

Association Analysis for Knowledge Imbalances in Wikidata = Analisis Asosiasi untuk Ketidakseimbangan Pengetahuan di Wikidata

Millenio Ramadizsa, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=9999920541944&lokasi=lokal>

Abstrak

Ketidakseimbangan pengetahuan merupakan fenomena yang semakin menonjol seiring berjalannya waktu, terutama dengan perkembangan Artificial Intelligence dan Data Science. Salah satu open knowledge base yang mengalami fenomena ini adalah Wikidata. Ketidakseimbangan pengetahuan dapat menyebabkan banyak hal negatif, contohnya adalah data yang tidak akurat dan kesimpulan yang bias. Untuk membantu mengatasi ketidakseimbangan pengetahuan kami mengusulkan sebuah solusi menggunakan association analysis. Kami menyediakan framework yang dapat mengidentifikasi gap properties, yang dapat digunakan untuk mengidentifikasi akar dari ketidakseimbangan pengetahuan di dalam Wikidata class (e.g. computer scientists, sovereign states). Melalui gap property ratio, kami dapat menghitung level ketidakseimbangan pengetahuan dalam Wikidata class. Semakin tinggi gap property ratio maka semakin tinggi tingkat ketidakseimbangan pengetahuan dalam suatu kelas. Untuk memvalidasi framework yang kami buat, kami melakukan analisis ketidakseimbangan pengetahuan di 20 kelas Wikidata. Kami harap hasil dari riset ini dapat membantu kontributor Wikimedia dalam menyelesaikan fenomena ketidakseimbangan pengetahuan lebih cepat dan akurat.

.....Knowledge imbalances are a phenomenon that has become more and more prominent over the years, especially with the growth of AI and data science. Wikidata is one of the open knowledge bases having this phenomenon. The growing number of items in Wikidata is not followed by an even distribution to every group and community. This phenomenon may have multiple negative implications, such as data inaccuracy and biased conclusions. In order to help in addressing knowledge imbalances in Wikidata we propose an approach using association analysis. We provide a framework that can identify gap properties, useful to pinpoint the root causes of knowledge imbalances in Wikidata classes (e.g., computer scientists, sovereign states). Furthermore, through the gap property ratio, we can quantify the knowledge imbalance level within Wikidata classes. The higher the gap property ratio, the larger the knowledge imbalance is for that class. To further validate our framework we conduct a knowledge imbalance analysis on 20 Wikidata classes. We hope that the result of this research can help Wikimedia contributors in addressing the knowledge imbalance phenomenon in Wikidata more swiftly and accurately.