A Literatural Overview on Explanations in Personalization and Possible Criterion for Designing a Short Explanation

Yifei Zhan

Abstract— This paper provides a literatural overview of explanations in personalization and recommender systems. It then introduced two standards/criterion on how to design and evaluate explanations in recommender systems. Which type of explanations the system should offer to users depends on the goal of the system. And in some scenarios, like in a driving context, pages of explanations might not be suitable. In this case, short explanations is preferred. In order to explore possible criterion of short explanations, two studies were introduced and what we have found is that **efficiency**, **persuasiveness** and **why** question-type-based explanations might be suitable for short explanations. However, some more studies need to be done in order to check the validity of other criterion in short explanations.

Index Terms—Explanation, Recommender system, Personalisation

1 Introduction

Personalization has already existed for a long time. The first emergence of this phenomena can be traced back to antiquity, when experienced merchants provided different customers with different products or services[1]. However, the real interest of personalization arose in the mid to late 1990s with the advancement of the Web technologies[11]. From then on, more efforts have been investigated in this field in order to decode the essence of personalization.

In the world of personalization, the most common form of it is recommender systems (RS), which can intelligently give users personalized suggestions. By adopting different algorithms, the recommender systems can also be different. Content-based recommendation and collaborative filtering recommendation are two popular types of recommendation, on which large amount of researches have been done. The first one, content-based recommendation, observes utility of items experienced previously and their attributes for a given user is taken as input to predict the utility of other items to that user[9]. A simpler explanation can be "User A has bought a novel. And based on this history, a magazine is more likely to be recommended to him than a pair of shoes, because novels and magazines are more in common than novels and shoes ". The second type, collaborative filtering, observes utility for user-item pairs is taken as input to predict utility for unobserved user-item pairs[9]. One example can be "User A(1,2) has bought items 1 and 2, user B(1,2,3) has bought items 1,2 and 3 and user C(2,4) has bought items 2 and 4. Thus, item 3 is more likely to be recommended to user A than item 4, because user A and user B are more similar than user A and user C".

A good recommendation algorithm is no doubt indispensable for a recommender system and the way of explaining its complex algorithm to users plays also an important role. With the development of the Internet, the privacy issue gains more and more importance in the society and people are willing to know which information is "consumed" by the system. They may lose their trust on the system if it behaves another way as they expected. In this case, a good explanation of the behaviors of a recommender system can help inspire user trust and loyalty, increase satisfaction, make it quicker and easier for users to find what they want[25].

The following content are divided into four parts. The first part provided an overview of different researches on explanations in recommender systems. The second part focused on two possible methods or standards, which are recently often used by researches to evaluate and design explanations of recommender systems. while the third extracts some factors from these two methods, which can be used to design "short explanations" in automotive scenario. And the last part discussed on some limitations in recent researches, proposed several research questions for future research and provided a summary of this whole paper.

- *Yifei Zhan is studying Media Informatics at the University of Munich, Germany, E-mail: yifei.zhan@campus.lmu.de*
- *This research paper was written for the Media Informatics Advanced Seminar ' Hauptseminar Medieninformatik', 2017*

2 Overview

User experience research is increasingly attracting researchers attention in the recommender system community. Some researchers, like Li Chen[21] and Bart P. Knijnenburg[15], focus on a larger scale and try to figure out what constitutes an effective and satisfying recommender system. Other researchers focus on a more detailed part, that is the explanation of the system. They have proved that a good explanation can boost users' trust towards the system and help them in decision making process[25][27][20].

David Mcsherry proposed a case-based reasoning (CBR) approach[19] to product recommendation that offers important benefits in terms of the ease with which the recommendation process can be explained and the system's recommendations can be justified.

Another perspective focused on how the soundness and completeness of the explanations impacts the fidelity of end users' mental models[17], where the soundness means the extend to which an explanation describes all of the underlying system and the completeness means how truthful each element in an explanation is with respect to the underlying system.

And Brian Y. Lim and Anind K. Dey dived into question types. In their work, Toolkit Support Intelligibility in Context-Aware Applications[18], they divided explanation types into eight factors, which are Why, Why Not, How To, What, What If, Inputs, Outputs and Certainty. More details are explained in later section.

Finally, a great breakthrough has been made by Nava Tintarev and Judith Masthoff. They have proposed what they called "seven possible advantages of an explanation facility"[25]. These seven advantages are often taken as a standard criterion to evaluate different types of explanation in other researchers' works.

