GDPR Stress-tests by AI: Towards a user-centric data protection approach

''For every action, there is an equal and opposite reaction''
Newton, 3rd law state

I. Introduction

We already live in the Fourth industrial revolution, an era where disruptive technologies like artificial intelligence ('AI'), connectivity and human-machine interaction[1] impact our daily lives. Especially generative AI, a type of AI that is based on deep learning systems[2] and can generate creative content as outputs[3] has grown so rapidly that McKinsey has declared 2023 as the year of generative AI[4]. While the full potential of AI will not be unleashed soon, the more sophisticated the AI models become, the more mature the market will be to realise the value this technology brings, and more companies will start integrating AI systems in their various business functions[5]. The advent of this new era where the processing of massive amounts of data[6] is now regarded as a prerequisite to take advantage of big data and unlock their full potential[7] comes with many legal implications, including personal data protection[8]. Failing to address these issues might result in a serious compromise of the key principles of the applicable General Data Protection Regulation[9] and to deprive data subjects of their enshrined right to

[1] McKinsey & Company 'What Are Industry 4.0, the Fourth Industrial Revolution, and 4IR?' (McKinsey & Company, 17 August 2022) < https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-are-industry-4-0-the-fourth-industrial-revolution-and-4ir > accessed 3 September 2023.

[2] The author would like to clarify that deep learning is a subtype of machine learning systems.

[3] IBM, 'What is Generative AI?' (IBM, 20 April 2023) < https://research.ibm.com/blog/what-is-generative-AI > accessed 20 September 2023.

[4] QuantumBlack AI 'The State of AI in 2023: Generative AI's Breakout Year ' (McKinsey & Company, August 2023) 3 < https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-in-2023-generative-ais-breakout-year#/ > last accessed 30 September 2023.

[5] According to Deloitte's survey, 76% respondents plan to invest more in AI next year, however 22% of those who already deployed AI applications were not satisfied with the outcome, see Deloitte, 'Deloitte's State of AI in the Enterprise, 5th Edition report' (Deloitte, October 2022) 6 < https://www2.deloitte.com/uk/en/pages/deloitte-analytics/articles/state-of-ai-in-the-enterprise-edition-5.html > last accessed 3 September 2023.

[6] Balazs Gati, 'Some Data Protection Issues of the EU Regulation of Artificial Intelligence' (2022) 2022 Collection Papers from Conf Org on Occasion Day Fac L 588,595.

[7] ICO, 'Big Data, AI, Machine Learning and Data Protection' (2017 version 2.2) para 11 ('ICO Big Data') < https://ico.org.uk/media/for-organisations/documents/2013559/big-data-ai-ml-and-data-protection.pdf > accessed 21 September 2023.

[8] Nikolaus Marsch, 'Artificial Intelligence and the Fundamental Right to Data Protection: Opening the Door for Technological Innovation and Innovative Protection' in T Wischmeyer and Timo Rademacher (eds), *Regulating Artificial Intelligence* (Springer 2020) 33.

[9] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data and repealing Directive 95/46/EC (General Data Protection Regulation) [2016] OJ L 119/1 (GDPR).

data protection[10], which is a prerequisite of other fundamental rights like freedom of expression and information[11].

Taking into account both the GDPR and the data protection-related provisions of the proposed AI Act[12], this paper aims to identify the main data protection challenges machine learning systems bring and to stress the importance of embracing legal design and utilizing technology to preserve privacy and ensure effective compliance. To this end Part II will start by distinguishing three main types of machine learning systems and map the data journey in machine learning models, while Part III will identify some important GDPR compliance issues arising from this data journey and critically analyse the proposed AIA through 'data protection lenses'. After building a theoretical foundation for the proposed solution in Part IV, the last Part V will then propose a way to ensure that technology can indeed facilitate data protection compliance for ML systems and enhance protection of end-users' rights.

II. The Basics: The Data Journey in Machine Learning Models

Under the Commission's proposed AIA[13], an AI system can be defined as '' any software that is developed with one or more of the techniques and approaches listed in Annex I and can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with''[14]. A subset of AI that is expressly included under Annex I of the EC's proposal is ''Machine Learning'' that ''*uses models, or algorithms, to analyze large amounts of complex data and identify patterns*''[15] ('ML'). ML can be further distinguished to 3 main different types: supervised learning, unsupervised learning and reinforcement learning depending on the

[10] Charter of Fundamental rights of the European Union [2012] OJ C326/391, ('CFR') arts 7,8.
[11] European Union Agency for Fundamental Rights, 'Getting the Future Right-Artificial Intelligence and Fundamental Rights' (2020) ('FRA') 61.
[12] European Commission, 'Proposal for a Regulation of the European Parliament and of the Council laying down harmonized rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain union legislative acts' COM (2021) 206 final ('AIA'), art 3(1).
[13] AIA, art 3(1).

[14]The author would like to highlight that this definition is not yet final, and both the EC and the EP have proposed substantial changes, see European Parliament 'Artificial Intelligence Act Amendments adopted by the European Parliament on 14 June 2023 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD))' (P9_TA(2023)0236), amendment 165 <https://www.europarl.europa.eu/doceo/document/TA-9-2023-0236_EN.pdf > accessed 20 September 2023; Council of the European Union 'Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts - General approach' (2021/0106(COD)), art 3(1) < https://www.consilium.europa.eu/en/press/press-releases/2022/12/06/artificial-intelligence-act-council-calls-for-promoting-safe-ai-that-respects-fundamental-rights/ > accessed 20 September 2023.
[15] Intel, 'Machine Learning' (Intel) <https://www.intel.com/content/www/us/en/developer/topic-technology/artificial-intelligence/training/machine-learning.html > accessed 21 September 2023.

training data and level of human intervention. The figure below, provides a compact overview of the 3 different types of ML[16].

