

Right to Explanation of AI Decisions

Elena Dubovitskaya, Gregor Bosold

Abstract *As artificial intelligence (AI) systems play an increasingly significant role in decision-making across domains like credit assessment, recruitment and healthcare, concerns about their transparency and accountability are becoming more pressing. This text examines the legal notion of a right to explanation. It analyzes the recognition of this right within the framework of European Union law, particularly under the General Data Protection Regulation (GDPR), the Consumer Credit Directive, and the AI Act, and explores whether this right may also be grounded in broader fundamental rights. The analysis addresses both technical and legal strategies for tackling the challenge of AI opacity, with a particular focus on the concept of local explanations, which clarify individual decisions rather than the overall system logic. Using the example of credit scoring, the text illustrates how personalized explanations can enhance transparency and empower those affected by automated decisions.*

1. Introduction

“Nothing in life is to be feared, it is only to be understood.” This famous saying, attributed to the French scientist Marie Curie, can easily be applied to the use of artificial intelligence in the 21st century. AI is now frequently used to make decisions on important and life-altering matters, such as determining credit eligibility or selecting candidates to job interviews. When a decision is negative, the affected person will often want to understand why. Only if the AI decision is adequately explained can the affected person truly comprehend it. The explanation of AI decisions is therefore an essential prerequisite for understanding their meaning. This aligns with the Capacity Approach, which emphasizes that individuals must not only have formal rights but also the actual ability to exercise them effectively.

From the legal perspective, a question arises as to whether explainability is merely a reasonable concern or a policy goal, or whether the affected individual has a legal right to know the reasons behind the decision. In practice, these reasons are often not disclosed, and legal scholarship has not paid enough attention to the right to explanation of algorithmic decisions. However, explainable and transparent AI

is essential for ensuring that people can trust this increasingly powerful technology while retaining control over it.

It is therefore not surprising that, at the insistence of the European Parliament, the right to explanation was incorporated into the AI Regulation (Article 86 of the AI Act). This right has a *residual* character, meaning it applies only when no equivalent right is provided under other Union law provisions (Article 86(3) of the AI Act). The primary applicable regulations include Articles 13 et seq. of the GDPR, which establish information rights and obligations in cases of automated decision-making under Article 22(1) of the GDPR. Additionally, Article 18(8)(a) of the Consumer Credit Directive contains a specific provision for consumer credit agreements. In Germany, the legislator plans to incorporate the right to explanation in the case of scoring into the Federal Data Protection Law.

What is the scope of these provisions, and how do they align with each other? What measures should be taken to achieve an effective right to an explanation in European law to enhance the capabilities of the affected individuals? Above all, however, it is questionable whether these provisions do establish entirely distinct and independent rights to an explanation, or whether they are specific manifestations of a general principle that may derive from fundamental rights. This article seeks to explore these questions in greater depth.

2. Capacity Approach, Right to Explanation, and Explainable AI

According to the previously developed capacity-based approach, a digital system can only be considered trustworthy if people, communities or organizations are able to sufficiently understand the functioning of digital systems. However, there are limits to this understanding of digital systems: It is conceivable that the technical functioning of the systems may not be reconstructed precisely in terms of model opacity. As explained above, we can only understand systems in terms of model opacity if we are able to identify the decisive factors that led to an actual approval or rejection. The term pragmatic opacity, on the other hand, describes the additional problem that the effects and side effects of digital systems cannot be understood under the conditions of use in the social world. Explanations of AI decisions benefit multiple stakeholders, but mainly those who rely on AI for decision-making (decision-makers) and those whose rights are affected by such decisions (decision subjects). Decision-makers include, for example, physicians using AI for prognosis or corporate executives relying on AI models for business decisions. Decision subjects include consumers who receive a credit score generated by an AI-based credit agency. The right to explanation is in particular a legal tool for decision subjects, allowing them to demand local explanations from the AI system operator to overcome the model opacity.

The right to explanation for AI decisions aligns well with the Capacity Approach because it transforms a formal right into an actual capability. AI decisions are not always comprehensible to those affected. The comprehensibility depends on the complexity of the model used. If the model is simply structured (a so-called white-box model, such as linear regression or a simple decision tree), its decisions are easily understandable for users. Large and powerful AI models, such as artificial neural networks or large, complex decision trees, are generally considered black boxes (Dubovitskaya and Buchholz 2023: 64–65). With increasing performance comes a loss of transparency. The output of such black-box models is not comprehensible for users, whether they are IT professionals, domain experts, or laypersons.

Without clear and understandable explanations, individuals affected by automated decisions may lack the knowledge or means to challenge unfair outcomes, thereby limiting their actual agency. Ensuring real access to justice and redress means that individuals can understand why a decision was made about them and act if necessary. Without this right, a person denied a loan due to an algorithmic credit score would have no real way to challenge or improve their situation. The right to explanation also promotes transparency and reduces power imbalances. Many algorithmic decisions are made by companies or institutions that hold significantly more knowledge and influence than the individuals affected. Providing explanations helps to reduce asymmetries of information and allows people to participate in decisions that impact their lives.

Another important aspect is supporting informed decision-making. If individuals understand the factors affecting their credit score or job application, they can make better choices about how to improve their situation. Additionally, protecting vulnerable groups is a crucial concern, as marginalized individuals, such as those with lower digital literacy or from disadvantaged socioeconomic backgrounds, are particularly at risk of being unfairly affected by algorithmic decisions. The right to an explanation ensures that such individuals are not left without recourse, thereby supporting a fairer distribution of capabilities.

However, the right to an explanation can only support affected individuals and enhance their capabilities if it is technically feasible. From a technical perspective, the problem of model opacity can generally be addressed in various ways. One approach to solving this problem is to seek global ex-ante explanations of the model's behaviour, which could make its decisions predictable. According to the current state of technology, this is largely infeasible, a point frequently emphasized in legal literature (Kumkar and Roth-Isigkeit 2020: 285). As a result, some propose significantly lowering the requirements for explainability, arguing that a rudimentary understanding of the basic functionality of an AI model suffices to establish “algorithmic control” (Weber, Kiefner and Jobst 2018: 1132). This perspective is found in both corporate and data protection law scholarship. Others counter that such

a superficial understanding is ineffective and advocate for the abandonment of explainability requirements altogether (Wischmeyer 2018: 53).