3 Explanation Factors

In this part, we take a look at explanation factors and summarize two different methods or standards that are frequently used in recent researches when people start to design or evaluate explanations. And an example is also provided in order to illustrate how these methods are used in modern web application.

3.1 Seven Explanatory Criterion

Like what we have already mentioned in the overview, that the "seven explanatory criterion"[25](see table 1) from Nava Tintarev and Judith Masthoff is the most popular standard used to evaluate the explanation.

It is comprehensive since it summarized all advantages which explanation can bring to the system. Thus this "seven explanatory criterion" will be useful if we want to design explanations.

Aim	Definition
Transparency	Explain how the system works
Scrutability	Allow users to tell the system it is wrong
Trust	Increase users' confidence in the system
Effectiveness	Help users make good decisions
Persuasiveness	Convince users to try
Efficiency	Help users make decisions faster
Satisfaction	Increase the ease of use or enjoyment

Table 1: Explanatory criteria and their definitions

Although they may called differently in different research works ("explanation attributes"[7], "seven possible aims for explanations"[26] or "quality factors"[13]). The basic idea is the same,that is to make system more understandable by users.

Transparency: Transparency in a recommender system is related to the capability of a system to expose the reasoning behind a recommendation to its users[14]. A recommender systems without explanation works like a "black box", which may lost trust from user. Thus, transparency is considered as an important factor to build user's trust in the system[24].

Scrutability: Scrutability means, in short, allow users to tell the system it is wrong. It can also be seen as a kind of User Control, which allows users to correct reasoning from system[10].

Trust: Trust is sometimes linked with transparency. Studies have already shown that transparency and the possibility of interaction with recommender systems increases user trust[12][23].

Effectiveness: An effective explanation can help users to make a better decision. If an item suggested by the system is the one the user really likes, such explanation can be considered as effective.

Persuasiveness: Persuasiveness, sometimes referred to as promotion, is strongly related to effectiveness and can be defined as the ability of an explanation type to convince the user to accept or disregard certain items[13].

Efficiency: An explanation is usually considered as efficient when it helps the user to decide more quickly or when it helps to reduce the cognitive effort required in the decision process[13]. The most common approach to measure it is to look at the iteration time between users and recommender system before users reach their goals.

Satisfaction: The users overall satisfaction with a recommender system is assumed to be strongly related to the perceived quality of its recommendations and explanations[24].

What we need to bear in mind, is that, it is hard to design an explanation which obeys all seven criterion. In most of the time, we need to have a trade-off. For instance, an explanation that offers good transparency may impede efficiency as the user may spend time taking in explanations[25]. In reality, using which criteria depends on the goal of the system.

3.2 Question-type-based Explanation

Another possible way to design explanations in context-aware applications is to focus on different question types offered by the system. Brian Y. Lim and Anind K. Dey have proposed eight question-type-based explanations[18] This method is quite useful when we intend to design short explanations for automobiles which are mostly used in a context-aware environment.

These eight question types are:

Input: Explanations inform users which information sources are used to generate the recommendation.

Output: Explanations inform users which output/result is generated by the system.

What: System explains the current state to users (what it is now).

What If: Which results will be generated, if users alter his/her input values(change the behavior).

Why: Explanations inform users why the application derived its output value from the current (or previous) input values[18].

Why Not: Explanations inform users why another output/result is not chose.

How to: Explanations tell users what they can do so that they can get another output/result.

Certainty: Explanations tell users, how certain/uncertain the output/result is.

In addition to context-aware applications, we can also find some hints of these question types in some other scenarios, like in recommendation systems of e-commercial websites or in some online applications, which are not necessarily "context-aware". The next section will show how it may look like in Online Advertisements.

3.3 Example – Explanations in Google Advertisement Recommendation Service

As users browse the Internet, read their emails and shop online, they see ads. For example, Google Search results page may show ads similar to a search phrase a user just typed, or user's favorite blog may show interactive ads related to the content on the page. These kinds of ads ads on Google services and millions of websites and apps that partner with Google are Google ads[6]. According this description, we can know that Google Ads is a personalized advertisement service. Figure 1 shows an example of Google Ads.

