Transparency on Paper, Black Box in Practice

Voice Assistants, the GDPR and the AI Act

ELENA DELLA VALENTINA*

Abstract: This work examines the European Union's regulatory framework on transparency, focusing in particular on the dimension of explainability, with specific reference to voice assistants. It analyzes the main legal sources governing this aspect, namely the General Data Protection Regulation and the Artificial Intelligence Act, with the aim of clarifying the scope and content of transparency and explainability obligations under both regulations. It investigates how these requirements operate in practice by applying the legal framework to a specific category of AI-based consumer devices: voice assistants. In this context, Amazon Alexa is used as a case study to determine the applicable legal regime and to assess how transparency obligations are translated into real-world interactions between users and AI systems. The analysis then considers Alexa+, the latest version of the technology, whose enhanced functionalities and increased autonomy raise some regulatory questions. In particular, the work explores whether its features could lead to its classification as a high-risk AI system under the AI Act, triggering stricter obligations.

Keywords: Transparency; Explainability; Voice Assistants; GDPR; AI Act.

1. *Introduction*

"*Designed to protect your privacy*". This is how Amazon titles its informative page[1] on Alexa's privacy standards, Amazon's most popular voice assistant. The voice assistant, which works on a variety of devices – the best-known being part of Amazon's smart speakers' family, Echo – is quite representative of one of the most popular applications of Internet of Things (IoT): voice assistants.

Like many other voice assistants, Alexa can perform various functions: it can turn your TV on, help you write down your grocery list, read the news, and play music[2]. It can even learn to recognize your voice, in order to offer you a personalized experience[3].

To do this, Alexa collects lots of data: like all Artificial Intelligence (AI) applications, voice assistants need data, not only to work, but also to

---

* Elena Della Valentina is a second-year Master's student in Global Law Making at the University of Trento, focusing on AI and law, with particular interest in AI and fundamental rights. She holds an LL.B. in Comparative European and International Legal Studies from the University of Trento.

[1] See *Alexa Privacy Hub* (Amazon.com), available at https://www.amazon.com/Alexa-Privacy-Hub/b?ie=UTF8&node=19149155011 (last visited March 17, 2026).

[2] See *Features for the Original Alexa* (Amazon.com), available at https://www.amazon.com/b?ie=UTF8&node=21576558011 (last visited March 17, 2026).

[3] See *Alexa Profiles* (Amazon.com), available at https://www.amazon.com/alexa-profiles/b?ie=UTF8&node=23804562011&ref=pe_alxhub_aucc_en_us_IC_HP_16_HUB_PROF (last visited March 17, 2026).

improve. And with a great collection of information (especially with personal data, such as voice recordings), come great privacy concerns. Alexa has been on the market for around a decade, yet it is often at the center of debates, news reports, and even court cases[4] concerning privacy issues. Amazon itself had to set up a page called "Separating fact from fiction when it comes to Alexa privacy"[5]. The intent was to reassure users that statements such as "Alexa is listening to your personal conversations" are false, whilst others – "I can see and delete everything Alexa heard me say[6]"– are indeed true.

As with many IoT devices, privacy concerns surrounding voice assistants arise from their characterizing lack of transparency. As the *Transparency in the consumer Internet of Things*[7] report underlines, "there is limited visibility over the nature of data processing that occurs in the consumer IoT". This is due to the 'black-box phenomenon', which is caused

[4] See Natasha Singer, *Amazon to Pay $25 Million to Settle Children's Privacy Charges* (The New York Times, May 31, 2023), available at https://www.nytimes.com/2023/05/31/technology/amazon-25-million-childrens-privacy.html (last visited March 17, 2026).

[5] See *Separating Fact from Fiction When It Comes to Alexa Privacy* (Amazon.com), available at https://www.amazon.com/b/?node=23608567011 (last visited March 19, 2026).

[6] Reading this statement in the EU on March 20, 2025, is a bizarre experience, considering the unfortunate update that U.S. Alexa users received just a couple of days earlier. See Scharon Harding, *Everything You Say to Your Echo Will Soon Be Sent to Amazon, and You Can't Opt Out* (Wired, March 17, 2025), available at https://www.wired.com/story/everything-you-say-to-your-echo-will-be-sent-to-amazon-starting-march-28/ (last visited March 19, 2026).

[7] See Anna Ida Hudig, Chris Norval and Jatinder Singh, *Transparency in the Consumer Internet of Things: Data Rights and Data Flows* at 3 (Information Commissioner's Office 2023), available at https://iot-transparency.org/Transparency%20in%20the%20consumer%20Internet%20of%20Things_files/iot-transparency.pdf (last visited March 19, 2026).

both by the intrinsic nature of AI itself, and conversely, by the unwillingness of companies to disclose details on their data processing, for reasons ranging from protecting reputation to preserving trade secrecy[8].

