The Pursuit of Transparency and Control: A Classification of Ad Explanations in Social Media

Daricia Wilkinson, Moses Namara, Lijie Guo, Karishma Patil, Apoorva Manda, Bart P. Knijnenburg
Clemson University
(dariciw, mosesn, lijieg, karishp, manda, bartk) @clemson.edu

Abstract

Online advertising on social media platforms has been at the center of recent controversies over growing concerns regarding users' privacy, dishonest data collection, and a lack of transparency and control. Facing public pressure, some social media platforms have opted to implement explanatory tools in an effort to empower consumers and shed light on marketing practices. Yet, to date research shows significant inconsistencies around how ads should be explained. To address this issue, we conduct a systematic literature review on ad explanations, covering existing research on how they are generated, presented, and perceived by users. Based on this review, we present a classification scheme of ad explanations that offers insights into the reasoning behind the ad recommendation, the objective of the explanation, the content of the explanation, and how this content should be presented. Moreover, we identify challenges that are unaddressed by either current research or explanatory tools deployed in practice, and we discuss avenues for future research to address these challenges. This paper calls attention to and helps to solidify an agenda for interdisciplinary communities to collaboratively approach the design and implementation of explanations for online ads in social media.

1. Introduction

Advertising has steadily grown to become an essential part of the business models of social media platforms and accounts for a significant proportion of the total revenue generated by these companies. Online advertising is usually highly dependent on personal data that the company has collected directly (through likes etc.) or indirectly (through partners or tracking on other websites). It can be argued that the large-scale collection of detailed personal data allows companies to serve ads that are better aligned with consumers' interests. However, adopting this approach often results in the sacrifice of privacy.

As a result, there have been growing concerns around the opacity and invasiveness of online advertising—people do not understand what data is being collected about them, how that data is being used to serve them ads, and what they can do when they are served with ads they do not like. Prior work has found that consumers found targeted ads to be useful but "creepy" [1, 2, 3, 4]. Moreover, people have reported fears over discrimination in ads [5], risks of being embarrassed [6], and concerns about the accuracy of inferences [7, 8]. As many use social media as a gateway to the Internet, providing ad transparency will have a significant impact on people's online experiences. Compared to traditional targeting ads on the web, social media companies have access to richer data sources and more detailed personally identifiable information (PII) that can be used as attributes to allow advertisers to target their users. Additionally, unbeknownst to many, data sharing among partners is extensive. For example, purchasing the data on 70 million US households enabled Facebook to tailor ads to specific audiences based on these users' purchasing history on other sites [9].

Existing work has explored methods for promoting transparency and explanations for targeted advertising on the web (see [10] for an overview). However, the tracking capabilities through social media could be extensive as information could be gathered over multiple devices (i.e. knowing your behavior on your tablet usage versus your phone), websites versus different platforms and inferences could be made from users' social networks (e.g. friends and family). Considering the invasive nature of such practices, both end users and companies are increasingly acknowledging the importance of ad transparency. Unfortunately, there has been limited work on ad explanations in social media platforms. The urgency of this type of work is further exacerbated by the major inconsistencies in explanation tools that are currently deployed commercially. Various platforms have voluntarily provided insights into the

HĭCSS

reasons behind the occurrence of individual ads via a "Why am I seeing this ad?" button, but the properties of the explanation provided by this feature vary significantly across platforms. Moreover, studies have found that current ad explanations in social media are often incomplete and at times misleading [11]. As such, researchers have conducted design probes to assist with the creation of ad explanations that more align with users' expectations [12].

Despite these limited but significant strides, there are still gaps around the properties that should be included when attempting to explain marketing practices in social media. This paper aims to take the first steps towards exploring the key characteristics that may be considered when creating transparent explanation tools for ads in social media. We present a literature analysis that explores 1) the objectives of the explanation (which influences the design choices), 2) the content of the explanation, and 3) how the explanation is presented to users. Although not exhaustive, this work contributes towards a better understanding of social media advertising, and it explores options for a more responsible and ethical online experience.

Our work offers contributions that could inform practice and research. To our knowledge, this is the first *systematic* review on this topic. We conducted a search process in a structured criteria-based manner which assisted in avoiding potential researcher bias. Our main contribution is our classification scheme which was constructed based on the results of our systematic approach. Rather than focusing on one aspect of explanations for ads, we adopted a holistic approach which resulted in a comprehensive overview that could inform the design and development of explanation or transparency tools. Finally, we discuss meaningful directions for future work based on the gaps identified in our review.

2. Background

To situate our research in the context of the existing literature, in this section we provide an overview of targeted advertising techniques and the state of the art in explanation tools in social media.

2.1. Targeted Advertising in Social Media

Many use social media technologies to connect with loved ones, stay updated with world news, and share personal updates, beliefs, and emotions [13]. The wealth of information collected on these platforms makes it possible for companies to develop extensive user profiles [11, 14] and allow marketers to target consumers based on fine-grained information [15, 16].

Fingerprinting and other techniques for tracking are widely adopted and combined with user-provided personal information and behavioral data from interaction within the platform to create comprehensive profiles. Prior work has shown that targeting based on seemingly neutral attributes can lead to advertising practices that are discriminatory against marginalized groups [17]. For a detailed overview of online advertising see Yuan et al. and Alavi et al. [18, 19].

2.2. Explanation Tools

Recently, there has been a surge in interpretability and explainability research in Machine Learning and similar domains, acknowledging the importance and benefits of more interpretable systems [20]. Moreover, some studies have viewed interpretability in terms of providing explanations around the input parameters that most impact the output. Other works have explored incorporating transparency at the early stages of design with the hope that it would result in more interpretable systems [20]. Regardless of the approach, the underlying goal of this research is to provide end-users with sufficient information to assist them in identifying misbehavior. However, there are few empirical studies that focus on how explanations of ads serve end-users and what factors influence the interpretability of marketing practices in social media.

