Legitimate Interest: A Legal Basis for AI?

Andrea Stazi*

The **legitimate interest** is one of the six legal bases for the processing of personal data set out in Art. 6, paragraph 1, letter f of the GDPR. For processing to be lawful, three conditions must be met:

1. **Existence of a legitimate interest:** The data controller must have a genuine and legitimate interest in processing the data. In the context of Artificial Intelligence, this translates into the need to train algorithms to develop new products, improve existing services, or conduct research and development.

2. **Need for treatment:** The processing of personal data must be necessary to pursue that interest. Training AI models on large volumes of data, often not directly anonymous, is technically essential for developing high-performance systems.

3. **Balancing with the interests and rights of the data subject:** The data controller's interests must prevail over the interests, rights, and fundamental freedoms of the data subject. This is the most delicate point. **Arguing that AI training is based on legitimate interest**, without requiring explicit consent, could be based on several key principles and tools:

- **Proportionality and Data Minimization:** First, legitimate interest could be supported when the training of models takes place using techniques that ensure the **pseudonymization** or the **anonymization**, like the **differential privacy** or the use of synthetic data. This approach would reduce the privacy risk and strengthen the argument that the controller's interests do not disproportionately undermine the rights of data subjects.

- **Practical Impact of Consent:** Requiring consensus for every single dataset used for training at scale is technically and logistically

unsustainable. AI, especially **generative models**, requires billions of data points to learn language or image patterns. Obtaining specific, informed, and revocable consent for every individual whose data was used to train a model would ultimately cripple innovation.

- **Risk Assessment (DPIA):** Companies could support legitimate interest through rigorous **Data Protection Impact Assessment**. A well-conducted DPIA would not only identify privacy risks but also propose appropriate mitigation measures, demonstrating the organization's commitment to respecting data subjects' rights. If the DPIA demonstrates that the risks have been minimized and that the company's economic interest is significant, the balance could be remedied in favor of legitimate interest.

- **The Thesis of the Draghi Report:** Some of the problems with the European data protection framework have been highlighted in the **Draghi Report**. Implementation uncertainties influence business decisions and have a direct impact on Europe's profitability, research and development, and competitiveness. Legitimate interest, if interpreted more flexibly and consistently, could be one of the "therapies" for these issues.

In conclusion, arguing that legitimate interest is the appropriate legal basis for AI training does not mean ignoring data protection, but rather recognizing the importance of innovation and competitiveness. The solution is not to abandon the GDPR, but to evolve its interpretation to more effectively balance data protection with the technical and competitive demands of the AI era.

* CEO and Co-Founder Techno Polis