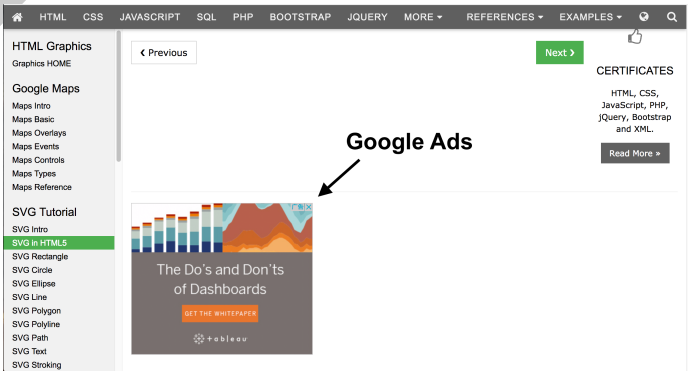

Figure 1: An example of Google Advertisement[5]

Not only the highly personalized service, Google Advertisement also provides users with pages of explanations. We can see here in figure 2. It is the first level of explanations and Google Ads mainly explains two things here. "how does it work" and "which information is used". The first one improves the **transparency** and increases users' **trust**, since it exposes the reasoning behind the recommendation to its customers. The second one tells users which information Google Ads uses in producing Advertisements. Such information can be seen as a **Input** type question because it explains users which information sources are used to generate the recommendation.

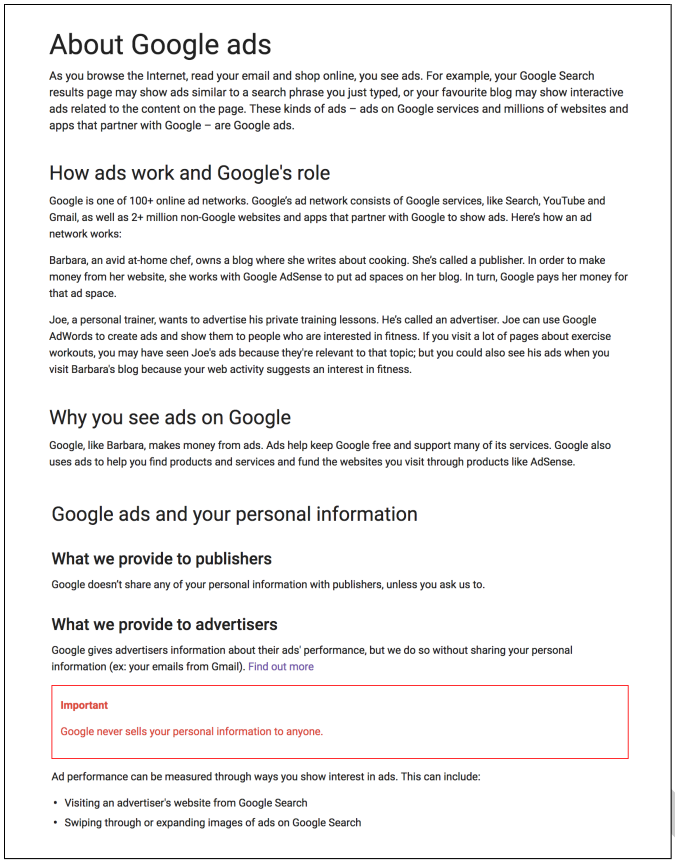

Figure 2: Explanations from Google Advertisement[2]

Figure 3 shows the second level of explanations in Google Ads. It again improves the **transparency** and increases users' **trust**. Meanwhile, "You may see ads on Google Search results pages and other Google services such as Google Maps. The ads you see may be based on what you searched for, your location and the time of day", such information can been seen as a **Why** and **Why Not** question types since it explains the reason why users see certain advertisements.

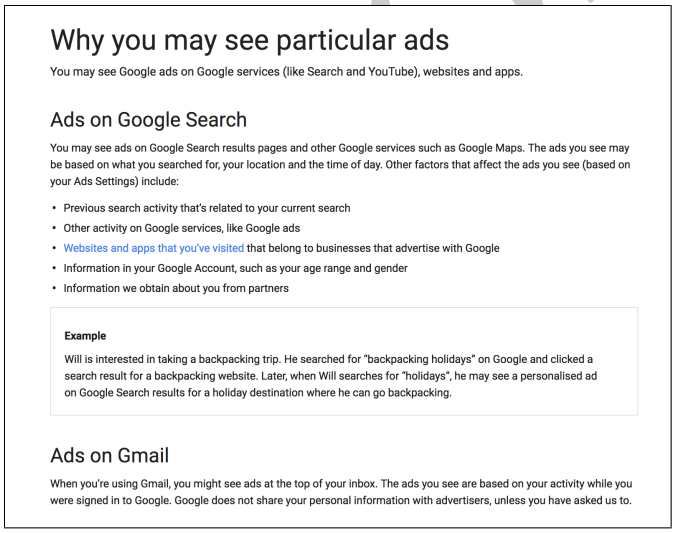

Figure 3: Explanations from Google Advertisement[3]