Prior work has explored methods to detect *why* users were being targeted with particular ads [21, 22, 23, 24]. Researchers have also investigated interfaces that explain online tracking. For instance, Weinshel et al. considered design dimensions for popular ad blocking tools (such as Ghostery) with a tool named "Tracking Transparency" [25]. They found that users of the tool were able to more accurately describe the extent of tracking, and they were more willing to engage in privacy-protecting actions. Although most of the existing work on explaining marketing practices is centered around general Web experiences, Andreou et al. investigated the effectiveness of ad explanations on Facebook [11]. They found that explanations on the platform were often incomplete, misleading or vague. Similarly, Eslami et al. found users preferred interpretable "non-creepy" explanations for ads on social media [12]. However, balancing *how* to provide explanations that align with what consumers want is a challenge, as revealing too much or too little about the algorithmic process has been shown to both negatively impact system trust [26].

3. Approach to Systematic Review

A number of survey papers from various disciplines have looked at the state-of-the-art concerning either

(Sub)-Discipline	Generated	Presented	Perceived	Total	% of Total
Computer Science	20	4	8	32	61.54
HCI	*3*	*4*	*7*	*14*	*26.92*
Security	*10*	*0*	*0*	*10*	*19.23*
AI/ML	*5*	*0*	*1*	*6*	*11.54*
Recommender Systems	*2*	*0*	*0*	*2*	*3.84*
Marketing	0	0	14	14	26.92
Information Systems	0	0	3	3	5.77
Communication	0	1	2	3	5.77
Total	20	5	27	52	100

Table 1.   Description of the core papers.

explanations in decision systems or online behavioral targeting. However, these surveys are often limited in their scope, they are restricted to a small subjective selection of papers, or the focus is specific to particular aspects of explanations. To our knowledge, no prior work has attempted to provide a comprehensive overview of existing work on explanations in the context of online advertising in social media.

3.1.   Research Goals

The interdisciplinary review provided in this paper aims to assist designers, developers, and researchers of social networking sites understand what has already been explored within the context of explanation mechanisms.

We choose to use the term 'classification scheme' opposed to terms such as 'framework', 'taxonomy', or 'typology' based on our analytical approach and the criteria identified from prior work [27, 28]. Bailey offers a comprehensive review of grouping systems used in social sciences [28]. Bailey explains that the term classification could be used to describe the "ordering entities into groups or classes on the basis of similarity" [28]. This could be accomplished either conceptually or empirically.

In this work, we adopted a conceptual approach followed by an analysis of empirical cases. Nickerson et al. describe best practices for classification within Information Systems [27]. We follow the development method for a conceptual-to-empirical approach as described by Nickerson et al. [27]. First, we conceptualize characteristics and dimensions of objects. In particular, we consider fundamental characteristics that would influence an ad explanation. We propose dimensions that investigate: (i) the underlying mechanisms that generate explanations; (ii) how the explanations are presented; (iii) how the explanations are perceived by users. Next, we examine objects for these characteristics. Lastly, we create or revise the classification. We aggregate the insights gathered

by our review in a classification scheme with the interrelated goals of a) helping stakeholders understand the multifaceted nature of explanation facilities and b) providing an aid for ad explanation designers.

3.2.   Inclusion and Exclusion Criteria

The identification of related work began with a search in the following databases and search engines: Google Scholar, ACM Digital Library, ScienceDirect, Springer Link, and IEEE Xplore Digital Library. Regarding the exclusion criteria, the language of the article had to be in English (one potential article was written in Mandarin and therefore it was excluded) and there was no limit on the publication year.

Our search string covered three main concepts. The first concept was *explanations*, which was the main topic of our work. We included the synonym 'justification' as well. The second concept was *online advertisement*, for which we included the synonyms 'online advertising', 'ads', and 'online behavioral advertising'. The third concept was *social media*, the specific context of our review. The final search string used is as follows:

> (explanation OR justification) AND (online advertising OR ads OR online advertisement OR online behavioral advertising) AND (social media)

The titles and abstracts for all relevant results were read to further determine relevance. Related articles and referred works that may not have been present in the original search were considered. They were retrieved from references cited in core papers and work that may have cited an article of interest. The inclusion criteria featured: studies that specifically mentioned explanations or transparency tools for social media, studies conducted with participants who tested explanations for ads in social media, and design proposals that were focused on explanations for ads in

Generated: papers focused on targeting mechanisms	[29], [30], [31], [32], [33], [34], [35], [36], [37], [38], [39], [40], [41], [42], [43], [44], [45], [46], [5], [47]
Presented: papers focused on design of explanations	[48], [8], [49], [50], [51]
Perceived: papers focused on consumer attitudes, beliefs and perceptions	[52], [53], [54], [55], [56], [57], [58], [3], [59], [60], [61], [62], [63], [64], [65], [66], [67], [68], [69], [70], [71], [72], [73], [74], [75], [76]

Table 2. Core papers and a description of their research contribution

social media.

If a paper was deemed to be relevant it was selected to be analyzed in detail. We classified each paper according to its relevance to how the ad was *generated*, *presented*, or *perceived*. Each paper was analyzed by one researcher who strictly followed the guidelines for inclusion. This procedure resulted in a total of 309 papers of which 52 core papers were selected after exclusion (see Table 2). Table 1 and Figure 1 feature a breakdown of the papers organized by discipline and focus (i.e. on how the ad was generated, presented, or perceived).