Against this background, this paper will focus on the relationship between the European Union legal framework for transparency and voice assistants, especially in the light of the General Data Protection Regulation[9] (GDPR) and of the EU's recent Artificial Intelligence Act[10]. The pages that follow will try to provide an answer to the following questions:

i. What is the current transparency framework for voice assistants in the European Union?
ii. How does the AI Act influence the explainability dimension of transparency?
iii. What challenges are associated with transparency in voice assistants, and how does the AI Act address these challenges?

In order to answer these questions, together with the general analysis of the legal framework (Section 2), this work presents a case study regarding Amazon's Alexa (Section 3). Alexa was chosen as a reference system because of its popularity, the amount of literature addressing it[11], the debates it has sparked in the last decade, and the characteristics of the recently announced Alexa+, which will be at the center of the last paragraph of this paper (Section 4).

---

[8] This paper will not touch upon the theme. However, for an analysis, see Guido Noto La Diega, *Internet of Things and the Law: Legal Strategies for Consumer-Centric Smart Technologies* at 264–69 (Routledge 2023).

[9] See EU reg. 27 April 2016, no. 679 (General Data Protection Regulation, GDPR); EU dir. 24 October 1995, no. 46.

[10] See 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).

[11] See also Noto La Diega, *Internet of Things and the Law* (cited in note 8).

2. *The EU Framework for Transparency*

2.1. *Transparency Under the GDPR*

Provided that the territorial scope requirement[12] is respected, the GDPR is particularly fit for data collected and processed by voice assistants[13]. Indeed, together with personal data provided by the user through, for example, the setting up of an account, voice assistants collect and process voice recordings, which not only are personal data *per se*[14], but may also contain personal data.

Transparency is one of the key principles of the EU's GDPR, applied mostly in providing data subjects with information to ensure fair processing of their personal data, in the way data controllers communicate with individuals regarding their rights under the GDPR, and in the

[12] In order for the GDPR to be applicable to a specific processing of personal data, it must fall under the cases under art. 3 GDPR, which describes the territorial scope of application of the regulation. Indeed, the GDPR applies not only to establishments in the EU, but also has an extraterritorial scope, being applicable to data processing operations conducted by non-EU establishments when it is related to the offer of goods or services or the monitoring of the behaviour of data subjects in the EU. It also applies extraterritorially where Member State law applies by virtue of public international law.

[13] In addition to the GDPR, certain aspects of voice assistants may also fall within the scope of the e-Privacy Directive, as voice assistants are software services operating through physical devices which constitute "terminal equipment" under the e-Privacy framework. In particular, according to art. 5(3), consent is required to store or gain access to information not necessary to execute users' request, while processing users' data to execute users' requests falls under one of the exceptions. For a more complete analysis, see European Data Protection Board, *Guidelines 02/2021 on Virtual Voice Assistants*, Version 2.0, available at https://edpb.europa.eu/our-work-tools/our-documents/guidelines/guidelines-022021-virtual-voice-assistants_en (last visited May 9, 2026). However, since this work focuses primarily on transparency and explainability obligations, the analysis will mainly concentrate on the GDPR and the AI Act.

[14] See Eoghan Furey and Juanita Blue, *Alexa, Emotions, Privacy and GDPR* (Proceedings of the 32nd International BCS Human Computer Interaction Conference, 2018).

methods data controllers use to enable individuals to exercise those rights[15].

While the regulation does not define transparency explicitly, as underlined by the Article 29 Working Party *Guidelines on transparency under Regulation 2016/679* [16], Recital 39 of the GDPR can be of help in clarifying the meaning and effect of the principle in the context of data processing. It states that "It should be transparent to natural persons that personal data concerning them are collected, used, consulted or otherwise processed and to what extent the personal data are or will be processed", further characterizing the principle of transparency by requiring that information and communications be "easily accessible and easy to understand" and in "clear and plain language". Notably, Recital 39 also includes in the principle of informing data subjects about the identity of the data controller, the purposes of processing, and other relevant details to ensure fair and transparent treatment of their data, as well as their right to access their personal information.

Under the GDPR, the principle of transparency finds its best expression in Article 12, which mandates the requirements for "transparent information, communication and modalities for the exercise of the rights of the data subject" [17]. The Article must be read in conjunction with Articles 13 and 14 as regards the provision of information related to the collection of personal data and with Articles 15 to 22 and 34 for communications to the data subject on the processing of their personal data, which must respect the criteria set out in article 12.

Requirements set out in Article 12 mandate, among other things, that information and/or communications be concise, transparent, intelligible and easily accessible (Article 12.1) and that clear and plain language be used, especially when providing information to children (Article 12.1).

Articles 13 lists "information to be provided where personal data are collected from the data subject", while Article 14 regards "information

[15] See Article 29 Working Party, *Guidelines on Transparency Under Regulation 2016/679* (2018).

