

Освоение моделирования как элемента подготовки студентов-гуманитариев в условиях развития технологии больших данных

Mastering Technologies for Working with Big Data As an Element of Modern Training of University Humanities Students

Получено 08.06.2024 Одобрено 18.06.2024 Опубликовано 25.06.2024

УДК 373 ББК 74.2

DOI: 10.12737/1998-1740-2024-12-3-11-14

М.И. ШУТИКОВА,
д-р пед. наук, профессор, заведующий кафедрой
информатики и управления, ФГКВОУ ВО «Военный
университет имени князя Александра Невского»,
г. Москва
e-mail: raisins_7@mail.ru

С.А. БЕШЕНКОВ,
д-р. пед. наук, профессор, профессор кафедры
информатики и управления, ФГКВОУ ВО «Военный
университет имени князя Александра Невского»,
г. Москва
e-mail: srg57@mail.ru

M.I. SHUTIKOVA,
Doctor of Pedagogical Sciences, Professor,
Head of Department of Informatics and Management,
Military University named after Prince Alexander Nevsky,
Moscow
e-mail: raisins_7@mail.ru

S.A. BESHENKOV,
Doctor of Pedagogical Sciences, Professor,
Head of Department of Informatics and Management,
Military University named after Prince Alexander Nevsky,
Moscow
e-mail: srg57@mail.ru

Аннотация

Большие данные являются качественно новым информационным объектом, который требует особых методов анализа и обработки. Особенность больших данных заключается в следующем: традиционные методы обработки в случае больших данных могут оказаться неэффективными, поскольку требуют очень больших временных рамок и вычислительных ресурсов; большие данные во многом теряют дискретную структуру и приближаются по своим свойствам к структурам непрерывности, что затрудняет применение традиционных методов дискретного анализа.

В статье рассматриваются подходы к изучению больших данных в гуманитарных вузах.

Обсуждаются источники возникновения больших данных в этой сфере как дисбаланс между синтаксическими семантическими аспектами информационного объекта. Обосновывается, что в гуманитарных вузах акцент целесообразно сделать на освоении методологии моделирования, которая во многом позволяет устранить этот дисбаланс.

Ключевые слова: большие данные, гуманитарное образование, информация, данные, моделирование.

Abstract

Big Data is a qualitatively new information object that requires special methods of analysis and processing. The peculiarity of Big Data is the following: traditional methods of data processing in the case of Big Data may not be effective, since they require very large time frames and computing resources; Big Data, in many ways, loses its discrete structure and approaches its properties to structures of continuity, which complicates the use of traditional methods of discrete analysis.

The article discusses approaches to the study of Big Data in humanitarian universities.

The sources of the emergence of big data in this area are discussed as an imbalance between the syntactic and semantic aspects of the information object. It is substantiated that in humanitarian universities it is advisable to place emphasis on mastering the modeling methodology, which, in many respects, allows eliminating this imbalance.

Keywords: Big Data, humanities education, information, data, modeling.

Феномен больших данных является одной из характерных черт четвертой промышленной революции. **Особенность больших данных** заключается, в частности, в следующем:

- традиционные методы обработки данных могут оказаться неэффективными, поскольку требуют очень больших временных рамок и вычислительных ресурсов;
- большие данные во многом теряют дискретную структуру и приближаются к структурам непрерывности, что затрудняет применение методов дискретного анализа.

Технологии извлечения информации из больших данных весьма разнообразны: от имитационного моделирования и теории принятия решений до визуализации больших данных средствами ассоциативной живописи. Владение технологиями извлечения информации из больших данных является на сегодняшний день неотъемлемым компонентом ИКТ-компетентности современного специалиста [3].

В настоящее время изучение технологии работы с большими данными входит в содержание практически всех курсов подготовки не только

IT-специалиста, но и широкого круга гуманитарных специальностей.

В рамках IT-специальностей акцент делается преимущественно на освоение программных инструментов работы с большими данными.

Примером этого могут служить учебные курсы по аналитике больших данных, реализуемые в НИУ ВШЭ.

Например, в Дополнительной программе профессиональной переподготовки «Аналитик данных» НИУ ВШЭ (544 часа) за год предполагается изучить **восемь программных блоков:**

- 1) «Python как инструмент анализа данных»;
- 2) SQL;
- 3) «Математическая статистика»;
- 4) «A/B-тестирование»;
- 5) Business Intelligence;
- 6) «Машинное обучение»;
- 7) «Продуктовая аналитика»;
- 8) Data Warehouse (DWH) [1].

В этой программе акцент сделан на программные инструменты анализа больших данных, применяемые преимущественно в бизнесе.