In the next level of explanation, it even gives users access to control their personalization. In figure 4, it divides the explanation into users' authority("What you can do"), users' controls("Your controls"), users' ability to turn on/off personalization("Ads Personalization") and how to alter the personalization("Where your Ads settings apply"). Such accessibility to control the personalization can be seen as a symbol of **Scrutability** since it allows users to avoid the Ads that they are not willing to see. And in the perspective of question-type-based Explanation, it provides users with **What If** and **How to** type explanations because it teaches users how to alter their personalization.

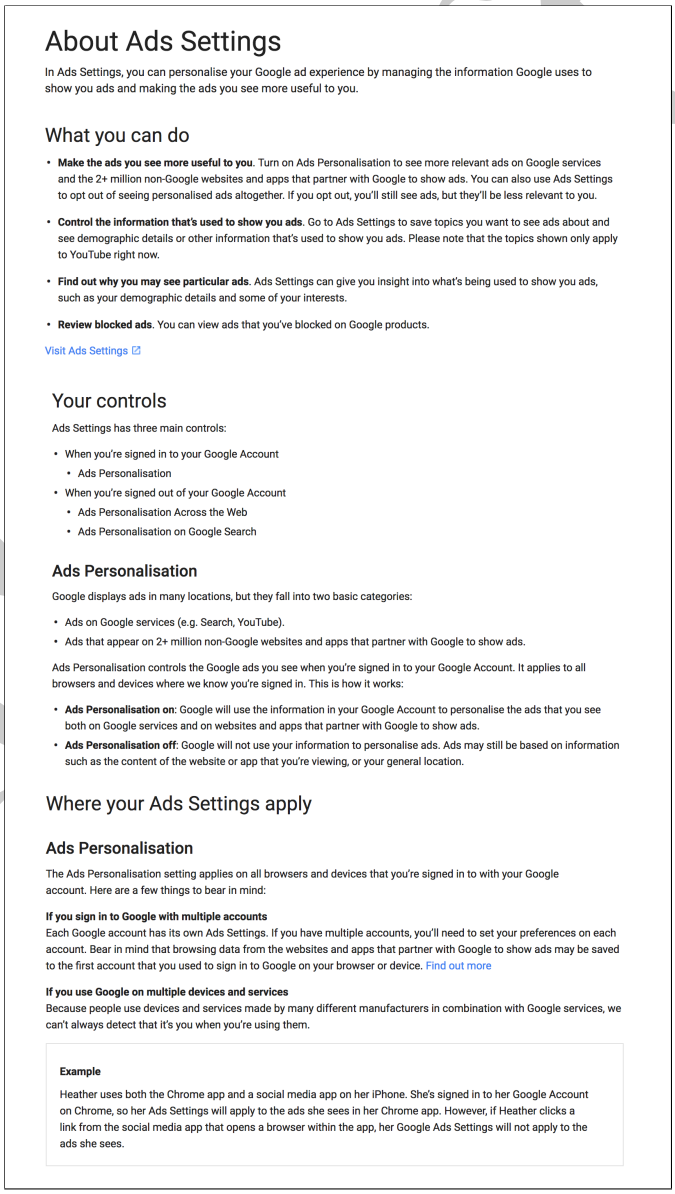

Figure 4: Explanations from Google Advertisement[4]

Although the explanation from Google Ads is transparent and comprehensive, it may not be suitable for some scenarios, like in a driving-context, where short explanation is preferred. The next part tries to figure out the connection between two standards(seven explanatory criterion and question-type-based explanation) and short explanations.

4 Short Explanation

In this section, we will dive into short explanation and introduce how this kind of explanations works in a driving-context and which criterion or factors are helpful when we try to design short explanation.

