



## UvA-DARE (Digital Academic Repository)

### Know your algorithm: what media organizations need to explain to their users about news personalization

van Drunen, M.Z.; Helberger, N.; Bastian, M.

**DOI**

[10.1093/idpl/ipz011](https://doi.org/10.1093/idpl/ipz011)

**Publication date**

2019

**Document Version**

Final published version

**Published in**

International Data Privacy Law

**License**

CC BY-NC-ND

[Link to publication](#)

**Citation for published version (APA):**

van Drunen, M. Z., Helberger, N., & Bastian, M. (2019). Know your algorithm: what media organizations need to explain to their users about news personalization. *International Data Privacy Law*, 9(4), 220-235. <https://doi.org/10.1093/idpl/ipz011>

**General rights**

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

**Disclaimer/Complaints regulations**

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

*UvA-DARE is a service provided by the library of the University of Amsterdam (<https://dare.uva.nl>)*

# Know your algorithm: what media organizations need to explain to their users about news personalization

M.Z. van Drunen\*, N. Helberger\* and M. Bastian\*

## Key Points

- If the right to an explanation is expected to effectively safeguard users' rights, it must be interpreted in a manner that takes the contextual risks algorithms pose to those rights into account.
- This article provides a framework of transparency instruments in the context of the news personalization algorithms employed by both traditional media organizations and social media companies.
- Explaining the impact on a user's news diet and the role of editorial values in the algorithm is especially important in this context.
- Conversely, explanations of individual decisions and counterfactual explanations face specific practical and normative barriers that limit their utility.

## Introduction

Citizens' online lives are shaped by algorithms that tailor their search results, rate their credit, or sort their news. Their knowledge of how, why, or even if they are affected by the algorithms that judge them is typically limited and

sometimes mistaken, however.<sup>1</sup> Moreover, they have traditionally lacked easy access to the information needed to remedy this situation. This is problematic. A lack of algorithmic transparency can hide discrimination, enable manipulation, or cause individuals to blindly (dis)trust algorithmic decision-making. However, making algorithms transparent can be costly, and organizations have few incentives to show users how their algorithms engage in manipulation or discrimination.<sup>2</sup> This incentive-structure has fuelled increasing calls for regulation.<sup>3</sup> In particular, an individual transparency mechanism in the General Data Protection Regulation (GDPR) that entitles users to information about these automated processes, commonly referred to as the right to an explanation, has recently been seized on to combat the lack of algorithmic transparency.<sup>4</sup>

However, which (if any) kind of individual transparency is a suitable policy tool to regulate algorithms remains a contested issue. Many critiques take an instrumental approach to transparency, arguing that transparency (especially with regard to private parties) is primarily valued because of the goals it advances. Making more information available does not automatically advance these goals, as users often lack the ability or motivation to process complex technical information.<sup>5</sup> Not only the form, but also the content of the information that is to be

\* Institute for Information Law, University of Amsterdam, Amsterdam, the Netherlands. E-mail: m.z.vandrunen@uva.nl

This article was made possible by the European Research Council under Grant 638514 (PersoNews). The authors report no conflict of interest. The authors would like to thank J.P. Quintais, prof. dr. T. McGonagle, and the anonymous reviewer for their insightful comments on this article.

- 1 See for example the 2018 Reuters Institute report, which found only 29% of respondents answered correctly that an algorithm takes most of the individual decisions regarding the stories to include in someone's newsfeed. Nic Newman and others, 'Reuters Institute Digital News Report 2018' (2018) 34 <<https://reutersinstitute.politics.ox.ac.uk/sites/default/files/digital-news-report-2018.pdf>> accessed 23 July 2019. See also Emilee Rader and Rebecca Gray, 'Understanding User Beliefs About Algorithmic Curation in the Facebook News Feed', *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (ACM Press, New York 2015) <<http://dl.acm.org/citation.cfm?doid=2702123.2702174>> accessed 23 July 2019.
- 2 Jenna Burrell, 'How the Machine "Thinks": Understanding Opacity in Machine Learning Algorithms' (2016) 3 *Big Data & Society* 1 <[http://](http://journals.sagepub.com/doi/10.1177/2053951715622512)

[journals.sagepub.com/doi/10.1177/2053951715622512](http://journals.sagepub.com/doi/10.1177/2053951715622512)> accessed 23 July 2019.

- 3 See, in the context of the media Nicholas Diakopoulos and Michael Koliska, 'Algorithmic Transparency in the News Media' (2017) 5 *Digital Journalism* 809 <<http://www.tandfonline.com/doi/full/10.1080/21670811.2016.1208053>> accessed 23 July 2019.
- 4 Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), OJ 2016 L 119/1.
- 5 See generally David Albert Heald, 'Varieties of Transparency' in Christopher Hood and David Heald (eds), *Transparency: The Key to Better Governance? Proceedings of the British Academy* 135 (OUP, Oxford 2006). Matteo Turilli and Luciano Floridi, 'The Ethics of Information Transparency' (2009) 11 *Ethics and Information Technology* 105, 222. Lilian Edwards and Michael Veale, 'Slave to the Algorithm: Why a Right to an Explanation Is Probably Not the Remedy You Are Looking For' (2017) 16 *Duke Law & Technology*

communicated can be unsuitable to reach the desired goal. Disclosures may not be relevant to the particular context in which individuals try to exercise their rights, may be shown not to enhance trust, or may be unable to produce accountability where individuals are unable to exert pressure on the party that should be held accountable.<sup>6</sup> Disregarding the connection between a disclosure and the goal it pursues risks creating a transparency fallacy, in which transparency functions as an ineffective solution that wards off stronger regulation.<sup>7</sup>

This article will argue that the goals of the right to an explanation and the information that is suitable to further them cannot be assessed in isolation. Instead, like other general open transparency norms in EU law, the right to an explanation of automated decision-making must be interpreted in the context of the decision-making that is to be explained.<sup>8</sup> This context shapes the goals the right to an explanation aims to promote. An algorithm that decides on credit applications, for example, puts different rights at risk than an algorithm that determines what news an individual is able to see. If the right to an explanation is expected to be a suitable safeguard against risks to individuals' rights and interests, such contextual risks must be taken into account. Existing sector-specific literature already contains well-developed insights on the exact nature of these risks and the corresponding suitable disclosures.<sup>9</sup> Drawing on these insights provides a better normative foundation for the information required by the right to an explanation, and ensures consistency between data protection law and the public values that are specific to the many different contexts in which data protection law applies.

To operationalize the context-specific approach, this article focuses on news personalization algorithms, a form of profiling that uses personal data to tailor the selection, content, or layout of news stories to an individual's characteristics or preferences.<sup>10</sup> Such algorithms are increasingly adopted by the legacy media and are already at the

core of the social media platforms where individuals spend much of their time.<sup>11</sup> This allows the right to an explanation to provide protection with regard to algorithms that play a large role in individuals' daily lives and that have the potential to significantly affect their fundamental rights, in particular their right to receive information. At the same time, legal obligations regarding the use of algorithms, including the disclosures required under the right to an explanation, must stay within the limits set by the freedom of expression. There is a rich history in media law and literature regarding disclosures that stay within these limits while allowing readers to safeguard their own interests, address declines in trust, and promote accountability—goals that feature heavily in the GDPR but are understood and achieved in a particular way in the media.

To understand why and how the right to an explanation needs to be tailored to the context in which an algorithm is implemented, this article asks what concrete disclosures are necessary to further the goals of the right to an explanation in the context of news personalization. It develops its context-specific approach in three steps. Firstly, it reviews existing work on the right to an explanation in order to identify the limits of the right, and argue why a contextual approach is necessary to understand which disclosures are suitable to advance its goals. Secondly, it analyses the goals of the right to an explanation and explores how they are understood in the context of the media. Finally, it draws on literature on media and algorithmic transparency to develop a conceptual framework of disclosures that are suitable to further the goals of the right to an explanation in the context of news personalization.

## The right to an explanation: consensus and gaps

To understand what the right to an explanation requires in a specific context, it is first necessary to outline its

Review 18, 39 <<https://scholarship.law.duke.edu/cgi/view-content.cgi?article=1315&context=dltr>> accessed 23 July 2019.

6 Mike Ananny and Kate Crawford, 'Seeing without Knowing: Limitations of the Transparency Ideal and Its Application to Algorithmic Accountability' (2016) 20 *New Media & Society* 973.

7 Edwards and Veale (n 5).

8 Art 7 of the Unfair Commercial Practices Directive, for example, prohibits misleading omissions, and covers information requirements ranging from the labelling of advertisements to the disclosure of shipping costs. Directive 2005/29/EC of the European Parliament and of the Council of 11 May 2005 concerning unfair business-to-consumer commercial practices in the internal market and amending Council Directive 84/450/EEC, Directives 97/7/EC, 98/27/EC and 2002/65/EC of the European Parliament and of the Council and Regulation (EC) No 2006/2004 of the European Parliament and of the Council (Unfair Commercial Practices Directive), OJ 2005 L149.

9 The Council of Europe's recommendation on profiling, referred to in the Commission's proposal for art 22 GDPR, similarly notes 'that it is

desirable to assess the different situations and purposes in a differentiated manner' Council of Europe, 'Recommendation CM/REC(2010)13 of the Committee of Ministers to Member States on the Protection of Individuals with Regard to Automatic Processing of Personal Data in the Context of Profiling' (2015) <[https://search.coe.int/cm/Pages/result\\_details.aspx?ObjectID=09000016805cdd00](https://search.coe.int/cm/Pages/result_details.aspx?ObjectID=09000016805cdd00)> accessed 23 July 2019. See also Andrew Selbst and Julia Powles, 'Meaningful Information and the Right to Explanation' (2017) 7 *International Data Privacy Law* 233 <<http://academic.oup.com/idpl/article/7/4/233/4762325>> accessed 23 July 2019.

