



CENTRE FOR IT & IP LAW

Kenniscentrum
Data & Maatschappij



News Recommenders and Cooperative Explainability Confronting the contextual complexity in AI explanations¹

M.Z. van Drunen, PhD researcher, Institute for Information Law, University of Amsterdam

J. Ausloos, Postdoctoral researcher, Institute for Information Law, University of Amsterdam

N.M.I.D. Appelman, PhD researcher, Institute for Information Law, University of Amsterdam

N. Helberger, Distinguished University Professor of Law and Digital Technology, University
of Amsterdam

¹ The views expressed in this paper are those of the authors and do not necessarily reflect the views or policies of the Knowledge Center Data & Society or CiTiP. The paper aims to contribute to the existing debate.

1. Introduction

Artificial Intelligence (AI) needs to be explainable. This is a key objective advanced by the [European Commission](#) (and its [high-level expert group](#)) throughout its AI policy, the [Council of Europe](#), and a rapidly growing body of academic scholarship in different disciplines, from [Computer Sciences](#) to [Communication Sciences](#) and [Law](#). This interest in explainability is in part fuelled by pragmatic concerns that some form of understanding is necessary for AI's uptake (and therefore economic success). But on a more fundamental level, there is a recognition that explainability is necessary to understand and manage the societal shifts AI triggers, and to ensure the continued agency of the individuals, market actors, regulators, and societies confronted with AI.

However, what does it mean for AI to be explainable? In this vision paper, we argue that the answers to this question must take better account of explainability's contextual and normative complexity. For our purposes here, and without discarding [the ample scholarly debate](#) around this notion, we understand explainability as "meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing" (as can be derived from a combined reading of Articles 13(2)f, 14(2)g, 15 (1)(h), Art. 22 and Recital 71 in the General Data Protection Regulation (GDPR)). As will become apparent in the following pages, explainability is relied on to perform a wide variety of functions with regard to a wide variety of actors involved in, and affected by, AI. This complexity is reflected in the patchwork of sector-specific legal frameworks, policy recommendations, ethics guidelines and self-regulatory instruments governing all kinds of decision-making processes now grouped under the common 'AI' denominator. Given these considerations we argue that approaches that focus on specific AI explanations or treat explainable AI as a general, abstract concept, however, cannot fully address its inherent complexity.

That complexity is lost when explainable AI is treated as an abstract, general concept that is treated generically and without attention for the intrinsics of a particular sector. What is needed instead, is to acknowledge this complexity head on. This requires us to understand normative demands and restrictions on explainability in specific contexts, the different capacities and responsibilities of the various stakeholders within these contexts, and how these can and should relate to one another. AI explainability, in other words, must be seen as part of the larger sectoral approach to governing specific technology-applications and stakeholders.

To illustrate this point, we will focus on the specific AI systems that recommend news on [social, and increasingly also on legacy](#) media. In the context of the media, AI is often [defined broadly](#) as "A collection of ideas, technologies, and techniques that relate to a computer system's capacity to perform tasks normally requiring human intelligence." Where AI is used to automate an editorial activity such as news dissemination, explainability is not just important to enable accountability vis-a-vis government and especially non-governmental actors, but also to afford individuals the means to control their media diets and generate the trust the media requires to fulfil its role in democracy.

Stressing the need for a more nuanced understanding of AI explainability, this Vision Paper draws on the concept of [cooperative responsibility](#), as developed by Helberger, Pierson, and Poell. For the purposes of this paper, we wish to highlight three key points:

- Cooperative responsibility addresses the idea that certain problems in complex (automated) systems (such as their impact on diversity, accountability, or privacy) are the result of the interplay of all stakeholders involved, and cannot be solved by, or attributed solely to one actor (the so called '[problem of many hands](#)').
- These problems cannot be fully resolved without considering the different roles, capabilities and responsibilities of all stakeholders and how these are meaningfully divided. The (degree of) responsibility of the respective stakeholders depends on, among other things, their respective knowledge, capacities, resources, incentives, and efficiency in addressing the problem. These factors are context-dependent.
- The responsibilities of each actor cannot be seen in isolation, but are interconnected. Platforms, for example, have a responsibility to ensure that the users whose interactions they facilitate can do their share to contribute to the realisation of public values (for example through the way they design and explain their systems). Individuals at the same time have an ability to exercise pressure on their governments and platforms to secure public values, and with that also a responsibility for the realisation of those public values. One critical precondition for all actors, including users, to be able to exercise their responsibility is explainability.

