Dr. Karina Grisse, LL.M.

"Recommender Systems: Useful Aids in Decision-Making or Autonomy-Impairing Tools of Manipulation? – A Civil Law Perspective"

Abstract

Recommender systems are omnipresent in digitalized environments. They exist in various forms and contexts and they influence behaviour and decision-making. Concerns have been raised that recommender systems might harm human autonomy.

This paper looks at recommender systems and autonomy from a civil law perspective, that is, with a view to private autonomy. Private autonomy is understood here as the right to shape one's own legal relationships according to one's will. Do recommender systems deployed by platforms impair the so understood autonomy of those to whom the recommendations are addressed? Or, more precisely: Which features of recommender systems manipulate users under which circumstances and harm the users' free and willing choice in the context of legally relevant decisions? The paper uses "influence" as the general term and "manipulation" for influences that are not compatible with the other's autonomy.

Private autonomy is a legal principle crucial to all western jurisdictions. Still, legal systems usually neither contain a description of when a decision is sufficiently autonomous nor a full-fledged or irrevocable account of what kind of influences cross the line and harm autonomous decision-making. Even though private autonomy is a legal concept, it is open to extra-legal influences and the understanding of autonomy and undue influence (manipulation) is subject to change over time, affected by social, economic and technological development. Philosophical concepts in particular have had a decisive influence on the legal understanding of the terms and should be taken into account when assessing new phenomena.

This paper draws on a number of concepts of manipulation to assess under which circumstances recommender systems influence legally relevant decision-making in a way that should be considered manipulative. The analysis concludes that, though influential, recommender systems are not generally manipulative. Recommendations that follow a search query may facilitate decision-making. Recommendations of similar items or complements based on a user's search query can also be useful. Though they might tempt users to buy more than originally planned, these kinds of offers are usually transparent and easy to resist. Even ad-hoc and out-of-context-recommendations, that is, recommendations that are displayed independently from a user's current search query, are not generally manipulative, even when based on profiling (i.e. on the knowledge of a user's specific personal circumstances). However, recommendations can be manipulative under certain circumstances.

The reasonable expectation for a recommendation is that the recommendation is based on some kind of quality or relevance criteria. When that is not the case – e.g. if a recommendation was made because someone paid for it – and the recommendation criteria are not made transparent, the recommendation is rather deceiving. Ad-hoc out-of-context recommendations should be considered manipulative when they are based on profiling and address issues that are typically perceived as negative (e.g. illness, financial distress) and usually evoke negative or dire emotions (e.g. fear, anxiety). Under such conditions, they are likely to cause and exploit decision-making vulnerabilities. They constitute harm-alleviating offers that are less easy to resist. Recommendations based on real-time emotion recognition should also be considered manipulative. They exploit emotional states for less rational decisions in favour of the platform deploying the recommender system or its partners.

The paper goes on looking at existing and forthcoming legal rules relevant to recommender systems and examines whether the regulation is sufficient to remediate manipulative recommendations. The existing and forthcoming rules consist mostly of transparency obligations. They are suitable to solve the problem of recommendations based on unexpected criteria. Art. 29 draft Digital Services Act also provides that users must be able to choose non-profiling-based recommendations. The analysis concludes that this is

an important step, but that the non-profiling option should be the obligatory default setting or at least that factors such as health and financial situation (and other conditions that typically go hand in hand with fear, anxiety etc.) should be excluded as recommendation parameters as a default. If a user actively chooses to receive recommendations based on these factors, she consents to the influence in a way that it cannot be considered manipulative or at least that it is not harmful to her autonomy. Ad-hoc recommendations based on emotions should not be allowed where recommendations aim at influencing decisions of legal relevance. Information about the recommendation parameter is unlikely to eradicate the manipulative effect of such recommendations. Even with consent, exploiting emotions in such situations leaves a hunch of discomfort regarding human dignity.