10 Neil Thurman and Steve Schifferes, 'The Future of Personalization at News Websites: Lessons from a Longitudinal Study' (2012) 13 *Journalism Studies* 775 <<http://www.tandfonline.com/doi/abs/10.1080/1461670X.2012.664341>> accessed 23 July 2019.

11 Nic Newman, 'Journalism, Media, and Technology Trends and Predictions 2018' (2018) <<https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2018-01/RISJ%20Trends%20and%20Predictions%202018%20NN.pdf>> accessed 23 July 2019.

more general applicability and limits. The right to an explanation is used as a catch-all term for the provisions in the GDPR that give users a right to information about the logic and consequences of automated decisions.<sup>12</sup> Versions of the right can be found in Articles 13–15 and 22 GDPR. Articles 13(2)(f) and 14(2)(g) and 15(h) uniformly require that data subjects are informed of ‘the existence of automated decision-making, including profiling, referred to in Article 22(1) and (4) and, at least in those cases, meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject’.

Article 22(1) gives users a 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’. While the wording is ambiguous, the provision is commonly interpreted as a prohibition, as opposed to a right that needs to be actively invoked.<sup>13</sup> Article 22(3) requires that safeguards are put in place when this prohibition does not apply because the decision is necessary for a contract or based on explicit consent. Recital 71 clarifies that the right ‘to obtain an explanation of the decision reached after such assessment’ may be one such safeguard. Additionally, Mendoza and Bygrave argue that the right to information about a decision is also necessary for the effective exercise of the right to contest it, as provided by Article 22(3) GDPR.<sup>14</sup>

Much has been written about the conditions under which the right to an explanation becomes applicable. The goal of this article is not to relitigate that debate, or explore the applicability of the GDPR more generally.

Instead, it assumes that news personalization, like online behavioural advertising, can likely fall under the relevant legislative provisions under certain circumstances.<sup>15</sup> On that basis, it will assess what specific kinds of information need to be communicated about news personalization to further the goals of the right to an explanation.

### The information to which data subjects are entitled: indications and limits

Neither the provisions themselves nor the relevant recitals indicate what concrete information data subjects are entitled to under their right to an explanation.<sup>16</sup> As a result, scholars are faced with the difficult task of determining what information is ‘meaningful’, and which (transparency) measures are ‘suitable’. There has been some debate as to whether suitable measures include a transparency requirement at all. The examples provided in Article 22 itself do not explicitly require any disclosures, and the recital that refers to the need for an explanation is, of course, non-binding.<sup>17</sup> Recitals can and have played an important role in the interpretation of EU (data protection) law, however. While the Article 29 Working Party (A29WP) appears careful not to state explicitly that Article 22(3) requires an explanation in all cases, it emphasizes the importance of transparency as an appropriate safeguard and notes that recital 71 interprets suitable safeguards as always including a right to obtain an explanation.<sup>18</sup> This article therefore operates under the assumption that Article 22 can impose transparency requirements, and is especially likely to do so

- 12 Coined by Goodman and Flaxman in reference to recital 71, it is now used more broadly to refer to the provisions in the GDPR that concern transparency in automated processing. This article is agnostic as to whether the term accurately describes the information users are entitled to under these provisions. Bryce Goodman and Seth Flaxman, ‘European Union Regulations on Algorithmic Decision-Making and a “Right to Explanation”’, *arXiv preprint* (2016). Selbst and Powles (n 9). 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’ (2017) 7 *International Data Privacy Law* 76 <<https://academic.oup.com/idpl/article-lookup/doi/10.1093/idpl/ix005>> accessed 23 July 2019; Gianclaudio Malgieri and Giovanni Comandé, ‘Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation’ (2017) 7 *International Data Privacy Law* 243 <<http://academic.oup.com/idpl/article/doi/10.1093/idpl/ix019/4626991>> accessed 23 July 2019.
- 13 Art 29 Working Party, ‘Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679 (WP251 Rev.01)’ 19 <[http://ec.europa.eu/newsroom/article29/item-detail.cfm?item\\_id=612053](http://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=612053)> accessed 23 July 2019. (Automated Decision-Making Guidelines). Isak Mendoza and Lee A Bygrave, ‘The Right Not to Be Subject to Automated Decisions Based on Profiling’ in TE Synodinou and others (eds), *EU Internet Law* (Springer International Publishing, Cham 2017) 85 <[http://link.springer.com/10.1007/978-3-319-64955-9\\_4](http://link.springer.com/10.1007/978-3-319-64955-9_4)> accessed 23 July 2019. BC van Breda, ‘Profiling in de AVG: Nieuwe Regels, Voldoende Bescherming?’ [2017] *Computerrecht* 1, 2. Bryan Casey, Ashkon Farhangi and Roland Vogl, ‘Rethinking Explainable Machines: The GDPR’s “Right to

Explanation” Debate and the Rise of Algorithmic Audits in Enterprise’ [2018] SSRN 1, 14 <[https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3143325#%23](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3143325#%23)> accessed 23 July 2019.

- 14 Mendoza and Bygrave, *ibid.* see similarly Automated Decision-Making Guidelines, *ibid.*
- 15 Automated Decision-Making Guidelines (n 13) 22. See also the suggestion that profiling can ‘undermine [data subjects] ability to choose, for example, . . . newsfeeds’ on p 5 of the same guidelines.
- 16 This issue was also flagged in the deliberations, most bluntly by the Slovak delegation: ‘It is not clear for us at all, so we appreciate clearer reformulation of this part in the way which [sic] could provide better and stricter interpretation’ Council of the European Union, ‘Comments from the Slovak Republic on Articles 11-27 of the Draft General Data Protection Regulation (14147/12 ADD 1)’ (2014) <<http://data.consilium.europa.eu/doc/document/ST-14147-2012-ADD-1/en/pdf>> accessed 23 July 2019.
- 17 Wachter, Mittelstadt and Floridi (n 12.); see for a different perspective Margot E Kaminski, ‘The Right to an Explanation, Explained’ (2018) <[https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3196985](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3196985)> accessed 23 July 2019.
- 18 Automated Decision-Making Guidelines (n 13) 27. Michael Veale and Lilian Edwards, ‘Clarity, Surprises, and Further Questions in the Article 29 Working Party Draft Guidance on Automated Decision-Making and Profiling’ (2018) 34 *Computer Law & Security Review* 398 <<https://www.sciencedirect.com/science/article/pii/S026736491730376X>> accessed 23 July 2019.

where such requirements are suitable to safeguard users' rights and the effectiveness of EU law.<sup>19</sup>

The breadth of the term 'meaningful' makes it difficult to determine what information it covers. Different language versions of the term have provided a useful first indication as to its meaning. Several authors have noted that the term can be understood as 'comprehensible', but that other language versions of the GDPR reflect different aspects of the term, namely significant and useful.<sup>20</sup> As article 12(1) already requires any information in the relevant articles is communicated to users in a way they can understand it, the information must also be significant and/or useful if the 'meaningful' is to have any added value.

This raises a follow-up question: meaningful or suitable for what? On this point, Article 22 is refreshingly clear: measures must at least be suitable 'to safeguard the data subject's rights and freedoms and legitimate interests'. Selbst and Powles advocate a functional approach, and argue that information should at a minimum enable data subjects to exercise their rights under the GDPR and human rights law.<sup>21</sup> Other scholars also point to other goals, such as accountability and trust.<sup>22</sup> The contextual nature of these goals and the information that is suitable to further them will be explored in the sections 'The goals of the right to an explanation in context' and 'A conceptual framework of algorithmic transparency in the media', respectively.

What information users are entitled to may also depend on the time at which information must be communicated. Wachter, Mittelstadt, and Floridi argue that the law cannot require that a specific decision is explained before it has been taken. Information that must be communicated before data is processed can therefore only concern the algorithm in general, and not its specific output.<sup>23</sup> Selbst and Powles point out that this distinction is not quite that stark with regard to deterministic models, as the specific decision they

will take can be predicted as soon as the input for that decision is known (and as such, before a final decision takes place). Information about the model will therefore allow users to understand specific decisions, and controllers can explain individual future decisions as long as they know the input they will be based on.<sup>24</sup> However, where users lack the time or understanding necessary to infer information about specific decisions from the model, explanations of individual decisions will have an added value and increase the likelihood that the rights provided by the GDPR are effective in practice.

Timing is not mentioned explicitly in the provisions that constitute the right to an explanation. However, recital 71's reference to an explanation after an assessment and transparency's function as prerequisite for the right to object both provide a strong indication that Article 22 requires that users also have access to information after a decision concerning them has been taken. In contrast, Articles 13 and 14 tie the moment when the data subject must be informed to the moment when his data is acquired.<sup>25</sup> As such, they will often require that users are informed before the processor knows what data will be used in a specific future decision. Finally, Article 15 is triggered upon a user's request and as such, could apply at any point in time after his data is collected. In practice, it is likely that a data subject would invoke the right after being confronted with a decision that affects him particularly negatively. Nevertheless, the A29WP asserts that while article 15 allows a user to become aware of a decision, it does not entitle him to an explanation of the factors that underlie that particular decision.<sup>26</sup> This considerably lowers the chances of article 15(h) providing any added value with regard to the information about the algorithm that must be provided.<sup>27</sup>

A number of concrete information requirements have been suggested in the context of these indications and limits. Disclosure of the source code is used as an

19 Koen Lenaerts and Jose A Gutierrez-Fons, 'To Say What the Law of the EU Is: Methods of Interpretation and the European Court of Justice' (2013) 20 *Columbia Journal of European Law* 3, 20 <[20 Malgieri and Comandé \(n 12\) 257. In contrast, 'suitable' is generally translated in comparable ways: \*angemessene\* \(German\), \*appropriées\* \(French\), \*appropriate\* \(Italian\), \*adecuadas\* \(Spanish\), \*właściwe\* \(Polish\), \*corespunzătoare\* \(Romanian\), \*passende\* \(Dutch\).](https://heinonline.org/HOL/Page?handle=hein.journals/coljeul20&id=183&div=&collection=> accessed 23 July 2019.</a></p>
</div>
<div data-bbox=)

21 Selbst and Powles (n 9) 236.

22 Kaminski (n 17). Tae Wan Kim and Bryan R. Routledge, 'Informational Privacy, A Right to Explanation, and Interpretable AI', *2018 IEEE Symposium on Privacy-Aware Computing (PAC)* (2018) <<https://ieeexplore.ieee.org/document/8511831>> accessed 23 July 2019.

