



HAL
open science

Another Path for AI Regulation: Worker Unions and Data Protection Rights

Thomas Le Bonniec

► **To cite this version:**

Thomas Le Bonniec. Another Path for AI Regulation: Worker Unions and Data Protection Rights. Italian Labour Law e-journal, 2024, 17 (2), pp.115-131. 10.6092/issn.1561-8048/20870. hal-04854809

HAL Id: hal-04854809

<https://hal.science/hal-04854809v1>

Submitted on 6 Jan 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Another Path for AI Regulation: Worker Unions and Data Protection Rights

Thomas Le Bonniec*

1. Introduction: how the EU regulation shapes personal data and AI markets. 2. The AI Regulation and the prevalence of Data Protection Regulation. 3. The Role of DPAs in Interpreting Data Protection Laws under the GDPR. 4. The shaping of markets by the Data Protection Agencies. 5. Workers as stakeholders in the interpretation of the GDPR. 6. Collective strategies for the reappropriation of the GDPR. 7. Concluding remarks.

Abstract

The Artificial Intelligence Regulation (EU Reg. 2024/1689) is widely regarded as the European Union's primary tool for regulating the market for AI systems. This paper, however, explores how the General Data Protection Regulation (EU Reg. 2016/679) takes precedence and remains a crucial legislation in determining how AI systems are produced and deployed in the EU. This research analyses how the GDPR's prevalence and application, however, is steered by interpretations put forward by different actors, with particular attention to Data Protection Authorities. In this context, I analyse how AI workers and (working) data subjects are essential stakeholders for the production of AI systems. I explore how their involvement in this regulatory standard could enhance the protection of their own interests and rights. More specifically, I focus on how consumer rights organizations and trade union cooperation could cooperate to challenge current predatory practices by the AI industry through the collective use of data protection rights.

Keywords: Artificial Intelligence; Data Protection; Digital Labour; Regulation; GDPR.

1. Introduction: how the EU regulation shapes personal data and AI markets.

Reinforced learning through human feedback (RLHF) means that humans are tasked with annotating, correcting and sometimes pretending to be “Artificial Intelligences”.¹ Examples

* Sociology PhD candidate, Numérique, Organisation et Société, Institut Interdisciplinaire de l'Innovation, Sciences Economiques et Sociales; Institute Polytechnique de Paris. This essay has been submitted to a double-blind peer review.

¹ Tubaro P., Casilli A.A., Coville M., *The Trainer, the Verifier, the Imitator: Three Ways in Which Human Platform Workers Support Artificial Intelligence*, in *Big Data & Society*, 7, 1, 2020.

of AI systems depending on RLHF include, non-exhaustively, content moderation, vocal assistants, computer vision for video-surveillance algorithms, self-check-out stores, robotic vacuum cleaners, Large Language Models and “autonomous cars”. Despite other methods for AI production existing, this paper will focus on *Machine Learning*² techniques that are data intensive and labour dependent.

The *digital labour* studies have developed a comprehensive picture of this type of work, which is characterised by its embeddedness in the platform economy,³ complex subcontracting global chains of value,⁴ and its contribution as an essential part of the production process of “Artificial Intelligence” systems, the *AI data pipeline*.⁵ For clarity purposes, the term *AI data work*⁶ will regroup these activities.

This paper argues that this production system and its regulation can be studied as a *Bourdieusian economic field*. This is what Pierre Bourdieu describes as a “*relatively autonomous subspace inside the global economical field*”.⁷ Defining this field would require a description of the agents structuring it and the power relations between them which determine their relative positions. I argue that the existence and composition of a market is ultimately dependent on the institutional forms that regulate it:⁸ in other words, a market only exists in the manner that economic, political and social actors (with enough influence) deem feasible. This is what I shall refer to as regulation: not the study and determination of the law, but a dynamic institutional setting. AI digital labour and platform labour are also part of this

²“For supervised machine learning, human engineers supply labeled training data to a computer. Two distinct types of algorithms then come into play: learners and classifiers. The learner is the algorithm that is trained on these labeled data examples; it then informs the classifier how best to analyze the relation between the new inputs and the desired target output (or prediction). It might be predicting whether a face is contained in an image or whether an email is spam. The more examples of correctly labeled data there are, the better the algorithm will be at producing accurate predictions. There are many kinds of machine learning models, including neural networks, logistic regression, and decision trees”. Crawford K., *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*, Yale University Press, New Haven, 2021, 99-100.

³ Tubaro P., *Disembedded or Deeply Embedded? A Multi-Level Network Analysis of Online Labour Platforms*, in *Sociology*, 55, 5, 2021, 927.

⁴ Schmidt F.A., *The Planetary Stacking Order of Multilayered Crowd-AI Systems*, in Graham M. Ferrari F. (eds.), *Digital Work in the Planetary Market*, The MIT Press, Cambridge, 2022.

⁵ Muldoon J., Callum C., Graham M., *A Typology of Artificial Intelligence Data Work*, in *Big Data & Society*, 11, 1, 2024. The AI data pipeline is the following series of tasks integrated in the AI production process: data collection, data curation, data annotation, model training to model evaluation and data verification.

⁶ *Ibidem*. “We define AI data work as the human labour required to support machine learning algorithms through the preparation and evaluation of datasets and model outputs that is often outsourced to low-paid and marginalised workers. Our definition does not include software developers and machine learning engineers; nor does it include content moderators for social media platforms whose work does not feed into AI production”.

⁷ *Cont*. “The analysis supposes first to characterize the existence of a field and its nature, by evidencing the role and relative weight of the actors and public institutions, even in the case of the most deregulated markets. Every economical field is politically constructed, and this at different levels, from the most global to the most local” (own translation). Lebaron F., in Sapiro G., *Dictionnaire international Bourdieu*, CNRS Editions, Paris, 2020.

⁸“The study of regulation of capitalism cannot be the search for abstract economic laws. It is the study of the transformation of social relations creating new shapes both economic and non-economic, shapes organised in structures and reproducing a determining structure, the mode of production” (own translation). (Aglietta M. quoted in *Glossaire*, in Boyer R., Saillard Y., *Théorie de la régulation, l'état des savoirs* (2e éd.), La Découverte, Paris, 2010.

scope. The participation or non-participation of workers to this setting is what this paper sets out to explore.

Inside the European Union, this economic field would be formally structured by the General Data Protection Regulation (GDPR), and the AI Act. The first, having come into effect in 2018 (EU Reg. 2016/679), is subject to divergent interpretations and its application is far from being a settled matter; the second is being progressively applied from August 2024 until 2026 (EU Reg. 2024/1689). Section 2 of this paper argues that data protection regulation remains prevalent for the AI market. These are not the only pieces of legislation that come into effect, but the interplay between data protection regulation and the AI market is a key element in the shaping of the AI market in the European Union. I set out to explore how AI data work is determined by this institutional setting, and how AI workers and data subjects can have a more active role in the application of this regulation.

