

О.Ю. Колесниченко

DATA SCIENCE (НАУКА О ДАННЫХ)

в становлении информационного общества

УЧЕБНОЕ ПОСОБИЕ



МЕДИТЕЛСТВО
Прометей

Колесниченко Ольга Юрьевна

DATA SCIENCE (НАУКА О ДАННЫХ) В СТАНОВЛЕНИИ ИНФОРМАЦИОННОГО ОБЩЕСТВА

Учебное пособие

118633

•TDIU
kutubxonasi

ОНТИ

ИЗДАТЕЛЬСТВО
Прометей

МОСКВА
2021

316(07)

+ 672.15(07)

УДК 001.5, 004.6, 004.8

ББК 1, 60, 30г, 30ф

К60 3

Автор:

Колесниченко Ольга Юрьевна – кандидат медицинских наук, старший преподаватель кафедры социологии медицины, экономики здравоохранения и медицинского страхования Института социальных наук ФГАОУ ВО Первого Московского государственного медицинского университета имени И.М. Сеченова Минздрава России.

Рецензенты:

Л.С. Мазелис – доктор экономических наук, профессор, заведующий кафедрой математики и моделирования Владивостокского государственного университета экономики и сервиса, Владивосток, Россия;

И.Н. Григорьевский – кандидат технических наук, старший научный сотрудник Института программных систем имени А.К. Айламазяна Российской академии наук; ответственный секретарь оргкомитета Национального Суперкомпьютерного Форума, Переславль-Залесский, Россия.

Колесниченко О.Ю.

К60 Data Science (наука о данных) в становлении информационного общества: учебное пособие. – М.: Прометей, 2021. – 52 с.: ил.

ISBN 978-5-00172-110-9

Книга посвящена науке о данных (Data Science) и предназначена для широкого круга читателей, особенно она будет полезной для обучающихся по направлению подготовки «Интеллектуальные системы в гуманитарной сфере». Кратко описана история появления новой науки, даны основные понятия, связанные с хранением информации, Большими данными, системами счисления, научными парадигмами. Показаны этапы становления кибернетики и моделирования в гуманитарных науках, рассмотрены подходы к кибернетическому моделированию в медицине. Дана базовая информация о суперкомпьютере и квантовом компьютере. Сформулировано определение науки о данных, представлен прогноз ее развития в будущем.

Книга может служить учебным пособием для студентов высших учебных заведений и будет интересна обучающимся по разным техническим и гуманитарным направлениям подготовки. Также книга рекомендована аспирантам, молодым ученым, исследователям в области применения новых цифровых технологий, специалистам и руководителям организаций в рамках цифровой экономики.

ISBN 978-5-00172-110-9

© Колесниченко О.Ю., 2021

© Издательство «Прометей», 2021

Содержание

§ 1. История появления новой науки.....	4
§ 2. Смена парадигм и периодов.....	8
§ 3. Системы счисления.....	15
§ 4. История моделирования и кибернетики.....	21
§ 5. Суперкомпьютер и моделирование.....	28
§ 6. Квантовый компьютер.....	35
§ 7. Определение науки о данных и Кривая Гартнера (Hurescycle).....	41
Список литературы.....	49

§ 1. История появления новой науки

Наука о данных (Data Science) как самостоятельная дисциплина для преподавания в системе высшего образования появилась недавно, в начале XXI века. Существуют разные взгляды на это научное направление, которые зависят от страны и принадлежности к разным научным школам. Англоязычные авторы, описывающие появление науки о данных, опираются на этапы последовательного развития западной статистики и кибернетики. В России и ранее в советское время были свои выдающиеся математики и кибернетики, которых также по праву можно считать основоположниками современной науки о данных.

Первое имя в истории, которое связывают с наукой о данных, – Джон Уайлдер Тьюки (John Wilder Tukey, 1915–2000 гг.). Американский математик ввел два базовых термина в новый компьютерный мир. Термин «software» (программное обеспечение) был придуман в 1958 г. и термин «бит» / «bit» (сокращение от «binary digit») был придуман ранее в 1946 г.

Binary digit – двоичное число, *базовая единица измерения количества информации*, равная одному разряду в двоичной системе счисления. Бит – это также игра слов, по-английски означает «частица». В двоичной системе числа записываются с помощью двух символов 0 и 1. Такое число, чтобы не путать с десятичной системой, обозначают впереди символом & (например, &1010) или указателем внизу в конце (например, 1010₂). Число в десятичной системе обозначают соответствующим указателем внизу в конце (например, 1010₁₀).

Еще одно важное понятие – это байт. Термин «байт» / «byte» (краткое обозначение «Б» / «В») впервые введен в 1956 г. американцем Вернером Бухгольцем (Werner Buchholz), одним из пионеров-разработчиков компьютеров IBM. Байт – *единица хранения и обработки цифровой информации* – отражает, сколько бит одновременно обрабатывает компьютер. Байт состоит из 8 битов (октет / octet). Байт служит мерой, позволяющей ориентироваться в объемах данных (см. табл. 1).



Джон Уайлдер Тьюки
Фото: Time Inc.
Alfred Eisenstaedt.



Вернер Бухгольц
Фото: cyclowiki.org.

Именно в связи с ростом объемов хранения информации в байтах, когда хранилища данных стали вмещать петабайты информации, и появилась судьбоносная для науки о данных статья американского ученого, профессора Калифорнийского университета в Беркли (University of California, Berkeley), Клиффорда Линча (Clifford Lynch). Статья была опубликована в журнале Nature в 2008 г., она называлась «Big Data: How do your data grow?» (Nature, Vol. 455, pp. 28-29., 4 September 2008). В этой статье Клиффорд Линч писал о нарастающих потоках разных данных (астрофизических, генетических, социальных и др.) и предложил называть такой объем данных «Большие данные» (Big Data) [1]. Ученый описал жизненный цикл Больших

данных, сделав упор на вопросах хранения их в дата-центрах и управления большими объемами данных.

Так появился *термин Big Data*, а затем стремительно стала «наращиваться» концептуальная канва. Клиффорд Линч инициировал появление нового научного направления по изучению больших объемов данных, что явилось прорывом в цифровом мире. Это новое направление опирается на более ранние концептуальные подходы к аналитике данных, которые Джон Тьюки описал в 1962 г. в статье «The Future of Data Analysis» в журнале «The Annals of Mathematical Statistics» [2]. Джон Тьюки подчеркивал важность наглядного представления данных, что сегодня называют «визуализацией», разнообразные виды которой делают современные компьютерные программы и языковые среды. Ученый впервые сформулировал мысль о ценности аналитики как таковой, подразделяя ее на этапы: разведочный и подтверждающий.

Таблица 1

Измерение количества обрабатываемой компьютером информации в байтах
Международная система единиц / International System of Units

Русское обозначение	Английское обозначение	Степень	Пример вместимости
килобайт (Кбайт)	Kilobyte (Kbyte, KB)	10^3	параграф текста
мегабайт (Мбайт)	Megabyte (Mbyte, MB)	10^6	аудиофайл с минутой музыкальной записи
гигабайт (Гбайт)	Gigabyte (Gbyte, GB)	10^9	10 метров книжных полок
терабайт (Тбайт)	Terabyte (Tbyte, TB)	10^{12}	250 тысяч музыкальных треков
петабайт (Пбайт)	Petabyte (Pbyte, PB)	10^{15}	20 миллионов шкафов картотеки
эксабайт (Эбайт)	Exabyte (Ebyte, EB)	10^{18}	объем телефонных разговоров страны за полгода
зеттабайт (Збайт)	Zettabyte (Zbyte, ZB)	10^{21}	объем всего мирового трафика Интернета за 2016 г.
иоттабайт (Ибайт)	Yottabyte (Ybyte, YB)	10^{24}	объем трудно представить
бронтобайт (Ббайт)	Brontobyte (Bbyte, Bb)	10^{27}	будет актуальным к 2025 г.
геопбайт (Гпбайт)	Geopbyte (Gpbyte, Gpb)	10^{30}	будет актуальным после 2025 г.



Клиффорд Линч
Фото: en.wikipedia.org.

Профессор Оксфордского Института Интернета (Oxford Internet Institute) Виктор Майер-Шенбергер (Viktor Mayer-Schoenberger) и редактор журнала The Economist Кеннет Кукьер (Kenneth Cukier) внесли существенный вклад в укрепление нового научного направления, написав в 2013 г. книгу «Большие данные. Революция, которая изменит то, как мы живем, работаем и мыслим» («Big Data: A Revolution That Will Transform How We Live, Work, and Think») [3] и статью в журнале Foreign Affairs «The Rise of Big Data» [4]. В статье, привлечшей внимание всего мира, авторы сформулировали емкую фразу: «The data take center stage» – данные занимают центральное место. Это краткое предложение может быть девизом нашего времени, в котором жизнь социума все больше наполняется цифровыми данными и все больше подчиняется решениям на основе аналитики Больших данных. А в своей книге они одними из первых обратили внимание, что *«Большие данные» – это социологический феномен нового технологического уклада*, меняющий взаимодействие человека с миром и подходы к принятию решений. Виктор Майер-Шенбергер и Кеннет Кукьер пишут: «Подобно тому, как

Интернет радикально изменил мир, добавив связь между компьютерами. Большие данные изменяют фундаментальные аспекты жизни, предоставив миру небывалые возможности количественного измерения. Данные порождают новые услуги и инновации. И очень многое ставят под угрозу» [3].



Виктор Майер-Шенбергер
Фото: esfr.eu.

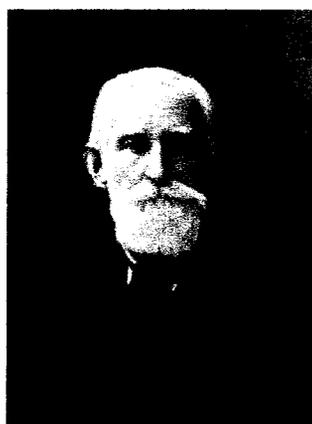
Кеннет Кукьер
Фото: fico.com.

Виктор Майер-Шенбергер и Кеннет Кукьер привели аналогию, которая отражает самую суть перемен: концепция Больших данных несет для истории науки такой же судьбоносный поворот, как изобретение микроскопа и телескопа в конце XVI века голландскими мастерами, отцом и сыном Янсенами. Микроскоп открыл человечеству невидимый микромир, телескоп позволил изучать макромир, также ранее недоступный для исследования. А Большие данные – то есть массивы данных, поступающие в электронные хранилища (дата-центры) – позволяют изучать время как отдельное измерение жизни человечества. В этом и заключена основная революционность Больших данных. Процесс оцифровки данных, меняющихся в динамике с течением времени, делает возможным исследование разных аспектов жизнедеятельности социума как динамических систем. Это, по сути, означает изучение эволюции систем во времени, что ранее осуществлялось только на теоретическом уровне, чем и занимались математика и кибернетика.

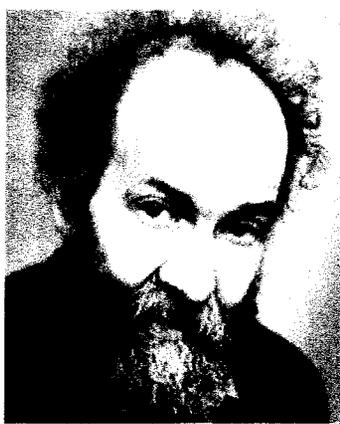


В России, а вернее в годы СССР, предтеча науки о данных – кибернетика – вначале была воспринята в штыки и даже запрещалась. В послевоенные годы (начиная с 1948 г.), в ответ на

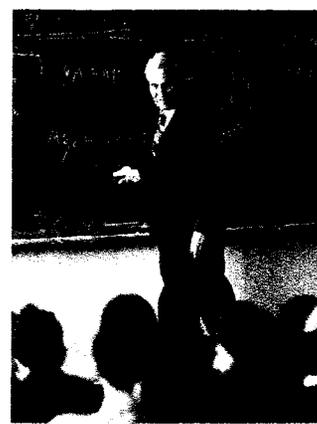
западные научные публикации, пропаганда в СССР обличала кибернетику как лженауку, нацеленную на построение рабовладельческого общества, где людей (трудовой пролетариат) заменят роботы. Критиков кибернетики пугала связь математики и инженерии с биологией, физиологией, психологией и социологией в единое учение. **Термин «кибернетика» происходит от древнегреческого выражения «искусство управлять».** Только на современном этапе, с появлением концепции Больших данных, Интернета вещей (Internet of Things, IoT), интеллектуальных сетей (Smart Grid), умных городов и т.д., кибернетические теоретические разработки стали повсеместно находить свое практическое применение.



Пафнутий Львович
Чебышев
Фото: История
Академии наук СССР,
М.–Л.: Наука, 1964. Т. 2.



Алексей Андреевич
Ляпунов
Фото: Архив музея
истории НГУ.



Андрей Николаевич
Колмогоров
Фото: «Новая газета».

Среди российских ученых, заложивших основу для развития направления, которое сегодня называют наукой о данных, необходимо выделить следующих. Пафнутий Львович Чебышев (1821–1894 гг.) внес существенный вклад в развитие теории вероятностей (центральная предельная теорема, закон больших чисел). Алексей Андреевич Ляпунов (1911–1973 гг.) считается одним из основоположников кибернетики, а также он занимался математической лингвистикой и математической биологией. Андрей Николаевич Колмогоров (1903–1987 гг.) – самый известный сегодня российский математик в мире. Его работы изучают в американских вузах и научных институтах программисты, которые создают передовое программное обеспечение для аналитики Больших данных и Искусственного интеллекта. А.Н. Колмогоров считается одним из основоположников теории вероятностей, он оставил научное наследие в таких областях, как математическая логика, функциональный анализ, теория сложности алгоритмов, теория информации, теория функций, теория меры, теория множеств, динамические системы и др.

§ 2. Смена парадигм и периодов

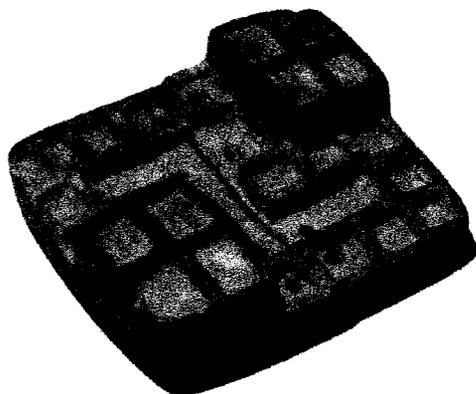
Парадигма – совокупность научных установок и взглядов. В ходе истории человечества наблюдалась неоднократная смена научной парадигмы. В древние времена знания строились на описании наблюдений – это *эмпирическая парадигма* (от греческого слова *empeiria*, что означает «опыт»). С середины прошлого тысячелетия (в средние века) стали появляться *теоретические модели* и обобщения. Изученные эмпирическим путем в естественных науках явления рассматривались как универсальные для понимания общества и процессов, происходящих в нем во времени. Возникали такие научные течения, как «социальная физика» и «социальная механика» [5]. Существовали противоположные упрощенные точки зрения на организацию общества как целого из отдельных частиц (индивидуумов). Голландский философ Гуго Гроций (Hugo de Groot, 1583–1645 гг.) считал, что люди стремятся к общению, результатом чего стало их объединение в общество. А английский философ Томас Гоббс (Thomas Hobbes, 1588–1679 гг.) наоборот утверждал, что людям свойственно взаимное отчуждение и неприятие друг друга, что требует установки правил поведения в обществе и порождает постоянные войны. В настоящее время, когда общество пронизано информационными системами со сбором Больших данных, эти два полюса мнений о природе человеческого общества можно обозначить как «маятник Гроция – Гоббса», постоянно раскачивающий общественную дискуссию то в сторону необходимости прозрачного общения и открытых данных, то в сторону права на личную тайну и закрытое общение через пароли. Нахождение и установление баланса представляет собой самую сложную проблему современности.

С середины XX века социализация научной парадигмы еще более усилилась. Изучение времени (исторические знания) превратилось в междисциплинарную область «социальной истории». Добавились возможности строить математические модели с использованием компьютеров. Данные стали оцифровываться – этот процесс называется «датафикация». *Вычислительная модельная парадигма* просуществовала вплоть до начала бурного роста объемов данных. Затем возникла новая, современная парадигма – *eScience*, с использованием больших объемов данных, облачных технологий, распределенных сетей и суперкомпьютеров для сложных вычислений. Термин «eScience» предложил англичанин Джон Тейлор (John Taylor), директор министерства науки и технологий Великобритании (United Kingdom's Office of Science and Technology) в 1999 г.

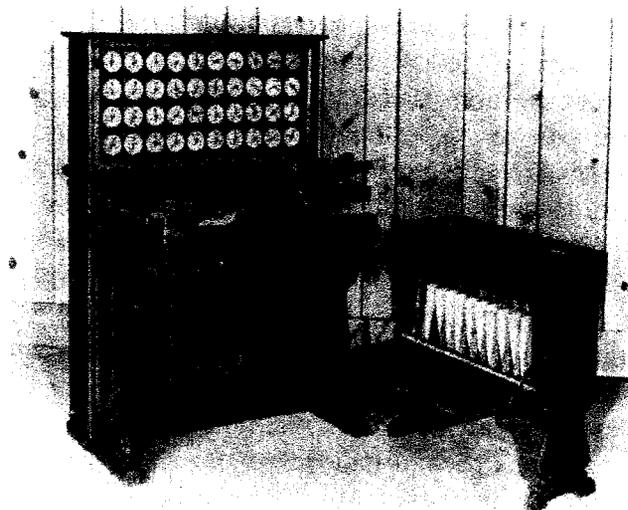
eScience – это наука, изучающая большие объемы оцифрованных (датафицированных) данных. Это понятие очень близко к определению Data Science, которое стало более популярным и вошло в описание новой профессии – научный специалист по данным или Data Scientist. Особых различий между *eScience* и *Data Science* нет, оба определения касаются концепции Больших данных. Но все же Data Science как «наука о данных» (а не наука о методах сбора и анализа оцифрованных данных) имеет всеобъемлющий характер и точнее отражает современную научную парадигму. С точки зрения Data Science все процессы рассматриваются как данные, и только потом разрабатываются методы их оцифровывания и записывания в электронные хранилища для дальнейшего анализа.

В процессе научного познания человечество прошло путь от наблюдения до осмысления собранных фактов, опираясь на свои природные умственные способности. Это первый виток спирали развития, когда человек использовал только свои умственные (вычислительные) способности и простые приспособления (юпана, кипу, абак, счеты и так далее). А затем начался новый виток спирали развития, человек придумал машину (компьютер) как «рычаг» для усиления своих ограниченных вычислительных возможностей. Наблюдения из окружающего мира стали записывать на цифровые носители информации. Компьютер, суперкомпьютер, Искусственный интеллект используют для анализа и осмысления оцифрованных наблюдений. Современный этап охватывает научный аппарат от оцифровывания всего, что только возможно оцифровать, до разработки теории «цифрового

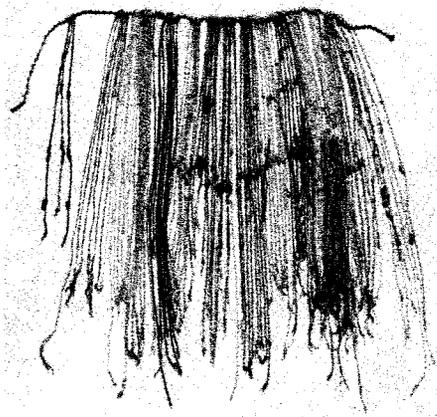
бессмертия», когда информация о живом, заключенная в цифровые хранилища данных, будет так же ценна, как и сама жизнь. Например, полная расшифровка генома живого объекта, с полной записью всей истории его жизни. Если это касается человека – то с полной записью социального пути человека, ежедневной записью параметров его здоровья и истории болезни. Такие данные представляют собой абсолютно новый виток научного познания, для которого потребуются новые концептуальные философские подходы.



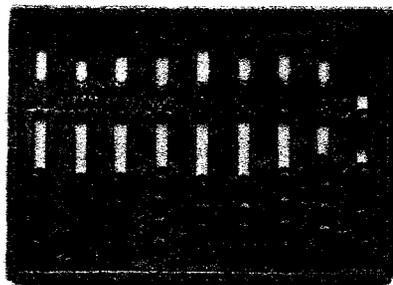
Юпана – древние счеты инков Южной Америки; считается, что для юпаны использовалась фибоначчиева система счисления. Изображение: artkhade.com.



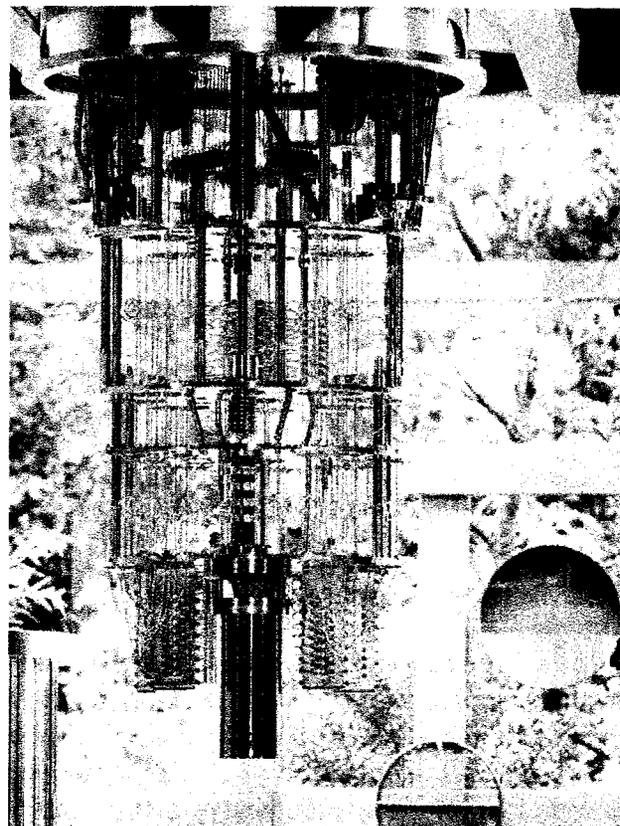
Табулятор (счетная машина) Германа Холлерита, 1890 г. Изображение: evmhistory.ru.