4. Literature Analysis

In the remaining sections we reflect upon the results of our literature analysis. In particular, we (i) propose a new classification scheme that represents various groupings that should be taken into consideration when designing explanation tools for ads and (ii) describe possible directions for future research.

4.1. Related Classifications

Researchers in various disciplines have proposed taxonomies with different perspectives on explanations in online systems [35, 77, 78]. These works capture important aspects of explanation facilities including the purpose, technical limitations and affordances, and information that should be conveyed to users. However, they are specifically focused on the recommendation process which is only a component of the entire ad experience. Our classification scheme, shown in Figure 2, includes dimensions that represent the various characteristics of a comprehensive explanation for online ads presented in social media platforms.

4.2. Major Facets of Ad Explanation

Our classification scheme includes three major facets that cover the mechanisms of social media advertising that can be featured in an explanation:

1. the reasoning behind **why** an ad was served;

2. **how** the data was collected or processed;

3. and **what** should be displayed to users.

Several different attributes can be identified as mechanisms that contribute to **why** a particular ad would be served to a user. The subset of empirical research that emerged under this facet were primarily focused on the information that would explain how the ad was *generated*. For this facet we consider three main aspects: the input parameters, context, and input source. The approaches that help to identify the reasons behind an ad may vary greatly as Andreou et al. indicated when they explored the input parameters used in ad explanations on Facebook [30]. Ad explanations that present the input parameters of the ad targeting algorithm may focus on *stated attributes* such as demographics provided by the user to the platform or *inferred attributes* such as an interest in cars inferred from the user's recent visits to car dealer Web sites. This information is often combined with contextual information (e.g. "You are seeing this ad because of your current location") and the source of the data (i.e. whether the data was collected on the platform or third party partners).

Unlike literature that focused on the reasoning, there are technical mechanisms that influence *how* information is collected, inferred or processed. Looking across the literature, it is unsurprising that different platforms would employ various strategies to create inferences. Increasingly, there are third-party tools that aim to understand the processes and algorithmic behaviors in an effort to provide more transparency [36].

Empirical work on how ad explanations were presented and perceived helped to distinguish the characteristics for the next facet. In the next major facet of our classification we acknowledge *what* is being displayed to users. This facet is organized by two major categories: (a) the *content* of the explanation itself, and (b) how that information is *presented*. Under the *content type* facet of the classification there are two sub-facets. First, the scope of the explanation may differ. Explanations may be focused on *individual ads* which is useful for when users are confronted with potentially problematic ads. Likewise, explanation tools may seek to offer an overview of advertising practices across all social media *platforms*. The scope could also be expanded beyond social media to explain
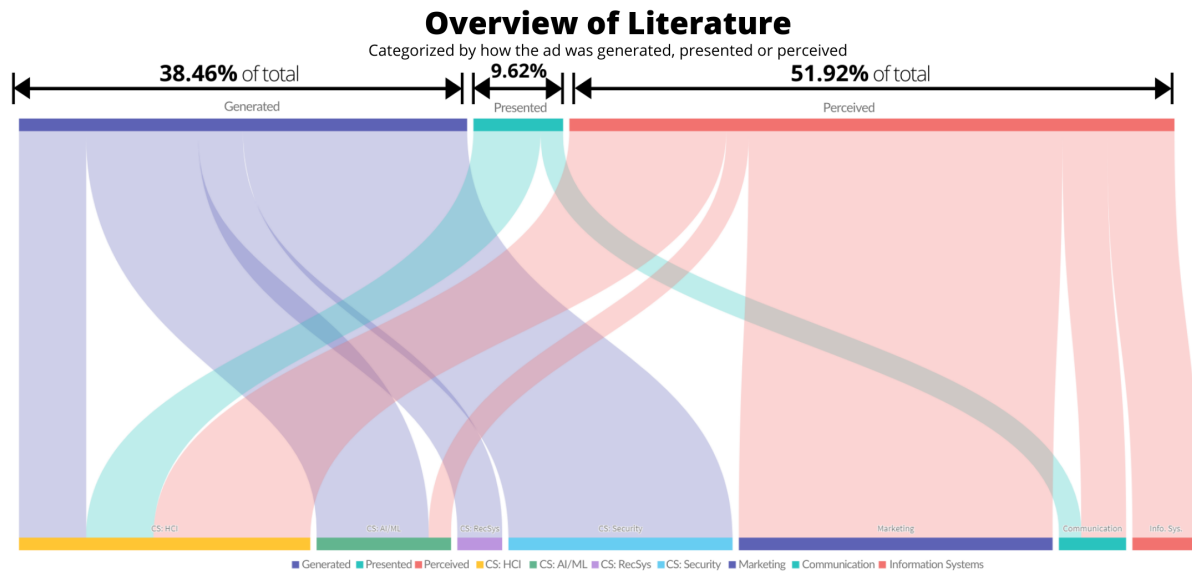
Figure 1. Core papers of the literature review categorized by the purpose of the paper and the discipline.

how information collected on platforms are connected to users' general internet behaviors and usage. This would empower users by educating them about online advertising at the *ecosystem* level. The effectiveness of the explanation may be influenced by its *interpretability*. For example, more complete explanations may feature more accurate representations of marketing and data collection practices but providing an extensive amount of information may become overwhelming to consume [79]. Similarly, the *format* of the explanation may be more appealing to consumers while the *level of engagement* would offer actionable options based on the needs of consumers.