Nevertheless, the problem of model opacity can be addressed not only globally but also locally. The latter approach focuses on explanations that do not predict future model decisions but instead clarify decisions already made (local explanations). This approach does not require reconstructing the technical workings of the model but merely ensures that individual decisions can be understood. Various technical methods for explainable AI (XAI) are already available for this purpose. XAI tools are generally software solutions, specifically developed to provide explanations for the decisions of AI models. These tools consist of algorithms that can either be integrated into existing AI models or used separately to interpret model outputs. Most of these tools are available as libraries or frameworks that data scientists and AI developers can incorporate into their workflows. These methods include, for example, LIME (Local Interpretable Model-Agnostic Explanations, see Ribeiro, Singh and Guestrin 2016) and SHAP (SHapley Additive exPlanations, Lundberg and Lee 2017). A major advantage of both methods is that they are model-agnostic, as indicated by the term “Model-Agnostic” in LIME. This means they can be used to explain any AI model, regardless of its structure. Additionally, they are capable of handling various data formats, including image, text, and tabular data. While LIME is computationally faster and easier to implement for quick local explanations, SHAP offers a more theoretically justified, consistent, and robust approach to AI explainability. SHAP is mathematically well-founded: The method is based on concepts from game theory, particularly the Shapley values, which were developed in the 1950s by mathematician Lloyd Shapley. These values provide a fair method for distributing gains (or contributions) among participants in a cooperative game. In practical terms, SHAP works by calculating the contribution of each feature, such as gender or age, to a specific model prediction. This is done by measuring the difference in the prediction when a particular feature is omitted or included. A positive SHAP value indicates that the given feature increases the prediction, while a negative value suggests that it decreases it. The magnitude of the SHAP value reflects the strength of the feature’s influence. This approach determines how a particular feature affects a specific model prediction, a process known as feature attribution. SHAP also considers interactions between different features, revealing dependencies where the contribution of one feature depends on whether other features are present or not.

Nonetheless, it is imperative to acknowledge that this XAI method also has weaknesses. For complex data, the computation can take a long time because the number of possible feature combinations increases exponentially with the number of features in the dataset. This issue has been partially mitigated with methods like Kernel-SHAP, which approximates Shapley values and thus speeds up the computation. Moreover, while long computation times can be problematic for time-critical

real-time applications, they are less of a concern when explaining decisions that have already been made.

A more fundamental problem, however, is that SHAP, like nearly all currently available XAI methods, does not guarantee that the generated explanations accurately reflect the behaviour of the AI model being explained. One possible solution could be to combine different methods, such as SHAP with counterfactual explanations. Counterfactual explanations (see, for example, in the context of GDPR Wachter, Mittelstadt and Russell 2017: 841 et seq.) provide insights into AI decisions by answering “what if” questions. Instead of describing why a particular decision was made, they explain how the decision could have been different. A counterfactual explanation identifies the smallest possible change in input data that would have led to a different outcome. For example, in a loan approval scenario, if an applicant is denied a loan, a counterfactual explanation might state: “If your annual income had been € 5,000 higher, your loan would have been approved.”

If both methods produce consistent results, the likelihood that the explanations are accurate increases. Additionally, combining both methods can improve the quality of explanations: SHAP can provide an overview of feature importance, while counterfactual explanations offer actionable insights for individuals affected by AI decisions. For example, in credit scoring, SHAP could explain that income is the most important factor, while a counterfactual explanation could specify exactly how much income needs to increase for loan approval.

Nonetheless, it should be noted that even if multiple XAI methods produce consistent results, they may still deviate from the actual decision-making process of the AI model. However, this residual uncertainty seems acceptable, as there is no guarantee of absolute correctness in human explanations either. When a person justifies their decision, we cannot know to what extent their explanation truly reflects their internal decision-making process. Even if the explanation is given honestly, it may merely be an attempt to retrospectively rationalize, what was actually the result of intuitive and instinctive decision-making.

3. Right to an Explanation in European Secondary Law

After demonstrating that the right to an explanation can, in principle, be technically implemented, it is worthwhile to examine its specific manifestations in European secondary law, such as EU Regulations and Directives, more closely.

3.1. Right to Explanation under the GDPR

3.1.1. Protection of individuals affected by automated decisions

The GDPR provides a special right of access for data subjects in cases where decisions “based solely on automated processing” have legal effects on them or significantly affect them in a similar manner (Article 22(1) of the GDPR). This right is established in Article 15(1)(h) of the GDPR, which requires that affected individuals receive “meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject”. The corresponding obligations for data controllers are set out in Article 13(2)(f) of the GDPR (for data collected directly from the data subject) and Article 14(2)(g) of the GDPR (for data collected from third parties).

The scope of the right under Article 15(1)(h) of the GDPR has always been controversial. The main point of uncertainty was the interpretation of the term “the logic involved”. Some scholars took the expression literally and sought to highlight the technical challenges of obtaining meaningful information about the logic of certain machine learning algorithms. Consequently, they concluded that “the game is not worth the candle” and that “a ‘right to an explanation’ is probably not the remedy you are looking for” (Edwards and Veale 2017: 65). Others went even further and argued that a right to explanation of automated decision-making does not exist in the GDPR (Wachter, Mittelstadt and Floridi 2017: 90). In line with this, some data protection scholars in Germany opposed an extensive right of access, arguing that a general description of the data processing was sufficient (Kumkar and Roth-Isigkeit 2020: 277; Kamlah 2023: 28; Bienemann, 2022: 38). Others countered that data subjects have a right to an explanation and should also be informed of the weighting of different factors that led to the decision, the identity of reference groups, and the reasons for their categorization within such groups (Buchner 2024: 35a; Dix 2019: 25).