Section 3 turns to the role and efficiency of the Data Protection Authorities (DPAs) in their diverging interpretations of the GDPR and how these are influenced by other agents. Section 4 elaborates on the central role of these DPAs in the shaping of this economic field through their interpretation and application of data protection regulation. I argue in Section 5 that, based on current examples, workers can and should participate in the DPA's interpretation and application of the GDPR. Section 6 explores the potential for cooperation between worker and consumer unions to challenge existing industry practices through collective action. Section 7 concludes by emphasising the role of worker participation in the evolving regulatory standard to steer it in a more favourable direction for them.

2. The AI Regulation and the prevalence of Data Protection Regulation.

Data annotation tasks, where humans train or correct AI systems on an example basis, are often based on information produced by users, which is regularly collected without their knowledge and consent. This has been evidenced with ChatGPT⁹ and vocal assistants,¹⁰ as well as devices depending on computer vision, where Personally Identifiable Information (PII) was collected. One such example is the Roomba robot vacuum which produced video recordings, then collected and later annotated by Venezuelan workers. A journalist from the MIT Technology Review who had access to some of these images describes: “the most intimate image we saw was the series of video stills featuring the young woman on the toilet, her face blocked in the lead image but unobscured in the grainy scroll of shots below”.¹¹

⁹ Perrigo B., *Exclusive: OpenAI used Kenyan workers on less than \$2 per hour to make ChatGPT less toxic*, in *TIME*, 18 January 2023, <https://time.com/6247678/openai-chatgpt-kenya-workers/> (last accessed 14 November 2024).

¹⁰ Tubaro P., Casilli A.A., nt. (1).

¹¹ Guo E., *A Roomba Recorded a Woman on the Toilet. How Did Screenshots End up on Facebook?*, in *MIT Technology Review*, 19 December 2022, available at <https://www.technologyreview.com/2022/12/19/1065306/roomba-robot-robot-vacuums-artificial-intelligence-training-data-privacy/> (last accessed 25 July 2024).

These data collection practices are commonplace across the AI industry, and affect hundreds of millions of individuals across the EU. Since these AI systems are heavily dependent on personal data, their producers become controllers or processors of personal data from a GDPR perspective,¹² and thus an important focus for Data Protection Agencies.

Before the draft on the AI Act was finalised, some DPAs had already taken a hands-on approach on the matter. On March 31st, 2023, the Italian DPA (GDPD) issued an “*immediate temporary limitation*” on OpenAI’s ChatGPT processing of personal data, arguing among other things that “*there appears to be no legal basis underpinning the massive collection and processing of personal data in order to ‘train’ the algorithms on which the platform relies*”¹³. This led to OpenAI’s decision to suspend its services in Italy until the GDPD considered its expectations had been met, a month later.

More recently, the European Data Protection Board (EDPB), which oversees all DPAs, published the findings of a European *taskforce* on ChatGPT’s compliance,¹⁴ stating:

“If ChatGPT is made available to the public, it should be assumed that individuals will sooner or later input personal data. If those inputs then become part of the data model and, for example, are shared with anyone asking a specific question, OpenAI remains responsible for complying with the GDPR and should not argue that the input of certain personal data was prohibited in first place.”

The sudden emergence of AI as generative tools became publicly available has also heavily influenced the DPA’s missions and agendas since 2022. Although some agencies had started engaging with it before 2022, there is a clear shift towards AI as one of the most important topics in data protection. This also comes with an effort to streamline their approach: the French DPA (CNIL) created a department dedicated to AI in early 2023.¹⁵

¹² EU Reg. 2016/679 (GDPR), Art. 4(7) and 4(8) GDPR: (7) “controller” means the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data; where the purposes and means of such processing are determined by Union or Member State law, the controller or the specific criteria for its nomination may be provided for by Union or Member State law; (8) “processor” means a natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller”.

¹³ Garante per la Protezione dei Dati Personali, *Artificial intelligence: stop to ChatGPT by the Italian S.A. Personal data is collected unlawfully, no age verification system is in place for children*, 31 March 2023: <https://www.garanteprivacy.it/home/docweb/-/docweb-display/docweb/9870847#english>. “*In its order, the Italian S.A highlights that no information is provided to users and data subjects whose data are collected by Open AI; more importantly*”.

¹⁴ European Data Protection Board (EDPB), *Report of the work undertaken by the ChatGPT Taskforce*, 23 May 2024, https://www.edpb.europa.eu/our-work-tools/our-documents/other/report-work-undertaken-chatgpt-taskforce_en (last accessed on 24 July 2024).

¹⁵ CNIL, *Création d’un service de l’intelligence artificielle à la CNIL et lancement des travaux sur les bases de données d’apprentissage*, 23 January 2023, <https://www.cnil.fr/fr/creation-dun-service-de-lintelligence-artificielle-la-cnil-et-lancement-des-travaux-sur-les-bases-de> (last accessed on 24 July 2024)

On April 4th 2024,¹⁶ the CNIL began publishing factsheets for AI companies, intending to break down how, in its view, data collection, processing, and all steps relative to AI systems should be carried in order to be GDPR compliant. Factsheet number 11 provides guidelines on how to make data annotation GDPR compliant, stating that personal annotation practices are considered akin to data processing.¹⁷

This document explicitly links personal data protection, AI production, and AI data work, by pointing out several steps that should be undertaken for private entities to remain compliant with European personal data protection law. Individuals whose information was collected (data subjects) may demand to know if their personal information has been subject to these practices on the grounds of the rights to access, rectification and suppression.

With the progressive implementation of the AI Regulation, EU Member States have until August 2025¹⁸ to designate a national “Market Surveillance Authority” or create one. Article 74(8) states that for “high-risk AI systems” in law enforcement, border control and justice administration, this role will be devoted to Data Protection Agencies¹⁹. The EDPB published a statement on July 16th, 2024, recommending that Member States expand this responsibility to other “high-risk AI systems”, in particular those impacting personal data.

Moreover, the prevalence of the GDPR is acknowledged several times in the preambles of the AI Regulation. Recital 10²⁰ reads as follows:

*“this Regulation does not seek to affect the application of existing Union law governing the processing of personal data, **including the tasks and powers of the independent supervisory authorities** competent to monitor compliance with those instruments”. Recital 69 adds: “The right to privacy and to protection of personal data must be guaranteed throughout the entire life cycle of the AI system. In this regard, the principles of data minimization and data protection by design and by default, as set out in Union data protection law, are applicable when personal data are processed.”*

Hence, in such instances, the GDPR takes precedence over the AI Act, which neither replaces nor undermines existing data protection requirements.