A significant number of works within social media advertising consider users' attitudes, beliefs, and perception. The creation of an ad explanation involves a number of design choices that significantly deviate depending on **the objective** of the explanation. For instance, previous works indicate that an explanation may be provided to promote factors such as trust or transparency [78, 80]. However, while that may be a surface-level goal, the underlying intention for a platform provider may be to increase user acceptance of advertising in the platform. We propose three levels of possible *objectives* for explanations:

1. *explanation purpose*, which are overall objectives of providing the tool;

2. *stakeholder goals*, e.g. to allow feedback that would improve recommended ads;

3. *subjective factors*, which are perceived aspects that contribute to the experience of using the

system such as trust.

The design of an explanation may vary depending on its purpose. For instance, explanation tools aimed at *transparency* might highlight the advertising practices of the platform and the goals of the advertiser to increase user acceptance of recommended ads [44, 79], while mechanisms for *self-actualization* may consider users' future ambitions or their desire to learn more about their inferred tastes or perceptions of their algorithmic selves [62, 81] (i.e. who advertisers/platforms think they are, based on inferences made).

In the next section, we discuss insights gathered from our review and development of the classification scheme.

5. Discussion and Future Directions

Although there have been substantial strides towards explanation and transparency tools within social media, there are a variety of challenges that still need to be addressed. In this section, we offer examples of possible directions for future research. We present these as "calls to action", as we highlight concerns that need investigation and important gaps in the literature.

5.1. Comprehensive Explanation Content

We discovered that there are wide disparities regarding what type of content is considered to be included in the explanation. This holds true for both research papers as well as in deployed tools in social media platforms. Explanations provided by existing tools vary greatly in their level of granular detail.
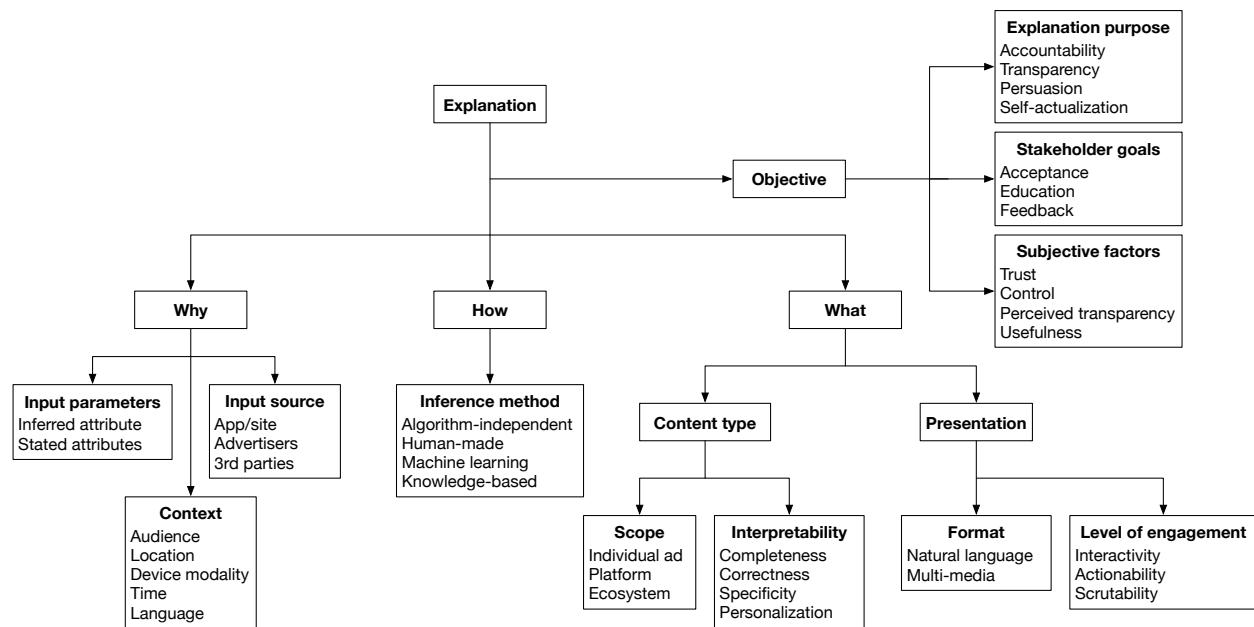
Figure 2. Ad explanation classification scheme

For some explanations the language is oversimplified, while other explanations are vague and incomplete [48, 51, 30]. In turn, this contributes to explanation tools generally being perceived as untrustworthy and uninterpretable [49, 51].

Ideally, an effective explanation tool would serve the interests of the stakeholders (i.e. the advertisers and the platform) while at the same time giving users agency over their data. Admittedly, designing non-creepy yet complete explanations is a formidable task. This is an understudied area which would have direct implications for the success of these types of tools.

5.2. Towards more Action

Offering a description of the process that led to an ad being served is helpful when people are inquisitive. However, if the user takes issue with an ad, then an explanation alone might not be enough. For example, if a user is uncomfortable with being targeted with health ads, then a mere explanation of why they were targeted would not be helpful. Instead, it might be more useful to provide direct links to taking action towards either deleting collected health data or opting out of targeting by this type of advertiser.

While many social network platforms have control mechanisms for ads, these are often physically and temporally divorced from the ads themselves. Future studies could explore how actionable options can be combined with explanations to create *in-situ* tools that would assist in addressing users' concerns about presented ads.

5.3. Expanding the Scope

Most of the papers reviewed in our analysis focused on explanations at the individual ad level. This is a reasonable research direction considering the design and layout of most social media platforms: it gives people easy access to information about the ads they see in order to control how their data is used in advertising on that platform. However, limiting the scope to this use case may restrict possibilities for privacy protection at a larger scale. Future work could explore tools at both the platform (i.e. all the ads on a specific platform) and ecosystem level (i.e. all the ads on all social media platforms). This would empower users by allowing them to investigate advertising behavior at both the micro and macro level. Additionally, tools that expand their scope would be useful to other stakeholders who conduct algorithmic audits. In this way, future work could help identify misbehavior and unfair practices that might be more widespread.