The European Court of Justice (CJEU) has addressed this issue in the context of credit scoring, which credit agencies perform on behalf of other companies, such as banks or mobile network providers. Credit agencies often use mathematical methods like logistic regression to predict the influence of certain variables, such as income, the number of loans, or changes of residence, on a person's creditworthiness. Strictly speaking, such methods do not fall under AI, nor do they involve machine learning algorithms. However, since the individual score is calculated automatically, this constitutes profiling within the meaning of Article 4(4) of the GDPR, which refers to a form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person, in particular to analyse or predict aspects concerning that person's creditworthiness.

In 2023, the CJEU ruled that the credit score calculated by Schufa Holding, Germany's largest credit agency, constitutes an automated decision under Article 22(1)

of the GDPR (CJEU 7.12.2023 – *SCHUFA Holding (Scoring)*, para 50). With this ruling, the CJEU rejected the prevailing view in Germany, which, in line with the wording of Article 22(1) of the GDPR, distinguishes between a “decision” on the one hand and the “profiling” on the other. According to this view, profiling, and thus credit scoring, is generally not considered a decision but merely a preceding data analysis process (profiling) that serves as a basis for the decision (BGH 28.1.2014, para. 34; Schulz 2022: 4, 17; Spindler and Horváth 2019: 12). Implicitly, the CJEU appears to have decided to distinguish between profiling (scoring) and its outcome (the score), interpreting the latter as a “decision”. According to the CJEU, this decision significantly affects the data subject within the meaning of Article 22(1) of the GDPR. It applies at least when the score is low, as companies that have commissioned SCHUFA to assess creditworthiness, in practice, almost always decide against entering into a contract with the data subject in such cases.

Following this decision, credit agencies are now required to provide explanations under Article 15(1)(h) of the GDPR (Dubovitskaya and Bosold 2024: 1809). In response to these extensive changes, SCHUFA has announced a transparency initiative aimed at revising its existing scoring practices (Schufa 2024).

Another significant ruling by the CJEU followed in February 2025 (CJEU, 27.2.2025 – *Dun & Bradstreet Austria*). The case was based on a dispute between an Austrian consumer and the credit reporting agency Dun & Bradstreet Austria. The consumer, whose creditworthiness was rated poorly by Dun & Bradstreet, claimed that her right of access under Article 15(1)(h) of the GDPR had been violated. In its decision, the CJEU provided a detailed interpretation of the scope of this right of access and addressed the question of how a balance should be struck between this right on the one hand and the interests of data processors, particularly the protection of their trade secrets, on the other.

The CJEU has decided that, in order to provide the “meaningful information about the logic involved”, the controller must explain to the data subject, in a concise, transparent, intelligible and easily accessible manner, the procedure and principles actually applied in the automated processing of his or her personal data in order to obtain a particular result – for example, a credit profile. The core element of the explanation under Article 15(1)(h) of the GDPR is therefore “the procedure and principles” applied in the automated data processing (CJEU 27.2.2025 – *Dun & Bradstreet Austria*, Ruling – Point 1). Advocate General de la Tour previously referred to “the essential elements of the method and the criteria used” (Advocate General de la Tour 12.9.2024 – *Dun & Bradstreet Austria*, para. 64).

The CJEU does not explicitly define what is meant by “the procedure and principles”. However, it establishes a clear link between Article 15(1)(h) of the GDPR and the rights of the data subject under Article 22(3) of the GDPR. The right of access under Article 15(1)(h) of the GDPR primarily serves to enable the data subject to effectively exercise their rights under Article 22(3) of the GDPR, namely the right to express

their own point of view and the right to challenge the decision (CJEU, 27.2.2025 – *Dun & Bradstreet Austria*, para. 55). In order to effectively exercise these rights, the data subject must be able to understand the reasons for the automated decision (CJEU, 27.2.2025 – *Dun & Bradstreet Austria*, para. 56). From this, the CJEU, for the first time in the history of the GDPR, derives a genuine right to an explanation: “[...] Article 15(1)(h) of the GDPR affords the data subject a genuine right to an explanation as to the functioning of the mechanism involved in automated decision-making of which that person was the subject and of the result of that decision” (CJEU, 27.2.2025 – *Dun & Bradstreet Austria*, para. 57).

In contrast, it is not necessary to provide the data subject with a complex mathematical formula (such as an algorithm) or to describe every step of an automated decision-making process in detail. Such information does not meet the requirements for explanations under Article 15(1)(h) of the GDPR, as it would not constitute a sufficiently precise and comprehensible clarification. Instead, the controller must describe the procedure and principles actually applied in a way that “the data subject can understand which of his or her personal data have been used in the automated decision-making at issue, with the complexity of the operations to be carried out in the context of automated decision-making not being capable of relieving the controller of the duty to provide an explanation” (CJEU, 27.2.2025 – *Dun & Bradstreet Austria*, para. 59, 61).

These findings of the Court provide valuable guidance for the right to an explanation and its technical implementation. First, the CJEU establishes that in the case of automated decisions, including AI decisions, there is a “genuine right to an explanation”. Second, this right does not relate to the general functioning of the AI model used, but rather to the specific decisions made by the model (local explanations). Third, the CJEU does not prescribe a specific XAI method for explanations, nor does it formulate concrete requirements that would indicate the necessity of a particular method, such as SHAP, for example. Instead, the Court limits itself to the requirement that the decision must be comprehensible to the data subject.

However, in its judgment, the Court states that in a case of profiling, as in this instance, it may be sufficiently transparent and intelligible to inform the data subject of the extent to which a variation in the personal data considered would have led to a different result (CJEU, 27.2.2025 – *Dun & Bradstreet Austria*, para. 62). This means, from a technical perspective, that counterfactual explanations may also be sufficient. In the Schufa case, however, Advocate General Pikamäe took the view that the controller must inform the data subject about “the factors taken into account for the decision-making process and their respective weight on an aggregate level” (Advocate General Pikamäe 16.3.2023 – *SCHUFA Holding (Scoring)*, para. 58). The mention of the “factors taken into account” (= features) and their weighting (= contribution to a specific model prediction) suggested that methods such as SHAP must be

used to explain ML decisions, as they reveal feature importance. The CJEU does not appear to require this.