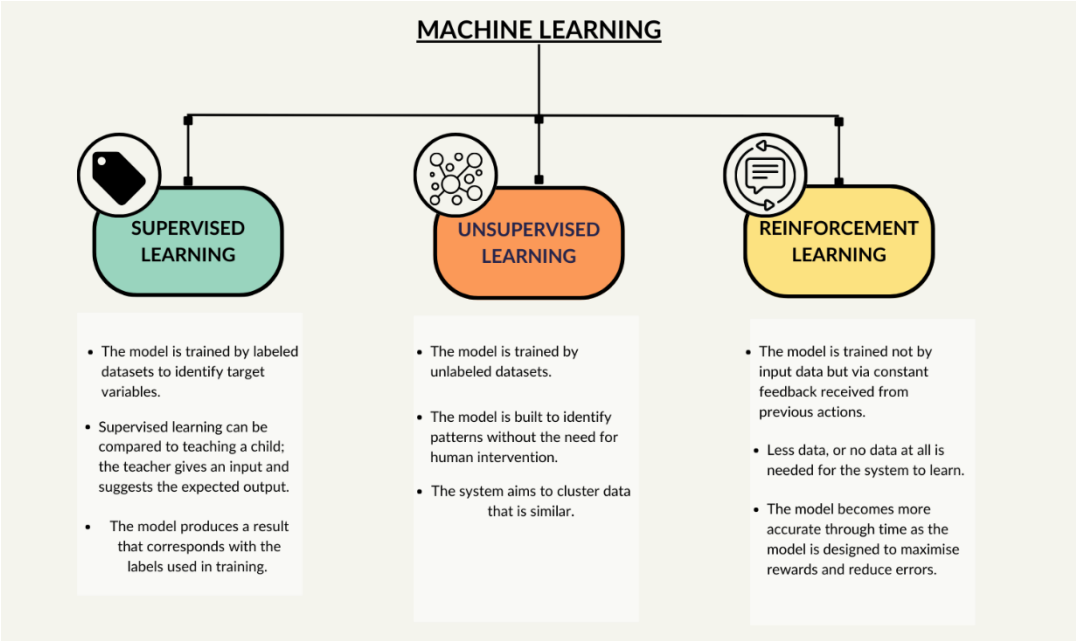

Figure 1: Types of Machine Learning Models

Although ML is not a monolithic concept[17] and as such, variations exist depending on the type of model and its application[18], in general, the following main data journey takes place in ML models, as illustrated in the graph below:

[16] A detailed analysis of the various types of ML can be found at Datatilsynet (The Norwegian Data Protection Authority), 'Artificial Intelligence and Privacy' (January 2018), 7-10 < https://www.datatilsynet.no/en/regulations-and-tools/reports-on-specific-subjects/ai-and-privacy/ > accessed 20 September 2023; IBM, 'Supervised vs. Unsupervised Learning: What's the Difference? ' (IBM, 12 March 2021) <https://www.ibm.com/blog/supervised-vs-unsupervised-learning/ > accessed 21 September 2023; ICO Big Data (n 7) 7-8.

[17] Thomas Wischmeyer, 'Artificial Intelligence and Transparency: Opening the Black Box' in Thomas Wischmeyer and Timo Rademacher (eds), Regulating Artificial Intelligence', (Springer, 2020) 81.

[18] For a specific illustration of the data journey per each type of AI, see Datatilsynet (n 16).

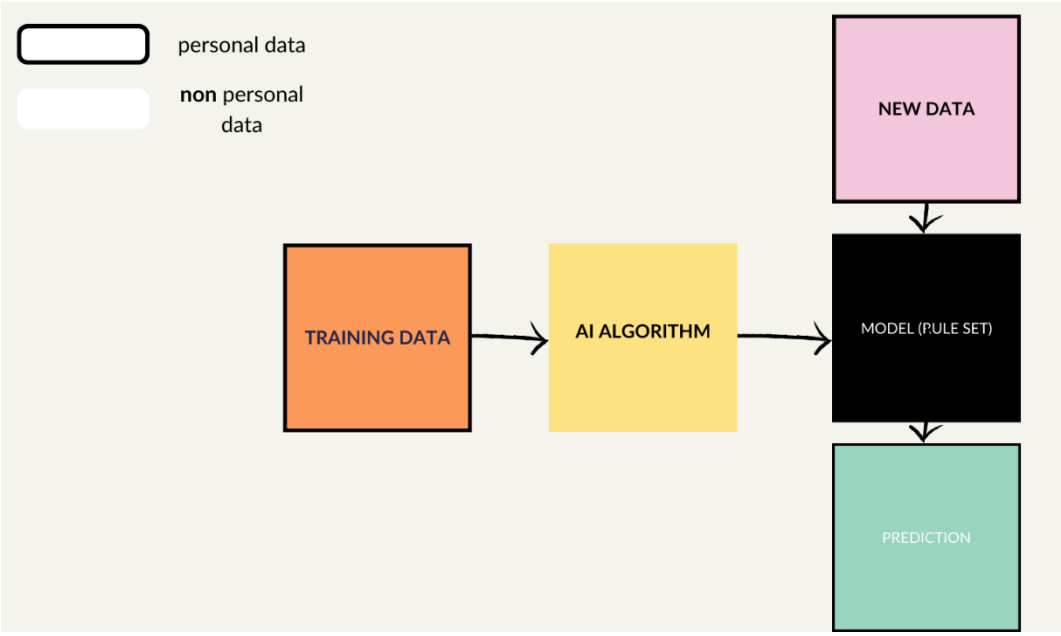

Figure 2: The Data journey in ML models[19]

Initially, training data is used, and a chosen AI algorithm is optimized with a set of human-defined objectives. The algorithm is then ''trained'' with this training data and a model is generated. Every time this model is deployed, new data is inserted that are similar to the training data. The model identifies the pattern and produces the requested outcome/prediction. From the above, it is obvious that all ML systems are relying on data and in fact the precision of the ML-based applications depends on both the volume of training data and its quality[20]. We now start living in a ''global interconnected data-processing infrastructure''[21] in which AI is a prerequisite to exploit big data, but the data protection challenges raised from the use of AI systems require an innovative way to ensure effective data protection.

III. The Data Protection Challenges

A. Are all this data personal data?

Although massive amounts of data is required in every stage of the above shown data journey, it is important to clarify that not all this data will fall into the definition of ''personal data'' under the GDPR[22], while in some cases even the outcomes of the ML tool would not even include personal data (e.g. weather predictions). However, in light of the broad interpretation of what constitutes personal data[23]

[19] Tiago Sergio Cabral, 'Forgetful AI: AI and the Right to Erasure under the GDPR' (2020) 6 Eur Data Prot L Rev 378, 387.

[20] Datatilsynet (n 6) 11.
[21] European Parliament, 'Artificial Intelligence: Challenges for EU Citizens and Consumers' (January 2019) < https://www.europarl.europa.eu/RegData/etudes/BRIE/2019/631043/IPOL_BRI(2019)631043_EN.pdf > accessed 20 September 2023.
[22] GDPR, art 4 (1).
[23] Case C-434/16 Peter Nowak [2017] EU:C:2017:582, para 34, note that even opinions related to a person are personal data.

and easily identifiable information[24] careful analysis before reaching this conclusion is necessary[25]. According to the GDPR, personal data is ''any information related to an identified or identifiable natural person…''[26] while pdeusonymized data fall are still considered personal data[27] and even the anonymization of personal data constitutes ''processing '' of personal data[28]. But even when correct anonymization has been achieved, the risk of reidentification arises[29] as large datasets are processed in the training stage to increase efficacy[30]. What is more, recital 26 of the GDPR further clarifies that the available technology at the time of processing shall be considered when assessing if the personal data can easily identify a data subject. It is thus logical to assume, that the more AI advances, the easier and quicker it could become to identify data subjects, making the GDPR applicable in more processing activities in the future[31]. In fact, since 2019 ML systems already facilitate re-identification attempts through other datasets[32], while differential privacy techniques proved inadequate to ensure anonymization, as now ML systems can re-identify data subjects even without processing any personal data[33]. Therefore, at least one stage of the data journey will most likely be subject to the GDPR, raising compliance issues as the next chapter will analyse.

