Abstract

Using big data on women and family for policy development and improvement (III)

Sung Mi Jung
You Kyung Moon
Soo Yeon Lee
Yoon Ji So
Yeon Gyu Lim
Song Yi Park
Jae Kyung Lee

This paper reports the findings of the last research of our 3-year research project (2017-2019). The goals of the research are the followings. First, it investigates a way to lay a foundation for the use of big data regarding women and family. In the 1st and 3rd year of the project, we conducted studies on how to improve, manage, and share data with an aim to facilitate the use of big data. Second, it explores the possible use of big data and its implications on policy improvement by performing a pilot analysis of big data in selected areas of women's policy. For the pilot analysis, a wide range of subjects were selected over the course of

three years to analyze related data.

Based on the findings of the first and second research, we examined the possibility of establishing a platform in the mid-to-long term for the use of big data on women and family. We also discussed the creation of a platform and governance with an aim to explore the prospect of the use of big data in the field of women and family. The pilot analysis used social media data in order to examine gender differences in political discourse, which is typically considered a male-dominant area, in online spaces. Furthermore, we reviewed the likelihood of gender discrimination in data and algorithm and conducted an experimental study in order to determine the potential reproduction of gender bias by algorithm.

According to our study on gender differences in online political discourse, political discussions were routinely engaged in online community sites. Our research has also showed that the nature of political discourse changed depending on the gender-related characteristics of the site, implying that there may be distinctive gender differences in political needs. In addition, we have identified a need to keep track of easily accessible, open community sites in order to avoid potential conflicts with research ethics and to create a platform to collect and store online postings over a long term to track changes over time.

Through a hiring experiment performed to examine the potential reproduction of gender discrimination by big data algorithm, we proved that there was a potential for gender discrimination by big data algorithm. In particular, it was confirmed that a hiring algorithm which used intrinsically discriminatory data for learning reinforced and reproduced a discriminatory hiring practice. On the contrary, a hiring algorithm

that learned from data clear from attributes pertaining to discrimination, such as gender and number of children, made non-discriminatory hiring recommendations.

Last, we performed research on measures to establish an infrastructure and governance to promote the use of big data on women and family and provided the following suggestions: to construct a platform for big data on women and family; and to link and complement the existing platform for big data on women and family. For the establishment of governance for big data on women and family, we recommend the Ministry of Gender Equality and Family to secure personnel dedicated to statistics and big data and reshuffle its organization. Second, it is necessary to introduce regulations on the management of statistics and data on women and family. Third, we recommend designating an organization (the Korean Women's Development Institute) to plan, evaluate, and strengthen statistics on women and family. Last, there is a need to strengthen cooperation with other relevant organizations and establish strategies for the step-by-step implementation of the governance structure.