

**Центр макроэкономического анализа и краткосрочного
прогнозирования**

Тел.: 8-499-129-17-22, факс: 8-499-129-09-22, e-mail: mail@forecast.ru

**Мониторинг и анализ технологического
развития России и мира.**

№ 34, 2 кв. 2023 г.



Август 2023

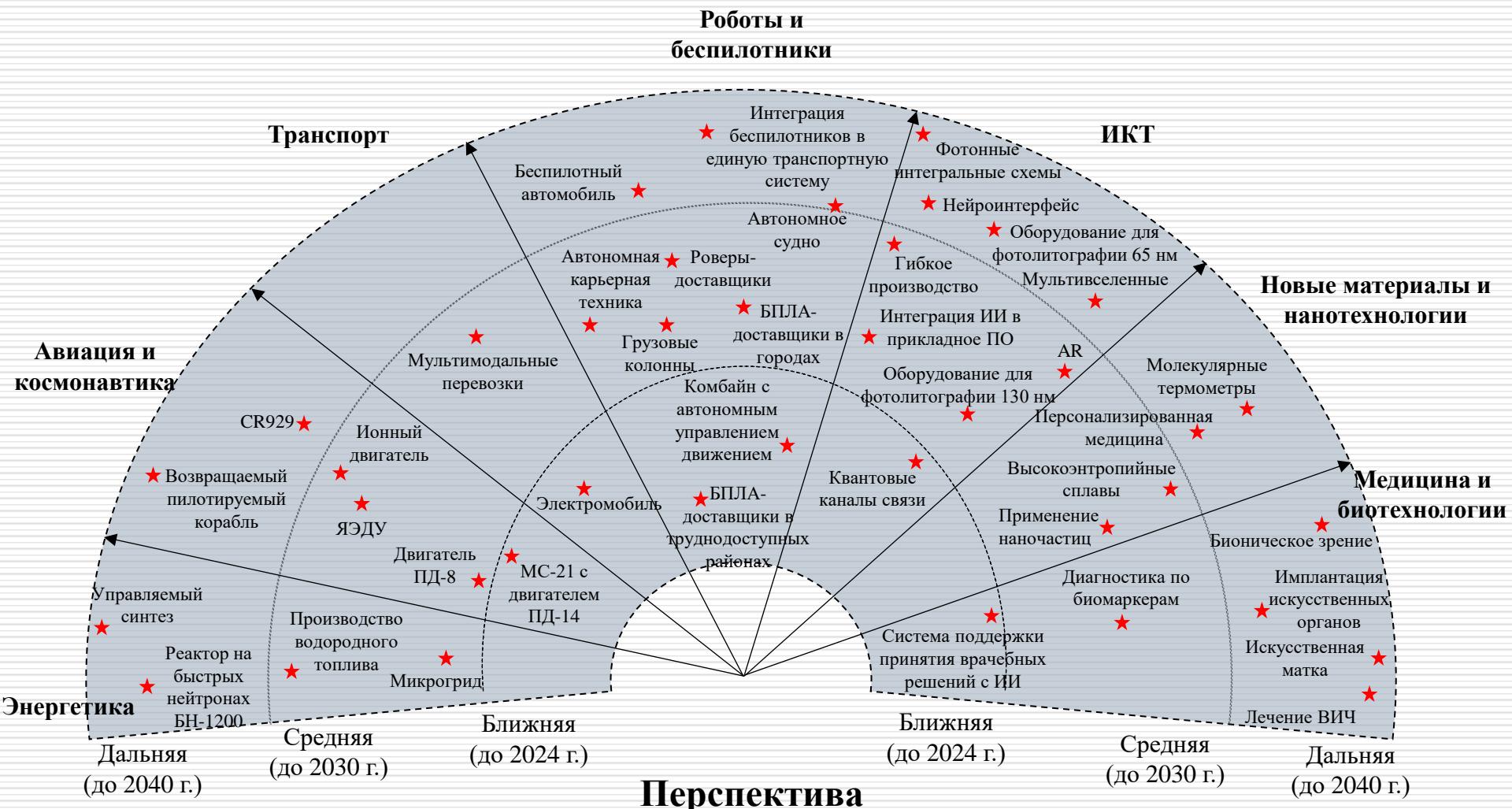
**В данной работе используются результаты проекта Т3-20, выполняемого в рамках
Программы фундаментальных исследований НИУ ВШЭ в 2023 г.**

