# Detecting and Visualizing the Knowledge Structure of AI Ethics in Wikipedia Data

Mengyi Wei [a]*, Chuan Chen [a], Chenyu Zuo [a], Liqiu Meng [a]*,

*a Chair of Cartography and Visual Analytics, Technical University of Munich, Munich, Germany, Mengyi Wei - mengyi.wei@tum.de, Chuan Chen – chuan.chen@tum.de, Chenyu Zuo – chenyu.zuo@yahoo.com, Liqiu Meng - liqiu.meng@tum.de*

\* Corresponding author

**Abstract:**

AI-related ethical issues, such as privacy leaks (Srivastava, 2013), algorithm discrimination (Žliobaitė, 2017), have an enormous impact on public life. Current research mostly focuses on the establishment and elaboration of guidelines of AI ethics (Jobin et al, 2019), which, however, are often too vague to implement. An in-depth analysis of public opinions about AI ethics is required in order to formulate actionable guidelines. This study employs the community detection to identify the hierarchical structure of knowledge about AI ethics, and explores the relationship between knowledge communities, using Wikipedia as a collaborative data source with sufficient cross-validation (Stephany, 2017). It aims to analyze the status quo of the public cognition of AI ethics with the community-based network, and provide the necessary support and reference for AI ethics policy makers.

The workflow of this study is shown in Figure 1. First, we used the entry "Ethics of Artificial Intelligence" as the source link to crawl the Wikipedia data. The collected dataset is expressed as unweighted directed graph. Then, the Infomap algorithm (Rosvall and Bergstrom 2008, Leskovec et al. 2010) was applied to detect hidden structures in the network, which combined community detection and information coding. The core idea of Infomap algorithm is to use the shortest codelength to describe the path generated by a random walk. Two types of information codes were required for coding – codes of different groups and codes of different nodes within a group. A shorter encoding length means a better community partition, and better results of community detection. Large networks contain plentiful information, so how to extract information from large network structures is an important task of visualization. We chose the method of map equation to find, evaluate, and visualize the modular organization of the network (Rosvall, 2009). The map equation framework is flexible and can be combined with the Infomap search algorithm to identify two-level, multi-level, and overlapping organizations in weighted, directed, and multiplex networks (Bohlin et al, 2014). In the end, we analyzed the visualization to explore the hierarchical community structure of AI ethics.

Some preliminary results are shown in Fig. 2. Fig. 2(a) is the network of the original data of Wikipedia, which is a direct display of crawler data. The red dot is seed link, namely ethics of artificial intelligence. We can get a gradually extend network from this dot, and the figure displays the labels of several important nodes. Fig. 2(b) shows the communities detected by Infomap. Different colors represent different communities which can be explicitly labelled if desired. Interactive visual analytical methods can then be applied to extract topics from the individual communities and explore their relationships. This may eventually lead to the creation of a hierarchical knowledge structure of AI ethics, allowing us to understand the main risks that raise public concerns about AI ethics, and enforce actionable AI ethics guidelines.
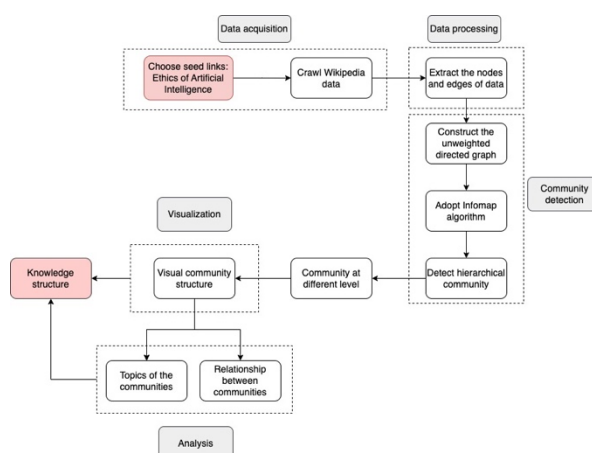
Figure 1. The workflow of detecting Wikipedia knowledge communities concerned with AI ethics.
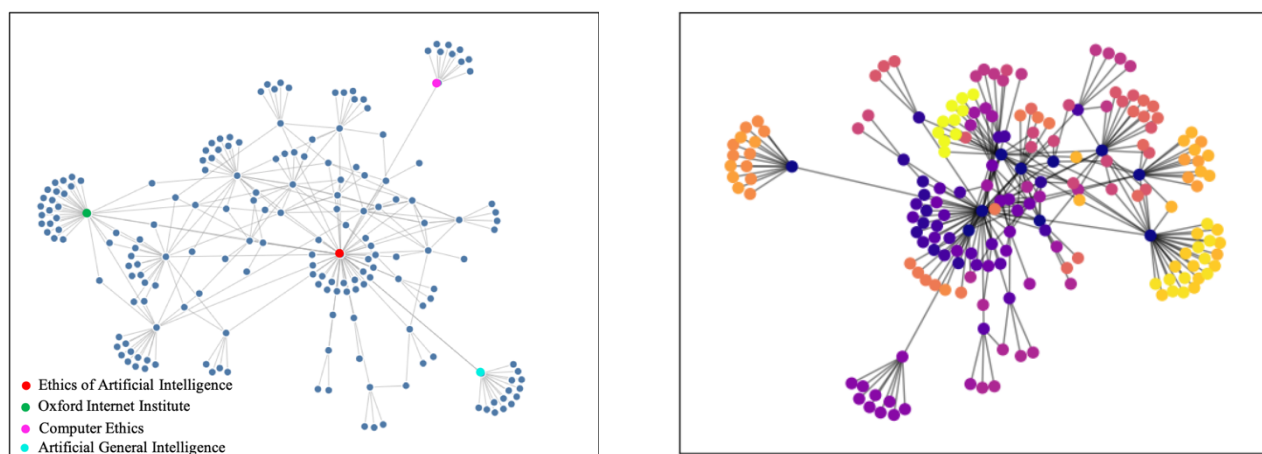


Figure 2. (a)The knowledge network of Wikipedia data related to AI ethics   (b) Community distribution of Wikipedia data

## References

Srivastava, A., & Geethakumari, G. (2013). Measuring privacy leaks in online social networks. In 2013 International Conference on Advances in Computing, Communications and Informatics (ICACCI) (pp. 2095-2100). IEEE.

Žliobaitė, I. (2017). Measuring discrimination in algorithmic decision making. Data Mining and Knowledge Discovery, 31(4), 1060-1089.

Jobin, A., Ienca, M., & Vayena, E. 2019. The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, 1(9), 389-399.

Stephany, F., & Braesemann, F. (2017). An exploration of wikipedia data as a measure of regional knowledge distribution. In International Conference on Social Informatics (pp. 31-40). Springer, Cham.

Rosvall, M., & Bergstrom, C. T. (2008). Maps of random walks on complex networks reveal community structure. *Proceedings of the national academy of sciences*, 105(4), 1118-1123.

Leskovec, J., Huttenlocher, D., & Kleinberg, J. 2010. Predicting positive and negative links in online social networks. *In Proceedings of the 19th international conference on World wide web*. pp. 641-650.

Rosvall, M., Axelsson, D., & Bergstrom, C. T. (2009). The map equation. The European Physical Journal Special Topics, 178(1), 13-23.

Bohlin, L., Edler, D., Lancichinetti, A., & Rosvall, M. 2014. Community detection and visualization of networks with the map equation framework. *In Measuring scholarly impact* (pp. 3-34). Springer, Cham.