



DOT Europe position paper

Algorithmic decision-making: transparency as explainability

Introduction

Public interest in the workings of algorithms is growing as algorithmic decision-making becomes more widespread across the economy. This is sometimes manifested in calls for insight into the functioning of algorithms by making information about code, design documentation, parameters and/or training data available to the public or for audit or certification by public authorities. In this position paper, DOT Europe, the leading voice of digital, online and tech companies in Europe, seeks to provide clarity on the workings of algorithms and propose a set of criteria for transparency in algorithms that is meaningful to citizens and public authorities alike.

What is algorithmic decision-making

Algorithmic decision-making systems are a fundamental aspect of all artificial intelligence (AI) systems; they use algorithms – series of instructions telling a computer how to complete a calculation – to generate outputs such as content, forecasts, recommendations or decisions for a given set of human-defined objectives. AI systems vary significantly in their complexity and techniques, ranging from rules-based algorithms used in email spam filters to complex machine learning, whereby the computer automatically draws inference when exposed to new data by feeding it back into the algorithm.

An example of algorithmic decision-making comes from traffic prediction, where data on historical traffic patterns is combined with live traffic data to generate suggestions on the quickest route for drivers to go from A to B, thereby minimising the risk of being caught in traffic. Similarly, in the health sector algorithms can be used to identify serious illnesses such as prostate cancer by analysing images taken from patients' biopsies. They are shown thousands of images of tissue (training data), before using an algorithm to identify commonalities to detect prostate cancer, which can then inform a doctor's diagnosis. The complexity of algorithmic decision-making can cause the perception of algorithms as being arbitrary and biased, which has in turn led to calls for greater transparency.

Algorithms in EU legislation

Different parts of the acquis deal with algorithmic transparency requirements from perspectives of data protection, B2B contract law and consumer protection. An example of the first approach is the General Data Protection Regulation (GDPR), which contains a "right to an explanation" provision, establishing that where a controller obtains personal data, it must inform the data subject as to "the existence of automated decision-making, including profiling", providing "meaningful information





about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject” (Article 13(2)(f))¹.

The Regulation on platform-to-business relations (P2B) requires that platforms set out in terms and conditions the ‘main parameters determining ranking and the reasons for the relative importance of those main parameters as opposed to other parameters’ (Article 5(1))². Similarly, in the consumer protection acquis the Unfair Consumer Practices Directive (UCPD) requires that the online interface make available general information ‘on the main parameters determining the ranking of products presented to the consumer as a result of the search query and the relative importance of those parameters, as opposed to other parameters’ (Article 7(4a))³. Similar requirements obtain in the revised Price Indication Directive (Article 6a(1)(a))⁴.

The draft Digital Services Act sets out in Article 29 that “[v]ery large online platforms that use recommender systems shall set out in their terms and conditions, in a clear, accessible and easily comprehensible manner, the main parameters used in their recommender systems, as well as any options for the recipients of the service to modify or influence those main parameters that they may have made available⁵.” Similar requirements obtain in respect of content moderation for providers of intermediary service (Article 12) and for online platforms (Article 23); and for hosting services who use automated means for processing notices of illegal content (Article 14(6) or removing or disabling access to specific items of information (Article 15(2)(c))⁶.

The draft AI Act’s Article 11 takes a different approach to the above-mentioned approaches to algorithmic transparency in that the documentation is intended to be read by market surveillance authorities rather than customer and business users. Here, high risk AI systems require technical documentation to be drawn up, which should contain “a detailed description of the elements of the AI system and of the process for its development” (Annex IV(2)) such as “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, also with regard to persons or groups of persons on which the system is intended to be used; the main classification choices; what the system is designed

¹ Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data <https://eur-lex.europa.eu/eli/reg/2016/679/oj>

² Regulation (EU) 2019/1150 of the European Parliament and of the Council of 20 June 2019 on promoting fairness and transparency for business users of online intermediation services <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32019R1150>

³ Directive (EU) 2005/29 of the European Parliament and of the Council of 11 May 2005 concerning unfair business-to-consumer commercial practices in the internal market <https://eur-lex.europa.eu/legal-content/en/ALL/?uri=CELEX%3A32005L0029>

⁴ Directive (EU) 2019/2161 of the European Parliament and of the Council of 27 November 2019 amending Council Directive 93/13/EEC and Directives 98/6/EC, 2005/29/EC and 2011/83/EU of the European Parliament and of the Council as regards the better enforcement and modernisation of Union consumer protection rules <https://eur-lex.europa.eu/eli/dir/2019/2161/oj>

⁵ Proposal for a Regulation of the European Parliament and of the Council on a Single Market For Digital Services (Digital Services Act) and amending Directive 2000/31/EC <https://eur-lex.europa.eu/legal-content/en/TXT/?uri=COM%3A2020%3A825%3AFIN>

⁶ *Ibid.*





to optimise for and the relevance of the different parameters” (Annex IV(2)(a))⁷. Similarly, in the Digital Markets Act there is a provision allowing the Commission to “request access to data bases and algorithms of undertakings and request explanations on those by a simple request or by a decision” (Article 19(1))⁸.