[16] See id.

[17] See art. 12, EU reg. 27 April 2016, no. 679 (GDPR).

to be provided where personal data have not been obtained from the data subject". Article 15 also concerns transparency, as it provides for "right of access by the data subject" to their data. Transparency is also mentioned in Article 34, on the "communication of a personal data breach to the data subject", and in Article 22 on "automated individual decision-making, including profiling", which will be extensively analyzed in paragraph 2.3.a.

2.2. *Explainability*

"Explainability"[18] of AI models is often characterized in the legal domain as a "right to an explanation"[19] with reference to algorithmic decision making. As noted by Wachter, Mittelstadt and Floridi[20], this "right to an explanation" can encompass different dimensions. Indeed, the "explanation" may refer both to explanations of the system's *overall functionality* (such as logic, criteria, and models used) as well as details regarding *specific decisions* (such as the rationale and individual factors considered)[21]. A further distinction must be made on the basis of when explanations are given: either before (*ex ante*), focusing only on overall functionality since

[18] Explainable AI is one of the most current topics in AI research. For a technical overview, see Rudresh Dwivedi, et al., *Explainable AI (XAI): Core Ideas, Techniques, and Solutions*, 55 ACM Computing Surveys 194:1 (2023); Andreas Holzinger, et al., *Explainable AI Methods—A Brief Overview*, in *XXAI—Beyond Explainable AI* 13 (Andreas Holzinger et al. eds., 2022); Plamen P. Angelov, Eduardo A. Soares, Richard Jiang, Nicholas I. Arnold and Peter M. Atkinson, *Explainable Artificial Intelligence: An Analytical Review*, 11(5) Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery (2021).

[19] In this work, "explainability" and "right to an explanation" are used with the sole distinction of the connotation of "right to an explanation" as, indeed, a right, while "explainability" simply refers to "the concept that a machine learning model and its output can be explained in a way that 'makes sense' to a human being at an acceptable level." See *Explainable AI* (C3 AI), available at https://c3.ai/glossary/machine-learning/explainability.

[20] See Sandra Wachter, Brent Mittelstadt and Luciano Floridi, *Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation*, 7 International Data Privacy Law 76, 82 (2017).

[21] See id.

the specific decision rationale isn't yet known, or after (*ex post*) a decision is made, potentially covering both the overall functionality and the rationale for the specific decision[22].

In the EU context, the right to an explanation is encompassed, at least to a certain extent, both in the AI Act and in the GDPR.

2.3. *The Right to an Explanation and the GDPR*

The existence (and the exact extent) of a right to explanation in the GDPR has long been a matter of debate. Literature has been divided, with scholars affirming (or refuting) the existence of the right to an explanation in the GDPR within two main legal bases: Articles 13, 14 and 15 GDPR (information to be provided for the collection of personal data and the right of access) and Article 22 (on automated individual decision-making, including profiling)[23].

While for the purposes of this paper we will offer a brief overview of both points of view in order to underline the unclear nature of the right to explanation in the GDPR, it must be highlighted that judgments such as the Dutch cases of *Uber and Ola*[24] seem to point towards the existence of a right to explanation in the GDPR[25].

---

[22] See id.

[23] This paper follows the division made in Ljubiša Metikoš and Jef Ausloos, *The Right to an Explanation in Practice: Insights from Case Law for the GDPR and the AI Act*, 17 Law, Innovation and Technology (2025), distinguishing between two main methodologies.

[24] See Gerechtshof Amsterdam, April 4, 2023, no. 200.295.742/01.

[25] See generally Raphaël Gellert, Marvin van Bekkum and Frederik Zuiderveen Borgesius, *The Ola and Uber Judgments: For the First Time a Court Recognises a GDPR Right to an Explanation for Algorithmic Decision-Making* (EU Law Analysis, 2021), available at https://eulawanalysis.blogspot.com/2021/04/the-ola-uber-judgments-for-first-time.html (last visited March 18, 2026); Damian Clifford, Jake Goldenfein, Aitor Jimenez and Megan Richardson, *A Right of Social Dialogue on Automated Decision-Making: From Workers' Right to Autonomous Right*, 2023 Technology and Regulation 1; Metikoš and Ausloos, *The Right to an Explanation in Practice* (cited in note 23).

2.3.1. *Article 22 GDPR*

In particular, positions affirming the existence of a right to explanation under Article 22 GDPR are based upon a series of different readings of the Article.

