

УДК 304.2

ТЕХНОЛОГИИ BIGDATA – КАК ФАКТОР КОНКУРЕНТОСПОСОБНОСТИ ПРЕПОДАВАТЕЛЯ ВУЗА В УСЛОВИЯХ СОВРЕМЕННОГО ИТ-ОБРАЗОВАНИЯ



В.В. Шаталова

*Заместитель декана факультета
компьютерного проектирования,
канд.техн.наук,
доцент*

*Учреждение образования Белорусский государственный университет информатики и радиоэлектроники, Республика Беларусь
E-mail: shatalova@bsuir.by*

В.В. Шаталова

Окончила Белорусский государственный университет информатики и радиоэлектроники, факультет компьютерного проектирования. Работает в должности заместителя декана факультета компьютерного проектирования.

Аннотация. В данной статье рассмотрена значимость технологии Big Data как фактор конкурентоспособности современного преподавателя, проанализированы перспективы применения данной технологии в сфере высшего образования. Показаны возможности применения технологий Big Data. Представлены современное состояние, тенденции развития технологий Big Data.

Ключевые слова: BigData, Datascience, конкурентоспособность преподавателя, компоненты профессиональной деятельности, ИТ-образование.

Символами первой промышленной революции были чугун и пар, второй – сталь и точное производство, третьей – полимерные материалы, алюминий и электроника, а очередная революция проходит под знаком композитных материалов и больших данных. Так что же такое BigData? Стихийный поток лавины данных, новые технологии, радикально изменяющие информационную среду или очередной этап в технологической революции? Скорее всего, и то, и другое, и третье, и еще пока что-то неведомое.

Последнее время большое внимание ведущих аналитиков уделяется умению работать с большими объемами данных для обеспечения своей конкурентоспособности.

В современных условиях образование давно ассоциируется с ИТ-образованием и является отражением мировых экономических реалий общества. Конкурентоспособность преподавателя учебных заведений зависит от быстроты реакции на изменения в сфере образования, а именно способностью достигать поставленные цели в разных, быстро меняющихся образовательных ситуациях за счет владения методами решения большого класса профессиональных задач, которые отражают, в том числе инновационные тенденции профессиональной деятельности преподавателя [1].

Деятельность конкурентоспособного преподавателя регламентирована и представлена компонентами его профессиональной деятельности:

- гностическим компонентом,
- конструктивным компонентом,
- коммуникативным компонентом,
- организаторским компонентом.

Конструктивные способности преподавателя являются решающими и составляют основу для других направлений профессиональной деятельности, а содержание организаторского компонента напрямую зависит от возраста преподавателя. Конкурентоспособному преподавателю в условиях быстро изменяющихся реалий необходимо постоянно оценивать свои профессионально-личностные возможности, возможности использования инновационных технологий для решения текущих профессиональных задач, потребности потребителей образовательных услуг (студентов) [2].

С точки зрения собственной самооценки профессиональной деятельности конкурентоспособному преподавателю вуза необходимо периодически обращать внимание на возможности своих коллег: на публикационную активность, на инновационные методики преподаваний, на проводимые и посещаемые мероприятия и т.д.

Конкурентоспособный преподаватель, должен быть ориентирован на личностные ценности, основу которых определяет «нравственный фундамент» - интеллигентность; должен обладать системой профессиональных и ценностных ориентаций, мобильностью, стремлением к высокому качеству результатов своего труда, а также способностью находить «динамический баланс» между данными компонентами и потребностями обучающихся в получении образования в удобной для него форме. Все это есть источники дополнительного профессионального роста, а также являются потенциальными составляющими BigData конкурентоспособного преподавателя вуза.

В качестве объекта развития конкурентоспособности необходимо рассматривать интегральные характеристики личности, такие как: направленность, компетентность, гибкость; в качестве фундаментального условия – переход на более высокий уровень самосознания; в качестве психологического механизма – превращение собственной жизнедеятельности в предмет практического преобразования; в качестве движущих сил – противоречивое единство Я-действующего, Я-отраженного и Я-творческого; в качестве результата развития – творческую самореализацию, достижение неповторимости личности [3].