The AI data pipeline is thus identified by the DPAs – and the AI regulation itself – as a domain where data protection regulation must be seen as a fundamental pillar.

¹⁶ See CNIL, *Les fiches pratiques IA*, <https://www.cnil.fr/fr/les-fiches-pratiques-ia> (last accessed on 24 July 2024).

¹⁷ See CNIL, *IA: Annoter les données*, 14 June 2024, <https://www.cnil.fr/fr/ia-annoter-les-donnees> (last accessed on 24 July 2024).

¹⁸ Article 70, Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act).

¹⁹ Points 6,7,8 of Annex III, AI Regulation.

²⁰ Recital 10, AI Regulation.

3. The role of DPAs in interpreting Data Protection Laws under the GDPR.

As established in the previous section, the data pipeline is engaging in a great deal of data collection, which means the companies that produce these systems become accountable under data protection laws.

This is true in particular of companies producing LLMs,²¹ which scrape all available information on the internet to train them.²² Whether their data collection practices are compatible with data protection law is down to competing interpretations by agents of this field. Data Protection Authorities (DPAs), in particular are tasked with providing an interpretation that tries to reconcile economic policies and the fundamental rights to data protection and privacy. The tension between these two competing interests is embedded in the GDPR, and thus provides room for differing, even contradictory interpretations: this is demonstrated through key examples, such as the “consent popups” case.

The “consent popups” issue clearly illustrates the underlying tension in the GDPR. Online advertising companies use – among other means – cookies to track online activities and create commercial profiles based on personally identifiable data. To prove they were GDPR compliant, a technical solution, called the *Transparency and Control Framework* (TCF), was proposed by the association representing the interest of these advertisement companies, the *Internet Advertisers Board* (IAB).²³

A group of European NGOs, among which the Irish Council for Civil Liberties (ICCL),²⁴ filed a complaint in 2018 to push back against what they considered abusive collection of personal data and a serious security risk. In 2022, the European Data Protection Agencies issued a joint decision determining the TCF was unlawful for its lack of transparency, confidentiality, and because it didn’t properly request consent from users accessing websites behind a TCF pop-up.²⁵ This decision was challenged by the IAB, which

²¹ Solove D.J., Hartzog W., *The Great Scrape: The Clash Between Scraping and Privacy*, in *California Law Review*, 3 July 2024 (draft), (forthcoming) 2025.

²² Koebler J., *The Backlash Against AI Scraping Is Real and Measurable*, in *404 Media*, 23 July 2024, available at <https://www.404media.co/the-backlash-against-ai-scraping-is-real-and-measurable/> (last accessed on 1 August 2024); Longpre S., et al., *Consent in Crisis: The Rapid Decline of the AI Data Commons*, in *ArXiv*, *abs/2407.14933*, 2024, available at <https://www.dataprovenance.org/consent-in-crisis-paper>.

²³ An in-depth historical background of how this standard managed to prevail can be found in Arrah-Marie J., Rossi J., *La rivalité très politique des standards de recueil du consentement sur le Web*, in *Quaderni*, 112, 2, 2024, 22-23.

²⁴ Ryan J., *GDPR enforcers rule that IAB Europe’s consent popups are unlawful*, in *Irish Council for Civil Liberties*, 2 February 2022, available at <https://www.iccl.ie/news/gdpr-enforcer-rules-that-iab-europes-consent-popups-are-unlawful/> (last accessed on 29 July 2024).

²⁵ Led by the Belgian DPA, which explains: «The draft decision that the BE DPA had prepared was examined within the cooperation mechanism of the GDPR (the “one-stop-shop mechanism”). After serious scrutiny, and two objections that the BE DPA incorporated in a new draft, the present decision was approved by all concerned authorities representing most of the thirty countries in the European Economic Area». See Belgian Data Protection Authority, *Belgian DPA sends its draft decision in the IAB Europe case to European counterparts*, 25 November 2021, <https://www.dataprotectionauthority.be/belgian-dpa-sends-its-draft-decision-in-the-iab-europe-case-to-european-counterparts>; Belgian Data Protection Authority, *The BE DPA to restore order to the online advertising industry: IAB Europe held responsible for a mechanism that infringes the GDPR*, 2 February 2022, <https://www.dataprotectionauthority.be/iab-europe-held-responsible-for-a-mechanism-that-infringes-the-gdpr> (last accessed on 24 July 2024).

appealed to the Belgian Market Court. As of November 2024, the verdict is still pending, and the application of the DPAs' decision was suspended.²⁶

This case demonstrates how organisations whose activities revolve around production, collection and processing of personal data, can propose their own interpretation of the GDPR and see it become prevalent until it is challenged by the DPAs. This leeway is possible due to a contradiction inherent in the GDPR itself, which leaves enough room for very different, and often conflicting interpretations of the regulation.

The GDPR's Recital 2 reads as follows:

*“The principles of, and rules on the protection of natural persons with regard to the processing of their personal data should, whatever their nationality or residence, respect their fundamental rights and freedoms, in particular their right to the protection of personal data. This Regulation is intended to contribute to the [...] **strengthening and the convergence of the economies within the internal market, and to the well-being of natural persons.**”²⁷*

This sentence in particular highlights the two different and sometimes contradictory goals that European data protection regulation intends to achieve. On the one hand, data protection legislation is rooted on the EU Charter,²⁸ but on the other, it sets out to achieve EU “*internal market*” integration, for which a data processing standard is necessary. This is where the DPAs and “*Market Surveillance Authorities*” created by the AI Act are supposed to intervene: their role is to find a balance between market logics and individual fundamental rights.

In other words, these two pieces of legislation acknowledge that personal data and AI production cannot be left to *self-regulating* markets, but rather they require intervention from public independent agencies in charge of policing and finding a balance between a strict interpretation as economic laws (similar to anti-trust and competition laws) intended to shape a market and their role as fundamental rights guarantors.

²⁶ For more on this, see Belgian Data Protection Authority, *LAB EUROPE case: The CJEU answers the questions referred for a preliminary ruling*, 7 March 2024, <https://www.dataprotectionauthority.be/citizen/iab-europe-case-the-cjeu-answers-the-questions-referred-for-a-preliminary-ruling> (last accessed on 24 July 2024).

²⁷ 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) (Text with EEA relevance) 2016 (OJ L). Emphasis added.