23 Wachter, Mittelstadt and Floridi (n 12) 83.

24 Selbst and Powles (n 9) 239–41.

25 Art 13 requires information is communicated as soon as data is collected from the data subject. Art 14(a)–(c) applies when data is acquired from a third party and provides for a more flexible timeframe, which the A29WP interprets as having a maximum limit of one month after data is acquired. Art 29 Working Party, 'Guidelines on Transparency under Regulation 2016/679 (WP260 Rev.01)' 16 <[http://ec.europa.eu/newsroom/article29/item-detail.cfm?item\\_id=622227](http://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=622227)> accessed 23 July 2019. (Transparency Guidelines).

26 Automated Decision-Making Guidelines (n 13) 25–27. Veale and Edwards (n 18) 399.

27 This in turn raises questions with regard to art 15's recital 63, which contains an exception for 'trade secrets or intellectual property and in particular the copyright protecting the software' that is omitted with regard to arts 13 and 14. If art 15(h) does not entitle users to additional information, the recital is either irrelevant or applies to other information that must be communicated under article 15.

example of information that is not required by the right to an explanation more often than it is seriously proposed, and indeed the costs of its disclosure seem to outweigh its contribution to the goals of the right to an explanation.<sup>28</sup> Proposed disclosures are typically more limited and more easily understandable, and include the data the algorithm takes into account and how it is weighted, details about the profile it is used to construct, and what the algorithm aims to do.<sup>29</sup> In particular the (weighting of) factors and input data are also commonly proposed as ways to better explain individual decisions. With regard to the consequences that must be explained, most opt for concrete ways in which the algorithm might impact the data subject in the future, such as the potential ineligibility for a loan and an increased insurance premium, or broader information on an algorithm's goals or statistical impact.<sup>30</sup>

### A contextual approach to the right to an explanation

The discussion above provides useful insights regarding the information that can be required under the right to an explanation. However, why (only) these specific disclosures are necessary to meet the requirements and further the goals of the right to an explanation is not always clarified. Moreover, the argument that a specific list of information must always be communicated to users assumes that the right to an explanation requires the same kind and level of transparency regardless of context. This assumption is valid when different algorithms have shared characteristics that create similar problems. As long as algorithms produce inaccurate results, for example, transparency about their output can promote accountability.

However, the decision-making processes algorithms now automate have traditionally fallen under open norms and sector-specific transparency policies that tailor the information that must be communicated to the unique characteristics of specific sectors. Glossing over these differences with a uniform approach to the right to an explanation would undermine the effectiveness of the right. In other words, while the GDPR may aim for a minimum level of protection across sectors, this

outcome cannot be realised with uniform measures. Instead, open norms such as the right to an explanation must be tailored to the context in which the algorithm is implemented. To facilitate this process, it is important that overarching discussions of the right are increasingly complemented by discussions of how it should be understood in specific sectors.

The right to an explanation's text and goals support such an approach. Because different algorithms affect different rights and interests, the information that is meaningful or suitable to safeguard these rights and interests is inherently contextual. This can be illustrated by comparing the different goals pursued by transparency in radically different sectors, such as the media and self-driving cars. For the media, transparency regarding the (commercial) source of content has traditionally been an important way of allowing the audience to independently evaluate the information that shapes their opinions. This concern is largely absent from discussions about self-driving cars, where information that allows an accurate judgment of their safety is prioritised. In other words, specific information that is relatively trivial in one context may be vital in another.

Not only do different contexts prioritize different goals of transparency; the same goal can also be understood and attained differently in different contexts. The role of trust in the algorithms that power self-driving cars and news personalization again illustrates this difference. Both kinds of algorithms benefit from being perceived as trustworthy, as individuals have to be willing to be vulnerable to the potential negative consequences of relying on them without being able to control or monitor them.<sup>31</sup> However, because they can impact individuals' lives in very different ways, they expose users to very different vulnerabilities. Definitions of trust in the media, for example, typically focus on the media's task to select the relevant facts and topics to report to users.<sup>32</sup> This type of trust is irrelevant in the context of self-driving cars. Similarly, different characteristics of the trustee can also be relevant to varying degrees depending on the context. If the risk of relying on an algorithm is high, trust in its ability to perform well is of great importance.

28 When source code is mentioned in the context of the right to an explanation, it is to mention the limited usefulness of its disclosure to data subjects or its incompatibility with trade secrets. Selbst and Powles (n 9) 239. Kaminski (n 17). Interestingly, the A29WP does not categorically exclude the possibility: Automated Decision-Making Guidelines (n 13) 25.

29 Wachter, Mittelstadt and Floridi (n 12). Automated Decision-Making Guidelines (n 13) 16. Maja Brkan, 'AI-Supported Decision-Making under the General Data Protection Regulation', *Proceedings of the 16th international conference on Artificial intelligence and law* (2017) 5 <<https://doi.org/10.1145/3086512.3086513>>. Edwards and Veale (n 5) 55.

30 Mendoza and Bygrave (n 13) 93. Malgieri and Comandé (n 12) 258. Wachter, Mittelstadt and Floridi (n 12) 84.

31 Kevin Anthony Hoff and Masooda Bashir, 'Trust in Automation' (2015) 57 *Human Factors: The Journal of the Human Factors and Ergonomics Society* 407 <<http://journals.sagepub.com/doi/10.1177/0018720814547570>> accessed 23 July 2019.

32 Matthias Kohring and Jörg Matthes, 'Trust in News Media: Development and Validation of a Multidimensional Scale' (2007) 34 *Communication Research* 231 <<http://crx.sagepub.com/content/34/2/231.abstract>>; Katherine M Grosser, 'Trust in Online Journalism' (2016) 4 *Digital Journalism* 1036 <<http://dx.doi.org/10.1080/21670811.2015.1127174>> accessed 23 July 2019.

Conversely, if an algorithm can perform well on a low-risk task while at the same time exploiting another vulnerability of the person who uses it, trust in its integrity can become highly relevant.<sup>33</sup>

Finally, different contexts impose different regulatory constraints that determine the scope and content of transparency obligations. In the context of news personalization, the freedom of expression in particular must be balanced with the right to data protection to minimize conflicts between the rights. To that end, Article 85 GDPR requires Member States to provide for derogations to certain provisions of the GDPR (including those that make up the right to an explanation) insofar as strictly necessary to safeguard the freedom of expression.<sup>34</sup> For example, journalists may under certain conditions need to be exempted from the obligation to disclose their identity or notify data subjects from which source their personal data was obtained in order to continue to be able to conduct investigative journalism.<sup>35</sup> Explaining news personalization may also create new tensions with the freedom of expression, for example when counterfactual explanations instruct individuals not to read certain stories, creating potential chilling effects. Conversely, as the section ‘The goals of the right to an explanation in context’ will argue, the right to an explanation can also alleviate the tension between the freedom of expression and the right to data protection by enabling users to better assess news recommendations on their value and helping them to make more informed choices regarding the information they consume. Whether one of data protection law’s transparency requirements promotes the freedom of expression, or restricts it to such an extent that an exemption for the media is strictly necessary depends on what specific information is disclosed, and potential conflicts will be highlighted throughout the article as they arise. In general, however, the impact of an explanation of news personalization on the freedom of expression is limited compared to other elements of data protection

law, as it stops short of directly limiting the media’s ability to gather or publish information and merely requires an explanation of the way in which content that has already been produced is disseminated to individual readers. This reduces the likelihood that exemptions from the right to an explanation for media organizations that use news personalization are strictly necessary to safeguard the freedom of expression.<sup>36</sup>

In short, this article argues that the concrete disclosures required by the right to an explanation must be identified by analysing not only its text and legislative history, but also the way in which its goals must be understood in a specific context. The following sections will explore what a contextual approach to the right to an explanation might look like with regard to news personalization. The section ‘The goals of the right to an explanation in context’ will therefore expand on the way the goals of the right to an explanation should be understood in the media. The section ‘A conceptual framework of algorithmic transparency in the media’ will analyse what information is traditionally considered meaningful or suitable to achieve these goals in media law and literature, and adapt these discussions to the algorithmic context.

## The goals of the right to an explanation in context

Before exploring the goals of GDPR, it is important to clarify what can and cannot be expected of the right to an explanation. Both algorithmic and media transparency can advance goals that fall outside the scope of the right to an explanation, such as the improvement of user satisfaction, the promotion of social cohesion, or the monitoring of media concentrations.<sup>37</sup> As these are not goals of the right to an explanation, they cannot guide a teleological interpretation of the right. Conversely, accountability, agency, and to a lesser extent trust feature heavily in the policy and literature

33 For example, an algorithm can use innocuous data for one task while at the same time inferring sensitive data that can be sold to third parties. See for the differentiated impact of privacy violations on the elements of perceived trustworthiness Gaurav Bansal and Fatemeh Mariam Zahedi, ‘Trust Violation and Repair: The Information Privacy Perspective’ (2015) 71 *Decision Support Systems* 62 <<https://www.sciencedirect.com/science/article/pii/S0167923615000196>> accessed 23 July 2019.

