

ПРОБЛЕМЫ АНАЛИЗА BIG DATA В СОЦИАЛЬНЫХ НАУКАХ

Широканова А. А.

*Белорусский государственный университет, Минск, Беларусь,
e-mail: shirokanova@bsu.by*

«Большие данные» становятся новой обсуждаемой темой в социальных науках. Начиная с 2012 г., многие определяют данное понятие через такие признаки, как объем (volume), скорость (velocity) и разнообразие (variety) [1; 2, с. 18]. Коммерческие компании, предоставляющие программы по сбору «больших данных», также подчеркивают точность (veracity) получаемой информации о поведении пользователей, по сравнению с опросами пользователей о мнениях и поведении [3].

Общую трудность работы с большими данными представляет собой, с одной стороны, хранение огромных массивов данных, вплоть до необходимости обработки потока данных без возможности их хранения; а с другой – необходимость фильтрации получаемых больших данных, чтобы найти и распознать паттерны, имеющие ценность, поскольку большие данные не создаются специально для исследовательских целей («терриконы информации»).

Специфические технические трудности в социальных науках работа с большими данными заключаются в широком освоении таких направлений, как: сбор информации из электронных социальных сетей и всемирной паутины в целом; интеллектуальный анализ данных; машинное обучение для автоматизированной классификации данных; сетевой, текстовый и семантический анализ.

Помимо технических, значительны и содержательные проблемы использования больших данных для социального анализа: прежде всего, это динамичные источники ошибок и погрешностей, зачастую не позволяющие обеспечить надежные результаты или анализ в нескольких временных точках [2, с. 37]. Этические проблемы анализа больших данных состоят в неопределенности соблюдения границ частной жизни и отсутствии общественно одобренных стандартов применения больших данных. В Беларуси также существует необходимость законодательной защиты персональных данных [4] и определения юрисдикции в регулировании сбора данных международными коммерческими компаниями.

Таким образом, появляется новое направление сотрудничества между компьютерными и социальными науками, связанное с анализом и использованием результатов анализа больших данных.

Литература

1. Dumbill, E. What is big data? [Electronic resource]. – 2012. – Mode of access: <http://radar.oreilly.com/2012/01/what-is-big-data.html>. – Date of access: 30.03.2015.
1. Отчет ААРОР о больших данных: 12 февраля 2015 / Л. Джапек [и др.]; Американская ассоциация исследователей общественного мнения; пер. с англ. Д. Рогозина, А. Ипатовой, Е. Вьюговской. – Москва, 2015. – Режим доступа: http://wciom.ru/fileadmin/nayka/grusha2015/AAPOR_big_data.pdf. – Дата доступа: 30.03.2015.
2. The FOUR V's of Big Data [Electronic resource]. – 2015. – Mode of access: <http://www.ibmbigdatahub.com/infographic/four-vs-big-data>. – Date of access: 30.03.2015.