²⁸ Articles 7 and 8 of the Charter of Fundamental Rights of the European Union, which guarantee that “*everyone has the right to respect for his or her private and family life, home and communications*” and “*the right to the protection of personal data concerning him or her*”.

4. The shaping of markets by the Data Protection Agencies.

This paper has argued so far that Data Protection Agencies have to be analysed in regards to their central position in the data processing and AI markets, but, as national agencies, they often publish diverse, and sometimes contradictory, positions.

There are 27 national Data Protection Agencies and 3 more belonging to the members of the European Economic Area: Norway, Iceland and Lichtenstein. On top of all of these are the European Data Protection Board, which coordinates them all, and the European Data Supervisor, which verifies EU institutions' compliance with the GDPR. DPAs have the same role all across the EU. As stated on article 51 of the GDPR, "*Each supervisory authority shall contribute to the consistent application of this Regulation throughout the Union. For that purpose, the supervisory authorities shall cooperate with each other and the Commission*".²⁹

However, these agencies have been found to be in disagreement over several cross-border cases, and some of them have been heavily criticized. The Irish Data Protection Commissioner (DPC) for instance has been accused of being too lenient in its enforcement of the GDPR,³⁰ and its decisions have been challenged by other DPAs in cross-border cases several times – such as Instagram's inappropriate collection of underage children's data.³¹ This is far from an exception; according to the Irish Council for Civil Liberties (ICCL), 67% of its cross-border decisions had been overruled as of 2023.³²

The former chair of the EDPB (2018-2023) and director of the Austrian DPA (DSB) Andrea Jelinek has contradicted this perspective in several public interventions. In 2022, she is quoted as saying: "*In the majority of cases, we have very good and calm cross-border cooperation. In a few cases, the dispute resolution mechanism is triggered*".³³ This affirmation highly contrasts with Ulrich Kelber, head of the German Federal DPA's, who accused his Irish counterpart,

²⁹ 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*).

³⁰ Dachwitz I., *Irland: Wenn Die Datenschutzbehörde Zur Facebook-Freundin Wird*, in *netzpolitik.org*, 6 December 2021, available at <https://netzpolitik.org/2021/irland-wenn-die-datenschutzbehoerde-zur-facebook-freundin-wird/> (last accessed on 30 July 2024).

³¹ See European Data Protection Board (EDPB), *Binding Decision 2/2022 on the Dispute Arisen on the Draft Decision of the Irish Supervisory Authority Regarding Meta Platforms Ireland Limited (Instagram) under Article 65(1)(a) GDPR*, 15 September 2022, https://www.edpb.europa.eu/our-work-tools/our-documents/binding-decision-board-art-65/binding-decision-22022-dispute-arisen_en (last accessed on 30 July 2024). See also Beesley A., *Irish watchdog opposed €1.2bn Meta fine, saying it would have no "meaningful dissuasive effect"*, in *The Irish Times*, 23 May 2023, <https://www.irishtimes.com/technology/2023/05/23/meta-fined-12bn-despite-irish-regulators-claim-no-financial-sanction-needed/> (last accessed on 30 July 2024).

³² As noted earlier, the ICCL also was one of the NGOs that put out the complaint in the "consent pop-ups" case. Ryan J., *5 Years: GDPR's crisis point. ICCL's 2023 report on EEA data protection authorities*, in *Irish Council for Civil Liberties*, 15 May 2023, <https://www.iccl.ie/digital-data/iccl-2023-gdpr-report/> (last accessed on 30 July 2024).

³³ Fanta A., *Europas Oberste Datenschützerin: „Ich Halte Überhaupt Nichts von Einer DSGVO-Reform“*, in *netzpolitik.org*, 29 August 2022, available at <https://netzpolitik.org/2022/europas-oberste-datenschuetzerin-ich-halte-ueberhaupt-nichts-von-einer-dsgvo-reform/> (last accessed on 30 July 2024) (own translation).

Helen Dixon, of making “*false statements*”³⁴ and refusing to collaborate, adding that his office had “*sent more than 50 complaints about WhatsApp to the Irish data protection supervisory authority (DPC), none of which has been closed to date*”.³⁵ Mr. Kelber goes on to add that the DPC is a major cause for obstruction for other EU regulators:

Ms Dixon [...] takes great care to be closely involved via the EDPB in any decisions taken by other EU supervisory authorities in these areas and thus to a certain extent prevents any initiatives by other supervisory authorities. Any lack of draft decisions is therefore primarily due to the DPC's extremely slow case handling, which falls significantly behind the case handling progress of most EU supervisors – especially German supervisors.

The Irish DPC is in fact a central player due to article 56 of the GDPR, which designates a “*lead supervision authority*” for cross-border cases. The lead is assigned to the DPA of the country where a company’s EU headquarters are found. Since most of the US tech giants have their EU operations in Ireland, the Irish DPC has become the primary regulatory hub for overseeing their compliance with EU data protection laws.

Beyond anecdotal evidence, the example of the DPC is telling of two structuring elements on AI and personal data regulation. The first one is that these regulators have extremely variable weight and importance. And the second is that these agencies have been following different strategies on personal data regulation on both compliance and enforcement.³⁶ These shifts also follow historical patterns, where a harmonization of different national institutions throughout the European Union is conducted for them to become EU market regulatory agencies.³⁷

³⁴ Fanta A., *Vorwurf von Ulrich Kelber: Irische Datenschutzbehörde macht „falsche Aussagen“*, in *netzpolitik.org*, 18 March 2021, <https://netzpolitik.org/2021/vorwurf-von-ulrich-kelber-irische-datenschutzbehoerde-macht-falsche-aussagen/> (last accessed on 30 July 2024).

³⁵ Official letter sent to the LIBE Commission of the European Parliament on March 17th, 2021. See Noyb, *DPC cancels Parliamentary Hearing on EU-US transfers*, in Noyb, 18 March 2021, <https://noyb.eu/en/dpc-cancels-parliamentary-hearing-eu-us-transfers> (last accessed on 2 August 2024).

³⁶ Sivan-Sevilla I., *Varieties of enforcement strategies post-GDPR: a fuzzy-set qualitative comparative analysis (fsQCA) across data protection authorities*, in *Journal of European Public Policy*, 31, 2, 2024, 552. Ido Sivan-Sevilla’s comparative study of when and how the DPAs decide to investigate and impose sanctions demonstrates this point. He concludes: The second contribution of this research is in providing a starting point for understanding the impacts of Europeanization on national enforcement strategies in the data protection case post-GDPR. Almost half of the studied agencies adopt a high-deterrence strategy of wide monitoring & supervision and high levels of coercion in their sanctioning approach. With a few outliers and two other clusters of less deterrent strategies, national divergence clearly arises, and it remains to be seen whether such divergence will erode over time.