B. **GDPR stress tests for AI**

The GDPR is one of the most influencing legislation at a global level[34], and it is now considered the benchmark for any data protection legislation. Although the impact of this legislation is significant, its long-term success depends to a great extent on whether it can ensure effective compliance even in privacy-intrusive technologies like AI. Almost every key principle set out under article 5 of the GDPR is challenged by AI[35]. However, this paper aims to analyse and address the following pain points whenever an ML-based application is being developed for commercial use:

---

[24] CJEU, Patrick Breyer v Bundesrepublik Deutschland,C582/14, paras 43,49; see also Jeffrey Bholasing, 'How Technological Advances in the Big Data Era Make It Impossible to Define the 'Personal' in GDPR's 'Personal Data'' (2022) 8 Eur Data Prot L Rev 346, 349-351.

[25] Bholasing (n 24) 357-358.

[26] GDPR, art 4(1).

[27] GDPR, art 4(5).

[28] ICO, 'Introduction to Anonymisation' (ICO, May 2021) 12 < https://ico.org.uk/media/about-the-ico/consultations/2619862/anonymisation-intro-and-first-chapter.pdf > accessed 30 September 2023.

[29] Arvind Narayanan and Vitaly Shmatikov, 'Myths and Fallacies of ''Personally Identifiable Information'', (2010) 53(6) Commun ACM 24, 26; But see also contra Select Committee on Artificial Intelligence, 'AI in the UK: Ready, Willing and Able?' (HL 2017-19,100) 31.

[30] EDPS, 'Opinion 7/2015 Meeting the Challenges of Big Data' (19 November 2015) 15; W Nicholson Price II, 'Black-box Medicine' (2015) 28(2) Harvard Journal of Law & Technology 420, 423.

[31] AEDP and EDPS, '10 Misunderstandings Related to Anonymisation' (27 April 2021), emphasis added on misunderstanding 4 < https://edps.europa.eu/data-protection/our-work/publications/papers/aepd-edps-joint-paper-10-misunderstandings-related_en > accessed 20 September 2023; Bholasing (n 24) 348.

[32] The Center for Open Data Enterprise, 'Balancing Privacy with Health Data Access' (September 2019) 9 < http://reports.opendataenterprise.org/RT2-Privacy-Report-Final.pdf> accessed 20 September 2023.

[33] Jayanth Kancherla, 'Re-identification of Health Data through Machine Learning' (30 November 2020) 5 emphasis added on studies analysed < http://dx.doi.org/10.2139/ssrn.3794927 > accessed 20 September 2023.

[34] Gati (n 6) 594.

[35] Marsh (n 8) 33, 36.

Identification of roles: The correct identification of which natural or legal person is a data controller and which acts as a data processor[36] is fundamental to ensure compliance under the GDPR. Although the EDPB has already provided guidelines on the concepts of controller and processor[37] these are functional concepts that need to reflect the actual roles of each party ad hoc[38]. The situation could be quite straightforward if a company decides to build an in-house ML tool to enhance its operations and use its own training datasets. However, there is a high chance that companies will start contracting with third-parties that specialize in developing AI tools and this is where things get blurry. As already explained in chapter II, the AI developer will decide on the input data and will set out the rules for the model. However, under the GDPR the data controller and not the processor is the one that determines the purposes and means of processing, making it unclear which party has the role of the controller[39]. The situation becomes even more complex if one considers that the training datasets could be owned by third-parties and shared to other controllers via data sharing agreements. In this case, the (joint) data controllers might be different depending on the timing of each processing activity. This is particularly problematic considering that under the GDPR, the contact points of each data controller must be available to the data subjects[40].

Lawfulness: Under the lawfulness principle, each processing activity requires a valid legal basis, from the ones expressly mentioned under arts 6 and 9 GDPR respectively. However, as already indicated in figure 2, at least 2 different processing activities take place in ML systems, thus two different legal basis shall exist; a legal basis for the processing of training personal data and another one for each new personal data input. Especially for the former, the lawfulness principle is usually violated, as most of the training data have been initially collected for other purposes[41]. Assuming that in many cases no processing of special categories of data[42] will take place[43], two possible legal basis[44] are available: the data subjects' consent and the legitimate interests of the data controller or of a third party[45], each of which raise the following problems:

a) Consent offers the advantage that it can be selected as a legal basis in various types of processing[46]. However, according to article 4(11) and 7 of the GDPR, multiple requirements should be met for a lawful consent. The basic elements of these are that it has to be freely

[36] See definitions under GDPR, art 4(7) and (8) respectively.
[37] EDPB, 'Guidelines 7/2020 on the Concepts of Controller and Processor in the GDPR' (7 July 2021, v. 2.1) < https://edpb.europa.eu/system/files/2021-07/eppb_guidelines_202007_controllerprocessor_final_en.pdf > accessed 21 September 2023.
[38] ibid, 9, para 12.
[39] ICO Big Data (n 7) 56-57; see also argument in favor of an actual joint controllership relationship in such cases in Julia Powels and Hal Hodson, 'Google DeepMind and Healthcare in an Age of Algorithms' (2017) Health Technol (7), 351–367, 358.
[40] GDPR, arts 13-14 (a) and (b).
[41] Marsch (n 8)36.
[42] As indirectly defined under GDPR, art 9(1).
[43] The author would like to point out that even if sensitive data will be processed, a lawful basis from the ones available in art. 6 GDPR is also necessary, besides the additional legal basis under art 9 GDPR, so even in this scenario the same problems arise.
[44] The author would like to clarify that the rest legal bases provided under art 6 of the GDPR would be unlikely to be valid in the context of ML models, see also Article 29 Working Party, ' Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679' (3 October 2017) 12-14; ICO Big Data (n 7) 29, para 55.
[45] GDPR, art 6 (1)(a) or (f).
[46] GDPR, arts 8(1), 9(2)(a), 22(2)(c) and49(1)(a) though explicit consent is needed for the last three arts.

given[47], specific, informed and unambiguous[48]. Though in theory at least all these requirements would empower data subjects, the reality is different. Considering that the fulfilment of the above general criteria has to be decided on a case by case[49] basis, controllers can never be certain[50] of the maturity and tech literacy of each data subject they request their consent from, thus of the consent's validity. Jones and Edenberg called this a ''consent crisis'' and claimed that consent is not functioning well, but merely creates the illusion of choice, without really providing sufficient information to data subjects or achieving the necessary understanding of their action. They identified four main issues : too many policies, lengthy and confusing terms, inability to assess the severity of harm consent might cause and limited alternative choices[51].