Кипу – древняя счетная система хранения, статистического анализа и воспроизведения информации у инков. Изображение: Collection of the National Museum of the American Indian.



Абак (лат. Abacus – доска) – счетная доска, применялась для арифметических вычислений с V века до н.э. в древних культурах, Древняя Греция и др. Изображение: ru.wikipedia.org.



Квантовый компьютер IBM, 50-кубитный. Изображение: innotechnews.com.

Можно ли предположить, каким будет следующий этап эволюции научной парадигмы в истории человечества? Пройдя этап тотального сбора и записывания цифровой информации, человек должен будет преодолеть вычислительный барьер (известный сегодня как закон Мура), чтобы можно было осмыслить все данные и перейти к масштабному моделированию жизненных процессов и повсеместным вычислениям (ubiquitous computing), то есть к управлению социумом на основе Больших данных. Это будет следующий виток спирали развития в ближайшей перспективе.

Компьютер по определению – это вычислитель (англ. compute – вычислять, считать), устройство для вычислений, использующее последовательные действия, операции. Человечество прошло путь от каменных и деревянных счет до программируемых логических интегральных кремниевых схем и даже до первых квантовых компьютеров. Все это путь вычислителей, сделанных из неживых материалов, использующих физические законы и химические реакции, но очень далеких от сложности устройства самого простого листа растения или даже всего одной клетки.

Нынешний этап развития компьютерной отрасли подошел к определенному барьеру. Американец Гордон Мур (Gordon Moore), основатель корпорации Intel, вывел в 1965 г. эмпирическое правило (его назвали «законом Мура»). Согласно закону Мура число транзисторов в интегральной микросхеме удваивается каждые два года, и этот рост имеет свой предел по вместимости, что привело к разработке параллельных вычислений, то есть к созданию суперкомпьютеров и квантовых компьютеров для продолжения наращивания вычислительных мощностей.

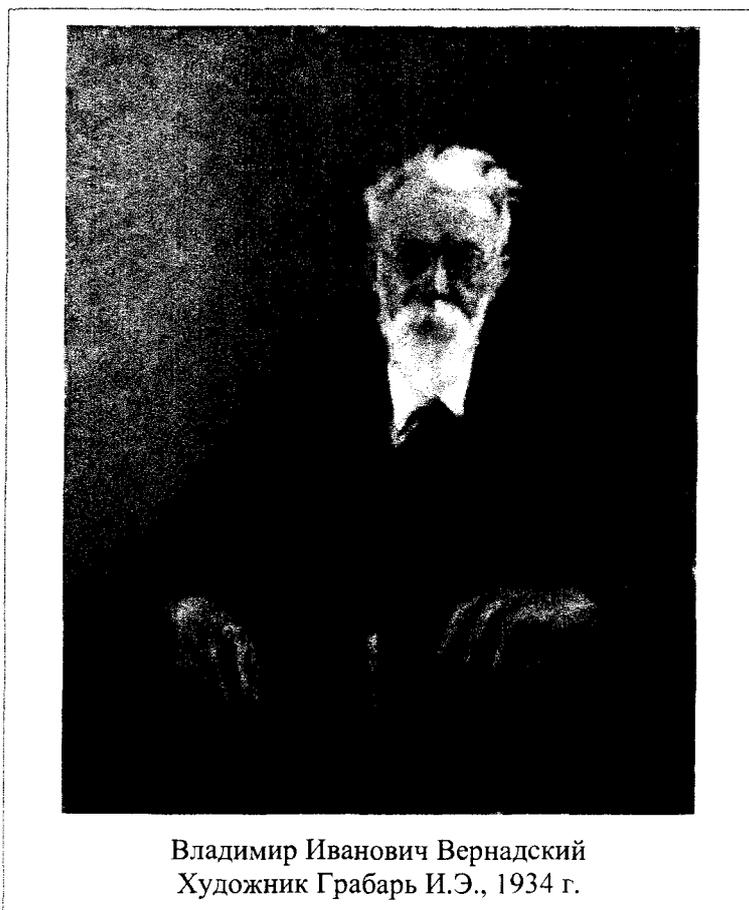


Эволюция вычислительных этапов в истории человечества. Инфографика: Колесниченко О.Ю.

Достигнутый на современном этапе предел, который надо преодолеть, касается глубоких перемен, и не только технологических, а уходящих корнями в мировоззрение. Параллельно с совершенствованием вычислительных устройств, расширяются знания о самом человеке, его биологическом устройстве. Эволюция научных познаний идет от исходной нулевой точки гармонии с природой первобытных людей до вершины этой гармонии, когда человек сможет настолько познать биологическую природу, что будет использовать ее как вычислитель, то есть запрограммировать биологические процессы и существовать на этом более быстром и более мощном вычислительном уровне живой природы.

Обратимся к философу Владимиру Ивановичу Вернадскому (1863–1945 гг.). В 1938 г. он написал статью «Научная мысль как планетарное явление» [6]. В.И. Вернадский в статье обсуждал «биосферу, переработанную научной мыслью», что ученый обозначил как «ноосфера» и относил ее к геологическому явлению, а ум, познания человека – к геологической силе. Стремление человека управлять биологическими процессами заложено в самой эволюции. Сегодняшний уровень знаний и умений в области компьютерных вычислений и техники можно образно сравнить с самым первым шагом, когда человек взял в руки камень в качестве орудия. Возможности по применению этого «кремниевого камня» почти исчерпали себя, и настало время сделать следующий эволюционный шаг.

В связи с этим возникают две проблемы, напрямую касающиеся перспектив Data Science. Первая проблема – это развитие суперкомпьютеров, основанных на кремниевых технологиях (по закону Мура предел этих технологий уже близок). Вторая проблема – переход на новые принципы компьютерного устройства, то есть создание квантовых компьютеров и биокомпьютеров. А в перспективе – создание фотонных компьютеров, более всего природоподобных, так как поток фотонов солнечного света является базовым механизмом возникновения и существования живой природы.



Научные парадигмы меняются вместе со сменой уклада жизни социума. Эволюция социума была очень точно и емко охарактеризована американским философом, социологом и футурологом Элвином Тоффлером (Alvin Toffler, 1928–2016 гг.). Элвин Тоффлер в 1980 г. опубликовал работу, в которой описал развитие человечества как последовательные три периода (три «волны»). Каждая «волна» полностью меняла социум, его культуру, законы и образ жизни. «Первую волну» ученый назвал сельскохозяйственной, в этот период главной движущей силой развития общества было владение землей. «Вторая волна» – индустриальная, развитие определяло производство промышленной продукции. «Третья волна» – информационная, переход к глобальному управлению [7].



Элвин Тоффлер
 Фото: The New York Times, Bettmann.

Владение землей, промышленное производство и информационные системы управления – это совершенно разные периоды развития человечества, но их можно логически связать в последовательную цепь технофункционализма. Каждый новый шаг в научно-техническом прогрессе порождает «рычаги» для достижения нужных целей. «Рычаги» нужны человеку для усиления той или иной своей функции или для приобретения новой функции. Каменные орудия, обработка металла, использование энергии воды и пара, а затем нефти, электричество и атомная энергетика, нынешние информационные технологии и NBIC-конвергенция – все это «рычаги». Они позволяют человеку усиливать разные функции: масштабная обработка земли, быстрое перемещение в пространстве, пролонгирование памяти, мощные вычисления и т.д. Иными словами, «рычаги» дают человечеству возможность выживать, приумножать свою популяцию и осваивать новые пространства для обитания. Информационное управление также относится к категории «рычага», так как позволяет на новом витке возможностей человечества укреплять и совершенствовать развитие социума. Наука о данных и кибернетика призваны совершенствовать управление для *устойчивого развития* человечества.

Идеи российского ученого Николая Дмитриевича Кондратьева (1892–1938 гг.), расстрелянного в годы сталинских репрессий, перекликаются с работой Элвина Тоффлера. Н.Д. Кондратьев первым создал теорию экономических циклов. Между арестами ученый успел опубликовать в 1925 г. свою основную работу «Большие циклы конъюнктуры», в которой была показана циклическая закономерность экономического, социального и культурного развития общества. Последователи Н.Д. Кондратьева уже в постсоветское время перенесли его теорию на современность, продолжив циклы дальше во времени [8, 9]. Л.Е. Гринин с соавторами подсчитали, что с 2020 г. начинается свой отсчет шестой цикл («волна») Кондратьева или К-волна, которая завершит ход кибернетической революции (идущей с середины XX века) и выведет социум на новый уровень развития.



Николай Дмитриевич
Кондратьев
Фото: ru.wikipedia.org.



Кибернетическая революция – это крупнейший в истории человечества технологический переворот, ведущий к появлению саморегулирующихся систем, то есть систем Искусственного интеллекта. Термин «Искусственный интеллект» объединяет разные технологии, основной чертой которых является самоуправляемость, то есть автономность от человека в принятии решений. Самоуправление, осуществляемое искусственными системами, на высшем иерархическом уровне будет касаться управления социумом как процессом, используя Большие данные и сверхмощные вычисления. Далее развитие человечества может столкнуться с **сингулярностью**, то есть с неуправляемым ускорением технологического прогресса и потерей человеком возможности контролировать эволюцию Искусственного

интеллекта. Этот прогноз является основным вызовом современности. Открытым остается вопрос, как избежать такого исхода истории, при котором человек станет инволюционным «придатком» эволюционирующего Искусственного интеллекта.

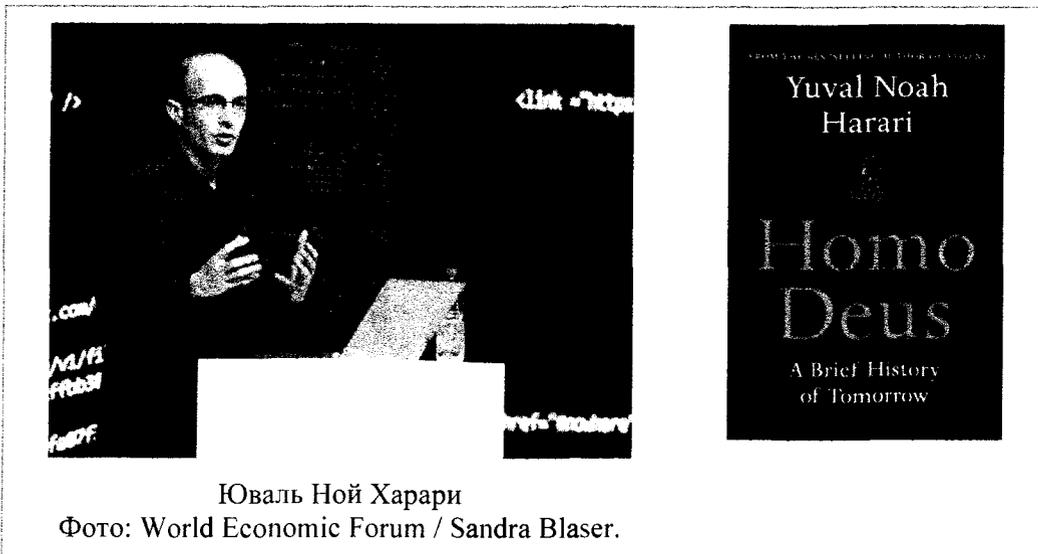
Шестая К-волна (2020–2070 гг.) будет характеризоваться преимущественно прорывом в медицинских технологиях, суммарно обозначенных как **МАНБРИК-технологии**: медико-аддитивно-нано-био-робото-инфо-когнитивные [9]. Медицина станет ведущим социально-экономическим сектором, что связано с возрастающими потребностями и возможностями в получении медицинских услуг; с вытекающей из этих возможностей большей продолжительностью жизни и, соответственно, с необходимостью поддерживать здоровье стареющего населения; с высокими шансами выживания и улучшения качества жизни инвалидов; с новыми репродуктивными и генными технологиями; с новыми технологическими условиями жизни, увлекающими человека в мир дополненной реальности и киборгизации, а также позволяющими расширять зоны обитания (океан и космос). Необходимо также учитывать и тот факт, что лидерство медицины в цифровом мире обусловлено подавляющим преимуществом в генерации Больших данных в сравнении с другими секторами экономики.

§ 3. Системы счисления

Знания о разных системах счисления имеют очень важное значение для специалиста в области науки о данных. Мера (единица измерения) чего-либо – основа организации жизни человеческого общества с самых древних времен. К этой теме можно отнести буквально все – измерение времени, формирование языка, торговые и денежные единицы, культуру, архитектуру, научное познание окружающего мира и т.д.

Мера множества – это область математического изучения. Она имеет свойства неотрицательной величины и аддитивности (значение величины, соответствующее целому объекту, равно сумме значений величин, соответствующих его частям). В математической теории есть сложные меры, объединяющие понятия объема, площади и длины для сложных множеств в трехмерном евклидовом пространстве. Например, мера Жордана, мера Лебега. Эта область математического изучения приобретает особую важность в эпоху появления конвергентных технологий.

Медицина с самого зарождения опиралась на все возможные системы счисления и единицы измерения, какие только могли существовать в рамках своих исторических этапов. Задача нынешнего этапа – наконец-то попытаться понять изученные процессы на уровне клетки с точки зрения математики. И не просто понять, а научиться оцифровывать эти процессы и далее – управлять ими.



Юваль Ной Харари

Фото: World Economic Forum / Sandra Blaser.

Израильский ученый Юваль Ной Харари (Yuval Noah Harari), профессор Еврейского университета в Иерусалиме (Hebrew University of Jerusalem), написал в 2015 г. книгу «Homo Deus: A Brief History of Tomorrow» («Homo Deus / человек божественный: краткая история завтрашнего дня»). В книге Харари рассуждает о том, что биологические организмы, в том числе человек – это биохимические алгоритмы, а происходящая технологическая революция ведет к появлению иного вида человека, отличающегося от Homo Sapiens / человека разумного. Ученый предлагает термин «датаизм», обозначающий власть данных, равную религии [10]. Все человеческое уходит в этой новой религии на второй план, так как путь развития проложен только через данные и вычисления. Датаизм как идеология лежит в основе создания нейроинтерфейсов и киборгизации.

Система счисления – это символический метод записи чисел, представление чисел с помощью письменных знаков, цифр. Число является мерой для описания множества, количества. **Цифры (числовые знаки)** могут быть разными, широко распространены арабские цифры (знаки 0, 1, 2, 3, 4, 5, 6, 7, 8, 9) и римские цифры (I, V, X, L, C, D, M). Известны также шестнадцатеричные цифры, относящиеся к шестнадцатеричной позиционной системе счисления (0, 1, 2, 3, 4, 5, 6, 7, 8, 9, A, B, C, D, E, F). В древних культурах существовали разные

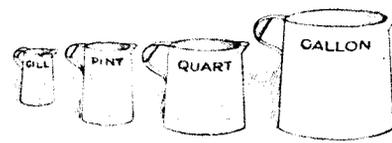
знаки, обозначающие числа – буквы алфавита, точки и линии, символы. Системы счисления различаются по правилам, на основе которых строится порядок записи чисел с помощью цифр и комбинации цифр.

Системы счисления разделяют на позиционные и непозиционные. **Позиционные системы счисления могут быть однородными и смешанными.** В позиционной системе значение каждой цифры зависит от позиции (разряда). Существуют следующие позиционные системы: двоичная (используется для программирования), троичная, четверичная, восьмеричная (используется для программирования), десятичная (широко применяется), двенадцатеричная (использовалась в древности), шестнадцатеричная (используется для программирования), двадцатеричная (использовалась у майя и ацтеков), сорокаичная (использовалась в древности), шестидесятеричная (использовалась древнегреческими астрономами, в настоящее время используется для измерения времени суток).

Сегодня повсеместно используется десятичная позиционная система счисления (целочисленное основание равно 10). В ней числа записывают арабскими цифрами 0, 1, 2, 3, 4, 5, 6, 7, 8, 9. Эта система связана с анатомией человека, так как она возникла от счета на пальцах рук. Например, число 555, в нем три одинаковые цифры, но в зависимости от позиции они соответствуют разным числам; число считывается слева направо от большего разряда к меньшему: пять сотен + пять десятков + пять единиц.

Двоичная система применялась с древних времен и использовалась в 200 г. до н.э. в Индии. В Средние века в Англии, например, разновидность двоичной системы закрепились как мера в торговле вином, что нашло отражение и в языке:

два джилла (Gill) = полуштоф (Stof), два полуштофа = пинта (Pint), две пинты = кварта (Quart), две кварты = потл (Potlle), два потла = галлон (Gallon), два галлона = пек (Peck), два пека = полубушель, два полубушеля = бушель (Bushel), два бушеля = килдеркин (Kilderkin), два килдеркина = баррель (Barrel), два барреля = хогзхед (Hogshead), два хогзхеда = пайп (Pipe), два пайпа = тан (Tun).



Двоичная система счисления является позиционной, с целочисленным основанием, равным 2. В ней числа имеют всего две цифры 0 или 1. Этот простой принцип используется в компьютерах, так как он позволяет минимизировать помехи и сбои. Если ток в микросхеме достигает пороговой величины, то это значение 1, если нет, то это значение 0. То же касается и индукции магнитного поля, которая может достигать пороговой величины или же нет.

Умножение в двоичной системе очень простое: $0 \times 0 = 0$; $0 \times 1 = 0$; $1 \times 0 = 0$; $1 \times 1 = 1$. Деление производится углом, как и в десятичной системе. При сложении двух единиц осуществляется перенос в более старший разряд $1 + 1 = 10$. В случаях сложения $0 + 0 = 0$, при сложении $0 + 1 = 1$; $1 + 0 = 1$. Вычитание: $0 - 0 = 0$; $1 - 0 = 1$; $1 - 1 = 0$; $0 - 1 = 1$ с заемом из более старшего разряда.

Современная двоичная система с 0 и 1 была описана Готфридом Вильгельмом Лейбницем (Gottfried Wilhelm von Leibniz, 1646–1716 гг.) в труде «Explication de l'Arithmétique Binaire». Использовать двоичную систему в компьютерах для представления чисел предложил Джон фон Нейман (John von Neumann, 1903–1957 гг.), работая в 1944–1945 гг. над первыми ламповыми компьютерами ENIAC (Electronic Numerical Integrator and Computer) и EDVAC (Electronic Discrete Variable Automatic Computer) в Электротехнической школе Мура Пенсильванского университета (Moore School of Electrical Engineering of University of Pennsylvania). До появления этих компьютеров числа обрабатывались вычислительными машинами в десятичном виде. Джон фон Нейман с коллегами показали удобство и простоту двоичной системы для арифметических и логических операций, выполняемых компьютерами. В более широком диапазоне все нововведения, предложенные Джоном фон Нейманом, известны сегодня как «Архитектура фон Неймана» (модель фон Неймана или Принстонская архитектура).

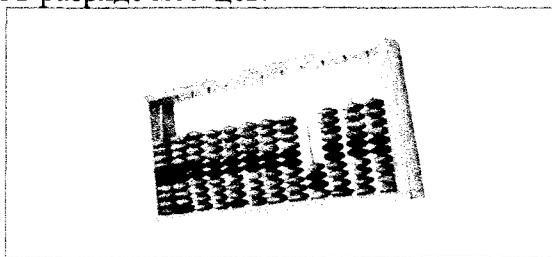
Перевод числа из десятичной позиционной системы в двоичную позиционную систему осуществляется через последовательное деление числа на 2 (получается «у» на каждом шаге), при этом в зависимости от условия «число₁₀» – 2у = 0 или 1, записывается соответственно 0 или 1 в каждую позицию до последнего шага деления на 2. В результате этого деления получается число из остатков от деления на каждом шаге, из нуля и единиц. Последняя единица будет старшим разрядом в двоичном числе, а все число нужно записать справа налево: 11111000. Например, число $248_{10} = 11111000_2$.

248	/ 2								
248	124	/ 2							
0	124	62	/ 2						
	0	62	31	/ 2					
		0	30	15	/ 2				
			1	14	7	/ 2			
				1	6	3	/ 2		
					1	2	1	/ 2	
						1	0	0	

Можно произвести обратный перевод числа из двоичной системы в десятичную.

разряд								
7	6	5	4	3	2	1	0	
1	1	1	1	1	0	0	0	$= 0 \times 2^0 + 0 \times 2^1 + 0 \times 2^2 + 1 \times 2^3 + 1 \times 2^4 + 1 \times 2^5 + 1 \times 2^6 + 1 \times 2^7$
Число:		0 +	0 +	0 +	8 +	16 +	32 +	64 + 128
								Число ₁₀ = 248

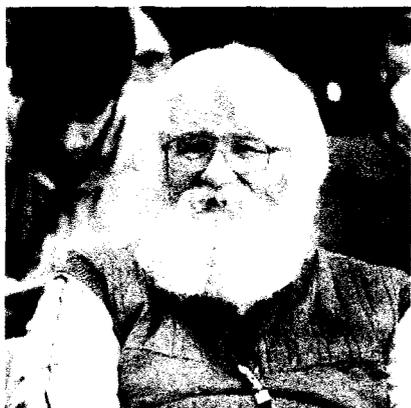
Двоичная и десятичная системы **однородные**, так как в них используется только одна цифра из первичного набора цифр для каждого разряда. К **смешанным системам** относят, например, систему измерения времени. В каждом разряде (позиции) набор символов (цифр) может отличаться от других разрядов. Так, в разряде секунд и минут используется 60 различных символов (от «00» до «59»), в разряде часов существует 24 разных символа (от «00» до «23»), в разряде суток применяются символы от «1» до «365» (в високосный год до «366»), в разряде недель добавляется кодирование последовательности семи дней словами, то же применяется и в разряде месяцев.



Унарная система счисления – это целочисленная система с основанием, равным 1. Русские счёты основаны на унарнодесятичной системе представления цифр с одной избыточной унарнодесятичной цифрой.