4.1 Short Explanation related to Seven Explanatory Criterion

Like it was already mentioned above, it is recommended to evaluate explanations with seven explanatory criterion. The same works with short explanations. Roland Bader and his colleagues have used this criterion and described a method based on knowledge-utility-based recommender systems to extract explanations automatically[8]. They proposed that proactively provided recommendations can reduce driver distraction while searching for information. However, such proactively delivered recommendations may not be accepted by the driver without a suitable explanation. Thus their goal was to enhance transparency of proactively delivered recommendations by means of explanations and tried to figure out the best way to explain.

They described a gas station recommender of a car and set up a user study with a desktop prototype, in which they provided explanations that were evaluated based on **efficiency** and **persuasiveness**. Besides what type of explanations should be presented to users, they also wanted to find out how long should an explanation be. Thus, they conducted a preliminary study in which they found out that **two arguments** seemed to be the a good size of explanation in the case of gas station(Arguments are like: (1)"this gas station provides gas with very low priced" or (2)"this gas station is on the route").

In their user study, explanations are distributed over several levels of detail. The lowest level (first phase) is provided automatically with the recommendations. Then gradually more and more information is accessible by the user manually. The elements in the first phase are short explanations for the situation and for the items[8].

In the phase of evaluation, the first criteria **persuasiveness** was measured by asking the subjects for their satisfaction with a selection in the first phase and if they need more information. And the second criteria **efficiency** was measured by looking at how often the subjects needed to switch to deeper phases with more information.

The result was as expected. Almost all the experimental subjects were satisfied with the first level explanation. And they found the information in this stage was adequate for decision-making in normal scenarios, which means that short explanation in this case meets the criterion of **persuasiveness**. The information in second level was selected only if special details were needed, e.g. an ATM or a shop, which means that short explanations in this study in this case meets the criterion of **efficiency**.

In conclusion, it seems that **persuasiveness** and **efficiency** play a quite important role in short explanation.

4.2 Short Explanation related to Question Type

Another possible approach to design short explanations is to derive explanations from different question types(Why, Why Not, How To, What, What If, Inputs, Outputs and Certainty). By Using a driving simulator with an auto-braking function, Martin Steinert and Larry John Leifer have explored how the content of the verbalized message accompanying the cars autonomous action affects the drivers attitude and safety performance[16].

Two types of information were provided to explain the auto-braking behavior to the driver during driving(simulated). The first one is "how" information, which tells the driver the current state of cars(like: "the car is braking"). We need to notice that this kind of explanation equals to **"what"** question type in Brian's Toolkit[18]. Thus we will exchange "how" to "what" in later descriptions. The second one is **"why"** information, which tells the driver the reason why it behaves in certain ways(like: "Obstacle ahead.").

In their experiment, they employed a two-by-two between-participants experimental design(see table 2). In table 2, external information is like "Obstacle ahead" while internal information described car activity("The car is braking"). The autonomous action(automatically braking) took place during driving(simulated) and each time, one of the four different types of explanations would be presented:

1. No explanation.
2. **"what"** explanation("The car is braking") short before the automatically braking.
3. **"why"** explanation("Obstacle ahead") short after the automatically braking.
4. A combination of **"what"** and **"why"** explanation("The car is braking due to obstacle ahead").

	Why message	**without Why message**
what message	Both internal and external information referring to automation	Internal (car activity) information
Without what message	External (situation) information	No information (control condition)

Table 2: Structure of study[16]

Drivers' attitudes and safety performance would be assessed in order to learn how different types of information affect the driver.

The result showed something different than the original hypothesis. A combination of explanation(**"why"** + **"how"**) was supposed to improve driving behaviors. However, the result indicated that it affected driver attitude negatively. When people were told both how and why the car was about to act on their behalf, they felt anxious. According to Martin Steinert and Larry John Leifer. One possible reason for this result was that the driver must process two types of information(the machines status and the situational status) at the same time, which created large cognitive load. Although it was perceived negatively, the combination of explanation contributed to safer driving by minimizing off-lane excursions. Thus for further studies, we need to anticipate a potential design trade-off between attitudinal preferences and driving safety.

By only providing **"how"** explanation, drivers performed the worst: they drifted out of their lane. It seems that drivers felt like they took an passive role in driving if they only received **"how"** explanation. Another possible reason is that we human beings expect a polite machine behaviors[22]. And only explaining the cars behavior, "Car is braking," without explaining the reason might be impolite.

