UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

Speaking the news

How conversational agents influence trust and issue salience

Reséndez, V.

Publication date
2024

Link to publication

Citation for published version (APA):
Reséndez, V. (2024). *Speaking the news: How conversational agents influence trust and issue salience*. [Thesis, fully internal, Universiteit van Amsterdam].

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, P.O. Box 19185, 1000 GD 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)

Download date:17 Nov 2026

Chapter IV

News recommenders explained

The effects of explainability and modality on trust through privacy concerns and perceived personalization

This chapter has been submitted for publication: Resendez, V., Araujo, T., Helberger, N., & de Vreese, C. (2024). News recommenders explained: The effects of explainability and modality on trust through privacy concerns and perceived personalization [Manuscript submitted for publication]

Abstract

The growing use of news recommenders within conversational agents and social media platforms has amplified concerns about their transparency. In response to these concerns, current regulations mandate platforms to offer more transparency through explanations about the algorithmic processes driving the recommendations. In this sense, different explanation types have been designed to convey the necessary information to users. However, two challenges have arisen in this context: first, identifying the level of information individuals require; second, implementing the explanations for various types of modalities (e.g., audio, text). Existing research has shown mixed findings on how explanations can influence trust in the news recommender system since it can operate through perceived personalization or privacy concerns. Additionally, certain modalities (e.g., audio) in the news recommender might evoke the perception of engaging in a conversation with another being. Thus, the modality of the information could also influence users' perceptions (i.e., perceived personalization, privacy concerns). To assess the impact of explainability and modality on trust, we conducted a 2x2 experiment. We tested the influence of explanations' specificity (global vs. local) and modality (text vs. audio and text) on trust, through the mediating role of perceived personalization and privacy concerns. Our results reveal that perceived personalization and privacy concerns significantly affect users' trust in the news recommender, while explanations and modality do not have a direct or moderating effect. However, our findings suggest that local explanations could potentially increase privacy concerns, specifically within text-based recommender systems. These findings contribute to the interpretation and implementation of regulatory frameworks such as the Digital Services Act (DSA).

As the usage of news recommenders on platforms such as conversational agents (e.g., Alexa, Siri, Google Assistant) and social media (e.g., Instagram, Facebook) continues to rise, the concerns regarding their lack of transparency and explainability have increased. News recommenders are information filters aiming to provide relevant information for users (Karimi et al., 2018). Most news recommenders informed by user data rely on machine learning algorithms to make personalized recommendations. Problematically, these recommenders are black boxes that do not reveal the logic or process behind the recommendations (Diakopoulos, 2019). To address this issue, different explainable AI (XAI) methods have been created to provide justifications for the cause of the algorithmic decisions (Coeckelbergh, 2020).

Additionally, current regulations such as the General Data Protection Regulation (GDPR) and Digital Services Act (DSA) have established that very large online platforms should provide explanations for their recommender systems (Digital Services Act, 2022; General Data Protection Regulation, 2016). However, opening the black box of news recommenders presents two specific challenges: firstly, identifying the optimal level of detail for explanations aimed at the general public (Büchi et al., 2021); and secondly, operationalizing these disclosures across different modalities (e.g., audio, text). In a multimodal information ecosystem where more audio interfaces are emerging (Newman, 2022), it is crucial to analyze the impact of various types of explanations based on the modality of interaction (e.g., text, audio). This is particularly relevant as previous research has established that modality is an important factor in deciding on the type of explanation provided to the users (Turró & Lapedriza, 2019). Moreover, current regulations encourage the use of evidence-based best practices to assess platforms on the information conveyed to the audience (Dobber et al., 2021).

Multiple researchers have evaluated different explanation types and tested their effects on users (Dodge et al., 2019; T. Ha et al., 2020; Lim & Zhang, 2022; Wilkinson et al., 2021; Zhang & Curley, 2018). A shared insight is that explanations can influence trust in different directions. Some authors indicate that explanations' specificity influences trust through perceived personalization since it helps users assess the extent to which the recommender system can address needs and preferences (Zhang & Curley, 2018). On the contrary, some scholars also argue that explanation specificity can erode trust by triggering privacy concerns, as users may perceive a loss of control over their personal information (Dogruel, 2019; Eslami et al., 2018). In other words, the phrasing of the explanation can trigger privacy concerns, regardless of whether these concerns are justified. Therefore, since distinct forms of explanations may elicit different effects on users' trust by affecting either perceived personalization or privacy concerns, it is necessary to consider the nuances when providing explanations.

This study seeks to investigate the influence of global and local explanations (i.e., level of detail) on trust in news recommenders, by assessing

the level of information specificity in the disclosure. In addition to the explanations, it is critical to consider that these disclosures are incorporated into various types of modalities. Due to the different affordances of each recommender system, various elements can influence the perception of the explanation (Ehsan et al., 2021). The modality of interaction is a crucial factor that can influence how users perceive the explanation. Despite previous studies exploring the effectiveness of explanations in conversational agents (Shin, 2021a), voice-based recommenders (Weitz et al., 2021), and textual news recommenders (Shin, 2020, 2021b), there is a lack of research comparing explanation types and modalities in the context of news recommendations.

This study aims to bridge this gap by (1) testing the effects that distinct explanation types (local vs. global) have on trust, (2) examining the mediating role of privacy concerns and perceived personalization in the context of news recommenders, and (3) evaluating the interaction effects between global and local explanation types and different modalities (e.g., audio vs. text) on privacy concerns and perceived personalization. To guide this research, we propose the following main research question:

RQ: *To what extent does the explainability of news recommendations through text and audio influence trust, perceived personalization, and privacy concerns with regard to the news recommender?*

Explanations

There is no single definition of explainability in news recommenders and AI systems. Previous studies defined explainability as a justification or reason for a belief or action (Miller, 2019). Here we define explainability as the level to which a system can provide justifications for the cause of its decisions or outputs (Coeckelbergh, 2020). Prior studies highlight users' desire for more detailed explanations of why certain information is recommended (Büchi et al., 2021). However, it is unclear how much information users need to assess the system appropriately. A way to overcome this challenge is to directly evaluate users' perceptions of the system when exposed to different explanations (Coeckelbergh, 2020).