Конкурентоспособного специалиста отличает умение преподнести учебный материал, грамотно структурировать и преобразовать массив научной информации в учебную лекцию, излагать предметный материал во взаимосвязи с дисциплинами, представленными в учебном плане, осваиваемом студентами, выявлять межпредметные связи. Кроме того, использовать в учебном процессе знание современных проблем и тенденций развития соответствующей научной области.

Система образования всегда генерировала значительный объем данных. Вопрос касался как работать с этими данными на системном уровне – анализировать их и принимать решения на их основе. Анализ данных помогает повышать эффективность работы преподавателей, индивидуализировать обучение и вооружать студентов навыками, которые бы соответствовали их будущей специальности.

Термин «большие данные» был впервые употреблен в 1998 году Джоном Мэши (главный ученый компании Silicon Graphics). Четкого определения bigdata не существует: неясно, где проходит граница между большими и просто данными. «Большие данные – это когда больше терабайта», «большие данные – это объем информации, который невозможно обработать на одном компьютере», «большие данные – это новая нефть». Каждый может дать свое определение и будет прав. Данных было много всегда. Стремительное развитие

информационных технологий, подключение практически всех девайсов к интернету, развитие сенсорных технологий на современном этапе увеличивает количество данных в геометрической прогрессии.

Терминология, которую необходимо знать, чтобы лучше ориентироваться в мире больших данных:

1. BigData: как правильно употреблять?

Правильно говорить «бигдэйта» (с точки зрения транскрипции), но если вы скажете «биг дата», криминала не будет. Скорей всего именно такой вариант произношения войдёт в словарь как нормативный вариант. Почему data, а не datas, ведь «данные» – их много? Слово data является заимствованием из латинского языка, где множественное и единственное число образуются по правилам, отличным от английского, и data – это как раз форма множественного числа от datum. Поэтому говорить «дата» можно, а «датас/дэйтас» – нет. Русский эквивалент – «большие данные» – тоже уверенно вошёл в лексикон, хоть он и в два раза длиннее.

2. Datascience: как употреблять?

Чаще всего используется английская версия - «дата сайнс» или «дэйтасайнс». Русский аналог – «наука о данных». Это наука о методах анализа данных и извлечения из них ценной информации. Datascience как академическая дисциплина формируется с начала 2010-х. Чтобы стать специалистом в этой области, необходимо прежде всего быть отличным математиком – знать математическое моделирование, математическую статистику, комбинаторику, теорию графов и многое другое. Ну и, конечно, уметь программировать. Спрос на Datascience сильно превышает предложение.

3. «Машинное обучение»: как употреблять?

Русское словосочетание «машинное обучение» используется так же часто, как английское machine learning (что-то вроде «мэшинлёрнинг»). «Именно благодаря машинному обучению поисковая машина понимает, какие результаты (и рекламу) показывать в ответ на ваш запрос.

4. Datamining: как употреблять?

Datamining переводится с английского как «обнаружение знаний в базах данных», что отражает суть понятия, но звучит уж слишком длинно. Поэтому принято говорить «дата/дэйтамайнинг», «майнить» – извлекать данные, «намайнить» – извлечь. Датамайнингом называют как технологии, так и процесс обнаружения в сырых данных неизвестной и полезной информации. Основу datamining составляют всевозможные методы классификации, моделирования и прогнозирования.

В научный обиход термин ввёл израильский математик Григорий Пятецкий-Шапиро – ещё в 1989 году.

5. «Облака»: как употреблять?

Говорят как «облако/облачный», так и cloud (например, cloudcomputing – облачные вычисления).

Держать в голове все задачи на день, месяц, год не очень-то удобно, поэтому мы записываем их в блокнот или заносим на виртуальную доску. Точно так же наш компьютер не может хранить на своём диске сотни гигабайт видео, фотографий и музыки – их приходится закачивать на такие сервисы, как GoogleDrive или Яндекс.Диск. Естественно, для работы с большими данными «облака» просто необходимы.

6. «Суперкомпьютер»: как употреблять?

Это слово начало входить в русский язык ещё в конце 1960-х, когда в СССР появился первый суперкомпьютер БЭСМ-6, способный выполнять 1 млн. операций в секунду.

Речь идёт о вычислительной машине, значительно превосходящей по техническим параметрам и скорости обработки данных обычные персональные компьютеры. Как правило, она представляет собой систему высокопроизводительных компьютеров, используемых для решения задач в самых разных областях науки и технологий: от разработки атомного оружия до моделирования новых лекарств.