Moreover, one of the most surprising findings of this review is that the majority of research on the development of explanation tools is conducted in isolation from actual future users of these tools. In other words, newly developed tools are rarely tested by their expected audience before publication/deployment. We believe this would be a fruitful area for future work. Ideally, such research could be executed by multidisciplinary teams, where a tool could be designed and developed by Computer Scientists, informed by IS or marketing researchers' understanding of human

behavior and reactions to advertising, and tested by Human-Computer Interaction specialists. Similarly, research focused on perception should expand beyond acceptance and ad avoidance and consider identifying features that are most important for adoption by consumers.

6. Conclusion

Based on our literature analysis, we propose a classification scheme that highlights three major facets: *why* an ad was served, *how* the data was collected or process, and *what* is being displayed to users. Moreover, we discuss open challenges and possible research directions such as exploring appropriate content that would inform but not overwhelm, providing more actionable options with explanations, expanding the scope of research beyond individual ads and fostering more multidisciplinary work.

References

[1] B. Ur, P. G. Leon, L. F. Cranor, R. Shay, and Y. Wang, "Smart, useful, scary, creepy: perceptions of online behavioral advertising," in *proceedings of the eighth symposium on usable privacy and security*, p. 4, ACM, 2012.

[2] L. Barnard, *The cost of creepiness: How online behavioral advertising affects consumer purchase intention*. PhD Thesis, The University of North Carolina at Chapel Hill, 2014.

[3] B. Zarouali, K. Ponnet, M. Walrave, and K. Poels, ""Do you like cookies?" Adolescents' skeptical processing of retargeted Facebook-ads and the moderating role of privacy concern and a textual debriefing," *Computers in Human Behavior*, vol. 69, pp. 157–165, 2017.

[4] H. Torkamaan, C.-M. Barbu, and J. Ziegler, "How can they know that?: a study of factors affecting the creepiness of recommendations," in *Proceedings of the 13th ACM Conference on Recommender Systems*, pp. 423–427, ACM, 2019.

[5] M. Ali, P. Sapiezynski, M. Bogen, A. Korolova, A. Mislove, and A. Rieke, "Discrimination through optimization: How facebook's ad delivery can lead to skewed outcomes," *arXiv preprint arXiv:1904.02095*, 2019.

[6] L. Agarwal, N. Shrivastava, S. Jaiswal, and S. Panjwani, "Do not embarrass: re-examining user concerns for online tracking and advertising," in *Proceedings of the Ninth Symposium on Usable Privacy and Security*, pp. 1–13, 2013.

[7] M. Eslami, K. Vaccaro, K. Karahalios, and K. Hamilton, ""be careful; things can be worse than they appear": Understanding biased algorithms and users' behavior around them in rating platforms," in *Eleventh international aaai conference on web and social media*, 2017.

[8] C. Dolin, B. Weinshel, S. Shan, C. M. Hahn, E. Choi, M. L. Mazurek, and B. Ur, "Unpacking perceptions of data-driven inferences underlying online targeting and personalization," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, p. 493, ACM, 2018.

[9] T. Wasserman, "Facebook to hit 1 billion user mark in august," *Last retrieved on 12th of April*, 2012.

[10] T. Ermakova, B. Fabian, B. Bender, and K. Klimek, "Web Tracking-A Literature Review on the State of Research," 2018.

[11] A. Andreou, G. Venkatadri, O. Goga, K. Gummadi, P. Loiseau, and A. Mislove, "Investigating ad transparency mechanisms in social media: A case study of facebook's explanations," 2018.

[12] M. Eslami, S. R. Krishna Kumaran, C. Sandvig, and K. Karahalios, "Communicating algorithmic process in online behavioral advertising," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, (New York, NY, USA), Association for Computing Machinery, 2018.

[13] S. Stieglitz and L. Dang-Xuan, "Emotions and information diffusion in social media—sentiment of microblogs and sharing behavior," *Journal of management information systems*, vol. 29, no. 4, pp. 217–248, 2013.

[14] M. A. Bashir, U. Farooq, M. Shahid, M. F. Zaffar, and C. Wilson, "Quantity vs. quality: Evaluating user interest profiles using ad preference managers.," in *NDSS*, 2019.

[15] J. Knoll, "Advertising in social media: a review of empirical evidence," *International Journal of Advertising*, vol. 35, no. 2, pp. 266–300, 2016.

[16] A. A. Alalwan, N. P. Rana, Y. K. Dwivedi, and R. Algharabat, "Social media in marketing: A review and analysis of the existing literature," *Telematics and Informatics*, vol. 34, no. 7, pp. 1177–1190, 2017.

[17] T. Speicher, M. Ali, G. Venkatadri, F. Ribeiro, G. Arvanitakis, F. Benevenuto, K. Gummadi, P. Loiseau, and A. Mislove, "Potential for discrimination in online targeted advertising," 2018.

[18] S. Yuan, A. Z. Abidin, M. Sloan, and J. Wang, "Internet advertising: An interplay among advertisers, online publishers, ad exchanges and web users," *arXiv preprint arXiv:1206.1754*, 2012.

[19] S. Alavi, I. Mehdinezhad, and B. Kahshidinia, "A trend study on the impact of social media on advertisement," *International Journal of Data and Network Science*, vol. 3, no. 3, pp. 185–200, 2019.

[20] M. Wieringa, "What to account for when accounting for algorithms: A systematic literature review on algorithmic accountability," in *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, FAT* '20, (New York, NY, USA), p. 1–18, Association for Computing Machinery, 2020.