Непозиционные системы не имеют позиций, то есть разрядов. Каждому числу множества соответствует какой-то один символ. Например, урожай из 100 плодов будет записан как 100 черточек-засечек. Одна черточка – один плод. Непозиционной была до н.э. древнеримская система счисления.

Российский математик Николай Николаевич Непейвода в докладе на Национальном Суперкомпьютерном Форуме в 2017 г. (г. Переславль-Залесский, Институт программных систем им. А.К. Айламазяна РАН) предложил **современную научную классификацию систем счисления** (доклад «Представление чисел с точки зрения вычислений» [11]). Классификация Н.Н. Непейводы включает: 1/ Позиционные системы традиционного типа; 2/ Вариации позиционных систем; 3/ Аддитивные системы; 4/ Представления, принципиально отличающиеся от систем счисления (логарифмическое представление и др.).



Николай Николаевич Непейвода
Фото: Pereslavl Foto.



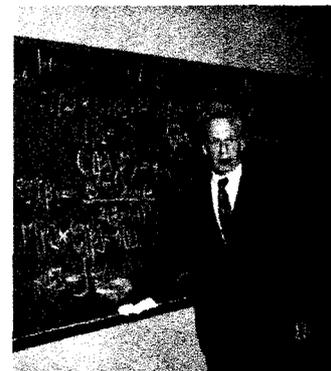
Национальный суперкомпьютерный форум
Переславль-Залесский, Фото: rg.ru.



Электронно-вычислительная машина
«Сетунь», 1959 г. Фото: auri.info.



Николай Петрович
Брусенцов
Фото: ru.wikipedia.org.



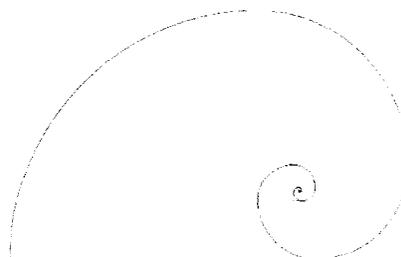
Сергей Львович Соболев
Фото: biblioatom.ru.

Позиционная целочисленная симметричная **троичная система счисления** имеет три цифры 0, 1, и -1 (основание равно 3). Итальянский математик Фибоначчи (Fibonacci – это псевдоним, настоящее имя – Леонардо Пизанский, Leonardo Pisano, около 1170–1250 гг.) создал троичную систему после изучения древних индийских и арабских рукописей. Эту систему в середине XX века применили советские ученые Сергей Львович Соболев (1908–1989 гг.), Николай Петрович Брусенцов (1925–2014 гг.) и другие соавторы при разработке российской троичной ЭВМ «Сетунь» [12]. Данный компьютер на основе троичной логики до сих пор не имеет аналогов в мире. Он работал с использованием двухбитных двоичнокодированных троичных цифр (2-Bit Binary Coded Ternary, 2B BCT) с использованием всех четырех кодов из четырех возможных (при этом два из четырех кодов кодируют одну и ту же троичную цифру из трех). Комбинации 2B BCT были следующими: (0,0) – «0»; (1,1) – «0» (не использовалась); (0,1) – «-1»; (1,0) – «+1». Единицей памяти ЭВМ «Сетунь-70» был **трайт**, который состоял из шести **трипов** (троичный код). Троичность или триалектика имеет логические подходы, которые следует учитывать в Data Science. Это три формы единства множества данных: дискретность, континуальность (множество связано в континуум) и иерархичность. Философское понимание множества как триединого заключается в следующем: всякое целое едино (1), оно одновременно является и частью большего целого (2), и само состоит из меньших частей (3) [13].

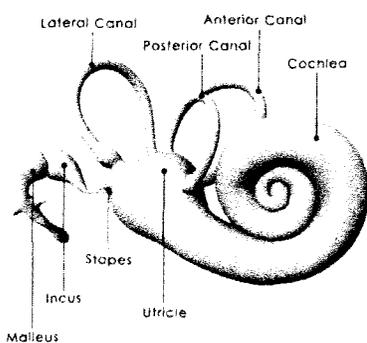
Фибоначчи, изучив манускрипты Древней Индии, вывел последовательность чисел, известную сегодня как последовательность **чисел Фибоначчи**: 0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, 377, 610, 987, 1597, 2584, 4181, 6765, 10946, 17711, и т.д. (Fibonacci numbers, A000045, The On-Line Encyclopedia of Integer Sequences, The OEIS Foundation). В этой последовательности каждое последующее число равно сумме двух предыдущих чисел. Гармонию последовательности Фибоначчи можно наглядно проиллюстрировать на спирали, в которой стороны квадратов соотносятся как числа Фибоначчи. Данный геометрический феномен широко распространен в природе (расположение зерен подсолнуха; строение сосновой шишки, ячеек ананаса, початка кукурузы; рост листьев и лепестков цветов; форма раковины моллюсков; пропорции строения насекомых и животных, пропорции тела человека).



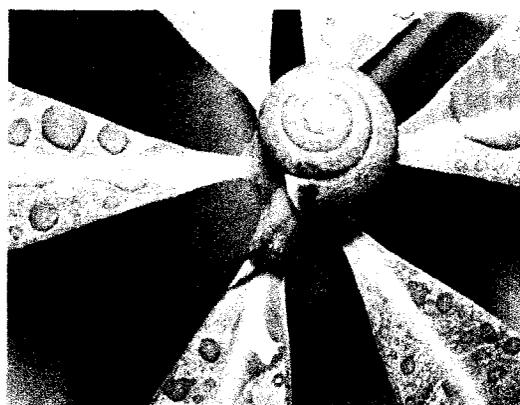
Фибоначчи (Леонардо Пизанский)
Изображение: ru.wikipedia.org.



Последовательность Фибоначчи.

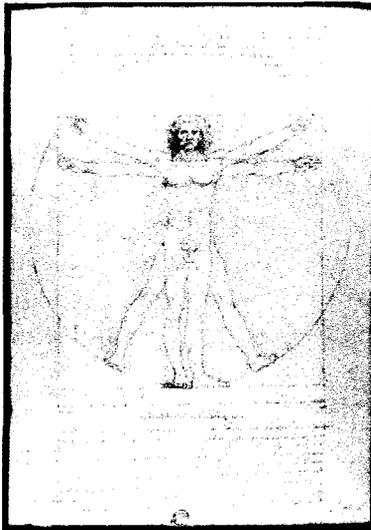


Улитка (cochlea) внутреннего уха человека.
Изображение: 7ActiveStudio/Fotolia.



Из последовательности Фибоначчи получено число Φ (фи), равное 1,618. Число Φ представляет собой универсальное соотношение частей в целом в разных природных явлениях, а также в древней архитектуре – золотая пропорция 1:1,618. Это соотношение найдено в известном рисунке Леонардо Да Винчи (Leonardo da Vinci, 1452–1519 гг.) «Витрувианский человек» (Homo vitruvianus). Рисунок является иллюстрацией для книги, посвященной трудам античного римского архитектора Витрувия (Vitruvius). Пропорции Homo vitruvianus считают «каноническими».

Числа представляются как суммы чисел Фибоначчи в аддитивной системе счисления. Такую систему называют «фибоначчиева система». По мнению математика Н.Н. Непейводы эта система имеет перспективы при применении для параллельных суперкомпьютерных вычислений. Н.Н. Непейвода считает, что на современном, меняющемся, переломном этапе науки особенно важно задумываться о том, что такое числа и какие числа применяются для вычислений [11].



Витрувианский человек,
рисунок Леонардо Да Винчи, 1490–1492 гг.
Фото: lucnix.be, Luc Viatour.

Французский математик Огюстен Коши (Augustin Louis Cauchy, 1789–1857 гг.) предложил в 1840 г. десятичную систему с отрицательными числами от -5 до 5 . Такая система нашла применение для параллельных вычислений с большими числами. Есть система счисления с цифрами 0 и 1 , в которой основанием является $\sqrt{2}$. Есть системы, в которых основания – это числа Пизо (французский математик Charles Pisot, 1910–1984 гг.) или корни уравнений.

Голландский математик Лейтзен Брауэр (Luitzen Brouwer, 1881–1966 гг.) предложил в 1921 г. систему с перекрытием представления чисел на отрезке (каждая цифра представляет собой не число, а интервал значений) [11]. Основание определяет, на сколько равных отрезков делится предыдущий разряд, а перекрытие – на какую долю пересекаются два соседних отрезка. Отрезок $[0, 1]$ делится на пересекающиеся отрезки $[0, 2/3]$, $[1/3, 1]$. Н.Н. Непейвода с соавторами разработал троичную систему с перекрытием отрезков соседних цифр на $1/2$ их длины [14, 15]. Она называется «три половинки» и имеет интервалы 0 : $[0, 1/2]$, $1/2$: $[1/4, 3/4]$, 1 : $[1/2, 1]$. Российский математик считает, что система с перекрытием отрезков соседних цифр на $1/2$ их длины оптимальна и надежна для быстрых операций.

Логарифмические системы представляют число как знак и логарифм его абсолютной величины. Такие системы используются для обработки сигналов. В 1970-х гг. российский ученый Дмитрий Александрович Поспелов предложил для компьютерных вычислений объединить логарифмическую систему с системой остаточных классов (система счисления, основанная на модулярной арифметике). Математик работал над удвоением памяти и увеличением скорости счета. Д.А. Поспелов более известен как исследователь в области Искусственного интеллекта, он был президентом советской (в последующем, российской) Ассоциации искусственного интеллекта.

§ 4. История моделирования и кибернетики

Моделирование отражает стремление исследователя систематизировать знания. Модель (от лат. *modulus* – «образец») создается для того, чтобы изучить какое-то явление, научно его описать и понять. Моделирование может быть разным, в зависимости от того, с какой целью и какими способами оно проводится и что изучается. Математическое и компьютерное (или суперкомпьютерное) моделирование в настоящее время широко применяется во всех областях и сферах, став самой важной частью **кибернетики**.

Начало кибернетики связывают с известным ученым Андре-Мари Ампером (Andre-Marie Ampere, 1775–1836 гг.). Французский физик, математик и естествоиспытатель более известен тем, что ввел в физику понятие электрического тока. Андре-Мари Ампер первым предложил термин «кибернетика» в своем научном труде «Опыт о философии наук, или аналитическое изложение естественной классификации всех человеческих знаний» (1834), рассматривая ее как науку об управлении государством.

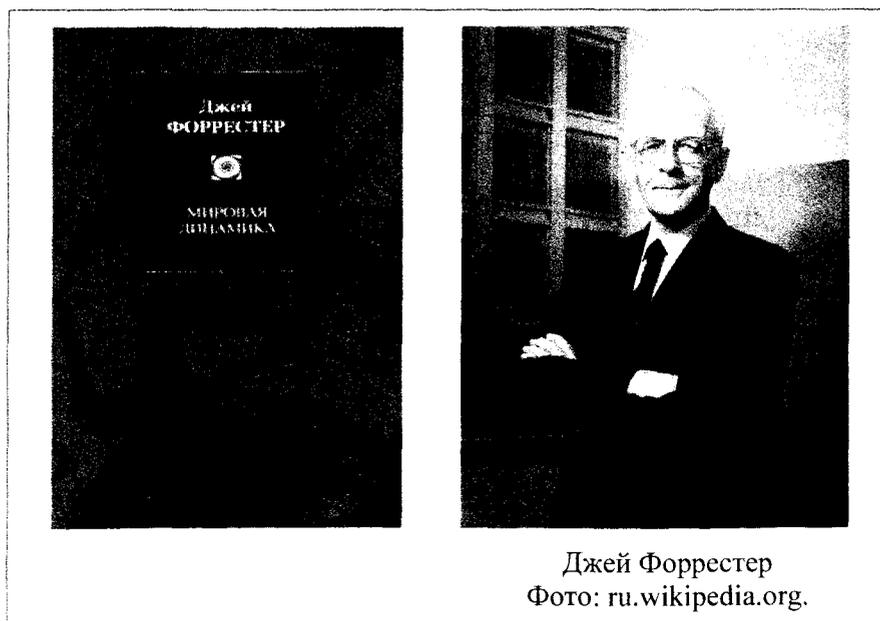
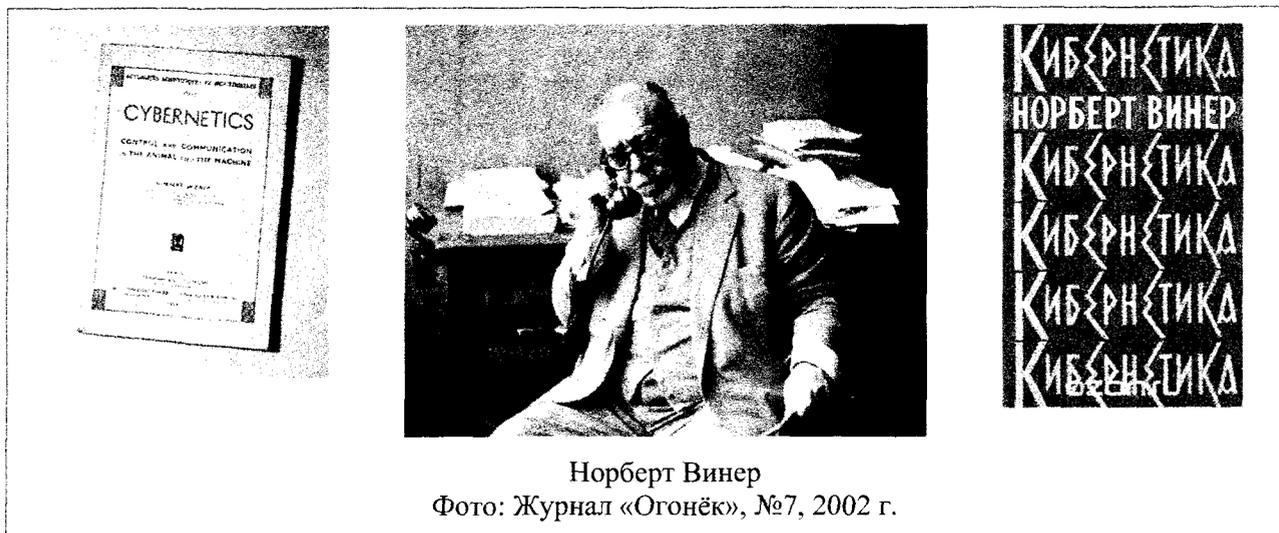


Андре-Мари Ампер
Изображение: Музей Галилея,
Флоренция, Италия

Второе «рождение» кибернетики произошло уже в XX веке благодаря американскому математику Норберту Винеру (Norbert Wiener, 1894–1964 гг.). Ученый написал в 1948 г. книгу «Кибернетика, или управление и связь в животном и машине» (*Cybernetics or Control and Communication in the Animal and the Machine*) [16]. Книга стала классическим учебником для тех, кто в последующем развивал кибернетическое направление. В своей книге Норберт Винер представил системы с обратной связью и самоорганизующиеся системы. Особое внимание ученый уделял универсальным кибернетическим подходам к пониманию устройства биологического организма, человека, его нервной системы и гомеостаза, а также разобрал и устройство общества. Норберт Винер писал: «Любой полный курс кибернетики должен включать в себя тщательный и подробный обзор гомеостатических процессов... Один из выводов настоящей книги состоит в том, что всякий организм скрепляется наличием средств приобретения, использования, хранения и передачи информации» [16]. В интервью советскому научно-популярному журналу «Природа», которое Норберт Винер дал в 1960 г. во время визита в Москву, где он принял участие в Конгрессе Международной федерации по автоматическому управлению, ученый сказал: «Различие между человеком и машиной, прежде всего, заключается в том, что в организме человека число элементов по порядку величин во много раз больше, чем обладает машина. Из этого непосредственно вытекает, что организация элементов в организме настолько сложна, что при помощи наших современных логических средств мы не можем еще овладеть этой сложностью. Я бы даже не решился сейчас дать определение понятия “сложность”» [17].

Кибернетика – это наука о закономерностях в процессах управления и передачи информации в машинах, живых организмах и обществе. В кибернетическом моделировании

особое внимание уделяется изучению обратной связи. Компьютеры стали инструментом для совершенствования методов кибернетического познания, а Большие данные вывели кибернетику на новый уровень возможностей изучения процессов в различных областях. Определенно, машинное моделирование с помощью Искусственного интеллекта позволит в будущем приблизиться к природной сложности биологических процессов в организме человека, что станет настоящим цивилизационным прорывом как для медицины, так и для робототехники. Возникшая во втором десятилетии XXI века концепция экономики, основанной на данных – Data driven economy, – полностью полагается на кибернетические методы аналитики Больших данных и управление сложными динамическими системами, в том числе с помощью Искусственного интеллекта. Применение кибернетических подходов в этой области позволит решить многие социальные проблемы.



Если Норберт Винер считается основателем фундаментальных теоретических подходов в кибернетике, то Джей Форрестер (Jay Wright Forrester, 1918–2016 гг.) был одним из первых и чрезвычайно смелых ученых, кто воплотил кибернетические подходы на практике. Джей Форрестер, будучи профессором Массачусетского технологического института (Massachusetts Institute of Technology, MIT) в Кембридже (США), написал в 1971 г. книгу «Мировая динамика» [18]. В книге он обобщил свой уникальный и пионерский опыт по разработке кибернетической социально-экономической модели динамики глобальных процессов «World–

1» и ее доработанной версии «World-2» («Мир-1» и «Мир-2»). Работа выполнялась по заказу Римского клуба. Модели «Мир-1» и «Мир-2» позволили ученому создать теорию *системной динамики*. Это направление изучает сложные системы, их поведение во времени, в зависимости от структуры элементов системы и от взаимодействия между ними (причинно-следственные связи, обратная связь, влияние среды и др.).

Проанализировав свои кибернетические модели «Мир-1» и «Мир-2», Джей Форрестер предложил внедрить на общемировом уровне глобальный контроль над промышленным загрязнением и рождаемостью, используя математическое моделирование глобального развития. Также Джей Форрестер известен как разработчик одного из первых универсальных компьютеров Whirlwind I («Вихрь-1»). Джей Форрестер показал, что *«мир вычисляем»*. Он писал: «Люди начали понимать, что взаимодействия внутри целого более важны, чем простая сумма отдельных его частей». Динамические системы или модели, системная динамика – это изучение процесса в пространстве и во времени как непрерывного континуума взаимодействий с обратной связью.

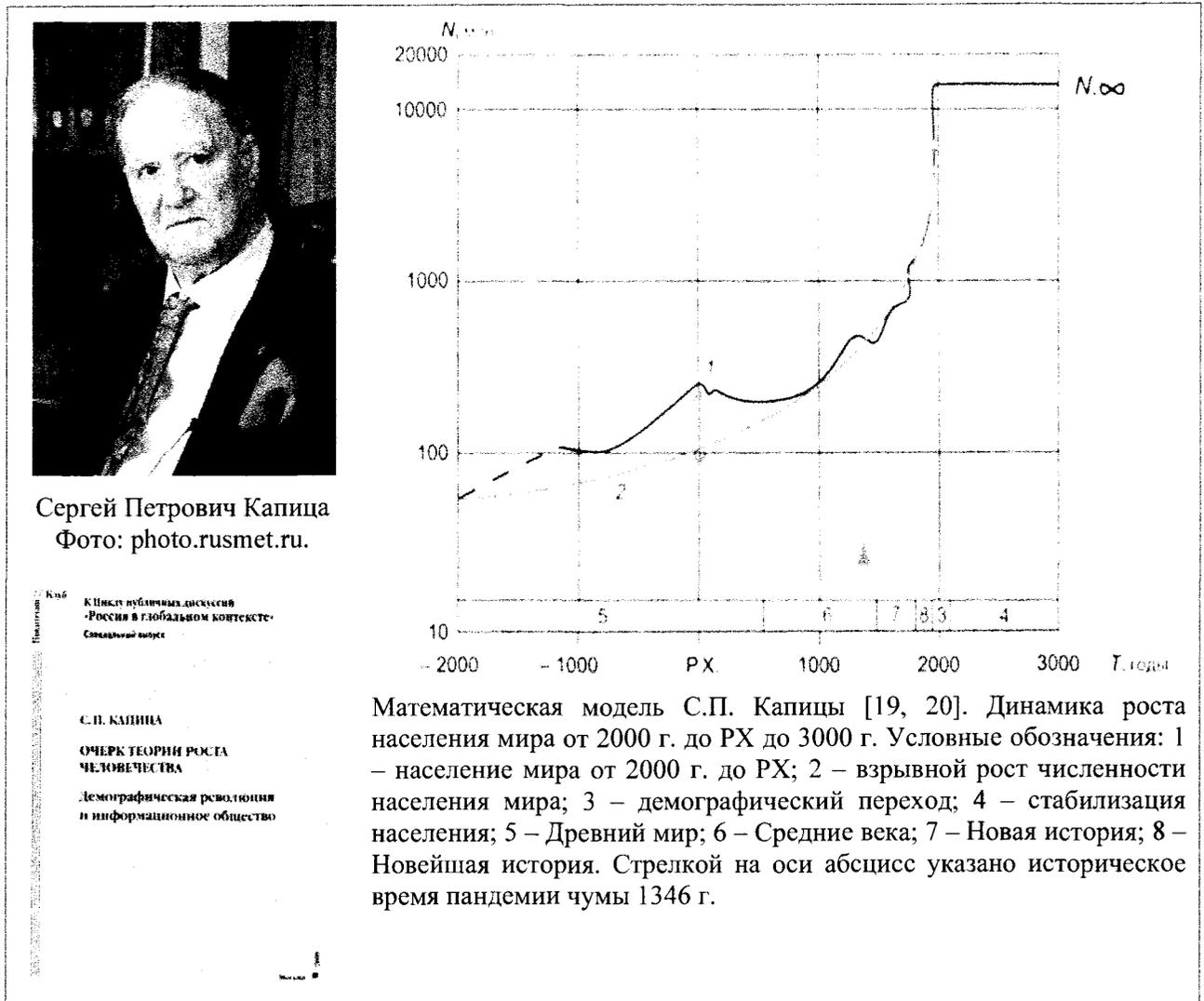
С выводами Джея Форрестера о необходимости ограничения роста и развития активно спорили советские ученые. Это самый первый в истории международный спор об этических аспектах кибернетических моделей, состоящих из последовательных уравнений, но не имеющих человеческой компоненты. Несмотря на то, что модель глобального развития Джея Форрестера появилась благодаря новым информационным технологиям, его алгоритмы опирались на старую идею Томаса Мальтуса (XVIII век), согласно которой предел роста численности человечества определяется ресурсами планеты.