34 Christopher Docksey, ‘Four Fundamental Rights: Finding the Balance’ (2016) 6 *International Data Privacy Law* 195, 204.; recital 153 GDPR; *Satamedia* case no. C-37/07 (ECJ, 16 December 2008), para. 56

35 David Erdos, ‘European Union Data Protection Law and Media Expression: Fundamentally Off Balance’ (2016) 65 *International & Comparative Law Quarterly* 139, 144. Dominic Brody and others, ‘Journalism and Media Privilege’ (Council of Europe 2017) <<https://rm.coe.int/journalism-and-media-privilege-pdf/1680787381>> *Buivids* case no. C-345/17 (ECJ, 14 February 2019); see analogously *IPI* case no. C-473/12 (ECJ, 7 November 2013).

36 See with regard to the overarching discussion that EU data protection law affords too little weight to the freedom of expression Maja Brkan, ‘The Unstoppable Expansion of the EU Fundamental Right to Data Protection: Little Shop of Horrors?’ [2016] *Maastricht Journal of European and Comparative Law* <<https://journals.sagepub.com/doi/abs/10.1177/1023263X1602300505>> accessed 23 July 2019.

37 See for an overview of the goals of explanations in recommender systems generally Nava Tintarev and Judith Masthoff, ‘Explaining Recommendations: Design and Evaluation’ in Francesco Ricci, Lior Rokach and Bracha Shapira (eds), *Recommender Systems Handbook* (Springer, New York 2015) <[http://link.springer.com/10.1007/978-1-4899-7637-6\\_10](http://link.springer.com/10.1007/978-1-4899-7637-6_10)> accessed 23 July 2019; Council of Europe, ‘Recommendation CM/Rec(2018)1 of the Committee of Ministers to Member States on Media Pluralism and Transparency of Media Ownership’ (2018) <<https://rm.coe.int/1680790e13>> accessed 23 July 2019.

discussions on the right to an explanation.<sup>38</sup> The following sections will draw on insights from media law and literature to explore how these goals of the right to an explanation should be understood in the context of news personalization algorithms.

## Accountability

The combination of algorithms' increasingly important role in everyday life and their potential for intended or unintended negative effects has fuelled increasing calls for algorithmic accountability. Accountability's status as a newly minted principle of data protection law reflects this focus. Under Article 5(2) GDPR, the principle requires that controllers not only adhere to the other principles of data protection law, but also that they be able to demonstrate such compliance. The recitals expand on this point, noting that users must, for example, be able to obtain information regarding the logic and consequences of processing so as to verify that it is occurring lawfully.<sup>39</sup>

While transparency and accountability are sometimes used interchangeably, it is important to distinguish the two.<sup>40</sup> Transparency is an important part of the process which may lead to accountability, as transgressors cannot be held accountable for mistakes that cannot be discovered.<sup>41</sup> However, transparency by itself is not sufficient to produce accountability. This also requires that the information recipient is able to process and act on the information, and that the accountable party is vulnerable to the pressure the information recipient can exert.<sup>42</sup>

Whether vulnerability to outside pressure is desirable takes on a special meaning for the news media due to the importance of a free press to democratic society. Accountability can be in tension with the freedom of the press, especially where political interference restricts the press' ability to inform the public. The accountability mechanisms through which responsible conduct can be compelled must therefore be tailored to the unique position of the press.<sup>43</sup> This is reflected in media accountability's focus on the pressure the market, other professionals, and the public can exert.<sup>44</sup> These parties cannot prohibit specific editorial choices, but can opt not to use types of news personalisation they do not perceive to be in their interest, or push a media organization to act in the interest of its audience or the public. In so doing, they can provide a form of algorithmic accountability that stops short of prescribing which editorial choices news organizations may or may not make, limiting the likelihood that an exemption is strictly necessary under Article 85 GDPR.

Understanding the specific nature of accountability in the media is therefore necessary to understand how transparency can enable accountability in the context of news personalization. Moreover, media law and literature has a number of well-developed insights on the relationship between transparency and accountability that can inform the interpretation of the right to an explanation in the context of news personalization. Firstly, instruments like ethics codes and editorial guidelines can clarify the norms media organizations intend to be accountable for upholding.<sup>45</sup> This facilitates a broader debate on the appropriateness of these values, as well as

38 Sandra Wachter, Brent Mittelstadt and Chris Russell, 'Counterfactual Explanations Without Openings the Black Box: Automated Decisions and the GDPR' (2018) 31 *Harvard Journal of Law & Technology* 23 <<https://jolt.law.harvard.edu/assets/articlePDFs/v31/Counterfactual-Explanations-without-Opening-the-Black-Box-Sandra-Wachter-et-al.pdf>> accessed 23 July 2019.

39 Recital 63 GDPR

40 Maayan Perel and Niva Elkin-Koren, 'Black Box Tinkering. Beyond Transparency in Algorithmic Enforcement' (2017) 69 *Florida Law Review* 181 <[http://www.floralawreview.com/wp-content/uploads/Perel\\_Elkin-Koren.pdf](http://www.floralawreview.com/wp-content/uploads/Perel_Elkin-Koren.pdf)> accessed 23 July 2019; Edwards and Veale (n 5) 41. Tal Zarsky, 'Transparent Predictions' [2013] *University of Illinois Law Review* 1503, 1533 <<https://illinoislawreview.org/wp-content/ilr-content/articles/2013/4/Zarsky.pdf>> accessed 23 July 2019.

41 See for accountability mechanisms that do not rely on transparency Joshua A Kroll and others, 'Accountable Algorithms' (2016) 165 *University of Pennsylvania Law Review* 633 <<https://heinonline.org/HOL/P?h=hein.journals/pnlr165&i=649>> accessed 23 July 2019.

42 Ananny and Crawford (n 6). See in the context of the traditional news media Stephen JA Ward, 'The Magical Concept of Transparency' in David Craig Lawrie Zion (ed), *Ethics for Digital Journalists Emerging Best Practices* (Routledge 2014) 49 <<https://www.taylorfrancis.com/books/e/9781135114237/chapters/10.4324%2F9780203702567-10>> accessed 23 July 2019. Of course users have several options under the GDPR to effect change when they perceive processing to be unlawful or unfair, including

direct control over their data and the possibility to lodge a complaint with a Data protection authority. Transparency Guidelines (n 25) 5.

43 Lisa H Newton, Louis Hodges and Susan Keith, 'Accountability in the Professions: Accountability in Journalism' (2004) 19 *Journal of Mass Media Ethics* 166, 173 <<http://www.tandfonline.com/doi/abs/10.1080/08900523.2004.9679687>> accessed 23 July 2019.

44 Denis McQuail, *Media Accountability and Freedom of Publication* (Oxford University Press, Oxford 2003) 204. On limited media accountability Vaira Viķe-Freiberga and others, 'A Free and Pluralistic Media to Sustain European Democracy' (2013) <[https://ec.europa.eu/digital-single-market/sites/digital-agenda/files/HLG\\_Final\\_Report.pdf](https://ec.europa.eu/digital-single-market/sites/digital-agenda/files/HLG_Final_Report.pdf)> accessed 23 July 2019. JLH Bardoel and LSJ D'Haenens, 'Media Responsibility and Accountability. New Conceptualizations and Practices' (2004) 29 *Communications: The European Journal of Communication Research* 5 <<https://www.degruyter.com/view/j/comm.2004.29.issue-1/comm.2004.007/comm.2004.007.xml>> accessed 23 July 2019. However, the media aims to guard its editorial independence not only from the government, but also from too much pressure from the audience: Kasper Welbers and others, 'News Selection Criteria in the Digital Age: Professional Norms versus Online Audience Metrics' (2016) 17 *Journalism: Theory, Practice & Criticism* 1037 <<http://journals.sagepub.com/doi/10.1177/1464884915595474>> accessed 23 July 2019.

45 Richard Van Der Wurff and Klaus Schönbach, 'Between Profession and Audience: Codes of Conduct and Transparency as Quality Instruments for off-and Online Journalism' (2011) 12 *Journalism studies* 407. Ward (n 42) 48.



the media's success in upholding them. Secondly, the media can justify individual moral decisions in the editorial process.<sup>46</sup> This can be as simple as explaining why a source remains anonymous, or more extensive, like explaining why a media organization focuses its attention on certain topics. Transparency subsequently enables the audience to give input by, for example, allowing them to stop their support for a news organization if they are particularly dissatisfied with its editorial decisions or to give direct feedback on (the reasoning behind) editorial decisions. Article 22(3) GDPR takes a similar approach to user input on automated decision-making by giving users a right to withdraw their consent or contest a decision, but not an independent right to have it changed. While this may provide users with an insufficient level of protection in certain contexts, it dovetails nicely with the concept of media accountability.

## Agency

Transparency is also advanced as a way to ensure individuals are not subjected to algorithms they cannot influence and that only see them as an incomplete set of statistics, as opposed to full-fledged individuals. Mendoza and Bygrave identify the decrease in an individual's influence over the algorithms that impact him as a key concern that led to the inclusion of the precursor of the right to an explanation in the data protection directive.<sup>47</sup> Similarly, the Council of Europe's (CoE) recommendation on profiling, referred to in the Commission's proposal for Article 22 GDPR, notes the risk that profiling techniques place individuals in

predetermined categories or negatively affect their rights and dignity.<sup>48</sup> The right to an explanation's facilitation of the exercise of fundamental rights reflects the GDPR's role as an instrument that not only protects the right to data protection, but also enables the exercise of other rights.<sup>49</sup>

When automated decision-making is used to personalize the news a reader gets to see, the right that is put most immediately at risk is the right to receive information. There is a strong connection between this right and individual transparency in the media, as illustrated by in the influential 'rule of transparency' formulated by journalism scholars Kovach and Rosenstiel in 2001: 'What does my audience need to know to evaluate this information for itself? And is there anything in our treatment of it that requires explanation?'<sup>50</sup> The Audiovisual Media Services Directive (AVMSD) similarly emphasizes individuals must be able to know who provides the information that shapes their opinions, as such information is 'directly linked to the freedom of expression'.<sup>51</sup> Council of Europe recommendations also highlight the importance of providing individuals with information that allows them to access a wide variety of sources and analyse this information for themselves, including with regard to the automatic processes that regulate the flow of online information.<sup>52</sup>

This focus on transparency's ability to enable individuals to access and evaluate different ideas reflects the right to receive information's function as a mechanism that allows readers to develop themselves and find the truth by weighing up competing perspectives.<sup>53</sup> The extent to which every reader engages in this process is, of course, doubtful.<sup>54</sup> It is however impossible if they are

46 Van Der Wurff and Schönbach (n 45) 417.

47 Mendoza and Bygrave (n 13).

48 'Recommendation CM/Rec(2010)13 of the Committee of Ministers to Member States on the Protection of Individuals with Regard to Automatic Processing of Personal Data in the Context of Profiling' (2010) <[https://search.coe.int/cm/Pages/result\\_details.aspx?ObjectID=09000016805cdd00](https://search.coe.int/cm/Pages/result_details.aspx?ObjectID=09000016805cdd00)>.

