EJLT European Journal of Law and Technology

The Right to Rectification and Inferred Personal Data

Andreas Häuselmann and Bart Custers*

Abstract:

Inferred data, such as predictions and emotional states, may have limited accuracy or may even be incorrect from the perspective of data subjects. According to Article 16 of the General Data Protection Regulation (GDPR), data subjects can then invoke their right to rectification. To rectify personal data, data subjects must provide objectively verifiable evidence that the personal data envisaged to replace the data currently processed by the data controller is accurate ('the standard of objective verifiability'). This causes three problems. First, whereas the standard of objective verifiability is easily met for factual data, predictions are not objectively verifiable, mainly because they relate to future conduct or events that have yet to happen (the verifiability problem). Second, the accuracy of subjective personal data, such as emotion data, cannot be proven objectively (the objectivity problem). Third, to effectively rectify inferred personal data, data subjects must be aware that data is inaccurate (the awareness problem). This is often not the case because inferred data are treated as trade secrets and are not shared with data subjects − even when they invoke their right of access.

Keywords: accuracy, AI, inferred personal data, rectification, machine learning, affective computing.

1. Introduction

The right to rectification provided by Article 16 GDPR is the first-choice remedy for inaccurate or incomplete data. This right enables the data subject to obtain 'the

* Andreas Häuselmann is an Assistant Professor of Privacy and Data Protection Law at Open Universiteit. Bart Custers is a full Professor of Law and Data Science at eLaw, the Center for Law and Digital Technologies at Universiteit Leiden.

rectification of inaccurate personal data' and 'to have incomplete personal data completed, including by means of providing a supplementary statement'.[1] As the name of the right indicates, rectification implicitly relies on the notion of verification in the sense that something may demonstrably be shown to be inaccurate or incomplete.[2] When invoking the right to rectification, a data subject has to provide accurate data that will replace the (presumed) inaccurate data. Here it should also be noted that the data controller has a responsibility in ensuring the accuracy of personal data. According to the accuracy principle in Article 5(1)(d) GDPR, the data controller must ensure the data are accurate and, where necessary, kept up to date. Furthermore, the principle of accountability requires data controllers to demonstrate compliance with data protection principles.[3] However, this does not change the fact that a data subject bears the burden of proof when enforcing the right to rectification.[4] As we demonstrate in this article, providing sufficient proof might be impossible when data subjects intend to rectify inferred personal data generated with artificial intelligence (AI).

There has been no substantial academic debate between computer scientists and legal scholars on information quality and accuracy.[5] Corresponding interdisciplinary research is a relatively recent development.[6] As we will show in this article, the dearth of interdisciplinary research on information quality and accuracy is a concern in the GDPR but will only intensify with the AI Act. The latter mentions accuracy several times, mainly in the context of high-risk systems.[7] As such, the AI Act, similar to the GDPR, requires an appropriate level of accuracy, which needs to be assessed in light of the AI system's intended purpose.[8]

This article provides an overview of the practical and conceptual challenges of the right to rectification in the context of inferred personal data. We focus on three significant issues limiting the effectiveness of the right to rectification. The first two are conceptual, and the third is practical. First, inferred data may not be verifiable (the verifiability problem). For instance, the correctness of many predictions can only be verified after they do or do not materialise. Although a prediction may be

[1] Article 16(1) General Data Protection Regulation (GDPR).
[2] Sandra Wachter and Brent Mittelstadt, 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' (2019) No. 2 *Columbia Business Law Review* 494, 548.
[3] Article 5(2) GDPR.
[4] Case C-247/23, *Deldits* [2024] ECR I-747, Opinion AG Collins para 47, see by analogy concerning the right to erasure Case C-460/20, *TU* [2022] ECR I-962 para 68.
[5] Dara Hallinan and Frederik Zuiderveen Borgesius, 'Opinions can be incorrect (in our opinion)! On data protection law's accuracy principle' (2020) Vol. 10 No. 1 *International Data Privacy Law* 1, 4.
[6] Burkhard Schäfer, 'Information Quality and Evidence Law: A New Role for Social Media, Digital Publishing and Copyright Law?' in Luciano Floridi and Phyllis Illari (eds), *The Philosophy of Information Quality* (Springer Nature 2014) 217.
[7] Articles 13(3), 15 and 58, and Annex IV paras 2–3 AI Act, see also Recitals 59, 60, 66, 74 and 122.
[8] Annex IV para 3, see also Recital 74 AI Act.

established with high accuracy and is valid from a statistical perspective, its substantive correctness can only be established in the future (if at all). Second, the accuracy of some personal data aiming to replace inaccurate personal data processed by the controller cannot be proven objectively (the objectivity problem). Emotion data are a typical example because emotions are highly subjective. Inferring a person's emotions or emotional state (e.g., from facial expression or speech) may be done objectively from a technical perspective. Still, at the same time, it does not correspond with the subjective emotions experienced by the data subject. In other words, people may strongly disagree with the emotions ascribed to them, which raises the question of which version is accurate. Third, individuals need to be aware of the accurate version of personal data they want to rectify, which may be complicated in the case of inferred data (the awareness problem).

This article is structured as follows. Section 2 introduces the concept of inferences, inferred personal data, and two specific types of inferred personal data: predictions and emotion data. Section 3 provides background on the accuracy principle enshrined in the GDPR, which is inextricably linked to the right to rectification. Section 4 outlines the scope of the right to rectification. Section 5 discusses the standard of objective verifiability that must be met to rectify inaccurate personal data. Section 6 investigates the conceptual and practical challenges of the right to rectification. These include the verifiability problem, the objectivity problem and the awareness problem. Section 7 discusses possible solutions for the problems identified. Section 8 provides conclusions.

2. Inferences and Inferred Personal Data

In this section, we provide some technological background on inferences, which helps define inferred personal data and contextualise the two specific types of inferred personal data we focus on. Section 2.1 discusses predictions, and Section 2.2 elaborates on emotion data. In Section 6, we rely on these two types of inferred personal data to highlight the conceptual and practical problems that arise when the right to rectification is applied to inferred personal data.

In everyday use, inferences are defined as 'a guess that you make or an opinion that you form based on the information you have'[9] or 'something that you can find out indirectly from what you already know'.[10] Inference is the process whereby a conclusion is drawn without complete certainty but with some degree of probability.[11] Statistical inference, or 'learning' in computer science, typically concerns finding patterns in data, such as correlations, classifications and clustering.[12]

[9] See <https://dictionary.cambridge.org/dictionary/english/inference?q=inferences> accessed 12 November 2024.
[10] See <https://www.oxfordlearnersdictionaries.com/definition/english/inference?q=inference> accessed 12 November 2024.
[11] Michael P Cohen, 'Inference' in Paul J Lavrakas (eds), *Encyclopedia of Survey Research Methods* (Sage Publication, Inc 2008) 334.
[12] Larry Wassermann, *All of Statistics* (Springer 2004) ix.

Correlations can be identified through regression, a statistical approach to identify the relationship between variables.[13] Classification orders data into exhaustive and exclusive groups or classes based on similarity.[14] Clustering identifies groupings in a dataset: similar patterns are placed in the same group, while all others are put in different groups.[15] All these types of inferences enable decision-making under conditions of uncertainty.[16] Any inferential method is built on assumptions[17] that may or may not be correct, which means reliability may be limited.

Machine learning, data mining and AI are typically tools that can be used for inferences in large datasets.[18] When applying these tools, a distinction can be made between the training and inferencing phase of models. The training phase is where the model learns the weights for the neural network that is being trained. In the inference phase, the model computes the weights via forward propagation.[19] Forward propagation is the running of a neural network from inputs to outputs.[20] In the inference phase, the model receives input data from the user, feeds it into the model, and exhibits output to users. Hence, AI inference refers to putting a trained model into production.[21] To avoid semantic discussions,[22] we define inferred personal data as data resulting from inferences generated by automated means. In this article, we focus on inferred personal data generated by AI.

