Calculating participation

How algorithmic processes impact
opportunities to be part of society

- English Summary of the full German Working Paper -

BertelsmannStiftung

Calculating participation

How algorithmic processes impact
opportunities to be part of society

- English Summary of the full German Working Paper -

Kilian Vieth
Dr. Ben Wagner
for the Bertelsmann Stiftung

1 Executive Summary

Algorithms are increasingly making decisions with us, for us and about us – thereby giving rise to new questions about participation. This expert opinion makes a number of initial suggestions for structuring and classifying the participation-related issues stemming from algorithmic decision-making (ADM). In today's digitized knowledge society, shaping technology is becoming a fundamental form of power. And if knowledge is power, then algorithms are becoming today's instruments of power. To what degree is it acceptable and desirable for algorithms to have an impact on the lives of individuals and on society as a whole? And which aspects of ADM must we consider more closely if we want to benefit from the potentials and minimize the risks to the greatest degree possible?

This expert opinion offers a general overview of ADM processes by explaining key concepts, delimiting the relevant area of analysis of ADM processes and presenting typical scenarios and functions relating to the use of these processes. With this as a basis, a system for classifying ADM processes is then proposed. The objective is to make it easier to evaluate and compare the potential impact of ADM processes on participation by using fewer criteria. This, in turn, will facilitate a prioritization and preparation of more in-depth research on the subject.

An evaluation of the impact on participation – a structure for which we are proposing here – is the point of departure for further steps, be they detailed analyses or regulatory measures. In those cases where ADM is expected to affect participation significantly, then a thoroughgoing effort must be made to ensure that positive impacts are exploited as much as possible, while negative impacts, such as ADM-driven discrimination, are prevented. When algorithms greatly influence participation, it is often necessary to counteract that influence by ensuring diversity, fair competition and principles of due process. The more prevalent the participation-relevant ADM processes are, the more comprehensive the potential responses must be.

Methodological approach

The potential impact on participation cannot be classified using binary decision-making criteria, since that would overlook the many socially relevant logics embedded in ADM processes.

Numerous criteria can affect the relevance to participation. The selection and weighting of these criteria, moreover, depend on how participation is defined. The point here is not to judge ADM processes as good or bad, but to evaluate their potential impact on participation in relative terms, regardless of the direction that might take.

The criteria can be divided into three groups: *actors, social embeddedness* and *consequences.* In other words, the groups do not reflect technological factors, but are an attempt to document political and social considerations.

The *actors*-related criteria examine the actual economic and political power of those supplying and/or operating the decision-making processes. The criteria used for *social embeddedness* reflect mutually reinforcing social interdependencies, both intended and unintended. Finally, the potential *consequences* – along with those that have already become apparent – are analyzed in terms of basic political and social rights. Existing legal norms are used as the guidelines for evaluating the impact on participation. The starting point is the German Antidiscrimination Act (§1 and 2 AGG) and its area of application. Discrimination is thus prohibited in the following areas in particular: recruitment, hiring, working conditions, membership in trade unions, and occupational training. Fair access to public goods such as education, social security, health care and housing must also be ensured.

We suggest assuming there is an impact on participation only after a certain threshold has been reached. Using the criteria sketched out below, the relative consequences for participation can be summarized and compared. This makes it possible to rank algorithms more simply and manageably in order to prioritize further steps, where needed.

Each of the criteria listed below can increase the *potential impact on participation* (Teilhabewirkungspotenzial, TWP; here, TWP=+1). If a criterion does not apply or cannot be evaluated as part of the quick-check process, no

point is given for the potential impact ($TWP_{neutral}=0$). Half points can also be given for each criterion ($TWP=+0.5$), and the maximum is two points ($TWP_{max}=+2$). The values thus range from 0 to 2 points per criterion. The higher the overall value, the greater the process's potential impact on participation.

Test questions are used to evaluate the criteria, and answering them is not always easy. The various applications are highly diverse and, on occasion, obscure, which is precisely why the system of abstract classifications is beneficial: It reduces the complexity of a highly complex network of issues. What might limit or promote participation? To answer that question, a person or organization uses the evaluative structure to analyze one or more relevant processes. After applying the criteria to a specific case and considering the results, they can assess which aspects of the decision-making process are most prominent; they can also compare different processes with each other.

GRAPHIC: **How we determine the potential impact on participation (TWP) of algorithmic decision-making systems**

Source: Vieth / Wagner (2017).

