Identifying Structural Barriers to Women's Participation in ICT via Topic Modeling

Prof. Dr. Chi-Yo Huang

Graduate Institute of Technology Management, National Taiwan University of Science and Technology, Taiwan

Abstract

The ICT (Information and Communication Technology) industry is important for the modern economy. The industry enables innovation and influences almost everything, from business to the daily life of human beings. Despite its importance, the ICT industry shows a significant gender imbalance, with women being notably underrepresented, especially in technical roles. This disparity is evident in both education and the workforce, with fewer women pursuing ICT-related studies and careers. Although various scholars have attempted to identify the obstacles, the results vary. This research aims to identify key factors that influence gender imbalances in the ICT industry by employing text mining methods to bridge the existing research gap. Using topic modeling techniques, this research analyzed academic works on gender imbalance in the ICT industry published over the past two decades with text mining methods. This research identified the major factors contributing to gender imbalance in ICT industries. The coherence analysis results indicate that five topics best represent the retrieved keywords from the literature. This research then further deduced the implications of the 5 topics by analyzing the keywords and their corresponding probabilities. According to the analytic results, major topics including career and industry integration, educational segregation and stereotyping, pipeline and entry into ICT careers, structural and global leadership barriers, and workplace experience and retention are major barriers that cause gender imbalance in ICT industries. Implementing this research successfully will advance gender equality in the ICT industry. The results can offer policy guidelines and suggestions to improve women's economic empowerment and social impact.

Keywords: ICT (Information and Communication Technology) industry; gender equality, text mining, Latent Dirichlet allocation (LDA), topic model