2.1 Machine Learning and Predictions

Put simply, machine learning is a set of computational methods that use experience to improve its performance or to make accurate predictions.[23] This is achieved by using algorithms that learn from experience.[24] Learning in this context is about making computers modify or adapt their performance (actions) so that these actions

---

[13] However, note that decision tree regression would not be considered as traditional statistics.
[14] Toon Calder and Bart Custers, 'What is Data Mining and How Does it Work?' in Bart Custers et al. (eds) *Discrimination and Privacy in the Information Society* (Springer 2013) 32.
[15] Vijay Kotu and Bala Deshpande, *Data Science* (2nd edn, Elsevier 2019) 11; Toshinori Munakata, *Fundamentals of the New Artificial Intelligence* (2nd edn, Springer 2008) 72.
[16] Lawrence Hazelrigg, 'Inference' in Melissa Hardy and Alan Bryman (eds), *Handbook of Data Analysis* (Sage Publications 2004) 14.
[17] Michael Betancourt, 'A Unified Treatment of Predictive Model Comparison' (2015) 1 <https://arxiv.org/pdf/1506.02273.pdf> accessed 12 November 2024.
[18] Wassermann (n 12) ix.
[19] Miro Hodak et al, 'Benchmarking AI Inference: Where we are in 2020' in Raghunath Nambiar and Meikel Poess (eds), *Performance Evaluation and Benchmarking* (Springer 2021) 93.
[20] Stephanie Kay Ashenden et al, 'Introduction to artificial intelligence and machine learning' in Stephanie Kay Ashenden (ed), *The Era of Artificial Intelligence, Machine Learning, and Data Science in the Pharmaceutical Industry* (Academic Press 2021) 15–26.
[21] Hodak et al (n 19) 96.
[22] Concerning the arguably different meanings of inferences in statistics, data science and computer science and varying definitions of AI.
[23] Mehryar Mohri, Afshin Rostamizadeh and Ameet Talwalkar, *Foundations of Machine Learning* (MIT Press 2012) 1.
[24] Ian Goodfellow, Yoshua Bengio and Aaron Courville, *Deep Learning* (MIT Press 2016) 97 <www.deeplearningbook.org> accessed 12 November 2024.

become more accurate.[25] The main goal of machine learning is to generate accurate *predictions* for unseen data and to design efficient algorithms to produce these predictions.[26] In essence, predictions inferred by machine learning constitute 'educated guesses or bets, based on large amounts of data'.[27]

Machine learning generates probable yet inevitably uncertain knowledge.[28] For this reason, predictions generated by machine learning create tensions with the accuracy principle enshrined in the GDPR. Such data is probabilistic by nature, uncertain, not based on human reasoning and can thus be inaccurate.[29] If predictions generated by machine learning are considered *facts*, despite their probabilistic nature, this will have a tangible impact on humans, mainly because such data relates to future conduct that has yet to happen. This might harm data subjects (e.g., when applying for a loan or insurance).

2.2 Affective Computing and Emotion Data

Affective computing, sometimes called 'emotion AI', is computing that relates to, arises from or influences emotion.[30] Affective computing is a scientific and engineering endeavour inspired by psychology, neuroscience, linguistics and related areas.[31] Affective computing distinguishes between single-modal and multi-modal affect recognition approaches. Single-modal approaches are divided into text sentiment analysis, audio emotion recognition, visual emotion recognition focusing on facial expression and body gestures, and physiological-based emotion recognition systems.[32] Affective computing infers *emotion data*, namely information about an individual's emotions. Several studies have questioned the accuracy of emotion data inferred using affective computing.[33] Thus, emotion data generated through affective

[25] Steven Marsland, *Machine Learning: An Algorithmic Perspective* (2nd edn, Chapman & Hall 2015) ch 1.2.1.
[26] Mohri, Rostamizadeh and Talwalkar (n 23) 2.
[27] Teresa Scantaburlo, Andrew Charleswoth, Nello Cristianini, 'Machine Decisions and Human Consequences' in Karen Yeung and Martin Lodge (eds) *Algorithmic Regulation* (OUP 2019) 57; Lee A Bygrave, 'Machine learning, Cognitive Sovereignty and Data Protection Rights with Respect to Automated Decisions' (2021) University of Oslo Faculty of Law Legal Studies Research Article Series No. 202-35, 5.
[28] Brent Daniel Mittelstadt et al, 'The ethics of algorithms: Mapping the debate' (2016) Vol. 3 No. 2 *Big Data & Society* 1, 4.
[29] Bart Custers, 'Effects of Unreliable Group Profiling by Means of Data Mining' in Gunter Grieser, Yuzuru Tanaka and Akihiro Yamamoto (eds), *Lecture Notes in Artificial Intelligence, Proceedings of the 6th International Conference on Discovery Science* (Springer 2003) 290–295.
[30] Rosalind W Picard, 'Affective Computing' (1995) MIT Media Laboratory Perceptual Computing Section Technical Report No. 321, 1 <https://hd.media.mit.edu/tech-reports/TR-321.pdf> accessed 12 November 2024.
[31] Rafael Calvo et al, 'Introduction to Affective Computing' in Rafael Calvo et al (eds), *The Oxford Handbook of Affective Computing* (OUP 2015) 2.
[32] Yan Wang et al, 'A systematic review on affective computing: emotion models, databases, and recent advances' (2022) Vols. 83/84 *Information Fusion* 19–52.
[33] Lisa Feldman Barrett et al., 'Emotional Expressions Reconsidered' (2019) Vol. 20 No. 1 *Psychological Science in the Public Interest* 1–68; Damian Dupré et al, 'A performance

computing create severe tensions with the accuracy principle and lead to issues regarding the right to rectification. Emotion data generated by affective computing systems represent unproven and factually uncertain information about the emotional states of individuals. When emotion data are considered as facts, despite their questionable accuracy, this might unduly impact people's lives and access to opportunities.[34]

3. The Accuracy Principle

The accuracy principle according to Article 5(1)(d) GDPR is inherently intertwined with the right to rectification. It states that the processing of personal data must be accurate and, where necessary, kept up to date. Data controllers must take every reasonable step to rectify or erase inaccurate personal data without delay, notably on the controller's own initiative. The term 'reasonable' arguably implies that it is legitimate for data controllers to consider cost and resource factors when deciding upon measures to rectify or erase inaccurate data.[35] The accuracy principle intends to protect the individual concerned from being irrationally or unfairly treated based on wrong and inaccurate representations.[36] According to the accountability principle and CJEU case law, the burden of proof regarding compliance with principles enshrined in Article 5(1) GDPR lies with the controller.[37] This also applies to the accuracy principle.

The exact substantive requirements of the accuracy principle still need to be explored. According to regulatory guidance, accurate means 'accurate as to a matter of fact'.[38] The need for personal data to mirror the reality regarding the data subject concerned is also stressed in academia:[39] personal data shall, at any given time, reflect reality.[40] In the *Nowak* case, however, the Court of Justice of the European Union (CJEU) has ruled that 'the assessment of whether personal data is accurate and

---

comparison of eight commercially available automatic classifiers for facial affect recognition' (2020) Vol. 15 No. 4 *PLoS ONE* 1, 10; Kate Crawford et al, 'AI Now Report' (2019) AI Now Institute 12 <https://ainowinstitute.org/publication/ai-now-2019-report-2> accessed 12 November 2024; Margaret Lech et al, 'Real-Time Speech Emotion Recognition Using a Pre-trained Image Classification Network: Effects of Bandwidth Reduction and Computing' (2020) Vol. 2 *Frontiers in Computer Science* 1, 3.

[34] Crawford et al (n 33).

[35] Lee A Bygrave, *Data Privacy Law: An International Perspective* (OUP 2014) 164.

[36] Hallinan and Zuiderveen Borgesius (n 5) 9.

[37] Case C-175/20 *'SS' SIA* [2022] ECR I-124 paras 77, 81.

[38] Art 29 Working Party, 'Guidelines Google Spain' (WP 225, 26 November 2014) 15.

[39] Ibid; Hallinan and Zuiderveen Borgesius (n 5) 4.