49 See on this goal in relation to the right to an explanation Selbst and Powles (n 9); see more generally Manon Oostveen and Kristina Irion, 'The Golden Age of Personal Data: How to Regulate an Enabling Fundamental Right?' in M Bakhom and others (eds), *Personal Data in Competition, Consumer Protection and Intellectual Property Law* (Springer, Berlin 2018).

50 Bill Kovach and Tom Rosenstiel, *The Elements of Journalism: What Newspapers Should Know and the Public Should Expect* (Crown, New York 2001). See also Patrick Lee Plaisance, 'Transparency: An Assessment of the Kantian Roots of a Key Element in Media Ethics Practice' (2007) 22 *Journal of Mass Media Ethics* 187, 203.

51 The matter has received increased attention in the revision of the AVMSD, which directly ties ownership transparency to the exercise of the freedom of expression, and has a stronger focus on the disclosure of the identity of individuals who own the media organisation. Directive (EU) 2018/1808 of the European Parliament and of the Council of 14 November 2018 amending Directive 2010/13/EU on the coordination of certain provisions laid down by law, regulation or administrative action

in Member States concerning the provision of audiovisual media services (Audiovisual Media Services Directive) in view of changing market realities, OJ 2018 L 303/69 (AVMSD revision) recitals 15–16, art 5.

52 Council of Europe, 'Recommendation CM/Rec(2018)1 of the Committee of Ministers to Member States on Media Pluralism and Transparency of Media Ownership' (n 37); Council of Europe, 'Recommendation CM/Rec(2012)3 of the Committee of Ministers to Member States on the Protection of Human Rights with Regard to Search Engines' (2012) <[https://search.coe.int/cm/Pages/result\\_details.aspx?ObjectID=09000016805caa87](https://search.coe.int/cm/Pages/result_details.aspx?ObjectID=09000016805caa87)> accessed 23 July 2019. Council of Europe, 'Recommendation CM/Rec(2007)2 of the Committee of Ministers to Member States on Media Pluralism and Diversity of Media Content' (2007) para III.2 <[https://search.coe.int/cm/Pages/result\\_details.aspx?ObjectID=09000016805d6be3](https://search.coe.int/cm/Pages/result_details.aspx?ObjectID=09000016805d6be3)> accessed 23 July 2019.

53 Sarah Eskens, Natali Helberger and Judith Moeller, 'Challenged by News Personalisation: Five Perspectives on the Right to Receive Information' (2017) 9 *Journal of Media Law* 259 <<https://www.tandfonline.com/doi/full/10.1080/17577632.2017.1387353>> accessed 23 July 2019. Brent Daniel Mittelstadt, 'Automation, Algorithms, and Politics| Auditing for Transparency in Content Personalization Systems' (2016) 10 *International Journal of Communication* 12, 4994.

54 Michael Karlsson and Christer Clerwall, 'Transparency to the Rescue?' (2018) 19 *Journalism Studies* 1923 <<https://www.tandfonline.com/doi/full/10.1080/1461670X.2018.1492882>> accessed 23 July 2019.

only given the bare facts of a story, or when the only information they have about the inner workings of an algorithm is that it aims to present relevant content. Relevance, after all, can mean many different things depending on for whom and for what purpose it is defined.<sup>55</sup> Further information concerning the sources, the methods of reporting, and potential biases of the publication is required to allow the audience to form an opinion on whether the content is produced in accordance with journalistic ethics and whether they can rely on it.<sup>56</sup>

## Trust

Finally, transparency is argued to increase trust in algorithmic decision-making.<sup>57</sup> Trust allows individuals to rely on others in situations where they are unsure whether the other party will exploit their vulnerabilities. As such, it is rather useful for individuals who aim to use algorithms without fully understanding them. Conversely, being perceived as untrustworthy can diminish an algorithm's success regardless of its actual accuracy or compliance with legal and ethical principles. While trust is a commonly featured goal in both discussions of algorithmic transparency as well as the GDPR and the EU's digital market strategy, it has thus far received limited attention in the context of the right to an explanation.<sup>58</sup>

Trust in the media is typically defined in relation to the media's function. Grosser, for example, defines it as 'the willingness of the recipient to be vulnerable to the journalistic system's selection and communication of current information'.<sup>59</sup> Transparency is expected to lower the barrier for trust by allowing users to more accurately assess the likelihood that their vulnerability will

be exploited. This allows for the repeated interactions through which long-term trust may be built.<sup>60</sup> There is some overlap in the role transparency plays in trust and enabling users to exercise their right to receive information. In both fields, transparency is used to allow users to form a more accurate judgment of the other party and the information they are provided with. There are two important differences, however. Firstly, the right to receive information is concerned with allowing individuals to assess their news diet in a broad sense. Conversely, a single news organization (such as a trade magazine) can still be perceived as trustworthy even if it only covers certain topics.

Secondly, trust is a psychological process, and therefore functions more as a heuristic than the exercise of the right to receive information. As a result, trust in the media may be affected by factors unrelated to the media's ability to correctly communicate current information. A particularly salient issue in the context of news personalization is the role of privacy. Research indicates tracking can negatively impact trust, especially if it violates consumer's privacy norms and reminds them of this fact with highly personalized cues.<sup>61</sup> While privacy concerns typically do not translate directly into a lower use of the service,<sup>62</sup> they may have an indirect impact on the media by spilling over into trust in the way the media organization informs its readers. Do readers trust an organization to tell them the truth while simultaneously finding the organization creepy for tracking every article they read? This is an empirical question that appears to be unanswered. The fact that privacy violations affect an organization's perceived integrity indicates this might be the case however; being perceived as lacking integrity reflects on the

55 Tarleton Gillespie, 'The Relevance of Algorithms' in Tarleton Gillespie, Pablo J Boczkowski and Kirsten A Foot (eds), *Media technologies: Essays on communication, materiality, and society* (MIT Press, Cambridge 2014).

56 Konstantin Nicholas Dörner and Katharina Hollnbuchner, 'Ethical Challenges of Algorithmic Journalism' (2017) 5 *Digital Journalism* 404. Turilli and Floridi (n 5); Van Der Wurff and Schönbach (n 45) 417. L Hitchens, 'Commercial Content and Its Relationship to Media Content: Commodification and Trust' in Monroe E Price and Libby Verhulst, Stefaan G Morgan (eds), *Routledge handbook of media law* (Routledge, Abingdon 2013) 94 <<https://www.routledge.com/Routledge-Handbook-of-Media-Law/Price-Verhulst-Morgan/p/book/9780415683166>> accessed 23 July 2019.

57 Brent Daniel Mittelstadt and others, 'The Ethics of Algorithms: Mapping the Debate' (2016) 3 *Big Data & Society* 1; Bruno Lepri and others, 'Fair, Transparent, and Accountable Algorithmic Decision-Making Processes' [2017] *Philosophy & Technology* 1. Marco Tulio Ribeiro, Sameer Singh and Carlos Guestrin, "'Why Should I Trust You?': Explaining the Predictions of Any Classifier" [2016] *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 1135 <<http://arxiv.org/abs/1602.04938>> accessed 23 July 2019.

58 David Wright and others, 'Privacy, Trust and Policy-Making: Challenges and Responses' (2009) 25 *Computer Law & Security Review* 69 <<https://www.sciencedirect.com/science/article/pii/S0267364908001672>> accessed

23 July 2019. Neil M Richards and Woodrow Hartzog, 'Taking Trust Seriously in Privacy Law' (2016) 19 *Stanford Technology Law Review* 431 <<https://www-cdn.law.stanford.edu/wp-content/uploads/2017/11/Taking-Trust-Seriously-in-Privacy-Law.pdf>> accessed 23 July 2019. See for a discussion on trust and the right to an explanation Kim (n 22).

59 Grosser (n 32) 1040. see similarly Kohring and Matthes (n 32) 240.

60 See also Kim (n 22).

61 See on highly personalised cues' impact on feelings of vulnerability Elizabeth Aguirre and others, 'Unraveling the Personalization Paradox: The Effect of Information Collection and Trust-Building Strategies on Online Advertisement Effectiveness' (2015) 91 *Journal of Retailing* 34, 41 <<https://www.sciencedirect.com/science/article/pii/S0022435914000669>> accessed 23 July 2019.

