SWE 577 Report

Ömer Faruk Çevik

2021719039

Table of Contents

[Introduction 1](#_Toc92110107)

[Motivation 3](#_Toc92110108)

[Framework 3](#_Toc92110109)

[Terms and Definitions: 4](#_Toc92110110)

[Reference Authors of Academic Papers 6](#_Toc92110111)

[Research Methodology: 6](#_Toc92110112)

[References 7](#_Toc92110113)

**Title: Explainable Recommendation Systems**

# Introduction

Internet usage increased dramatically in recent decades [1]. People have been spending more time on the internet and on what is called social media in general [2]. The people who use the internet are exposed to a number of suggested news, articles and videos depending on the platform they are visiting. Almost all websites including search engines, shopping portals, newspaper sites, learning portals and social media use recommender systems [11]. Actually, recommender systems (RS) are not new, and not unique to websites or information technologies in general. Humanity used recommendations of certain people or systems throughout history. Modern RS are inspired by basic human behavior; "individuals often rely on recommendations provided by others in making routine, daily decisions" [12,13,14]. However, in the information technology field, academic papers started to be published and RS emerged as an independent research area in the mid-1990s, in addition to the other classical information system tools and techniques (e.g., databases or search engines) [12,16,17].

The technology giants developed social media platforms which enabled creation of content or items, and publishing them to the whole world by anyone. The content or item is suggested to the other users of the platform depending on algorithm and user interactions. There are nine types of social media platforms [15].

|  |  |  |  |
| --- | --- | --- | --- |
| No | Type | Examples | Used For |
| 1 | Social audio platforms | Clubhouse, Twitter Spaces, Spotify | Listening to live conversations on specific topics. |
| 2 | Video social media platforms | YouTube, TikTok, Instagram Stories and Reels, Facebook Watch | Watching videos in short and long formats. |
| 3 | Disappearing content formats | Snapchat, Instagram Stories, Facebook Stories, LinkedIn Stories | Sending ephemeral messages privately and publishing timely, in-the-moment content for all of your followers to view for up to 24 hours. |
| 4 | Discussion forums | Reddit, Quora | Asking and answering questions, networking, forming communities around niche- and interest-based topics. |
| 5 | Shoppable social media platforms and features | Pinterest Product Pins, Facebook Shops, Instagram Shops, TikTok, Shopify, Douyin, Taobao | Researching and purchasing products from brands directly through social media platforms. |
| 6 | Social media live streams | Twitch, YouTube, Instagram Live Rooms, Facebook Live, TikTok | Broadcasting live video to many viewers. Live video streams can range from one person showing themselves and what they’re doing on their screen to professionally organized panels with multiple speakers. |
| 7 | Business social media platforms | LinkedIn, Twitter | Connecting with professionals in your industry or potential clients. |
| 8 | Closed/private community social media platforms | Discourse, Slack, Facebook Groups | Creating communities, with the possibility of requiring registration or other screening measures for new members. |
| 9 | Inspirational social media platforms | Pinterest, YouTube, Instagram, blogs | Searching for information and finding inspiration for anything from cooking to travel to decorating to shopping and more. |

Table 1: Types of Social Media Platforms [15]

Although some of the content in social media is created by experts of that topic, there are dozens of misinformation, manipulation and fraud-oriented content. The recommendation system or recommender system participates in suggestions to users in those platforms. The recommended content can either come from the user's activities (known as post-hoc) or from an algorithm [7]. What is more, there are even recommendation systems based on knowledge graphs which associate content based on properties or metadata of the content such as director, actors and category of the movie [5]. Besides, multi-agent system technologies can provide recommendations by combining trust and argumentation [10]. The recommendation system’s artificial intelligence (also AI in general) is criticized for their black-box design, but adding fairness, accountability, transparency, and explain ability (FATE) are linked to AI to increase liability [8]. The importance of explaining the recommended content to users of the platform by categorizing recommendation into the 5W (what, when, who, where, and why) has been researched by academicians [7].