In addition, consent has to be easily withdrawn[52]and be written in 'clear and plain language'[53], while the burden of proof relies on the controller[54], due to the principle of accountability[55]. The data controller has thus to take proper organizational measures[56] and constantly be aware of the data subjects who have withdrawn their consent and stop processing their data, as this will no longer be lawful[57]. Especially for unsupervised algorithms that continue to learn through time this is exceptionally problematic, as the improvement of the model is highly dependent on the data subject's decision to withdraw their consent or not[58].

b) Legitimate interests can play an important role in digital innovation, as it is the only legal basis under which the fundamental right to conduct business[59] might override data protection rights. To ensure this will be a valid legal basis, three main criteria under 6(1)(f) GDPR shall be met: the controller has to a) pursue legitimate interests, b) the processing has to be necessary for the pursued interests[60] and c) a balancing test is required, as the

---

[47] GDPR, recitals 42-43 and art 7(4); EDPB, 'Guidelines 05/2020 on Consent under Regulation 2016/679' [2020] (version 1.1) 7-19 < https://edpb.europa.eu/our-work-tools/our-documents/guidelines/guidelines-052020-consent-under-regulation-2016679_en > accessed 24 September 2023.

[48] GDPR Recitals 32,42; Case C-92/09 Volker und Markus Schecke GbR v Land Hessen [2010] ECR I-11063, Opinion of AG Sharpston, para 79 ; Case C-673/17 Planet49 GmbH v Bundesverband der Verbraucherzentralen und Verbraucherverbände – Verbraucherzentrale Bundesverband e.V. [2019] published in the electronic Reports of Cases (Court Reports - general), Opinion of AG Szpunar, paras 66,68,112.

[49] Planet 49 (n 48) paras 97,99.

[50] Eleni Kosta, *Consent in European Data Protection Law*, (Martinus Nijhoff Publishers 2013) 209 emphasis added on ''it is practically impossible to come up with detailed instructions on what information should be provided… ''.

[51] Meg Leta Jones and Elizabeth Edenberg 'Troubleshooting AI and Consent ' in Markus D Dubber, Frank Pasquale and Sunit Das (eds), *The Oxford Handbook of Ethics of AI* (OUP 2020) 359-361.

[52] GDPR, art 7(3).

[53] ibid art 7(2).

[54] ibid art 7(1); Case C-61/19 Orange România SA v Autoritatea Naţională de Supraveghere a Prelucrării Datelor cu Caracter Personal (ANSPDCP) [2020] ECLI:EU:C:2020:158 Opinion of AG Szpunar, para 54.

[55] GDPR, art 5 (2).

[56] ibid, art 24(1).

[57] ibid art 6(1).

[58] Matthew Humerick, 'Taking AI Personally: How the E.U. Must Learn to Balance the Interests of Personal Data Privacy & Artificial Intelligence' (2018) 34(4) Santa Clara High Tech. L.J. 393, 406.

[59] CFR, art 16 (1).

[60] ICO, 'Legitimate Interests' < https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/lawful-basis-for-processing/legitimate-interests/ > accessed 07 September 2023.

interests of the controller should not override interests or fundamental rights of the data subjects[61]. Although 'interests' is a broad term that includes both private and public interest, recital 47 GDPR narrows the scope as legitimate interests among others include 'a relevant and appropriate relationship between the data subject and the controller'. In the context of AI though, two main issues arise. Firstly, at the time of collection of the personal data, the data controller has to name the legitimate interests it relies upon in the privacy notice[62], in light of the principle of purpose limitation. The data controller's legitimate interests have therefore to be known before this processing takes place. If no specific mentioning of data processing for training AI models is expressly included in the privacy notice, then the data subjects would hardly expect this processing to take place[63]. What is more, under art 21(1) GDPR the data subject has the right to object to any processing that relies on the legitimate interests, which if exercised can also trigger the exercise of the right to restrict the processing of its personal data[64] until the controller demonstrates that its legitimate interests override those of data subjects. Therefore legitimate interests face similar issues with consent withdrawal, and proper organisational measures must be implemented to promptly detect such requests and restrict processing.

Transparency:

Regardless of which legal basis is selected, the key principle of transparency inter alia[65] requires data controllers to provide data subjects at the time they obtain this personal data with strictly enumerated information[66]. This information which must be available in 'clear and plain language'[67] aim to give data subjects an overview of who is the data controller, what types of personal data will be processed, why etc. This in practice requires a privacy notice, either at the company's website or via a link within the application. In the context of AI three main issues arise:

A) Under the EDPS' opinion, the information shall refer to both observed (training data or data input) and inferred data (data derived from the outcome/prediction)[68], while the logic behind the chosen algorithm must also be disclosed. Data controllers have thus to find innovative ways to explain complex processing activities and the logic behind AI models in a way that is easy for data subjects to understand.

B) When it comes to training data, usually personal data will not be obtained directly by the data subjects but by third parties who own the datasets and share them with other data controllers. In some cases the processed personal data may not include contact details of data subjects. According to art 14 GDPR, if the disclosure of information requires disproportionate effort by the data controller, then the obligation to provide this information is not mandatory, but a publicly available privacy notice will suffice. Data subjects at the end

[61] Case C-13/16, Valsts policijas Rīgas reģiona pārvaldes Kārtības policijas pārvalde v Rīgas pašvaldības SIA [2017] ECLI:EU:C:2017:336, para 28.

[62] GDPR, art 13(1)(d) and 14(2)(b).

[63] Article 29 WP, ' Opinion 06/2014 on the Notion of Legitimate Interests of the Data Controller under Article 7 of Directive 95/46/EC' (9 April 2014) 40, point iv.

[64] GDPR, art 18 (1)(d).

[65] Ex-post transparency about the outcome of the AI application is out of this paper's scope.

[66] GDPR, arts 13-14.

[67] ibid, art 12(1).

[68] EDPS (n 30) 4.

might not even be aware that their personal data is processed for training AI models or not, continuing the 'information imbalance' that exists already.

C) According to FRA's report only 22% of consumers read the terms and conditions (including privacy notices[69]), while 27% of those who actually read them do not understand them[70]. A static privacy notice that is updated on an annual basis[71] will hardly be able to fully reflect all dynamic processing activities that take place at the time the data subjects seek this information, let alone to properly identify and address understanding issues.

It is important to note that besides violating the GDPR, the lack of transparency also decreases consumer's trust, and companies might risk losing a competitive advantage[72].

