

УДК 1:004

ФИЛОСОФИЯ БОЛЬШИХ ДАННЫХ

Круговой В.Н., магистрант

*Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь*

Малыхина Г.И. – канд. фил. наук, доцент

В статье рассматриваются философские аспекты феномена больших данных. Анализируются онтологические основания цифровых данных, их влияние на классические эпистемологические принципы (проблема индукции, соотношение корреляции и причинности), а также этические вызовы, связанные с алгоритмической обработкой информации. Особое внимание уделяется трансформации научного метода в эпоху data-driven исследований и формированию новой парадигмы познания. Показано, что большие данные не только изменяют инструментарий науки, но и ставят под сомнение традиционные представления о субъекте, объекте и границах научного знания.

Стремительное развитие цифровых технологий, повсеместная цифровизация социальных, экономических и научных процессов привели к возникновению феномена больших данных (big data). Под большими данными принято понимать массивы информации, характеризующиеся большим объёмом (volume), высокой скоростью обновления (velocity) и разнообразием структур (variety) [1]. Однако за этими количественными характеристиками скрывается качественно новый феномен, который требует философской рефлексии. Философия больших данных представляет собой междисциплинарную область, находящуюся на пересечении эпистемологии, онтологии, этики и философии науки. В настоящей статье мы рассматриваем ключевые философские проблемы, возникающие в связи с широким внедрением методов анализа больших данных в научное познание и социальную практику.

Современные философские определения больших данных демонстрируют разнообразие подходов. В рамках data science (в частности, публикаций arXiv) большие данные рассматриваются как масштабные и высокоразмерные массивы информации, которые невозможно обработать традиционными методами из-за объёма, скорости обновления и структурной сложности [2]. В философии искусственного интеллекта Ramon Alvarado интерпретирует Big Data как эпистемическую среду, в которой знание формируется через вычислительные процедуры [3]. Подходы Explainable AI (Long et al.) акцентируют внимание на том, что большие данные представляют собой многослойные и неструктурированные информационные потоки, требующие автоматизированных моделей для выявления устойчивых зависимостей [4]. В философии науки William Pietsch подчёркивает, что Big Data выступают новой формой эмпирического материала, где статистические зависимости могут предшествовать теоретическим моделям [5].

Инженерные науки (в частности, стандарты IEEE) рассматривают большие данные как информацию, требующую распределённых вычислений и новых архитектур хранения [6]. В когнитивной информатике (исследования arXiv) они описываются как массивы, выходящие за пределы человеческой интуитивной обработки [2]. В философии техники (журнал AI & Society) большие данные трактуются как социотехнические конструкции [7]. В вычислительной социологии (издания Springer) они понимаются как цифровые следы человеческой деятельности [8]. В машинном обучении (публикации ACM) подчёркивается эффект дрейфа данных, указывающий на нестабильность статистических свойств [9], а в теории сложных систем (исследования arXiv) — их многомасштабность и нелинейность [10].

Совокупность современных определений больших данных показывает, что речь идёт не об одном феномене, а о многослойной концепции, объединяющей технические, эпистемологические и социальные измерения. В рамках data science большие данные прежде всего описываются как количественно и структурно избыточные массивы информации, выходящие за пределы традиционных методов обработки [2]. Однако уже на этом уровне становится очевидно, что их специфика определяется не только масштабом, но и динамикой — скоростью обновления и изменчивостью структуры.

Философские интерпретации смещают акцент с характеристик данных на способы работы с ними. Так, Ramon Alvarado рассматривает большие данные как эпистемическую среду, где знание формируется через вычислительные процедуры, а не через классические теоретические конструкции [3]. Близкую позицию занимает William Pietsch, подчёркивая, что в условиях больших данных статистические зависимости могут предшествовать теории и фактически направлять исследование [5]. Это означает сдвиг от объяснения к выявлению закономерностей, где корреляция становится не побочным результатом, а отправной точкой научного поиска.

Подходы Explainable AI акцентируют внимание на внутренней сложности данных: они понимаются как многослойные, неструктурированные потоки, требующие автоматизированных моделей для извлечения устойчивых зависимостей [4]. При этом сами модели становятся необходимым посредником между данными и знанием. В инженерных дисциплинах (например, IEEE и ACM) подчёркивается инфраструктурный аспект: большие данные требуют распределённых вычислений и новых архитектур