Explainability is subjected to be added to software development processes as a nonfunctional requirement [9]. Moreover, even inspection and approval of such recommendation algorithms by independent auditors will be in question in near future [9]. Especially after the exposure of using social media for manipulation of the US election in 2016 [3,4] and fake social media groups managed by foreign troll factories [6].

# Motivation

Actually, the explainability issue is not limited to recommendation systems in information technologies. It has been more than half a century since the development of the first software, but there is not even an agreed definition on what explainability is in the literature. In other words, not only recommender systems but also software development in general lack of explainability. Moreover, Köhl et al.(2019)[29] stated that, in addition to lack of consensus on definition of explainability in software development, there is no metrics to evaluate a system’s explainability performance, no explicit specification of explainability to take them in to account during development, the concept itself and which techniques are appropriate remains underspecified, even domains experts struggle to understand certain aspect of a system. What is more, the explainability of the overall system can be satisficed, not satisfied. In the bargain, Larissa Chazette, together with other researchers, also proposed three academic studies related to adding explainability to the software development as a non-functional requirement in the years 2019, 2020 and 2021 [9],[30],[31]. Chazette also demonstrated that explainability is currently under-researched in the domain of requirements engineering and there is a lack of conceptual models and knowledge catalogs that support the requirements engineering process and system design [30]. The lack of explainability in general is added to the complex nature of recommender systems as motivation for the research.

# Framework

First of all, the explainability should be defined. After an agreed explainability definition, the definition should be tailored to recommender systems. Of course, we need definition of a non-functional requirement, and it is stated as a constraint on how such services (the functional requirements) should be provided. According to Köhl et al. (2019) [29], the explanation varies according to stakeholder type. Meaning that an explanation to an engineer might mean nothing to an end user. Köhl et al. [29] stated that the explanation is ‘answer to certain questions, in particular Why question’. Because it contains the form and attributes defined by technical accounts, a response to a question can constitute an explanation. It's also important to consider the context in which an explanation is given. Köhl et al. [29] proposed three possible definition:

* *Definition 1 (Explanation For): E* is an *explanation* of explanandum *X* with respect to aspect *Y* for target group *G*, in context *C*, if and only if the processing of *E* in context *C* by any representative1 *R* of *G* makes *R* understand *X* with respect to *Y*
* *Definition 2 (Explainable System):* A system *S* is *explainable* by means *M* with respect to aspect *Y* of an explanandum2 *X*,for target group *G* in context *C*, if and only if *M* is able to produce an *E* in context *C* such that *E* is an explanation of *X* with respect to *Y* , for *G* in *C*.
* *Definition 3* (Explainability Requirement): A system S must be explainable for target group G in context C with respect to aspect Y of explanandum X.

Secondly, elicitation of explanation should be performed by tailoring general explainability guidelines. Köhl et al. [29] offered three questions for elicitation of explainability. Again, those questions are not limited to recommender systems, they are applicable in all systems.

1) What are the relevant target groups G, e.g., engineers, end users, or lawyers, and which traits characterize each group’s representatives R, e.g., specific background knowledge or cognitive capacities?  
2) What are the explananda X, e.g., events or decisions?  
3) Which aspects Y of the explananda X must be explained to which target group G, e.g., why is a decision justified, which causal chain of internal system events led up to it, why did some event e happen instead of event e?  
4) In which context C may an aspect Y need explanation, and what are the implied constraints? For example, explanations might have to be aural in a driving situation

Thirdly, after elicitation of explainability, specification of explainability comes into the scene. Köhl et al. [29] suggested Softgoal Interdependency Graphs [T12] for decomposition and illustration of explainability requirements during specification. This modeling tool should be used according to the recommender system’s application domain. In other words, the diagram depends on the nature of the platform, stakeholder types, purpose and events.

Lastly, verification or validation of explainability should be performed. Recall from Köhl [29], Explainability of the overall system is then satisficed by satisficing the resulting explainability sub-softgoals. The authors of the paper suggest selecting a sample of a particular stakeholder with different characteristics such as selecting lawyers of different age, gender, experience with a given problem, and provide them with explanations. Then, collecting feedback from the participants to check whether or not they understood the explanation in the desired way. It seems necessary to gain deeper insights into the representatives’ processing of explanations. For instance, one could use the think-aloud technique trying to understand how people perceive a given explanation. After processing the explanation, the representatives could try to use self-explanation to answer their own questions based on the explanation.