Purpose limitation: Under the GDPR, all personal data must be collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes[73]. Purpose limitation is crucial, as it allows data subjects to continue having control over the usage of their personal data by data controllers[74]. Considering that personal data could be initially processed for a different reason other than training an ML system[75], let alone by a different initial data controller[76], the risk of incompliance with this principle is high, as data subjects did not reasonably expected that their personal data will be further processed for training algorithms[77]. Although article 6 (4) of the GDPR allows the further processing of personal data for other purposes than the initial one, the compatibility test that is required[78] under this exception and the possible consequence that data subjects may eventually lose control over which data controllers can process their data and for which reason, does not leave big room for interpretation[79]. In an example where a company collects personal data to fulfil a contract, the further processing of this data to train algorithms or as new data input is problematic, because the distance between the purposes of collection and the purposes of further processing is far[80]. In fact, since article 6(4) GDPR is an exception of the purpose limitation principle, it should be interpreted narrowly[81]. If the new purpose seems incompatible

[69] ICO Big Data (n 7) 62-63.
[70] FRA, 'Your Rights Matter: Data Protection and Privacy' (2020) 9 < https://fra.europa.eu/en/publication/2020/fundamental-rights-survey-data-protection > accessed 30 September 2023.
[71] Although updates to privacy notices are possible, in practice this requires a lengthy internal approval process so the privacy notices cannot easily reflect the dynamic nature of all currently processing activities.
[72] ICO Big Data (n 7) 27, para 53.
[73] GDPR, art 5 (1)(b).
[74] Datatilsynet (n 16)16.
[75] European Parliamentary Research Service, ' The Impact of the General Data Protection Regulation (GDPR) on Artificial Intelligence' (EPRS, June 2020) 45 < https://www.europarl.europa.eu/RegData/etudes/STUD/2020/641530/EPRS_STU(2020)641530_EN.pdf> accessed 29 September 2023.
[76] ICO Big Data (n 7) para 18.
[77] Bholasing (n 24) 357.
[78] For an analysis of the compatibility test see Article 29 WP, ' Opinion 3/2013 on Purpose Limitation' (April 2013), 23-26 < https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2013/wp203_en.pdf> accessed 30 September 2023.
[79] The author would like to note that recital 159 of the GDPR calls for a broad interpretation of scientific research, however it takes the view that commercial use of an ML system cannot be considered privately funded research.
[80] Article 29 WP (n 78) 24.
[81] Datatilsynet (n 16) 17; But see contra Article 29 WP that suggests that art 6(1)b is not an exception but a specification, thus flexible interpretation would be possible, Article 29 WP, (n 78) 13.

with the initial one, the further processing must be relied on a new legal basis and a relevant change to the privacy notice must be made and provided promptly to data subjects[82]. In any case, generic reference to future purposes like ''future research'' or ''improving users experience'' most likely will not meet the level of specificity required[83].

Storage limitation and right to erasure: AI also raises substantial issues with regards to the storage limitation principle and the subsequent right to erasure[84]. Assuming that the initial processing of personal data either for training ML models (training data) or for using an ML model to generate a desired outcome ('new data input') was lawful, the question remains as to what extend both training personal data and each new personal data input can be kept and for how long. The storage limitation principle allows personal data to be kept for longer than necessary but only for scientific research purposes or statistical ones. Although one can argue that if a university or research institute tries to develop an ML model this exception would apply, the minute this ML model is deployed for commercial use, this exception would no longer apply.[85] For models that are static and stop learning from each new data input, the storage limitation principle hardly raises any issues, as the training data can and should be deleted. However, unsupervised models continue to learn through time, making the lines between research development and actual commercial use to blur[86]. A strict interpretation of this principle could thus stifle innovation, as unsupervised algorithms will not be able to improve through time[87]. A limited right to erasure[88] could perhaps be a viable solution, but data subjects need to receive a warning that informs them in advance that the processing of their personal data with this particular ML model will result only in a limited, ex nunc right to erasure.[89]

[82] GDPR, arts 13(3) and 14(4).
[83] Article 29 WP (n 78) 16.
[84] GDPR, arts 5(1)(e) and 17.
[85] See also CEDPO AI Working Group, 'AI and Personal Data:A Guide for DPOs 'Frequently Asked Questions'' (12 June 2023) 9.
[86] Datatilsynet (n 16) 18.
[87] Humerick (n 58), 408.
[88] ibid, 416.
[89] Christina Varytimidou, 'Looking through Black Boxes in Medical Diagnosis: Is the Upcoming Three-Dimensional European Regulatory Framework Ready, Willing and Able?' (2022) 6 EHPL 24, 34-35.

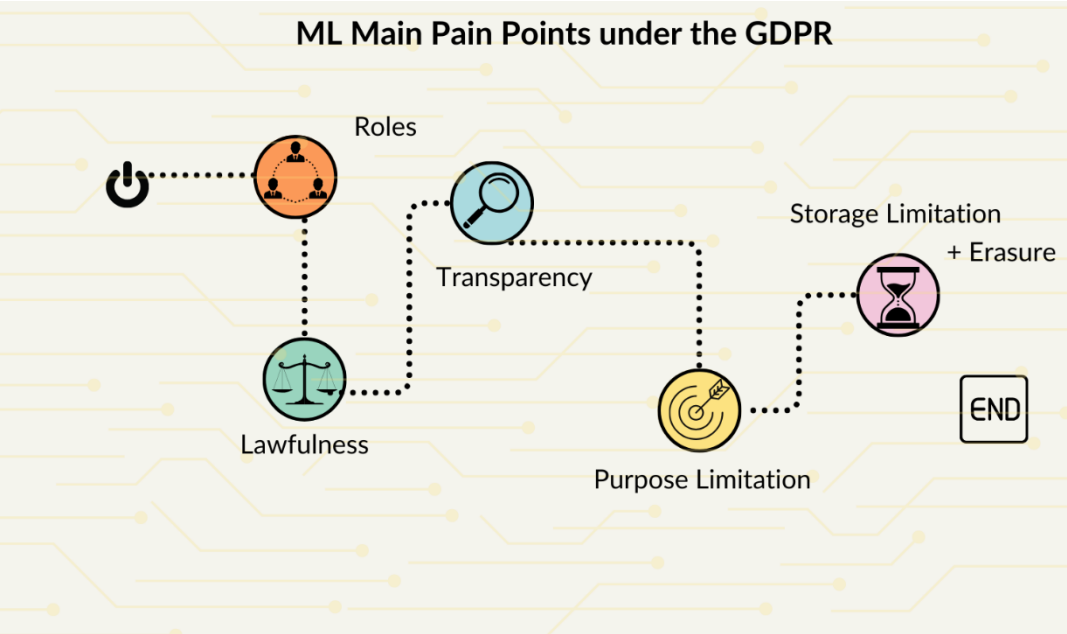

Figure 3: Illustration of the main pain points under the GDPR explained in this chapter.