Информационные технологии (ИТ) можно рассматривать как «рычаг», который не только позволяет людям общаться друг с другом на дальних расстояниях (что невозможно с точки зрения биологической природы человека) или пролонгировать свою память (благодаря записи всех знаний в пространстве Интернета с открытым доступом к ним), но самое главное, ИТ стали «рычагом», усиливающим такое свойство людей, как коллективное взаимодействие, связанность. Физик и философ с мировым именем, представитель России в Римском клубе Сергей Петрович Капица (1928–2012 гг.) создал ставшую известной математическую модель глобального развития, в основе которой заложено информационное коллективное взаимодействие людей (цепная реакция информации) [19, 20].

С.П. Капица подошел к построению модели гиперболического роста численности населения Земли не с точки зрения ресурсов планеты, оборота товаров, продуктов и финансовых потоков, а с точки зрения социального развития человеческого общества. Ученый показал, как социально-информационные процессы меняли изначально биологическую предопределенность человечества. Так, первый миллиард людей на протяжении всего периода существования цивилизации появился лишь к началу XIX века, с наступлением информационного бума и научно-технического прорыва (развитие связи и печати, появление телеграфа), укрепившего коммуникации между людьми. А уже к 1930 г. население Земли достигло 2 миллиардов человек. Это «ускорение» продолжается, что является характерной особенностью третьего тысячелетия (сегодня численность населения Земли превысила 7 миллиардов человек).

С.П. Капица утверждал, что информационная эра показала несостоятельность теории Томаса Мальтуса (популяционный принцип Мальтуса, или мальтузианский популяционный принцип). Общество представляет собой социокультурную систему коммуникаций. С.П. Капица называл общество сложной динамической системой в пространстве и времени и подчеркивал, что рост и развитие популяции человечества основаны на коллективном взаимодействии, происходящем по законам физики сложных систем [20]. Ученый связал демографический рост населения планеты со становлением информационного общества и информационной эры, с развитием сознания общества, культуры, базы знаний. Через передачу информации от поколения к поколению проходила социальная эволюция. По мере совершенствования технологий связи и передачи информации человечеству стала доступна передача информации не только вертикально из поколения в поколение, но и горизонтально

внутри социума на любом расстоянии. Общество становится все более устойчивым как система.

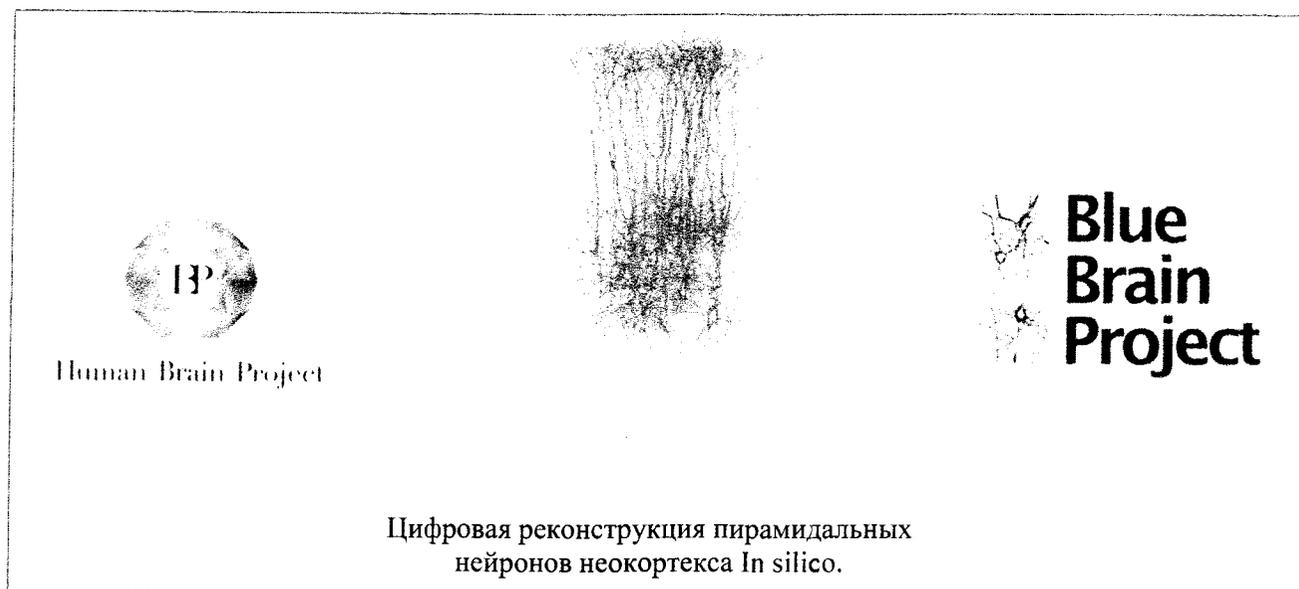


На основе своей модели С.П. Капица ввел понятие ответственного управления обществом. По мнению С.П. Капицы, в отличие от Джея Форрестера, остановку роста и развития населения планеты может вызвать не дефицит ресурсов, а информационный кризис управления и ценностей. Ресурсы всегда будут вторичной проблемой, а первичная проблема – уровень коллективного сознания человечества, видение людьми перспектив развития и *образа будущего*, ответственное глобальное управление. С.П. Капица предупреждал, что в XXI веке человечество ожидает резкая смена жизненных ценностей, изменение гуманитарной парадигмы, что и вызывает больше всего опасений в сравнении даже с экологическими и энергетическими проблемами [20]. Связанность всех процессов в единое информационное поле – это то, что дают современные ИТ. Усиление в информационном мире связанности социума, культура и ценности, коллективное взаимодействие определяют ход человеческой истории – развитие или кризис.

С подходами Джея Форрестера спорил и другой российский ученый с мировым именем – Никита Николаевич Моисеев (1917–2000 гг.). Ученый руководил лабораторией в Вычислительном центре Академии Наук СССР, в которой была разработана математическая модель экологических последствий ядерной войны – ядерной зимы. Эта модель оказала существенное влияние на заключение договоров между СССР и США по ограничению гонки ядерных вооружений. Сам термин «ядерная зима» появился в 1983 г. после выступления в США с докладом физиков Георгия Сергеевича Голицына и Карла Сагана (1934–1996 гг.).

Следует упомянуть об американском нобелевском лауреате Джоне Нэше (John Nash, 1928–2015 гг.), который внес большой вклад в развитие междисциплинарной кибернетической теории игр. Это направление моделирования приобретает особую актуальность сегодня при создании Искусственного интеллекта, так как он рассматривается как виртуальный агент или набор виртуальных агентов, взаимодействующих между собой.

Моделирование в медицине в настоящее время очень широко распространено. Например, крупнейший в мире международный проект по моделированию мозга Blue Brain Project, который осуществляет компания IBM. Для вычислений и моделирования структур головного мозга человека используется суперкомпьютер Blue Gene. Другой аналогичный проект – The Human Brain Project – финансируется Евросоюзом. Оба эти проекта связаны с Федеральной политехнической школой Лозанны в Швейцарии (Ecole polytechnique federale de Lausanne).



Основоположниками кибернетического моделирования и системного подхода в медицине стали российские физиологи Петр Кузьмич Анохин (1898–1974 гг.) и Константин Викторович Судаков (1932–2013 гг.). Ученые работали в 1-м Московском медицинском институте имени И.М. Сеченова. П.К. Анохин создал теорию функциональных систем (кибернетическая физиология), а его ученик К.В. Судаков развил это учение, сделав его масштабной дисциплиной, которая вначале перевернула подходы в физиологии, а затем и в клинической медицине.

По определению К.В. Судакова, функциональная система – это саморегулирующаяся динамическая организация, деятельность всех компонентов которой содействует достижению полезных для организма в целом приспособительных результатов [21]. Обратная афферентация (поступающая с периферии в нервные центры сигнализация о состоянии различных физиологических функций и достигнутом результате), представляющая собой информационный процесс в функциональных системах организма, была открыта П.К. Анохиным еще до того, как Норберт Винер описал кибернетическую обратную связь в качестве общей закономерности. К.В. Судаков предложил общие принципы взаимодействия отдельных элементов в живом организме – **корреляция** (сосуществование отдельных элементов в целом организме); **регуляция** (подчинение структур или процессов в целях достижения интересов целого организма); **саморегуляция** (возвращение отклонившегося показателя от нормы к исходному нужному уровню). Эти три принципа являются базовыми с точки зрения кибернетического моделирования любого процесса, связанного с живыми организмами.



Константин Викторович
Судаков
Фото: uzrf.ru.

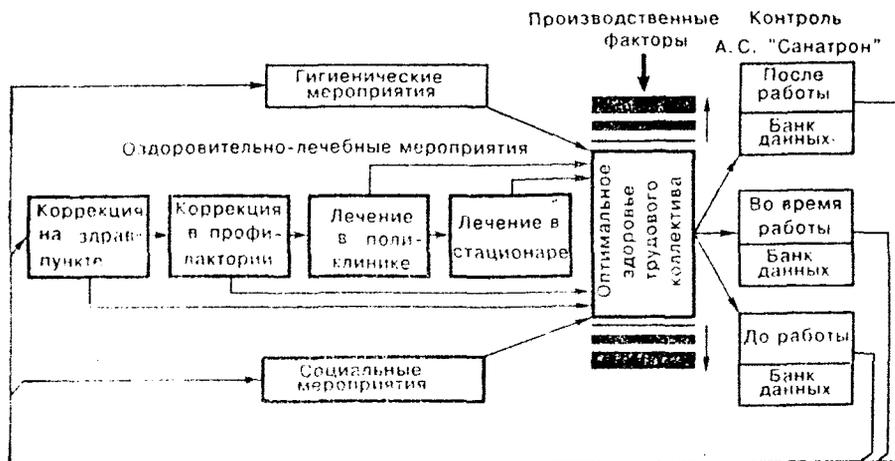
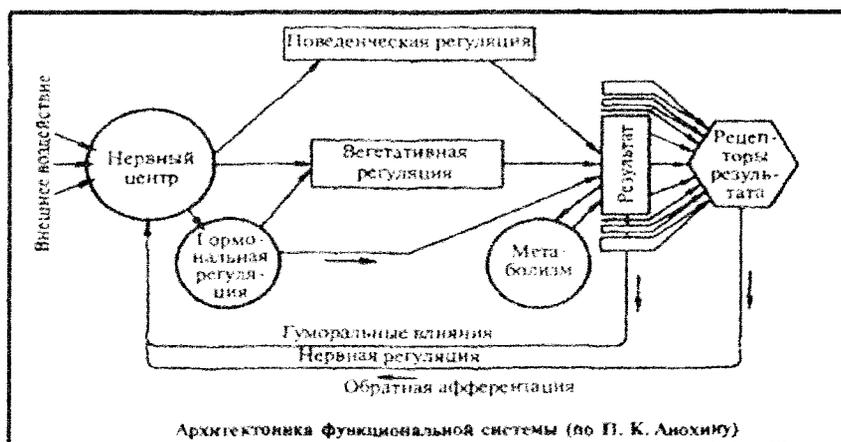


Схема функциональной социомедицинской системы для работающего населения, разработанная К.В. Судаковым [21].



Петр Кузьмич Анохин
Фото: ru.wikipedia.org.



Архитектура функциональной системы (по П. К. Анохину)

Новое слово в моделировании на современном этапе сказал американский физик Альберт-Ласло Барабаши (Albert-Laszlo Barabasi), изучающий науку о сетях (Network Science и Human dynamics). Альберт-Ласло Барабаши вместе с Река Альберт (Reka Albert) в 1999 г. разработали алгоритм генерации безмасштабных сетей с использованием принципа предпочтительного присоединения – Модель Барабаши-Альберт (Barabasi-Albert model). Безмасштабные сети – это графы с вершинами и ребрами, которые встречаются как в живой природе, так и в социальной жизни людей, особенно в сфере коммуникаций. Модель Барабаши-Альберт применима к изучению связей разных заболеваний через общие генетические поломки или коморбидности заболеваний (то есть связей заболеваний, объединенных патогенетическим механизмом) через общие биохимические каскады.

Моделирование, особенно в медицине, опирается на причинно-следственную связь. Модель может быть эталоном для сравнения, или отражением существующего процесса как он есть, или же несуществующим прогнозом развития процесса в будущем. В Data Science имеет значение и построение эталона какого-то процесса, чтобы равняться на него; и создание реальной модели существующего процесса, чтобы его оценить и изучить; и построение прогностической модели. Сложность моделирования в медицине требует применения суперкомпьютерных мощностей.

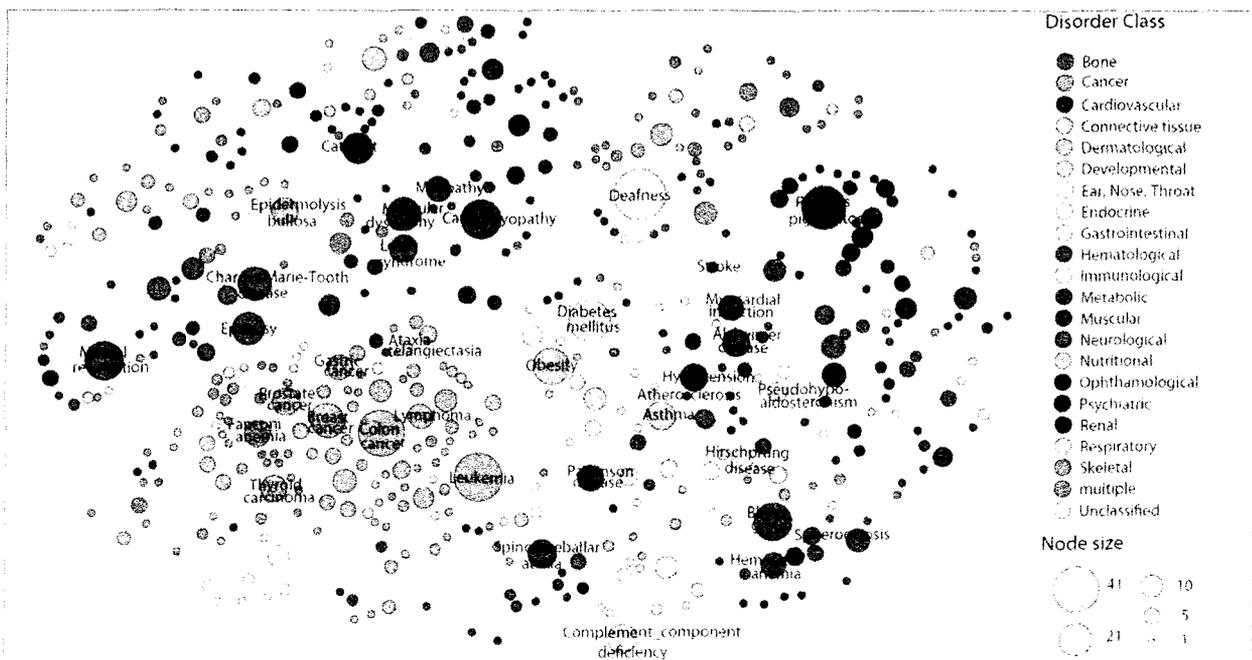


Альберт-Ласло Барабаши
Northeastern University,
Harvard University
Фото: citatis.com.



Река Альберт
Pennsylvania State University
Фото: ralbert.me.

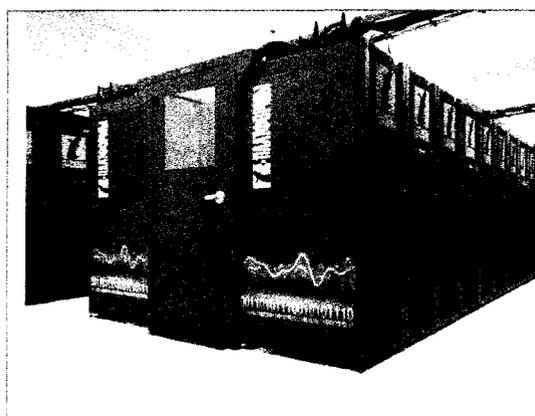
По мере становления квантовых вычислений причинно-следственные подходы перестали быть господствующими в науке, так как развитие квантовой теории разрушает ранее существовавшие научные понятия. Например, квантовая суперпозиция – это неопределенное состояние кванта. Квант может быть одновременно в разных, взаимоисключающих состояниях (как если бы человек стоял одновременно на ногах и вверх ногами). А результат какого-то квантового процесса всегда случаен и определяется только по факту его фиксации. То есть, с точки зрения квантовой физики, на данном этапе знаний человечества нельзя предугадать, что будет в случае фиксации, то есть измерения результата. Образно, человек будет стоять на ногах или вверх ногами – определится только в момент, когда исследователь будет фиксировать это состояние, то есть посмотрит на этого человека. Квантовый этап научного познания требует новых концептуальных философских разработок, которые не будут теоретически отстраненными от практических аспектов Data Science, а напрямую определяют судьбу квантовой информатики и всех вытекающих из этого кибернетических методов моделирования.



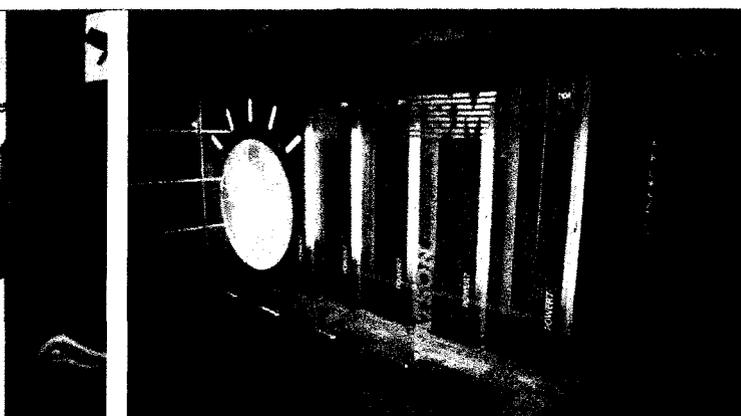
Disease или Human disease networks (HDN). Безмасштабная сеть Барабаши-Альберт. Построена на данных OMIM (2005), Online Mendelian Inheritance in Man (Законы наследования Менделя), An Online Catalog of Human Genes and Genetic Disorders (Каталог генов человека и генетических расстройств). Вершины (node, узел) графа обозначают различные заболевания. Вершины связаны ребрами (линиями) в случаях, когда имеются общие генетические нарушения для двух болезней. Источник: bfg.oxfordjournals.org.

§ 5. Суперкомпьютер и моделирование

Самый передовой рубеж ИТ – это суперкомпьютеры (High performance computing, HPC). Существуют разные суперкомпьютерные системы, они создаются под очень широкий спектр задач. Термин «суперкомпьютер» или «супервычислитель» можно считать относительным и зависящим от уровня технического прогресса. В начале XX века, в годы становления компании IBM (International Business Machines), усовершенствованные табуляционные машины называли машинами для супервычислений. На современном этапе термином HPC обозначают вычислительные машины, значительно превосходящие по своим вычислительным параметрам и скорости вычислений большинство существующих в мире компьютеров – то есть это всегда «верхушка пирамиды», малая и лучшая часть от всего объема мирового парка вычислительных машин. В основном суперкомпьютеры состоят из множества высокопроизводительных серверных компьютеров, соединенных между собой локальной высокоскоростной магистралью. Такая архитектура мультикомпьютерной системы позволяет применять распараллеливание вычислительной задачи и достигать максимально возможной производительности. Есть HPC классические, которые выполняют очень сложные и массивные вычисления. Есть HPC неклассические – это современные гигантские дата-центры и такие системы, как Google, Facebook, Twitter, Яндекс и т.д. Передовой суперкомпьютерный уровень подразумевает не только максимально возможную производительность, но и максимально возможный объем памяти (оперативной и дисковой) и самое современное программное обеспечение, а также энергосберегающие технологии.



Суперкомпьютер SKIF (РФ).
Фото: Т-Платформы.



Суперкомпьютер IBM Watson (США).
Фото: IBM.

Производительность суперкомпьютеров оценивают в флопсах (см. табл. 2). FLOPS (flops, flop/s, флопс, флоп/с) – Floating-point Operations Per Second, количество операций над числами с плавающей запятой в секунду. FLOPS служит мерой сравнения суперкомпьютеров с точки зрения их возможностей производить вычисления очень высокой точности для моделирования. Для простых компьютеров применяют показатель IPS (Instructions Per Second) или MIPS (million IPS), который отражает число инструкций, выполняемых процессором компьютера за одну секунду (мера быстродействия процессора). За годы развития компьютерных процессоров показатель MIPS вырос с 0,002 (MIPS at 2.25 MHz, один из первых американских компьютеров UNIVAC I, 1950-е гг.) до 2356230 (MIPS at 4,35 GHz, микропроцессор AMD Ryzen Threadripper 3990X, 2020 г.). Время такта процессора от начального в 2 микросекунды ускорилось до современного уровня, достигающего 0,25 наносекунды. Выражаясь образно, человек с простым калькулятором в руках должен считать более 100 тысяч лет, чтобы выполнить тот объем вычислений, который способны сделать сегодня суперкомпьютеры за 1 секунду [22]. За период эволюции компьютеров скорость вычислений выросла в десятки тысяч раз, а производительность увеличилась почти в миллиард раз. Удалось добиться такого роста производительности при относительно

медленном улучшении скоростных показателей благодаря архитектурным решениям, то есть путем создания суперкомпьютерных систем для параллельной и конвейерной обработки данных (когда одновременно выполняется много действий для решения одной задачи).