The requirement that the explanation has to be provided in a concise, transparent, intelligible, and easily accessible form aligns well with the Capacity Approach, which emphasizes that individuals must not only have formal rights but also the real ability to exercise them effectively. Merely providing information is insufficient if people struggle to comprehend it. Therefore, the CJEU's decision bridges the gap between formal rights and real capabilities, making the right to an explanation practically accessible and meaningful.

3.1.2. Protection of trade secrets

The other major part of the CJEU judgement in *Dun & Bradstreet* case addresses the conflict between the right to an explanation and the protection of trade secrets. The CJEU emphasizes that the right of access must not infringe on the rights and freedoms of others, such as trade secrets or intellectual property rights, including copyright on software. On the other hand, these rights cannot be used as a justification to completely deny the data subject any access to information (CJEU, 27.2.2025 – *Dun & Bradstreet Austria*, para. 69, 70).

For this reason, the CJEU seeks to establish a fair balance. In doing so, it adopts the solution advanced by the referring Austrian court in its request for a preliminary ruling, as subsequently supported by Advocate General de la Tour in his Opinion. If the controller considers that the information required to be provided to the data subject under Article 15(1)(h) of the GDPR constitutes a trade secret, they must submit this information to the competent supervisory authority or court, which must then weigh the competing rights and interests to determine the extent of the data subject's right of access under Article 15 of the GDPR (CJEU, 27.2.2025 – *Dun & Bradstreet Austria*, Ruling – Point 2). Especially when the court or the competent authority concludes that the information is protected as a trade secret, a decision must be made based on the principle of proportionality regarding which information, and to what extent, must be disclosed to the data subject (see CJEU, 27.2.2025 – *Dun & Bradstreet Austria*, para. 73).

The recognition that fundamental rights must be balanced with conflicting rights on the other side is not new in the CJEU's case law (CJEU 29.1.2008 – *Promusicae*, para. 65 et seq.; CJEU 29.7.2019 – *Spiegel Online*, para. 51 et seq.; CJEU 4.5.2023 – *Austrian Data Protection Authority*, para. 44). It is only questionable whether the CJEU achieves a fair and practical balance in *Dun & Bradstreet*. The CJEU is correct in stating that balancing these conflicting rights may require limiting explanation obligations to the minimum information necessary to protect affected individuals' rights. However, since this does not require disclosing the actual algorithm or source code (see above), direct exposure of trade secrets is unlikely. Only indirect exposure through reverse engineering based on explanations could pose a risk.

Against this background, it is a positive development that the burden of proof for the risk of trade secret disclosure now lies with the controller, not the affected individual. The controller can no longer simply claim that explaining the automated decision would reveal trade secrets as a blanket defence. Instead, they must convince the court or the competent authority that the information in question is indeed protected. On the other hand, the court or the competent authority will often lack the technical expertise to thoroughly assess the controller's claims (Metikoš and Ausloos 2025: 26; Langenbucher and Bauer 2025). This is particularly true for the question of whether and when the reconstruction of the original algorithm through reverse engineering is possible. This will lead to the involvement of expert opinions, which, in turn, could result in delays and legal uncertainty.

3.2. Right to Explanation in the Consumer Credit Directive

The Consumer Credit Directive stipulates that the creditor must be required under national law to carefully assess the consumer's creditworthiness before concluding a credit agreement (Article 18(1) of the Consumer Credit Directive). The assessment of creditworthiness shall be carried out on the basis of relevant and accurate information regarding the consumer's income, expenses, and other financial and economic circumstances, which must be necessary and proportionate to the nature, duration, value, and risks of the credit for the consumer. The information shall be obtained from relevant internal or external sources (Article 18(3) of the Consumer Credit Directive).

Automated processing of personal data (internal or external scoring) is also permitted in this context. However, in such cases, the consumer has the right under Article 18(8)(a) of the Consumer Credit Directive to "request and obtain from the creditor a clear and comprehensible explanation of the assessment of creditworthiness, including the logic and risks involved in the automated processing of personal data, as well as its significance and effects on the decision." This provision constitutes a specific regulation for consumer credit agreements, which takes precedence over the right of access under Article 15(1)(h) of the GDPR (see also Buck-Heeb 2023: para. 41).

Since the wording of Article 18(8)(a) of the Consumer Credit Directive is similar to that used in Article 15(1)(h) of the GDPR, the scope of this right of access can be interpreted in light of the CJEU's decision in the *Dun & Bradstreet Austria* case (see above). This means that the consumer must be provided with an explanation of the automated decision regarding their creditworthiness in a way that allows them to understand it.

It should be noted that Article 18(8)(a) of the Consumer Credit Directive covers a broader range of automated decisions than the GDPR. The GDPR requires a decision to be based "solely" on automated processing (see Article 22(1) of the GDPR,

as referenced by Articles 13(2)(f), 14(2)(g), and 15(1)(h) of the GDPR). Recital 71 of the GDPR refers in this context to decisions made “without any human intervention”. However, decisions in which human involvement is merely formal are also included, such as cases where a person processes a decision but has no ability to deviate from an automatically generated outcome (Buchner 2024: 15). It should also be noted that, according to the Schufa judgement of the CJEU, the results of fully automated profiling could also constitute “decisions” within the meaning of Article 22(1) of the GDPR (see above). This significantly expanded the scope of the provision.

Under the Consumer Credit Directive, it is sufficient that the creditworthiness assessment “involves” the use of automated processing of personal data (Article 18(8) of the Consumer Credit Directive). This means that the right to an explanation of the creditworthiness assessment applies even if automated data processing is merely a component of the assessment.

3.3. Right to Explanation in the AI Act

For certain high-risk AI systems, the right to an explanation is also provided for in Article 86(1) of the AI Act. However, this right to explanation has a supplementary character: It applies only insofar as no equivalent right is provided under other provisions of Union law (Article 86(3) of the AI Act).