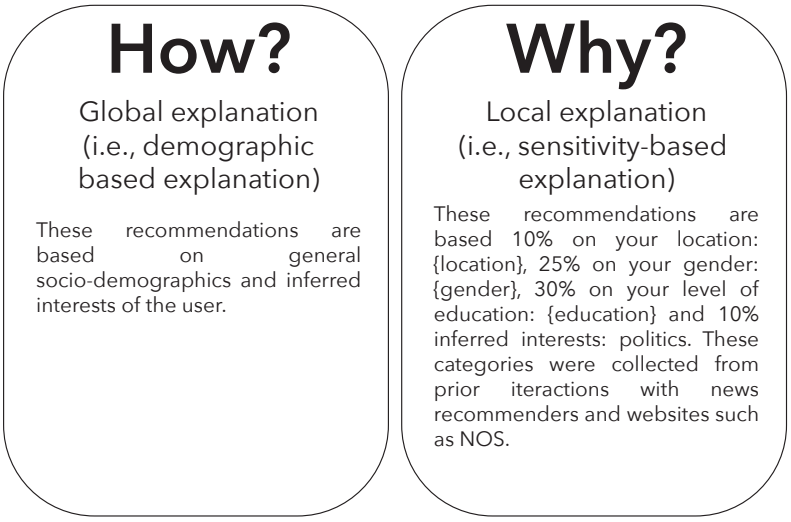
There are different taxonomies for explanations (e.g., post-hoc explanations, local explanations, feature relevance, explanation by example, simplification, case-based reasoning, deep formulation, attention networks) (Arya et al., 2019; Barredo Arrieta et al., 2019). These explanations usually follow a what, why, what for, and how structure to disclose the utility of systems' different features (Barredo Arrieta et al., 2019). In this study, we focus on global (i.e., how explanations) and local (i.e., why explanations) since these disclosures have a complementary nature that helps users

with their information needs (Dodge et al., 2019; Zhang & Curley, 2018). Global explanations help users understand how the algorithm works in general, which can vary from sharing source code and training data to mentioning the overarching features in the model. In contrast, local explanations provide details on each specific output the user is exposed to in the model. For example, in social media, a global explanation would be a general disclosure of the parameters used to create the news feed, while a local explanation would be "Why are you seeing this ad" in each of the recommendations made by the system, including specifications on the reasoning behind the recommendation (Kim et al., 2019; Wilkinson et al., 2021). These explanations are highly relevant since these disclosures can influence trust and fairness (Dodge et al., 2019; Eslami et al., 2018; Wang & Benbasat, 2016; Wilkinson et al., 2021; Zhang & Curley, 2018).

Explanations and Trust

Explanations are often assumed to have a direct influence on users' trust in the recommender system since there is a transference of knowledge from the system to the user (Miller, 2019; Shin, 2020). In a way, explanations

Figure 4.1.
Global and local explanations.



could be perceived as a cue for transparency, potentially increasing trust in the system (Eiband et al., 2019). Aligning with prior studies that test trust in recommender systems, we understand trust as users' perception of the news recommenders' competence, benevolence, and integrity (Benbasat & Wang, 2005; Wang & Benbasat, 2016). Previous studies examining the effects of explainability on trust in various domains, such as news recommenders and autonomous driving, have found mixed results (Dogruel, 2019; Eslami et al., 2018; T. Ha et al., 2020; Wilkinson et al., 2021; Zhang & Curley, 2018). Eslami et al. (2018) conducted a study on local explanations with varying degrees of information specificity.

They found that providing specific explanations increased users' trust in the system as it helped users understand the system's inner logic. However, Eslami et al. (2018) also noted that explanations might increase privacy concerns for users who prioritize their privacy. A similar study using local explanations with varying degrees of detail by Dogruel (2019) indicated that highly specific information could increase users' privacy concerns and reduce users' trust in the system as it might remind users of surveillance practices. T. Ha et al. (2020) further explored the effectiveness of local explanations, finding that attributional explanations were helpful in low-risk situations but not high-risk situations. They hypothesized that users' cognitive load might affect the impact of explanations on trust.

In contrast to these results, Wilkinson et al. (2021) evaluated two types of local explanations and found that "why" explanations increase users' sense of control compared to "why not" explanations. Similarly, Zhang and Curley (2018) investigated the impact of providing global explanations versus not providing any on user trust. Their findings indicate that explanations can influence multiple aspects of trust, including competence and benevolence. According to Zhang and Curley (2018), explanations increase trust by potentially enhancing users' understanding of the system and their sense of control over it. Moreover, they advocate for incorporating both global and local explanations to enhance users' trust when engaging with a news recommender.

While these findings provide insight into the relationship between explanations and trust, the extent to which two different types of explanations influence trust remains a question. Therefore, we propose the following research sub-question:

RQ1: *To what extent do local explanations lead to more trust in the news recommender compared to global explanations?*

The Effect of Explanation Types on Perceived Personalization

IV

Explanations can influence perceived personalization according to the information disclosed in the system (Zhang & Curley, 2018). We refer to personalization as the extent to which the content has been tailored to the individual (Vesanen, 2007). In addition, we understand perceived personalization as the extent to which users believe or feel that a system, service, or product has been tailored to their individual needs, preferences, or characteristics (Li, 2016). As a consequence, the information can be personalized by the recommender system without the user perceiving it as personalized, and vice versa (Li, 2016).

In the context of advertisements, De Keyzer et al. (2022) describe that including elements such as age, location, and interests in an ad leads to users' perceived personalization. Consequently, perceived personalization can foster a sense of acknowledgment, enhancing message processing and persuasion (De Keyzer et al., 2015; Li, 2016). This mechanism can potentially apply to the context of explanations in news recommenders. Since the goal of a news recommender is to filter information and present users with content tailored to their inferred preferences or explicit characteristics (Shin, 2020), disclosing the personalization process of recommendations can further enhance the sense of a tailored service, highlighting the system's intent to meet user needs (Zhang & Curley, 2018).

Additionally, disclosing the personalization process of recommendations can further enhance the sense of a tailored service, highlighting the system's intent to meet user needs (Zhang & Curley, 2018). Consequently, explanations that address users' personal information (local explanations) will lead to a stronger perception of personalization than generic explanations about the recommendation process (global explanations). Particularly, since personalized messages foster a deeper sense of relevance and attention among users, we hypothesize:

H1a: *Local explanations in comparison to global explanations increase perceived personalization.*

Perceived Personalization and Trust

Higher levels of perceived personalization have been linked to higher levels of trust in various technologies, such as virtual reality and recommender systems (Alimamy & Gnoth, 2022; Komiak & Benbasat, 2006; Zhang & Curley, 2018). Such personalization has been found to impact different aspects of trust, including competence and integrity (Zhang & Curley, 2018). In the context of news recommenders, perceived personalization might play

a crucial role. Personalization might foster the impression that the system can understand and serve users' needs and preferences, which, in turn, enhances trust in the recommender system (Komiak & Benbasat, 2006). A news recommender that can effectively communicate its ability to provide news items based on users' characteristics can improve users' perception of the system's value in filtering information (Alimamy & Kuhail, 2023; Zhang & Curley, 2018). We hypothesize that users' perceived personalization will increase users' trust in the news recommender system, as users value being recognized and served according to their needs. Hence, we propose the following hypothesis:

H1b: *Perceived personalization will positively influence trust in the news recommender.*

The Effect of Explanation Types on Privacy Concerns

As explanations may reveal the manner in which information is tailored to individual users, such disclosures could potentially heighten privacy concerns among users (Dogruel, 2019; Eslami et al., 2018). Privacy concerns refer to the worry that institutions are capitalizing on users' personal information (Dinev & Hart, 2006; Gouthier et al., 2022). Since each explanation can contain distinct levels of detail on the data collected, prior studies on explainability argue that the specificity of the data provided in the explanation can influence users' privacy concerns (Dogruel, 2019; Eslami et al., 2018). Moreover, the disclosure of personal information can heighten the perception of creepiness, thereby increasing privacy concerns within the system (Dogruel, 2019; Eslami et al., 2018). Based on these insights, local explanations (i.e., sensitivity-based explanations) may increase privacy concerns for users since more personal information is revealed to the user. We therefore hypothesize:

H2a: *Local explanations in comparison to global explanations will increase privacy concerns.*

Privacy Concerns and Trust

Multiple studies have found that privacy concerns can impact users' behavior on online platforms. Specifically, heightened privacy concerns have been linked to behaviors such as avoiding online shopping, reluctance to share personal information, and terminating online interactions (Q.-A. Ha, Victor Chen, et al., 2021; Liu & Tao, 2022; Smit et al., 2014; Vimalkumar et al., 2021).