[40] Paul Voigt and Axel von dem Bussche, *The EU General Data Protection Regulation (GDPR): A Practical Guide* (Springer 2017) 91; similarly, Tobias Herbst, 'Art. 5 Grundsätze für die Verarbeitung personenbezogener Daten' in Jürgen Kühling and Benedikt Buchner (eds), *DatenschutzGrundverordnung/BDSG* (2nd edn, Beck 2018) 229, para 60; Sebastian Dienst, 'Lawful Processing of Personal Data in Companies under the GDPR' in Daniel Rücker and Tobias Kugler (eds), *New European General Data Protection Regulation: A Practitioner's Guide* (Beck/Hart/Nomos 2018) 68, para 326.

complete must be made in the light of the purpose'.[41] In this case, the purpose of processing consisted of evaluating the level of knowledge and competence of a candidate during a professional examination. That level is revealed precisely by any errors in the candidate's answers. Therefore, the CJEU ruled that errors do not represent inaccuracy within the meaning of the accuracy principle.[42] Hence, the level of accuracy of personal data is determined by the purpose of processing. This is emphasised in the wording of the accuracy principle in Article 5(1)(c) GDPR and the wording of the right to rectification in Article 16 GDPR.[43] Thus, two distinct types of accuracy can be derived. We call them absolute and relative accuracy.[44] *Absolute accuracy* refers to 'accurate as a matter of fact' aiming to reflect reality[45] (e.g., data of birth) as regulatory guidance suggests.[46] *Relative accuracy* is more nuanced and determines accuracy based on the purpose of processing[47] (e.g., the accuracy needed to get accurate results).

In this article, we focus on relative accuracy because the CJEU's interpretation of accuracy is legally binding as opposed to regulatory guidance. Relative accuracy is a convincing concept because it considers the context of processing. It makes sense to set a high level of accuracy for processing occurring in contexts that may have adverse effects on individuals. Consider situations entailing power inequalities, such as employment, access to certain services (e.g., to obtain a loan or insurance), or opportunities (e.g., university admission). By contrast, significantly lower levels of accuracy are acceptable where effects for individuals are insignificant. An example is Netflix's recommender system, which predicts whether someone is interested in a movie or series.[48] With targeted advertising, accuracy only needs to be higher than untargeted advertising to make it attractive to companies, leaving aside societal problems that may arise.[49]

The unclear substantive requirements of the accuracy principle are problematic when considering the developments in AI and its significance concerning the right to rectification.[50] The accuracy principle does not outline specific levels of accuracy that

[41] Case C-434/16, *Nowak* [2017] ECR I-994 para 53; see also Case C-434/16, *Nowak* [2017] ECR I-994 Opinion AG Kokott para 35.
[42] Ibid.
[43] Hallinan and Zuiderveen Borgesius (n 5) 4.
[44] See also Andreas Häuselmann, 'EU Privacy and Data Protection Law Applied to AI: Unveiling the Legal Problems for Individuals' (PhD thesis, Leiden University 2024) 134 <https://scholarlypublications.universiteitleiden.nl/handle/1887/3747996> accessed 12 November 2024.
[45] Voigt and von dem Bussche (n 40) 91.
[46] Art 29 Working Party (n 38) 15.
[47] Case C-434/16, *Nowak* [2017] ECR I-994 para 53.
[48] Harald Steck et al, 'Deep learning for recommender systems: A Netflix case study' (2021) Vol. 42 No. 3 *AI Magazine* 7.
[49] For instance filter bubbles; see Eli Pariser, *The Filter Bubble* (Penguin Books 2012).
[50] Diana Dimitrova, 'The rise of the personal data quality principle: is it legal and does it have an impact on the right to rectification?' (2021) Vol. 12 No. 3 *European Journal of Law and Technology* 2.

personal data processed in the context of AI must reach. There is no one-size-fits-all approach[51] considering that the level of accuracy depends on the purpose of processing when interpreted as relative accuracy.[52] In addition, regulators have neglected the accuracy principle by not providing substantive guidance apart from the statement that accuracy means 'accurate as a matter of fact'.[53]

When looking for more specific approaches that are helpful to interpret the accuracy principle in the context of inferred data, it is impossible to simply refer to the concept of accuracy or information quality in computer science.[54] Information quality goes far beyond the accuracy principle contained in the GDPR.[55] In computer science, information quality is a multidimensional concept[56] and covers at least four dimensions: intrinsic; contextual; representational; and accessibility information quality. What exactly falls under the scope of these four dimensions varies from the perspectives of academics and practitioners.[57] Further clarification and formalisation of these dimensions are required.[58] Nevertheless, accuracy is often explicitly[59] considered as an intrinsic information quality dimension[60] and is, therefore, particularly interesting in the context of the accuracy principle.

Hence, it is obvious that more interdisciplinary research is needed to develop an interpretation of the accuracy principle, which is valid and practical from both a legal and computational perspective. Interdisciplinary research into information quality and accuracy is a relatively recent development.[61] This causes problems regarding the GDPR, but these issues will only intensify with the AI Act, as the latter mentions accuracy several times, mainly in the context of high-risk systems.[62] The AI Act requires an appropriate level of accuracy, which needs to be assessed in light of the

---

[51] Hallinan and Zuiderveen Borgesius (n 5) 4.
[52] Case C-434/16, *Nowak* [2017] ECR I-994 para 53.
[53] Art 29 Working Party (n 38) 15.
[54] Hallinan and Zuiderveen Borgesius (n 5) 4.
[55] Dimitrova (n 50) 9–10; Luciano Floridi and Phyllis Illari, 'Information Quality, Data and Philosophy' in Luciano Floridi and Phyllis Illari (eds), *The Philosophy of Information Quality* (Springer Nature 2014) 6.
[56] Floridi and Illari (n 55) 6; Leo Pipino et al, 'Developing Measurement Scales for Data Quality Dimensions' in Richard Y Wang et al (eds) *Information Quality* (1st edn, Routledge 2005) 37.
[57] Yang W Lee et al, 'AIMQ: a methodology for information quality assessment' (2002) Vol. 40 No. 2 *Information & Management* 133, 134, 136; Floridi and Illari (n 55) 6.
[58] Carlo Batini, Matteo Palmonari and Gianluigi Viscusi, 'Opening the Closed World: A Survey of Information Quality Research in the Wild' in Luciano Floridi and Phyllis Illari (eds), *The Philosophy of Information Quality* (Springer Nature 2014) 44.
[59] Lee et al (n 57) 133, 134, 136; Carlo Batini and Monica Scannapieco, *Data Quality* (Springer 2006) 20–27; Floridi and Illari (n 55) 7.
[60] Also, contextual information quality is at least partially relevant for the accuracy principle as it often refers to the term completeness. However, it also contains other less relevant aspects such as timeliness; see also Lee et al (n 57) 133, 134, 136; Batini, Palmonari and Viscusi (n 58) 60.
[61] Schäfer (n 6) 217.
[62] Articles 13(3), 15 and 58, and Annex IV paras 2–3 AI Act, see also Recitals 59, 60, 66, 74 and 122.

AI system's intended purpose.[63] Since this is similar to accuracy in data protection law, the shortcomings outlined in this section are likely to apply also to accuracy under the AI Act.

4. The Scope of the Right to Rectification

The right to rectification in Article 16 GDPR enables the data subject to 'the rectification of inaccurate personal data' and 'to have incomplete personal data completed, including by providing a supplementary statement'.[64] Providing a supplementary statement adds missing elements rather than rectifying inaccurate personal data.[65] It is unclear what specific obligations such a supplementary statement imposes on the data controller.[66] The right to rectification is an underexplored provision in academia and regulatory guidance. The same applies to CJEU case law: only three rulings explicitly deal with the right to rectification.[67] One case, *Deldits*, is pending at the CJEU (discussed in Section 5).[68] Nevertheless, in our view, the right to rectification will play a more prominent role in the future due to the developments in AI.

For the discussion in Section 6, it is necessary to outline the proper scope of Article 16 GDPR. Based on a teleological interpretation, it is clear that the right to rectification should cover everything that constitutes personal data.[69] This includes inferred personal data, such as predictions and emotion data. In this section, we discuss why we disagree with limiting the scope of this right as suggested within academia[70] and the opinions of the CJEU's Advocates-General (AGs).