Article 86(1) of the AI Act applies to so-called high-risk AI systems listed in Annex III of the AI Act, with the exception of AI systems intended to be used as safety components in the management and operation of critical digital infrastructure, road traffic, or the supply of water, gas, heating, or electricity. In contrast, the high-risk AI systems covered include those in the areas of biometrics, education and vocational training, employment, workers’ management and access to self-employment, essential services (such as healthcare, evaluation of creditworthiness, life and health insurance, emergency calls and services), law enforcement, migration, asylum and border control management, as well as the administration of justice and democratic processes.

If a decision is taken based on the output of such a high-risk AI system, any affected person subject to the decision has 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. The prerequisite is that the decision produces legal effects or similarly significantly affects the person in a way that they consider to have an adverse impact on their health, safety, or fundamental rights.

The obligation to provide an explanation applies to deployers, meaning persons, public authorities, agencies, or other bodies using an AI system under their authority, except where the AI system is used in the course of a personal, non-professional activity (Article 3(4) of the AI Act). The right to an explanation is available to every af-

affected person. Despite the broad wording (“person”), it is generally understood that only natural persons are covered by this provision (Hartmann 2024: 12).

Article 86(1) of the AI Act applies to both fully and partially automated decision-making processes, as it refers to decisions made by the deployer “on the basis” of the system’s output. The scope of the right to an explanation is therefore broader than under the GDPR and similar to that in the Consumer Credit Directive. Similar to Article 22(1) of the GDPR, Article 86(1) of the AI Act requires that the decision has legal effects on a person or similarly significantly affects them. Given the nature of high-risk AI systems to which this provision applies, this will regularly be the case. These are AI systems used in highly sensitive areas.

Additionally, from the perspective of the affected individual, the impact must relate to health, safety, or fundamental rights. The range of potentially affected fundamental rights is broad. It includes human dignity, respect for private and family life, protection of personal data, freedom of expression and information, freedom of assembly and association, the right to non-discrimination, the right to education, consumer protection, workers’ rights, the rights of persons with disabilities, gender equality, intellectual property rights, the right to an effective remedy and a fair trial, the right to a defence, the presumption of innocence, and the right to good administration (see Recital 48 of the AI Act). According to the wording of Article 86(1) of the AI Act, it is sufficient that the affected person perceives the impact to be adverse. Therefore, it should be enough for the individual to credibly demonstrate the impairment (Hartmann 2024: 13).

As indicated by Article 86(1) of the AI Act, the affected person has the right to receive a clear and meaningful explanation from the operator regarding the role of the AI system in the decision-making process and “the main elements” of the decision made. However, the regulation does not specify what exactly is meant by “the main elements” (critically also Merkle 2024: 42). There are good reasons to interpret this as referring to the most important factors (Hartmann 2024: 13), including their weighting. It remains to be seen whether the CJEU will extend its interpretation of Article 15(1)(h) of the GDPR, as formulated in *Dun & Bradstreet*, to Article 86(1) of the AI Act.

4. Statutory Right to Explanation in German Law

The German legislator intends to take the new European developments into account when reforming the Federal Data Protection Act (FDPA). Increasing transparency in credit scoring had already been planned in Germany for some time (Coalition Agreement 2021–2025: 17). The regulation of credit scoring is outlined in the new § 37a of the FDPA, which is designed as an exception to the prohibition under Article 22(1) of the GDPR (Parliament 2024: 22). It should be remembered that Article 22(1) of

the GDPR generally prohibits decisions based solely on automated data processing that produce legal effects concerning the data subject or similarly significantly affect them. Article 22(2)(b) of the GDPR provides an exception to this prohibition if the decision is authorized by Union or Member State law to which the controller is subject, and which also lays down suitable measures to safeguard the data subject's rights, freedoms, and legitimate interests. The new § 37a of the FDPA will provide such authorization.

As safeguards for the data subject's rights, freedoms, and legitimate interests, § 37a of the FDPA (draft version) introduces new substantive and procedural requirements for scoring. In particular, § 37a(4) of the FDPA (draft version) regulates the disclosure obligations of controllers in the context of scoring, given that scoring involves "special risks for the data subject, the impact and scope of which often cannot be understood without additional information" (Parliament 2024: 24). The disclosure obligations cover four categories of data: (1) the personal data and criteria used to generate the probability values; (2) the weighting of categories and criteria, as well as the relationships between individual criteria, that have the greatest influence on the probability value; (3) the significance of the specific probability value; (4) the generated probability values and their recipients.

The explanatory notes to the law provide further clarification only with regard to the third category, the significance of the specific probability value. To make this understandable, specifying a Gini coefficient may be considered; additionally, the probability value should be compared to reference values from other segments of the population (Parliament 2024: 24).

According to the proposed § 37a(4) of the FDPA, the information must be made available to the data subject upon request and "in a precise, transparent, understandable, and easily accessible form in clear and simple language", if necessary, in a language tailored to the target group. Just like similar statements made by the CJEU in the *Dun & Bradstreet* case (see above), this can be seen as a positive aspect from the perspective of the Capacity Approach. Ensuring that information about credit scoring is communicated in a way that different groups can understand strengthens individuals' capability to make informed financial decisions. A target-group-specific approach takes into account varying levels of financial literacy and language proficiency, helping to reduce structural inequalities and ensuring that transparency measures lead to actual empowerment.

The right to explanation under the proposed § 37a(4) of the FDPA generally exists alongside the right to explanation under Article 15(1)(h) of the GDPR. While this duplication presents certain legal and practical disadvantages, the creation of a clearer and more detailed regulation at the national level is nevertheless welcome, given the ongoing uncertainties in the interpretation of the European provision. At least in the area of scoring, the proposed § 37a(4) sentence 1 no. 2 of the FDPA clarifies that the weighting of data that most significantly influences the probability value (feature

importance) must be disclosed. Additionally, to explain the significance of an individual score, it must be placed in relation to reference values from other segments of the population, which implies the disclosure of comparison groups.