# Terms and Definitions:

T1: Recommendation Systems or Recommender Systems (RS): There are dozens of definitions for the term. Two of them were selected. The other definitions are listed in reference 22 link.

T1.1 ‘Recommender systems are information filtering systems that deal with the problem of information overload [20] by filtering vital information fragment out of a large amount of dynamically generated information according to user’s preferences, interest, or observed behavior about item [21]’ [19].

T1.2 A recommender system or a recommendation system is a subclass of information filtering system that seeks to predict the “rating” or “preference” that a user would give to an item [22].

T2: Social Media: Social media are interactive technologies that facilitate the creation and sharing of information, ideas, interests, and other forms of expression through virtual communities and networks [23].

T3: Technology Giants: Or Big Tech is a term that refers to the most dominant and largest technology companies in their respective sectors. Although most of the companies are in the US, there are big tech companies in other countries like China [24].

T4: Search Engine: A search engine is a software program that helps people find the information they are looking for online using keywords or phrases [25].

T5: Artificial Intelligence: The theory and development of computer systems able to perform tasks normally requiring human intelligence, such as visual perception, speech recognition, decision-making, and translation between languages [26].

T6: Content: Or “Item” is the general term used to denote what the system recommends to users [15]. It can be video, news, articles, pictures, etc.

T7: Item: Synonym for content.

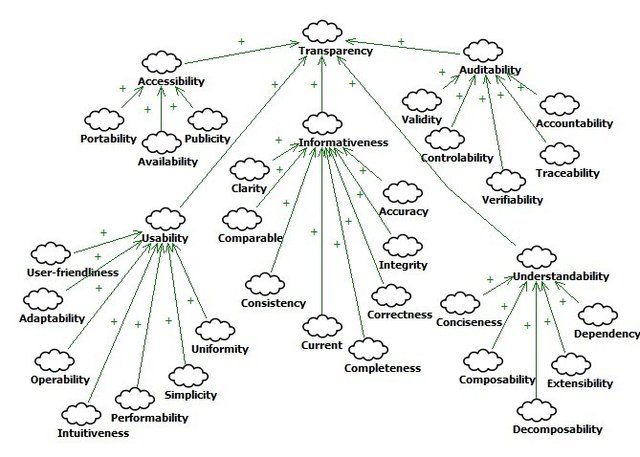
T8: Platform: The website, application or software which the RS is used in.

T9: Multi-agent system: allow many intelligent agents to interact with each other [27].

T10: Black-box design: a black box is a device, system, or object which can be viewed in terms of its inputs and outputs, without any knowledge of its internal workings.

T11: Troll factory: an organization set up in order to publish a large number of messages or posts on the internet, that often appear to be from people who do not really exist, and that are intended to cause trouble, influence political views, etc. [28]

T12: A graphical representation, proposed in (Chung et al. 2000), to model non-functional requirements by decomposing them on other non-functional requirements; specifying the interdependencies among them, and determining how to operationalize them. An example is below [32].



# Reference Authors of Academic Papers

Introduction to Recommender Systems Handbook by Francesco Ricci, Lior Rokach and Bracha Shapira (reference number 12) is a good start to be familiar with RS. That paper mentions the history and development of RS in addition to providing terms and definitions. It has references to fundamental studies in the references section.

Generally, papers can be classified as technical papers and research papers. The technical studies describe how to use the technology to develop a recommendation system by using a certain technology or combination of certain technology methods, and how, when, where to use a certain technology to add explainability to the RS. In other words, technical papers illustrate design or set-up of a RS, and its explainability. Also, technical researchers focus on selecting a technology according to the target set of items or content, nature of the website, and the user profile. For example, which method is superior if the recommender system will suggest books or associate uploaded videos based on their metadata. In technical perspective a researcher in the RS field should know about knowledge graphs, artificial intelligence, agent-based modeling. Research paper more focuses on explainability of the RS to the users, inspectors and other developers. Basically, fairness, accountability, transparency, and explain ability are top concerns of researchers.