Article 22 establishes a data subject's "*right not to be subject* to a *decision based solely on automated processing, including profiling*, which produces *legal effects concerning him or her* or *similarly significantly affects him or her*"[26] (paragraph 1), with the three exceptions (paragraph 2) of necessity for contract fulfillment between the data subject and controller (letter a), authorization under EU or Member State law with safeguards for the data subject (letter b) and explicit consent of the data subject (letter c). Additionally, in situations where automated decision-making is necessary for contract fulfillment or based on explicit consent, the data controller must put in place measures to protect the rights, freedoms, and legitimate interests of the data subject including at least the right to human intervention by the controller, the ability for the data subject to express their viewpoint, and the option to challenge the decision (paragraph 3). Finally, a ban (with exceptions) is established for decisions based on what the GDPR defines as "special categories of personal data" (paragraph 4).

The explicit mention of the right to explanation is, on the other hand, present in Recital 71 of the GDPR which enlists "the right […] to obtain an explanation of the decision reached" among the "suitable safeguards" to which automated individual decision-making should be subject.

Arguments favoring the existence of a right to explanation under Article 22 can rely on the European Commission's recognition of the crucial role that recitals play in interpreting EU act provisions[27]. In this view,

[26] See art. 22, para. 1, EU reg. 27 April 2016, no. 679 (emphasis added).

[27] See Roberto Baratta, *Complexity of EU Law in the Domestic Implementing Process*, 19th Quality of Legislation Seminar, *EU Legislative Drafting: Views from Those Applying EU Law in the Member States* (2014), cited in Guido Noto La Diega, *Against the Dehumanisation of Decision-Making*, 9 Journal of Intellectual Property, Information Technology and Electronic Commerce Law 3 (2018).

the mention of the right to an explanation in a recital can be leveraged to interpret Article 22 in a way that reflects the broader objectives of the GDPR, specifically to enhance the protection of data subjects' rights, thus pushing for an interpretation that includes the right to an explanation as a part of Article 22[28].

However, the contrary can be argued, also considering that the explicit right to explanation present in some of the drafts of the GDPR was ultimately moved to Recital 71[29]. Indeed, Wachter Mittelstadt, and Floridi are of the opinion that "the omission of a right to explanation from Article 22 appears to be intentional"[30], considering both the historical evolution of the provision in drafts and negotiations and that "the safeguards specified in recital 71 are almost identical to those in Article 22(3)" except for the inclusion of a right to explanation.

Other approaches[31], either argue that recitals must guide the interpretation of Article 22 GDPR in conjunction with Articles 13, 14, and 15[32]; or connect the right to explanation to an effective exercise of the right to contest automated decisions contained in Article 22(3)[33]; or link Article 22

[28] See Noto La Diega, *Internet of Things and the Law* at 23 (cited in note 8).

[29] See Wachter, Mittelstadt and Floridi, *Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation* at 81–82 (cited in note 20).

[30] See id.

[31] The approaches that follow can be found in the literature review in Metikoš and Ausloos, *The Right to an Explanation in Practice* at 7 (cited in note 23).

[32] See Gianclaudio Malgieri and Giovanni Comandé, *Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation*, 7 International Data Privacy Law 243 (2017).

[33] See Metikoš and Ausloos, *The Right to an Explanation in Practice* at 7 (cited in note 23), referring to Ljubiša Metikoš, *Explaining and Contesting Judicial Profiling Systems*, Technology and Regulation 188 (2024); Marco Almada, *Human Intervention in Automated Decision-Making: Toward the Construction of Contestable Systems*, in *Proceedings of the Seventeenth International Conference on Artificial Intelligence and Law* (2019); Emre Bayamlıoğlu, *The Right to Contest Automated Decisions Under the General Data Protection Regulation: Beyond the So-Called "Right to Explanation"*, 16 Regulation and Governance 1058 (2021); Claudio Sarra, *Put Dialectics into the Machine: Protection Against Automatic Decision-Making Through a Deeper Understanding of Contestability by Design*, 20 Global

with fundamental rights such as the right to an effective remedy and a fair trial, underscoring the importance of AI transparency in protecting these rights[34].

2.3.2. *Articles 13, 14 and 15 GDPR*

Arguments interpreting Articles 13, 14 and 15 GDPR as establishing a right to explanation rely on the fact that the three articles mention the obligation to provide "meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject" in the case of "automated decision-making, including profiling, referred to in Article 22(1) and (4)". This formulation is regarded by many authors as a right to an explanation[35], but this is not uncontested. Notably, Wachter, Mittelstadt, and Floridi, despite not denying completely the validity of a right to explainability as per these articles, underline that they seem to require only an *ex-ante* explanation of the functioning of the system[36].

It is important to underline that, in both cases we have analyzed, the right to an explanation as outlined by the GDPR is considerably limited. Indeed, even if we acknowledge its existence, this right is only applicable if the specific case falls under the provisions of Article 22 of the GDPR, which represents a particularly difficult test requiring that a "decision" is "based solely on automated processing" and that it produces "legal effects" or similarly affects the data subject.

Jurist 1 (2020). This is also echoed in Article 29 Working Party, *Guidelines on Transparency Under Regulation 2016/679* (cited in note 15).