62 Spyros Kokolakis, 'Privacy Attitudes and Privacy Behaviour: A Review of Current Research on the Privacy Paradox Phenomenon' (2017) 64 *Computers & Security* 122 <<https://www.sciencedirect.com/science/article/pii/S0167404815001017>> accessed 23 July 2019. Tom Evens and Kristin Van Damme, 'Consumers' Willingness to Share Personal Data: Implications for Newspapers' Business Models' (2016) 18 *International Journal on Media Management* 25 <<http://www.tandfonline.com/doi/full/10.1080/14241277.2016.1166429>> accessed 23 July 2019.

trustworthiness of an actor more broadly than being perceived to lack a certain ability, which is more restricted to a single area in which another party cannot perform adequately.<sup>63</sup>

Of course, personalization may also affect trust in the media directly due to the changes in the way the media selects what information to show to readers. Fears over filter bubbles remain prevalent and could decrease trust in personalizing media organizations, regardless of whether such fears are justified. Conversely, personalization could also improve the media's ability to inform readers by allowing it to tailor the information provided to each individual's prior knowledge.<sup>64</sup> There is indeed some evidence that users can prefer algorithms to human editors.<sup>65</sup>

It should be noted that from a policy perspective, more trust is not necessarily better. Rather, what is important is that individuals are able to correctly assess another party's trustworthiness.<sup>66</sup> Transparency can play a role in this process by correcting the uncertainties or misconceptions that prevent users from trusting parties they otherwise would trust (and vice versa). Indeed, certain kinds of information, such as how the system arrives at a recommendation or why a recommendation was made, have been shown to improve trust in recommender systems.<sup>67</sup> In line with policy desires, this effect appears to be predicated on users seeing the system as more competent after learning more about it.<sup>68</sup> The available research on the effects of transparency on trust in recommender systems more typically concerns eCommerce than news recommenders, however.<sup>69</sup> As a result, transparency may affect trust in news recommenders in different and as of yet unknown ways.

This uncertain relationship between transparency and the goals of the right to an explanation is inherent in the use of emerging technologies in new contexts. However, it is not necessary to develop the disclosures that are suitable to safeguard the goals of the right to an explanation in such contexts from scratch. Literature on media transparency already contains a wealth of research regarding the individual transparency mechanisms that are suitable to further the goals explored above. The following section will adapt this research to the automated decision-making processes covered by the GDPR.

## A conceptual framework of algorithmic transparency in the media

### Current approaches to media and algorithmic transparency frameworks

Transparency covers a wide range of information that could be used to safeguard accountability, agency, and trust. Literature on media transparency law and ethics categorizes this information in largely similar ways. Groenhart and Evers, for example, categorize traditional media transparency instruments into actor, source, and process transparency. Actor transparency refers to the parties who are able to influence editorial decisions. Source transparency covers the parties who provide the information on which a story is based, as well as the source material itself. Process transparency concerns the editorial process, including the mechanisms and justifications of editorial decisions.<sup>70</sup>

63 Bansal and Zahedi (n 33).

64 Judith Möller and others, 'Do Not Blame It on the Algorithm: An Empirical Assessment of Multiple Recommender Systems and Their Impact on Content Diversity' (2018) 21 *Information, Communication & Society* 959.

65 Nic Newman and others, 'Reuters Institute Digital News Report 2017' (2017) 15 <[https://reutersinstitute.politics.ox.ac.uk/sites/default/files/Digital News Report 2017\\_web\\_0.pdf](https://reutersinstitute.politics.ox.ac.uk/sites/default/files/Digital%20News%20Report%202017_web_0.pdf)> accessed 23 July 2019.

66 Frank B Cross, 'Law and Trust' (2004) 93 *Georgetown Law Journal* 1457 <<https://heinonline.org/HOL/P?h=hein.journals/glj93&i=1471>> accessed 23 July 2019.

67 Peter A Hancock and others, 'A Meta-Analysis of Factors Affecting Trust in Human-Robot Interaction' (2011) 53 *Human Factors: The Journal of the Human Factors and Ergonomics Society* 517 <<http://journals.sagepub.com/doi/10.1177/0018720811417254>> accessed 23 July 2019.

68 Henriette Cramer and others, 'The Effects of Transparency on Trust in and Acceptance of a Content-Based Art Recommender' (2008) 18 *User Modeling and User-Adapted Interaction* 455. Weiquan Wang and Izak Benbasat, 'Recommendation Agents for Electronic Commerce: Effects of Explanation Facilities on Trusting Beliefs' (2007) 23 *Journal of Management Information Systems* 217. Fatih Gedikli, Dietmar Jannach and Mouzhi Ge, 'How Should I Explain? A Comparison of Different Explanation Types for Recommender Systems' (2014) 72 *International Journal of Human Computer Studies* 367 <<http://linkinghub.elsevier.com/retrieve/pii/S1071581913002024>> accessed 23 July 2019.

69 See an exception M ter Hoeve and others, 'Do News Consumers Want Explanations for Personalized News Rankings', *FATREC Workshop on Responsible Recommendation Proceedings* (2017) 39 <<https://pdfs.semanticscholar.org/6563/457311cbaa14ebb9ff438279bde918292e64.pdf>> accessed 23 July 2019. The study found newsreaders wanted explanations of individual decisions, but found no effect on click-through rates; the authors speculate this is because users did not pay attention to the explanations.

70 Some categorizations omit source transparency or include a category for reader interaction entitled interactive, dialogue, or participatory transparency. The model of Groenhart & Evers that omits this category is better suited to the right to an explanation, as interactive transparency does not cover disclosures but instead concerns mechanisms that allow readers to act on information they have learned. H Groenhart and H Evers, 'Media Accountability and Transparency – What Newsrooms (Could) Do' in S Fengler and others (eds), *Journalists and media accountability: An international study of news people in the digital age* (Peter Lang, New York 2014).

See for other (largely similar) frameworks Klaus Meier and Julius Reimer, 'Transparenz Im Journalismus' (2011) 56 *Publizistik* 133 <<http://link.springer.com/10.1007/s11616-011-0116-7>> accessed 23 July 2019; David Domingo and Heikki Heikkilä, 'Media Accountability Practices in Online News Media' in *The Handbook of Global Online Journalism* (Wiley-Blackwell, Chichester 2012) <<http://doi.wiley.com/10.1002/9781118313978.ch15>> accessed 23 July 2019. Michael Karlsson, 'Rituals of Transparency' (2010) 11 *Journalism Studies* 535 <<http://www.tandfonline.com/doi/abs/10.1080/14616701003638400>> accessed

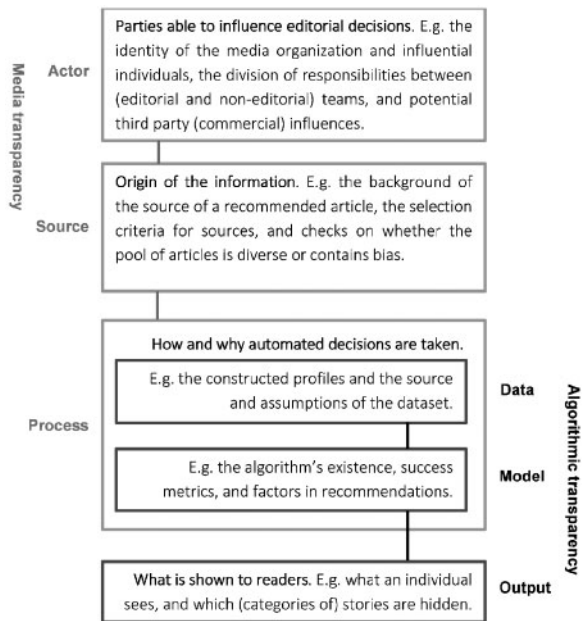


Figure 1. An analytical framework of the disclosures that make up algorithmic transparency in the media.

Categorizations of algorithmic transparency commonly centre on the data that is used in decision-making, how this data is processed, and the output.<sup>71</sup> Integrating the two frameworks produces the model for algorithmic media transparency shown in Figure 1. On the one hand, media transparency contains aspects that are affected by algorithms, but that are not traditionally part of the discussion on the right to an explanation. On the other hand, algorithms shape the way traditional editorial processes are carried out, or introduce new aspects. The following section will expand on

these interactions between algorithmic and media transparency and the resulting required disclosures in the context of the right to an explanation.

## Actor transparency

Actor transparency instruments inform the audience about those inside and outside the organization who may be able to influence editorial decisions. On the organizational level, it involves disclosure of not only the identity of the company, but also information regarding its ownership, and parties that have a financial interest in the organization.<sup>72</sup> In the media, such information is expected to allow users to access diverse sources and make more informed decisions about the information they are provided with.<sup>73</sup> Similarly, the law requires that users are informed when third parties pay for editorial content to promote certain products. Additionally, self-regulatory instruments such as ethics codes inform users of the ethical standards media organizations hold themselves to, and enable criticism when they fail to live up to these standards.<sup>74</sup> Transparency can also provide readers with a more fine-grained view of media organizations. The public may, for example, be informed of the division of responsibilities within an organization or the background of those who contribute to a specific story or can exercise significant influence over editorial choices.<sup>75</sup> Such disclosures can conflict with the freedom of expression, as anonymity can promote the free flow of ideas by protecting authors from personal retaliation.<sup>76</sup> While the large media organizations that have the resources to implement news personalization rarely rely on such anonymity and have instead increasingly

23 July 2019; HP Groenhart and J LH Bardoel, 'Conceiving the Transparency of Journalism: Moving towards a New Media Accountability Currency' (2012) 12 *Studies in Communication Sciences* 6 <<https://www.sciencedirect.com/science/article/pii/S1424489612000045>> accessed 23 July 2019.

71 Analysed in most concrete terms by Diakopoulos and Koliska (n 3) 9. See also Zarsky (n 40) 1521. Danielle Keats Citron and Frank Pasquale, 'The Scored Society: Due Process for Automated Predictions' (2014) 89 *Washington Law Review* 1, 20 <<https://heinonline.org/HOL/P?h=hein.journals/washlr89&i=8>> accessed 23 July 2019.