C. AIA through data protection lens

Unlike the GDPR that took a tech-neutral approach[90] the AIA chose to address horizontally the risks of AI and is inspired by both the GDPR[91] and by the new legislative framework[92] on products (medical devices, machinery, etc.)[93]. The AIA adopts GDPR's risk-based approach and distinguishes between 4 levels of risk: 1) unacceptable, 2)high-risk, 3)low-risk and 4)minimal risk[94], each of them with different obligations and restrictions. As far as data protection is concerned, the following new obligations will supplement the GDPR:

a) whenever data subjects interact with AI, they need to be informed in advance thereof unless this is obvious[95];

b) for high risk systems[96] technical documentation that includes inter alia the ''general logic of the AI system'', the training datasets used and their origin, and the ''foreseeable unintended outcomes… to fundamental rights…'' of the intended use of the AI system shall be available before the placement of the market[97]. This information will include any other information required under the NLF, but the mandatory information under the GDPR is not expressly included in it.

---

[90] GDPR, recital 15.
[91] Vera Lucia Raposo, 'Ex Machina: Preliminary Critical Assessment of the European Draft Act on Artificial Intelligence' (2022) 30 International Journal of Law and Information Technology 88,89.
[92] Listed in AIA, Annex II ('NLF').
[93] AIA, 4.
[94] AIA, arts 5,6,52; for a compact overview see European Commission, 'Regulatory Framework Proposal on Artificial Intelligence', emphasis added on pyramid of criticality <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai > accessed 29 September 2023.
[95] AIA, art 52 (1).
[96] AIA, Annex III.
[97] AIA, art 11 and Annex IV s 2(b) (d)and s 3.

c) the contact details of the AI provider, but not of the user[98].

Unfortunately, the AIA failed to solve any of the abovementioned data protection problems and in fact it will result in more information overload for data subjects, without being enshrined any new rights[99] or have anyone besides the DPO to complain to[100]; the contact details of the AI provider will be available, but in most cases the data controller will act as a user, under AIA's definition[101]. As AEPD correctly identified, transparency under AIA involves, different stakeholders, and is intended for different recipients[102].

IV. The Answer to the Machine is in Humans

A. Lessig Revisited

From the above it is obvious that the advent of AI acts as a magnifying glass of already existed compliance issues under the GDPR. To resolve this, this paper argues that we should not just focus on the legislation itself to solve everything, but rather acknowledge and take advantage of the rest three modalities of regulation. Specifically, Lessig identified four modalities that regulate each individual; The Law, the market, the embedded norms and the architecture. All forces interact with each other and affect the rest[103] to regulate effectively the pathetic 'dot' that lies in the centre, which represents every individual. Murray went a step further and suggested that this 'dot' is not that pathetic, but it is an active part of the regulatory process and interacts with the rest modalities[104]. In light of Murray's approach, this paper suggests that especially for AI, the 'dot' does not just affect the rest modalities. All the rest modalities are first and foremost completely dependent on this 'dot', ergo humans.

According to Bryson ''*Artificial intelligence only occurs by and with design*''[105]. AI is produced intentionally to achieve a specific, human-defined purpose[106] and humans make design choices in each development stage. Humans developed the algorithms, chose the training data, set the rules for the model and decided for which purposes each AI model will be used for[107]. Thus, if we accept that humans can exercise substantial influence on the rest modalities, then the root of all the current problems -and their solution- starts from each and every one of us. We therefore need to realise the control and responsibility we have for the development of AI and exploit it to ensure effective data protection.

---

[98] AIA, art 13 (3)(a); See definitions under AIA, art 3(2) and(4).

[99] EDPB-EDPS, 'Joint Opinion 5/2021 on the Proposal for a Regulation of the European Parliament and of the Council Laying down Harmonised Rules on Artificial Intelligence' (June 2021), para 18.

[100] Varytimidou (n 89),40.

[101] EDPB-EDPS (n 99), para 20.

[102] AEPD (Spanish Data Protection Authority), 'Inteligencia Artificial: Transparencia' (20 September 2023) <https://www.aepd.es/prensa-y-comunicacion/blog/inteligencia-artificial-transparencia?mkt_tok=MTM4LUVaTS0wNDIAAAGOUydDjJzoET3dgoOoE8o5wjZ6xmlVqlp_SdN8eL0u5MZXwrcUaaBivI8HxUvFgunO-5vyxsmHnhKW23UGFGzA07hAdcO9DiXczIto63ovV4Mn > accessed 3O September 2023.

[103] Lawrence Lessig, *Code Version 2.0,* (Basic Books 2006)122-123.

[104] Andrew Murray, *Information Technology Law: The Law and Society* (4th edn, OUP 2019) 66-67.

[105] Joanna J Bryson 'The Artificial Intelligence of the Ethics of Artificial Intelligence: An Introductory Overview for Law and Regulation' in Markus D. Dubber, Frank Pasquale, Sunit Das, *Oxford Handbook of Ethics of AI* (OUP, 2020) 3, 6.

[106] AIR, art 3(1), emphasis added on ''for a given set of human-defined objectives''.

[107] On the human centricity of algorithms see Jack M Balkin, 'The Three Laws of Robotics in the Age of Big Data'(2017) Ohio State Law Journal (78) 12.