The exclusion of the right of access provided for in the new § 34(1) of the FDPA to protect trade secrets shall not apply to scoring, as stated in § 37a(5) of the FDPA (draft version). The legislative materials justify this by referring to the particular risks to which the affected individuals are exposed in the context of scoring (Parliament 2024: 24). In light of the CJEU's decision in the *Dun & Bradstreet* case, completely eliminating the protection of trade secrets in the context of scoring goes too far. Such a provision would disregard the fundamental rights of scoring providers, which are enshrined in the EU Charter of Fundamental Rights. Consequently, it would not be compatible with higher-ranking European law.

5. Fundamental Rights Foundations of the Right to Explanation

Given the various statutory manifestations of the right to explanation, the question arises as to whether these different forms share a common origin in the fundamental rights of the affected individual. In this respect, it is worth taking a closer look at the Charter of Fundamental Rights of the European Union (CFR) and the European Convention on Human Rights (ECHR).

Article 8 of the CFR guarantees the right to the protection of personal data. The connection between Article 8 CFR and the GDPR is regularly emphasized by the CJEU. The aim of the GDPR is to ensure a high level of protection for the right to privacy in the processing of personal data (CJEU 4.10.2024 – *Schrems*, para. 45). The right to the protection of personal data includes the principle of transparency, ensuring that individuals are informed about the processing of their data and have control over it. The right to an explanation emerges as a prerequisite for effectively exercising this control. The issue of control is particularly pressing when data processing is carried out by AI, especially through machine learning algorithms.

There are several reasons for this. In human understanding, people should control machines, not the other way around. The reversal of this dynamic evokes fear, a theme explored in numerous science fiction novels and films (Seng 2018: 60 et seq.). The fear of being reduced to an object by a machine can only be mitigated by a sense of control, which transforms us into active subjects and restores our human dignity. The comprehensibility of machine decisions helps to regain real control, rather than merely the illusion of control (Dubovitskaya and Buchholz 2023: 71).

Another reason why the risk of losing control is particularly high lies in the specific characteristics of machine learning algorithms. Their decisions are not causal conclusions, but rather probabilistic calculations based on the analysis of vast amounts of data. As a result, machine learning applications can make errors

that a human would not. These errors may arise from poor data quality, faulty training, or even from the system becoming the target of a cyberattack – a risk that should not be underestimated. Further, there is a risk of discrimination due to biased algorithms or inaccurate data. These particular risks can be mitigated by the right to an explanation, as it can reveal hidden deficiencies in decision-making and thus protect affected individuals from the risks described. In particular, the right to an explanation can help uncover and challenge discriminatory patterns in decision-making. In this way, the right to an explanation also contributes to enforcing the prohibition of discrimination, which is enshrined in Article 21 of the CFR.

It is worth noting that the German Federal Constitutional Court (FCC) already addressed the issue of loss of control in connection with the novel possibilities of data processing in its decision “Right to Be Forgotten I” (FCC 6.11.2019 – *Right to be forgotten I*, para. 90). The FCC considers it the task of fundamental rights, to protect individuals from such a loss of control. In this context, the court refers to the general right of personality, which is derived from Article 2(1) and Article 1(1) of the German Constitution. The general right of personality protects, among other things, informational self-determination and thus represents a national counterpart to the right to the protection of personal data under Article 8 of the CFR.

According to the “Right to Be Forgotten I” ruling, the right to informational self-determination is primarily to be understood as a guarantee that protects individuals against non-transparent data processing and usage. The court states: “It provides protection against third parties seizing individual data and using it in an incomprehensible manner as an instrument to categorize affected individuals into characteristics, types, or profiles over which they have no control, but which are of significant importance for the free development of personality and equal participation in society” (FCC 6.11.2019 – *Right to be forgotten I*, para. 90).

In the cited part of the “Right to Be Forgotten I” ruling, the FCC refers to profiling. It is important to note that profiling is not entirely identical to AI-based decision-making: AI decisions can be made with or without profiling, and profiling can occur without AI. However, at the very least, when an AI decision involves profiling – which is often the case in practice – the FCC likely takes the position that algorithmic calculations must be made comprehensible. This position is also valuable from the perspective of the Capacity Approach. By explicitly protecting individuals against opaque data processing practices that can categorize them without their control or understanding, it directly supports individuals’ capability to actively shape their own lives and participate equally in society. Ensuring transparency and comprehensibility in data-driven decisions empowers individuals with genuine agency and reduces structural inequalities stemming from asymmetric information and power dynamics.

Furthermore, there is a close connection between the right to an explanation and the fundamental right to an effective remedy and to a fair trial (Article 47 of the CFR).

In order for an affected person to exercise their right to an effective remedy, they must be able to understand the basis on which the decision was made. The right to an explanation thus serves as a key instrument in upholding the right to be heard. In an earlier decision, the CJEU recognized the tension between the right to effective remedy and the opacity of conventional AI technologies. The lack of transparency in AI decision-making systems, the Court noted, “could [...] deprive the affected individuals of their right to an effective judicial remedy enshrined in Article 47 of the CFR” (CJEU 21.6.2022 – *Ligue des droits humains*, para. 195; Bäckér 2024: 270). In *Dun & Bradstreet*, the CJEU also refers to Article 47 of the CFR when balancing disclosure and the protection of trade secrets, stating that disclosure is necessary to ensure the effective exercise of the right to an effective remedy (CJEU, 27.2.2025 – *Dun & Bradstreet Austria*, para. 73).

In conclusion, it must be recognized that the right to an explanation is not merely guaranteed by ordinary (European and national) laws but has constitutional roots. However, it would go too far to regard it as an independent fundamental right. Rather, it is an expression of various rights, such as the right to the protection of personal data and the right to an effective remedy.

6. Conclusion, Critique, and Outlook

Europe has made significant progress in establishing the right to explanation of AI Decisions. This development is crucial for fostering trustworthy AI “made in Europe.” The right to an explanation also plays a central role within the Capacity Approach. The Capacity Approach emphasizes that people should not only have formal rights but must also be genuinely able to exercise them effectively. The right to an explanation serves this very purpose by ensuring that affected individuals can understand and, if necessary, challenge automated decisions.