"Why" explanation alone created the least anxiety and highest trust. This kind of behavior from the system can be seen as **transparent**, which can in turn gain more **trust** from users. Compared with the combination of explanation(**"why"** + **"how"**), the amount of time that it takes to hear and cognitively process the why-only message is shorter, which allowed a quicker response/reaction from the driver. Besides, drivers might find the data valuable and credible when the data content satisfies their cognitive need (explain the system behavior). For this reason, the **"why"** explanation is more important to drivers and safer because drivers can anticipate upcoming events[16].

5 Discussion and Limitations

With the development of automobile industry and mobile markets, more and more recommender systems need to work in a context-aware environment and compared with the desktop environment, users can not fully concentrate on the device because they also need to pay attention to environment at the same time. In such cases, a clear and brief explanation which informs the system's behaviors to users is especially important.

However, unlike the general explanations of recommender systems (explanations for desktop recommender systems), which has already some "standard" methods (seven explanatory criterion or question-based explanations). There is no "standard" approach to design and evaluate short explanations.

And compared with explanations of a recommender system in desktop scenario, which can contain pages of explanations, the explanation for automobile scenario is quite limited in the perspective of information amount. Thus, we need to extract some criterion from the already existed standards of general explanations and check their validity for short explanations.

In Roland Bader's study, although he has successfully extracted two criterion (**efficiency** and **persuasiveness**) for short explanations, he did not explain why he choose these two criterion specifically. We still do not know whether some other criterion from the seven explanatory criterion are also suitable for short explanations. Thus, it might be helpful if further studies can examine and evaluate other criterion (**transparency**, **scrutability**, **effectiveness**, **trust** and **satisfaction**) for short explanations.

The same goes with the study of Martin Steinert. Only two types of explanations (**what** and **why**) were evaluated. Therefore, some further studies can focus on other question-type-based explanations (**why not**, **how to**, **what if**, **inputs**, **outputs** and **certainty**).

6 Conclusion

In this paper, it firstly provided a literatural overview of explanations in personalization and recommender systems. Then two kind of standards were introduced in order to gain a better understanding of the methods used to design and evaluate explanations. Thirdly, an example of explanations on Google Advertisement was presented so that readers can understand how these two methods/standards look like in reality. In the next, we introduced two studies of short explanations in automobile scenario which extracted some useful information about how to especially design and evaluate short explanations. In the end, we discussed the limitations of both two studies on short explanations and proposed some possible directions for future studies.

References

[1] Gediminas Adomavicius, Zan Huang, and Alexander Tuzhilin. "Personalization and recommender systems". In: *State-of-the-Art Decision-Making Tools in the Information-Intensive Age*. INFORMS, 2008, pp. 55–107.

[2] Google Advertisement. *Explanations of personalization from Google Advertisement, the first level*. A snapshot, [Online; accessed 1-Dec-2017]. 2017. URL: https://support.google.com/ads/answer/1634057?hl=en-GB&ref_topic=7049263.

[3] Google Advertisement. *Explanations of personalization from Google Advertisement, the second level*. A snapshot, [Online; accessed 2-Dec-2017]. 2017. URL: https://support.google.com/ads/answer/7029660.

[4] Google Advertisement. *Explanations of personalization from Google Advertisement, the third level*. A snapshot, [Online; accessed 2-Dec-2017]. 2017. URL: https://support.google.com/ads/answer/2662856.

[5] Google Advertisement. *Google Advertisement in w3cschool website*. a snapshot of google advertisement, [Online; accessed 1-Dec-2017]. 2017. URL: https://www.w3schools.com/graphics/svg_inhtml.asp.

[6] Google Advertisement. *Introduction of Google Advertisement*. [Online; accessed 1-Dec-2017]. 2017. URL: https://support.google.com/ads/answer/1634057?hl=en-GB&ref_topic=7049263.

[7] MOHAMMED Z Al-Taie. "Explanations in recommender systems: overview and research approaches". In: *Proceedings of the 14th International Arab Conference on Information Technology, Khartoum, Sudan, ACIT*. Vol. 13. 2013.

[8] Roland Bader12, Andreas Karitnig, Wolfgang Woerndl, and Gerhard Leitner. "Explanations in proactive recommender systems in automotive scenarios". In: *First International Workshop on Decision Making and Recommendation Acceptance Issues in Recommender Systems (DEMRA 2011)*. 2011, p. 11.