[34] See C-817/19, *Ligue des droits humains ASBL v Conseil des ministres*, ECLI:EU:C:2022:491, para. 194.

[35] Among others, see Andrew D. Selbst and Julia Powles, *Meaningful Information and the Right to Explanation*, 7 International Data Privacy Law 233 (2017); Bryce Goodman and Seth Flaxman, *European Union Regulations on Algorithmic Decision-Making and a "Right to Explanation"*, 38 AI Magazine 50 (2017).

[36] See Wachter, Mittelstadt and Floridi, *Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation* at 82–84 (cited in note 20).

2.4. *Transparency, Explainability and the AI Act*

Transparency and explainability are fundamentally important within the framework of the EU's AI Act, which integrates these concepts in various ways through its risk-based approach.

For systems classified as "high-risk"[37], transparency and explainability are present in various provisions, and framed by the AI Act as objectives to be achieved primarily by providing users with relevant documentation and instructions for using the AI system, rather than by mandating specific transparency-by-design models or compulsory use of explainable AI tools[38]. This is perfectly reflected in Article 11, which, together with Annex IV, mandates technical documentation for high-risk systems. It includes, among other information on the functioning of the system, "the design specifications of the system, namely the *general logic* of the AI system and of the algorithms; the key design choices including the rationale and assumptions made, including with regard to persons or groups of persons in respect of whom, the system is intended to be used"[39]. As Panigutti and others notice, it is likely that the term "general logic" refers to the core principles guiding a system's decision-making, with technical documentation unlikely to explain specific outcomes or decisions[40], and thus being classified, in the terms of Wachter, Mittelstadt, and Floridi[41], as an *ex-ante*, *overall functionality* kind of explanation. Article 13 explicitly addresses transparency, by stating that "high-risk AI systems

[37] For the purposes of this paper, we will not dive deeply into the classification rules in the AI Act. The chapter of reference for high-risk systems in the AI Act is Chapter III.

[38] See Cecilia Panigutti, et al., *The Role of Explainable AI in the Context of the AI Act*, in *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency* 1144 (June 12–15, 2023).

[39] See Annex IV(2)(b), AI Act (emphasis added).

[40] See Panigutti, et al., *The Role of Explainable AI in the Context of the AI Act* at 1144 (cited in note 38).

[41] As discussed in paragraph 2.1, with reference to Wachter, Mittelstadt and Floridi, *Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation* (cited in note 20).

shall be designed and developed in such a way as to ensure that their operation is sufficiently transparent to enable deployers to interpret a system's output and use it appropriately" and mandating the content of instructions for the use of high-risk systems to be provided. Also, Article 14 establishes that "high-risk AI systems shall be designed and developed in such a way, including with appropriate human-machine interface tools, that they can be effectively overseen by natural persons during the period in which they are in use". It also ensures that operators for human oversight are in a position fit to exercise effective oversight through the "checklist" set forth in paragraph 4. These elements somehow imply a certain level of transparency (e.g. the human oversight operator must be able to "properly understand the relevant capacities and limitations of the high-risk AI system[…]", as per paragraph 4, letter a)[42].

As regards the ex-post, case-specific right to an explanation, the AI Act seems to be less ambiguous than the GDPR[43], establishing, in Article 86, that an individual affected by decisions taken "on the basis of the output from a high-risk AI system […], which produces legal effects or similarly significantly affects that person in a way that they consider to have an adverse impact on their health, safety or fundamental rights shall have the right to obtain from the deployer clear and meaningful explanations of the role of the AI system in the decision-making procedure and the main elements of the decision taken", with a possibility of exceptions established by "Union or national law in compliance with Union law".

Transparency requirements for the so-called "limited risk" systems are contained in Article 50, and mostly address the need to make users

[42] See Panigutti, et al., *The Role of Explainable AI in the Context of the AI Act* at 1144 (cited in note 38).

[43] For a more complete comparison between art. 86 AIA and the GDPR framework, see, among others, Krystyna Nizioł, *The Right to Obtain an Explanation of the Decision-Making Procedure in an Individual Case Under Article 86(1) of the Artificial Intelligence Act—Selected Problems*, 4 Studia Prawnicze KUL 77 (2025). For an analysis of the interplay between arts. 15(1)(h) and 22 GDPR and art. 86 AIA, see Blendi Sheremeti, *Can the AI Act Fill the GDPR's Black Box? A Comparative Analysis of Articles 15(1)(h) and 22 of the GDPR and Article 86 of the AI Act on the Right of Access in AI Decision-Making* (Sept. 12, 2025) (SSRN), available at https://ssrn.com/abstract=6697239.

aware that they are interacting with an AI system, or viewing AI-generated content.[44] Providers must indeed design systems that interact directly with natural persons in a way in which concerned users are informed of interacting with an AI system, except when it is apparent or if the AI is employed for certain legal purposes, such as crime detection. Additionally, AI systems (including general-purpose AI systems) that produce synthetic content, such as deepfakes, must label their outputs as artificially generated. Furthermore, deployers must disclose when AI is used for emotion recognition or biometric categorization, except when used for certain legal purposes, such as crime detection. Additional rules and exceptions are provided for deep fakes or informative text.

