
Transparency and the Mediation of Meaning in Algorithmic Systems

Jesse Josua Benjamin

Human-Centered Computing
Institute of Computer Science
Freie Universität Berlin, Germany
jesse.benjamin@fu-berlin.de

Claudia Müller-Birn

Human-Centered Computing
Institute of Computer Science
Freie Universität Berlin, Germany
clmb@inf.fu-berlin.de

Rony Ginosar

Design Hybrids Lab
School of Computer Science
The Hebrew University of
Jerusalem, Israel
rony.ginosar@mail.huji.ac.il

Author Keywords

Post-phenomenology, qualitative research, explicitation interview.

Challenges in Algorithmic Systems

Today's algorithmic systems are increasingly ubiquitous, while advancements in machine learning, especially deep learning, have made them nearly incomprehensible in their functioning. With the growing presence of these algorithmic systems for supporting or even replacing human decision making, contemporary designers and engineers in both academia and industry are confronted with novel challenges, especially regarding the disclosure of the system's inner functioning. Design practitioners are already urged to consider explaining the presence and function of algorithms to end users [2]. However, the dynamic nature of algorithmic systems, as described by Redström and Wiltse [6] for example, complicates current design approaches that evolve around questions such as: What can (or should) be made present to users when considering algorithmic systems, while still maintaining an actionable visual presentation—the algorithm, the data, the confidence level of the results? These open questions highlight only some of the challenges that occur when considering transparency in algorithmic systems.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

Copyright held by the owner/author(s).

For our workshop contribution, we refer to philosopher of technology Verbeek, who argues that technology *mediates* human relations to the world (e.g., other humans, artifacts, values) [7]. Consequently, participation in this mediation is not only a concern for engineers and designers to be considered at the start of development. On the contrary, it is a continuous end in itself with a particular moral and ethical urgency [8]. In the case of algorithmic systems as a mediating technology, we argue that transparency is fundamental for sparking participation. Against this background, we seek to understand how meaning can be generated in a participatory manner, both in the interaction of users with an algorithmic system as well as through the development process of an algorithmic system.

To this end, we introduce an *expert explicitation interview* in this contribution; in the context of an algorithmic visualization system that employs a *topic modelling pipeline*. This particular interviewing method has previously been applied by researchers to understand decision-making processes in algorithmic governance [5] as well as to disclose the concrete experiences humans have with technologies [3]. Following a brief introduction into the role of the topic modelling pipeline in the system's development, we discuss preliminary insights we have gathered from our expert explicitation interview. In closing, we outline our aims in participating in the 'Participation+Algorithms' workshop.

Expert Explicitation Interview

In our algorithmic visualization system, we employ a topic modelling pipeline (TMP) to compute and visualize clusters of research project data. Our goal is to provide users with a visualization system on past and present research data, and to raise their awareness for potential knowledge transfer. However, during the development of our algorithmic visualization system, the role of the TMP changed. This

was triggered by the fact that we, as non-experts in the application domain of the visualization system (predominantly natural and life sciences), doubted that we were capable of an informed judgment on the meaning of the clusters generated by the TMP. Subsequently, we questioned how decisions on meaning had been made in the development of the TMP in the first place.

Taking inspiration from the discourse on algorithm awareness [1], we hypothesized that making the TMP interactive would allow a less-prescriptive definition of meaning, thus possibly allowing our non-technical experts to generate their own understanding. To identify possible interactions, we conducted an expert explicitation interview with the developer of the TMP. Our goal was to understand the chronology of decisions made in the development, and how the developer accounted for them. We assumed that this would provide us with actionable insights into how to make the TMP interactive. We will briefly describe the technical setup of the TMP. We then proceed to provide key quotes from the expert elicitation interview, followed by preliminary insights.

Prelude: The Topic Modelling Pipeline

The developer separated the TMP into four main components.¹ Firstly, the descriptions of the research project are transferred into a term frequency–inverse document frequency (tf-idf) matrix. In the next component, a Latent Semantic Analysis (LSA) is used to identify possible (n) topics. Subsequently, the clustering method k-means is employed to search the resulting word-space for k clusters. Finally, the developer uses either the t-Stochastic Neighbor Embedding (t-SNE) or Linear Discriminant Analysis (LDA) algorithm to compute a representation of research projects as points in a 2D-space.

¹A complete overview is given here: <http://bit.ly/nlpflowchart>

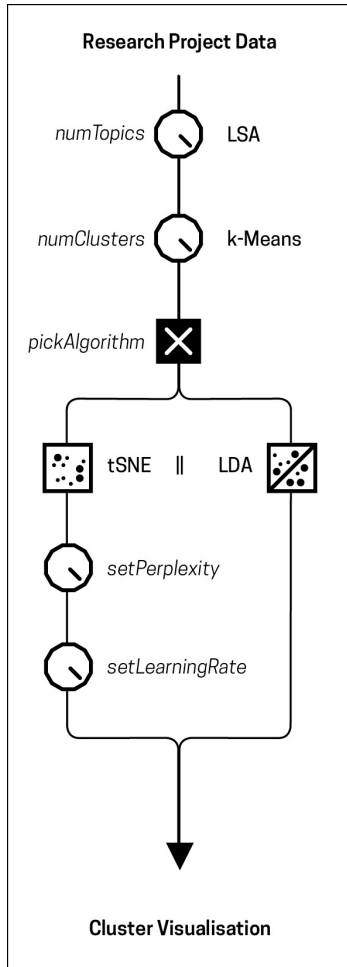


Figure 1: Possible access points for participation in the algorithmic generation of meaning in our Topic Modelling Pipeline.