72 Domingo and Heikkilä (n 70). Groenhart and Evers (n 70).

73 It is a staple of Recommendations by the Council of Europe, eg Council of Europe, 'Recommendation No. R (94) 13 of the Committee of Ministers to Member States on Measures to Promote Media Transparency' (1994) <<https://rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?documentId=09000016804c1bdf>>. Council of Europe, 'Recommendation CM/Rec(2018)1 of the Committee of Ministers to Member States on Media Pluralism and Transparency of Media Ownership' (n 37). See also in the EU context Council of the European Union, 'Conclusions of the Council and of the Representatives of the Governments of the Member States, Meeting within the Council, on Media Freedom and Pluralism in the Digital Environment (2014/C 32/04)' (2014) para 13 <<https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52014XG0204%2802%29>> accessed 23 July 2019. In national law, it typically results from non-media law sources Izabela Korbiel and Katharine Sarikakis, 'Between Two (in)Competencies: A Critical View on Media Ownership Regulation in the EU' (2017) 13 *MCP Intellect Limited International Journal of Media & Cultural Politics* 183 <[http://www.ingentaconnect.com/content/10.1386/macp.13.1-2.183\\_7](http://www.ingentaconnect.com/content/10.1386/macp.13.1-2.183_7)> accessed 23 July 2019.

74 Groenhart and Bardoel (n 70) 8. Domingo and Heikkilä (n 70) 278.

75 Arthur S Hayes, Jane B Singer and Jerry Ceppos, 'Shifting Roles, Enduring Values: The Credible Journalist in a Digital Age' (2007) 22 *Journal of Mass Media Ethics* 262. Ward (n 42) 51. H Heikkilä and others, 'Innovations in Media Accountability and Transparency' in S Fengler and others (eds), *Journalists and Media Accountability: An International Study of News People in the Digital Age* (Peter Lang, New York 2014) 58. Groenhart and Bardoel (n 70) 8. Council of Europe, 'Recommendation CM/Rec(2007)2 of the Committee of Ministers to Member States on Media Pluralism and Diversity of Media Content' (n 52) para III. Such information also allows other professionals in the media system to hold each other accountable (Bardoel and D'Haenens (n 44)); such accountability falls outside the scope of the right to an explanation however, as it focuses on allowing individuals to hold organizations accountable

76 *Delfi v Estonia* app no 64569/09 (EctHR, 16 June 2015) para 147.

embraced transparency, the possibility for exceptions must be maintained.

The implementation of news personalization often requires organizational changes that affect this division of responsibilities. While writing and selecting articles are editorial tasks, journalists and editors typically lack the funding and expertise necessary to build the personalization algorithms that automate such tasks. As a result, personalization creates new roles within an organization and requires collaboration between editors, engineers, and publishers.<sup>77</sup> This can change or obscure who is able to exert influence on (and has final responsibility for) the editorial decisions in news personalization. Disclosing any such changes prevents users from (potentially mistakenly) assuming that the editorial staff remains fully in control of the editorial decisions in news personalization. The New York Times, for example, outlines the role of its editorial staff in the design of its news personalization systems. When explaining how its recommendation process works, the organization emphasizes that the editorial team advises engineers how personalization fits into the Times' editorial mission, determines what news content is to be used for personalization, and incorporates feedback from readers.<sup>78</sup>

Personalization can also be influenced by outside parties. The clearest example of such a case is when a media organization rents out space on its site to third party algorithms that recommend third party content, like those of Outbrain or Taboola. This kind of news personalization comes close to online behavioural advertising and its transparency requirements regarding the commercial nature of the provided information, and typically carries a 'sponsored by' label. However, intermediate versions of outside influence are also possible. A company might allow the marketing department influence over the design of a recommender system, or might simply use (a modified version of) a third-party algorithm to recommend its articles. None of these situations are inherently problematic. However, unless the third party and any payments are disclosed, users are confronted with a third-party system that takes editorial

decisions they would logically assume are taken by the media organization in line with its editorial standards.<sup>79</sup>

Finally, personalization affects the values an actor holds himself accountable for. Not in the sense that values like newsworthiness or diversity lose their relevance, but by changing the way in which they can be furthered or safeguarded. Ethics codes, mission statements, and policy documents have traditionally provided insight into how the media and individual organizations believe this should occur. As more editorial tasks are automated, it will become increasingly important that the media also provides insight into the ethical standards it believes should guide the use of algorithms. Failing to clarify such norms would hamper accountability in light of the media's relative freedom to regulate its own behaviour, as it is difficult to hold a media organization accountable to a non-existent standard. Conversely, outlining the values an actor holds himself to allow users to object in a way that is most likely to resonate with the media organization.

### Source transparency

Source transparency traditionally concerns (the interests of) the sources that supplied the information on which an article is based. When the digitalization of the news removed space constraints, simply linking to the full source material directly also became more common.<sup>80</sup> Such disclosures allow readers to take the sources' potential interests into account, check whether the media selected all the relevant facts, or simply investigate further to satisfy their curiosity.<sup>81</sup>

News personalization makes use of some third-party sources of information. Traditional media organizations increasingly draw on information and statistics provided by other parties to automatically generate (personalized) niche articles, and intermediaries use the articles created by traditional media organizations to fill their personalized news feeds.<sup>82</sup> The controversies surrounding fake news have prompted calls for users to be better informed regarding the identity and intentions of the source behind articles recommended to them on platforms. In response, many recent projects research what information is necessary to ensure readers do not

77 Balázs Bodó, 'Means, Not an End (of the World) The Customization of News Personalization by European News Media' [2018] SSRN <<https://www.ssrn.com/abstract=3141810>> accessed 23 July 2019. Matthew S Weber and Allie Kosterich, 'Coding the News' (2018) 6 *Digital Journalism* 310, 323.

78 New York Times, 'Personalization' <<https://help.nytimes.com/hc/en-us/articles/360003965994-Personalization>> accessed 23 July 2019.

79 See similarly in the context of transparency in the algorithms that rank search results Council of Europe, 'Recommendation CM/Rec(2012)3 of the Committee of Ministers to Member States on the Protection of Human Rights with Regard to Search Engines' (n 52).

80 Of course, where the source material is another article this is also a matter of giving credit where it is due; Hayes, Singer and Ceppos (n 75). Heikkilä and others (n 75) 59.

81 Society of Professional Journalists, 'SPJ Code of Ethics' (2014) <<https://www.spj.org/ethicscode.asp>> accessed 23 July 2019.

82 Tal Montal and Zvi Reich, 'I, Robot. You, Journalist. Who Is the Author?' (2017) 5 *Digital Journalism* 829 <<https://www.tandfonline.com/doi/full/10.1080/21670811.2016.1209083>> accessed 23 July 2019.

trust information from inaccurate sources. Preliminary results indicate that users quickly forget the source of the information they have learned; current projects aim to be more direct in providing third party labels which indicate whether a story is trustworthy.<sup>83</sup> While this may prove to be more effective, government regulation that establishes criteria by which particular sources are labelled as (un)trustworthy could cause significant issues related to freedom of expression. Simply disclosing the factors by which an algorithm assesses a particular source and determines whether (or how much) it will be recommended may prove to be a less problematic, albeit also less effective proxy. It also avoids the potential conflicts with the media's right to the confidentiality of their sources that a blanket requirement to disclose the source of information might cause. This right is vital to the media's continued ability to act as a public watchdog.<sup>84</sup>

In addition to information on individual sources and articles, information on the aggregate pool of information news personalization can prove useful, as these ingredients determine what news personalization are and are not able to do for their users. If the available articles only reflect one viewpoint, for example, the reader will not be able to get a diverse news diet even if he is able to instruct the algorithm to provide it. Transparency regarding the types of articles the personalization algorithm is able to show readers therefore gives them the opportunity to determine how using it will fit into their news diet. Concretely, this could require that media organizations outline the steps they take (if any) to ensure that the information news personalization draws on fits into their editorial mission. For example, organizations could disclose the existence of any monitoring of the pool of articles itself (eg for bias or diversity), ex ante or ex post selection criteria for potential contributors, or assurances that certain particularly important content will not be personalized. The latter would allow readers to use personalization in the knowledge that they are not at risk of missing out on particularly important information.

Finally, content personalization makes it possible that each reader is presented with a different version of the same story. This touches on a central component of trust in the media, namely their ability to select the relevant facts of a story to communicate to readers. One

way media organizations have addressed the suspicions of their readers with regard to this issue in the past is by including links to the full source material on which an article is based, and disclosing sources' interests. Similarly, readers of a personalized story could be provided with a way to access a version that contains all the information it could have used, and provide reasons why particular pieces of information were omitted in the version initially provided to them.

### Process transparency

Process transparency concerns the disclosure, explanation, and justification of editorial choices, both in relation to broader issues like topic selection as well as small-scale decisions in specific stories.<sup>85</sup> Several instruments are used to enable this kind of transparency. Newsroom blogs or livestreamed editorial meetings can give individuals a detailed account of the reasons behind broader and particularly significant editorial decisions. Conversely, smaller editorial decisions (like the anonymization of a source) are increasingly woven into the news story itself.

Much of the discussion on algorithmic transparency fits into this category. When media organizations implement news personalization, they use information about readers to automatically take editorial decisions regarding the selection, arrangement, and production of news content. The simplest disclosure, and the one which is seemingly most common in privacy notices, is that this process is taking place at all. Informing users of the existence of automated-decision-making is of course also required explicitly by Articles 13–15. The common phrasing that a news site contains personalized content is too broad to be of any use, however. Readers can only start to determine how news personalization affects them if they know which parts of the site are personalized.<sup>86</sup> Analogously, disclosing that the newspaper includes advertisements or opinion pieces does not help readers if the opinion section, advertisements, and news stories themselves are indistinguishable.

Disclosing which of a particular user's characteristics the algorithm takes into account (and how they are weighted) would address the risk that motivated the Commission to regulate automated processing in the original data protection directive, namely that decisions will be taken on the basis of users' 'data shadow'

83 See, for example, the indicators developed by Santa Clara University's Markkula Center for Applied Ethics, 'The Trust Project' (2018) <<https://thetrustproject.org/>> accessed 23 July 2019. See also European Commission, 'Final Report of the High Level Expert Group on Fake News and Online Disinformation' (2018) 23 <<https://ec.europa.eu/digital-single-market/en/news/final-report-high-level-expert-group-fake-news-and-online-disinformation>> accessed 23 July 2019.