Аналогичным образом построен обязательный курс «Методы и средства обработки больших данных» (Московский институт электроники и математики им. А.Н. Тихонова в рамках направления «Интернет вещей и киберфизические системы»).

Содержание этого курса следующее:

- «Основы технологии работы с большими данными»;
- «Архитектура обработки больших данных MapReduce в рамках платформы Hadoop»;
- «Процесс обработки больших данных на примере Apache Spark»;
- «Анализ и хранение данных на платформе Hadoop»;
- «Визуализация больших данных».

Важным аспектом этого курса является обсуждения происхождение самого феномена больших данных. Подчеркивается, что программно-аппаратные комплексы интернета вещей и различные киберфизические системы генерируют очень большой объем неструктурированных данных, которые необходимо оперативно обработать.

Названный подход достаточно хорошо отражает особенности деятельности с большими данными, которую осуществляют специалисты IT-сфера.

Однако применительно к гуманитарным специальностям этот подход оказывается малоэффективным.

Чтобы выстроить стратегию освоения больших данных для студентов-гуманитариев, необходимо проанализировать особенности их будущей деятельности в рамках цифрового социума.

Главная особенность этой деятельности состоит в том, что специалист гуманитарных специальностей имеет дело прежде всего с разнообразными текстами, то есть с совокупностью организованных знаков, которые несут в себе определенную информацию. Такую информацию целесообразно рассмотреть в трех плоскостях по **методу треугольника Фреге, в котором фиксируются:**

- объекты внешнего мира;
- знаки, которые замещают эти объекты в процессе познания или коммуникации;
- смыслы, которые несут эти знаки.

Как показывает практика, успешность профессиональной деятельности непосредственно зависит от соблюдения баланса этих компонентов. Феномен больших данных в теоретическом смысле это и есть дисбаланс между объемом (скоростью, разнообразием) поступающих данных и их осмысливанием. Взаимодействие субъекта с цифровым социумом в значительной мере сводится к непосредственному восприятию различных знаковых систем без существенного развития их смысловой составляющей. Из этих систем человек создает свой индивидуальный мир (или даже много миров), который имеет очень слабую связь с реальностью. При этом такие миры, очень часто взаимоисключающие, существуют в сознании человека одновременно, как равноправные псевдореальности. Эта ситуация создает сложности принципиального характера. В процессе профессиональной деятельности человек так или иначе обращается к реальным объектам. Однако при наличии множества равноправных миров возникает вопрос: какой из этих миров считать настоящей реальностью? Ответ на этот, казалось бы, очевидный вопрос для современного человека далеко не однозначен.

В современном цифровом социуме для человека существует множество конкурирующих между собой реальностей: средства массовой информации, социальные сети и т. д. Физическая реальность в этом списке играет далеко не определяющую роль [7].

Система образования должна в определенной мере исправить эту ситуацию. Привлечение настоящей реальности, которая присутствует в схеме Фреге, дает возможность представить множество миров как совокупность различных

моделей реальности. Это дает исходную установку для формирования содержания обучения для гуманитариев по освоению феномена больших данных.

Идея модели формирует определенную иерархию познания реальности, где существуют ее различные описания. Это первый шаг к тому, чтобы перевести разрозненные, но равноправные по представлению данные в информацию, несущую определенный смысл.

Понятие модели является фундаментальным для процесса познания, общения и практической деятельности. В современном мире это понятие становится особенно важным.

Существует множество определений этого понятия, которые в целом сводятся к тому, что моделирование – это замена объекта некоторым другим объектом, который обладает существенными для целей моделирования свойствами и в рамках этих целей полностью заменяет исходный объект [5].

Использование моделей как инструмента противодействия размытию реальности и конструированию всевозможных виртуальных миров опирается на ряд принципов, которые целесообразно рассмотреть более подробно.

Цифровой социум – исключительно сложное общественное явление. Стоит рассмотреть систему моделей, в каждой из которых четко обозначен угол зрения на этот социум и определены границы применимости модели.

При обсуждении этих моделей нужно опираться на ряд важнейших общенаучных принципов, которые, по мнению ведущих исследователей В.Д. Руднева, В.В. Налимова и др. [4; 6], составили основу научной методологии XX в. и которые, по нашему мнению, не утратили своего значения и в начале XXI в.

В частности, одним из таких принципов является принцип дополнительности Н. Бора. Впервые он был сформулирован в рамках квантовой механики и очень скоро приобрел обще-теоретический характер. На языке моделей это означает возможность строить и обсуждать модели, в которых противоположности не противоречат, а дополняют друг друга.