Interview: Meaning in Developing an Algorithmic System
 Following Mager [5] and Light [3], we conducted a qualitative interview with the developer of the TMP. We asked general questions that aimed to disclose the chronology of decision making during the development process, spontaneously interspersed by questions on the specific circumstances (i.e., situation, location, emotion, reasoning). This combination is designed to elicit explicit and detailed statements from the developer of the TMP pipeline.

In the following, we showcase four key quotes from the interview. Firstly, when asked about a crucial choice in the development of the TMP, the use of the LDA, the developer expressed a preference for it:

LDA is a pretty nice method because you're able to actually create neat clusters because of its nature, so it's basically, um a way to reduce the dimensionality of a data set in order to for example visualize it.

The 'neat' nature of LDA was a re-occurring theme in the interview, and extended to how the developer assessed meaning in the TMP:

I was concerned with making the visualization neat and I liked it, because it somehow conveys the feeling of having the optimal set of parameters in order to extract the most information out of the data sets.

To the developer, this 'neat' quality of both the LDA algorithm and the resulting visualization stands in sharp contrast to employing the alternative algorithm for representing clusters in a 2D space, the tSNE algorithm, which was initially dismissed:

Um, I was not really happy with the results because it looked like [expletive] and there wasn't like really a meaning like behind it, at least not in my opinion.

Finally, when reflecting on possible new, interactive role of the TMP, the developer explicates the key difference:

Um, if you do it like the way I did it's basically a set extraction, extracting information and presenting it in a usable fashion to the user. The [interactive use] concerns itself with how do I generate knowledge from the person itself [...], it's not really about visualization itself. It's just [...] a helper in order to facilitate the process in the person itself [sic!].

Interview: Preliminary Insights

We employed open coding on the conducted explication interview. Firstly, we discovered how the interviewed developer has found meaning in the TMP. For the developer, meaning arises from a mathematically elegant solution. The solution was brought about by using the LDA algorithm. The result of this solution, i.e. the visualization of 'neat' clusters, is indeed a site for participation for users. However, participation in this case does not extend to the process of how the result has been generated. In this first version of the TMP, as information extraction, the developer was guided by his assessment of meaning. In the second, interactive use of the TMP, participation in the process became key. A conflict, then, lay in the developer's subjective assessment of meaning versus the necessity to provide interactive access to the TMP. As a result, the previously discarded tSNE algorithm became relevant again, because its parameters (e.g. perplexity and learning rate) could be

easily manipulated via frontend solutions, in contrast to the LDA algorithm. This leads us to consider possible "access points" [4] (Figure 1) for the TMP. As a result, we consider the paradox relationship between the developer's preferred meaningful solution, the LDA algorithm, and the potential for higher participation by our key stakeholders with the tSNE algorithm, especially relevant: it highlights how differently various stakeholders may assess meaning in algorithmic systems. This could potentially be operationalized in order to strengthen transparency, while at the same time allowing interactivity. For example, in our use case, "access points" can be provided if users can interactively choose n topics and k clusters, switch between the LDA and tSNE algorithms, and/or manipulate the parameter settings of the latter—thereby influencing how meaning is generated in and received from the visualization system.

At the Workshop

In our contribution to the 'Participation+Algorithms' workshop, we will present the theoretical foundation for our qualitative research, namely post-phenomenology as discussed by Verbeek. We see our qualitative research as a first step towards *mediating* meaning for various stakeholders in an algorithmic system. We then provide details on our interview technique and the first results, and outline future work. This will include further explicitation interviews, and bespoke research-through-design studies that aim to tackle the challenge of meaningful participation in algorithmic systems via "access points". We are particularly eager to discuss our approach with CSCW practitioners, as our research project ultimately aims at a holistic approach to collaborative sensemaking.

Acknowledgements

This work is supported by the German Federal Ministry of Education and Research, grant 03IO1633 ("IKON – Wis-

senstransferkonzept für Forschungsinhalte, -methoden und -kompetenzen in Forschungsmuseen").

REFERENCES

1. Kevin Hamilton, Karrie Karahalios, Christian Sandvig, and Motahhare Eslami. 2014. A Path to Understanding the Effects of Algorithm Awareness. In *Extended Abstracts CHI 2014*. ACM, 631–642.
2. Cliff Kuang. 2017. The Next Great Design Challenge: Make AI Comprehensible To Humans. (Dec. 2017). <http://bit.ly/2nH2WGN>
3. Ann Light. 2006. Adding method to meaning: A technique for exploring peoples' experience with technology. *Behaviour & Information Technology* 25, 2 (March 2006), 175–187.
4. Davor Löffler. 2018. Distributing Potentiality. Post-capitalist Economies and the Generative Time Regime. *Identities: Journal for Politics, Gender and Culture* 15, 1-2 (2018), 8–44.
5. Astrid Mager. 2018. Internet governance as joint effort: (Re)ordering search engines at the intersection of global and local cultures. *New Media & Society* (Feb. 2018), 1461444818757204.
6. Johan Redström and Heather Wiltse. 2015. Press Play: Acts of defining (in) fluid assemblages. *Nordes* 1, 6 (May 2015).
7. Peter-Paul Verbeek. 2006. *What things do : philosophical reflections on technology, agency, and design*. Penn State Press, University Park.
8. Peter-Paul Verbeek. 2011. *Moralizing Technology : Understanding and Designing the Morality of Things*. University of Chicago Press.