7. «Интернет вещей»: как употреблять?

Популярен и русский вариант, и английский – *internet of things*, а также аббревиатура ИОТ.

Вслед за компьютерами и смартфонами в Сеть вышли фитнес-трекеры, чайники, стиральные машины, телевизоры, датчики и сенсоры. Есть масса примеров внедрения технологий ИОТ в медицине, спорте, сельском хозяйстве, промышленности [4].

Изначально понятие *bigdata* описывалось через три V: *volume* (объём) – очень много информации; *velocity* (скорость) – данные быстро увеличиваются и обрабатываются; *variety* (многообразие) – в работу идут и числа, и тексты, и графические образы, и другие виды данных, в том числе неструктурированных.

Потом количество V увеличилось: *veracity* – достоверность; *viability* – жизнеспособность; *value* – ценность для экономики, науки и общества; *variability* – переменчивость; *visualization* – возможность образного представления.

Весь массив данных можно условно разделить на две группы. К первой относятся **структурированные данные**, то есть те, которые ручным способом вносятся в электронные таблицы, традиционные базы данных, подвергаемые обработке и анализу. Вторую группу составляют **не структурированные данные**, то есть те, которые появляются в виде различных «цифровых следов» взаимодействия человека с электронной средой и с другими людьми в электронной среде. Количество структурированных составляет лишь 5% от всех данных.

В сфере образования выделяются пять основных типов данных: *персональные данные*; *данные о взаимодействии студентов с электронными системами обучения и друг с другом* (электронными учебниками, онлайн-курсами, показатели отказов, скорости просмотра страниц, возвраты к страницам, количество связей, расстояние связей, количество просмотров страниц одним пользователем и т.д.); *данные об эффективности учебных материалов* (какой тип ученика с какой частью контента взаимодействует, результаты взаимодействия, образовательные результаты и т.д.); *административные (общесистемные) данные* (посещаемость, пропуски по болезни, количество проведенных уроков и т.д.); *прогнозные (предполагаемые) данные* (какова вероятность участия ученика в той или иной деятельности, какова вероятность выполнения задания и т.д.).

Очевидно, что сегодня все образовательные организации работают только с малыми данными. Это связано с тем, что в образовательных организациях отсутствует специальная электронная среда, которая содержит много онлайн-контента и как следствие – большое

количество взаимодействий пользователей с этим контентом и между собой относительно него. По сути дела, с большими данными работают авторы онлайн-курсов, электронного обучения [5].

Если традиционно преподавательская методика создаётся на основе персонального опыта одного или нескольких учителей, то на основе больших данных методика становится продуктом массового опыта. Помимо повышения качества и эффективности создаваемых методик, большие данные помогают персонализировать контент под потребности каждого студента, позволяя сделать обучение более индивидуальным. Преподаватели смогут лучше помогать отстающим студентам, так как программа укажет, в каких именно областях есть проблемы, станет меньше отстающих, так как технологии позволят заранее выявлять студентов, которые могут оказаться в группе риска. Студенты будут иметь возможность подобрать свою программу курсов, выполнить отдельное домашнее задание, получить более подробные рекомендации.

Система будет выбирать наилучшие места для будущих студентов, а к окончанию вуза у каждого студента будет цифровое портфолио, которое поможет молодым специалистам ориентироваться при выборе карьеры.

Образовательная аналитика дает ответ на вопрос: как работа с большими данными в ближайшее время изменит образовательный процесс?

Во-первых, анализ данных позволяет работать с индивидуальными программами обучающихся, персонализировать обучение. Данные показывают, какой тип учащегося с какой частью контента взаимодействует, как происходит это взаимодействие, где он проявил интерес, а где ему было скучно, с кем и как он взаимодействовал в процессе обучения, как прохождение того или иного курса повлияло на образовательные результаты, на каком этапе обучения ему нужна помощь. Обучение становится адаптивным.