BertelsmannStiftung
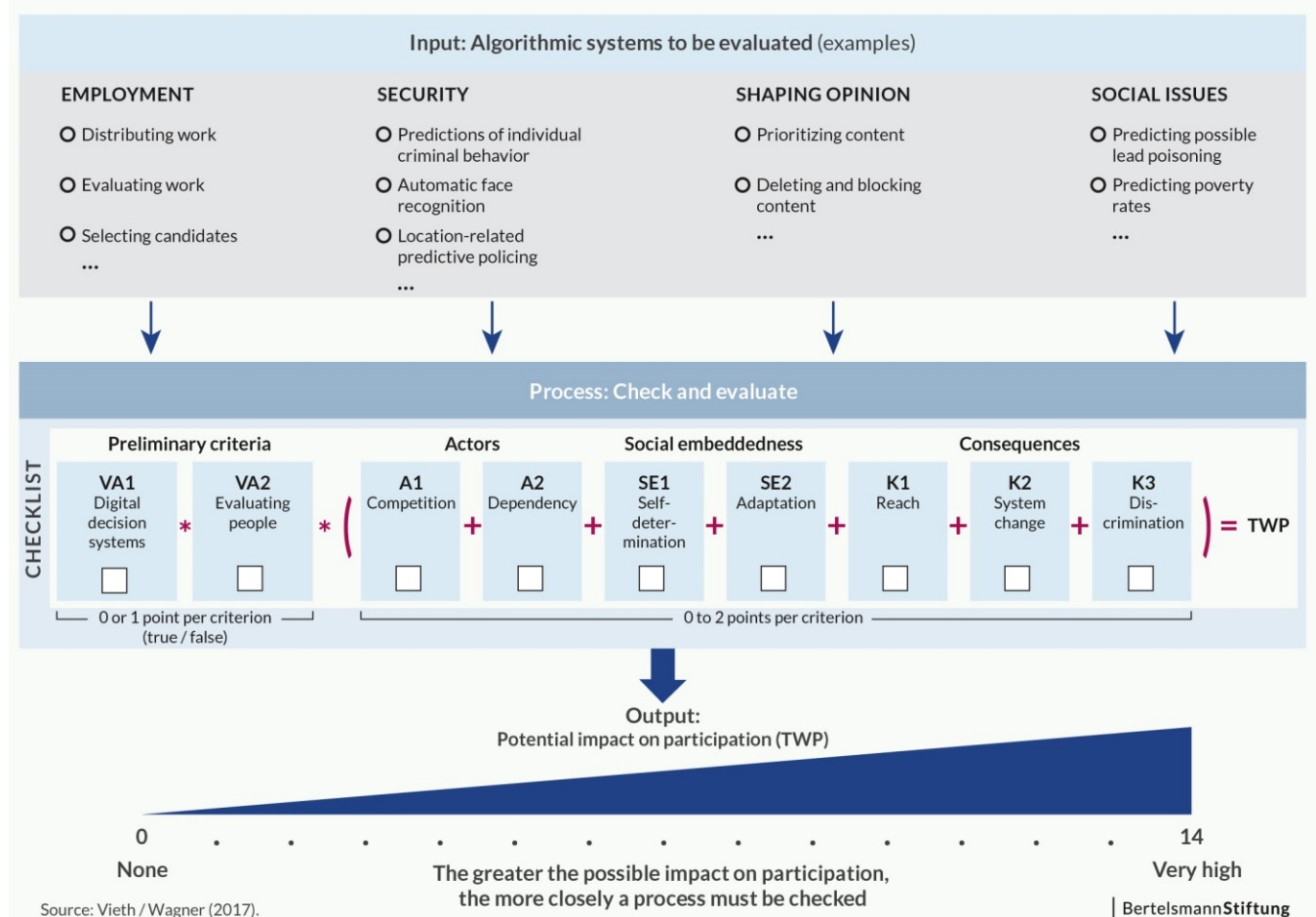
1.1 Actors

A1 Competition: Who is supplying/operating the algorithm? How much political and economic power do they have? Is there real competition between multiple suppliers/operators (oligopolies, monopolies, cartels, dominant market position, sovereign power)?

Example – Border controls: Even if private companies produce or supply software for granting visas, the software's use is an expression of sovereign power. No one can escape the state's power to grant visas, and to that extent a private supplier working on behalf of the state must also be seen as a monopoly. The actual organizational type or legal status of the supplier/operator is therefore less relevant for the impact on participation.

A2 Dependency: Is it a problem when the process or product is no longer available or access to it is denied? Can the product be readily replaced? Do substitute processes/products exist? What would it cost to switch to another process? Which organizations or groups are structurally dependent on the process? Do the system's construction and functioning result in an informal dependency? Which technical or social lock-in effects exist?