Таблица 2

Производительность суперкомпьютеров в флопсах

Русское обозначение	Английское обозначение	Степень	Годы достижения уровня производительности
килофлопс (Кбайт)	KiloFLOPS (kFLOPS)	10^3	1949 г.
мегафлопс (Мбайт)	MegaFLOPS (MFLOPS)	10^6	1964 г.
гигафлопс (Гбайт)	GigaFLOPS (GFLOPS)	10^9	1987 г.
терафлопс (Тбайт)	TeraFLOPS (TFLOPS)	10^{12}	1997 г.
петафлопс (Пбайт)	PetaFLOPS (PFLOPS)	10^{15}	2008 г.
эксафлопс (Эбайт)	ExaFLOPS (EFLOPS)	10^{18}	2019 г.
зеттафлопс (Збайт)	ZettaFLOPS (ZFLOPS)	10^{21}	2030 г.
иоттафлопс (Ибайт)	YottaFLOPS (YFLOPS)	10^{24}	будет актуальным после 2030 г.
бронтофлопс (Ббайт)	BrontoFLOPS (BFLOPS)	10^{27}	будет актуальным после 2030 г.
геофлопс (Гпбайт)	GeopFLOPS (GFLOPS)	10^{30}	будет актуальным после 2030 г.

Скорость процессоров после 2000 года:

Машинный такт (частота тактового генератора) 1 миллиард импульсов в 1 секунду = 1 ГГц (GHz)

Длительность машинного такта = 1 миллиардная часть секунды, т.е. 1 наносекунда (нс, ns)

Скорость процессоров после 2010 года:

Машинный такт (частота тактового генератора) 4 миллиарда импульсов в 1 секунду = 4 ГГц (GHz)

Длительность машинного такта = 0,25 наносекунды (нс, ns)

Принципы распараллеливания вычислительной задачи разрабатываются с самого начала кибернетической революции. Так, известно, что в начале 1950-х гг. советский математик Александр Андреевич Самарский (1919–2008 гг.) при моделировании взрыва водородной бомбы применял «метод сеток» для ускорения расчетов, которые выполнял коллектив математиков на арифмометрах – настольных механических вычислительных машинах для сложения, вычитания, умножения и деления. Кстати, первый в России арифмометр был создан в 1850-х гг. П.Л. Чебышевым. Каждый стол с несколькими арифмометрами был «узлом» общей вычислительной сетки, цифры передавались на словах от человека к человеку. Так удалось рассчитать эволюцию ядерной взрывной волны.

Более того, еще до зари компьютерной эры можно найти примеры применения распараллеливания, один из таких – в известной экономической теории Адама Смита (Adam Smith, 1723–1790 гг.) о разделении труда, изложенной в его работе «Исследование о природе и причинах богатства народов». Адам Смит – шотландский экономист и философ, он был профессором логики в Университете Глазго (University of Glasgow). Его экономическая теория до сих пор служит классической основой для современных научных школ. Адам Смит описал идею распараллеливания на примере роста производительности на булавочной фабрике более чем в 100 раз. Если разделить процесс производства булавок на 18 операций, при этом каждую операцию будет делать отдельная группа работников, то получится параллельная технологическая цепочка, благодаря которой минимум 18 рабочих смогут сделать в день 48 тысяч булавок, а 36 рабочих сделают за один рабочий день 96 тысяч булавок. То есть при разделении труда на одного рабочего за один рабочий день приходится 2667 готовых булавок, в то время как без деления задачи один рабочий самостоятельно успевал сделать всего 20 булавок.

Суперкомпьютеры, выполняющие параллельные вычисления, появились в середине прошлого века. В начале 1960-х гг. в США были созданы первые суперкомпьютеры с параллельной памятью (IBM STRETCH) по заказу Агентства национальной безопасности и Комиссии по атомной энергетике США. В этот же период в Великобритании разработали конвейерный суперкомпьютер ATLAS, также по правительственному заказу для оборонных целей. В 1960-е гг. корпорация Control Data Corporation выпустила суперкомпьютеры CDC-6600 и CDC-7600. Их производительность была на уровне мегафлопсов. В СССР в эти годы также шла работа по повышению производительности компьютеров. В начале 1960-х гг. был разработан компьютер Т340-А, имевший мегафлопсовую производительность. Т340-А использовался для противоракетной обороны. Далее в мире стали появляться более мощные суперкомпьютеры линейки Cray. Их разрабатывал американский инженер Сеймур Роджер Крэй (Seymour Roger Cray, 1925–1996 гг.), основоположник американской суперкомпьютерной индустрии. Сегодня корпорация-наследник компании Сеймура Крэя – Cray Inc. – продолжает выпускать современные суперкомпьютеры передового уровня. С суперкомпьютером Cray-2 (1985 г.) соперничал советский суперкомпьютер М-13 (1984 г.). Машина М-13 – многопроцессорная векторно-конвейерная ЭВМ (электронно-вычислительная машина), которая обрабатывала большие объемы данных в режиме реального времени. Этот суперкомпьютер, так же как и его предшественник Т340-А, использовался для противоракетной обороны. Главным конструктором М-13 был Михаил Александрович Карцев (1923–1983 гг.).

Петафлопсный рубеж первым преодолел в 2008 г. суперкомпьютер IBM Roadrunner. С 2007 г. IBM работает над проектом Watson. Этот суперкомпьютер назван в честь первого президента IBM Томаса Уотсона (Thomas John Watson, 1874–1956 гг.). IBM Watson включает в себя технологии Искусственного интеллекта – вопросно-ответная система, обработка естественного языка (Natural Language Processing). IBM Watson «просматривает» текстовые Большие данные и моментально находит нужный ответ в режиме реального времени. Этот суперкомпьютер получил мировую известность в 2011 г., когда в американской телевикторине Jeopardy («Риск», аналог программы «Своя игра») обыграл популярных в США рекорсменов игры, находя ответы на вопросы быстрее «умников». При этом IBM Watson использовал массивы текстов, хранящиеся в его памяти. В настоящее время среди сфер, в которых его применяют как облачный сервис, на первое место выходит медицина – IBM Watson Health. По данным 9-го президента IBM Джинни Рометти (Ginni Rometty), озвученным в 2017 г. на международной конференции HIMSS (The Healthcare Information and Management Systems Society), суперкомпьютер Watson в 30% сложных клинических случаев находит из больших текстовых данных решения лучше, чем таковые предлагают именитые американские доктора, соревнующиеся с этой ЭВМ.

В 2020 г. в мировом рейтинге суперкомпьютеров Top 500 (<https://www.top500.org>) первое место занял японский суперкомпьютер Fugaku, его пиковая производительность достигает 1000 petaflops (1 exaflops). В третьем десятилетии XXI века начался период гонки на уровне «эксафлопсных» суперкомпьютеров. На второй и третьей позиции рейтинга Top 500 в 2020 г. находятся суперкомпьютеры IBM Summit и Sierra. В России решением научных задач в области создания суперкомпьютеров занимается Институт программных систем (ИПС) им. А.К. Айламазяна РАН (город Переславль-Залесский). Ежегодно в ИПС РАН проходит Национальный Суперкомпьютерный Форум (НСКФ), на котором освещаются все вопросы, касающиеся суперкомпьютерной отрасли [23]. Директор ИПС РАН и председатель Национального Суперкомпьютерного Форума, член-корреспондент РАН Сергей Михайлович Абрамов руководит научной программой по созданию серии суперкомпьютеров СКИФ. Самые мощные суперкомпьютеры в России – «Ломоносов» и «Ломоносов-2», установленные в Московском государственном университете имени М.В. Ломоносова (производитель «Т-Платформы»), также содержат разработки ИПС РАН. Оба суперкомпьютера входят в лист Top 500. Другие российские суперкомпьютеры, произведенные компанией «Т-Платформы»: суперкомпьютер Росгидромета, МИИТ Т-4700 (расположен в Российском университете

транспорта, МИИТ), СКИФ МГУ «Чебышев», СКИФ Урал (Южно-Уральский государственный университет), СКИФ Cyberia (Томский государственный университет) и др. (представлены на сайте <https://t-platforms.ru/projects>).

В мировом рейтинге суперкомпьютеров Top 500 учитывается то, какая машина быстрее решает систему линейных уравнений в рамках теста LINPACK Benchmark. Показатель «пиковая производительность» (произведение тактовой частоты процессора на число операций, выполняемых за такт) означает максимальную теоретическую производительность, которую может достичь компьютер. Однако в рейтинге совсем не учитываются цена и энергозатраты для лидеров по скорости вычислений. В ИПС РАН рассматривают работу суперкомпьютеров не только с точки зрения скорости, но и в аспекте энергозатрат. Электричество требуется для суперкомпьютеров в двойном объеме – на вычислительную работу самой машины и на работу систем охлаждения. Заместитель директора ИПС РАН, профессор Сергей Анатольевич Амелькин, выступая на НСКФ в 2016 г., отметил, что, сдав тест на скорость и производительность, машина может просто перегреться и выйти из строя (так называемое «тепловое загрязнение» суперкомпьютера). А прекращение работы перегретого суперкомпьютера в мире ИТ, Больших данных и Интернета вещей может обернуться настоящей катастрофой с человеческими жертвами (например, если суперкомпьютер обеспечивает работу воздушных или железнодорожных перевозок, или регулирует автомобильное движение и т.д.). Существуют разные системы охлаждения суперкомпьютеров: воздушные, гибридные (используется воздух и жидкость) и только жидкостные (погружение деталей суперкомпьютера в специальную жидкость). Уникальный способ охлаждения путем погружения в жидкость применяется в отечественных инновационных суперкомпьютерных вычислительных комплексах ИММЕРС.



Сергей Михайлович Абрамов
Фото: НСКФ.

Суперкомпьютерный блок с системой ИММЕРС.
Фото: Департамент инвестиционной политики
правительства Ярославской области, Наталья
Гребнева, <http://smartnews.ru/regions/yar/14110.html>.

Суперкомпьютеры используют для моделирования разных процессов. Без суперкомпьютерного моделирования невозможны исследования климата и прогноз погоды. Например, созданы модели глобального изменения климата Coupled Model Intercomparison Project, к 2020 г. завершилась уже 6-я фаза этого международного проекта. Суперкомпьютеры нужны для моделирования и создания новых материалов и лекарств, для изучения физических и биологических процессов и многого другого (сейсморазведка, добыча полезных ископаемых, проектирование разных объектов и машин, астрофизические исследования и изучение космоса). Исследования, основанные на Больших данных в медицине, также проводятся с использованием суперкомпьютеров.



Рабочее место исследователя в области климатического моделирования. Met Office, Великобритания. Фото: Met Office. <https://www.carbonbrief.org/qa-how-do-climate-models-work>.



Суперкомпьютерная система Cray / T-Platforms в Гидрометцентре России. Фото: <https://meteoinfo.ru/novosti/15610-rosgidromet-predstavil-vysokoproizvoditelnyj-vychislitelnyj-klaster>.

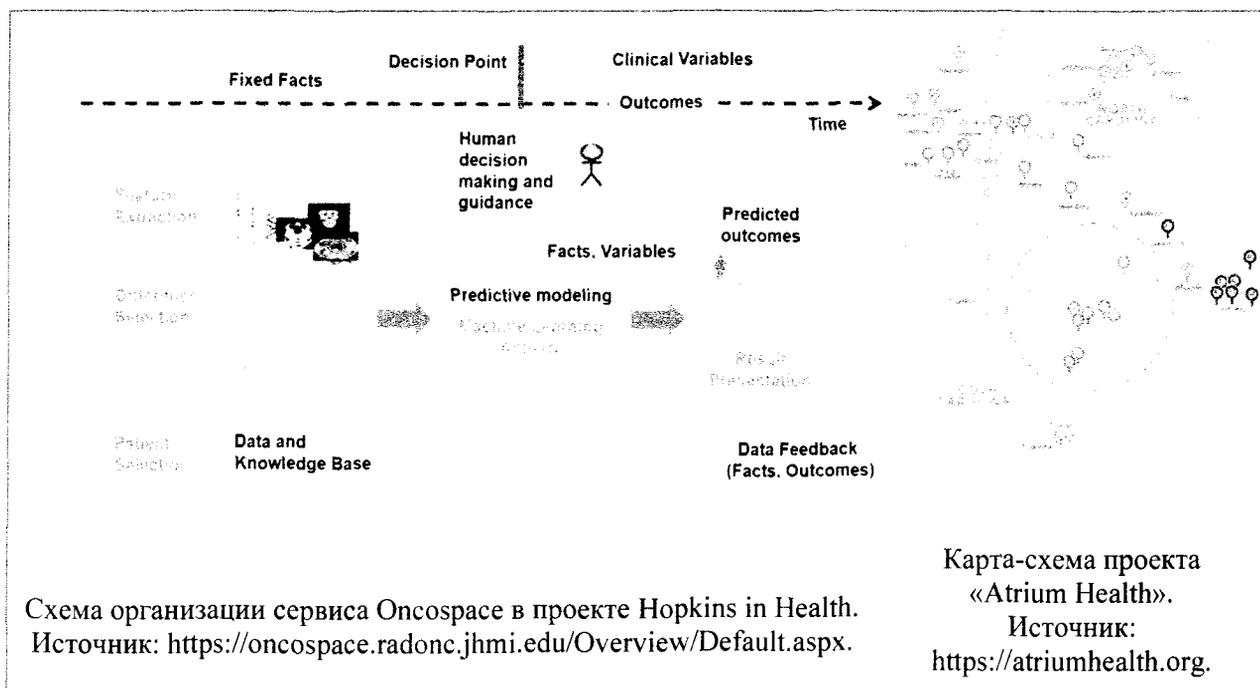
Интересный суперкомпьютерный проект по Большим данным и персонализированной медицине «Johns Hopkins Individualized Health Initiative», кратко «Hopkins in Health», организован в Университете Джонса Хопкинса (Johns Hopkins University, Балтимор, США). Госпиталь университета (Johns Hopkins Hospital) и еще 5 больниц и 40 амбулаторных центров вовлечены в него с 2012 г. В проекте «Hopkins in Health» через облачный сервис Microsoft's Azure Cloud Platform собирают всю информацию из электронных медицинских баз данных, включая медицинские диагностические изображения (медицинская визуализация), данные секвенирования генома и записи с мониторирующих показатели приборов. Цель проекта – поиск новых персонализированных моделей ведения пациентов на основе прогноза (построение прогностических моделей). Одна из таких прогностических (или предиктивных) моделей называется Oncospace. Собираемые данные по ведению онкологических пациентов (изображения компьютерной томографии, записи врачей из электронных карт, опрос пациентов) подвергаются анализу алгоритмами машинного обучения (Machine Learning). Результат машинного анализа данных врач получает на свой смартфон в виде подсказок, как спланировать лучевую терапию больному с точным подсчетом дозы радиации для всех тканей и органов в зоне облучения, и с прогнозом последствий лучевой терапии для конкретного больного.

Еще один интересный проект, в котором используются суперкомпьютерные мощности, – это американская региональная медицинская логистическая система «Atrium Health» (ранее система называлась Carolinas Health Care System). Этот цифровой сервис объединяет два штата North and South Carolina. Дата-центр, созданный на основе технологий IBM, собирает Большие данные. Алгоритмы уровня Искусственного интеллекта координируют потоки информации от 47 госпиталей и множества разных клиник. Обработывается более 20 миллионов транзакций данных в сутки. Это медицинские данные пациентов, записи врачей по результатам обследований, данные секвенирования генома, санитарно-эпидемиологические данные по районам о загрязнении окружающей среды, социологические данные (по демографии и т.п.), данные из соцсетей. Сервис «Atrium Health» перераспределяет пациентов в соответствии с ситуацией, позволяя быстро получить нужную медицинскую помощь. Создаваемые математические предиктивные модели помогают сокращать случаи повторной госпитализации после выписки из стационара.

Моделирование в медицине, биохимии, биологии требует перевода логических подходов в математический язык и правила программирования. Эта область изучения близка к области создания Искусственного интеллекта. Ведь для того чтобы Искусственный интеллект мог

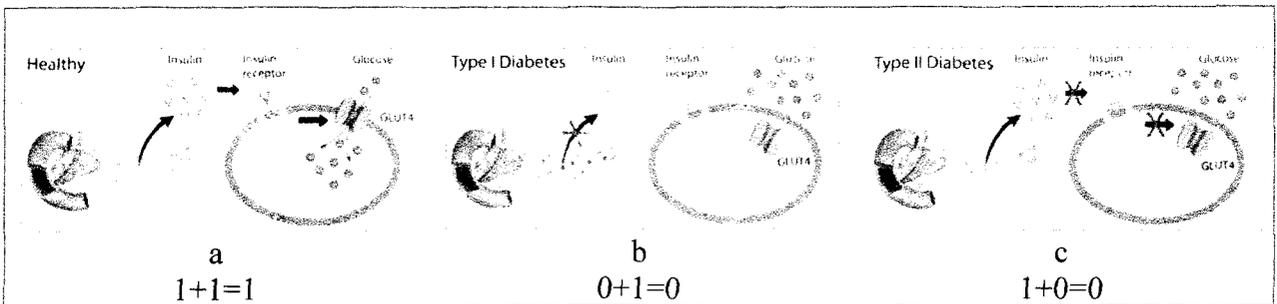
стать настоящим помощником врача, простого накопления базы знаний недостаточно. Нужны логические схемы и алгоритмы, применяющие накопленные знания. Собственно, это же относится и к человеку – недостаточно выучить и помнить какой-то объем знаний, для работы врачом нужно умение правильно логически мыслить, что принято называть «клиническим мышлением врача». При моделировании разных процессов поиск соответствующей математической логики открывает новые горизонты для понимания. Например, для понимания того, как развивается болезнь или как работает та или иная физиологическая система организма. Уже нашли широкое применение в изучении различных биохимических процессов бинарные логические операции (Logic gates). Основным интересом ученых приковано к нескольким из них.

Из бинарных логических операций можно складывать простые схемы (patterns) и сложные цепи из десятков операций (complex patterns), для вычисления которых требуется суперкомпьютер. Используются символы, соответствующие стандартам American National Standards Institute (ANSI) и Institute of Electrical and Electronic Engineers (IEEE). Для каждой логической операции разработана таблица истинности (Truth Table), описывающая логическую функцию на основе Булевой алгебры (Boolean algebra) с применением знаков «0» и «1». **Математическая логика в виде бинарных логических операций может применяться для изучения этиологии и патогенеза болезней.** Если отталкиваться не от классификации заболеваний (основная из классификаций – Международная классификация болезней), а от физиологической нормы какого-то параметра и его отклонений, то можно подобрать бинарную логическую операцию, описывающую происходящие патологические изменения. Системный подход, с ориентацией на поддержание физиологической нормы какого-то выбранного параметра (уровень глюкозы в крови, уровень артериального давления, насыщение крови кислородом), был заложен К.В. Судаковым и отражен в разработанных им разнообразных функциональных системах организма человека [21]. Выбранный параметр представляется как суммарный результат регуляции со стороны всех вовлеченных физиологических (или патологических) процессов. А если подбирать подходящие бинарные логические операции к клиническим ситуациям, в которых меняется какой-то определенный физиологический показатель, то можно описать таблицу истинности. Например, разные по природе заболевания (сахарный диабет 1 типа и сахарный диабет 2 типа) могут быть вписаны в единую бинарную операцию конъюнкции, взяв за ориентир уровень глюкозы и инсулина в крови.



 <p>Инвертор</p>	<table border="1"> <tr><th>A</th><th>\bar{A}</th></tr> <tr><td>0</td><td>1</td></tr> <tr><td>1</td><td>0</td></tr> </table>	A	\bar{A}	0	1	1	0	 <p>Конъюнкция</p>	<table border="1"> <tr><th>A</th><th>B</th><th>AND</th></tr> <tr><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>1</td><td>0</td></tr> <tr><td>1</td><td>0</td><td>0</td></tr> <tr><td>1</td><td>1</td><td>1</td></tr> </table>	A	B	AND	0	0	0	0	1	0	1	0	0	1	1	1									
A	\bar{A}																																
0	1																																
1	0																																
A	B	AND																															
0	0	0																															
0	1	0																															
1	0	0																															
1	1	1																															
 <p>Дизъюнкция</p>	<table border="1"> <tr><th>A</th><th>B</th><th>OR</th></tr> <tr><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>1</td><td>1</td></tr> <tr><td>1</td><td>0</td><td>1</td></tr> <tr><td>1</td><td>1</td><td>1</td></tr> </table>	A	B	OR	0	0	0	0	1	1	1	0	1	1	1	1	 <p>Инверсия конъюнкции</p>	<table border="1"> <tr><th>A</th><th>B</th><th>NAND</th></tr> <tr><td>0</td><td>0</td><td>1</td></tr> <tr><td>0</td><td>1</td><td>1</td></tr> <tr><td>1</td><td>0</td><td>1</td></tr> <tr><td>1</td><td>1</td><td>0</td></tr> </table>	A	B	NAND	0	0	1	0	1	1	1	0	1	1	1	0
A	B	OR																															
0	0	0																															
0	1	1																															
1	0	1																															
1	1	1																															
A	B	NAND																															
0	0	1																															
0	1	1																															
1	0	1																															
1	1	0																															
 <p>Инверсия дизъюнкции</p>	<table border="1"> <tr><th>A</th><th>B</th><th>NOR</th></tr> <tr><td>0</td><td>0</td><td>1</td></tr> <tr><td>0</td><td>1</td><td>0</td></tr> <tr><td>1</td><td>0</td><td>0</td></tr> <tr><td>1</td><td>1</td><td>0</td></tr> </table>	A	B	NOR	0	0	1	0	1	0	1	0	0	1	1	0	 <p>Равнозначность</p>	<table border="1"> <tr><th>A</th><th>B</th><th>XNOR</th></tr> <tr><td>0</td><td>0</td><td>1</td></tr> <tr><td>0</td><td>1</td><td>0</td></tr> <tr><td>1</td><td>0</td><td>0</td></tr> <tr><td>1</td><td>1</td><td>1</td></tr> </table>	A	B	XNOR	0	0	1	0	1	0	1	0	0	1	1	1
A	B	NOR																															
0	0	1																															
0	1	0																															
1	0	0																															
1	1	0																															
A	B	XNOR																															
0	0	1																															
0	1	0																															
1	0	0																															
1	1	1																															
 <p>Инверсия равнозначности</p>	<table border="1"> <tr><th>A</th><th>B</th><th>XOR</th></tr> <tr><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>1</td><td>1</td></tr> <tr><td>1</td><td>0</td><td>1</td></tr> <tr><td>1</td><td>1</td><td>0</td></tr> </table>	A	B	XOR	0	0	0	0	1	1	1	0	1	1	1	0																	
A	B	XOR																															
0	0	0																															
0	1	1																															
1	0	1																															
1	1	0																															

Часто применяемые бинарные логические операции в экспериментах с биохимическими реакциями. Другие, не показанные на рисунке, операции называются: прямая импликация от А к В, обратная импликация от В к А, декремент – инверсия импликации от А к В, инкремент – инверсия импликации от В к А.