Во-вторых, образовательная аналитика на основе больших данных меняет представление о формате образовательных программ. Тексты, используемые в образовательном процессе, могут быть не только оцифрованы, но и переведены в числовые данные. Пользователи двигаются по материалу с большой свободой, затем осуществляется анализ, как пользователи взаимодействовали с материалом: что оказалось эффективным, что неэффективным. Результатом такой аналитики является изменение контента. Поэтому образовательная программа превращается из формата утверждаемого текста в формат некоторой совокупности онлайн-контента, которая динамически изменяется через анализ данных, появляющихся в результате взаимодействия с онлайн-контентом обучающихся. Появляются так называемые «умная программа», «умный учебный план». Можно предположить, что программы учебных курсов также претерпят изменения: они могут стать метапредметными.

В-третьих, изменение подходов к мониторингу и оценке, как самого образовательного процесса, так и образовательных результатов. Мониторинг становится постоянным. Заинтересованность студентов в постоянном мониторинге связана с тем, что анализ данных позволяет сделать его учебный план индивидуальным, заинтересованность преподавателей связана с возможностью получения информации о продуктивных группах, обратной связи от учащихся к создаваемому контенту (интересно/не интересно, сложно/ легко и т.д.), для преподавателей – эффективное распределение ресурсов.

Оценка образовательных результатов может быть самостоятельной и/или коллективной, агрегированной на основе всех данных студента, полученных из всех взаимодействий. Оценка осуществляется для того, чтобы грамотно расширить образовательную программу обучаемого. Динамика образовательных результатов фиксируется постоянно, на основе этих данных формируются паттерны (повторяющиеся шаблоны), по которым можно судить о развитии учащегося.

В-четвертых, составной частью образовательной аналитики станут новые методы: а) прогноз, когда комбинация известных данных позволит прогнозировать искомое неизвестное; б) метод выявления структуры и кластеризация; в) сетевой анализ.

В то же время анализ научных и публицистических работ позволяют выделить два риска использования больших данных на современном этапе.

Во-первых, это угроза конфиденциальности. Большая часть создаваемых данных содержит личную информацию, мы можем больше собирать данных и дольше их хранить. Получение согласия на использование данных, анонимизация (удаление персональных данных) бесполезны. Повторное использование данных может восстановить личность человека. Важно подчеркнуть, что данные, сгенерированные учащимся, принадлежат только ему.

Во-вторых, данные могут превратиться в орудие репрессий и наказаний. Прогнозы, основанные на больших данных, могут повлечь за собой наказание лиц за несовершенно совершенные действия. В этом смысле важно формировать «культуру анализа данных» и у тех, кто управляет образованием, и у тех, кто непосредственно осуществляет образовательный процесс. Имеющиеся в арсенале и широко применяемые сейчас два управленческих решения на основе анализа данных – «наказание» и «поощрение» – однозначно не будут работать в мире больших оперативных данных [5].

Образование – это отрасль деятельности человека, в которой всегда производилось больше данных, чем в какой-либо другой области: огромное количество часов занятий за семестр, большое количество заданий, выполняемых обучающимися, многообразный образовательный контент и количество взаимодействий с ним, бесчисленное множество взаимодействий участников образовательного процесса между собой – все это дает основания утверждать, что в современном образовании уже имеются те самые большие данные. Образование является одной из сфер, в которой новые технологии «Больших Данных» являются очень актуальными, обеспечивают возможность перехода к новым, более эффективным образовательным моделям.

Широкое использование информационных технологий в современном образовательном процессе позволяет собирать, хранить и обрабатывать большие объемы данных, в том числе и в оперативном режиме. Аналитика этих данных позволяет решать новые задачи, обеспечивающие поддержку таких важных образовательных трендов, как использование среды открытых знаний, внедрение адаптивных систем обучения, реализация коллаборативного обучения.

Приход в образовательную сферу технологий BigData означает существенное усиление роли в ней информационных технологий, что способствует повышению конкурентоспособности преподавателей, и как результат – развитию научного и инновационного потенциала университетов. Развитие концепции образовательной аналитики предполагает создание систем, функционал которых охватывает задачи извлечения данных из разнородных источников, моделирование образовательных процессов, накопление баз педагогических практик. Включение в образовательный процесс курсов, формирующих знания и умения, использование технологии обработки BigData, внедрение инструментария, обеспечивающий возможность аналитики больших объемов информации, все это является важным условием формирования у студентов компетенций, соответствующих современному уровню знаний в различных областях деятельности.