Following this logic, we will use section 2 to sketch the complex web of actors involved in providing and receiving explanations on AI in news recommender systems, and section 3 to highlight the normative complexities of AI explainability in this specific context. Section 4 will chart a path toward a more comprehensive form of cooperative AI explainability

2. Complex web of Actors

The digitization and platformization of our society has precipitated a complex infrastructure of actors intermediating information flows. This is also true for the way we consume news today. Consider the following fictional example:

A news website 'NewsAmsterdam' wants to use AI to improve news recommendations, which is arguably one of its [most important functions in the media industry](#) and [can be considered](#) one of the many editorial decisions the news website makes. NewsAmsterdam hopes to draw more readers to its website with the personalisation such a recommender system offers and which will, subsequently, make the news website more attractive for more advertisers. As the organisation lacks the required technical expertise, it seeks help from a promising startup 'RecommendingAI' to build an AI-driven recommender system. The startup is aware that a [large audience encounters news through online platforms](#), and its AI accordingly recommends articles that users are likely to share, which in turn is expected to increase NewsAmsterdam's (advertisement and subscription) revenue stream. People increasingly read AI-recommended articles, and share them on Facebook, where the articles are spread through news recommender systems. Eventually, many more people read the articles via Facebook than via NewsAmsterdam itself.

From this fairly straightforward example we can already observe the complex web of actors involved in the recommendation and dissemination of news content. The newspaper NewsAmsterdam itself is an important actor, as is RecommendingAI, the startup that created the recommender system. At the same time, individuals not only read AI-recommended articles, but also disseminate them to large audiences that are exposed to news through social media's AI recommenders. Finally, it is also important to recognise the role played by advertisers (and the [underlying industry](#)), and how they impact the optimisation goals of news recommender systems (e.g. engagement, sharing and/or time spent). Any form of explainable AI in this context, should acknowledge and even embrace this complexity, engaging the different actors in order to operationalize explainability. How these different actors are connected will be explored in the following pages.

2.1. Individuals

As reflected in the example, an individual can simultaneously take in several positions in today's news ecosystem, that each come with different rights and responsibilities when it comes to AI explainability. We see at least three important roles that an individual can have vis-a-vis news content: *consumer*, *disseminator* and *(co-)producer*.

As news consumers, individuals are mainly on the receiving end of explanations. In the example, NewsAmsterdam would primarily direct information about their AI recommender system to the individuals reading their articles. Indeed, a significant portion of the literature on AI explainability is specifically aimed at informing 'end-users'. At the same time, explanations that target end-users are also among the most heavily [criticized](#) legal information requirements, as they often are based on unrealistic assumptions about people's time, motivation, knowledge, and capacity to act on the information provided. That said, we still consider individual-targeted explanations - if done properly and [taking full advantage](#) of the interactive environment in which news is often consumed today - to be an important component to sensitize news consumers in general about the existence, scope and impacts of news recommendation systems. As section 3 will highlight, individuals play an important normative role in the specific context of the media. The choices they make as news consumers (and that are facilitated by AI explainability) cannot be easily outsourced to another party.

Apart from passively consuming news, individuals are also actively involved in disseminating news. As we saw in the NewsAmsterdam example, people started sharing the articles on Facebook. Online platforms, and social media in particular, are spaces where individuals interact with (news) content in many different ways. They do so, most notably perhaps, by sharing news stories with friends, followers or to the public at large. But individuals also affect how (news) content propagates in a more subtle manner. The platforms through which news is consumed (from content providers directly, to social media platforms) actively monitor individuals, in order to dynamically adjust their recommendation systems, often to boost attention and engagement. This can affect recommendations to the individual surveilled (in what is often referred to as the [algorithmic feedback loop](#)), but also to the recommender system more broadly. A simple example would be the 'most read' tab on many newspaper websites, but social media functionalities (liking, sharing, commenting, reacting, etc.) combined with granular surveillance (e.g. time spent, [user behavior outside of the platform](#), etc) affect how news is recommended to individuals in general, in much more fine-grained

ways. These new and opaque mechanisms enabling content to 'go viral', have been criticised for generating important risks, such as the [spread of disinformation](#).

