November 14th,2024

Growth and competitiveness are fueled by data

As the CEE Digital Coalition, an entity gathering leading ICT trade associations from across Central Eastern Europe, we are committed to building a competitive and innovative digital economy in Europe and a prosperous, safe, modern society in our region. To that end, we need Europe to be a front runner in the ongoing AI revolution. As rightly pointed out in a recent letter signed by dozens of companies "Europe can't afford to miss out on the widespread benefits from responsibly built open AI technologies that will accelerate economic growth and unlock progress in scientific research. For that we need harmonised, consistent, quick and clear decisions under EU data regulations that enable European data to be used in AI training for the benefit of Europeans"[1].

Not only AI is becoming increasingly important for Europe's competitiveness. But it is playing a key role in safety and local cultural heritage. CEE languages and cultures must embrace AI. For this to happen we need AI models to be trained on CEE data.

This is where the European Data Protection Board's (EDPB) opinion on the use of personal data for AI model training comes into play. The upcoming EDPB opinion on data processing for AI model training in Europe will have overarching consequences to how companies will be able to develop and deploy AI technology in Europe, within the remit of the GDPR. As the EDPB aims to take a general approach, introducing guidelines that every European company wanting to implement AI will need to follow, all European governments should follow the process closely, get informed and support a process that is enabling progressive development of AI. This crucial opinion is about the interpretation of the GDPR legal basis for AI model training. DPA's from the whole Europe can choose to interpret the GDPR in a way that supports development of AI use-cases in Europe, or in a way which puts additional burden to

---

[1] An Open Letter: Europe needs regulatory certainty on AI, available at: https://www.euneedsai.com

companies from Europe or operating in Europe. In extreme scenarios (e.g. consent required for both 3PD/1PD, impractical safeguards), the ability for the EU to develop AI technology would be severely impaired.

While we believe in the importance of protecting individuals' privacy and ensuring that their personal data is used in a responsible manner, privacy - as a fundamental right under the Charter and the GDPR - must be balanced against other fundamental rights such as freedom of expression and information, and the freedom to run a business. For this reason we believe that the use of publicly available first-party data for training large language models (LLMs) based on legitimate interest should be allowed.

Firstly, it is important to note that LLMs are not personal databases of user information. They are mathematical objects that do not store copies of the data on which they were trained. Instead, they aggregate learnings from across the training data to predict typical language patterns and logical entailments between words. As such, the legal basis needs to reflect the need for volume for both technical efficacy and data subject protection.

Secondly, the use of publicly available first-party data for training LLMs based on legitimate interest is essential for the development of AI technology in Europe. It allows companies to build accurate business decisions, optimise warehouse management, and develop innovative products that can compete with those offered by companies in other regions. Limiting the use of publicly available first-party data for training LLMs would put European companies at a disadvantage and hinder the development of AI technology in Europe.

Thirdly, the use of publicly available first-party data for training LLMs based on legitimate interest is consistent with individuals' previous decision to make their thoughts public. It is also consistent with individuals' will regarding the use of their personal data in AI training datasets through their ability to object. The focus should be on the implementation of the right to object.

In conclusion, we urge decision makers to consider the importance of a strong AI sector in Europe and for models to be trained on European data. For this to happen, we call decision makers and more specifically DPA's gathered within the EDPB to make a balanced opinion including by allowing the use of publicly available first-party data for training LLMs based on legitimate interest. This is essential for the development of AI technology in Europe, preserving the Union's cultural heritage, and ensuring that European companies remain competitive.

AAVIT - Association for Applied Research in IT (Czechia)

ANIS - Employers' Association of the Software and Services Industry (Romania)

APDETIC - Association of Producers and Dealers of ICT (Romania)

BAIT – Bulgarian Association of Information Technologies (Bulgaria)

Digital Poland Association (Poland)

DNA - Digital National Alliance (Bulgaria)

INFOBALT (Lithuania)

ICT Association of Slovenia (Slovenia)

LIKTA – Latvian Information and Communications Technology Association (Latvia)

SAPIE - Slovak Alliance for Innovation Economy (Slovakia)