The *first limitation* suggested in the literature and case law reduces the scope of the right to rectification to factual[71] and input data,[72] thereby excluding inferred personal data.[73] At least implicitly, AG Pikamäe significantly limits the right to rectification concerning the automated establishment of a credit score. According to the AG, data subjects may enforce their right to rectification 'if the *personal data used to carry out*

[63] Annex IV para 3, see also Recital 74 AI Act.
[64] Article 16(1) GDPR.
[65] Cécile de Terwangne, 'Commentary of Article 16' in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 473.
[66] Dimitrova (n 50) 27.
[67] As previously adopted in Member States under the Data Protection Directive, but none under the GDPR.
[68] Case C-247/23, *Deldits.*
[69] Bart Custers and Helena Vrabec, 'Tell me something new: data subject rights applied to inferred data and profiles' (2024) Vol. 52 *Computer Law & Security Review* 2, 9, 10.
[70] Wachter and Mittelstadt (n 2) 550.
[71] Joined Cases C-141/12 & C-372/12, *YS, M* [2014] ECR I-2081, Opinion of AG Sharpston para 56; European Data Protection Supervisor, 'Guidelines on the Rights of Individuals with regard to the Processing of Personal Data' (25 February 2014) 18.
[72] Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-220, Opinion AG Pikamäe para 50.
[73] Wachter and Mittelstadt (n 2) 550.

the scoring should prove to be inaccurate'.[74] This reduces the right to rectification to input data, i.e., to the personal data used to establish the credit score. Simultaneously, it excludes the output in the form of the calculated credit score (inferred personal data). We disagree with this suggested limitation. In the corresponding ruling, the CJEU rightly did not follow the AG's opinion.[75] According to the wording enshrined in Article 16 GDPR, this right applies to the rectification of 'inaccurate personal data'. In our view, it is irrelevant whether such personal data constitutes factual data, input data or inferred data (output) *as long as it is personal data*. Nothing in the preparatory documents of the GDPR indicates the legislator's intention to limit this right to factual and input data. Besides, such a limitation would contradict the CJEU's contextual and teleological approach to interpreting data subject rights.[76] Our broad interpretation is also in line with regulatory guidance, according to which the right to rectification applies to 'input data' and 'output data'.[77]

Limiting the right to rectification to input data, as suggested by the AG,[78] indicates that the data subject should not try to rectify the output (inferred data) but rather rectify the input. For instance, if the data subject is convinced that it should have a more favourable credit score, this should be remedied by rectifying or supplementing the input data instead of the inferred personal data (output). Rectifying inaccurate input data and providing additional input data might make the inferred personal data more accurate. However, this would not solve the problem that predictions always relate to the future and are unverifiable (see Section 6). Arguably, the AG's limitation intends to avoid that data subjects can change inferred data that they may not like. Obviously, this could raise practical concerns, as data controllers may have to deal with many such requests, leading to debates on what the 'correct' score should be. We argue, however, that the scope of Article 16 GDPR is defined by whether the data is accurate. If a credit score is indeed accurate, even if a data subject may not like that score, the data controller is not required to adjust the credit score. Obviously, the question here is what defines a correct credit score: a correct calculation based on accurate data (i.e., a more technical perspective) or a fairly assessed score in which data subjects can recognise themselves (i.e., a more ethical and legal perspective)? As already pointed out in Section 3, interdisciplinary research is needed to interpret accuracy in a way that is valid and practical from both a legal and computational perspective.

The *second limitation* suggested in literature and case law excludes inferred personal data in the form of opinions or assessments from the scope of Article 16 GDPR.

---

[74] Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-220, Opinion AG Pikamäe para 50, emphasis added.

[75] The Court neglected the right to rectification because this case deals with Article 22 GDPR; see Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-957.

[76] Case C-434/16, *Nowak* [2017] ECR I-994 paras 53–54.

[77] Art 29 Working Party 'Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679' (WP251rev.01, 6 February 2018) 8–9, 18.

[78] Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-220, Opinion AG Pikamäe para 50, emphasis added.

Referring to CJEU case law, Wachter and Mittelstadt argue that inferred personal data cannot be rectified under data protection law as such data constitute opinions and assessments.[79] This view is based on a non-contextual reading of the CJEU's case law and is wrong from our perspective. In *Nowak*, the CJEU held that the right to rectification may also be asserted concerning written answers submitted by the candidate in the context of a professional examination, including comments made by an examiner.[80] However, the right to rectification must be interpreted teleologically. Obviously, the right to rectification should not result in situations where a candidate for a professional examination would be allowed to correct his answers in an exam retroactively.[81] In joined cases *YS*, the CJEU ruled that a person involved in an immigration case cannot rectify the content of a legal analysis by enforcing the right to rectification.[82] These two contextual and normative limitations adopted by the CJEU are justified and necessary to avoid an interpretation of the right to rectification that is excessively broad or 'over-inclusive'.[83] The right to rectification is also not intended to change value judgments[84] as this would contradict the freedom of expression and information according to Article 11 EUCFR.[85] Invoking the right to rectification would involve trying to forcibly change the data controller's assessment or opinion, arguably interfering with the rights and freedoms of the data controller.[86]

Apart from these justified limitations, opinions and assessments relating to a data subject fall under Article 16 GDPR. Opinions and assessments relating to a particular data subject constitute personal data, according to the CJEU.[87] In the words of the CJEU, personal data 'encompasses all kinds of information, not only *objective* but also *subjective*, in the form of *opinions* and *assessments*, provided that it "relates" to the

[79] Wachter and Mittelstadt (n 2) 550.
[80] Case C-434/16, *Nowak* [2017] ECR I-994 para 51.
[81] Ibid para 54.
[82] Joined Cases C-141/12 & C-372/12, *YS, M and S v Minister voor Immigratie, Integratie en Asiel* [2014] ECR I-2081, para 45.
[83] Koen Lenaerts and José A Gutiérrez-Fons, 'To Say What the Law of the EU Is: Methods of Interpretation and the European Court of Justice' (2013) European University Institute Working Article AEL 2013/9 at 27.
[84] Case C-460/20, *TU* [2022] ECR I-962 para 66. In common language usage, value judgments are 'a personal opinion about whether something is good or bad' based on 'personal opinion rather than facts'; see <https://dictionary.cambridge.org/dictionary/english/value-judgment> accessed 12 November 2024.
[85] It seems questionable whether data controllers can rely on the freedom of expression and information concerning inferred personal data generated by 'machines' using AI.
[86] In practice it may be very difficult to clearly distinguish between value judgments, assessments, opinions and facts. The right to rectification does not formally require making these distinctions, contrary to the Law Enforcement Directive. The latter requires data controllers to distinguish facts from opinions when storing and processing data. This requirement is highly problematic in practice; cf. Marc Leiser and Bart Custers, 'The Law Enforcement Directive: Conceptual Issues of EU Directive 2016/680' (2019) Vol. 5 No. 3 *European Data Protection Law Review* 367.
[87] Case C-434/16, *Nowak* [2017] ECR I-994 paras 35, 46.

data subject'.[88] In our view, it seems likely that the CJEU will subsume inferred personal data generated by AI under the scope of Article 16 GDPR by applying a functional interpretation ('effet utile').[89] If personal data in the form of opinions or assessments established by humans falls under the right to rectification,[90] the same must apply to opinions and assessments established by machines.[91] In addition, it might be premature to qualify inferred personal data generated by AI as opinions when considering that AI systems have been called clueless[92] in understanding cause and effect and lack common sense[93] and legal personality. In our view, inferred data generated by AI should be considered personal data that fall under the scope of Article 16 GDPR.

5. The Standard of Objective Verifiability

The right to rectification constitutes an essential element of the fundamental right to data protection.[94] In its case law, the CJEU has repeatedly emphasised the significance of this right.[95] However, neither the GDPR, case law of the CJEU, nor regulatory guidance yields insights about the standard of proof that needs to be applied to rectify personal data. Arguably, this is caused by the Member States' *procedural autonomy*: in the absence of EU procedural law, Member States may set up any procedural system they deem fit.[96] Thus, regulating procedural law is generally considered a matter of Member State autonomy as far as it satisfies the minimum principles of effectiveness and equivalence.[97]

---

[88] Ibid para 34, emphasis added.
[89] Lenaerts and Gutiérrez-Fons (n 83) 25.
[90] Without prejudice to justified limitations as in Case C-434/16, *Nowak* [2017] ECR I-994 and Joined Cases C-141/12 & C-372/12, *YS, M and S v Minister voor Immigratie, Integratie en Asiel* [2014] ECR I-2081.
[91] A credit score generated by automated means essentially constitutes an assessment of a person's credit-worthiness and is thus a form of inferred personal data that should fall under the right to rectification; see by analogy Case C-203/22, *Dun & Bradstreet Austria* [2024] ECR I-745, Opinion AG de La Tour para 46.
[92] Brian Bergstein, 'What AI still can't do' *MIT Technology Review* (31 January 2020) https://www.technologyreview.com/2020/02/19/868178/what-ai-still-cant-do/ accessed 12 November 2024.
[93] Brandon Bennet and Anthony G Cohn, 'Automated Common-sense Spatial Reasoning: Still a Hughe Challenge' in Stephen Muggleton and Nicholas Chater (eds), *Human-Like Machine Intelligence* (OUP 2021) 405.
[94] Terwangne (n 65) 473.
[95] Case C-434/16, *Nowak* [2017] ECR I-994 para 49; Case C-362/14 *Schrems v Data Protection Commissioner* [2015] ECR I-650 para 95; Case C-553/07 *Rijkeboer* [2009] ECR I-03889 para 51.
[96] Bart Krans and Anna Nylund, 'Aspects of Procedural Autonomy' in Bart Krans and Anna Nylund (eds) *Procedural Autonomy Across Europe* (Intersentia 2020) 1.
[97] Anna Wallerman, 'Towards an EU law doctrine on the exercise of discretion in national courts? The Member States' self-imposed limits on national procedural autonomy' (2016) Vol. 53 No. 2 *Common Market Law Review* 339. These minimum principles appear in numerous cases, for instance Case C-353/20, *Skeyes* [2022] ECR I-423; Case C-497/20, *Randstad Italia SpA* [2021] ECR I-1037; Joined Cases C-222/05 and C-225/05, *Van der Weerd* [2007] ECR I-4233.