[Metrics about news consumption](#) are also directly affecting news production by shaping the way in which individuals understand what content finds an audience on platforms. Which brings us to the third role of individuals in today's news-landscape; as content producers. Indeed, legacy media are increasingly incorporating individuals' content - from tweets to pictures and videos - into their news coverage. Moreover, as the [European Court of Human Rights](#) emphasises, bloggers and (popular) social media users can also be considered to fulfil the 'public watchdog' role, traditionally bestowed on legacy media alone.

Especially individuals' role as disseminators and (co-)producers of content and recommendations is strongly affected by their reach and influence, which can significantly vary. For example, political figures or semi-professional users such as influencers act as powerful disseminators/creators of content (and as such are [possibly also subject to ensuing transparency obligations](#) under advertising law) while private individuals posting about their personal life are not expected to exert much influence. When acting together, even people with an individually insignificant reach/influence, can have considerable impact on how the narrative of a news story evolves ([Pizzagate](#), being a cautionary tale in this regard). In sum, the position of individuals vis-a-vis explaining news recommender AI is complex to say the least. They are both important receivers of explanations, but their actions may also affect the content of explanations and/or necessitate them to provide explanations as well.

2.2. Legacy Media

It goes without saying that the media themselves, NewsAmsterdam in the example, occupy a central role in the discussion on news recommender explainability. They are still the primary sources of news content, and make important editorial decisions as to what content is generated and disseminated. These editorial decisions are [to varying degrees](#) outsourced when AI recommender systems are employed. This turn to algorithmically curated news recommendations, automating editorial decisions, often without much *a priori* human oversight or insight, poses new and important questions. The question of what editors' and journalists' role should be regarding AI that takes editorial decisions is very much one about their continued ability to exercise editorial control, and the new dependencies on engineers, data providers, external software providers, and commercial interests that they encounter in doing so.

It follows that in our example the relationship between NewsAmsterdam and RecommendingAI, should be taken into consideration with regard to the editorial control of NewsAmsterdam which is connected to the possible needs for explainability. In order to continue to claim both editorial control, and to be able to explain their editorial decisions and strategy to their audience, it is crucial that editors and journalists [understand how](#) the AI systems they use affect editorial values. In general, any discussion on explainable news recommenders should therefore also consider these complex dependencies, where content producers themselves also have legitimate claims to information and explanations.

2.3. Platform operators

In the example, more people eventually came into contact with the NewsAmsterdam's articles through Facebook than through its own portal. This is in line with [studies repeatedly demonstrating](#) how people are increasingly consuming news through social media and other online platforms. Importantly, online platforms such as Facebook heavily rely on AI-driven recommendation algorithms to sort and rank content, from news to personal messages, adding another layer of AI recommendation to our example. Automating editorial decisions on [what content \(not\) to show](#) on a mass scale, these algorithms have a subtle but important impact on what (news) content we consume and thus how we perceive the world. Indeed, it has been shown that algorithmic curation of content on social media may have [discriminatory effects](#) and result in [hiding critical social and political news](#) from users.

To complicate things even further, platforms can be considered to serve as gatekeepers or chokepoints of an otherwise highly dynamic and modular technological environment. Put differently, while platforms are often the front-end interface that users interact with, they generally rely on a constantly changing constellation of interacting service providers (e.g. hosting providers, content providers, advertisers, etc.). This [dynamic and complex reality of interacting online services](#) renders it even harder to properly map the distribution of rights and responsibilities in relation to explainable AI. Indeed, an effective explanation by one provider will often require information from other entities in order to be complete.

2.4. Advertisers

Often overlooked in the context of explainable AI but not less relevant are the actors that are ultimately driving much of the adoption and adaption of the technology: advertisers and [the AdTech industry](#) more broadly. AI systems, and news recommenders in general, are often [optimised to also realize economic goals](#) of a particular news outlet or a platform. Indeed, the interests of advertisers, which are often [closely entangled](#) with the interests of commercial media actors in particular, can be even the primary [reason](#) for implementing certain (recommender) technologies in the first place. And yet, it is these influences on news recommendations that are often also [particularly opaque and unseen](#).