Схематичное изображение прохождения глюкозы в клетку в норме в присутствии инсулина (лиганд) как ключа при контакте с рецептором (а), и невозможность глюкозы попасть в клетку при сахарном диабете 1 типа в условиях отсутствия инсулина (б), или при сахарном диабете 2 типа в условиях инсулинорезистентности тканей с «поломкой» рецепторов клетки (с). Подписи к рисункам – бинарная операция конъюнкции AND. Для иллюстрации использован художественный ресурс Shutterstock.

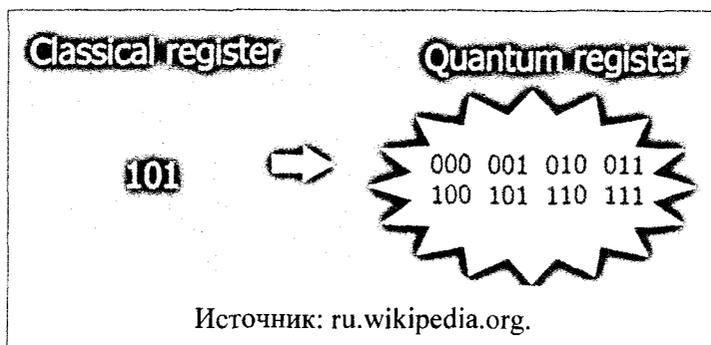
	Инсулин	Рецептор клетки	AND	Состояние
	0	0	0	Сахарный диабет 2 типа
	0	1	0	Сахарный диабет 1 типа
	1	0	0	Сахарный диабет 2 типа
	1	1	1	Нормогликемия

Применение бинарной операции конъюнкции AND для описания трех состояний – нормогликемии, сахарного диабета 1 типа и сахарного диабета 2 типа.

§ 6. Квантовый компьютер

Сегодня в мире ученые работают над созданием квантового компьютера, что стало возможным благодаря теоретическим открытиям в области квантовой физики. История квантовой физики начинается с работ немецкого ученого Макса Планка (Max Planck, 1858–1947 гг.). В 1918 г. Макс Планк получил Нобелевскую премию за открытие квантов энергии и развитие теории квантов. Идею о возможности осуществлять квантовые вычисления и создать квантовый компьютер первыми высказали в 1980-е гг. советский математик Юрий Иванович Манин (с 1993 г. русский ученый работает в Математическом институте Макса Планка в Бонне, Max Planck Institute for Mathematics) и американский физик, нобелевский лауреат Ричард Фейнман (Richard Feynman, 1918–1988 гг.), который разработал первую теоретическую модель квантового компьютера.

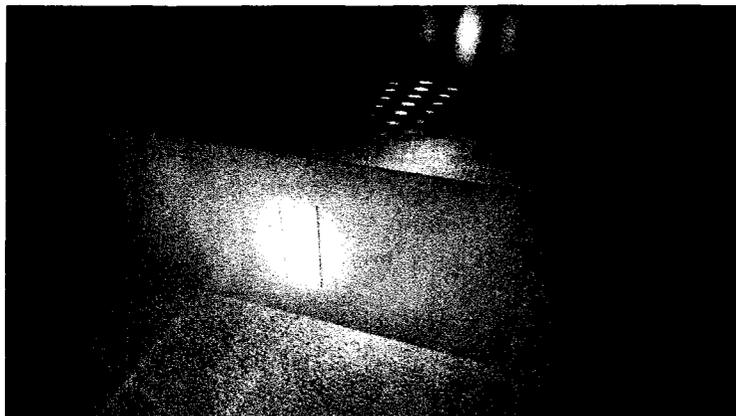
В обычном компьютере бит имеет значение либо 0, либо 1, а q-бит (quantum bit) или *кубит* (квантовый разряд, наименьший элемент для хранения информации в квантовом компьютере) одновременно имеет значение 0 и 1. Это одновременное нахождение кванта в двух состояниях называют *квантовой суперпозицией* (superposition). Квантом (от лат. quantum – «сколько») обозначают неделимую часть величины в какой-то системе, которая принимает определенные значения (квантуется). Например, фотон – квант электромагнитного поля; гравитон – квант гравитационного поля; фонон – квант колебательного движения кристалла; хронон – квант времени; бозон Хиггса – квант поля Хиггса; глюон – квант векторного (глюонного) поля сильного взаимодействия.



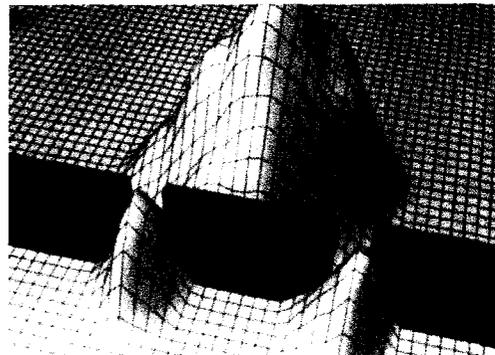
Объяснить квантовую суперпозицию можно при помощи опыта Юнга. Это классический и известный физикам эксперимент на двух щелях, продемонстрированный в 1803 г. английским физиком и врачом Томасом Юнгом (Thomas Young, 1773–1829 гг.). Опыт Юнга показывает интерференцию и дифракцию света и доказывает волновую теорию света. Фотоны, проходя через две щели, распространяются как волна и имеют случайное распределение. С равной долей вероятности фотон может как волна пройти и через правую, и через левую щель экрана (ширина щелей равна длине волны излучаемого света). Образующиеся две вторичные волны когерентного света видны на проекционном экране в виде интерференционной картины (чередование максимумов и минимумов яркости света). Неопределенность – через какую щель прошел отдельный фотон – остается до тех пор, пока специальные детекторы не зафиксируют факт прохождения этого фотона через одну из двух щелей. Получается, что до экрана с щелями можно наблюдать квантовую суперпозицию фотона. После прохождения щели с детектором фотон будет подчиняться принципу случайности в условиях интерференции – с большей вероятностью фотон окажется в зоне максимума яркости (зона сложения амплитуд двух вторичных волн), но точно предсказать место падения фотона на экран невозможно. Это новое понятие в физике – местонахождение объекта остается неопределенным, пока его не зафиксируют (до фиксирующего детектора точного местонахождения объекта не существует).

В квантовой физике также есть такое явление, как *квантовая запутанность* (entanglement) – когда два кванта (квантовых объекта) взаимосвязаны, но при этом они могут находиться друг от друга на большом расстоянии (предел этого расстояния не определен, то есть предела нет).

Фотоны в запутанном состоянии (одновременно сгенерированные) ведут себя по-иному: их спины меняются одновременно и всегда противоположны друг другу. Альберт Эйнштейн называл такую взаимосвязь «жуткое действие на расстоянии» (spooky action at a distance). **Спин** (от англ. spin, вращение) – это упрощенно вектор момента импульса фотона (вращательного движения фотона) или поляризация фотона. Определив направление спина одного фотона в запутанной паре фотонов детектором, можно утверждать, что направление спина второго фотона будет противоположно первому, даже если второй фотон находится на недостижимом для измерения расстоянии. Как и в опыте Юнга с детекторами, спин можно определить только при замере и это будет случайный результат, но обязательно противоположный направлению спина второго запутанного фотона (положительный и отрицательный, и наоборот). Квантовая механика заставила ученых воспринимать реальность как совокупность вероятностей. В квантовых компьютерах кубиты могут быть запутанными.



Опыт Юнга. Две щели на ближнем экране и интерференционная картина (чередование максимумов и минимумов яркости света) на дальнем экране. Иллюстрация BBC.



Образование вторичных волн при прохождении фотонов через щели в опыте Юнга. Кадр из фильма Джеральда Каргла (Gerald Kargl) «Das geheimnisvolle Reich der Quanten», 2006.



Квантовый компьютер IBM «Q».
Фото: <https://www.research.ibm.com/ibm-q>.



Помещение, где работает квантовый компьютер IBM «Q».
Фото: <https://www.research.ibm.com/ibm-q>.

Если обычные компьютеры делают вычисления и хранят информацию в двоичной системе 0 и 1, то квантовые компьютеры должны использовать в работе для вычислений и хранения информации два физических феномена – квантовую суперпозицию (superposition) и квантовую запутанность (entanglement). Несмотря на сложность и необъяснимость многих квантовых явлений, квантовые компьютеры уже существуют. IBM успешно протестировала в 2001 г. 7-кубитный квантовый компьютер «Q». Российские ученые объявили о создании

2-кубитного квантового компьютера в 2005 г. В Калифорнии в 2012 г. создали 2-кубитный квантовый компьютер на кристалле алмаза. В 2017 г. группа профессора Гарвардского университета (Harvard University) Михаила Дмитриевича Лукина сделала 51-кубитный квантовый симулятор, в то же время группа Кристофера Монро (Christopher Monroe) из Университета Мэриленд (University of Maryland, College Park) создала 53-кубитный квантовый симулятор. В 2017 г. в IBM сделали 50-кубитный процессор, а в Intel к началу 2018 г. – 49-кубитный компьютер «Tangle Lake». В 2018 г. в Google построили 72-кубитный квантовый процессор «Bristlecone», в 2019 г. компания Google представила квантовый процессор «Sycamore». 54-кубитный «Sycamore» выполнил за 200 секунд задание, которое один из лучших суперкомпьютеров мира «Summit» выполняет за 2,5 суток [24]. Канадская компания D-Wave Systems в 2007 г. разработала компьютер «D-Wave», в 2015 г. компания заявила о достижении мощности процессора до 1000 кубитов. Квантовые компьютеры «D-Wave» приобрели корпорация Lockheed Martin, DARPA, Google, космическое агентство NASA и разрабатывающая ядерное оружие Лос-Аламосская национальная лаборатория США (Los Alamos National Laboratory).



Квантовые компьютеры необходимы для создания и функционирования Искусственного интеллекта и обработки Больших данных в режиме реального времени. С помощью запутанных кубитов можно будет производить параллельные вычисления в несравнимо большем объеме и с несравнимо большей скоростью, чем это могут делать самые продвинутые современные суперкомпьютеры. По оценке экспертов квантовые компьютеры уже соответствуют верхним рейтинговым строчкам суперкомпьютеров в Top 500. Для оценки квантовой производительности придуман свой тест qFlex (random quantum circuits, имитация случайных квантовых схем), который позволяет сравнивать обычные суперкомпьютеры и квантовые компьютеры [25, 26]. Далее, по мере развития квантовой отрасли, обычные суперкомпьютеры навсегда сдадут позиции, что называют новым термином **«квантовое превосходство»** (quantum supremacy). Лучшие суперкомпьютеры мира сегодня требуют энергообеспечения, как целый город среднего размера, на уровне нескольких мегаватт, в то время как аналогичные по мощности квантовые компьютеры потребляют энергию на уровне киловатт, то есть в тысячу раз меньше. Эксперты Gartner считают, что 49 кубитов – это порог, начиная с которого вычисления могут быть сделаны только квантовыми компьютерами [27]. Тем не менее существующие сегодня квантовые компьютеры пока не обходят суперкомпьютерный лист Top 500, это показало тестирование 72-кубитного квантового процессора «Bristlecone» от Google. Однако это временная ситуация, так как

суперкомпьютеры имеют два непреодолимых порога – огромное энергопотребление и закон Мура.

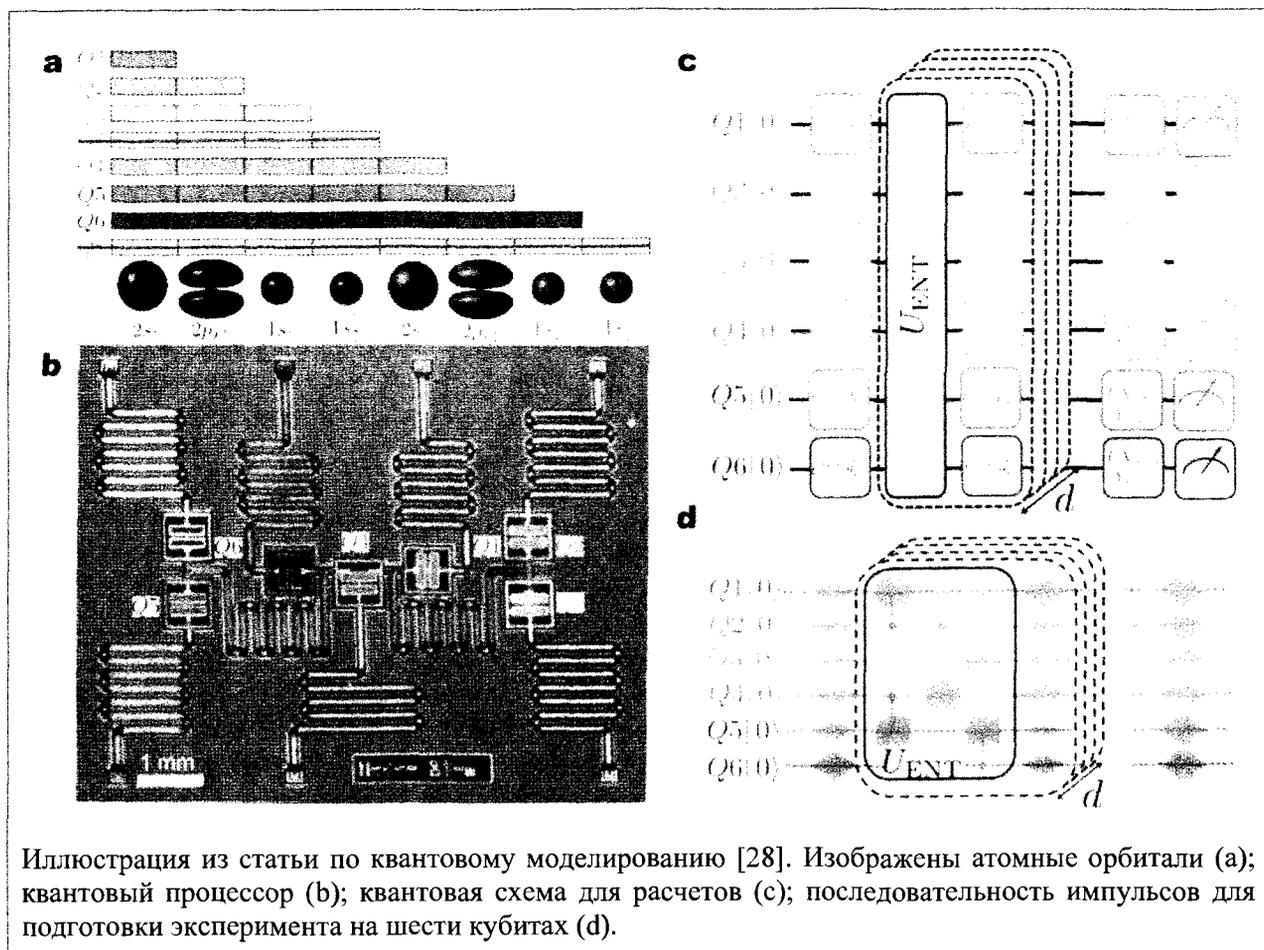


Иллюстрация из статьи по квантовому моделированию [28]. Изображены атомные орбитали (a); квантовый процессор (b); квантовая схема для расчетов (c); последовательность импульсов для подготовки эксперимента на шести кубитах (d).

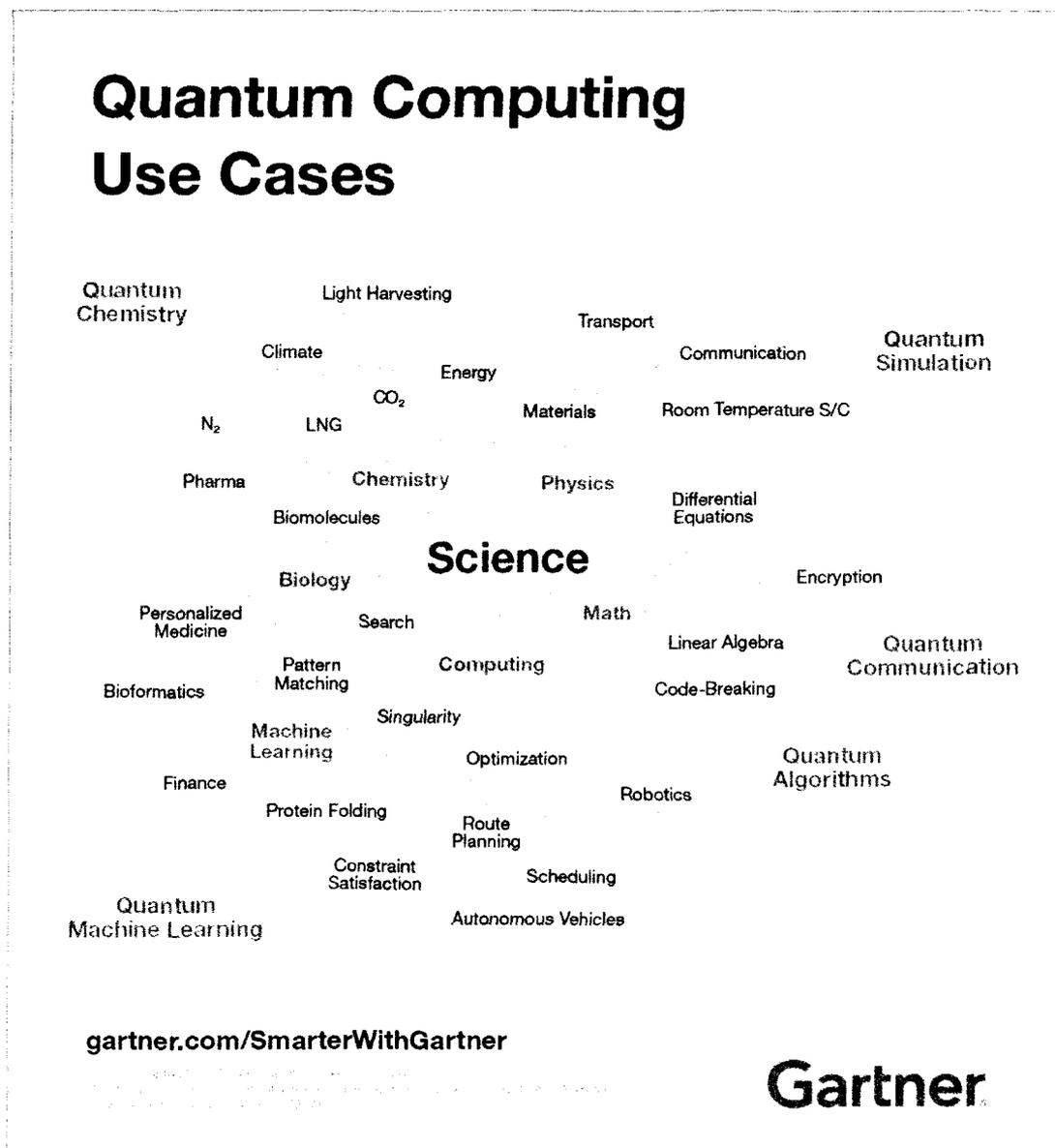
На сегодняшний день существующие квантовые компьютеры пока не заменяют полностью классические суперкомпьютеры, они могут выполнять лишь узкий круг задач. Развитию квантовых вычислений будет способствовать использование мировой профессиональной аудиторией этого дорогого оборудования удаленно, через Интернет – Quantum computing as-a-Service (QCaaS). Так, IBM уже предоставила открытый доступ к своему 16-кубитному квантовому компьютеру в качестве бесплатного облачного сервиса Quantum Information Science Kit, Qiskit (<https://qiskit.org/>). Еще одна американская компания по производству квантовых компьютеров Rigetti также создала облачный сервис Quantum Cloud Services (QCS) platform (<http://www.rigetti.com/>), который позволяет осуществлять квантовое программирование удаленно на 8-кубитном квантовом компьютере.

На современном этапе сети IoT могут генерировать в секунду объем данных, исчисляемых в петабайтах [27]. Миллионы датчиков необходимы для сетей IoT для мониторинга погоды, обеспечения производственного контроля на большом заводе, управления инфраструктурой города и т.п. Сети IoT внедряются в крупные больницы, транспортные маршруты, да и в целом, в города и дома («умный город», «умный дом»). Такой растущий объем данных и скорость их накопления уже требуют применения квантовых технологий.

Квантовые вычисления, как ожидается, могут обеспечить более быструю и более качественную работу нейросетей (Machine Learning / Deep Learning). Так, было заявлено о создании квантовой нейросети Quantum Boltzmann Machines [29] по аналогии с Машиной Больцмана (Boltzmann machine), являющейся одним из видов рекуррентной нейронной сети (Recurrent Neural Network, RNN, с направленной последовательностью связей). Машина Больцмана использует для обучения алгоритм имитации отжига (Simulated annealing) и решает сложные комбинаторные задачи. Основанная на квантовых вычислениях Quantum Boltzmann

Machine доказала свои преимущества перед обычной нейросетью, а также математики выяснили, что квантовая нейросеть имеет большой неизученный потенциал применения, так как она не просто лучше обычной, а способна решать иные, недоступные для обычной нейросети, задачи [29].