The *Deldits* case[98] will shed some light on the burden of proof concerning the right to rectification. *Deldits* concerns a data subject that wants to rectify its gender recorded in a national register of refugees. AG Collins acknowledges that the GDPR does not specify what proof a data subject must submit when requesting the rectification of inaccurate personal data. In his view, a case-by-case analysis is needed because accuracy depends on the purpose of processing. Hence, a data subject should 'produce evidence that may be reasonably required to establish the inaccuracy' of personal data in light of the purpose of processing.[99] It remains to be seen whether the CJEU follows the AG's threshold of 'reasonable proof' and how the latter can be applied to inferred personal data.

Another CJEU case, *TU*, relating to the request to erase inaccurate personal data according to Article 17(3)(a) GDPR, provides some insights about the standard of proof to be met to establish the inaccuracy of personal data processed.[100] According to the CJEU, the data subject bears the burden of proof to establish the manifest inaccuracy of the information in question. However, the CJEU sets limits to this burden of proof. To avoid an excessive burden, the data subject must provide evidence that can *reasonably be required* – similar to the threshold suggested by AG Collins in *Deldits*. The data subject must submit '*relevant* and *sufficient* evidence capable of substantiating his or her request and of establishing the *manifest inaccuracy* of the information'.[101] However, the context of this case must be borne in mind. It does not concern inferred personal data, but deals with articles published online that show pictures of the data subjects suggesting a luxury lifestyle and criticise their investment companies. In addition, the case relates to the right to erasure and balancing the fundamental rights to privacy and the protection of personal data against the fundamental right to freedom of expression and information. Therefore, we will not give much weight to the 'manifest inaccuracy' standard.

Instead, we rely on the CJEU's ruling in *Ligue des droit humains*[102] and the related PNR opinion.[103] Said case relates to 'predictive algorithms' involving the processing of passenger name records to 'identifying anyone who may be involved or might engage in criminal activities'.[104] The CJEU stressed that such automated processing relies on 'unverified personal data' based on 'pre-determined models and criteria' with a 'margin of errors' and a 'fairly substantial number of false positives'.[105] In the preceding Canada Passenger Name Records (PNR) opinion, the CJEU has pointed to the significant 'margin of error' inherent to the automated processing of personal data, particularly if such processing is carried out based on '*unverified* personal data

[98] Case C-247/23, *Deldits*.
[99] Case C-247/23, *Deldits* [2024] ECR I-747, Opinion AG Collins para 47.
[100] Case C-460/20, *TU* [2022] ECR I-962 para 68.
[101] Ibid paras 68, 72.
[102] Case C-817/19, *Ligue des droits humains* [2022] ECR I-491.
[103] Opinion 1/15 CJEU [2017] ECR I-592.
[104] Case C-817/19, *Ligue des droits humains* [2022] ECR I-491 para 58.
[105] Ibid paras 106, 124 and 178.

[…] and pre-established *models* and *criteria*'.[106] Therefore, the *Ligue des droit humains*[107] ruling and the related PNR opinion[108] suggest that rectification somehow relates to *verification* due to the terms 'verified' and 'unverified' personal data. According to the CJEU, competent authorities should give 'preference to the result of individual review conducted by non-automated means' (verified data) over that 'obtained by automated processing' (unverified data).[109] In other words, verification can be seen as a step that data controllers could take when assessing accuracy in the context of a rectification request. To facilitate such verification, the data subject needs to provide evidence allowing the controller to verify whether personal data is inaccurate and needs to be rectified.

Given the absence of CJEU case law, European Court of Human Rights (ECtHR) rulings are also worth considering. According to the ECtHR, individuals should adduce 'objectively verifiable evidence' for having personal data relating to them changed.[110] Case law on the right to rectification in the Netherlands applies a similar standard: inaccuracies in personal data must be 'easily' and 'objectively' verifiable.[111] In Germany, the standard concerning the right to rectification amounts to 'objective reality'. Correct is data that reflects reality; data is incorrect if it does not correspond with reality.[112] However, the context of these cases should also be borne in mind here. They deal with personal data that can be objectively verified (e.g., ethnicity, deregistration from school, date of birth), meaning they might significantly differ from cases with inferred personal data.

From ECtHR case law,[113] the CJEU's *Ligue des droit humains* ruling, and the related PNR opinion,[114] it can be concluded that the right to rectification relies on verification. The CJEU's ruling dealing with the erasure of inaccurate personal data also depends on verification. The data subject must submit '*relevant* and *sufficient* evidence capable of substantiating his or her request'.[115] Thus, when data subjects dispute the accuracy or completeness of personal data processed by the data controller ('current data'), they must provide objectively verifiable evidence that the 'new' personal data envisaged to replace the current data is accurate. We call this 'the standard of objective verifiability'.

---

[106] Opinion 1/15 CJEU [2017] ECR I-592 paras 169 and 170, emphasis added.
[107] Case C-817/19, *Ligue des droits humains* [2022] ECR I-491.
[108] Opinion 1/15 CJEU [2017] ECR I-592.
[109] Case C-817/19, *Ligue des droits humains* [2022] ECR I-491 208.
[110] This case concerns the claimants ethnic identity entry *Ciubotaru v Moldov* App No 27138/04 (ECtHR 27 July 2010) para 59.
[111] This case concerns the allegedly incorrect deregistration date from a school; *Raad van State*, ECLI:NL:RVS:2021:1020, 20 February 2019 para 5.1.
[112] This case concerns an allegedly incorrect data of birth contained in the data subjects passport; BVerwG – 6 C 7.20 [2022], ECLI:DE:BVerwG:2022:020322U6C7.20, para 32.
[113] *Ciubotaru v Moldov* App No 27138/04 (ECtHR 27 July 2010) para 59.
[114] Case C-817/19, *Ligue des droits humains* [2022] ECR I-491; Opinion 1/15 CJEU [2017] ECR I-592.
[115] Case C-460/20, *TU* [2022] ECR I-962 paras 68, 72.

The standard of objective verifiability is met easily when personal data is verifiable by nature (such as a name, date of birth, or email address).[116] Section 6 explains that this differs from inferred personal data generated by AI.

6. Conceptual and Practical Issues

To rectify personal data in line with the standard of objective verifiability introduced in Section 5, data subjects need to provide evidence that personal data is inaccurate or incomplete. As we will show in this section, this leads to two conceptual issues (the verifiability problem and the objectivity problem) and practical issues (the awareness problem). The verifiability problem focuses on (establishing) the inaccuracy of the personal data that needs to be replaced. The objectivity problem focuses on proving the accuracy of the personal data, aiming to replace the inaccurate personal data processed by the data controller. The awareness problem is one of the practical issues of the right to rectification. To effectively rectify inferred personal data, a data subject needs to be aware that data is inaccurate. This, however, is often not the case.

6.1 The Verifiability Problem

The verifiability problem refers to the problem of verifying the accuracy of inferred personal data in the form of predictions. If it cannot be established that the data is inaccurate, the controller can reject a data subject's rectification request.