Additionally, these concerns have been directly associated with diminished trust (Hoffman et al., 1999). Notably, privacy concerns can originate from users' perceptions of lacking control over their personal data, leading to negative evaluations of online systems (Boerman et al., 2021; Eslami et al., 2018). This sense of unease is often exacerbated when platforms mismanage users' privacy expectations (Kim et al., 2019) or when there are breaches of user information, with users interpreting these breaches as the institutions acting in their own self-interest rather than prioritizing users' privacy (Dinev & Hart, 2006; Kim et al., 2019). Ultimately, these factors can reduce trust in online systems, including recommender systems. Therefore, we hypothesize that privacy concerns in the news recommender can decrease trust in the news recommender:

H2b: *Privacy concerns will negatively affect trust in the news recommender.*

The Role of Modality

In addition to the explanations, it is important to consider these disclosures are embedded in different platforms. Prior studies had tested explanation in conversational agents (Shin, 2021a), voice recommenders (Weitz et al., 2021) and textual news recommenders (Shin, 2021b), however, the comparison between explanation types and modalities has not been researched in the context of news recommendations. Commonly, explanations are provided via text or visual formats (Turró & Lapedriza, 2019). Since there is a rise in news consumption through audio (Newman, 2022), it is important to understand the way auditive explanations can influence users' perceptions of the system.

Modality and Personalization

Information modality refers to the way in which the information is conveyed and how this affects the way messages are processed by individuals (Sundar et al., 2015). Modality can act as a tool for increasing perceived personalization (Verhagen et al., 2014), as conveying information via audio can increase the feeling of having a social conversation leading to a sense of social presence (Cho et al., 2019). Social presence is defined as the feeling that another being interacts with you, even though the being is non-human or non-intelligent (Lee & Nass, 2003). By eliciting social presence via audio, the information conveyed by the recommender system can appear to be shared by a social being aiming to serve the user. Consequently, the combination of local explanations and audio modality can increase the sense of a human being that acknowledges and serves users' needs. This suggests

that perceived personalization will be enhanced when the explanations are provided through audio modality, leading to the following hypothesis:

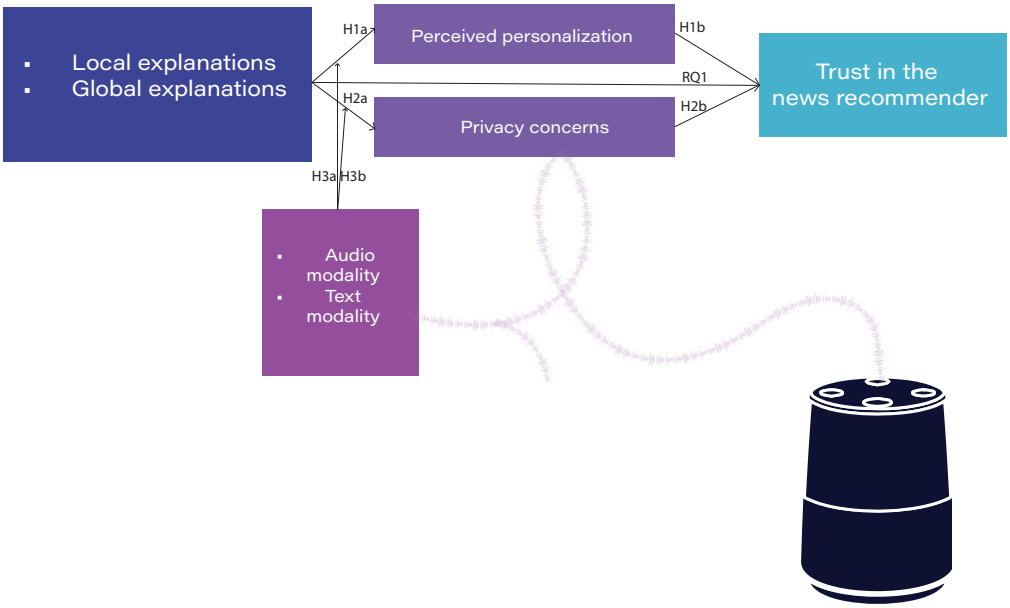
> **H3a:** *Modality moderates the effect of local explanations (vs. global explanations) on perceived personalization, such that the effect is stronger for individuals exposed to audio modality (vs. text only).*

Modality and Privacy Concerns

In addition, since interactions via audio can create a feeling of sharing information with another being, modality can reinforce concerns related to privacy in the interaction (Cho et al., 2019; Q.-A. Ha, Victor Chen, et al., 2021). From this perspective, including personal information through local explanations in audio news recommenders can increase the effect on privacy concerns in comparison to textual news recommenders.

Notably, since local explanations may trigger alarmistic perceptions (Dodge et al., 2019; Shin, 2021b), the combination of features between

Figure 4.1.
Summary of the model



local disclosures and audio news recommenders (e.g., aural conversational agents) can generate stronger negative influence on privacy concerns than in social media. Conversely, local explanations might be preferably disclosed through visual or textual modalities since the way of sharing the information is less privacy sensitive (i.e., less intrusive) than through audio modality (Cho et al., 2019; Easwara Moorthy & Vu, 2015). This suggests that privacy concerns will be more pronounced in the presence of audio modality, leading us to propose the following moderated mediation hypothesis:

H3b: *Modality moderates the effect of local explanations (vs. global explanations) on privacy concerns, such that the effect is more prominent for individuals exposed to audio modality (vs. text only).*

Methodology

To test the hypotheses regarding the moderating role of explainability in conversational agents (e.g., voice assistants) and social media, we conducted a between-factors experiment in February 2023. This study investigated the influence of explainability and modality (audio and text vs. text only) on trust through privacy concerns and perceived personalization. The experiment was pre-registered in OSF: rb.gy/lr5ax.

Pretest

We conducted two pretests to evaluate the manipulation of explanations. In the first pretest (N = 103), we tested the explanations suggested by Dodge et al. (2019), which included two types of local (i.e., sensitivity-based and influence-based) and global explanations (i.e., demographic-based and case-based inference). However, the manipulation checks did not show significant differences across conditions.

In the second pretest (N = 136), we evaluated only one local (i.e., sensitivity-based) and global explanation (i.e., demographic-based), incorporating additional recommendations from prior studies on explainability and information specificity (Büchi et al., 2021; Dogruel, 2019; Kyu Kim et al., 2020). Particularly, the local explanation included percentages of each parameter in the model and causes that helped determine the inferred interests (Figure 4.1). Examples of the stimuli can be found in the OSF link: rb.gy/lr5axa.