Finally, for general purpose AI systems[45], in addition to the requirements that are applicable, when appropriate, as per Article 50, the AI Act establishes some transparency requirements, mostly in terms of technical documentation, as established by Article 53 and Annex XI, without, however mentioning the "general logic" mentioned in Annex IV, and thus being clearly less inclined towards explainability.

3. *Transparency in Alexa: A Case Study*

After this legal overview on the transparency framework that can potentially be applicable to IoT devices, including smart speakers featuring voice assistants, it is now time to move to our case study: Amazon's Alexa.

The analysis that follows will consider first a more general dimension of transparency, then the more specific dimension of explainability, starting from the GDPR framework. After this, the potential impact of the AI Act in this framework will be addressed. These first two points exclude

[44] For a complete analysis of art. 50, see Thomas Gils, *A Detailed Analysis of Article 50 of the EU's Artificial Intelligence Act*, in *The EU Artificial Intelligence (AI) Act: A Commentary* 776 (C.N. Pehlivan, N. Forgó and P. Valcke eds., 2025).

[45] For the purposes of this paper, we will not dive deeply into the classification rules in the AI Act. The chapter of reference for general-purpose AI systems in the AI Act is Chapter V.

from their analysis Alexa+, which will be at the center of a final paragraph addressing speculations on the most recent models and the implications for the transparency legal framework.

3.1. *The GDPR Framework*

Transparency concerns in Amazon's Alexa are far from being something recent[46], yet, after a decade they are far from being solved. A large portion of these issues finds their origin in what Noto La Diega defines as the "contractual quagmire" of the legal documentation referring to the system. The analysis conducted by the author[47] on Amazon Alexa's legal documentation and especially on its privacy notice underlines the misalignment with the principle of transparency and, more in general, with provisions contained in Articles 12, 13 and 14 of the GDPR. Indeed, Amazon's privacy notice is quite generic, disclosing only the "types" of information gathered (distinguishing between data provided by users, automatic information, and data from unspecified other sources), providing "examples" for each category[48]. Additionally, it does not feature an exhaustive list of purposes for data collection and processing, but simply examples, often formulated in a quite generic manner.

---

[46] Among the many newspaper articles on the matter, see Dorian Lynskey, *Alexa, Are You Invading My Privacy? – The Dark Side of Our Voice Assistants* (The Guardian, October 9, 2019), available at https://www.theguardian.com/technology/2019/oct/09/alexa-are-you-invading-my-privacy-the-dark-side-of-our-voice-assistants (last visited March 26, 2026).

[47] The analysis is contained in Noto La Diega, *Internet of Things and the Law* (cited in note 8).

[48] As Noto La Diega's analysis was performed on UK legal documentation two years ago, it was necessary to compare it to an EU country. For this paper, the Italian version was used: *Informativa sulla Privacy* (Amazon.it), available at https://www.amazon.it/gp/help/customer/display.html?nodeId=GX7NJQ4ZB8MHFRNJ (last visited October 18, 2025). As regards the aspects taken into account here, Noto La Diega's considerations remain valid.

There is a specific webpage[49] especially dedicated to "Alexa, Echo Devices, and Your Privacy", where aspects such as personalization (which seems to imply profiling, also considering the privacy policy[50]) and the use of recordings for training of Amazon's models are briefly explained. However, the overall documentation seems quite generic as to the exact extent of the collection and processing of data collected by Alexa.

Noto La Diega's "experiment" on the right to access was not any luckier from the point of view of transparency. After the submission of a subject access request, the right to access was granted by Amazon only to a limited part of his personal data, presented in the form of "hundreds of obscure spreadsheets, without any explanation and in a format that is hard to decipher"[51], and, notably, lacking the data subject's "digital twin" profile, construed through inferences made on the user's data in order to predict their behavior[52].

With regard to the dimension of explainability, all the hypotheses we have discussed in paragraph 2.3 require the conditions envisioned in Article 22(1) or 22(4) in order to try to argue for a right to an explanation (both *ex ante* and *ex post*; both on the overall functioning and on the specific decision), Therefore, using this methodology to attempt to look inside Alexa's black box is unlikely to produce any results. This is true especially considering that, in order to trigger Article 22, Alexa's "decisions" would not only need to be "based solely on automated processing", but also produce "legal effects" on the data subject or "similarly significantly affecting him or her" – a scenario which seems quite far from the actual capabilities that Alexa, even in the smart home integrations, has had up until now.