[21] M. Lecuyer, R. Spahn, Y. Spiliopolous, A. Chaintreau, R. Geambasu, and D. Hsu, "Sunlight: Fine-grained targeting detection at scale with statistical confidence," in *Proceedings of the 22nd ACM SIGSAC*, pp. 554–566, 2015.

[22] A. Datta, M. C. Tschantz, and A. Datta, "Automated experiments on ad privacy settings," *Proceedings on privacy enhancing technologies*, vol. 2015, no. 1, pp. 92–112, 2015.

[23] C. Iordanou, N. Kourtellis, J. M. Carrascosa, C. Soriente, R. Cuevas, and N. Laoutaris, "Beyond content analysis: Detecting targeted ads via distributed counting," in *Proceedings of the 15th International Conference on Emerging Networking Experiments And Technologies*, pp. 110–122, 2019.

[24] P. Barford, I. Canadi, D. Krushevskaja, Q. Ma, and S. Muthukrishnan, "Adscape: Harvesting and analyzing online display ads," in *Proceedings of the 23rd international conference on World wide web*, pp. 597–608, 2014.

[25] B. Weinshel, M. Wei, M. Mondal, E. Choi, S. Shan, C. Dolin, M. L. Mazurek, and B. Ur, "Oh, the places you've been! user reactions to longitudinal transparency about third-party web tracking and inferencing," in *Proceedings of the 2019 ACM SIGSAC*, pp. 149–166, 2019.

[26] R. F. Kizilcec, "How much information? effects of transparency on trust in an algorithmic interface," in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pp. 2390–2395, 2016.

[27] R. C. Nickerson, U. Varshney, and J. Muntermann, "A method for taxonomy development and its application in information systems," *European Journal of Information Systems*, vol. 22, no. 3, pp. 336–359, 2013.

[28] K. D. Bailey, *Typologies and taxonomies: An introduction to classification techniques*. No. 102, Sage, 1994.

[29] R. R. Hoffman, S. T. Mueller, G. Klein, and J. Litman, "Metrics for Explainable AI: Challenges and Prospects," *arXiv:1812.04608 [cs]*, Feb. 2019. arXiv: 1812.04608.

[30] A. Andreou, G. Venkatadri, O. Goga, K. Gummadi, P. Loiseau, and A. Mislove, "Investigating ad transparency mechanisms in social media: A case study of Facebook's explanations," 2018.

[31] A. Ghosh, G. Venkatadri, and A. Mislove, "Analyzing political advertisers' use of facebook's targeting features," in *IEEE Workshop on Technology and Consumer Protection (ConPro'19)*, 2019.

[32] M. Silva, L. Santos de Oliveira, A. Andreou, P. O. Vaz de Melo, O. Goga, and F. Benevenuto, "Facebook ads monitor: An independent auditing system for political ads on facebook," in *Proceedings of The Web Conference 2020*, pp. 224–234, 2020.

[33] J. G. Cabañas, Á. Cuevas, and R. Cuevas, "Unveiling and quantifying facebook exploitation of sensitive personal data for advertising purposes," in *27th {USENIX} Security Symposium ({USENIX} Security 18)*, pp. 479–495, 2018.

[34] P. G. Leon, B. Ur, Y. Wang, M. Sleeper, R. Balebako, R. Shay, L. Bauer, M. Christodorescu, and L. F. Cranor, "What matters to users? factors that affect users' willingness to share information with online advertisers," in *Proceedings of the ninth symposium on usable privacy and security*, pp. 1–12, 2013.

[35] I. Nunes and D. Jannach, "A systematic review and taxonomy of explanations in decision support and recommender systems," *User Modeling and User-Adapted Interaction*, vol. 27, no. 3-5, pp. 393–444, 2017.

[36] G. Venkatadri, A. Mislove, and K. P. Gummadi, "Treads: transparency-enhancing ads," in *Proceedings of the 17th ACM Workshop on Hot Topics in Networks*, pp. 169–175, 2018.

[37] M. Wieringa, "What to account for when accounting for algorithms: a systematic literature review on algorithmic accountability," in *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, pp. 1–18, 2020.

[38] B. Liu, A. Sheth, U. Weinsberg, J. Chandrashekar, and R. Govindan, "AdReveal: Improving Transparency into Online Targeted Advertising," in *Proceedings of the Twelfth ACM Workshop on Hot Topics in Networks*, HotNets-XII, (New York, NY, USA), pp. 12:1–12:7, ACM, 2013. event-place: College Park, Maryland.

[39] P. Barford, I. Canadi, D. Krushevskaja, Q. Ma, and S. Muthukrishnan, "Adscape: Harvesting and analyzing online display ads," in *Proceedings of the 23rd international conference on World wide web*, pp. 597–608, ACM, 2014.

[40] A. Datta, M. C. Tschantz, and A. Datta, "Automated experiments on ad privacy settings," *Proceedings on privacy enhancing technologies*, vol. 2015, no. 1, pp. 92–112, 2015.

[41] M. M.-d. Tran and C. Castelluccia, "Betrayed by your ads! reconstructing user profiles from targeted ads," in *The 12th (PETS 2012) Privacy Enhancing Technologies Symposium, Vigo, Spain*, 2012.

[42] P. Papadopoulos, N. Kourtellis, and E. Markatos, "Cookie synchronization: Everything you always wanted to know but were afraid to ask," in *The World Wide Web Conference*, pp. 1432–1442, ACM, 2019.

[43] A. Smith-Renner, R. Fan, M. Birchfield, T. Wu, J. Boyd-Graber, D. S. Weld, and L. Findlater, "No explainability without accountability: An empirical study of explanations and feedback in interactive ml," in *Proceedings of the 2020 CHI*, pp. 1–13, 2020.