Example – Crowdwork: Robust competition exists in the market for crowdworking platforms. Multiple platform operators compete both for contract suppliers and crowdworkers. Participants could nonetheless become dependent if job histories and reputations become linked to the platform's specific algorithmic assessment processes. A similar lock-in effect can also arise with other assessment systems or with social media.

1.2 Social embeddedness

SE1 Self-determination: To what extent does the process serve (only) as a preliminary step for making a decision? Does the user retain any right to influence the process? Does the system decide (de facto or officially) on its own? How much freedom to change or manage an algorithmic decision does the user have? Do time pressure and the process's practical implementation affect the user's autonomy?

Example – Content control: The way software is designed for monitoring potentially illegal content can have a major impact on user autonomy. Time pressure (e.g. fast pace) and technical limitations can, de facto, greatly reduce the possibility of human input, even if a human will officially be the one to make the final decision.

SE2 Adaptation: How do people adapt to the algorithmic process? What impact does the decision-making system have on a practical level? Which interactions between humans and computers can potentially change the outcome? What is the process's actual power in social terms? Which (social) significance is ascribed to the process? Does it define and/or lead the discussion? How inherently influential are the process outcomes?

Example – Trending topics: The "trending topics" on Twitter are not infrequently seen as a social "reality." Although it is not generally known how the ranking is created, it still affects the public discourse and different actors adjust their behavior to reflect the logic used by the "trending topics" algorithm.

1.3 Consequences

K1 Reach: How many people are affected by the decision-making process (e.g. user population)? Is the process's reach known and/or can it be limited? How great is it?

Example – Candidate selection: If only one company uses a software application for choosing job candidates, then the application's reach is relatively small. If, however, it becomes a core decision-making tool for many companies, its impact on participation increases accordingly, for example if the logic used by one system predominates when employees are hired in a particular region or industry.

K2 System change: Does the decision-making process undermine principles of solidarity? Would individualization change the system? Could the process result in an unintended transformation of the impacted (social) environment?

Example – Health insurance: If the costs for public health insurance are partially decided by algorithms and are subsequently individualized using a new logic, then a system which has been based on social solidarity will no longer be to some degree. Instead of having everyone pay only according to their ability to do so, at least part of each person's contributions will be determined by individual factors.

K3 Discrimination: To what degree could people be disadvantaged by the decision-making process? Do the results exhibit a pattern of discrimination or might such a pattern be expected, e.g. based on race, ethnic or cultural background, gender, sexual identity, physical disability, age, religion or world view?

Example – Automated job listings: An analysis of Google's employment advertisements revealed that users classified as men were shown more highly paid positions than were women. In this case, the process has an impact on job market participation.

1.4 Summary: Determining the potential impact on participation

The degree of the potential impact on participation (TWP) of a random new ADM process can thus be summarized as follows:

$$TWP = VA1 * VA2 * (A1 + A2 + SE1 + SE2 + K1 + K2 + K3)$$

For this expert opinion, ADM processes are fundamentally relevant when a digital decision-making system is used and when people or characteristics attributed to people are being evaluated by that system. These two filtering criteria – "digital decision-making systems" (VA1) and "evaluation of people" (VA2) – can have either a value of 0 (not true) or 1 (true). The degree of the potential impact on participation (TWP) is the product of both preliminary criteria and the sum of the impact criteria. The TWP value is the sum of the test criteria (A1 + A2 + SE1 + SE2 + K1 + K2 + K3). Expressing the extent to which an ADM process can affect participation, the overall result (TWP) can range from 0 to 14. This testing structure makes it possible to filter out processes which do affect participation but whose potential impact is limited. This could be the case when TWP is less than 3, for example, although this threshold value can be adjusted to reflect different definitions of participation.

The above criteria can be used to develop a fast and relatively simple testing procedure, allowing companies and evaluators to check their ADM processes using specific criteria in just two or three pages.[1] Individuals and organizations can employ such criteria to evaluate ADM processes for their relevance to participation. To that extent, the process is a simple and quick introduction to understanding ADM processes. At the same time, it is not a substitute for a more comprehensive evaluation.

You can find the full version of the Working Paper (in German) here: DOI 10.11586/2017027

---

[1] See the "Quick Check" developed by the Danish Institute for Human Rights: https://hrca2.humanrightsbusiness.org/docs/file/HRCA%20Quick%20Check_English.pdf

Address | Contact

Bertelsmann Stiftung
Carl-Bertelsmann-Straße 256
33311 Gütersloh
German
Phone   +49 5241 81-81216

Konrad Lischka
Taskforce Digitalisierung
Phone   +49 5241 81-81216
konrad.lischka@bertelsmann-stiftung.de

www.bertelsmann-stiftung.de

BertelsmannStiftung