From this point of view, while GDPR's transparency rights that should be granted under Articles 12, 13, 14 and 15 seem to be not fully

[49] *Alexa, Dispositivi Echo e la tua privacy* (Amazon.it), available at https://www.amazon.it/gp/help/customer/display.html/?nodeId=GVP69FUJ48X9DK8V&ie=UTF8.

[50] Noto La Diega, *Internet of Things and the Law* at 262 (cited in note 8).

[51] See id. at 261.

[52] See id. at 262.

respected in practice, in the case of explainability and the right to an explanation, the GDPR seems not to be the right tool to open Alexa's black box.

3.2. *The AI Act Framework*

Trying to understand the risk level of Alexa under the AI Act is not an easy task. With scarce literature addressing the topic and the AI Act being currently only partially implemented, certainty about the classification of systems as complex as Alexa is something difficult to achieve.

Considering Article 6 paragraph 1 and 2, in conjunction with Annex III and Alexa's features (excluding Alexa+, that will be dealt with in the following paragraph), it would be quite difficult to classify Alexa as a high-risk system under the AI Act framework. Thus, we cannot rely on the transparency requirements for high-risk systems that we have described in paragraph 2.4.

Conversely, the classification of Alexa as one of the limited-risk AI systems subject to the transparency requirements of Article 50 is almost intuitive, as it constitutes a clear example of an AI system "intended to interact directly with natural persons"[53]. As discussed previously, this means that the system "designed and developed in such a way that the natural persons concerned are informed that they are interacting with an AI system, unless this is obvious from the point of view of a natural person who is reasonably well-informed, observant and circumspect, taking into account the circumstances and the context of use"[54]. One could also safely hypothesize that the case of Alexa may constitute an "obvious" instance of an AI system, however this is yet to be confirmed.

Finally, Alexa's language model also seems to fit under the umbrella of general-purpose AI models[55]. Article 3(63) defines a general-purpose AI model as "an AI model, including where such an AI model is

[53] See art. 50, para. 1, AI Act.

[54] See id.

[55] It is important to remind that the classification of a system as a general-purpose AI system and its classification in one of the AI Act's risk levels are not mutually exclusive.

trained with a large amount of data using self-supervision at scale, that displays significant generality and is capable of competently performing a wide range of distinct tasks regardless of the way the model is placed on the market and that can be integrated into a variety of downstream systems or applications, except AI models that are used for research, development or prototyping activities before they are placed on the market". Alexa seems to adhere to this definition, meaning that the technical documentation required by Article 53 and detailed in Annex XI could be required if this classification is correct. This would have an impact on the explainability dimension, as specified in paragraph 2.4, even if not as deep as a potential classification as a high-risk system.

4. *Future Challenges and Opportunities: Alexa+ and the Next Generation of Voice Assistants*

At the end of February 2025, Amazon officially announced Alexa+, a "next-generation assistant powered by generative AI" [56]. Alexa+ features improved, including more natural conversational capabilities, enhanced smart home features, more personalization and even some "agentic" capabilities[57].

While it is safe to say that the new features of Alexa+ would not be enough to trigger Article 22 GDPR, the situation might be different for the AI Act.

Indeed, while we are still lacking precise technical information, it seems that Alexa+ is able to "understand user's tone […] and adapt its response accordingly"[58]. If this was the case, while this would probably

[56] Panos Panay, *Introducing Alexa+, the Next Generation of Alexa* (Amazon News, February 26, 2025), available at https://www.aboutamazon.com/news/devices/new-alexa-generative-artificial-intelligence (last visited March 19, 2026).
[57] See id.
[58] Lisa Eadicicco, *Amazon's Alexa Is Getting a Major Upgrade for the AI Chatbot Era* (CNN, February 26, 2025), available at https://edition.cnn.com/2025/02/26/business/amazon-alexa-plus/index.html (last visited March 19, 2026).

not have an impact on Alexa's inability to trigger Article 22 GDPR[59], one can safely argue that this "understanding of the user's tone" can be classified as emotion recognition[60], which would place Alexa among high-risk systems in the AI Act. Indeed, Article 6(2) AI Act classifies as high-risk systems contained in Annex III, which includes, under point 1(c) "AI systems intended to be used for emotion recognition"[61]. Article 3(39) AI Act defines an emotion recognition system as "an AI system for the purpose of identifying or inferring emotions or intentions of natural persons on the

[59] See also Ayça Atabey, Burkhard Schafer and Lachlan Urquhart, *How Do You Solve a Problem Like Alexa?*, Jusletter IT 185, 187 (March 2023), available at https://www.research.ed.ac.uk/en/publications/how-do-you-solve-a-problem-like-alexa/ (last visited March 19, 2026).