Данные, которые накапливались в образовательной системе в течение многих лет, носили преимущественно структурированный характер и были представлены, как правило, в форме отчетов и статистики. В большей степени они относились к так называемой академической аналитике, ориентированной на интегральную оценку образовательных учреждений.

Трансформация образовательной сферы под влиянием ИТ привела к внедрению новых, интерактивных форм обучения, которые обеспечивают поток неструктурированных данных, порождаемых интеллектуальными системами обучения, моделирующими программами и обучающими играми. Новые неструктурированные данные несут возможности изучения реальных результатов обучения, позволяют оценить компетенции студентов. Отмечается необходимость учета иерархической структуры этих данных, их зависимость от ситуации, в которой они возникают, а также их временные характеристики, которые являются важными для интерпретации данных. В целом использование аналитики этих данных позволяет реализовать идеи персонализированного, адаптивного обучения.

Создание адаптивных систем обучения связано с моделированием как самих образовательных процессов, так и поведения студентов. Основное назначение этих систем – формирование индивидуального подхода к обучению, который позволил бы с учетом индивидуальных особенностей обучаемого и конкретных задач, заложенных в программе обучения, обеспечить синтез необходимых знаний и навыков на основе получаемой информации. Важную роль здесь играет активное вовлечение обучаемого в процесс формирования его компетенций. Новые возможности образовательной аналитики связаны со сбором данных о поведении обучаемого на всех этапах его жизненного цикла, связанных с различными видами и программами обучения, и предоставлением этих данных по запросам образовательных учреждений (при условии соблюдения всех правовых и этических норм).

Образовательная аналитика нацелена на сбор и исследование данных на уровне образовательной программы, отдельных занятий. Ее возможности обеспечивают повышение качества обучения за счет постоянного мониторинга обучения и глубинного анализа причин, которые приводят к плохим результатам обучения, демонстрируемым на итоговом тестировании или других видах аттестации. Этот анализ позволяет обеспечить своевременную коррекцию процессов, причем персонализированно по отношению к отдельному студенту. Он также позволяет сформулировать рекомендации как для преподавателя, так и для студента. Технологии «Больших Данных» позволяют активно использовать данные, собираемые в рамках задач, решаемых образовательной аналитикой, для анализа деятельности университета в целом – как на уровне внутренней администрации, так и на уровне внешних организаций[6].

Задачи анализа таких данных для целей внутреннего администрирования направлены на оценку образовательных программ, преподавательского состава, образовательной среды университета. На внешнем уровне данные используются для сравнительного анализа университетов, вычисления рейтингов и т. д.

Одной из критических задач внедрения образовательной аналитики является сбор данных. Необходимыми элементами здесь являются некоторая стандартизация и категоризация данных, а также наличие у образовательных учреждений регламентов и инструментария для организации сбора таких данных и необходимость решения этических проблем – например, получение согласия обучаемого и преподавателя на постоянный мониторинг их деятельности. Важным условием является понимание участниками процесса полезности для них результатов анализа собираемых данных, а также знакомство с технологиями и аналитическим инструментарием, примерами полезного использования.

Для обоснования значительных затрат на внедрение технологий и инструментария Больших Данных в образовательных учреждениях необходимо создание новых метрик качества и эффективности образовательного процесса, которые были бы связаны с показателями, используемыми в интегральной академической аналитике, внедрение этих метрик для оценки образовательных учреждений, в том числе и на уровне государственных структур[7].

На сегодняшний день ещё не достаточно хорошо решены задачи использования Big Data. Один из примеров - задачи, связанные с образованием: автоматический анализ успеваемости, выработка индивидуальных программ и рекомендаций, прогнозирование показателей, социологические исследования коллективов студентов. В нашей стране пока ещё даже нет доступных и удобных массивов информации для решения подобных задач. Очень хочется надеяться, что у нас к Big Data будут относиться с должным вниманием. Что касается методов, которые здесь используются, то они достаточно стандартные для математика-специалиста в анализе данных. Как ни удивительно, с точки зрения статистики, результаты ответов на вопросы тестов очень похожи, например, на степень удовлетворённости просмотрами фильмов. Поэтому не важно, что рекомендовать: фильмы для просмотра или темы для повторения - механизм рекомендаций одинаков[8].