An analysis of variance (ANOVA) on the second pretest indicated a successful manipulation of the global and local explanations ($F(3, 133) = 21.021$, $p < .01$). Further, the post hoc comparisons utilizing the Tukey HSD test suggest that the local conditions, in both the audio (M = 5.24, SD = 1.34) and the text only (M = 5.23, SD = 1.11) modalities, were significantly perceived as more specific than the global conditions, in the audio (M =

3.73, *SD* = 0.90) and the text only (*M* = 3.71, *SD* = 0.97) modalities.

Procedure

After the pretests, we proceeded with a between-subjects online experiment. Participants were first requested to provide informed consent and complete a survey to gather demographic information, including age, gender, and city of residence. This information was used to tailor the explanations incorporated into the news recommender system.

Participants then interacted with the system by reading the explanation and reviewing a list of news items (Figure 4.3). Following this interaction, they completed a survey that included two attention checks, a manipulation check, and items related to privacy concerns, perceived personalization, and trust.

Participants

From February 8th to March 7th, 2023, we recruited a total of 604 participants for our study using Prolific. The final sample included 592 participants, excluding those who failed the attention checks and removing outliers based on cutoff points previously pre-registered rb.gy/lr5axa. The sample was limited to individuals residing in the Netherlands and aged 18 years or older, as the news recommender system was designed to operate in Dutch. We selected the Netherlands as a case study due to its high levels of connectivity and proclivity towards embracing novel technologies.The sample involved 327 male participants and 269 female participants, ranging in age from 18 to 73 years (*M* = 28.41, *SD* = 8.45).

Our initial aim was to recruit 1,300 participants, as determined by an a priori power analysis using a Monte Carlo simulation. However, since we only managed to include 592 participants in our final sample, we conducted a post hoc power analysis for the moderated mediation model using this sample (*N* = 592). The analysis indicated sufficient power ($\alpha = .05$, power $\approx .80$) to detect the direct effects of explanations, modality, perceived personalization, and privacy concerns on trust. Nevertheless, the power analysis revealed that the power might not be sufficient to detect moderated mediation effects of .3, as a sample of 592 participants only provided .52 power ($\alpha = .05$). To confirm the sufficiency of our sample size for assessing the impact of modality, we performed a post hoc power analysis for a two-way ANOVA using G*Power.

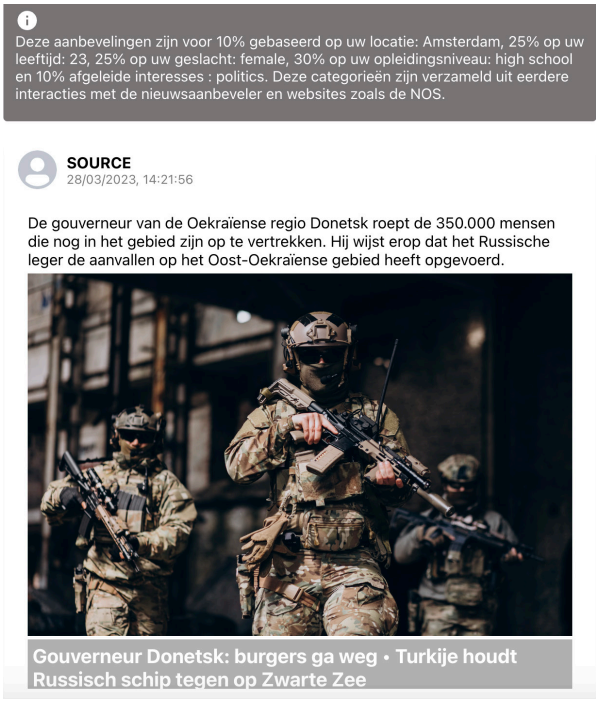
The analysis was conducted with a sample size of 592 participants and an estimated effect size of 0.25. The results revealed that our study had sufficient statistical power ($\alpha = .05$, power = .99) to detect significant direct effects of modality and explanations on trust, perceived personalization, and privacy concerns.

Measures

In this study, we conducted a confirmatory factor analysis (CFA) to evaluate the factor validity of all variables (Table 4.1). We assessed the factor loading estimates of the latent variables and only used items with a loading higher than 0.70 (Hair et al., 2010). Subsequently, we examined the construct reliability of all measures, which showed scores above 0.80 for all variables (Cronbach, 1951). Details of the constructs can be found in the OSF link: rb.gy/lr5axa.

Explanations. We used the framework developed by Dodge et al. (2019) to create global and local explanations for our recommendation system (Figure 4.1). Global explanations offered an overview of the data used in the recommendation process, while local explanations focused on specific features of the data. Both types of explanations were presented once to ensure comparability. We evaluated their effectiveness using a semantic differential scale assessing information specificity. The scale included statements such as, "The explanation disclosing why these news items were recommended to me was very vague – very concrete." Results indicated

Figure 4.3.
Example of the news recommender



that local explanations (M = 4.76, SD = 1.25, α = .81) were perceived as significantly more specific than global explanations (M = 3.54, SD = 1.29, α = .81). The efficacy of these manipulations was confirmed through a two-way analysis of variance (ANOVA), which yielded a statistically significant result ($F(1, 218) = 135.37$, $p < 0.01$). Post hoc comparisons using Tukey's honestly significant difference (HSD) test showed that local explanations, presented in both the audio (M = 4.80, SD = 1.30) and text-only (M = 4.73, SD = 1.22) modalities, were significantly more specific than global explanations, presented in the audio (M = 3.46, SD = 1.26) and text-only (M = 3.61, SD = 1.31) modalities. For the primary analysis, we used moderated mediation with PROCESS Model 7, incorporating a dummy variable for local (1) and global (0) explanations to assess differences between the groups.

Trust. We measured trust in the news recommender using 11 items (Wang & Benbasat, 2016). The items included statements such as "This news recommender is like a real expert in assessing news" and "This news recommender has the expertise to understand my needs and preferences about the news" (1 = strongly disagree, 7 = strongly agree). In the confirmatory factor analysis, only items with loadings higher than .70 were retained, resulting in 7 items with statements such as "This news recommender wants to understand my needs and preferences." The details of these items and the construct can be found in the OSF link: rb.gy/lr5axa. The combined reliability of these items was high (α = .87, M = 3.77, SD = 1.04).

Perceived Personalization. Perceived personalization was measured using 3 items (Zhang & Curley, 2018). We included statements such as: "This system understands my needs", "This system knows what I want". These items were measured on a 7-point Likert scale ranging from 1 (Strongly disagree) to 7 (Strongly agree). After combining the items, we see high reliability on these items (α= .87, M = 3.68, SD = 1.27).