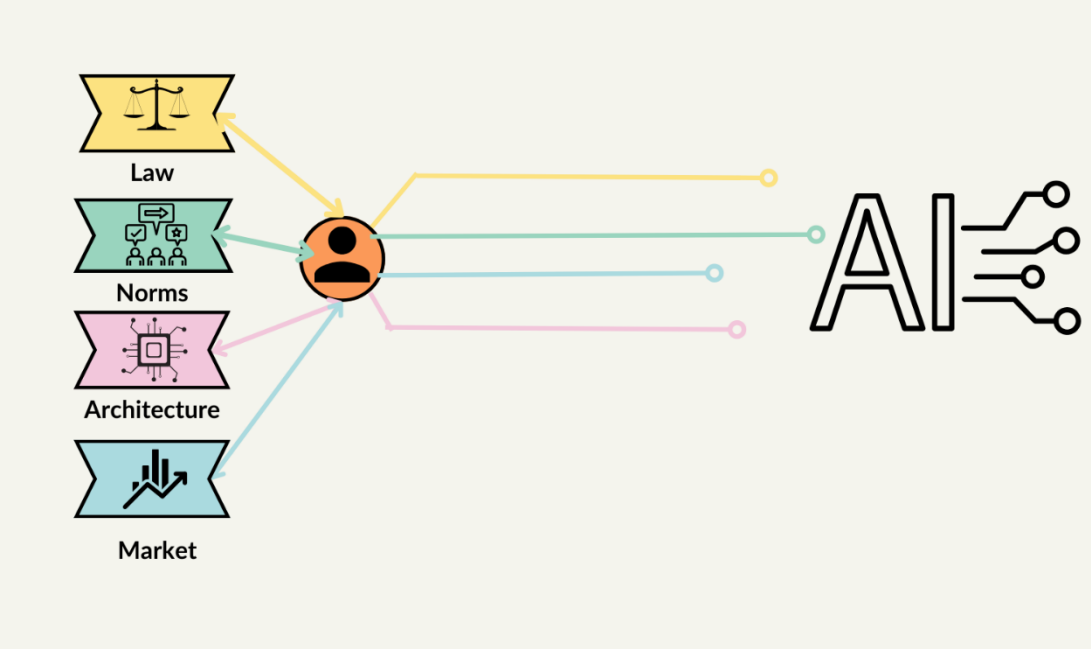

Figure 4: Lessig revisited

B) Legal design as a humanity embracer tool

After identifying that besides legislation there are also other forces we can exploit for our benefit[108], the next step is to start identifying the solutions that could actually solve the above-analysed data protection challenges. In this effort, legal design can play an important role. Legal design is a human-centered approach that intends to solve legal issues by focusing on the people who have to face these issues and aims to find a solution that suits their actual needs.[109] The aftermath of legal design is that it enhances proactive legal care, as it mostly focuses on preventing issues to arise[110]. Since AI raises more complex issues, the need for proactive legal care grows[111]. There are three layers of preventive law: The primary cause prevents the cause of the harm to arise, the secondary prevents the cause from doing harm and the third aims to mitigate damages[112]. This paper focuses on the first layer and aims to prevent the problems from arising. The main problem the next chapter will try to resolve via legal design is how to find ways to ensure data subjects are provided with meaningful information and are not losing control when ML systems are deployed. The choice to resolve this issue first lies on its importance and impact it has for the rest identified data protection challenges. Almost all data protection issues require adequate

[108] This paper will mostly focus on architecture as a solution, however embedding data protection norms in ML systems and incentivizing research on PETS could also be simultaneously explored, see Karen Yeung, Andrew Howes and Ganna Progrebna, 'AI Governance by Human Rights- Centered Design, Deliberation and Oversight: An End to Ethics Washing' in Markus D. Dubber, Frank Pasquale, Sunit Das, *Oxford Handbook of Ethics of AI,* (OUP 2020) 87; Committee on Science and technology, 'H.R.4755 - the Privacy Enhancing Technology Research Act' (20 July 2023) < https://republicans-science.house.gov/2023/7/h-r-4755-to-support-research-on-privacy-enhancing-technologies-and-promote-responsible-data-use-and-for-other-purposes > accessed 30 September 2023.

[109] Ashley Treni and Georges Clement, 'Co-designing digital tools for 21st-century tenant organizing' in Marcelo Corrales Compagnucci et al (eds), *Legal Design Integrating Business, Design and Legal Thinking with Technology,* (Edward Elgar Publishing 2021) 134-135**.**

[110] Helena Haapio, Thomas D Barton and Marcelo Corrales Compagnucci, ' Legal Design for the Common Good: Proactive Legal Care by Design' in Marcelo Corrales Compagnucci et al (eds), *Legal Design Integrating Business, Design and Legal Thinking with Technology*, (Edward Elgar Publishing 2021) 56,64.

[111] ibid.

[112] Ibid 61, emphasis added on ''the three domains of prevention'' pyramid.

and meaningful information to: a) ensure data subjects know who is the data controller in each data journey stage, b) to consent properly or to be aware of their right to objection, c) to have an overview of the data processing activities that will take place in the context of an ML model, and d) be informed that their data to erasure might be limited[113].

The GDPR already embraced the power of design[114], and obliges data controllers to take measures ''…which are designed to implement data protection measures…'' and suggested visualization to enhance understanding[115], while Article 29 WP also suggested ''interactive techniques to aid algorithmic transparency''[116]. By wearing 'legal design lens' to analyse this obligation, we can identify two main practical issues we need to resolve: a) the myth of an average data subject b) that dynamic technologies like AI, demand dynamic transparency techniques[117].

a) Legal design focuses on the end user needs. However, depending on the AI application, different types of data subjects will need to be informed. The well-established myth of an average consumer[118] who is reasonably well -informed and observant, needs to be demolished as no one-size fits all approach for privacy notices would never be effective[119]. To merely present all necessary information under arts 12-14 GDPR in a standardized privacy notice would exclude a large amount of data subjects from understanding and would lead to an invalid legal basis and increase information asymmetry[120]. Instead, we need to aim at a 'usable transparency'[121]' understand what is the maturity of the data subjects and present the information accordingly. Though ideally the privacy notice should be tailored to each data subject[122], considering the cost this would have for data controllers[123], a more cost-effective and standardized solution would be to create 4-5 different personas of data subjects that will most likely be subject to the particular processing activity[124], identify their abilities and informational needs, and create different notices per each type[125]. This would practically require from data subjects to first reply in one or

[113] As proposed under Part B.
[114] The 'founder' of data protection by design was Cavoukian, Anna Cavoukian, 'Privacy by Design: the Definitive Workshop. A foreword by Ann Cavoukian', (2010) IDIS (3), 247.
[115] GDPR, art 25 (1), recitals 58,60. ICO Big Data (n 7) 62.
[116] Article 29 WP (n 44) 31.
[117] N Diakopoulos, 'Accountability, Transparency and Algorithms', in Markus D Dubber, Frank Pasquale and Sunit Das(eds), *Oxford Handbook of Ethics of AI* (OUP 2020) 208; this is similar to a KYC exercise used successfully in marketing that helps salespeople identify target audiences and needs, see Damian Hodgson 'Know Your Customer: Marketing, Governmentality and the 'New Consumer' of Financial Services' (2002) Management Decision (40)3.
[118] C-26/13 Árpád Kásler and Hajnalka Káslerné Rábai v OTP Jelzálogbank Zrt [2014]ECLI:EU:C:2014:282, paras 73-75; C-430/17 Walbusch Walter Busch GmbH & Co. KG v Zentrale zur Bekämpfung unlauteren Wettbewerbs Frankfurt am Main eV [2019] ECLI:EU:C:2019:47,para 39; The notion of the average data subject is similar to that of the consumer, see Gianclaudio Malgieri, *Vulnerability and Data Protection Law* (OUP 2023) 40.
[119] Joasia Luzak, 'Tailor-made Consumer Protection: Personalisation's Impact on the Granularity of Consumer Information' in Marcelo Corrales Compagnucci et al (eds), *Legal Design Integrating Business, Design and Legal Thinking with Technology*, (Edward Elgar Publishing 2021), 107-109.
[120] ibid, 123.
[121] Diakopoulos (n 117) 204.
[122] Through time, a tailored privacy notice might be possible through generative AI tools, by asking data subjects if they understood the information or if they need either more details or more simple wording to understand the received information.
[123] Luzak (n 119) 123.
[124] The author would like to highlight that depending on the stage of the data journey, different types of data subjects might be identified.
[125] Diakopoulos (n 117) 204 emphasis added on ''…*different presentations of transparency information can be produced for different audiences*…''.

two questions to identify to which persona they belong to in order to provide them with the relevant privacy notice that was drafted for this specific category taking into account among others their age, maturity and tech literacy. This is aligned with Article 29 WP that demanded clarity on the data processing purposes for everyone, regardless of their linguistic/understanding abilities or special needs[126]. Especially for children and people with disabilities, UX designers must collaborate with DPOs to make different types of user experiences that are useful to each one[127].

b) Besides tailoring information per each type of data subject, the information must take into account the variations of AI models and the likelihood that substantial changes might occur in the future. The former would require adjustments to privacy notices per each ML type, the latter would require a way to ensure almost real-time amendments to privacy notices whenever the data processing activities change substantially[128]. Agility across time is also necessary when it comes to children data. As the child grows, so does its maturity and level of understanding[129], so if children data are being processed, then the privacy notice must gradually change to provide children with more information and details as they grow.