There are many examples of predictions about individuals, as introduced in Section 2.1, including predictions on life events such as pregnancy and divorce, likelihood of success, involvement in accidents, likelihood of behaving in antisocial ways, or likelihood of committing a crime.[117] In this article, predictions refer to statements that say what someone thinks will happen in the future.[118] Predictions produced by machine learning are educated guesses based on large amounts of data.[119] Systems that predict the future behaviour of individuals cannot be designed with absolute accuracy.[120] Inferences and predictions are closely intertwined.[121] Inferences refer to the processes whereby a conclusion is drawn without complete certainty but with some degree of probability.[122] As outlined in Section 2, any inferential method is built on assumptions[123] that may be correct or incorrect. Inference is helpful for decision-making under conditions of uncertainty.[124] Nonetheless, an inference 'is always an

[116] Wachter and Mittelstadt (n 2) 548.
[117] Hideyuki Matsumi, 'Predictions and Privacy: Should There Be Rules About Using Personal Data to Forecast the Future?' (2018) Vol. 48 No. 1 *Cumberland Law Review* 149, 153.
[118] Cf. <https://www.oxfordlearnersdictionaries.com/definition/english/prediction> and <https://dictionary.cambridge.org/dictionary/english/prediction> accessed 12 November 2024.
[119] Scantaburlo, Charleswoth and Cristianini (n 27) 57; Bygrave (n 27) 5.
[120] Dimitrova (n 50) 21.
[121] Nathan Sanders, 'A Balanced Perspective on Prediction and Inference for Data Science in Industry' (2019) Vol. 1 No. 1 *Harvard Data Science Review* 1, 7, 21.
[122] Cohen (n 11) 334.
[123] Betancourt (n 17).
[124] Hazelrigg (n 16) 14.

invasion of the unknown, a leap from the known'.[125] Thus, the very nature of inferences and predictions increases the risk of inaccuracy, given their probabilistic nature.[126]

Often, predictions are not objectively verifiable,[127] mainly because they relate to future conduct that has yet to happen. Predictions are poorly verifiable in the sense that they cannot be verified immediately after they are created (e.g., the individual is a 'high credit risk' or 'likely to buy a house in two years') or not at all.[128] In this sense, predictions are neither true nor false as the asserted matter has not yet materialised.[129] As described by Matsumi and Solove, predictions are merely projections of a *possible* future from the viewpoint of the past and the present.[130] Due to the lack of a 'ground truth' as a baseline for comparison,[131] predictions are not verifiable – they merely represent unverifiable personal data concerning the data subject's future life. This leads to the verifiability problem: the accuracy of predictions cannot be verified. Consequently, data controllers can reject a data subject's rectification request, although predictions are likely inaccurate due to their probabilistic nature. This is highly problematic, particularly when predictions are considered *facts.* The latter might harm data subjects (e.g., when applying for a job or a loan). Experimental evidence shows that humans follow algorithmic output *closely* and cannot correctly *assess* the quality thereof.[132] Relying on merely probabilistic data could lead to severe consequences for data subjects. For instance, predictions may propagate existing biased patterns, leading to disparate impact.[133]

---

[125] John Dewey, *The Middle Works of John Dewey, Volume 9, 1899–1924* (Carbondale Southern Illinois University Press 1980) 165.

[126] Christopher Burr and Nello Cristianini, 'Can machines read our minds?' (2019) Vol. 29 No. 3 *Minds and Machines* 461, 483.

[127] Jef Ausloos, Michael Veale and René Mahieu, 'Getting Data Subject Rights Right' (2019) Vol 10 Iss 3 JIPITEC 283, 302.

[128] Wachter and Mittelstadt (n 2) 510.

[129] Hideyuki Matsumi and Daniel J Solove, 'The Prediction Society' (2023) GWU Legal Studies Research Paper No. 2023-58, 26 <https://papers.ssrn.com/sol3/articles.cfm?abstract_id=4453869> accessed 12 November 2024. Compare this with Schrödinger's cat in quantum physics: in this thought experiment, there is a cat in a box and cannot be seen from outside of the box. If you want to know whether the cat is alive or dead, the only option is to open the box. As long as the box is not opened, essentially the cat can be both alive and dead (i.e., both can be true statements) from the outsider's perspective. See Amit Goswami, 'The Paradox of Schrödinger's Cat' in Amit Goswami, *The Physicists' View of Nature Part 2* (Springer 2001) 139–146.

[130] Matsumi and Solove (n 129) 21.

[131] Dimitrova (n 50) 21, Wachter and Mittelstadt (n 2) 548.

[132] Jan Biermann, John Horton and Johannes Walter, 'Algorithmic Advice as a Credence Good' (2022) Centre for European Economic Research Discussion Article No. 22-071, 2, 14 <https://articles.ssrn.com/sol3/articles.cfm?abstract_id=4326911> accessed 12 November 2024.

[133] Bart Custers, 'Profiling as inferred data. Amplifier effects and positive feedback loops' in Emre Bayamachine learningıoğlu, Irina Baraluic, Liisa Janssens and Mireille Hildebrandt (eds), *Being Profiled: Cogitas Ergo Sum* (Amsterdam University Press 2018) 115.

Despite these consequences, individuals lack any meaningful ability to challenge predictions,[134] even when invoking their right to rectification.

According to the CJEU, data subject rights contained in the GDPR must be effective.[135] However, the right to rectification is ineffective when applied to predictions and other types of inferred personal data. Due to the standard of objective verifiability, data subjects cannot enforce their right to rectification. According to the CJEU, merely facts are susceptible of proof.[136] However, concerning predictions, factual evidence eligible to verify inaccuracy is absent in most cases due to the unverifiable nature of predictions.[137] Due to the lack of such evidence, data subjects cannot effectively enforce their right to rectification for personal data which might be inaccurate due to its probabilistic nature. To use the words of the CJEU: it becomes '*virtually impossible or excessively difficult*'[138] for data subjects to enforce the right to rectification. Matsumi and Solove note that unverifiable predictions are not inaccurate.[139] According to them, the real issue is whether such predictions are fair and not causing unwarranted harm – considerations that do not fit with the 'true–false binary' underlying the right to rectification. However, as explained in Section 2, accuracy in data protection law is not a 'true–false binary'. Instead, accuracy is a relative concept because it depends on the purpose for processing.[140] Thus, in our view, the real issue is the lack of clarity regarding the substantive requirements of the accuracy principle, without which it is impossible to assess the accuracy of unverifiable personal data.

6.2 The Objectivity Problem

The objectivity problem arises when data subjects have to prove the accuracy of personal data envisaged to replace the inaccurate personal data currently processed by the data controller. To meet the standard of objective verifiability (Section 5), data subjects must provide objectively verifiable evidence to rectify inaccurate personal data. This standard can be easily met in case of a name or a date of birth. With inferred personal data, this is fundamentally different: there is a lack of 'ground truth' serving as a baseline comparison for predictions. Due to the unverifiable nature of predictions (Section 2.1), data subjects cannot provide objectively verifiable evidence that their own predictions of the future are accurate. The objectivity problem also occurs when other types of inferred personal data are used. Here, we illustrate the objectivity problem through the example of emotion data.

Affective computing systems generate emotion data (Section 2.2). In this article, we define emotion data as information relating to an individual's emotions. Emotions

---

[134] Matsumi and Solove (n 129). See also Custers and Vrabec (n 69).
[135] Case C-154/21 *Österreichische Post AG* [2023] ECR I-3 para 39.
[136] Case C-460/20, *TU* [2022] ECR I-962 para 66.
[137] Matsumi (n 117) 150, 205.
[138] Case C-353/20, *Skeyes* [2022] ECR I-423 para 52; Case C-497/20, *Randstad Italia SpA* [2021] ECR I-1037 para 58, Joined Cases C-222/05 and C-225/05, *Van der Weerd* [2007] ECR I-4233 para 28; *Jeroen van Schijndel et al* [1995] ECR I-4705 para 17.
[139] Matsumi and Solove (n 129).
[140] Case C-434/16, *Nowak* [2017] ECR I-994 para 53.

refer to the six emotion categories[141] commonly used in emotion research: anger; disgust; fear; happiness; sadness; and surprise.[142] These 'basic emotions'[143] categories have received the most attention in scientific research.[144] Most approaches deployed by affective computing rely on basic emotion categories[145] or alterations,[146] although the basic emotion taxonomy was quickly subject to substantial disagreement.[147] The corresponding critique continued steadily.[148]

Different studies have rebutted that a person's emotional state can accurately be inferred from their facial movements.[149] It is impossible to confidently infer happiness from a smile, anger from a scowl, or sadness from a frown because these emotion categories are more variable in their facial expressions.[150] Another study revealed that the accuracy levels of eight commercial automatic classifiers used for facial affect recognition were consistently lower when applied to spontaneous affective behaviours when compared to 'posed' affective behaviour. Validation accuracy rates of the tested classifiers varied from 48% to 62%.[151] Also, other means to detect emotions, for example, based on speech and physiological data, have been called into question due to a lack of scientific consensus on whether such methods can ensure accurate or even valid results.[152] Studies indicate that speech compression, filtering,

---

[141] These six emotions refer to research conducted by psychologists in the early seventies that developed the methodology of 'basic emotions'; see Paul Ekman and Wallace v Friesen, 'Constants across cultures in the face and emotion' (1971) Vol. 17 No. 2 *Journal of Personality and Social Psychology* 124.