Table 4.1.
Model fit indices

Fit statistics	Model
X2 (df)	183
p-value	0.000
RMSEA	0.093
CFI	0.863
TLI	0.842
Akaike Information Criterion (AIC)	38356.576

Privacy Concerns. We used 4 items to measure privacy concerns (Dinev & Hart, 2006). Some of the items were: "I am concerned that the information I submit on the news recommender could be misused" and "I am concerned about submitting information on the news recommender because it could be used in a way I did not foresee". We combined the items to create the construct. The items used showed high reliability ($\alpha = .91$, $M = 3.53$, $SD = 1.45$).

Results

To assess our research question and hypotheses, we utilized Model 7 in the PROCESS macro for IBM SPSS version 29. This approach allowed us to estimate conditional indirect effects within moderated mediation models (Hayes et al., 2011) while simultaneously testing both mediation and moderation effects (Hayes, 2012). To further understand the influence of the mediators, we conducted two exploratory analyses. First, we examined the indirect effect of the explanations without the presence of the moderator (modality) using Model 4 in PROCESS. Additionally, we conducted three separate ANOVAs to evaluate the influence of modality and explanations on trust, perceived personalization, and privacy concerns. A summary of the results is reported in Tables 4.2 and 4.3.

Direct effects

An open research question (RQ1) was posed to investigate the impact of local and global explanations on trust. The results showed a non-significant effect of explanations on trust ($B = .069$, $SE = .057$, 95% CI [-.042, .181]). Therefore, there is no direct influence of explanations on trust.

Hypothesis 1a suggested that exposure to local explanations would lead to a greater perception of personalization than exposure to global explanations. However, the findings suggest that explanations do not have an effect on perceived personalization ($B = .1575$, $SE = .1603$, 95% CI [-.1574, .4723]). Thus, Hypothesis 1a is not supported. Hypothesis 1b proposed that individuals with a higher perceived level of personalization would show greater trust in the news recommender. The results indicate that perceived personalization had a significant impact on trust ($B = .5577$, $SE = .0244$, 95% CI [.5098, .6057]). Therefore, Hypothesis 1b is supported.

Hypothesis 2a predicted that individuals exposed to local explanations would have a higher level of privacy concerns than those exposed to global explanations. However, the results demonstrate that the explanations did not have a significant effect on privacy concerns ($B = .0965$, $SE = .1826$, 95% CI [-.2621, .4551]). Hence, Hypothesis 2a is not supported.

Hypothesis 2b suggested that individuals with higher levels of

Table 4.2.
Summary of results for research question and hypotheses

Direct effects	B	SE	95% CI	Supported
RQ1: Explanation → Trust	.0399	.0623	-.0825, .1622	-
H1a: Explanation → Perceived personalization	.1575	.1603	-.1574, .4723	No
H1b: Perceived personalization → Trust	.5577	.0244	.5098, .6057	Yes
H2a: Explanation → Privacy concern	.0965	.1826	-.2621, .4551	No
H2b: Privacy concerns → Trust	-.0795	.0214	-.1217, -.0374	Yes
Index of moderated mediation	**Index**	**SE**	**95% CI**	**Supported**
H3a: Explanation*Modality → Perceived personalization	-.0346	.1209	-.2786, .2025	No
H3b: Explanation*Modality → Privacy concerns	-.0253	.0217	-.0747, .0111	No

*Note: This is the summary of Model 7 utilizing PROCESS Macro and including modality as a moderator.

privacy concerns would exhibit less trust in the news recommender. The results reveal that privacy concerns had a negative impact on trust (*B* = -.0795, *SE* = .0214, 95% *CI* [-.1217, -.0374]). As a result, Hypothesis 2b is supported.

Test of Conditional Indirect Effects

Hypothesis 3a indicated that individuals who were exposed to local explanations in the audio condition would report a higher level of perceived personalization compared to those who were exposed to local explanations in the text condition. The results showed that the modality of the explanation did not moderate the effect of the explanations on perceived personalization (index = -.0346, *SE* = .1209, 95% *CI* [-.2786, .2025]). Explanations did not influence trust through perceived personalization in the audio and text condition (effect = .0878, *SE* = .0898, 95% *CI* [-.0885, .2655]) nor the text-only condition (effect = .0532, *SE* = .0805, 95% *CI* [-.1067, .2126]). Thus, Hypothesis 3a is not supported.

Conversely, Hypothesis 3b proposed that individuals exposed to local explanations in the audio condition would have a higher level of privacy

concerns than those exposed to local explanations in the text condition. The findings show there is no significant difference in privacy concerns between those in the local explanation and audio condition compared to those in the text-only condition (index = -.0253, *SE* = .0217, 95% *CI* [-.0747 .0111]) .

Table 4.3.
Details on conditional indirect effects of explanation on trust

Paths	Effect	SE	95% CI
Explanation → Perceived personalization → Trust			
Audio and text modality (0)	.0878	.0898	-.0885 .2655
Text only modality (1)	.0532	.0805	-.1067 .2126
Explanation → Privacy concerns → Trust			
Audio and text modality (0)	-.0077	.0155	-.0406 .0221
Text only modality (1)	-.0330	.0168	-.0709 -.0063

*Note: This is the summary of Model 7 utilizing PROCESS Macro including modality as a moderator.

Although it seems privacy concerns influence trust in the text-only condition (effect = -.0330, *SE* = .0168, 95% *CI* [-.0709 -.0063]) compared to the audio and text condition (effect = -.0077, *SE* = .0155, 95% *CI* [-.0406, .0221]), the difference was no statistically significantly. Therefore, Hypothesis 3b is not supported.

Exploratory Indirect Effects

Additionally, as a robustness check, we tested the effect of explanations on trust using Model 4 in PROCESS Macro. In this model, we only included modality as a covariate to control for the treatment effects. Table 4.4 shows the direct, indirect and total effects of the exploratory analysis. The direct effect of explanations on trust was not significant (*B* = .0352, *SE* = .0629, 95% *CI* [-.0885, .1588]). The indirect effect of explanations via perceived personalization without the presence of the moderator (modality) was not statistically significant (*B* = .0682, *SE* = .0583, 95% *CI* [-.0455, .1839]) . The indirect effect of explanations through privacy concerns was significant (*B* = -.0218, *SE* = .0120, 95% *CI* [-.0483, -.0027]) . Those exposed to local explanations have higher privacy concerns than those exposed to global explanations. Finally, we tested the total effect. However, the results suggest a non-significant total effect (*B* = .0815, *SE* = .0871, 95% *CI* [-.0895 .2526]) of this model.

Table 4.4.
Direct, indirect and total effects

Paths	Effect	SE	95% CI
Direct effect: Explanation → Trust	.0352	.0629	-.0885, .1588
Indirect effect: Explanation → Perceived personalization → Trust	.0682	.0538	-.0455 .1839
Indirect effect: Explanation → Privacy concerns → Trust	-.0218	.0120	-.0483 -.0027
Total effect	.0815	.0871	-.0895 .2526

*Note: This is the summary of Model 4 utilizing PROCESS Macro and including modality only as a covariate.

Exploratory Two-way ANOVAs

Furthermore, we performed an exploratory analysis to address the issue of insufficient statistical power in the moderated mediation and to determine whether the absence of an effect was attributable to the limited power or to the absence of influence. We conducted three distinct two-way ANOVAs to evaluate the impact of both modality and explainability, as well as their potential interaction effects.