[44] J. Parra-Arnau, J. P. Achara, and C. Castelluccia, "Myadchoices: Bringing transparency and control to online advertising," *ACM Transactions on the Web (TWEB)*, vol. 11, no. 1, pp. 1–47, 2017.

[45] S. Z. M. Zain, M. H. Selamat, *et al.*, "Derivation of hashtag (#) factors for hashtag marketing model (hashmam) in social media platform," in *2019 6th International Conference on Research and Innovation in Information Systems (ICRIIS)*, pp. 1–6, IEEE, 2019.

[46] B. Hu, Q. Lin, Y. Zheng, Q. Yan, M. Troglia, and Q. Wang, "Characterizing location-based mobile tracking in mobile ad networks," in *2019 IEEE Conference on Communications and Network Security (CNS)*, pp. 223–231, IEEE, 2019.

[47] A. A. Galán, J. G. Cabañas, Á. Cuevas, M. Calderón, and R. C. Rumín, "Large-scale analysis of user exposure to online advertising on facebook," *IEEE Access*, vol. 7, pp. 11959–11971, 2019.

[48] M. Eslami, S. R. Krishna Kumaran, C. Sandvig, and K. Karahalios, "Communicating algorithmic process in online behavioral advertising," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, p. 432, ACM, 2018.

[49] L. Dogruel, "Too much information Examining the impact of different levels of transparency on consumers' evaluations of targeted advertising," *Communication Research Reports*, vol. 36, pp. 383–392, Oct. 2019.

[50] E. Rader, K. Cotter, and J. Cho, "Explanations as mechanisms for supporting algorithmic transparency," in *Proceedings of the 2018 CHI*, pp. 1–13, 2018.

[51] T. Urban, M. Degeling, T. Holz, and N. Pohlmann, "" your hashed ip address: Ubuntu." perspectives on transparency tools for online advertising," in *Proceedings of the 35th Annual Computer Security Applications Conference*, pp. 702–717, 2019.

[52] H. Boateng and A. F. Okoe, "Consumers' attitude towards social media advertising and their behavioural response: The moderating role of corporate reputation," *Journal of Research in Interactive Marketing*, vol. 9, pp. 299–312, Jan. 2015.

[53] T. Kim, K. Barasz, and L. K. John, "Why Am I Seeing This Ad? The Effect of Ad Transparency on Ad Effectiveness," *Journal of Consumer Research*, vol. 45, pp. 906–932, Feb. 2019.

[54] K. Sokol and P. Flach, "Explainability fact sheets: a framework for systematic assessment of explainable approaches," in *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, pp. 56–67, 2020.

[55] T. H. Baek and M. Morimoto, "Stay away from me," *Journal of advertising*, vol. 41, no. 1, pp. 59–76, 2012.

[56] E. G. Smit, G. Van Noort, and H. A. Voorveld, "Understanding online behavioural advertising: User knowledge, privacy concerns and online coping behaviour in europe," *Computers in Human Behavior*, vol. 32, pp. 15–22, 2014.

[57] H. Soh, *Measuring trust in advertising: development and validation of the ADTRUST scale*. PhD thesis, University of Georgia, 2006.

[58] C. E. Tucker, "Social networks, personalized advertising, and privacy controls," *Journal of marketing research*, vol. 51, no. 5, pp. 546–562, 2014.

[59] A. Stevens and C. Newmeyer, "Creepy and intrusive: A consumer's perspective of online personalized communications," in *Contemporary Issues in Social Media Marketing*, pp. 172–183, Routledge, 2017.

[60] F. Li and P. W. Miniard, "On the potential for advertising to facilitate trust in the advertised brand," *Journal of Advertising*, vol. 35, no. 4, pp. 101–112, 2006.

[61] E. Herder and B. Zhang, "Unexpected and Unpredictable: Factors That Make Personalized Advertisements Creepy," in *Proceedings of the 23rd International Workshop on Personalization and Recommendation on the Web and Beyond*, pp. 1–6, ACM, 2019.

[62] C. A. Summers, R. W. Smith, and R. W. Reczek, "An audience of one: Behaviorally targeted ads as implied social labels," *Journal of Consumer Research*, vol. 43, no. 1, pp. 156–178, 2016.

[63] T. Dehling, Y. Zhang, and A. Sunyaev, "Consumer perceptions of online behavioral advertising," in *2019 IEEE 21st Conference on Business Informatics (CBI)*, vol. 1, pp. 345–354, IEEE, 2019.

[64] S. Youn and W. Shin, "Teens' responses to facebook newsfeed advertising: The effects of cognitive appraisal and social influence on privacy concerns and coping strategies," *Telematics and Informatics*, vol. 38, pp. 30–45, 2019.

[65] J. Strycharz, G. van Noort, E. Smit, and N. Helberger, "Consumer view on personalized advertising: Overview of self-reported benefits and concerns," in *Advances in Advertising Research X*, pp. 53–66, Springer, 2019.

[66] H. R. Gaber, L. T. Wright, and K. Kooli, "Consumer attitudes towards instagram advertisements in egypt: The role of the perceived advertising value and personalization," *Cogent Business & Management*, vol. 6, no. 1, p. 1618431, 2019.

[67] J. Han, M. Drumwright, and W. Goo, "Native advertising: Is deception an asset or a liability?," *Journal of Media Ethics*, vol. 33, no. 3, pp. 102–119, 2018.

[68] S. Youn and S. Kim, "Understanding ad avoidance on facebook: Antecedents and outcomes of psychological reactance," *Computers in Human Behavior*, vol. 98, pp. 232 – 244, 2019.