Media and consumer laws require explanations about the [commercial influences](#) vis-a-vis particular items of media content, as well as the combination of editorial and commercial content. Yet, no clear corresponding obligations exist for the commercial considerations that have influenced particular news recommendations or the safeguards that the media have in place to separate commercial from editorial influences, respectively ensure editorial integrity.

2.5. Civil Society

A final and diffuse group, not directly reflected in the example, concerns key societal actors such as researchers and NGO's. As the AI systems that are part of news recommender systems and the regulation governing them are increasingly complex, civil society (interpreted broadly) fulfils an important function in both analysing these systems, making them understandable and helping people exercise their rights. As such, these societal actors can

function as middlemen helping individuals, the media, industry, and the general public to understand how these systems work, and what the societal consequences are.

An example of an NGO helping people navigate their rights, including helping them better understand news recommender systems, is [Bits of Freedom](#)'s "[My Data Done Right](#)" tool. This tool helps people in different Member States to exercise their GDPR rights, and notably their 'right of access' ([implying a right to explanation](#) as well), enabling better insight into why certain content was recommended to them. However, in many cases going from information to understanding requires expertise. This underlines the importance of research into the internal workings and external implications of AI systems. In the specific case of AI news recommender systems, this might mean giving some researchers privileged access to the system itself and the data it uses. Currently, [online platforms raise numerous legal, technical and economic obstacles to access](#) enabling external scrutiny, leaving researchers to try and develop all kinds of workarounds (e.g. Mozilla's [YouTube Regrets](#) project). A robust dialogue between developers and civil society, combined with independent research into the (wider implications of) AI news recommender systems is vital in the quest for effective and operational explanations.

AI news recommenders thus involve a wide variety of stakeholders including civil society, advertisers, platform operators, legacy media, and individuals. All perform a different role, with different relationships to one another, and different capacities to provide explanations. However, their responsibility to contribute to the larger effort to make AI-systems explainable cannot be considered without first acknowledging the different underlying values such explanations aim to advance.

3. Why explain?

Discussions on the need for explainable AI were catalysed by the debate on [algorithmic accountability and transparency](#) more broadly. We consider the normative claim to explainable AI - including when used for news recommendation - an important subset of this wider conversation. An explanation requirement goes further than a mere transparency requirement, in that it provides added detail and actively aims to make its target audience *understand* the respective AI system, whether detailed or only high-level. We consider explainability to primarily fulfil an instrumental role, safeguarding other rights, interests and values. For our purposes and in the context of news recommenders, it is worth highlighting three of these in particular: accountability, empowerment and trust.

3.1. Accountability

Generally speaking, explainability can enable accountability by allowing stakeholders to become aware of, understand, and ultimately contest AI systems or the decisions it produces. Of course, explanations do not produce accountability in isolation. Indeed, explanations need to be formulated and delivered in such a way that recipients can [understand them](#) and [act upon them](#), i.e. exert pressure on the accountable party based on what they have learnt.

Within the specific context of AI news recommendations, which parties have (or should have) the capacity to exert pressure is contentious due to the complex web of actors involved, and

the nature of decisions made by the AI. Even before the wide-spread adoption of AI, media policy debates have long struggled with the question of who should be able to demand accountability (if anyone at all), for what are ultimately editorial decisions. To safeguard freedom of expression and limit political interference, governments have relatively limited power to hold the media accountable. Correspondingly, there is an increased focus on accountability to [the market, other professionals, and the public](#), as well as a focus in the media to clarify the standards it holds itself to and explain individual editorial decisions. How these practices can be adapted to AI news recommenders and the complexity and scale of their decision-making, as well as the size and power of the main platforms that operate them, is subject to vigorous debate in policy-making circles.

Yet, explainable news recommender AI does not only serve holding to account legacy media. Making understandable why and how news is recommended has been a primary driver behind calls for so-called platform accountability as well. In the past half decade or so, social media operators have become an important [target of policy makers](#), trying to achieve more accountability over their content recommendation and moderation practices. Transparency and explainability - the latter giving shape to the former in particular contexts - constitute key elements in these [policy efforts](#).