³⁷ Bellon A. France P., *Chapitr 2. La CNIL: de l'institution militante à l'agence de régulation numérique*, in Vauchez A., *Le moment régulateur*, Presses de Sciences Po, Paris, 2024. Anne Bellon and Pierre France determine that the history of the French DPA (CNIL) can be presented as a mutation from a “militant” national institution preoccupied with limiting administrative and State overreach, into an “expert agency” From 2010, the “transformative effect of the field of regulation inside which the CNIL is progressively ‘caught’ also changes its focus and raison d’être”. Bellon adds: “While central administrations were the main personal data producers until the 2000s, now it is the private actors of the web that stock and accumulate information and traces of individual online activities, to the point that it radically transforms the CNIL’s audience whose deliberations concerned at 90% the public sector in 2003, and relate at 90% to the private sector in 2008”. Bellon A., *Protéger les données, entre mission politique et pratiques administratives, enquête sur le travail de la CNIL*, in Theviot A., *Gouverner par les données ? Pour une sociologie politique du numérique*, ENS Éditions,

From this evidence, I conclude that the European effort to unify national markets into one, by unifying national legislations, is not finalised. Although integration efforts and regulators affirm progress is continually made, this process will continue to take years, especially, as Jelinek puts it, because:

“Of course, large companies do not accept penalties, but fight every decision. Every one. Every one of these big decisions: The decision of the Irish authority, the decision of the European Data Protection Board. Everything is being fought. And as soon as the European Court of Justice has finally ruled - and that will of course take some time - only then will the matter be finally clear.”³⁸

The Data Protection Authorities have to manage contradicting forces between a *substantive*³⁹ “data economy” mindful of fundamental rights and the push for a *market economy*⁴⁰, focused primarily on the valuation of personal data and AI systems. In the first case, a DPA would be considered a watchdog guaranteeing oversight over predatory practices of personal data collection and use, with a focus on AI systems. In the second case, their role would merely be to provide the technical and legal guidance for any economic model based on personal data collection and usage. The practical approach of these regulatory agencies falls somewhere in between those poles, as they must both limit and facilitate data flows for economical purposes. As they are national independent organisations, these interpretations often vary according to their own priorities. The variations between regulations agencies must be explained accordingly: the proposal outlined in this article puts forward an interpretation in terms of institutional forms and contexts.

Lyon, 2023. Similar shifts are observed in other EU DPAs: a more complete panorama is presented by Gloria González-Fuster’s account of the merging of the national privacy laws into one EU data protection law from the 1970s to the GDPR creation. González Fuster G., *The Emergence of Personal Data Protection as a Fundamental Right of the EU*, Springer International Publishing, Switzerland, 2014.

³⁸ Brandtner B., *Andrea Jelinek: “Die DSGVO war und ist ein Erfolgsmodell”*, in *DER STANDARD*, 22 May 2023:

<https://www.derstandard.de/story/3000000021161/andrea-jelinek-die-dsgvo-war-und-ist-ein-erfolgsmodell> (last accessed on 31 July 2024).

³⁹ Cangiani M., *Karl Polanyi’s Institutional Theory: Market Society and Its “Disembedded” Economy*, in *Journal of Economic Issues*, 45, 1, 2011, 177. “The substantive meaning of economic” – Polanyi writes (1957, 243) – “derives from man’s dependence for his living upon nature and his fellows. It refers to the interchange with his natural and social environment, in so far as this results in supplying him with the means of material want satisfaction”.

⁴⁰ Polanyi K., *The Great Transformation: The Political and Economic Origins of Our Time*, Beacon Press, Boston, 2001. “A market economy is an economic system controlled, regulated, and directed by market prices; order in the production and distribution of goods is entrusted to this self-regulating mechanism. An economy of this kind derives from the expectation that human beings behave in such a way as to achieve maximum money gains. It assumes markets in which the supply of goods (including services) available at a definite price will equal the demand at that price. It assumes the presence of money, which functions as purchasing power in the hands of its owners”.

5. Workers as stakeholders in the interpretation of the GDPR.

The previous sections presented a conceptual framework in which interpretation and enforcement of the GDPR are influenced by various actors: private companies and NGOs in particular try to influence DPAs' decision-making. Moreover, DPAs across the EU follow different interpretations which are tied to their national contexts. This section's purpose is to explore how workers, and in particular AI data workers, can become active participants of this regulatory process, and influence it in their favour.⁴¹

Christina Colclough⁴² argues that with the *datafication* of work, all processes must involve workers – from data collection and analysis to storage and off-boarding – to let them negotiate and protect their rights. From knowing what data is collected at work and for what ends, to having a say in who can access it or buy it, all these are acknowledged as individual rights in the GDPR. This process of datafication of work is not necessarily new, but it is now heavily associated with AI systems as a means to create value from it.

One of the demonstrations brought about by digital labour researchers is how platform users not only produce personal data that is later used to train AI systems, but *also* contribute to its annotation, training and correction process.

This trend is encompassed under the concept of platformization, defined “*not as the disappearance of work, but as a modification of the ratio between implicitly productive work/ formally recognizable job*”.⁴³ Platform users perform activities that are paramount to “*hidden click work*”,⁴⁴ participating in the same production as AI data workers in diverse employment settings. Meta uses the photos and captions of Instagram and Facebook users to train its AI image generation models;⁴⁵ Reddit strikes a deal with Alphabet to use the content on its platform;⁴⁶ research articles published in reviews owned by Taylor & Francis are now collected by Microsoft for AI purposes.⁴⁷ Personal data protection is thus becoming one of

⁴¹ See for example Le Ludec C., Cornet M. and Casilli A.A., *The Problem with Annotation. Human Labour and Outsourcing between France and Madagascar*, in *Big Data & Society*, 10, 2, 2023; Casilli A.A. et al., *Le Micro-Travail en France. Derrière l'automatisation, de nouvelles précarités au travail ? (Report, Projet de recherche DiPLab 2019)*, HAL Open Science, 2019; Gray M.L., Suri S., *Ghost Work: How to Stop Silicon Valley from Building a New Global Underclass*, Harper Business, New York, 2019.

⁴² Colclough C.J., *Righting the wrong: putting workers' data rights firmly on the table*, in Graham M., Ferrari F. (eds.), *Digital Work in the Planetary Market*, The MIT Press, Cambridge, 2022.

⁴³ Casilli A.A., nt. (41).

⁴⁴ Casilli A.A., *La plateformisation comme mise au travail des usagers. Digital labor et nouvelles inégalités planétaires*, in Coriat B., Alix N., Bancel J.-L., Sultan F. (eds.), *Vers une République des Biens Communs?*, Les Liens qui Libèrent, Paris, 2018.