Содержание

- 1. Технологические тренды в России**
 - I. Технологический радар России
 - II. Мониторинг технологического развития в России
- 2. Технологические тренды в мире**
 - I. Мировой технологический радар
 - II. Мониторинг технологического развития в мире
- 3. Тема номера: «Обработка естественного языка»**

1. Технологические тренды в России

1.1.1. Технологический радар России



1.2.1. Мониторинг технологического развития в России: энергетика

Углеводородная энергетика

- Ученые Самарского государственного технического университета разработали метод совершенствования углеводородной энергетики за счет возобновляемого источника – солнечной энергетики. Они предложили использовать солнечную энергию для термохимической трансформации углеводородного топлива – метана. Получаемый таким образом обогащённый водородом газ позволит повысить эффективность работы газотурбинных установок на 30% за счет одновременного снижения потребления углеводородного топлива и расширения использования низкопотенциальной солнечной энергии.

Ядерная энергетика

- Ученые Томского политехнического университета нашли способ увеличить длительность работы торий-уранового топлива на 75%. Такой эффект обеспечивает увеличение диаметра тепловыделяющего элемента (ТВЭЛ). Теплофизические расчеты показали практическую возможность модернизации ядерного реактора и увеличения срока службы и выгорания топлива. Это позволит повысить экономическую эффективность ядерных энергоблоков.

Термоядерный синтез

- НИЦ «Курчатовский институт» Михаил Ковальчук сообщил об успешном получении первой термоядерной плазмы на токамаке Т-15МД, который был запущен в мае 2021 г. и входит в структуру международного термоядерного проекта ITER (International Thermonuclear Experimental Reactor).

1.2.2. Мониторинг технологического развития в России: авиация и космонавтика

Космические технологии

- ГК «Роскосмос» полностью развернула на геостационарной орбите группировку гидрометеорологической космической системы «Электро» в составе трёх спутников «Электро-Л». Это дает возможность проводить космическую съёмку поверхности Земли с геостационарной орбиты с частотой до 15 минут. Также аппараты «Электро-Л» ретранслируют сигналы от аварийных радиобуев международной спутниковой поисково-спасательной системы КОСПАС-САРСАТ.
- Компания «БЮРО 1440» (ИКС Холдинг) успешно запустила первую испытательную миссию «Рассвет-1» и вывела на орбиту три космических аппарата, предназначенных для широкополосной передачи данных на высоких скоростях с минимальными задержками на базе низкоорбитальной спутниковой группировки.
- НПО «Энергомаш» (ГК «Роскосмос») произвело первый лётный двигатель РД-171МВ с тягой 800 т, что делает его самым мощным в мире. Двигатель предназначен для перспективной ракеты-носителя «Союз-5».

Авиационная техника

- «ОДК-Пермские моторы» запустило серийное производство турбовентиляторного двигателя ПД-14 для самолёта МС-21, что значительно повышает импортонезависимость производства среднемагистрального узкофюзеляжного самолёта МС-21.
- На заводе «Аэрокомпозит» (г. Ульяновск) началось серийное производство крыла для самолёта МС-21. В текущем году планируется выпуск 4 самолетокомплектов для МС-21, в 2024 г. – 6 самолетокомплектов. Производство практически полностью осуществляется из отечественных комплектующих и материалов, что снижает импортозависимость российского авиастроения.