Taking a step back even further, it is also worth highlighting that a robust explainability framework for (news) content recommendation systems will also help holding to account other societal actors. It might become easier, for example, to detect troll armies and other (groups of) individuals with nefarious intentions trying to game the recommender system. Currently, we have to rely almost exclusively on platforms themselves to detect and prevent malicious behaviour. More openness about their internal workings for independent scrutiny is hoped to [improve their accountability](#), as well as a healthy (news) content ecosystem more broadly.

3.2. Empowerment

The growing intermediation of our daily activities through AI, in particular news consumption, risks constraining individual autonomy. Indeed, opaque recommendation systems thwart individual agency over their information/news diet. As such, ill-designed, AI-infused recommender systems [risk affecting fundamental rights and freedoms](#) of news consumers, including their rights to data protection (Art.8 [Charter](#)) and freedom of expression and information (Art.11 [Charter](#)). A minimal understanding of the existence and operation of news recommendation systems is critical for users to be able to control how they receive and interact with information, as well as the recommendation systems behind it. In light of this, explaining AI-driven decisions and how they impact individuals is a necessary precondition for empowering users in the exercise of their fundamental rights.

Importantly, the growing reliance on AI-driven recommendation systems for news delivery also affects [the autonomy](#) of content providers themselves. Indeed, as illustrated by the example above, news consumption increasingly occurs through online platforms such as social media or news aggregation services, content providers have less control over the reach of their content and the context in which it is displayed. Recent policy developments, such as the [‘Platform-to-Business’ Regulation](#) and the German [Medienstaatsvertrag](#), appear to recognise these growing power asymmetries and actively promote improved explainability for business users (including news content providers) of online intermediation services.

3.3. Trust

Finally, explainability also constitutes a precondition for individuals, and society in general, to [trust](#) the use of AI in specific contexts such as news recommendations. Indeed, as [emphasised by the European Commission](#), “[b]uilding an ecosystem of trust is a policy objective in itself, and should give citizens the confidence to take up AI applications and give companies and public organisations the legal certainty to innovate using AI.” The major societal impact of AI necessitates earning trust through explicated respect for European values and fundamental rights. [Research has shown](#) that when it comes to news recommendations, people can place relatively more trust in algorithmic content curation, than (expert) human selection. Yet, such greater trust appears to be due to a perceived - [though often unfounded](#) - neutrality and objectivity of automated news curation. Explainable news recommender AI appears vital in verifying whether this trust is indeed deserved. Explanations about the inner workings of the recommender algorithm enable critical reflection, allowing trust to align with trustworthiness.

How explainability will do so is an empirical question that is more easy to assert in policy than to prove in practice. Over the past decade, the media has increasingly [started to explain](#) its editorial decisions to counteract [falling levels of trust](#) (at least in certain countries). [Research into trust](#), and the relationship between transparency and trust in news recommendations suggests AI explainability’s relationship to trust encounters many of the same complexities as empowerment and accountability. What individuals trust AI to do is inherently context-specific, and [whether AI is developed by a trusted brand](#) or influenced by [a complex web](#) of unknown actors appears to be a significant obstacle to trust.

4. Towards Cooperative Explainability

Discussing AI explanations in isolation or in the abstract is inherently limited. In our example, AI news recommenders further complicate an already complex web of actors involved in online information flows by transforming relationships and introducing new actors. Efforts that address the need for AI explanations in one of these relationships, or the need for explainable AI in general, are ultimately futile if they are not folded into a larger discussion that considers a particular sector, as well as the ways in which different stakeholders and their responsibilities interact with one another in that sector.

Existing approaches to AI explainability have predominantly dealt with this complexity in an ad hoc fashion only. In an environment that largely lacks legally binding norms, explainability is often (ab)used as an easy solution to address controversies as they come up. The few legally binding norms that do exist (e.g. in the [GDPR](#) and the [platform-to-business regulation](#)) are also too often presented and viewed as isolated from a comprehensive effort to make AI explainable.

This is problematic, because the need for different stakeholders to understand AI’s impact is pressing. Not only to fulfil their responsibilities to one another, but also to understand their own relationship to AI and be able to explain this (as necessary) to other stakeholders with whom they interact. Explaining the logic and effects of news recommendations thus involves different explanations, targeted at interconnected actors, with (at least in the media context) the overall

goals to enable these actors to act, trust, and hold one another accountable. Below, we illustrate how we can begin to address this complexity in the context of news recommenders.