Анализ Big Data позволяет ускорить решение различных исследовательских, научных и педагогических недочетов. При изучении статистики, можно работать и с индивидуальными векторами, и с образовательными системами глобального уровня. Технологии Big Data помогают значительно улучшить так называемый «дизайн педагога» – в тех местах, где студент «засыпает от скуки», такую систему образования лучше заменить. Эта технология автоматизирует поведение образовательной системы и даёт фактические рекомендации и подсказки, если студенты застревают на одном месте. И если эти рекомендации и подсказки не помогают, – технология оповещает преподавателей и родителей, что таким студентам нужна дополнительная помощь.

Активное внедрение Big Data в образовательный процесс позволяет обеспечить переход к качественно новому уровню педагогической деятельности, значительно увеличивая ее дидактические, информационные, методические и технологические возможности, что в целом способствует повышению профессионального мастерства преподавателей. С повышением уровня информационно-коммуникационной компетентности изменяются мотивационные стимулы: преподаватели стремятся к профессиональному росту и личностной самореализации, включаются в поиск новых путей совершенствования процесса обучения. Но для этого необходимо создать все необходимые условия. Сегодня каждый преподаватель сам определяет наиболее важные аспекты совершенствования своего мастерства. Выбор есть и довольно большой: это прохождение курсов повышения квалификации и самообразование, участие в работе сетевых педагогических сообществ, различных научно-методических объединений, научно-педагогических конференций, семинарах и т.д. Все формы повышения педагогической квалификации преследуют благую цель – содействовать в повышении компетентности в предметной области и методики его преподавания. Технологии «Большие данные» способны превратить современное высшее образование в продукт с отличным предсказуемым результатом, а преподавателя ВУЗа – в суперпрофессора.

Литература:

- [1]. Фромм, Э. Человек для себя / Эрих Фромм; пер. с англ. Л. А. Чернышёвой. – М.: АСТ: ХРАНИТЕЛЬ, 2006. – 314 с.
- [2]. Андреев, В.И. Конкуренология: учеб. курс для творческого саморазвития конкурентоспособности / В. И. Андреев. – Казань: Центр инновационных технологий, 2009. – 468 с.
- [3]. Бибик, И.А. Big Data конкурентоспособного преподавателя вуза // Аспекты и тенденции педагогической науки: материалы II Междунар. науч. конф. – Санкт-Петербург : Издательский дом «Свое издательство», 2017. – 110 с.
- [4]. Соколова-Михайлова, С. 7 слов биг даты Краткий справочник терминов, которые нужно знать, чтобы лучше ориентироваться в мире больших данных [Электронный ресурс] / С. Соколова-Михайлова // Журнал «Кот Шрёдингера». – 2017. – № 10. – Режим доступа: <https://kot.sh/statya/3862/7-slov-big-daty>.
- [5]. (<http://www.edutainme.ru/post/big-data-edu/>)
- [6]. <https://predu.livejournal.com/1196996.html>
- [7]. Мальцева, С.В. Большие Данные в образовании: новые возможности и новые вызовы // Большие Данные в национальной экономике: тезисы докладов конференции «Большие Данные в национальной экономике» – Москва : Издательский дом «Открытые системы», 2013. – 53 с.

[8]. Михнев И.П. Технологии Big Data и их применение в сфере современного высшего образования / И.П. Михнев, А.Д. Челнокова, А.Д. Реут // Развитие современного образования: от теории к практике : материалы IV Междунар. науч.-практ. конф. (Чебоксары, 19 март 2018 г.) / редкол.: О.Н. Широков [и др.] – Чебоксары: ЦНС «Интерактив плюс», 2018. – С. 14-18.

BIG DATA TECHNOLOGIES – AS THE FACTOR OF COMPETITIVENESS OF A UNIVERSITY TEACHER IN THE MODERN IT-EDUCATION

V.V. SHATALOVA

Deputy Dean of the Faculty of Computer-Aided Design, Ph.D., Associate Professor

Belarusian State University Informatics and Radioelectronics, Republic of Belarus

E-mail: shatalova@bsuir.by

Abstract. This article discusses the importance of Big Data technology as a factor of competitiveness of the modern teacher, analyzes the prospects of using this technology in higher education. The results of research on the use of Big Data technologies are shown. The current state and development trends of Big Data technologies are presented.

Keywords: BigData, Datascience, teacher competitiveness, components of professional activity, it education.