[69] D. Belanche, I. Cenjor, and A. Pérez-Rueda, "Instagram Stories versus Facebook Wall: an advertising effectiveness analysis," *Spanish Journal of Marketing-Esic*, 2019.

[70] S. Youn and W. Shin, "Adolescents' responses to social media newsfeed advertising: the interplay of persuasion knowledge, benefit-risk assessment, and ad scepticism in explaining information disclosure," *International Journal of Advertising*, vol. 39, no. 2, pp. 213–231, 2020.

[71] S. Youn and S. Kim, "Newsfeed native advertising on facebook: young millennials' knowledge, pet peeves, reactance and ad avoidance," *International Journal of Advertising*, vol. 38, no. 5, pp. 651–683, 2019.

[72] E. Van den Broeck, B. Zarouali, and K. Poels, "Chatbot advertising effectiveness: When does the message get through?," *Computers in Human Behavior*, vol. 98, pp. 150–157, 2019.

[73] J. Jacobson, A. Gruzd, and Á. Hernández-García, "Social media marketing: Who is watching the watchers?," *Journal of Retailing and Consumer Services*, vol. 53, 2020.

[74] E. Van den Broeck, K. Poels, and M. Walrave, "An experimental study on the effect of ad placement, product involvement and motives on facebook ad avoidance," *Telematics and Informatics*, vol. 35, no. 2, pp. 470–479, 2018.

[75] A. A. Alalwan, "Investigating the impact of social media advertising features on customer purchase intention," *International Journal of Information Management*, vol. 42, pp. 65–77, 2018.

[76] A.-R. Jung, "The influence of perceived ad relevance on social media advertising: An empirical examination of a mediating role of privacy concern," *Computers in Human Behavior*, vol. 70, pp. 303–309, 2017.

[77] G. Friedrich and M. Zanker, "A taxonomy for generating explanations in recommender systems," *AI Magazine*, vol. 32, pp. 90–98, Jun. 2011.

[78] N. Tintarev and J. Masthoff, "Evaluating the effectiveness of explanations for recommender systems," *User Modeling and User-Adapted Interaction*, vol. 22, pp. 399–439", issn=, Oct 2012.

[79] T. Kim, K. Barasz, and L. K. John, "Why am i seeing this ad? the effect of ad transparency on ad effectiveness," *Journal of Consumer Research*, vol. 45, no. 5, pp. 906–932, 2019.

[80] D. Jannach and G. Adomavicius, "Recommendations with a purpose," in *Proceedings of the 10th ACM Conference on Recommender Systems*, pp. 7–10, 2016.

[81] B. P. Knijnenburg, S. Sivakumar, and D. Wilkinson, "Recommender systems for self-actualization," in *Proceedings of the 10th ACM Conference on Recommender Systems*, pp. 11–14, 2016.

A. Appendix
A.1. Classification Description

(Sub) Category	Property	Description
Objective facet (higher level goals of the explanations about the ads served.)		
Explanation Purposes	Transparency	Providing information to help users understand the inner workings of how ads are served on the platform
	Persuasiveness	Providing information that convinces users that the recommended ads are appropriate
	Self-Actualization	Promoting the exploration of preferences or learning about personal preferences versus the preferences of your "algorithmic self"
Stakeholder Goals	Acceptance	Providing information about predictions to increase the likelihood of users accepting the suggestions
	Education	Offering insights to users so they could better understand the system/platform
	Feedback	Providing an opportunity for two-way communication about the ad recommendation
Subjective factors	Trust	Being perceived as a platform that is trustworthy
	Control	Being perceived as a system that allows control over the ads that are recommended
	Perceived Transparency	Being perceived as a platform that provides a good understanding of the ad recommending process
	Usefulness	Being perceived as a useful platform
WHY facet (the reasons as to why the ads are served.)		
Context	Audience	The target audience of the ad.This may differ based on demographic characteristics such as age, gender, etc.
	Location	The user location for which an ad recommendation is made.
	Device Modality	The type of device for which an ad is served to
	Time	The time of day the user is on the platform
	Language	The language(s) that the user is known to use
Input Parameters	Inference Attribute	Information, attitudes, and behaviors inferred about the user based on their interaction with the platform that is used to target the ad.
	Stated Attribute	Information, attitudes, and behaviors explicitly provided by the user used to target the ad
Input Source	App or platform	Medium through which the users interact
	Advertisers	Parties who pay and place ads on the platform/app to attract potential clients/customers
	Partners and third parties	Parties that the platform/app developer works closely with to ensure the easy placement of ads and attraction of users to the platform.
HOW facet (how data is collected or processed.)		
Decision Inference Method	Machine learning	Data models reliant on various user attributes used to effectively serve ads based on user preferences
	Algorithm-independent	Decisions devoid of any machine learning models in the process of serving ads
	Human-made	Decisions made by humans on what ads to serve users
	Knowledge-based	Decisions made based on the research about users
WHAT facet (what is being displayed to users.)		
Content: Scope	Individual Ad	Advert of a product/service served to users
	Platform	Central application through which advertisers are able to reach users.
	Ecosystem	A collection of related products through which ads can be served
Content: Interpretability	Completeness	Explanations that include the complete reasoning behind how the ad was served.
	Correctness	The accuracy of the explanations offered.
	Specificity	The specific individual user attributes used to in the decision making of tar getting the ad.
	Personalization	Content is crafted to users' personal behavior
Presentation: Format	Natural Language	Text-based easy to understand language through which the explanations are written.
	Multimedia	Features graphics and icons that convey information
Presentation: Level of Engagement	Interactivity	Mediums through which users can offer feedback about the ads they see
	Actionability	Ads that call for particular actions from users.

Table 3. Description of the facets in the explanation classification.