⁴⁵ Edwards B., *Meta's new AI image generator was trained on 1.1 billion Instagram and Facebook photos*, in *Ars Technica*, 6 December 2023, available at <https://arstechnica.com/information-technology/2023/12/metas-new-ai-image-generator-was-trained-on-1-1-billion-instagram-and-facebook-photos/> (last accessed on 12 January 2024).

⁴⁶ Roth E., *Google cut a deal with Reddit for AI training data*, in *The Verge*, 22 February 2024: <https://www.theverge.com/2024/2/22/24080165/google-reddit-ai-training-data> (last accessed on 31 July 2024).

⁴⁷ Battersby M., *Academic authors “shocked” after Taylor & Francis sells access to their research to Microsoft AI*, in *The Bookseller*, 19 July 2024, <https://www.thebookseller.com/news/academic-authors-shocked-after-taylor--francis-sells-access-to-their-research-to-microsoft-ai> (last accessed on 31 July 2024). See also Informa, *Informa PLC: Press Release*, 8 May 2024, available at <https://www.informa.com/globalassets/documents/investor>

the cornerstones of a number of social and economic fields, not least in labour-related situations.

I have mentioned several examples of how the regulatory process can be influenced by different agents; this role should also be played by workers and unions. For example, the complaints put forward by the NGO None of your business (Noyb)⁴⁸ with several European DPAs challenge OpenAI's compliance with data protection law. By doing so, these campaigning efforts have a practical effect on policy, either by strengthening a DPA's position, or by pushing for the opening of an investigation.

Another significant tactic is adopted by Trade Unions, which can use GDPR provisions to fight back against invasive data collection practices. In January 2024, the Dutch DPA (Autoriteit Persoonsgegevens) issued a 10 million euro fine against Uber, resulting from an agglomerated 170 complaints by Uber drivers, who used their right to access the data that was collected through the app, which they claimed was not respected.⁴⁹ These tactics could be generalised, inciting Trade Unions to fight back against invasive data collection practices. These could range from a right to veto intrusive software in the workplace to a default protection from data scraping in work situations. If a company was unable to provide proof of GDPR compliance, unions could consider turning to other types of industrial action beyond filing complaints in order to force companies that engage with personal data collection – in particular when integrating AI systems – to integrate other sociopolitical risks, such as reputational damage, in their calculations.

There is room for them: the GDPR is increasingly claimed by Courts, DPAs and trade unions⁵⁰ as a legal tool to push back against invasive tools and data collection practices. The example of Uber drivers pooling their complaints provides a blueprint for the exercise of data protection rights in a labour relationship. And those rights, under Article 80 of the GDPR, can be exerted collectively. In particular, its French transposition specifically introduces the notion that unions can exercise this power in a collective manner.⁵¹

As trade unions put forward a divergent interpretation of what data protection law says, they may also reshape GDPR interpretation and application, and with it the institutional forms that determine the production process of AI. Challenging the interpretation of personal data as a commodity is potentially a major issue for companies engaging in data collection and/or AI production.

[relations/2024/informa-plc---market-update.pdf](#) (last accessed on 31 July 2024). Parent company Informa states that the partnership with Microsoft will “Provide non-exclusive access to Advanced Learning content and data to help improve relevance and performance of AI systems”.

⁴⁸ Noyb, *ChatGPT provides false information about people, and OpenAI can't correct it*, 29 April 2024, available at <https://noyb.eu/en/chatgpt-provides-false-information-about-people-and-openai-cant-correct-it> (accessed on 1 August 2024).

⁴⁹ Autoriteit Persoonsgegevens, *Uber fined €10 million for infringement of privacy regulations*, 31 January 2024, <https://www.autoriteitpersoonsgegevens.nl/en/current/uber-fined-eu10-million-for-infringement-of-privacy-regulations> (last accessed on 31 July 2024).

⁵⁰ IndustriAll, *IndustriAll Europe's new “GDPR Toolbox for Trade Unionists”*, in *IndustriAll Europe News*, 5 July 2023, <https://news.industriall-europe.eu/Article/926> (last accessed on 31 July 2024).

⁵¹ Art. 37, Loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés, version en vigueur depuis le 01 juin 2019.

6. Collective strategies for the reappropriation of the GDPR.

Personal data regulation and its application offer some opportunities for both workers and consumers, organized in unions or collectives. As noted before, Uber drivers pooling their complaints to exercise access to their personal data is one such example. Unionized effort across Europe culminated in the Platform Work Directive (PWD)⁵² which sets limits to automated decision-making and algorithmic management by work platforms, with some of the supervision power allocated to European DPAs (Article 24 PWD).

In the case of consumers, the “*Bureau Européen des Unions de Consommateurs*” (BEUC), which regroups 44 consumer organizations in the EU has also put forward several collective complaints. Examples include one against Google in 2018, for a “breach of GDPR” for undue geolocation tracking,⁵³ and another against Meta’s “pay-or-consent” model in 2023. The latter required users to agree to tracking cookies or pay a subscription to access Facebook and Instagram. The complaints were lodged with Consumer Practices Authorities on the grounds of “*unfair commercial practices*”, while still “*assessing whether Meta is infringing data protection law (the GDPR)*”.⁵⁴

As mass producers and annotators of data, workers and consumers provide substantial contributions to AI production, but rarely, if ever, have a say in the data collection practices they are subject to. The potential is there for worker unions to expand their collective bargaining efforts to integrate AI training as part of the scope of personal data regulation. Moreover, if companies fail to demonstrate GDPR compliance, unions could consider more assertive forms of industrial action, pressuring employers to account for the broader sociopolitical risks, associated with data collection and AI integration. By following the example set by consumer organisations, workers and their unions could strengthen their collective power and reshape the data economy.

7. Concluding remarks.

AI data work is a planetary market⁵⁵ where workers around the globe are forced to compete and accept whatever income and work conditions platforms or Business Process Outsourcing companies impose on them.⁵⁶

⁵² Directive (EU) 2024/2831 of the European Parliament and of the Council of 23 October 2024 on improving working conditions in platform work. Article 24(1): “The supervisory authority or authorities responsible for monitoring the application of Regulation (EU) 2016/679 shall also be responsible for monitoring and enforcing the application of Articles 7 to 11 of this Directive as far as data-protection matters are concerned, in accordance with the relevant provisions in Chapters VI, VII and VIII of Regulation (EU) 2016/679”.

⁵³ BEUC – The European Consumer Organisation, *Consumer groups across Europe file complaints against Google for breach of GDPR*, 27 November 2018, available at <https://www.beuc.eu/press-releases/consumer-groups-across-europe-file-complaints-against-google-breach-gdpr> (last accessed on 1 August 2024).