[9] Marko Balabanović and Yoav Shoham. "Fab: content-based, collaborative recommendation". In: *Communications of the ACM* 40.3 (1997), pp. 66–72.

[10] Marek Czarkowski and Judy Kay. "A scrutable adaptive hypertext". In: *International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems*. Springer. 2002, pp. 384–387.

[11] Peppers Don and Rogers Martha. *Managing customer relationships: a strategic framework*. 2011.

[12] Alexander Felfernig, Erich Teppan, and Bartosz Gula. "Knowledge-based recommender technologies for marketing and sales". In: *International Journal of Pattern Recognition and Artificial Intelligence* 21.02 (2007), pp. 333–354.

[13] Fatih Gedikli, Dietmar Jannach, and Mouzhi Ge. "How should I explain? A comparison of different explanation types for recommender systems". In: *International Journal of Human-Computer Studies* 72.4 (2014), pp. 367–382.

[14] Jonathan L Herlocker, Joseph A Konstan, and John Riedl. "Explaining collaborative filtering recommendations". In: *Proceedings of the 2000 ACM conference on Computer supported cooperative work*. ACM. 2000, pp. 241–250.

[15] Bart P Knijnenburg, Martijn C Willemsen, Zeno Gantner, Hakan Soncu, and Chris Newell. "Explaining the user experience of recommender systems". In: *User Modeling and User-Adapted Interaction* 22.4-5 (2012), pp. 441–504.

[16] Jeamin Koo, Jungsuk Kwac, Wendy Ju, Martin Steinert, Larry Leifer, and Clifford Nass. "Why did my car just do that? Explaining semi-autonomous driving actions to improve driver understanding, trust, and performance". In: *International Journal on Interactive Design and Manufacturing (IJIDeM)* 9.4 (2015), pp. 269–275.

[17] Todd Kulesza, Simone Stumpf, Margaret Burnett, Sherry Yang, Irwin Kwan, and Weng-Keen Wong. "Too much, too little, or just right? Ways explanations impact end users' mental models". In: *Visual Languages and Human-Centric Computing (VL/HCC), 2013 IEEE Symposium on*. IEEE. 2013, pp. 3–10.

[18] Brian Y Lim and Anind K Dey. "Toolkit to support intelligibility in context-aware applications". In: *Proceedings of the 12th ACM international conference on Ubiquitous computing*. ACM. 2010, pp. 13–22.

[19] David McSherry. "Explanation in recommender systems". In: *Artificial Intelligence Review* 24.2 (2005), pp. 179–197.

[20] Pearl Pu and Li Chen. "Trust-inspiring explanation interfaces for recommender systems". In: *Knowledge-Based Systems* 20.6 (2007), pp. 542–556.

[21] Pearl Pu, Li Chen, and Rong Hu. "A user-centric evaluation framework for recommender systems". In: *Proceedings of the fifth ACM conference on Recommender systems*. ACM. 2011, pp. 157–164.

[22] Byron Reeves and Clifford Nass. *How people treat computers, television, and new media like real people and places*. CSLI Publications and Cambridge university press, 1996.

[23] Rashmi Sinha and Kirsten Swearingen. "The role of transparency in recommender systems". In: *CHI'02 extended abstracts on Human factors in computing systems*. ACM. 2002, pp. 830–831.

[24] Kirsten Swearingen and Rashmi Sinha. "Interaction design for recommender systems". In: *Designing Interactive Systems*. Vol. 6. 12. 2002, pp. 312–334.

[25] Nava Tintarev and Judith Masthoff. "A survey of explanations in recommender systems". In: *Data Engineering Workshop, 2007 IEEE 23rd International Conference on*. IEEE. 2007, pp. 801–810.

[26] Nava Tintarev and Judith Masthoff. “Evaluating the effectiveness of explanations for recommender systems”. In: *User Modeling and User-Adapted Interaction* 22.4 (2012), pp. 399–439.

[27] Jeroen Van Barneveld and Mark Van Setten. “Designing usable interfaces for TV recommender systems”. In: *Personalized Digital Television: Targeting Programs to Individual Viewers, Kluwer Academic Publishers, Dordrecht, The Netherlands* (2004), pp. 259–285.