[60] See Noto La Diega, *Internet of Things and the Law* (cited in note 8) (noting that Amazon was granted a patent in 2018 for "Indirect feedback systems and methods," thus having "a monopoly on a technology that allows the company to detect users' physical, emotional, and behavioral states"; however, emotion recognition seems to be systematically implemented only in Alexa+). See also Jenny Kennedy and Yolande Strengers, *Alexa's Got a Hunch: The Human Decisions Behind Programming Emotion-Sensing and Caregiving into Digital Assistants*, in *Everyday Automation* 92–93 (Routledge 2022) (mentioning a "frustration mode," where Alexa will apologise if she detects a user becoming frustrated with ineffective request responses, implemented in Alexa in 2019; this feature would only represent a limited instance of such a technology, difficult to qualify exactly as "emotion recognition"; moreover, in the absence of any official sources addressing this feature, it seems to resemble more of a rumor or a project under development rather than a fully implemented feature).

[61] While explaining the general rules for classification under the AI Act is outside the scope of this work, it might be useful to explore the possibility of the exceptions under art. 6(3), which excludes from the high-risk classifications those systems that do not pose a significant risk of harm to the health, safety or fundamental rights of natural persons. The fact that the system performs personalization already excludes the application of this exception, as specified in the last subparagraph of art. 6(3). Additionally, none of the conditions listed in art. 6(3)(a)–(d) seems to be appropriate for the case of Alexa+. Furthermore, the agentic capabilities of the system point towards a level of risk that is difficult to disregard.

basis of their biometric data"[62]. As recalled by the EDPB with reference to Article 4(14) GDPR[63], "voice data is inherently biometric personal data"[64], thus it seems probable that if Alexa+ features this kind of capability, it will be classified as a high-risk system. Of course, ultimately this classification depends on the actual features presented by Alexa+ and on their technical details.

From a transparency perspective, this would mean that the technical requirements of Article 11 and Annex IV would apply, along with the *ex ante* right to an explanation of the overall functioning of the system established by Articles 13 and 14. For what concerns the *ex post* specific right to an explanation established by Article 86, it would still be an improbable occurrence in the case of Alexa+. However, the enhanced "agency" capabilities of Alexa+, coupled with the fact that the formulation of Article 86 seems to suggest that it depends on each individual to determine whether they are significantly affected by a decision encompassed by Article 86, potentially offer more opportunities than Article 22 of the GDPR.

5. *Conclusions*

This work focused on the transparency legal framework applicable to voice assistant powered IoT devices, using Amazon's Alexa as a case study.

[62] While the AI Act does not provide an exact definition of emotion, recital 18 states: "The notion refers to emotions or intentions such as happiness, sadness, anger, surprise, disgust, embarrassment, excitement, shame, contempt, satisfaction and amusement."

[63] See also recital 14 of the AI Act ("The notion of 'biometric data' used in this Regulation should be interpreted in light of the notion of biometric data as defined in Article 4, point (14) of Regulation (EU) 2016/679, Article 3, point (18) of Regulation (EU) 2018/1725 and Article 3, point (13) of Directive (EU) 2016/680. Biometric data can allow for the authentication, identification or categorisation of natural persons and for the recognition of emotions of natural persons.").

[64] European Data Protection Board, *Guidelines 02/2021 on Virtual Voice Assistants*, Version 2.0 (cited in note 13).

In Section 2, we analyzed the current framework on transparency in the GDPR and the AI Act, with particular attention to "the debate on the right to an explanation" and to how explainability is integrated into the AI Act.

Section 3 focused on the Alexa case study, presenting transparency challenges that have been widely analyzed in literature, also in light of the applicable GDPR framework, and on the right to an explanation, concluding that the GDPR does not properly grant this right in the case of Alexa (paragraph 3.1). Paragraph 3.2 analyzed the impact of the AI Act on the transparency dimension, attempting to categorize Amazon's Alexa and concluding that it is likely to fit into the "limited risk systems" and "general purpose AI systems categories", and thus it is subject to their transparency requirements.

Finally, Section 4 focused on the new generation of voice assistants, using Alexa+ as an example, defining the applicable legal framework and hypothesizing its classification as an "high-risk system", based on the possibility it recognizes emotions and defining the implications of this classification on transparency.

What seems to be evident from this framework is that in the context of voice assistants, although transparency requirements seem to be clearly established by relevant regulations, they are, first of all, not necessarily fully respected in practice (even by tech giants such as Amazon), and secondly, able to establish only a limited framework for the right to an explanation. While the AI Act seems to show some improvements in respect to a right to an explanation, both *ex-ante* and *ex-post*, these improvements are mostly limited to high-risk systems. It follows that the degree to which we will be able to open voice assistants' black boxes in the future will largely depend on how they are classified according to the AI Act. The hope is that transparency will not merely remain on paper but will also allow us to look into voice assistants' black boxes in practice.