International Requirements Engineering Conference (RE) held each year. The conference has researches about explainability in software general too. The explainability researches mostly focuses on creating norms for adding explainability to software development processes as a non-functional requirement. Those studies can be applied to recommender systems’ explainability.

Dedicated workshops and conferences related with RS: ACM Recommender Systems (RecSys) [11,18].

Traditional conferences includes RS in the area of databases, information systems and adaptive systems: ACM SIGIR Special Interest Group on Information Retrieval (SIGIR), User Modeling, Adap- tation and Personalization (UMAP), and ACM’s Special Interest Group on Man- agement Of Data (SIGMOD).[11,18]

The book “The People Vs Tech: How the internet is killing democracy” provides contemporary cases related to manipulations in social media.

# Research Methodology:

Mainly ieee, science direct and scholar.google.com used as search engines to search for published papers about explainable recommendation systems. The evolution of RS seems important. Firstly, the early years are searched. There are papers even in the 1950s about recommendation systems but all of them not about computer science, most of them about medicine and health in those early days. The RS in computer science emerged in the mid 1990s. Since then, there have been dozens of studies about the topic. Up to the mid-2000s, they were all about how to build recommendation systems. In other words, technical methods discussed to build RS are examined up-to mid-2000s. Since the mid-2000s explainability has been added to academic studies in addition to designing of RS.

In ieee, there are 66 studies for the search “explainable recommendation systems” in the date range 2016 and 2022. For the date interval 2016 and 2022, there are about 17.500 results for explainable recommendation systems search and 67 results for ”explainable recommendation systems” in scholar.google.com.

Yongfeng Zhang is among the top performing researchers (Computer Science, Rutgers University, TheUS).

<https://scholar.google.com/citations?user=A66WefUAAAAJ&hl=en&oi=sra>

Xu Chen is producing lots of researches in the year 2021 (Renmin University of China).

<https://scholar.google.com/citations?hl=en&user=loPoqy0AAAAJ&view_op=list_works&sortby=pubdate>

The universities in the USA, Israel, UAE and China have leading roles in the RS field. The universities in Germany focused on defining explainability and adding them as a non-functional requirement to software development processes. Köhl from Saarland University and Chazette from Leibniz University provided researches about expainability issues in information technologies. The two German researchers have high quality researches about general explainability issues related with software systems.

<https://scholar.google.com/citations?hl=en&user=xtjSVlcAAAAJ>

<https://scholar.google.com/citations?hl=en&user=zjxwI3kAAAAJ>

# References

1. Internet Growth Statistics

https://www.internetworldstats.com/emarketing.htm,2021,accessed

2.

https://www.statista.com/statistics/433871/daily-social-media-usage-worldwide/,2021,accessed

3.

https://www.nytimes.com/2018/07/23/books/putin-trump-russian-interference-election.htm/, 2018,accessed on 6 Nov 2021

4. Bartlett, J.,

The People Vs Tech: How the internet is killing democracy (and how we save it) 2018

5. Rose Catherine, Kathryn Mazaitis, Maxine Eskenazi, William Cohen, Explainable Entity-based Recommendations with Knowledge Graphs, [cs.IR] 2017

6. secession in Texas

https://edition.cnn.com/2017/10/05/politics/heart-of-texas-russia-event/index.html, 2017,accessed on 6 Nov 2021

7. Yongfeng Zhang and Xu Chen, “Explainable Recommendation: A Survey and New Perspectives”, Foundations and Trends® in Information Retrieval: Vol. 14, No. 1, pp 1–101. DOI: 10.1561/1500000066,2020

8. Donghee Shin, The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI, International Journal of Human - Computer Studies  
,2021

9. Chazette, L., &Schneider, K. (2020). Explainability as a non-functional requirement.

Require. Eng.10.1007/s00766-020-00333-1.

