwise be the ones required to bring the systems in line with the requirements for high risk AI. This might be an enormous burden, especially for SMEs and startups, or perhaps even prove to be technically impossible. Even in cases where the general purpose AI developer is willing to support downstream users in their efforts to comply with the AI Act, the latter are rendered fully dependent on the developers and are left without the appropriate means to seek redress when the general purpose AI system in question causes damage.

A third-party conformity assessment would guarantee to European users that these systems are accurate, robust, capable of withstanding cyberattacks and can be trusted for use.

ARTICLE 4C

NEW ARTICLE
Article 4c: Conditions for other persons to be subject to the obligations of a provider Any person who places on the market, puts into service or uses a general purpose AI system in any of the circumstances listed in Article 28 shall be considered one of the providers of the system subject to the provisions of this Regulation. The developer of the general purpose AI system will be considered the provider unless the Article 28 conditions apply.

While various stakeholders in the supply chain should have specific responsibilities, the original creators of these systems (almost exclusively large multinational companies headquartered in the US or China) should be treated as providers because they are best suited to conduct assessments of their systems. This will create lower regulatory burdens for the hundreds of potential companies using those general purpose systems as foundations for their narrower applications.

When general purpose AI systems are used in the AI supply chain, it should be possible for multiple companies to be considered providers under the regulation: the initial builder of the general-purpose system and the company or companies adapting it downstream.

17 AIA in-depth #1: Objective, Scope, Definition: https://allai.nl/wp-content/uploads/2022/03/AIA-in-depth-Objective-Scope-and-Definition.pdf