⁵⁴ BEUC – The European Consumer Organisation, *Consumer groups file complaint against Meta’s unfair pay-or-consent model*, 30 November 2023, available at <https://www.beuc.eu/press-releases/consumer-groups-file-complaint-against-metas-unfair-pay-or-consent-model> (last accessed on 1 August 2024).

⁵⁵ Graham M., Ferrari F. (eds.), *Digital Work in the Planetary Market*, The MIT Press, Cambridge, 2022.

This paper aimed to show a theoretical framework to study data protection regulation as it relates to “Artificial Intelligence” production from a socioeconomic perspective, where data producers are acknowledged as stakeholders, in order to explore their current position, and their potential role in shaping a different regulation path.

In sections 1 and 2, I insisted on the prevalence of the General Data Protection Regulation despite the upcoming AI law in the European Union. Section 3 provides an outline of the importance of DPAs in the interpretation and application of the GDPR. Section 4 underlines the importance of these agencies in the shaping of the EU market, and how they are influenced by outer actors. The GDPR offers room for competing interpretations: on sections 5 and 6, I show this is an opportunity for worker and consumer unions to steer regulation in a different direction.

In practical terms, data protection law could be leveraged to do more than grant minimal protections for workers in the EU and for EU data subjects. GDPR requirements in terms of data processing open up several opportunities to negotiate a better working situation or to severely limit ongoing abusive situations. Data minimization, data access, limitation of processing and data portability (Articles 5, 15, 18, 20 GDPR), for example, offer a series of initial rights that can be used collectively, possibly through an organisation, in order to change the power dynamic a bargaining situation.

For companies collecting and using data, regulation is now an extremely important issue: lobbying,⁵⁷ engaging with regulators and assuring the general public of its good practices⁵⁸ are all indicators of how crucial it has become for the AI industry to make its ongoing data collection practices acceptable.⁵⁹

AI companies are also heavily dependent on personal data they have collected, and on unrecognized AI data labour: if data subjects start agglomerating instead of being treated as atomized individuals, they improve their bargaining position in order to challenge the current regulatory standard. The existing European data protection, used this way, could emerge as a further challenge to *surveillance capitalism*,⁶⁰ by relying at least partially in the ability of NGOs, workers and consumers, to pool resources and effectively make use of the institutional context in which they are embedded. Proposals for a collective data protection right already exist in order to push regulation towards a *substantive data economy*, one where

⁵⁶ Miceli M., Posada J., *The Data-Production Dispositif*, in *Proceedings of the ACM on Human-Computer Interaction*, 6, 1, 2022, 460. “Many machine learning (ML) models are built from training data previously collected, cleaned, and annotated by human workers. Companies and research institutions outsource several of these tasks through online labour platforms and business process outsourcing (BPO) companies. In these instances, outsourcing organization and their clients regard workers as independent contractors, considering them factors of production, and their labour a commodity or a product subject to market regulations”.

⁵⁷ Perrigo B., *Exclusive: OpenAI Lobbied E.U. to Water Down AI Regulation*, in *TIME*, 20 June 2023, available at <https://time.com/6288245/openai-eu-lobbying-ai-act/> (last accessed on 29 July 2024).

⁵⁸ OpenAI, *OpenAI and Journalism*, 8 January 2024, available at <https://openai.com/index/openai-and-journalism/> (last accessed on 29 July 2024).

⁵⁹ Longpre S., et al., *Data Authenticity, Consent, and Provenance for AI Are All Broken: What Will It Take to Fix Them?*, in *An MIT Exploration of Generative AI – MitGenial*, 27 March 2024, available at <https://mit-genai.pubpub.org/pub/uk7op8zs/release/2> (last accessed on 22 July 2024).

⁶⁰ Zuboff S., *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*, PublicAffairs, New York, 2019.

personal data collection and AI production are integrated (one could say embedded) in other social functions.⁶¹

Bibliography

- Arrah-Marie J., Rossi J., *La rivalité très politique des standards de recueil du consentement sur le Web*, in *Quaderni*, 112, 2, 2024;
- Aufrère L., Maurel L., *Pour Une Protection Sociale Des Données Personnelles*, HAL Open Science, 2018;
- Autoriteit Persoonsgegevens, *Uber fined €10 million for infringement of privacy regulations*, 31 January 2024;
- Battersby M., *Academic authors “shocked” after Taylor & Francis sells access to their research to Microsoft AI*, in *The Bookseller*, 19 July 2024;
- Beesley A., *Irish watchdog opposed €1.2bn Meta fine, saying it would have no “meaningful dissuasive effect”*, in *The Irish Times*, 23 May 2023;
- Belgian Data Protection Authority, *Belgian DPA sends its draft decision in the LAB Europe case to European counterparts*, 25 November 2021;
- Belgian Data Protection Authority, *LAB EUROPE case: The CJEU answers the questions referred for a preliminary ruling*, 7 March 2024;
- Belgian Data Protection Authority, *The BE DPA to restore order to the online advertising industry: LAB Europe held responsible for a mechanism that infringes the GDPR*, 2 February 2022;
- Bellon A., France P., *Chapitr 2. La CNIL: de l’institution militante à l’agence de régulation numérique*, in Vauchez A., *Le moment régulateur*, Presses de Sciences Po, Paris, 2024;
- Bellon A., *Protéger les données, entre mission politique et pratiques administratives, enquête sur le travail de la CNIL*, in Theviot A., *Gouverner par les données ? Pour une sociologie politique du numérique*, ENS Éditions, Lyon, 2023;
- BEUC – The European Consumer Organisation, *Consumer groups across Europe file complaints against Google for breach of GDPR*, 27 November 2018;
- BEUC – The European Consumer Organisation, *Consumer groups file complaint against Meta’s unfair pay-or-consent model*, 30 November 2023;
- Boyer R., Saillard Y., *Théorie de la régulation, l’état des savoirs* (2e éd.), La Découverte, Paris, 2010;
- Brandtner B., *Andrea Jelinek: “Die DSGVO war und ist ein Erfolgsmodell”*, in *DER STANDARD*, 22 May 2023;
- Cangiani M., *Karl Polanyi’s Institutional Theory: Market Society and Its “Disembedded” Economy*, in *Journal of Economic Issues*, 45, 1, 2011;
- Casilli A.A. et al., *Le Micro-Travail en France. Derrière l’automatisation, de nouvelles précarités au travail? (Report, Projet de recherche DiPLab 2019)*, HAL Open Science, 2019;
- Casilli A.A., *La plateforme comme mise au travail des usagers. Digital labor et nouvelles inégalités planétaires*, in Coriat B., Alix N., Bancel J.-L., Sultan F. (eds), *Vers une République des Biens Communs?*, Les Liens qui Libèrent, Paris, 2018;
- CNIL, *Création d’un service de l’intelligence artificielle à la CNIL et lancement des travaux sur les bases de données d’apprentissage*, 23 January 2023;
- Colclough C.J., *Righting the wrong: putting workers’ data rights firmly on the table*, in Graham M., Ferrari F. (eds), *Digital Work in the Planetary Market*, The MIT Press, Cambridge, 2022;

⁶¹ Aufrère L., Maurel L., *Pour Une Protection Sociale Des Données Personnelles*, HAL Open Science, 2018, available at <https://hal.science/hal-01903526> (last accessed on 4 April 2024).