Our primary analysis was centred on the influence of explanations and modalities on trust. We employed a two-way ANOVA with trust as the dependent variable and the explanations as well as the modalities as the independent variables. This allowed us to assess their individual and interaction effects. The results revealed that neither explanations, $F(1,588) = 1.415$, $p = .235$ nor modalities, $F(1,588) = 2.981$, $p = .085$ had a significant main effect on trust. Additionally, their interaction failed to show a statistically significant effect, $F(1, 588) = 2.521$, $p = .113$. Our findings demonstrate that trust levels remained relatively consistent across all conditions. Particularly the audio and text modality with a global explanation ($M = 3.59$, $SD = 1.04$) was not significantly different than in the local explanation ($M = 3.83$, $SD = 1.06$). Similarly, the text-only modality with a global explanation ($M = 3.86$, $SD = 1.04$) was closely similar to the local explanation ($M = 3.82$, $SD = 1.02$).

In a subsequent analysis, our focus shifted to perceived personalization. Using a similar two-way ANOVA structure, we assessed the effects of explanations and modalities on perceived personalization. The results indicated that the effect of explanations on perceived personalization was non-significant $F(1, 588) = 2.006$, $p = .157$. In addition, the effect of modalities on perceived personalization was not statistically significant, $F(1, 588) = 3.126$, $p = .078$. The interaction between explanations and

modalities also lacked significance, F (1, 588) = 0.085, p = .771. From our analysis, we observed relatively low levels of perceived personalization across conditions. To further illustrate, within the audio-text modality with a global explanation, the mean perceived personalization was (M = 3.51, SD = 1.25); and with a local explanation, it was (M = 3.67, SD = 1.27). In the case of the text-only modality, the global explanation indicated a perceived personalization of (M = 3.72, SD = 1.22), while for local explanations it was (M = 3.82, SD = 1.32).

In our third analysis, we assessed the impact of explanations and modalities on privacy concerns, including their potential interaction effects. Results revealed a significant influence of explanations on privacy concerns, F (1, 588) = 4.549, p = .033. However, the effect of modality was not significant, F(1,588) = 1.566, p = .211. Furthermore, there was no significant interaction effect between explanations and modalities F (1, 588) = 1.711, p = .191. Subsequently, we conducted a post hoc comparison employing Turkey's honestly significant difference (HSD). The results revealed a significant difference in privacy concerns between the global (M = 3.27, SD = 1.44) and local (M = 3.68, SD = 1.40) explanations in the text-only modality. Conversely, the combined audio and text modality showed non-significant differences in privacy concerns between the global explanations (M = 3.56,SD = 1.49) and local explanations (M = 3.65, SD = 1.43).

Discussion

The growing prevalence of news recommenders on various platforms, including social media (e.g., Instagram, Facebook) and conversational agents such as voice assistants (e.g., Alexa, Siri, Google Assistant) have raised concerns regarding their insufficient transparency and explainability. The Digital Service Act (DSA) now requires platforms with recommender systems to ensure transparency through explainability. Despite current regulations, we still lack an understanding of how to effectively implement explainability to achieve normative objectives that empower individuals in their informed decision-making process. To bridge this gap, our study focuses on studying the effects of two explanation types (local and global) and two modalities (audio and text) on users' trust, as well as the underlying mechanisms, such as the relationship between explanations, modality, privacy concerns and perceived personalization. Therefore, we conducted an experiment to examine users' trust in the news recommender when presented with different explanations in different modalities (Coeckelbergh, 2020).

First, our analysis revealed that perceived personalization and privacy concerns influence trust in the news recommender. This finding is consistent with previous research in the context of e-commerce and advertisement (Alimamy & Kuhail, 2023; Dogruel, 2019; Eslami et al., 2018;

Li, 2016; Zhang & Curley, 2018), extending this relationship to the context of news recommender systems. Since perceived personalization refers to the extent to which users perceive that the recommender system understands and represents their personal needs and preferences, such perception can provide a sense of acknowledgment, enabling users to rely more on the recommender system (Komiak & Benbasat, 2006). Similarly, privacy concerns influence users' feelings of diminished control over their personal data, reducing trust (Dogruel, 2019; Eslami et al., 2018). Future research can go beyond perceived personalization and privacy concerns to evaluate other antecedents of trust such as perceived transparency and fairness, particularly in the context of explainability.

Second, our findings suggest that in text-only news recommenders, local explanations increase privacy concerns compared to global explanations. This indicates that providing more detailed information in text format about the data used for the news recommendations might increase privacy concerns. Potentially, local explanations via text can heighten users' awareness of the data collection process by institutions operating the recommender systems (Baruh & Popescu, 2017). Consequently, disclosing this information can create a feeling of 'creepiness' about the breadth of data collected (Dogruel, 2019; Eslami et al., 2018).

However, in audio news recommenders, we did not find a significant difference in privacy concerns between global and local explanations. Interestingly, our post hoc analysis revealed that the privacy concerns elicited from both local and global explanations in the audio and text modality paralleled the concerns evoked by the local explanation within the text-only condition. This might hint into a situation where the interaction with an audio news recommender might be enough to make users aware of third-party participation in data gathering (Q.-A. Ha, Chen, et al., 2021), regardless of the details in the explanation. Future studies should examine this observation further by comparing different types of explanations, such as counterfactual explanations or why and why-not explanations, within both audio and text-only news recommender systems to further understand their impact on privacy concerns.

Moreover, our results revealed that local explanations do not influence trust when compared to global explanations. In addition, although the type of explanation seems to influence privacy concerns in text-based news recommenders as discussed above, the overall results reveal a lack of significant total effect on trust. This aligns with a prior study from Dogruel (2019), which indicated that providing more detailed explanations does not necessarily influence trust. Instead, other factors might influence users' trust in the recommender. For instance, users might value that recommender systems filter the information for them (Dogruel, 2019). Additionally, prior studies suggest that when providing explanations providing ways to control the output of the recommendations might instead influence trust (Miller, 2019). In this context, future research should evaluate other factors in the

recommender system that influence trust.

Our findings come with certain limitations that must be acknowledged. First, our study was mainly conducted in a controlled environment using visual devices (e.g., computers, smartphones), the level of engagement in processing the information may have mitigated the effect of modality. Thus, future studies should consider investigating the impact of modality on purely audio-based devices recreating scenarios with lower information involvement to gain a more comprehensive understanding of the modality effects. Second, our experiment only focused on one user interaction with one news recommender, which may limit the ecological validity of our findings since users interact with various types of recommenders in their daily lives. In addition, this one interaction and our limited sample size might have impacted the statistical power of the moderated mediation. In this sense, users' frequent interactions with other types of recommender systems could alter their perceptions. Therefore, future studies could consider conducting longitudinal experiments incorporating explanations in different types of news recommenders with distinct modalities to evaluate users' perceptions of the explanations and the recommender system over time.