However, the current state is not yet optimal. Although the right to explanation is now recognized in various EU legal acts, such as the GDPR, the Consumer Credit Directive, and the AI Act, as well as in the case law of the CJEU, its current design remains inadequate and problematic. The existing regulations leave several critical questions unanswered, particularly regarding their coherence, practical implementation, and enforceability.

First, there is still no precise definition of what constitutes an “explanation.” In the *Dun & Bradstreet* ruling, the CJEU refers to “the procedure and principles” applied in automated data processing without defining what these “procedure and principles” actually are. The court merely clarifies that the decision must be made comprehensible to the affected person without formulating concrete requirements for explanations. This could incentivize companies to provide formal but largely uninformative explanations.

Second, the tension between transparency and the protection of trade secrets remains partially unresolved. The CJEU delegates this issue to national courts and supervisory authorities, which, due to a lack of technical expertise, may not be able to assess whether the controller's claim that explaining automated decisions would lead to the disclosure of trade secrets is valid. This could result in courts and supervisory authorities increasingly relying on expert opinions, leading to delays in proceedings and legal uncertainty. Additionally, it is likely that different judicial and administrative practices will emerge across Member States, further complicating legal consistency on this issue.

Third, there is an inconsistency between different EU laws. While the right to an explanation under Article 15(1)(h) of the GDPR requires a decision to be based "solely" on automated processing, the Consumer Credit Directive and the AI Act consider it sufficient if the decision is made with the use of automated systems. As a result, the level of protection for affected individuals varies depending on the applicable legal framework. On the other hand, the Consumer Credit Directive and the AI Act are narrower in scope than the GDPR. Article 18(8)(a) of the Consumer Credit Directive applies only to creditworthiness assessments in consumer credit agreements. Similarly, Article 86(1) of the AI Act applies only to certain high-risk AI systems. In contrast, the right to an explanation under Article 15(1)(h) of the GDPR generally applies to any automated decision, meaning that, in theory, an affected person could even request an explanation for personalized advertising. The necessary limitation is then established through the criterion of legal effect or a similarly significant effect (Article 22(1) of the GDPR). It remains to be seen whether this is sufficient to prevent the abusive exercise of the right to explanation under the GDPR.

In the AI Act, the European legislator missed the opportunity to establish clear regulatory requirements for explaining AI decisions. Explainability must be defined and implemented not only from a legal perspective but also from a technical standpoint. Without concrete technical guidelines, the right to an explanation runs the risk to remain a vague and unenforceable concept.

One possible solution is the mandatory use of recognized explainable AI (XAI) methods. AI providers should be required to employ a certified XAI methodology to generate explanations that meet established transparency and comprehensibility standards. This would ensure that affected individuals receive meaningful and understandable explanations rather than generic or overly complex technical descriptions.

Additionally, standardized technical evaluation procedures should be introduced. Regulatory authorities could conduct regular audits of AI explanation mechanisms to verify their effectiveness and quality. These audits would help maintain consistent standards and prevent companies from providing insufficient or misleading explanations. Implementing such measures would enhance trans-

parency, strengthen legal certainty, and ensure that the right to an explanation in AI decision-making is both meaningful and enforceable.

Table 1: Overview of the Key Provisions.

Fundamental Rights	
Article 8 CFR	Guarantees the right to the protection of personal data
Article 21 CFR	Prohibits discrimination based on a wide range of criteria
Article 47 CFR	Guarantees the right to an effective remedy and to a fair trial
Article 2(1) and 1(1) of the German Constitution	Guarantees the general right of personality
General Data Protection Regulation (GDPR)	
Article 22(1) GDPR	Prohibits decisions based solely on automated data processing
Article 22(2)(b) GDPR	Provides an exception to the prohibition of Article 22(1) GDPR
Article 22(3) GDPR	In the case of an automated decision, it requires the implementation of suitable measures to safeguard the data subject's rights and freedoms and legitimate interests
Article 15(1)(h) GDPR	Provides a special right of access for data subjects in cases where decisions based solely on automated processing have legal effects on them or significantly affect them in a similar manner
Article 13(2)(f) and 14(2)(g) GDPR	Contain obligations for data controllers
Consumer Credit Directive (CCD)	
Article 18(1) CCD	Contains obligations to assess the creditworthiness of the consumers
Article 18(8)(a) CCD	Contains the right to request and obtain a clear and comprehensible explanation of the assessment of creditworthiness from the creditor - takes precedence over the right of access under Article 15(1)(h) of the GDPR - covers a broader range of automated decisions than the GDPR

AI Act	
Article 86(1) AI Act	Provides a right to explanation and applies to the high-risk AI systems
Article 86(3) AI Act	Determines the <i>residual</i> character of Article 86 (1) AI Act, meaning it applies only when no equivalent right is provided under other Union law provisions
Federal Data Protection Act (FDPA)	
§ 37a FDPA	Designed as an exception to the prohibition under Article 22(1) of the GDPR
§ 37a(4) FDPA	Regulates the disclosure obligations of controllers in the context of scoring