10. Punam Bedi and Pooja Vashisth, Empowering recommender systems using trust and argumentation, / Information Sciences 279 (2014) 569–586 (2014)

11. Lalita Sharma and Anju Gera, A Survey of Recommendation System: Research Challenges, nternational Journal of Engineering Trends and Technology (IJETT) - Volume4Issue5- May 2013 ISSN: 2231-5381

12. Francesco Ricci, Lior Rokach and Bracha Shapira, Introduction to Recommender Systems Handbook, DOI 10.1007/978-0-387-85820-3\_1, © Springer Science+Business Media, LLC 2011

13. Mahmood, T., Ricci, F.: Improving recommender systems with adaptive conversational

strategies. In: C. Cattuto, G. Ruffo, F. Menczer (eds.) Hypertext, pp. 73–82. ACM (2009)

14. McSherry, F., Mironov, I.: Differentially private recommender systems: building privacy into

the net. In: KDD ’09: Proceedings of the 15th ACM SIGKDD international conference on

Knowledge discovery and data mining, pp. 627–636. ACM, New York, NY, USA (2009)

15. [https://blog.hootsuite.com/types-of-social-media/#Types\_of\_social\_media\_platforms\_and\_formats\_you\_should\_know\_in\_2021](https://blog.hootsuite.com/types-of-social-media/" \l "Types_of_social_media_platforms_and_formats_you_should_know_in_2021) , accessed on 27.11.2021

16. Goldberg, D., Nichols, D., Oki, B.M., Terry, D.: Using collaborative filtering to weave an information tapestry. Commun. ACM **35**(12), 61–70 (1992)

17. Anand, S.S., Mobasher, B.: Intelligent techniques for web personalization. In: Intelligent Techniques for Web Personalization, pp. 1–36. Springer (2005)

18. <https://recommender-systems.com/publish/conferences/> accessed on 28.11.2021

19. F.O.IsinkayeaY.O.FolajimibB.A.Ojokohc, Recommendation systems: Principles, methods and evaluation, Egyptian Informatics Journal Volume 16, Issue 3, November 2015, Pages 261-273

20. J.A. Konstan, J. Riedl Recommender systems: from algorithms to user experience

User Model User-Adapt Interact, 22 (2012), pp. 101-123

21. C. Pan, W. Li Research paper recommendation with topic analysis In Computer Design and Applications IEEE, 4 (2010) pp. V4-264

22. <https://www.igi-global.com/dictionary/recommender-system/24730> accessed on 28.11.2021

23. wiki [https://en.wikipedia.org/wiki/Social\_media#cite\_note-Kietzmann-1](https://en.wikipedia.org/wiki/Social_media" \l "cite_note-Kietzmann-1) accessed on 28.11.2021

24. <https://whatis.techtarget.com/definition/Big-Tech> accessed on 28.11.2021

25. <https://www.bdc.ca/en/articles-tools/entrepreneur-toolkit/templates-business-guides/glossary/search-engine> accessed on 28.11.2021

26. Oxford dictionary accessed on 28.11.2021.

27. Michael Wooldridge, An introduction to multi-agent systems, Second Edition, 2009

28. Cambridge dictionary <https://dictionary.cambridge.org/dictionary/english/troll-factory> accessed on 28.11.2021.

29. Maximilian A Köhl, et al., Explainability as a Non-Functional Requirement, IEEE 27th International Requirements Engineering Conference (RE) (2019)

30. Larissa Chazette, et al, Exploring Explainability: A Definition, a Model, and a Knowledge Catalogue, IEEE 29th International Requirements Engineering Conference (RE) (2021)

31. Larissa Chazette, Wasja Brunotte, Timo Speith , Do End-Users Want Explanations? Analyzing the Role of Explainability as an Emerging Aspect of Non-Functional Requirements, IEEE 27th International Requirements Engineering Conference (RE) (2019)

32. <https://www.researchgate.net/profile/Roxana-Portugal-2/publication/319890469/figure/fig2/AS:542606352281600@1506378843378/Transparency-Softgoal-Interdependency-Graph-SIG_W640.jpg> accessed on 03.01.2021.