За секунды квантовый компьютер смог бы обработать геном человека (секвенирование ДНК), в котором определено около 25 тысяч генов и около 3,1 миллиардов пар оснований: аденин (A) – тимин (T), гуанин (G) – цитозин (C). Или квантовый компьютер был бы востребован для моментального анализа данных сканирования головного мозга методом магнитно-резонансной томографии или компьютерной томографии, о работах в этом направлении докладывают эксперты Microsoft [30]. Особые надежды возлагаются на новый уровень Искусственного интеллекта, который будет достигнут благодаря квантовым вычислениям. Самой главной областью, в которой будут активно использоваться квантовые компьютеры, является криптография. Алгоритмы шифрования информации станут намного сложнее, что обеспечит надежную безопасность в цифровом мире. И наоборот, квантовый компьютер сможет быстро расшифровать любой секретный код.



Прогнозируемое применение квантовых компьютеров по данным компании Gartner. Источник: <https://www.gartner.com/SmarterWithGartner> [27].

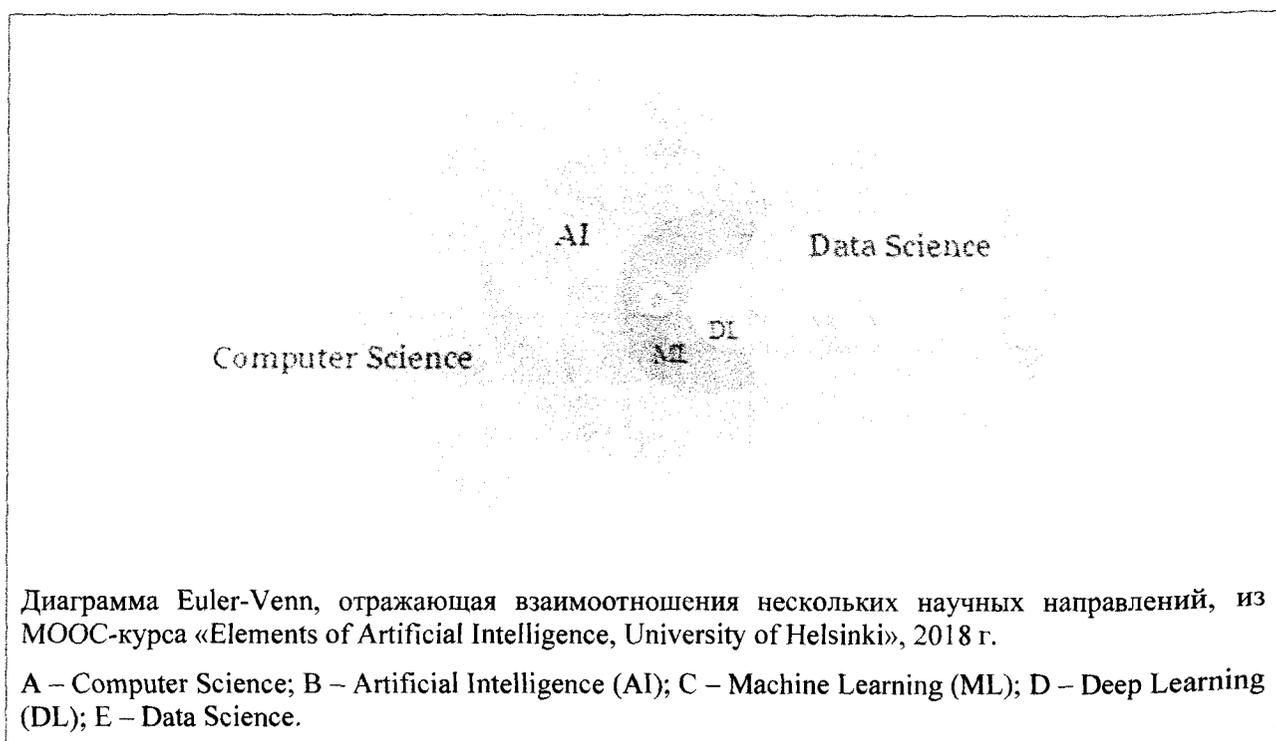
Квантовый компьютер IBM «Q» (7-кубитный чип) уже используется для моделирования молекул и химических реакций. Например, удалось рассчитать модель конфигурации трехатомной молекулы гидрида бериллия BeH_2 [28, 31]. Для построения стабильной молекулы необходимо смоделировать, как электроны и ядра атомов (которые состоят из протонов и нейтронов) будут взаимодействовать друг с другом, включая квантовые эффекты. Интерес к гидриду бериллия обусловлен тем, что он используется как ракетное топливо.

Молекулярное моделирование становится сложнее с увеличением числа атомов в молекуле, поэтому без квантового компьютера не обойтись ни в молекулярном моделировании, ни в создании новых лекарств и проверке всех взаимодействий лекарств и их эффектов. Однако пока остается неразрешенной одна серьезная проблема – как сохранить стабильным на время, достаточное для вычислений, состояние кубита. Удалось сохранить кубит стабильным только на 90 микросекунд. Увеличение количества кубитов лишь усложняет эту проблему.

В 2017 г. исследовательская и консалтинговая компания Gartner опубликовала свое видение востребованности квантовых компьютеров. Прежде всего, квантовые вычисления имеют высочайшую научную ценность. Определены пять кластеров применения квантовых вычислений – квантовая химия, квантовое моделирование (симуляция динамических процессов), квантовое машинное обучение, квантовые алгоритмы и квантовые коммуникации. Следует обратить внимание, что среди самых передовых областей, где нужны квантовые вычисления, таких как робототехника и криптография, эксперты Gartner упомянули фармацевтическую индустрию и персонализированную медицину.

§ 7. Определение науки о данных и Кривая Гартнера (Hurescycle)

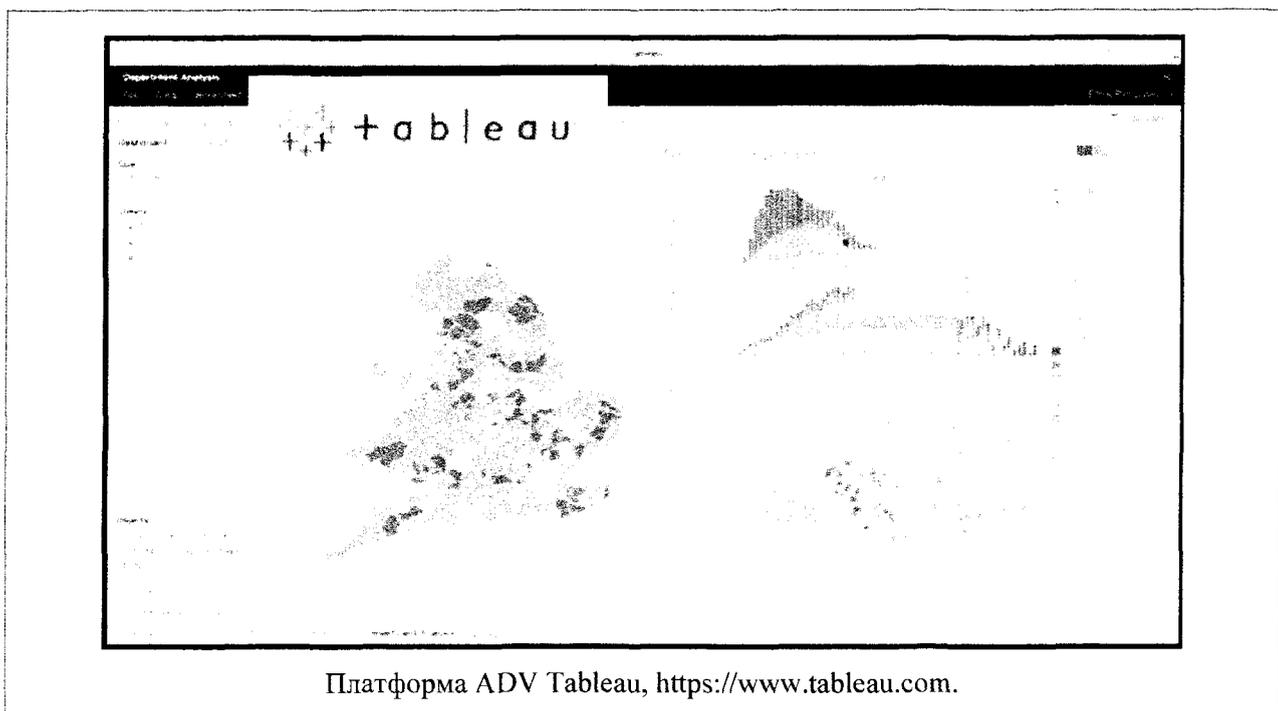
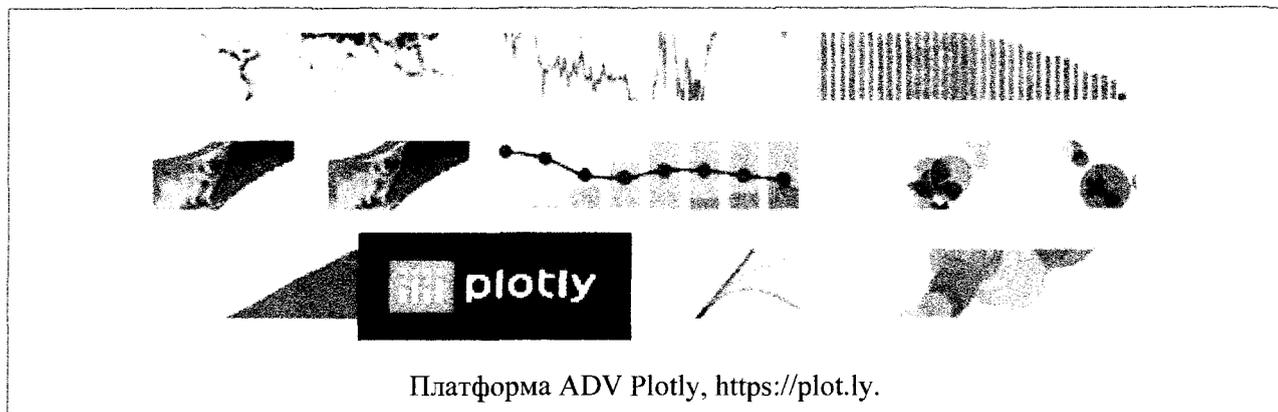
Data Science, или Наука о данных – это наука, которая объединяет все, что связано с ИТ, с одной стороны, и все, что касается сбора, обработки и анализа данных, с другой стороны. Термин «Data Science» относят к категории «зонтичных». От области компьютерных наук «Computer Science» наука о данных отличается тем, что фокусируется на данных, а не на программировании. Хотя эти две области – Computer Science и Data Science – пересекаются в методах и инструментах в вопросах обработки данных, а также всего, что относится к Искусственному интеллекту и кибернетике, Data Science обязательно касается вопросов управления процессами и принятия стратегических решений на основе данных (Data Governance). То есть наука о данных пересекается с такими направлениями, как менеджмент, политические, социальные и экономические науки (особенно цифровая экономика), международные отношения и глобалистика, любая область гуманитарных наук. Специалист, работающий в области науки о данных, называется *Data Scientist (научный специалист по данным)*.



Большие данные трансформируют медицину, которая становится областью, основанной на данных, – Data-Based Medicine. Этот термин отличается от более известного термина «Data driven» (например, его применяют к экономике – Data driven economy), так как акцент делается не на новых кибернетических технологиях и их влиянии, а на новых медицинских подходах, основанных на постоянно накапливающимися Большими данными. Data Science охватывает все сферы жизни социума. Уже сегодня без инструментов Data Science невозможно развитие ни одной компании или государственной отрасли.

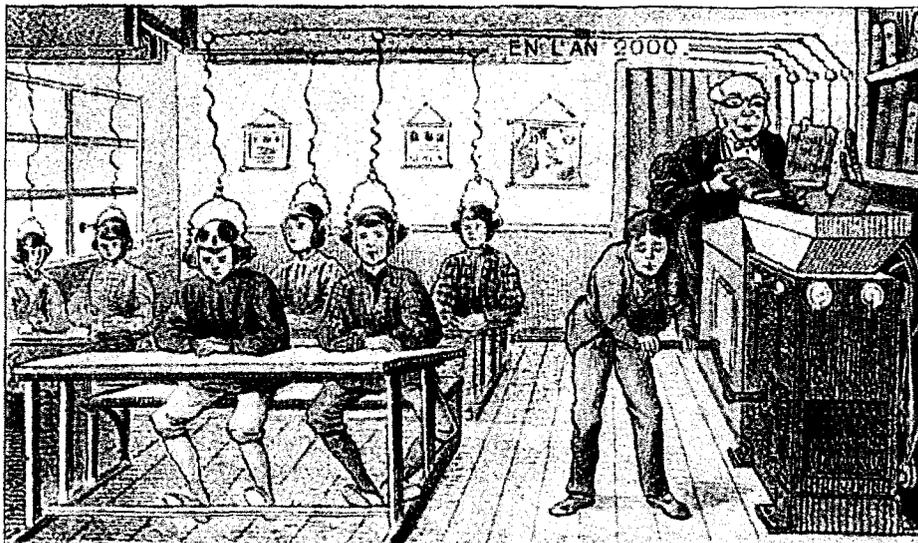
Важными инструментами Data Science являются платформы визуализации данных. Эти платформы вышли из классической статистической аналитики Business Intelligence. Сегодня к ним добавились потоки Больших данных и алгоритмы Искусственного интеллекта, включая обработку естественного языка (Natural Language Processing, NLP) и машинное обучение (Machine Learning), что позволило называть такую аналитику расширенной или дополненной (Augmented Analytics). Платформы визуализации данных функционируют на постоянной основе в режиме реального времени по принципу встроенной аналитики в процесс деятельности компании или организации (Embedded Analytics). В целом такие аналитические

платформы обозначают термином «*продвинутая визуализация данных*» (*Advanced Data Visualization, ADV*). Технологической основой для таких платформ служат сложные и многоуровневые ИТ-архитектуры типа «Data Lake». Выстраивание потоков данных в нужном порядке, а также в целом организацию сбора данных, выбор методов и инструментов аналитики и формирование единой ИТ-архитектуры называют процессом управления данными (Data Governance или Data Management). Пользователь может иметь доступ к платформам ADV с разных устройств (компьютер, смартфон, смарт-часы и т.п.). Такие платформы могут быть представлены на больших настенных экранах или интерактивных досках и щитах в общественных местах и в офисах, а также они могут быть в форматах дополненной или виртуальной реальности (Augmented Reality, Virtual Reality).



Наука о данных интенсивно развивается, все больше данных оцифровывается, все больше методов аналитики появляется. Становятся более совершенными такие инструменты Data Science, как алгоритмы Искусственного интеллекта, искусственные нейронные сети, машинное обучение. Насколько трудно предсказать облик науки о данных в будущем, можно понять на примере следующей иллюстрации. Французский художник-футурист Жан-Марк Кот (Jean-Marc Côté) с соратниками в начале XX века размышлял о том, какими будут технологии в начале XXI века. Он отразил свои предположения в серии рисунков для почтовых открыток. Сама суть переработки информации компьютерными программами была художником угадана. На рисунке «At school» (В школе) очень примитивно изображен процесс

работы по аналогии с «поисковым браузером», который имеет доступ к информации из многих книг и доставляет ее по запросу каждому персонально.



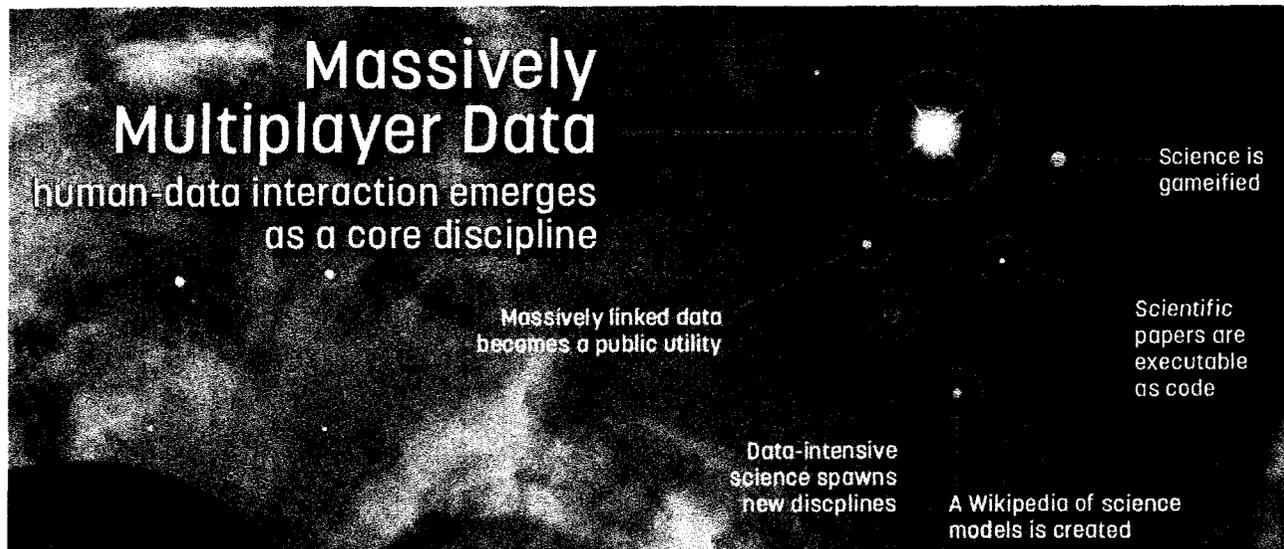
Открытка, сделанная в 1910 г. Название: «At school». Из коллекции National Library of France, En L'An 2000 (In the Year 2000, or France in the 21st Century), Villemard. Опубликовано в 1986 г. Айзеком Азимовым (Isaac Asimov, 1920–1992 гг.) в книге «Future days: A Nineteenth Century Vision of the Year 2000».

Не вызывает сомнений, что на художника оказала влияние изобретенная американцем Германом Холлеритом (Herman Hollerith, 1860–1929 гг.) в конце XIX века счетная электромеханическая машина с перфокартами (табулятор). Созданная в 1896 г. Германом Холлеритом компания «Tabulating Machine Company» в 1924 г. была переименована в «International Business Machines» или IBM. В 1941 г. IBM собрала первый компьютер – «Mark I» (Automatic Sequence Controlled Calculator, автоматический вычислитель, управляемый последовательностями). Можно также отметить, что «образ будущего», нарисованный художниками век назад, преимущественно основан на проводной технологии, которая в те времена была на пике популярности. Во второй половине XIX века в Европе и США шаг за шагом создавали проводную связь. В 1858 г. был достигнут большой успех – установлена прочная трансатлантическая телеграфная связь, провод протянули между Европой и Америкой по дну Атлантического океана.

В XXI веке разные прогнозы получили широкое распространение, в том числе есть прогнозы на короткие периоды времени и на целый век вперед. Краткосрочный прогноз Института будущего (The Institute for the Future, <http://www.iftf.org>, создан RAND Corporation) содержит несколько трендов развития. Один из крупных трендов – Большие данные (обозначены в прогнозе как «Massively Multiplayer Data»), с акцентом на формирование науки о данных как целостной экосистемы.

Долгосрочный прогноз на весь XXI век представлен интеллектуальной платформой «Envisioning.io» (руководитель Майкл Заппа / Mitchell Zappa). Основной технологический тренд – возникновение неосоциума, объединенного цифровыми нейротехнологиями в единую нейросеть. Такие технологии продуцируют Большие данные и являются новыми каналами социальных коммуникаций между людьми. Ожидается конвергенция социальных наук и нейронаук, а главной наукой становится Data Science. На основе данных и обратной связи будет происходить регуляция жизнедеятельности неосоциума.

Если попытаться представить, как художник-футурист Жан-Марк Кот мог бы изобразить прогноз развития технологий к 2100 г., то на его рисунке шлемы каждого ученика были бы соединены проводами не только с машиной, перерабатывающей книги со знаниями, но и друг с другом, что выглядело бы как сложная сеть (нейросеть), объединяющая людей.



Прогноз «Technology Horizons. The Future of Science 2021», The Institute for the Future, 2011 г., www.iff.org.