References

- Buck-Heeb, Petra (2023): “Rechtsunsicherheiten beim KI-Einsatz im Bankensektor”, in: Wertpapiermitteilungen 77(35), pp. 1625–1636.
- Coalition Agreement 2021–2025 between the Social Democratic Party of Germany (SPD), ALLIANCE 90/THE GREENS, and the Free Democrats (FDP), available at: https://www.fdp.de/sites/default/files/2021-11/Koalitionsvertrag%202021-2025_o.pdf, last access: July 30, 2025.
- Dubovitskaya, Elena and Buchholz, Annika (2023): “Die Geschäftsleitung und der Rat des Algorithmus”, in: Zeitschrift für Wirtschaftsrecht 44(2), pp. 63–73.
- Dubovitskaya, Elena and Bosold, Gregor (2024): “Die Schufa, der EuGH und das Recht auf Erklärung – Zugleich Besprechung von EuGH v. 7.12.2023 – C-634/21”, in: Zeitschrift für Wirtschaftsrecht 45(32), pp. 1805–1816.
- Edwards, Lilian and Veale, Michael (2017): “Slave to the Algorithm? Why a ‘Right to an Explanation’ Is Probably Not the Remedy You Are Looking For”, in: Duke Law & Tech. Review 16, pp. 18–84.
- Kumkar, Lea Katharina and Roth-Isigkeit, David (2020): “Erklärungspflichten bei automatisierten Datenverarbeitungen nach der DSGVO”, in: Juristenzeitung 75(6), pp. 277–286.
- Langenbacher, Katja and Bauer, Kevin (2025): “Erklärung von Kreditscores – Der Europäische Gerichtshof urteilt über automatisierte Bonitätsprüfungen”, SAFE Finance Blog, May 28, <https://safe-frankfurt.de/de/aktuelles/safe-finance-blog/details/explaining-credit-scores-the-european-court-of-justice-rules-on-automated-credit-assessments.html>, last access: July 30, 2025.
- Lundberg, Scott M. and Lee, Su-In (2017): “A Unified Approach to Interpreting Model Predictions”, in: Advances in Neural Information Processing Systems 30, pp. 1–10.

- Merkle, Marieke Luise (2024): "Transparenz nach der KI-Verordnung – von der Blackbox zum Open-Book", in: *Recht Digital*, pp. 414–420.
- Metikoš, Ljubiša and Ausloos, Jef (2025): "The Right to an Explanation in Practice: Insights from Case Law for the GDPR and the AI Act", in: *Law, Innovation and Technology* 17(1), pp. 1–36.
- Ribeiro, Marco T., Singh, Sameer and Guestrin, Carlos (2016): "Why Should I Trust You? Explaining the Predictions of Any Classifier", in: *KDD '16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1135–1144.
- Schufa (2024): "SCHUFA arbeitet an neuer Score-Generation", May 21, <https://www.schufa.de/newsroom/bonitaet/neuer-schufa-score-faq/>, last access: July 30, 2025.
- Seng, Leonie (2018): "Mein Haus, mein Auto, mein Roboter? ", in: Rath, Matthias, Krotz, Friedrich and Karmasin, Matthias (eds.), *Maschinenethik. Normative Grenzen autonomer Systeme*, Wiesbaden: Springer VS, pp. 57–72.
- Wachter, Sandra, Mittelstadt, Brent and Floridi, Luciano (2017): "Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation", in: *International Data Privacy Law* 7(2), pp. 76–99.
- Wachter, Sandra, Mittelstadt, Brent and Russell, Chris (2017): "Counterfactual Explanations Without Opening the Black Box. Automated Decisions and the GDPR", in: *Harvard Journal of Law & Technology* 31(2), pp. 841–887.
- Weber, Robert, Kiefner, Alexander and Jobst, Stefan (2018): "Künstliche Intelligenz und Unternehmensführung", in: *Neue Zeitschrift für Gesellschaftsrecht*, pp. 1131–1136.
- Wischmeyer, Thomas (2018): "Regulierung intelligenter Systeme", in: *Archiv des öffentlichen Rechts (AöR)* 143(1), pp. 1–66.

Legal Resources

- Bäcker, Matthias (2024), in: Kühling/Buchner/Bäcker, *Datenschutz-Grundverordnung BDSG*, 4. Auflage, Art. 15 DSGVO.
- BGH 28.1.2014 – VI ZR 156/13, *NJW* 2014(17), pp. 1235–1238.
- Bienemann, Linda (2022), in: Sydow/Marsch, *Datenschutz-Grundverordnung BDSG*, 3. Auflage, Art. 15 DSGVO.
- Buchner, Benedikt (2024), in: Kühling/Buchner/Bäcker, *Datenschutz-Grundverordnung BDSG*, 4. Auflage, Art. 22 DSGVO.
- CJEU 29.1.2008 – C-275/06 (*Promusicae*), *EU:C:2008:54*.
- CJEU 29.7.2019 – C-516/17 (*Spiegel Online*), *EU:C:2019:625*.
- CJEU 21.6.2022 – C-817/19 (*Ligue des droits humains*), *ECLI:EU:C:2022:491*.
- CJEU 4.5.2023 – C 487/21 (*Austrian Data Protection Authority*), *ECLI:EU:C:2023:369*.

- CJEU 7.12.2023 – C-634/21 (SCHUFA Holding (Scoring)), ECLI:EU:C:2023:957.
- CJEU 4.10.2024 – C-446/21 (Schrems), ECLI:EU:C:2024:834.
- CJEU 27.2.2025 – C-203/22 (Dun & Bradstreet Austria), ECLI:EU:C:2025:117.
- Dix, Alexander (2019), in: Simitis/Hoparaung/Spiecker gen. Döhmnn, Datenschutzrecht, 1. Auflage, Art. 15 GDPR.
- FCC 6.11.2019 – 1 BvR 16/13 (Right to be forgotten I), ECLI:DE:BVerfG:2019:rs20191106.1bvro01613.
- Hartmann, Sarah (2024), in: Martini/Wendehorst, Verordnung über künstliche Intelligenz, Art. 86 KI-VO.
- Kamlah, Wulf (2023), in: Plath et al., DSGVO/BDSG/TTDSG: Kommentar zu DSGVO, BDSG, TTDSG, 4. Auflage, Art. 13 DSGVO.
- Opinion of Advocate General Pikamäe, delivered on March 16, 2023 – C-634/21 (SCHUFA Holding (Scoring)), ECLI:EU:C:2023:220.
- Opinion of Advocate General de la Tour, delivered on September 12, 2024 – C-203/22 (Dun & Bradstreet Austria), ECLI:EU:C:2024:745.
- Parliament of the Federal Republic of Germany 27.03.2024 – Printed Matter 20/10859, p. 1–40.
- Schulz, Sebastian (2022), in: Gola/Heckmann Datenschutz-Grundverordnung, 3. Auflage, Art. 22 DSGVO.
- Spindler, Gerald/Horváth, Anna Zsófia (2019), in: Spindler/Schuster, Recht der elektronischen Medien, 4. Auflage, Art. 22 DSGVO.