The rise of news recommenders on platforms like social media and voice assistants has opened multiple questions regarding transparency and explainability. This study provides empirical insights into users' perceptions by examining different explanations and modalities. Notably, local explanations do not have a direct influence on trust compared to global explanations, challenging commonly held notions about providing more information for users. Instead, the study shows the role of perceived personalization and privacy concerns in influencing trust. This finding emphasizes the importance of a balance between tailoring information and users' privacy. Additionally, our findings provide empirical insights that can guide the interpretation and implementation of the Digital Services Act (DSA) and inform the implementation of legal requirements for explainability. In this regard, it is important users understand the data and underlying mechanisms for news recommendations. Although detailed explanations may increase privacy concerns, the concerns could serve as a catalyst for both users and platforms to refine the recommender system and ensure it better aligns with user requirements. Moreover, it is essential explanations go beyond offering an overview of the reasons behind the recommendations. Instead, explanations should also provide actionable information that empowers users to modify the received output.

References

Alimamy, S., & Gnoth, J. (2022). I want it my way! the effect of perceptions of personalization through augmented reality and online shopping on customer intentions to co-create value. *Computers in Human Behavior, 128*, 107105. https://doi.org/10.1016/J.CHB. 2021.107105

Alimamy, S., & Kuhail, M. A. (2023). I will be with you Alexa! The impact of intelligent virtual assistant's authenticity and personalization on user reusage intentions. *Computers in Human Behavior, 143*, 107711. https://doi.org/10.1016/j.chb.
2023.107711

Arya, V., Bellamy, R. K., Chen, P. Y., Dhurandhar, A., Hind, M., Hoffman, S. C., Houde, S., Liao, Q. V., Luss, R., Mojsilović, A., et al. (2019). One explanation does not fit all: A toolkit and taxonomy of AI explainability techniques. *arXiv preprint arXiv*:1909.03012.

Barredo Arrieta, A., Díaz Rodríguez, N., Ser, J. D., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil Lopez, S., Molina, D., Benjamins, R., Chatila, R., & Herrera, F. (2019). Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, 82–115.

Baruh, L., & Popescu, M. (2017). Big data analytics and the limits of privacy self-management. *New media & society, 19*(4), 579–596.

Benbasat, I., & Wang, W. (2005). Trust in and adoption of online recommendation agents. *Journal of the Association for Information Systems, 6*(3), 72–101. https://doi.org/10.17705/1jais.00065

Boerman, S. C., Kruikemeier, S., & Bol, N. (2021). When is personalized advertising crossing personal boundaries? How type of information, data sharing, and personalized pricing influence consumer perceptions of personalized advertising. *Computers in Human Behavior Reports, 4*, 100144. https://doi.org/10.1016/J.CHBR.2021. 100144

Büchi, M., Fosch Villaronga, E., Lutz, C., Tamò Larrieux, A., & Velidi, S. (2021). Making sense of algorithmic profiling: User perceptions on Facebook. *Information, Communication Society*. https://doi.org/10.1080/136911 8X.2021.1989011

Cho, E., Molina, M. D., & Wang, J. (2019). The effects of modality, device, and task differences on perceived human likeness of voice- activated virtual

assistants. *Cyberpsychology, Behavior, and Social Networking, 22*(8), 515–520.

Coeckelbergh, M. (2020). Artificial intelligence, responsibility attribution, and a relational justification of explainability. *Science and Engineering Ethics, 26*, 2051–2068. https://doi.org/10.1007/s11948- 019-00146-8

Cronbach, L. (1951). Coefficient alpha and the internal structure of tests. Psychometrika, 16(3), 1951.

De Keyzer, F., Dens, N., & de Pelsmacker, P. (2015). Is this for me? How consumers respond to personalized advertising on social network sites. *Journal of Interactive Advertising, 15*(2), 124–134. https: //doi.org/10.1080/15252019.2015.1082450

De Keyzer, F., Dens, N., & De Pelsmacker, P. (2022). How and when personalized advertising leads to brand attitude, click, and WOM intention. *Journal of Advertising, 51*(1), 39–56.

Diakopoulos, N. (2019). *Automating the news: How algorithms are rewriting the media*. Harvard University Press.

Dinev, T., & Hart, P. (2006). An extended privacy calculus model for e-commerce transactions. *Information Systems Research, 17*(1), 61–80. https://doi.org/10.1287/ISRE.1060.0080

Dobber, T., Kruikemeier, S., Goodman, E. P., Helberger, N., & Minihold, S. (2021). Effectiveness of online political ad disclosure labels: Empirical findings.

Dodge, J., Liao, Q. V., Zhang, Y., Bellamy, R. K. E., & Dugan, C. (2019). Explaining models: An empirical study of how explanations impact fairness judgment. *Proceedings of the 24th international conference on intelligent user interfaces*, 275–285. https://doi.org/ 10.1145/3301275.3302310

Dogruel, L. (2019). Too much information!? Examining the impact of different levels of transparency on consumers' evaluations of targeted advertising. *Communication Research Reports, 36*(5), 383–392.

Easwara Moorthy, A., & Vu, K. P. L. (2015). Privacy concerns for use of voice activated personal assistant in the public space. *International Journal of*

Human-Computer Interaction, 307–335. https: //doi.org/10.1080/10447318.2014.986642

Ehsan, U., Liao, Q. V., Muller, M., Riedl, M. O., Weisz, J. D., Vera Liao, Q., Muller, M., & Riedl, M. O. (2021). Expanding explainability: Towards social transparency in AI systems. *Conference on Human Factors in Computing Systems - Proceedings*, 19. https://doi.org/ 10.1145/3411764.3445188

Eiband, M., Anlauff, C., Ordenewitz, T., Zürn, M., & Hussmann, H. (2019). Understanding algorithms through exploration: Supporting knowledge acquisition in primary tasks. *ACM Reference Format: Malin Eiband, 19 b*(10), 1-19. https://doi.org/10.1145/3340764.3340772

Eslami, M., Krishna Kumaran, S. R., Sandvig, C., & Karahalios, K. (2018). Communicating algorithmic process in online behavioral advertising. *Proceedings of the 2018 CHI conference on human factors in computing systems*, 1–13.

European Parliament and the Council of the European Union. (2016). 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 - GDPR). *Official Journal of the European Union*, L119, 1–88. Retrieved August 2, 2023, from https://gdpr-info.eu/

European Parliament and the Council of the European Union. (2022, October). Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a single market for digital services and amending Directive 2000/31/EC (Digital Services Act). *Official Journal of the European Union*. Retrieved August 9, 2023, from http://data.europa.eu/eli/reg/2022/2065/oj/eng

Gouthier, M. H., Nennstiel, C., Kern, N., & Wendel, L. (2022). The more the better? Data disclosure between the conflicting priorities of privacy concerns, information sensitivity and personalization in e-commerce. *Journal of Business Research, 148*, 174–189. https: //doi.org/10.1016/J.JBUSRES.2022.04.034

Ha, Q.-A., Chen, J. V., Uy, H. U., & Capistrano, E. P. (2021). Exploring the privacy concerns in using intelligent virtual assistants under perspectives of information sensitivity and anthropomorphism. *International Journal of*

Human–Computer Interaction, 37 (6), 512– 527.