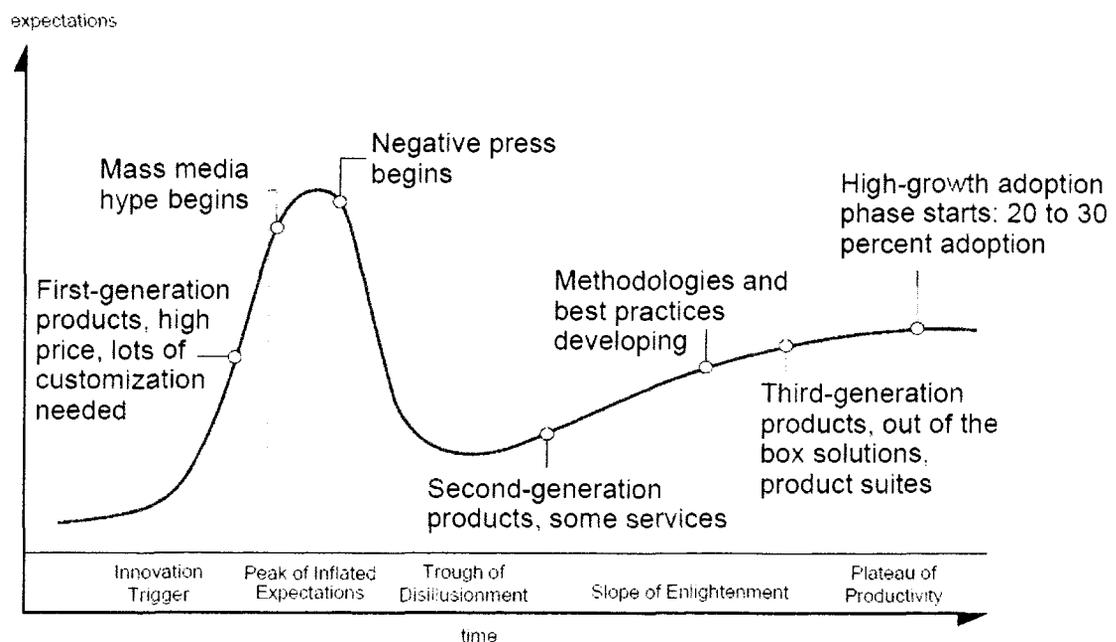
2030 – 2040 годы	Искусственный интеллект Artificial Intelligence	Расширенное познание Machine Augmented cognition
	Интернет Internet	Экзокортекс Exocortex
	Интерфейсы Interfaces	Полное погружение в виртуальную реальность Immersive virtual reality
	Сенсоры Sensors	Нейроинформатика Neuroinformatics
	Повсеместные вычисления Ubiquitous computing	Проецирование изображения на сетчатку глаза Retinal screens
	Робототехника Robotics	Самоорганизующаяся колония нанороботов Utility fog
	Биотехнологии BIOTECH	Лекарство от старости Antiaging drugs
Материалы Materials	Программируемая материя Programmable matter	

Прогноз появления новых технологий в 2030–2040 гг., сделанный в 2012 г. компанией Envisioning.io под руководством Майкла Заппа (Michell Zappa), <https://www.envisioning.io>.

Общемировую известность имеют прогнозы компании Gartner. Кривая Гартнера (Gartner Hype Cycle, или цикл общественного интереса к технологиям) – это признанное аналитическое графическое представление вхождения новых ИТ на мировой рынок, обновляемое ежегодно.

Первый график кривой Гартнера был разработан в 1995 г. Каждая новая технология проходит несколько этапов становления при меняющемся интересе к ней со стороны общества. Этап Innovation Trigger характеризуется первым появлением инновационной технологии, первыми публикациями о ней. Peak of Inflated Expectations – этап, когда новую технологию переоценивают, от нее ожидают кардинальных изменений в разных областях, ее активно обсуждают в прессе и в научном мире. Trough of Disillusionment – разочарование, избавление от иллюзий, выявляются недостатки новой технологии, переоцениваются ожидания, появляются негативные, критические публикации. Далее следует этап преодоления недостатков Slope of Enlightenment, просветление. Интерес к технологии возвращается, и она начинает внедряться в практику. Плато продуктивности Plateau of Productivity – это «золотой час» технологии или ее «зрелость», когда она востребована на рынке и широко используется до появления новых технологических трендов.

The Hype Cycle of Innovation



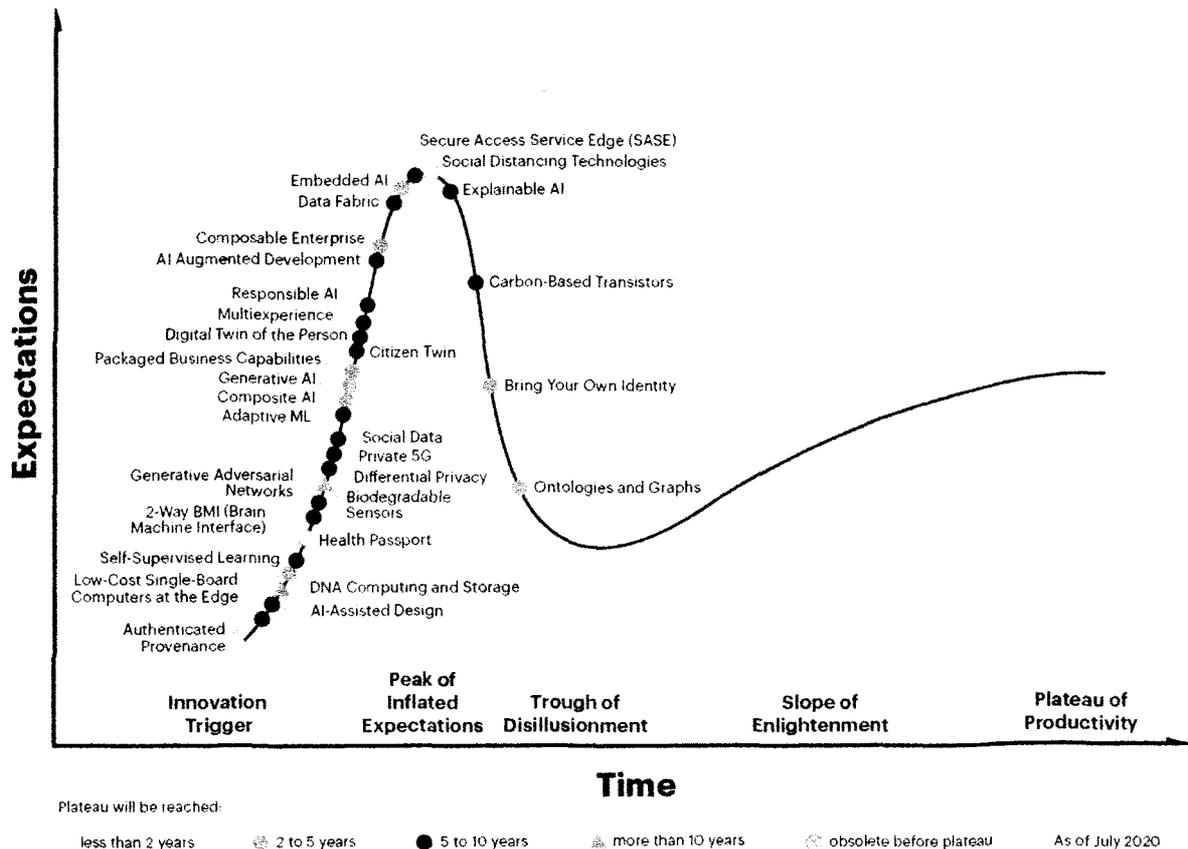
Кривая Гартнера (Gartner Hype Cycle). Описание см. в тексте. Источник: <https://www.gartner.com>.

Термин «Big Data» появился на кривой Гартнера в секторе Innovation Trigger в 2011 г., через три года после того как Клиффорд Линч предложил называть этим термином большие массивы данных и подходы к их сбору, хранению и аналитике. А последний раз Big Data / Большие данные были отмечены на кривой Гартнера в 2015 г., с указанием, что плато продуктивности будет достигнуто через 5–10 лет. Прогноз оказался верным, к 2020 г. Большие данные уже вошли в повседневный лексикон во всех отраслях экономики и стали неотъемлемой частью социальной жизни, бизнес-процессов и управления разными организациями, в том числе медицинскими. Аналитику Больших данных используют как обязательное условие роста и развития в медицине, бизнесе, на производстве, в сфере транспорта, в социальной сфере и т.д.

На кривой Гартнера за последние годы можно увидеть «вхождение» в цикл развития таких технологий, как искусственные ткани и 4D-печать, Общий (сильный) Искусственный интеллект (General Artificial Intelligence), квантовый компьютер, и др. На плато продуктивности успешно выходит дополненная реальность (Augmented Reality). Как указано на ресурсе «Smarter With Gartner», Искусственный интеллект, технология «блокчейн» (blockchain) и квантовый компьютер признаны экспертами Gartner базовыми

трансформирующими трендами, которые кардинально меняют экономику в ближайшее время.

Hype Cycle for Emerging Technologies, 2020



[gartner.com/SmarterWithGartner](https://www.gartner.com/SmarterWithGartner)

Source: Gartner
© 2020 Gartner, Inc. and/or its affiliates. All rights reserved. Gartner and Hype Cycle are registered trademarks of Gartner, Inc. and its affiliates in the U.S.

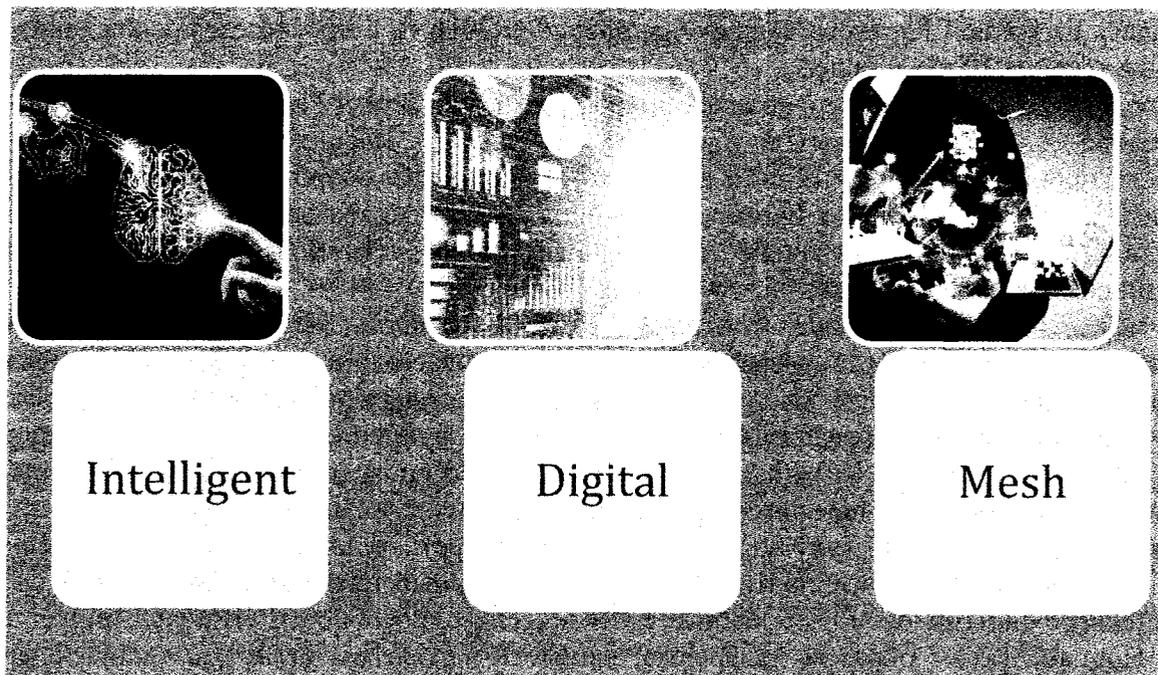
Gartner.

Кривая Гартнера (Gartner Hype Cycle) по итогам 2020 г.

Источник: <https://www.gartner.com/SmarterWithGartner>.

Вице-президент Gartner Дэвид Керли (David Cearley), выступая на конференции «Gartner Symposium / ITхро» в 2018 г. сказал, что современные технологии развиваются с учетом трех базовых характеристик:

- Intelligent, интеллектуальность технологий, создающих новые знания и новые возможности получать знания;
- Digital, цифровые технологии, проникающие в физический реальный мир;
- Mesh, сетевая связанность, позволяющая объединять людей, сервисы, приборы, инфраструктуру и машины в единое целое.



Три базовые характеристики новых технологий, предложенные Gartner.

В таком гибридном мире наука о данных превзойдет все другие традиционные научные направления, формируя новые конвергентные области. Социум все больше погружается в мир дополненной и виртуальной реальности (Augmented Reality, Virtual Reality). IoT формирует особую среду обитания человека, которую называют *умным пространством или интеллектуальной экосистемой* (Smart Space, Intelligent digitalized Ecosystem), куда входят умное рабочее место (Smart Workspace), умный дом (Smart Home), умный город (Smart City). Умное пространство – это совершенно новая социальная динамичная коммуникативная среда, где в непривычной нам форме смешаны разные сферы жизнедеятельности человека, где нет медицины «внутри стен больниц», а есть медицинское повсеместное окружение человека в повседневной жизни. Основные компоненты умного пространства – связанный социум, высокая мобильность, облачные сервисы, информационный обмен. Управление социальными экосистемами будет осуществляться через ИТ-платформы (или платформы ADV) с использованием аналитики Больших данных. Данные, обработанные на таких платформах, называют умными данными или Smart Data.

Аналитические платформы, предоставляющие для обычных пользователей готовые автоматизированные решения с использованием машинного обучения, нейросетей и Искусственного интеллекта (дополненная аналитика, Augmented Analytics), неизбежно привлекают все больше людей, не имеющих специального образования уровня Data Scientist. В связи с этим появился термин *«Citizen Data Science»* (что можно перевести как «обыватель в науке о данных»). А в будущем, по мере наполнения повседневной жизни человека цифровыми технологиями, каждый станет Citizen Data Science и будет ориентироваться в быту и на работе на выкладки различных аналитических автоматических платформ.

Такие аналитические платформы могут иметь и игровые формы взаимодействия с пользователем (геймификация жизни). Например, голографические экраны, жестовые интерфейсы, носимые и связанные с Интернетом устройства, виртуальный голосовой персональный помощник (Virtual Personal Assistants). Эти технологии не только смешивают физический и виртуальный мир вместе, но и могут представить этот гибридный мир в виде игры, которую будет сложно отличить от реальности. Еще больше данных появится по мере распространения такой технологии, как «умная пыль» (Smart Dust). Это самоорганизующиеся крошечные, размером с песчинку, групповые роботы (моты, от англ. mote «пылинка»). Они обмениваются беспроводными сигналами и данными, работают как единая система (гибкая сеть), имеют свои сенсоры, вычислительные узлы и системы энергоснабжения. Smart Dust

появилась на кривой Гартнера, в самом ее начале, в 2016 г., с прогнозом прохождения всех этапов в течение 10 лет и более.

Эксперты Gartner заявляют о появлении профессии Data Broker PaaS (Platform-as-a-Service), в переводе «посредник данных на платформе как услуга». Накопление данных в облачных сервисах происходит непрерывно, и актуальным становится оборот этих данных на рынке. Например, при помощи таких брокеров больница сможет зарабатывать на продаже наборов Больших данных фармкомпаниям или научным институтам больше, чем на оказании медицинских услуг пациентам.

В области Искусственного интеллекта, проникающего в цифровые экосистемы жизнедеятельности человека, разрабатывается так называемый «эмоциональный интеллект» (Affective computing, Artificial Emotional Intelligence). Благодаря попыткам понять, как устроен мозг человека, все больше расширяется направление по созданию искусственных нейронных сетей (Artificial Neural Networks), нейронных процессоров, мемристоров (Neuromorphic Hardware). Разработчики пытаются приблизить машину к человеку. Но есть и встречное движение – дополненный (улучшенный) человек. То есть человек стремится при помощи новых технологий (нейроимплантаты, нейрокомпьютерный интерфейс, экзоскелет) сделать более совершенными свои умственные и физические способности (Human Augmentation, Trans-Human). Таким образом, цифровые данные и алгоритмы станут важнейшей частью жизни человека и социума в новом, гибридном мире. Все инновационные тренды, формирующие неосоциум, будут способствовать становлению Data Science или науки о данных в качестве фундаментального научного направления XXI века.

Список литературы

1. Lynch C. How do your data grow? // *Nature*. – 2008. – Vol. 455. – Pp. 28–29.
2. Tukey J.W. The Future of Data Analysis // *The Annals of Mathematical Statistics*. – 1962. – Vol. 33. – № 1. – Pp. 1–67.
3. Майер-Шенбергер В., Кукьер К. Большие данные. Революция, которая изменит то, как мы живем, работаем и мыслим / пер. с англ. – М.: Манн, Иванов и Фербер, 2014. – 240 с.
4. Cukier K.N., Mayer-Schoenberger V. The Rise of Big Data, How it's changing the way we think about the world // *Foreign Affairs*. – May/June 2013. – Vol. 92. – № 3. – Pp. 28–41.
5. Репина Л.П., Зверева В.В., Парамонова М.Ю. История исторического знания: учебник для академического бакалавриата / под ред. Л.П. Репиной. – 4-е изд. – М.: Юрайт, 2018. – 258 с.
6. Вернадский В.И. Философские мысли натуралиста. – М.: Наука, 1988. – 520 с.
7. Тоффлер Э. Третья волна / пер. с англ. – М.: АСТ, 2004. – 781 с.
8. Гринин Л.Е. Кондратьевские волны, технологические уклады и теория производственных революций // *Кондратьевские волны. Аспекты и перспективы* / под ред. А.А. Акаев, Р.С. Гринберг, Л.Е. Гринин, А.В. Коротаев, С.Ю. Малков. – Волгоград: Учитель, 2012. – С. 222–262.
9. Гринин Л.Е., Гринин А.Л. Кибернетическая революция и шестой технологический уклад // *Кондратьевские волны: наследие и современность. Ежегодник* / под ред. Л.Е. Гринин, А.В. Коротаев, В.М. Бондаренко. – Волгоград: Учитель, 2015. – С. 83–106.
10. Харари Ю.Н. Homo Deus. Краткая история будущего / пер. с англ. – М.: Синдбад, 2018. – 496 с.
11. Непейвода Н.Н. Представление чисел с точки зрения вычислений, Доклад // *Национальный Суперкомпьютерный Форум, 2017. Переславль-Залесский, Институт программных систем им. А.К. Айламазяна РАН*. – URL: <https://www.youtube.com/watch?v=sLbjatWU8TA>.
12. Брусенцов Н.П. Об использовании троичного кода и троичной логики в цифровых машинах // *Вычислительная техника и вопросы кибернетики*. – 1970. – Вып. 7. – М.: Изд-во Моск. ун-та. – С. 3–33.
13. Сергиенко П.Я. Синтетическая геометрия. – Пушкино: ОНТИ ПНЦ, 2003. – 28 с.
14. Непейвода Н.Н., Григорьевский И.Н., Лилитко Е.П. О представлении действительных чисел // *Программные системы: теория и приложения*. – 2014. – Т. 5. – № 4 (22). – С. 105–121.
15. Непейвода Н.Н. О некоторых возможностях локальных вычислений в теории систем и базах данных // *Программные системы: теория и приложения*. – 2016. – Т. 7. – № 4 (31). – С. 145–160.
16. Винер Н. Кибернетика, или Управление и связь в животном и машине / пер. с англ. – М.: Советское радио, 1958. – 216 с.
17. Винер Н. Кибернетика и человек // *Природа*. – 1960. – № 8. – С. 68–69.
18. Форрестер Дж. Мировая динамика / пер. с англ. – М.: Terra Fantastica; АСТ; Neoclassic, 2003. – 384 с.
19. Капица С.П. Демографическая революция и будущее человечества // *В мире науки*. – 2004. – № 4. – С. 82–91.
20. Капица С.П. Очерк теории роста человечества. Демографическая революция и информационное общество // *К циклу публичных дискуссий «Россия в глобальном контексте». Никитский клуб ученых и предпринимателей. Специальный выпуск*. – М., 2008. – 64 с.
21. *Функциональные системы организма: Руководство* / Под ред. К.В. Судакова. – М.: Медицина, 1987. – 432 с.
22. Воеводин В.В. Суперкомпьютеры: вчера, сегодня, завтра // *Наука и жизнь*. – 2000. – № 5. – С. 76–83.
23. Абрамов С.М. Июнь 2019: анализ развития суперкомпьютерной отрасли в России и в мире // *Программные системы: теория и приложения*. – 2019. – Т. 10. – № 3 (42). – С. 3–40.

24. Pednault E., Gunnels J., Maslov D., Gambetta J. On “Quantum Supremacy” // IBM Research Blog. October 21, 2019. – URL: <https://www.ibm.com/blogs/research/2019/10/on-quantum-supremacy/>.
25. Villalonga B., Lyakh D., Boixo S., Neven H., Humble T.S., Biswas R., Rieffe E.G., Ho A., Mandrà S. Establishing the quantum supremacy frontier with a 281 Pflop/s simulation // Quantum Science and Technology. – 2020. – Vol. 5. – № 3. – 034003.
26. Emerging Technology from the arXivarchive. The new benchmark quantum computers must beat to achieve quantum supremacy // MIT Technology Review. – May 9, 2019. – URL: <https://www.technologyreview.com/2019/05/09/135440/the-new-benchmark-quantum-computers-must-beat-to-achieve-quantum-supremacy/>.
27. Panetta K. The CIO’s Guide to Quantum Computing // Smarter With Gartner. – November 29, 2017. – URL: <https://www.gartner.com/smarterwithgartner/the-cios-guide-to-quantum-computing/>.
28. Kandala A., Mezzacapo A., Temme K., Takita M., Brink M., Chow J.M., Gambetta J.M. Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets // Nature. – 2017. – Vol. 549. – Pp. 242–246.
29. Amin M.H., Andriyash E., Rolfe J., Kulchytskyy B., Melko R. Quantum Boltzmann Machine // Physical Review X. – 2018. – № 8. – 021050.
30. Langston J. How the quest for a scalable quantum computer is helping fight cancer // Innovation Stories, Microsoft, July 15, 2019. – URL: <https://news.microsoft.com/innovation-stories/quantum-computing-mri-cancer-treatment/>.
31. Kandala A., Mezzacapo A., Temme K. How to measure a molecule’s energy using a quantum computer // IBM Research Blog, September 13, 2017. – URL: <https://www.ibm.com/blogs/research/2017/09/quantum-molecule/>.

Для записей

Колесниченко Ольга Юрьевна

**DATA SCIENCE
(НАУКА О ДАННЫХ)
В СТАНОВЛЕНИИ
ИНФОРМАЦИОННОГО ОБЩЕСТВА**

Учебное пособие

Публикуется в авторской редакции

Дизайн обложки *Т.В. Середы*

Издательство «Прометей»
119002 Москва, ул. Арбат, д. 51, стр. 1
Тел.: +7 (495) 730-70-69
E-mail: info@prometej.su

Подписано в печать 11.12.2020.
Формат 60×84/8. Объем 6,5 п. л.
Тираж 500 экз. Заказ № 1262

ISBN 978-5-00172-110-9



9 785001 721109 >

45,000c.