

Trust in algorithmic recommendations: Evidence from two survey experiments

Esther N. Nieuwenhuizen, Utrecht University, e.n.nieuwenhuizen@uu.nl

Albert J. Meijer, Utrecht University, a.j.meijer@uu.nl

Stephan G. Grimmelikhuijsen, Utrecht University, s.g.grimmelikhuijsen@uu.nl

Floris J. Bex, Utrecht University, f.j.bex@uu.nl

Keywords: explainability, transparency, artificial intelligence, police

Abstract

Public sector organizations are increasingly providing algorithmic recommendations to citizens to improve their service delivery (Androutsopoulou et al., 2019). Algorithms are used, for example, to support how citizens fill out online tax forms or how they file a report of online fraud to the police. These decision-support algorithms come with many assumed benefits, such as the reduction of administrative burden or more accurate decision making. However, these algorithms can also pose potential challenges to public organizations. They could lead to, for instance, a decline of citizen trust because decision-support algorithms often lack transparency (Gupta 2019).

A core criticism on this use of algorithms has been their lack of *explainability* (Giest and Grimmelikhuijsen 2020). Providing explainable algorithmic recommendations is considered crucial to foster trust amongst citizens and for them to follow up on what the algorithm recommends (Shin 2020). However, there is limited empirical evidence for this assumption in public sector service delivery.

To fill this knowledge gap, we constituted the following research question: *What are the effects of different explanations on trust in algorithmic recommendations?*. In addition to contributing to the academic literature on citizen trust in algorithm use by public organizations, we aim to provide public organizations with practical knowledge on how to make their algorithm use more trustworthy.

Previous research shows that public organizations can be transparent about a decision process -transparency in process- and transparent about the motivation of a decision -transparency in rationale- or a combination of these two. A growing number of researchers are adopting the same distinction for explaining algorithmic decisions or recommendations (e.g. de Fine Licht and de Fine Licht 2020). We will build on this distinction by testing the effects of a procedural, a rationale and a combined explanation on trust in algorithmic recommendations.

We employed a mock version of a real algorithmic recommendation system: the intelligent crime reporting tool of the Dutch police. Participants will be asked to file a report of online fraud through this mock recommendation system. They will be randomly assigned to one of the experimental groups. After seeing an algorithmic recommendation (for every participant exactly the same) followed by their treatment-explanation, they have to decide whether or not they want to file a report of online fraud.

We will present a paper with independent two survey experiments ($N = 1722$). In the first experiment ($N = 717$) we tested the effect of three types of explanations - a procedural, a rationale and a combined explanation – on citizen trust in algorithmic recommendations. The results showed that providing an explanation causes a major increase in citizen trust, and that there is limited variation between the different types of explanations. In the second experiment ($N = 1005$) we investigated whether adding an *actionable* explanation- i.e. providing citizens with an alternative way to act - increases trust in algorithmic recommendations. We found that such an action perspective has a limited but real effect on trust. We conclude that explaining algorithmic recommendations in service delivery - in any form - is crucial to foster trusting perceptions and behavior of citizens.

Literature

Fine Licht, Karl de, and Jenny de Fine Licht. 2020. “Artificial Intelligence, Transparency, and Public Decision-Making: Why Explanations Are Key When Trying to Produce Perceived Legitimacy.” *AI & SOCIETY* 35 (4): 917–26. <https://doi.org/10.1007/s00146-020-00960-w>.

Giest, Sarah, and Stephan Grimmelikhuijsen. 2020. “Introduction to Special Issue Algorithmic Transparency in Government: Towards a Multi-Level Perspective.” Edited by Sarah Giest and Stephan Grimmelikhuijsen. *Information Polity* 25 (4): 409–17. <https://doi.org/10.3233/IP-200010>.

Gupta, Kriti, P. 2019. “Artificial Intelligence for Governance in India: Prioritizing the Challenges Using Analytic Hierarchy Process (AHP).” *International Journal of Recent Technology and Engineering* 8 (2): 3756–62. <https://doi.org/10.35940/ijrte.B3392.078219>.

Shin, Donghee. 2020. “User Perceptions of Algorithmic Decisions in the Personalized AI System: Perceptual Evaluation of Fairness, Accountability, Transparency, and Explainability.” *Journal of Broadcasting & Electronic Media* 64 (4): 541–65. <https://doi.org/10.1080/08838151.2020.1843357>.