В формулировке В.В. Налимова этот принцип обозначает, что для воспроизведения в знаковой системе целостного явления необходимы взаимоисключающие, дополнительные классы понятий. Ю.М. Лотман [2] в формулировке этого принципа использует несколько иную аргументацию. Он утверждает, что сколь ни распространяли бы мы круг наших сведений, потребность

в информации будет развиваться, обгоняя темп нашего научного прогресса. Следовательно, по мере роста знания незнание будет не уменьшаться, а возрастать, а деятельность, становясь более эффективной, – не облегчаться, а затрудняться. Этот принцип достаточно полно описывает ситуацию с интеллектуальными технологиями четвертой промышленной революции и роль пользователя этих технологий.

В целом, содержание обучения моделированию в рамках освоения технологий работы с большими данными, ориентированное на студентов-гуманитариев, выглядит следующим образом.

Модель как упрощенное подобие реального объекта. Модели в окружающей жизни. Их назначение и области применения. Модель как заменитель объекта в процессе познания, общения, практической деятельности. Виды моделей: натуальные и информационные.

Основные приемы моделирования внешнего вида, структуры, поведение объекта.

Информационные модели. Их виды и особенности. Формы представления информационных моделей.

Этапы моделирования. Формализация как важный этап моделирования.

Объект, задача, модель. Модель объекта, модель решения. Построение модели как важный элемент информационной технологии решения задачи.

Информационная модель как схема, изображение или описание изучаемого объекта с помощью определенной системы знаков. Получение информации. Адекватность модели объекту и целям моделирования.

Информационные модели в математике, физике, биологии, литературе и пр. Использование информационных моделей в познании, общении и практической деятельности.

Этапы построения информационных моделей.

Формализация как важный этап построения модели.

Формализация и структурирование задачи из различных предметных областей в соответствии с поставленной целью. Построение информационной модели, отвечающей данной задаче (словесное описание, таблица, график, диаграмма, формула, чертеж, алгоритм и пр.).

Моделирование видится одним из ключевых направлений информационной подготовки гуманитариев, поскольку методология моделирования является универсальной методологией познания окружающего мира, общения и практической деятельности. Понятие модели форми-

ирует единый взгляд на различные предметные области, позволяет обозначить междисциплинарные связи между различными областями, например, математикой и лингвистикой или математикой и музыкой.

Очень важным классом моделей являются модели управления. Понятие модели позволяет

с единой точки зрения подойти к освоению языков программирования. Каждый такой язык (C++, Python и др.) представляет собой информационную модель, ориентированную на решение определенного класса задач. Значительное число программных сервисов также являются программами моделирования.

Список литературы

1. Аналитик данных. Программа профессиональной педагогики. <https://www.hse.ru/edu/dpo/557704084>.
2. Лотман Ю.М. Семиосфера. – СПб., 2010. – 700 с.
3. Майер-Шенбергер В. Большие данные. Революция, которая изменит то, как мы живем, работаем и мыслим / В. Майер-Шенбергер, К. Кукиер. – Манн, Иванов и Фербер, 2014. – С. 240.
4. Налимов В.В. Вероятностная модель языка. – М.: Наука, 1974. – 272 с.
5. Ракитина Е.А., Бешенков С.А. Информационное образование в России // Знание. Понимание. Умение. – 2013. – № 3. – С. 42–51.
6. Руднев В.М. Словарь культуры XX века. – М.: Академический проект 2020. – 314 с.
7. Shutikova M.I., Beshenkov S.A., Mindzaeva E.V. Information and cognitive technologies in the context of the 4th technological revolution: education aspects// Journal of Siberian Federal University - Humanities and Social Sciences. 2019. № 9. Pp. 1694–1713.

References

1. Data analyst. Professional retraining program. <https://www.hse.ru/edu/dpo/557704084>.
2. Lotman Yu.M. Semiosphere. St. Petersburg, 2010, 700 p.
3. Mayer-Schoenberger V. Big data. The revolution that will change the way we live, work and think / V. Mayer-Schoenberger, K. Kukier. Mann, Ivanov and Ferber, 2014. p. 240.
4. Nalimov V.V. Probabilistic model of language. M.: Nauka, 1974. 272 p.
5. Rakitina E.A., Beshenkov S.A. Information education in Russia // Knowledge. Understanding. Ability. 2013. No. 3. pp. 42–51.
6. Rudnev V.M. Dictionary of culture of the twentieth century. M.: Academic project 2020. 314 p.
7. Kulikova M.Yu., Beshenkov S.A., Mindzaeva Yu.V. Information and cognitive technologies in the context of the 4th technological revolution: aspects of education // Bulletin of the Siberian Federal University - Humanities and Social Sciences. 2019. No. 9. pp. 1694–1713.