V. Technology as an enabler

While in theory tailored privacy notices for each type of data subject might solve some key transparency issues, this will not become a reality unless we utilize technology for our benefit. Even since 2015, the EDPS proposed as a solution ''personal data spaces'' that will provide a place to store and share real-time collected personal data with other parties to actively allow data subjects to participate in the data sharing process [130] and manage their data. The recently proposed regulation on financial data access[131] takes a step further and obliges all relevant data holders[132] like credit institutions to provide permission dashboards[133] to customers[134] that not only provide the customer with an overview, but also allows it to withdraw any given permission. If changes are made to a given permission, then the data holder shall try to inform the customer in real-time, while the user interface must make it easy for customers to locate the permissions[135].

A similar solution could also resolve the basic data protection challenges AI raise in the context of data protection. A real-time data processing dashboard, that is adapted to each different type of data subject could firstly include the relevant privacy notices, but also the rest technical documentation required[136],

---

[126] Article 29 WP (n 78) 17.
[127] Yeung et al (n 108) 77, 98.
[128] See also Varytimidou (n 89) that proposed agility across time, 28.
[129] Article 29 WP, 'Opinion 2/2009 on the Protection of Children's Personal Data', (February 2009) 4,6,10.
[130] EDPS Big Data (n 30) 13; See also European Commission, 'Communication 'Towards a Thriving Data-driven Economy', COM(2014) 442 final, point 4.2.3 < https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52014DC0442 > accessed 23 September 2023; ICO Big Data (n 7) 83-84.
[131] European Commission, 'Proposal for a Regulation of the European Parliament and of the Council on a Framework for Financial Data Access and Amending Regulations (EU) No 1093/2010, (EU) No 1094/2010, (EU) No 1095/2010 and (EU) 2022/2554' COM (2023)260 final ('FIDA').
[132] ibid, arts 2(2), 3(5).
[133] ibid, art 8, recital 22.
[134] As defined under FIDA, art 3(2).
[135] Ibid, art 8 (3) and(4).
[136] AIA, art 11, Annex IV.

while more information will become available to children as they grow[137]. This will resolve the transparency issues identified in Part III, providing that data controllers will try to find innovative ways to present the information to each type of identified data subject. As the roles of each party might be different depending on the stage of the data journey, swimlane diagrams[138] could also be used to clearly indicate in an engaging way[139] who is the data controller(s) per each data process and data journey stage.

What is more, this dashboard can also support the lawfulness principle, since it could be used to keep a record of the data subjects' consent and would allow him to easily withdraw at any time. The same applies for the right of objection, whenever the processing relies under the controller's legitimate interests. As far as storage limitation is concerned, FIDA's dashboard includes a warning of the possible consequences of permission withdrawal[140]. Similarly, the proposed data processing dashboard can warn in advance data subjects, when applicable, that if they exercise a right of erasure, this will be limited for unsupervised ML models. The warning can also further extend to provide a brief overview of the Data Protection Impact Assessment outcome, along with the foreseeable risks to the rights and safety of data subjects[141] each data processing activity entails, using a traffic light system to make it easier for data subjects to identify major privacy concerns. Lastly, the dashboard could also help data controllers to notify data subjects for any further processing activity they might decide on a later stage, and depending on the legal basis would either require them to consent to such use or will inform them about their right to object to such further processing, in line with the purpose limitation principle.

The suggested dashboard would thus provide a significant preventive measure that will mitigate the chances of ML models to violate the GDPR principles before any personal data processing activity takes place and will hopefully allow data subjects to take back control of their data. At the same time, data controllers would reduce the risk of GDPR incompliance -and of the envisaged fines- and will gradually gain data subject's trust. As this dashboard would in the future require significant time and effort by the data subjects to fully control all data processing requests, a privacy personal assistant that is already used in mobile apps could also be deployed, to reduce significantly the time requests by data subjects if the prediction rate is highly accurate.[142]. Specifically, the privacy personal assistant would not only identify and provide an overview of the basic privacy issues of each data processing request[143], but also via deploying ML it will be able to predict each data subject's privacy preferences and decide to requests, having the data subjects' tailored interests in mind.

VI. Conclusion

[137] For this, the collection of the child's birthdate would be necessary.

[138] Margaret Hagan, ' Exploding the Fine Print: Designing Visual, Interactive, Consumer-Centric Contracts and Disclosures' in Marcelo Corrales, Mark Fenwick and Helena Haapio (eds), *Legal Tech, Smart Contracts, and Blockchain* (Springer 2019) 93, 104.

[139] Stefania Passera, 'Flowcharts, Swimlanes, and Timelines: Alternatives to Prose in Communicating Legal–Bureaucratic Instructions to Civil Servants' (2018) J Bus Tech Commun 32(2), 229.

[140] FIDA, recital 22.

[141] AIA, Annex IV, s 3.

[142] As suggested by Pardis Emami-Naeini et al 'Privacy Expectations and Preferences in an IoT World' Proceedings of SOUPS 2017 Thirteenth Symposium on Usable Privacy and Security, 399,410; Also suggested by Jones and Edenberg (n 51) 367.

[143] See Pribot as an example that simplifies privacy notices Michihan Engineering, ' Simplifying Privacy Policies Using Artificial Intelligence' (YouTube, 10 April 2018) <https://www.youtube.com/watch?v=qe0oNDQGBs0 > accessed 30 September 2023.

AI is not just another privacy-intrusive technology. It can be used in a way that improves the quality of life overall and also further protect fundamental rights[144]. Although it poses many data protection challenges, humans are enabled in every AI development process, and it is high time we 'reacted' and impact the way AI will be process personal data. For this, we need to embrace humanity, wear our ''legal design lens' and take utmost advantage of the already enshrined data protection rights and obligations to ensure data protection will not become obsolete. As technology can be ''..applied in the service of interests that it concurrently threatens.''[145], a real-time data processing dashboard that would allow data subjects maintain control over who, why and what types of personal data each data controller can process, will hopefully be a promising start towards dynamic user empowerment that can cope with dynamic technologies like AI.

[144] Raposo, (n 91) 101.

[145] Lee A Bygrave, 'Hardwiring Privacy' in Roger Brownsword et al (eds), *The Oxford Handbook of Law, Regulation and Technology*, (OUP 2016) 504.