Principles for meaningful transparency

In light of the above we propose a set of criteria to be kept in mind when drafting provisions in this area.

a. Proportionality

Consistent with the approach taken in EU legislation in this area, it is important that **transparency obligations be proportionate to the objective at hand and be linked to a specific right such as consumer- or fundamental rights**. Access to data should be proportionate and serve the overarching goals of shedding light into algorithms and their results while maintaining users’ privacy and safety. Only regulators should have access to that sensitive data with the appropriate safeguards, including strict confidentiality protections. It should be clear that regulators’ powers extend only to information that is necessary to investigate specific, reasonable suspicions of an infringement. And a flexible open line of communication is necessary as companies may not have all the data requested at hand first but can suggest alternative data fulfilling the same goals.

b. Consistency

The code of an algorithm is subject to continuous changes and updates which may have little to no effect on users or are aimed at improving the algorithm’s effectiveness vis-à-vis its function. Access to the algorithmic code will therefore only provide a snapshot, with no guarantees for how long it will stay the same. Transparency should revolve around the need to ensure overall consistency of the algorithms with its overarching principles and design over time and **requiring disclosure when those principles are being changed or revised**.

c. Security and algorithmic integrity

Knowledge of the algorithm’s workings or access to training datasets used for illegal content detection could allow bad-faith actors to circumvent protective measures which are in place to limit the spread of content such as terrorist content and child sexual abuse material. By gaining access to those training datasets or learning how the model behaves, they could modify their content to avoid it being detected by algorithms. Extensive information disclosure requirements around the details of model operation could risk creating similar security vulnerabilities. Similarly, it is possible that bad-faith business users operating on goods or app store marketplace traders will use knowledge of the algorithm to benefit their fraudulent, illegal, unsafe, or simply inferior listings, adversely impacting

⁷ Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206>

⁸ Proposal for a Regulation of the European Parliament and of the Council on contestable and fair markets in the digital sector (Digital Markets Act) <https://eur-lex.europa.eu/legal-content/en/TXT/?uri=COM%3A2020%3A842%3AFIN>





consumer experience. **For this reason, the insights provided should not allow bad-faith actors to undermine the algorithm's function.**

d. Intellectual property (IP) and trade secrets

Provisions on transparency should be consistent with IP rights and avoid discouraging development and innovation. In particular, the Trade Secrets Directive affords protection to information that (i) is secret; (ii) is commercially valuable because of its secrecy; and (iii) has been subject to reasonable steps to keep it a secret (Article 2.1)⁹. That said, the Directive provides exceptions in the case of public, administrative or judicial authorities requesting information “for the performance of the duties of those authorities” (Article 1.2).

Algorithmic explainability

It is appropriate that any information be disclosed in a format relevant to the recipient. Algorithms are mathematical solutions that often consist of complex models such as neural networks implemented by code written in programming language whose functioning is tightly intertwined with the inputted data. Access to the algorithmic code may not provide meaningful insight on the purposes and functioning of the algorithm to laypeople or specialists. Source code is not the best way to measure output of an AI system, and use of regulatory sandboxes would be better placed to demonstrate performance.

For this reason, information made generally available to consumers and business users should describe the algorithm's functioning and objectives in an accessible way intelligible to laypeople, drawing on the concept of explainability, i.e. **providing insights into the functioning of AI models in a format that makes sense to the reader**. In a business-to-business or business-to-customer context this implies an accessible summary of what the programme is doing and of decisions that were made during the algorithm's development, including information about how parameters are weighted. Such provisions would for example allow users of social media services to understand why they are being shown certain types of content, while for P2B contract relations it would give businesses clarity on how to optimise rankings.

On the other hand, where the object of transparency changes from providing information to business and consumer users to assessing compliance with EU regulation on algorithms, the AI providers or deployers should support and equip market surveillance authorities with the necessary facilities to carry out robust testing (e.g., input/output auditing) as well as provide documentation on how to use the product. Regulators should seek to acquire insight into why an AI system behaves in a certain way (e.g. which parts of an image trigger a medical diagnosis tool to identify a disease) based on observing relationships between the inputs and the outputs of the system.

⁹ Directive (EU) 2016/943 of the European Parliament and of the Council of 8 June 2016 on the protection of undisclosed know-how and business information (trade secrets) against their unlawful acquisition, use and disclosure <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32016L0943>





Conclusion

Existing and proposed legislation in this area should find a balance between providing sufficient information to consumers and business users, making this information intelligible while not disclosing sensitive information or providing so much information that bad-faith users can “game the system”. More thoroughgoing transparency requirements would not necessarily increase the general public’s understanding of automated decision-making technology, while increasing the risk to innovation and public policy goals such as privacy and security online. For this reason, DOT Europe calls on decision makers to maintain consistency with existing regulations by promoting forms of algorithmic transparency based on providing general information regarding objectives, parameters and functioning of algorithms, while avoiding transparency requirements that dilute existing IP rights.