[142] Feldman Barrett et al (n 68) 52.

[143] Eiman Kanjo et al, 'Emotions in context: examining pervasive affective sensing systems, applications, and analyses' (2015) Vol. 19 *Personal and Ubiquitous Computing* 1197, 1204 <https://link.springer.com/content/pdf/10.1007/s00779-015-0842-3.pdf> accessed 12 November 2024.

[144] Feldman Barrett et al (n 68) 3.

[145] Sidney D'Mello and Rafael A Calvo, 'Beyond the basic emotions: what should affective computing compute?' (2013) CHI Changing Perspectives Conference, Paris, April–May, 2289 <https://dl.acm.org/doi/pdf/10.1145/2468356.2468751> accessed 12 November 2024.

[146] Andrius Dzedzickis, Artūras Kaklauskas and Vytautas Bucinskas, 'Human Emotion Recognition: Review of Sensors and Methods' (2020) Vol. 20 No. 3 *Sensors* 1, 2; Klaus Scherer, 'Emotions are emergent processes: they require a dynamic computational architecture' (2009) Vol. 364 No. 1535 *Philosophical Transactions of the Royal Society* B 3459, 3462.

[147] See for an overview: Andrew Ortony and Terence J Turner, 'Whats Basic About Basic Emotions?' (1990) Vol. 97 No. 3 *Psychological Review* 315–331; Richard S Lazarus, *Emotion and Adaption* (OUP 1991) 71; James A Russell, 'Core Affect and the Psychological Construction of Emotion' (2003) Vol. 110 No. 1 *Psychological Review* 145; Lisa Feldman Barrett, 'Are emotions natural kinds?' (2006) Vol. 1 No. 1 *Perspectives on Psychological Science* 28.

[148] Russell (n 147); Feldman Barrett (n 147) 8; Feldman Barrett et al (n 68).

[149] Feldman Barrett et al (n 68); Sara Preto, 'Emotion-reading algorithms cannot predict intentions via facial expressions' *USC News* (Los Angeles, 4 September 2019) <https://news.usc.edu/160360/algorithms-emotions-facial-expressions-predict-intentions/> accessed 12 November 2024.

[150] Feldman Barrett et al (n 68) 46.

[151] Dupré et al (n 33) 10.

[152] Crawford et al (n 33).

band reduction and the addition of noise reduce accuracy significantly in speech emotion recognition.[153] Despite this, such systems are already protected by patents[154] or applied 'into the wild'.[155]

Consequently, emotion data inferred by affective computing systems are likely inaccurate. Emotion data represent unproven and factually uncertain information about individuals' emotional states. Now, the question arises of how data subjects may rectify inaccurate emotion data according to the standard of objective verifiability when enforcing the right to rectification. This is a rather tricky endeavour due to emotion data's subjective nature.

The partial perspective is a basic characteristic common to emotions. Partial, in this sense, means that emotions always express a personal perspective.[156] Emotions feel a *certain way* for the individual experiencing the emotion.[157] For instance, someone feels or reacts angrily at a particular time and place.[158] Therefore, emotions can only be a fact for the individual experiencing the emotion: every individual has their own *personal* experience of emotions.[159] This is due to the *subjective* perception of emotions: emotions feel a certain way uniquely for the person undergoing an emotional experience.[160] Therefore, information about someone's emotional state is difficult to assess objectively.[161] Emotions are subjective in the sense that they express an exclusively personal perspective.[162] Consequently, emotion data is not objectively verifiable because every individual has their own personal experience of emotion.[163] Instead, it is subjectively verifiable: emotion data can uniquely be verified by the individual experiencing the emotion in question.

Due to the subjective nature of emotion data, no facts are available to objectively prove the accuracy of the personal data envisaged to replace the personal data currently processed by the controller. The data subject can merely provide a simple statement indicating that it has experienced another emotion. However, such a 'self-

---

[153] Lech et al (n 33) 3.
[154] Daniel Bromand et al, 'Systems and Methods for Enhancing Responsiveness to Utterances Having Detectable Emotion' US Patent Number US 10566010 B2 (Assignee: Spotify AB) February 2020, 11; Josh Mandell, 'Spotify Patents A Voice Assistant That Can Read Your Emotions' *Forbes* (12 March 2020) <https://www.forbes.com/sites/joshmandell/2020/03/12/spotify-patents-a-voice-assistant--that-can-read-your-emotions> accessed 12 November 2024.
[155] See for instance <https://www.ventureradar.com/keyword/Affective%20computing> accessed 12 November 2024.
[156] Aaron Ben-Ze'Ev, *The Subtlety of Emotions* (MIT Press 2000) 13, 35.
[157] Joel Smith and Catharine Abell, 'Introduction: Emotional Expression' in Catharine Abell, Joel Smith (eds) *The Expression of Emotion* (CUP 2016) 1.
[158] Lazarus (n 147) 46, 47.
[159] Jennifer Healey, 'Physiological Sensing of Emotion' in Rafael Calvo et al (eds), *The Oxford Handbook of Affective Computing* (OUP 2015) 213, 214.
[160] Smith and Abell (n 157) 1–3.
[161] Michèle Finck, 'The Limits of the GDPR in the Personalisation Context' (2021) Max Planck Institute for Innovation and Competition Research Article No. 21-11, 9.
[162] Ben-Ze'Ev (n 156) 35.
[163] Healey (n 159) 213, 214.

declaration'[164] does not constitute objectively verifiable evidence. On the contrary, it constitutes an exclusively subjective statement. Hence, data subjects cannot adhere to the standard of objective verifiability. As a consequence, data subjects cannot enforce their right to rectification. This is highly problematic, particularly because emotion data as such do not constitute special data as defined in Article 9 GDPR, despite its sensitive nature.[165]

6.3 The Awareness Problem

Beyond the conceptual issues discussed in the previous sections, there are also practical issues. Most notably, to effectively rectify inferred personal data, a data subject needs to be aware that the data is inaccurate (the awareness problem). This may not be obvious to a data subject. A data subject must know about the data, but inferred data are often treated as trade secrets and not shared with data subjects, even when they invoke their right of access.[166]

Machine learning and affective computing enable controllers to infer personal data such as predictions or emotion data based on personal data provided by data subjects or obtained from third-party sources. Transparency obligations contained in the GDPR do not require data controllers to inform data subjects about the specific personal data inferred (e.g., predictions or specific emotional states) if the data are processed for compatible purposes.[167] Consequently, data subjects are unaware of such data and cannot exercise their right to rectification.[168] Data controllers are not obliged to inform data subjects about specifics of inferred personal data (e.g., specific prediction or detected emotional state) if the initial personal data ('input') are directly collected from data subjects.[169] Regulatory guidelines on transparency[170] confirm that informing data subjects about the category of inferred personal data is sufficient according to Article 13 GDPR. As a consequence, data subjects cannot enforce their right of rectification regarding inferred personal data in the form of predictions and emotion data, nor can they assess the accuracy thereof. Data subjects could enforce their right of access, but empirical research shows that inferred data are often treated as trade secrets and not shared with data subjects, even when they invoke their right of access.[171]

---

[164] Finck (n 161) 9.
[165] Häuselmann (n 44) 151; Andreas Häuselmann, 'Fit for purpose? Affective Computing meets EU data protection law' (2021) Vol. 11 No. 3 *International Data Privacy Law* 245, 249.
[166] Bart Custers and Anne-Sophie Heijne, 'The Right of Access in Automated Decision-Making: The Scope of Article 15(1)(h) GDPR in theory and practice' (2022) Vol. 46 *Computer Law and Security Review* <https://doi.org/10.1016/j.clsr.2022.105727> accessed 12 November 2024.
[167] Only Case C-169/23 *Másdi* touches upon Article 14 GDPR, see following paragraph.
[168] Häuselmann (n 44) 111.
[169] See Article 13(1) GDPR and regulatory guidance which confirms that controllers must not provide individuals about the categories of personal data processed. See Art 29 Working Party, 'Guidelines on transparency under Regulation 2016/679' (WP 260 rev.01, 11 April 2018) 36.
[170] Art 29 Working Party (n 169) fn 30, emphasis added by the authors.
[171] Custers and Heijne (n 166).