- Arguably, as central chokepoints and main interface for end-users, **platforms** carry the primary responsibility for explaining how content is recommended through their platform, including how those recommendations are affected by the various actors involved - individuals, advertisers, other media organisations.
- From the perspective of **individuals**, explanations must not only show them how AI-driven recommendations affect the information they are seeing, but (to enable accountability, trust, and empowerment) also how their explicit and implicit choices shape this process, and how this relates to the role of third parties like advertisers or platforms.
- In order to be able to do so, **media organisations that deploy news recommenders** also need explanations about the algorithms involved and which entities are responsible for designing them (including, for instance, their underlying ideas and conceptualisations of public values when designing the recommender algorithms).
- **Third party vendors (notably of recommender-AI)** may thus be required to explain their AI to a media organisation, so it in turn can explain its use of AI as part of its broader editorial communication strategy to its audience. In case of media organisations buying off-the-shelf software, those media organisations would need explanations at least about what goals the AI system has been optimised for. Vice versa, producers of AI systems also need explanations from the media organisations they provide their software to, notably on how they conceptualise their editorial values, such as diversity or objectivity, in order to understand, and explain, the logic behind their AI systems and how those affect the workings of the media.
- This in turn requires **explanations within media organisations** about, on the one hand, what kinds of logics algorithms should operate in, what the role of commercial interests and the interests of advertisers are, and on the other hand, how news recommendations affect the editorial values and missions of a particular news outlet. Then there are advertisers who are still grappling with understanding how algorithms affect the placing of their advertisements next to specific contents, which again affects the kind of recommendations users get to see.
- Finally, there is a role for **independent research** - by civil society and academia - that is able to map and monitor [the highly dynamic interactions in this complex web of actors](#), and their implications for explainability's normative functions. Moreover, centralising explanation flows through one actor puts the governance system at risk of errors or abuse. In view of cooperative responsibility's acknowledgement of the capacity and responsibility of multiple stakeholders to address a service's impact on public values, explanations provided actors not directly involved can thus function as an important safeguard. In this sense, [enabling research](#) can be considered in line with the overall idea of cooperative explainability. For their part, academia, the media and civil society as well as the fact that these entities have not only a [right to access information](#), but also a (cooperative) duty to explain.

It becomes clear at that point that ‘understanding the logic involved as well as the significance and the envisaged consequences of such processing’ for individuals and society is far from a straightforward process, but rather a process that entails the interplay between a complex web of actors. It is also clear that ‘**cooperative explainability**’ as understood here goes far beyond the scope of the existing ‘right to explanation’ as currently implied in the GDPR. Before any such general forms of explainable AI can be meaningfully deployed, it is important to identify the stakeholders that shape AI in specific sectors, the relationships between these stakeholders, and the way explainability can, does, and should shape these relationships. Imposing general norms without understanding the shared demands of explainable AI across different sectors simply shifts the uncertainty forward and potentially constrains explainability discussions in sectors that do not fit the general norm.

The proliferation of AI-driven recommendations in the media have profound effects on the sector as a whole, the relationship and power dynamics between different actors, and the way the media performs its democratic functions. Those processes and dynamics are not easily captured by an individual right to explanation. The ecosystem involves a complex network of actors, each with different roles, necessitating a cooperative approach to governance and realising public values. What is needed, in addition, is a right to explanation regarding the broader political-economical logic involved in news recommendations and the way they affect not only individuals but society as a whole. Such a ‘right to understand’ currently does not correspond to a concrete legal right, but does constitute an important precondition for a healthy democratic society. And again, this obligation affects all actors in the ecosystem - to observe and consider the broader societal implications of AI-driven news recommendations and to enable each other to play that role.

In sum, a robust approach to explainable AI (in the news recommender context) requires us to recognise and embrace the complexity of the underlying ecosystems. Tunnel vision and over-simplifications fail to consider the multitude - and constantly moving interplay - of actors and values. We believe the model of cooperative responsibility can serve as a useful reference point in these explainability efforts.