- Crawford K., *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*, Yale University Press, New Haven, 2021;
- Dachwitz I., *Irland: Wenn Die Datenschutzbehörde Zur Facebook-Freundin Wird*, in *netzpolitik.org*, 6 December 2021;
- Edwards B., *Meta's new AI image generator was trained on 1.1 billion Instagram and Facebook photos*, in *Ars Technica*, 6 December 2023;
- European Data Protection Board (EDPB), *Binding Decision 2/2022 on the Dispute Arisen on the Draft Decision of the Irish Supervisory Authority Regarding Meta Platforms Ireland Limited (Instagram) under Article 65(1)(a) GDPR*, 15 September 2022;
- European Data Protection Board (EDPB), *Report of the work undertaken by the ChatGPT Taskforce*, 23 May 2024;
- Fanta A., *Europas Oberste Datenschützerin: „Ich Halte Überhaupt Nichts von Einer DSGVO-Reform“*, in *netzpolitik.org*, 29 August 2022;
- Fanta A., *Vorwurf von Ulrich Kelber: Irische Datenschutzbehörde macht „falsche Aussagen“*, in *netzpolitik.org*, 18 March 2021;
- Garante per la Protezione dei Dati Personali *Artificial intelligence: stop to ChatGPT by the Italian S.A. Personal data is collected unlawfully, no age verification system is in place for children*, 31 March 2023;
- González Fuster G., *The Emergence of Personal Data Protection as a Fundamental Right of the EU*, Springer International Publishing, Switzerland, 2014;
- Graham M., Ferrari F. (eds), *Digital Work in the Planetary Market*, The MIT Press, Cambridge, 2022;
- Gray M.L., Suri S., *Ghost Work: How to Stop Silicon Valley from Building a New Global Underclass*, Harper Business, New York, 2019;
- Guo E., *A Roomba Recorded a Woman on the Toilet. How Did Screenshots End up on Facebook?*, in *MIT Technology Review*, 19 December 2022;
- IndustriAll, *IndustriAll Europe's new "GDPR Toolbox for Trade Unionists"*, in *IndustriAll Europe News*, 5 July 2023;
- Informa, *Informa PLC: Press Release*, 8 May 2024;
- Koebler J., *The Backlash Against AI Scraping Is Real and Measurable*, in *404 Media*, 23 July 2024;
- Le Ludec C., Cornet M., Casilli A.A., *The Problem with Annotation. Human Labour and Outsourcing between France and Madagascar*, in *Big Data & Society*, 10, 2, 2023;
- Longpre S., et al., *Consent in Crisis: The Rapid Decline of the AI Data Commons*, in *ArXiv*, *abs/2407.14933*, 2024;
- Longpre S., et al., *Data Authenticity, Consent, and Provenance for AI Are All Broken: What Will It Take to Fix Them?*, in *An MIT Exploration of Generative AI – MitGenial*, 27 March 2024;
- Miceli M., Posada J., *The Data-Production Dispositif*, in *Proceedings of the ACM on Human-Computer Interaction*, 6, 1, 2022;
- Muldoon J., Callum C., Graham M., *A Typology of Artificial Intelligence Data Work*, in *Big Data & Society*, 11, 1, 2024;
- Noyb, *ChatGPT provides false information about people, and OpenAI can't correct it*, 29 April 2024;
- Noyb, *DPC cancels Parliamentary Hearing on EU-US transfers*, in *Noyb*, 18 March 2021;
- OpenAI, *OpenAI and Journalism*, 8 January 2024;
- Perrigo B., *Exclusive: OpenAI Lobbied E.U. to Water Down AI Regulation*, in *TIME*, 20 June 2023;
- Perrigo B., *Exclusive: OpenAI used Kenyan workers on less than \$2 per hour to make ChatGPT less toxic*, in *TIME*, 18 January 2023;

-
- Polanyi K., *The Great Transformation: The Political and Economic Origins of Our Time*, Beacon Press, Boston, 2001;
- Roth E., *Google cut a deal with Reddit for AI training data*, in *The Verge*, 22 February 2024;
- Ryan J., *5 Years: GDPR's crisis point. ICCL's 2023 report on EEA data protection authorities*, in *Irish Council for Civil Liberties*, 15 May 2023;
- Ryan J., *GDPR enforcers rule that LAB Europe's consent popups are unlawful*, in *Irish Council for Civil Liberties*, 2 February 2022;
- Schmidt F.A., *The Planetary Stacking Order of Multilayered Crowd-AI Systems*, in Graham M., Ferrari F. (eds), *Digital Work in the Planetary Market*, The MIT Press, Cambridge, 2022;
- Sivan-Sevilla I., *Varieties of enforcement strategies post-GDPR: a fuzzy-set qualitative comparative analysis (fsQCA) across data protection authorities*, in *Journal of European Public Policy*, 31, 2, 2024;
- Solove D.J., Hartzog W., *The Great Scrape: The Clash Between Scraping and Privacy*, in *California Law Review*, 3 July 2024 (draft), (forthcoming) 2025;
- Tubaro P., Casilli A.A., *Human Listeners and Virtual Assistants: Privacy and labour Arbitrage in the Production of Smart Technologies*, in Graham M., Ferrari F. (eds), *Digital Work in the Planetary Market*, The MIT Press, Cambridge, 2022;
- Tubaro P., Casilli A.A., Coville M., *The Trainer, the Verifier, the Imitator: Three Ways in Which Human Platform Workers Support Artificial Intelligence*, in *Big Data & Society*, 7, 1, 2020;
- Tubaro P., *Disembedded or Deeply Embedded? A Multi-Level Network Analysis of Online Labour Platforms*, in *Sociology*, 55, 5, 2021;
- Zuboff S., *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*, PublicAffairs, New York, 2019.

Copyright © 2024 Thomas Le Bonniec. This article is released under a Creative Commons Attribution 4.0 International License