Ha, Q.-A., Victor Chen, J., Uy Uy, H., Paolo Capistrano, E., Chi Minh City, H., & Nam, V. (2021). Exploring the privacy concerns in using intelligent virtual assistants under perspectives of information sensitivity and anthropomorphism. *International Journal of Human-Computer Interaction, 37*(6), 512–527. https://doi.org/ 10.1080/10447318.2020.1834728

Ha, T., Kim, S., Seo, D., & Lee, S. (2020). Effects of explanation types and perceived risk on trust in autonomous vehicles. *Transportation Research Part F: Traffic Psychology and Behaviour, 73*, 271–280. https://doi.org/10.1016/J.TRF.2020.06.021

Hair, J., Black, W., Babin, B., & Anderson, R. (2010). *Multivariate data analysis* (7th ed.). Pearson Prentice Hall. https://www.drnishikantjha.com/papersCollection/Multivariate%20Data%20Analysis.pdf

Hayes, A. F. (2012). *Process: A versatile computational tool for observed variable mediation, moderation, and conditional process modeling*.

Hayes, A. F., Preacher, K. J., & Myers, T. A. (2011). Mediation and the estimation of indirect effects in political communication research. In *Sourcebook for political communication research: Methods, measures, and analytical techniques* (pp. 434–465).

Hoffman, D. L., Novak, T. P., & Peralta, M. (1999). Building consumer trust online. *Communications of the ACM, 42*(4).

Karimi, M., Jannach, D., & Jugovac, M. (2018). News recommender systems–survey and roads ahead. *Information Processing & Management, 54*(6), 1203–1227.

Kim, T., Barasz, K., & John, L. K. (2019). Why am I seeing this ad? The effect of ad transparency on ad effectiveness. *Journal of Consumer Research, 45* (5), 906–932. https://doi.org/10.1093/JCR/UCY039

Komiak, S. Y. X., & Benbasat, I. (2006). The effects of personalization and familiarity on trust and adoption of recommendation agents. *MIS quarterly*, 941–960.https://doi.org/https://doi.org/10. 2307/25148760

Kyu Kim, Y., Yim, M. Y. C., Kim, E., & Reeves, W. (2020). Exploring the optimized

social advertising strategy that can generate consumer engagement with green messages on social media. *Journal of Research in Interactive Marketing, 15*(1), 30–48. https: //doi.org/10.1108/JRIM-10-2019-0171

Lee, K. M., & Nass, C. (2003). Designing social presence of social actors in human computer interaction. *Conference on Human Factors in Computing Systems - Proceedings*, 289–296. https://doi.org/10. 1145/642611.642662

Li, C. (2016). When does web-based personalization really work? The distinction between actual personalization and perceived personalization. *Computers in Human Behavior, 54*, 25–33. https://doi. org/10.1016/J.CHB.2015.07.049

Lim, J. S., & Zhang, J. (2022). Adoption of AI-driven personalization in digital news platforms: An integrative model of technology acceptance and perceived contingency. *Technology in Society*, 69, 101965. https://doi.org/10.1016/J.TECHSOC.2022.101965

Liu, K., & Tao, D. (2022). The roles of trust, personalization, loss of privacy, and anthropomorphism in public acceptance of smart healthcare services. *Computers in Human Behavior, 127*, 107026. https://doi.org/10.1016/J.CHB.2021.107026

Miller, T. (2019). Explanation in artificial intelligence: Insights from the social sciences. *Artificial Intelligence, 267*, 1–38. https://doi.org/ 10.1016/j.artint.2018.07.007

Newman, N. (2022). *Journalism, media, and technology trends and predictions 2022*.

Resendez, V., Araujo, T., Helberger, N., & de Vreese, C. H. (2023). *News recommenders explained: The effects of explainability and modality on trust through privacy concerns and perceived personalization* [Manuscript submitted for publication], Amsterdam School of Communication Research, University of Amsterdam.

Shin, D. (2020). User perceptions of algorithmic decisions in the personalized AI system: Perceptual evaluation of fairness, accountability, transparency, and explainability. *Journal of Broadcasting Electronic Media, 64*(4), 541–565. https://doi.org/10.1080/08838151. 2020.1843357

Shin, D. (2021a). The perception of humanness in conversational journalism:

An algorithmic information-processing perspective. *New Media and Society*. https://doi.org/10.1177/1461444821993801

Shin, D. (2021b). The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI. *International Journal of Human-Computer Studies, 146*, 102551. https://doi.org/10.1016/J.IJHCS.2020.102551

Smit, E. G., van Noort, G., & Voorveld, H. A. (2014). Understanding online behavioural advertising: User knowledge, privacy concerns and online coping behaviour in europe. *Computers in Human Behavior, 32*, 15–22. https://doi.org/10.1016/J.CHB.2013.11.008

Sundar, S. S., Jia, H., Waddell, T. F., & Huang, Y. (2015). Toward a theory of interactive media effects (time) four models for explaining how interface features affect user psychology. *The Handbook of the Psychology of Communication Technology*, 47–86. https: //doi.org/10.1002/9781118426456.ch3

Turró, M. R., & Lapedriza, A. (2019). Can we do better explanations? A proposal of user-centred explainable AI. ACM IUI 2019 Workshops.

Verhagen, T., van Nes, J., Feldberg, F., & van Dolen, W. (2014). Virtual customer service agents: Using social presence and personalization to shape online service encounters. *Journal of Computer-Mediated Communication, 19*(3), 529–545. https://doi.org/10.1111/JCC4. 12066

Vesanen, J. (2007). What is personalization? A conceptual framework. *European Journal of Marketing, 41*(5–6), 409–418. https://doi.org/10.1108/03090560710737534/FULL/PDF

Vimalkumar, M., Sharma, S. K., Singh, J. B., & Dwivedi, Y. K. (2021). 'Okay google, what about my privacy?': User's privacy perceptions and acceptance of voice based digital assistants. *Computers in Human Behavior, 120*, 106763. https://doi.org/10.1016/J.CHB. 2021.106763

Wang, W., & Benbasat, I. (2016). Empirical assessment of alternative designs for enhancing different types of trusting beliefs in online recommendation agents. Journal of Management Information Systems, 33 (3), 744–775. https://doi.org/10.1080/07421222.2016. 1243949

Weitz, K., Schiller, D., Schlagowski, R., Huber, T., & André, E. (2021). "Let me explain!": Exploring the potential of virtual agents in explainable AI

interaction design. *Journal on Multimodal User Interfaces, 15*(2), 87–98. https://doi.org/10.1007/S12193-020-00332-0/

Wilkinson, D., Alkan, Ö., Liao, Q. V., Mattetti, M., Vejsbjerg, I., Knijnenburg, B. P., & Daly, E. (2021). Why or why not? The effect of justification styles on chatbot recommendations. *ACM Transactions on Information Systems, 39*(4), 1–21. https://doi.org/10. 1145/3441715

Zhang, J., & Curley, S. P. (2018). Exploring explanation effects on consumers' trust in online recommender agents. *International Journal of Human-Computer Interaction, 34* (5), 421–432. https://doi. org/10.1080/10447318. 2017.1357904