84 Broy and others (n 35).

85 Groenhart and Evers (n 70). Meier and Reimer (n 70). Ward (n 42). Society of Professional Journalists (n 81).

86 However, labelling algorithmically produced content is not yet the norm: Montal and Reich (n 82).

without their knowledge. It also lets readers know how their actions influence the algorithm, and what they could read to shape their news diet differently. However, in light of the number of recommendations made by news personalization algorithms, explaining individual decisions is unlikely to be an effective way of advancing these goals. Providing more general information about the way the algorithm converts an individual's data into recommendations would save users from having to build an understanding of these matters from explanations of individual decisions. Dashboards that allow users to see and adjust what data is used could allow them to easily put this understanding into practice.

Counterfactual explanations face similar context-specific barriers.<sup>87</sup> News recommendations are not binary decisions; there are many other articles that could have been recommended, and each alternative recommendation requires the user to provide different input data. Of course, a user can be informed of the changes he can make to ensure a specific type of article will not be recommended in the future. Doing so in the context of news personalization is problematic, however, as the changes a user can make will often concern his reading behaviour. Instructing him to refrain from reading certain articles creates the risk of chilling effects, especially if it leads readers to fear they will be inundated with articles on a topic they are not generally interested in after clicking on a headline that sparked their curiosity.

Explaining the ethical basis of editorial decisions is an important aspect of media transparency, both to ensure the accountability of media organizations and to promote trust and understanding on the part of the audience. The extent to which editorial values are reflected in the success metrics of a news personalization algorithm therefore appears particularly important. While the role of editorial values in personalization algorithms often remains unclear (this is part of the problem, after all), for some organizations they do not appear to play a role at all. For example, DeVito's analysis of patents, press releases, and SEC filings indicates that Facebook's news curation is based on criteria that bear little resemblance to traditional news values, such as a user's social relationships or preferences.<sup>88</sup> Legacy European media

organizations also use personalization to promote different values, ranging from the commercial (eg to maximize the time spent on a site) to the editorial (eg to better deliver news to niche audiences).<sup>89</sup>

Personalization algorithms can, in short, promote a wide range of objectives, not all of which resemble traditional news values and each of which will shape a reader's news diet in a different way. Transparency about the specific objective(s) of a personalization algorithm would allow users to distinguish between editorial and non-editorial applications of news personalization. It would also correct some misconceptions on the basis of which readers currently evaluate news personalization. For example, although the research above shows algorithms are used to advance concrete objectives, the public still often sees algorithms as inherently objective or neutral.<sup>90</sup> In contrast, for news organizations that use personalization to advance traditional media ethics but whose readers remain suspicious of the new technology, transparency about the values built into their algorithms could distinguish them from more commercial organizations like Facebook.

## Output transparency

Media transparency typically does not focus on information in the post-publication phase. This is unsurprising, since what is published offline is transparent by default; if this information fails to reach users this is a circulation problem, not a transparency problem. Online, however, stories can be changed post-publication to include new information or correct errors. Disclosing such changes is an increasingly common way to ensure transparency even post-publication.<sup>91</sup>

This need for post-publication transparency is magnified by personalization. When a media organization personalizes its output, it inherently affects each individual in a unique way. A personalizing news provider may be able to give in depth, niche stories to a reader interested in government corruption, but not to one interested in climate change. Individuals' (inferred) preferences may also expose them to a different and less diverse collection of stories.<sup>92</sup> In short, a single personalization algorithm may pose different risks to different individuals. Moreover,

87 See on the usefulness of counterfactual explanations generally Wachter, Mittelstadt and Russell (n 38).

88 Michael A DeVito, 'From Editors to Algorithms' (2017) 5 *Digital Journalism* 753 <<https://www.tandfonline.com/doi/full/10.1080/21670811.2016.1178592>> accessed 23 July 2019; Weber and Kosterich (n 77) 322.

89 Bodó (n 77).

90 This danger was already flagged during the drafting of the provisions on automated decision-making in the DPD, Mendoza and Bygrave (n 13)

84. Min Kyung Lee, 'Understanding Perception of Algorithmic Decisions: Fairness, Trust, and Emotion in Response to Algorithmic

Management' (2018) 5 *Big Data & Society* 1 <<http://journals.sagepub.com/doi/10.1177/2053951718756684>>.

91 Kalyani Chadha and Michael Koliska, 'Newsrooms and Transparency in the Digital Age' (2015) 9 *Journalism Practice* 215. Karlsson (n 70).

92 Eytan Bakshy, Solomon Messing and LA Adamic, 'Exposure to Ideologically Diverse News and Opinion on Facebook' (2015) 348 *Science* 1130. See conversely, with regard to Google News Mario Haim, Andreas Graefe and Hans-Bernd Brosius, 'Burst of the Filter Bubble?' (2018) 6 *Digital Journalism* 330 <<https://doi.org/10.1080/21670811.2017.1338145>> accessed 23 July 2019.

because the same news provider shows each of its readers a unique collection of information, they cannot draw on the knowledge of others (ie a news providers' reputation) to assess the risks a particular personalization algorithm. This cuts off an avenue of information users typically rely on to evaluate media organizations, creating problems for even the most media-literate newsreaders.

Different types of information can be used to address this problem. The extent to which these are required by the right to an explanation is problematized by the future facing language of 'envisaged consequences' and the need not to interpret 'suitable measures' in an overly extensive way. The narrowest interpretation is that users must be informed of the basic consequence of personalization, namely that certain (types of) information will be filtered out. This is inherent in personalization, however, and as a result does not provide readers with any information they do not already gain from the disclosure of the existence of profiling. To be of added value, the information provided must therefore at least give some indication as to which stories will likely be filtered out, or how the algorithm will perform. As a concrete example, news organizations could inform readers that their algorithm has been known to steer its users towards more extreme content, or show that their algorithm outperforms human curation in terms of diversity.<sup>93</sup>

Both Articles 13–15 and 22 also mean to give the data subject information as to how his rights and interests are at risk, however.<sup>94</sup> Information about an algorithm's impact on the general population is only useful to further this goal to the extent that it overlaps with the impact on the data subject. To get to the heart of the problem, information on the consequences of personalization would have to be personalized much like the proposed explanations of specific decisions would have to be. Such transparency could be provided in two ways. The first is to inform a reader directly of the information that has been filtered out of his news diet. Tintarev, Rostami and Smyth, for example, have developed a method to visualize users' profiles so as to allow them to identify blind spots in their consumption of the available content.<sup>95</sup> Alternatively, readers could be informed of the most important stories they may have missed, or be provided with access to a non-personalized version of the site. Providing such information would also dovetail with CoE recommendations that aim to promote the findability of content.<sup>96</sup>

The second is to disclose the differences between a user's personalized news diet and his diet prior to personalization (or, if this is not available, that of the average or comparable reader) into account. This way, personalization's impact on, for instance, the diversity of a reader's personal news diet could be revealed to him. For sites that personalize the news in a way that is radically different from their readers believe it to be, such disclosures might be especially important if they draw their attention to unknown negative or positive effects of the algorithm.

## Conclusion

This article has argued that the concrete disclosures required by the right to an explanation should be identified by analysing the context of the algorithm that is to be explained. Such an approach matches the specific disclosures that make up the right to an explanation to its goals. It also promotes a more consistent legal approach to explainability that takes sector-specific tensions between the right to an explanation and other values, in the case of news personalization the freedom of expression, into account. To operationalize the context-specific approach, the article identified the limits of the right to an explanation, drew on insights from media law and literature to determine how its goals should be understood in the context of news personalization, and outlined which concrete disclosures are suitable to advance these goals in this context.

Stronger collaboration between supervisory and regulatory authorities would facilitate the operationalization of this approach and prevent the emergence of conflicting or overlapping approaches to algorithmic transparency.<sup>97</sup> Data protection authorities are well-placed to create an approach to algorithms that is consistent and comprehensive on a general level, but do not need to recreate existing wisdom on the risks that are likely to be triggered in specific contexts. Sector-specific supervisors and regulators already possess much of the expertise necessary to understand how values within their fields can be put at risk, and how individual transparency can be used to mitigate such risks. Drawing on this expertise promotes efficiency and consistency.

Of course, not all disclosures discussed in the article are relevant for each personalization algorithm, for the simple reason that not all personalization algorithms are

93 Derek O'Callaghan and others, 'Down the (White) Rabbit Hole The Extreme Right and Online Recommender Systems' (2015) 33 *Social Science Computer Review* 459. Möller and others (n 64).

94 Automated Decision-Making Guidelines (n 13) 20, 26.

95 Nava Tintarev, Shahin Rostami and Barry Smyth, 'Knowing the Unknown: Visualising Consumption Blind-Spots in Recommender Systems' [2018] *Proceedings of the 33rd Annual ACM Symposium on Applied Computing - SAC '18* 1396.

96 Council of Europe, 'Recommendation CM/Rec(2018)1 of the Committee of Ministers to Member States on Media Pluralism and Transparency of Media Ownership' (n 37) para 2.3.

97 European Data Protection Supervisor, 'EDPS Opinion 3/2018 on Online Manipulation and Personal Data' (2018) 3 <[https://edps.europa.eu/sites/edp/files/publication/18-03-19\\_online\\_manipulation\\_en.pdf](https://edps.europa.eu/sites/edp/files/publication/18-03-19_online_manipulation_en.pdf)> accessed 23 July 2019.



subject to the same risks. Not all personalization algorithms are influenced by third parties, for example. However, in light of their importance in the media particular importance should be afforded to the role of editorial values in personalization algorithms. Additionally, evidence about the way news personalization affects a particular reader could prove to be an especially useful

way to promote trust, accountability, and agency. It directly provides data subjects with information they would otherwise have to infer from other disclosures, and answers the question they are likely interested in the most: how will using this algorithm affect me?

*doi:10.1093/idpl/ipz011*

*Advance Access Publication 7 August 2019*