

ICT20-079 - Humanized Algorithms: Identifying and Mitigating Algorithmic Biases in Social Networks

Abstract

Many ranking and recommender algorithms rely on user-generated social network data. For example, social media platforms such as Twitter or LinkedIn use the information on social networks to rank people and recommend new social links to the users. These networks that are generated by people are driven by fundamental social mechanisms such as popularity and homophily and they often contain diverse socio-demographic attributes of people. These attributes play an important role in the way individuals interact with others, and thus they determine the structure of networks. More importantly, the structure of networks has a crucial role in dynamical processes on networks such as diffusion of information, formation and evolution of biases, norms, and culture. However, very little is known about the effect of network structure on algorithms, and to what extent machine learning methods magnify social biases and practical approaches to mitigate algorithmic biases. The overarching aim of this project is to study the role of recommender and ranking algorithms in reinforcing biases in social networks with a specific focus on minorities. To this end, the research plan contains three crucial components (1) identifying structural conditions in which biases emerge in attributed networks, (2) exploring the impact of different ranking algorithms and recommendation systems on reinforcing bias in social networks and (3) developing methods to mitigate the algorithmic bias.

Scientific disciplines:

Statistical physics (20%) | Social psychology (40%) | Information technology (20%) | Data science (20%)

Keywords:

Algorithmic Bias, Social Networks, Minorities, Ranking algorithms, Recommender Systems

Principal Investigator:	Fariba Karimi
Institution:	Complexity Science Hub Vienna
Co-Principal Investigator(s):	Markus Strohmaier (RWTH Aachen) Julia Anna Koltai (Center for Social Sciences)

Status: Ongoing (01.01.2022 - 28.02.2026)

GrantID: 10.47379/ICT20079

Further links to the persons involved and to the project can be found under

<https://www.gmbh.wwtf.at/funding/programmes/ict/ICT20-079/>