Arguably, inferred data constitute 'new' personal data not collected from the data subject, triggering the transparency obligations contained in Article 14 GDPR.[172] The CJEU's ruling in *Másdi* supports this view. According to the CJEU, the dichotomy between Article 13 and 14 GDPR is straightforward. All situations in which data are not collected from the data subject fall under the scope of Article 14 GDPR.[173] This includes personal data generated by the controller.[174] However, in *Másdi* the controller generated personal data based on data obtained from third parties, so it makes sense that Article 14 GDPR applies. Whether Article 14 GDPR should apply to personal data derived from data collected from the data subject[175] seems highly questionable.[176] In any case, data subjects are best advised to enforce their right of access. The CJEU has clarified that the scope of a copy under Article 15(3) GDPR includes personal data *generated* by the *controller*[177] and thus inferred personal data.

7. Potential Solutions

The issues surrounding the rectification of inferred personal data generated by AI have not remained unnoticed.[178] Wachter and Mittelstadt claim that inferences increasingly determine how data subjects are viewed and evaluated and that the GDPR attributes only limited rights over inferences to data subjects.[179] Therefore, they suggest closing this gap by proposing the 'right to reasonable inferences'. This right should apply to 'high-risk' inferences that cause damage to privacy or reputation, or have low verifiability in the sense of being predictive or opinion-based while being used for 'important decisions'. The proposed right is an important contribution to the field and contains several valid points and suggestions. However, it is unlikely that it solves the verifiability problem. This right would oblige data controllers to establish whether an inference is reasonable by disclosing, among other information, whether

[172] As discussed in Custers and Vrabec (n 69) Computer Law & Security Review 2, 7.
[173] Case C-169/23 *Másdi* [2024] ECR I-988 para 48.
[174] Ibid para 49.
[175] Governed by Article 13 GDPR.
[176] Particularly from a systematic interpretation. Generating inferred personal data constitutes 'further processing' mentioned in Articles 13(3) and 14(4) GDPR. Article 13(3) GDPR would be obsolete if Article 14 GDPR applies in this scenario. It would be odd if the AI system that generated the inferred data constitutes a 'third-party' source mentioned in Article 14(2)(f) GDPR and Recital 61 GDPR where the 'initial' personal data have been directly collected from the data subject. Also, this would suggest that the controller 'self-obtains' the inferred data from its own processing system while no other third party is involved in the processing (i.e. generating the inferred personal data).
[177] Case C-487/21, *F.F.* [2022] ECR I-1000 para 21; see also the opinion of AG Pitruzella paras 45 amd 70.
[178] We focus on solutions within data protection law. Other areas of law offer complementary protection and prevent that public and private actors treat unreliable predictions about people as facts, e.g., consumer law protecting economic interests of consumer, or product safety law and particularly the AI Act.
[179] Wachter and Mittelstadt (n 2) 611 and 613.

the data and methods used to draw the inferences are accurate and statistically reliable.

Data controllers may easily claim that the methods used to draw the inferences are accurate and statistically reliable. If not, data controllers would incriminate themselves and indicate non-compliance with the accuracy principle, which could lead to regulatory and private enforcement. Consequently, data subjects may, in practice, not receive information that empowers them to effectively enforce their right to rectification concerning inferred personal data generated by AI. It will arguably become even more difficult for data subjects to enforce this right because data controllers, when confronted with a rectification request, can claim that the methods used to draw the inferences are accurate and statistically reliable, and refer to the information already disclosed in the context of the right to reasonable inferences. Also, the suggested right contains several ambiguous terms, such as 'high-risk' inferences causing 'damage to privacy or reputation' and 'important decisions'.

Ausloos, Veale and Mahieu propose another solution for the verifiability standard problem. They suggest construing the right to rectification as an addendum rather than a data replacement. In contentious cases, neither the data subject nor the controller should act as 'the arbiter of truth'. Instead, when the data controller has 'good reasons' to disagree with the data subject concerning a requested rectification, both views co-exist in the data processing system. The data controller must consider both the suggested rectification and the original data.[180] In fact, the data subject already has a right to provide a 'supplementary statement' as enshrined in the second sentence of Article 16 GDPR. However, this 'co-existence' may not change much, since it is unclear what specific obligations such a supplementary statement imposes[181] on the data controller.[182] Thus, the right to complete incomplete personal data does not prove to be particularly helpful in the context of inferred personal data because it does not solve the problem of inaccuracy. Furthermore, the proposed solution does not effectively protect the data subject.

Another solution we have in mind is perhaps more straightforward.[183] In essence, we suggest reversing the burden of proof regarding rectifying inferred personal data. This could be done by adding an additional paragraph in Article 16 GDPR that broadens the right to rectification regarding the processing of inferred personal data generated by automated means. It empowers data subjects to contest the accuracy of such personal data easily. When data subjects do so, the data controller shall either cease processing or rectify the personal data as requested by the data subject unless it can demonstrate that its own interests in processing the personal data in the form contested by the data subject prevail. This is comparable to the balancing assessment

[180] Ausloos, Veale and Mahieu (n 127) 283, 302.
[181] Dimitrova (n 50) 27.
[182] Regulatory guidance simply states that Article 16 GDPR contains a right for the data subject to complement the personal data with additional information see Art 29 Working Party (n 77) 18.
[183] This solution builds on what has been suggested in literature by: Häuselmann (n 44) 332.

the controller needs to perform when a data subject objects to processing according to Article 21(1) GDPR. Thus, it is the data controller that bears the burden of proof. Further, other scholars suggest placing the burden of proof on the developers of the corresponding systems as well as the entities that rely on algorithmic predictions.[184] The reversal of the burden of proof makes the right to rectification more effective regarding inferred personal data generated by AI. It would also contribute to transparency requirements in the GDPR, as data controllers would have to show why inferred data are accurate.

8. Conclusions

Several issues arise when the right to rectification enshrined in Article 16 GDPR is applied to inferred personal data. This provision is intended to protect data subjects by offering them the right to have inaccurate personal data concerning them rectified and the right to have incomplete data completed. The underlying idea is that individuals will not be assessed or addressed based on inaccurate or incomplete data, which could easily lead to injustices and harm.

For factual data obtained from the data subject, such as an address or date of birth, this provision is relatively straightforward. However, in the data economy, many types of data are by nature less straightforward. In the case of inferred data, such as predictions, rectification rapidly becomes complicated. Some types of data are more nebulous, for instance, due to error margins or because they are not dichotomously right or wrong. Also, the accuracy of inferred personal data (such as emotion data) can be highly subjective.

To rectify inaccurate personal data, data subjects must provide objectively verifiable evidence. This is what we call the objective verifiability standard. Applying this standard to inferred personal data, such as unverifiable predictions or subjective emotion data, causes two conceptual and at least one practical problem.

The first conceptual problem is the verifiability problem. This problem arises for types of inferred personal data for which it may be challenging to assess what is accurate. The second conceptual problem is the objectivity problem, which arises in the context of inferred personal data that are highly subjective (e.g., emotion data). The (in)accuracy of subjective data cannot be proven objectively. As a consequence, data controllers may reject a rectification request. The most prominent practical problem is the awareness problem. This problem arises when the data subject is unaware that rectification is needed.

A potential solution to address these issues of the right to rectification could be to reverse the burden of proof in cases where data subjects contest the accuracy of

[184] Angelina Wang et al, 'Against Predictive Optimization: On the Legitimacy of Decision-Making Algorithms that Optimize Predictive Accuracy' (2023) 1, 4 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4238015> accessed 12 November 2024; Matsumi and Solove (n 129).

inferred personal data. This would balance the power and knowledge asymmetries between data subjects and controllers, and contribute to effective protection from risks associated with processing inferred personal data.