1.2.3. Мониторинг технологического развития в России: транспорт

Беспилотные автомобили

- Первые беспилотные грузовики КамАЗ на базе серийных автомобилей КАМАЗ-43118 завершили опытно-промышленную эксплуатацию на дорогах-зимниках в Арктике. Два беспилотных автомобиля в течение нескольких месяцев перевозили грузы по 140-км зимнику от автономного нефтепромысла до посёлка Тазовский. Автомобили оснащены несколькими типами сенсоров, которые осуществляют мониторинг обстановки вокруг: строят цифровую карту дорог, распознают различного рода препятствия в радиусе 200 метров, определяют статичные и движущиеся объекты. Успешный опыт эксплуатации открывает возможности по расширению использования беспилотного автотранспорта.
- СберАвтоТех и Globaltruck сообщили о запуске коммерческих беспилотных грузоперевозок по маршруту Москва-Санкт-Петербург-Москва в пределах трассы М-11 «Нева». Поездка осуществляется в автономном режиме, однако на данном этапе в кабине постоянно находится водитель-испытатель.

ДВС

- Ученые РГАУ-МСХА имени К.А. Тимирязева разработали новую модель двигателя внутреннего сгорания — его себестоимость ожидается примерно на 30% ниже существующих конструкций (за счет применения штамповки, а не литья), ресурс работы, напротив — в два раза выше за счет того, что разработчикам удалось заметно снизить трение между рабочим поршнем двигателя и рабочим цилиндром. В качестве топлива может использоваться бензин или газ. Двигатель может использоваться для сельскохозяйственной техники и других машин.

1.2.4. Мониторинг технологического развития в России: работы и БПЛА

БПЛА

- «Лаборатория Касперского» сообщила о создании мобильного радиочастотного детектора для обнаружения беспилотников. По словам разработчика, детектор способен определять точное местоположение наиболее распространённых моделей дронов в радиусе до 1 км, а также определять положение оператора, который управляет БПЛА. Устройство весит менее 5 кг, помещается в кейс размером с чемодан для ручной клади и может работать в автономном режиме порядка двух часов. Основное направление использования – обеспечение безопасности и контроля коммерческих и государственных структур.
- Компания «Беспилотные аппараты» представила прототип реактивного беспилотника самолетного типа, способного развивать скорость до 400 км/ч (крейсерская скорость 200 км/ч). В аппарате применяется система искусственного интеллекта, позволяющая распознавать тепловые сигнатуры. Дальность полета – 90 км, масса полезной нагрузки – 6 кг. Дрон может использоваться для быстрой доставки грузов, мониторинга и контроля скоростных объектов и процессов, а также в военной сфере.

Работы

- В Москве запустили первого в России робота-уборщика «Пиксель», разработанного компанией «Автономика» (резидент кластера "Ломоносов" в научно-технологической долине МГУ). Программная часть, кузов, шасси и конструкция "Пикселя" разработаны российской компанией, элементы дизайна и системы управления – совместная разработка с белорусскими партнёрами. Робот может делать влажную уборку и подметать тротуары в течение 16 ч. Сменное навесное оборудование позволяет использовать робота круглый год. По состоянию на апрель 2023 г. робот работал в тестовом режиме в сопровождении оператора.

1.2.5. Мониторинг технологического развития в России: ИКТ

Искусственный интеллект

- Компания «Яндекс» внедрила большую языковую модель нового поколения YandexGPT в виртуального помощника «Алиса». С помощью нейросети «Алиса» может генерировать тексты и предлагать различные идеи. По данным разработчиков, «Алиса» справляется с поставленными задачами почти так же хорошо, как разбирающийся в теме человек. По состоянию на май 2023 г. система запущена в тестовом режиме. Технология может расширить возможности «Алисы» в качестве виртуального ассистента.
- Компания «Alliesverse» разработала виртуального юриста на базе искусственного интеллекта – LawAl. Он умеет отвечать на вопросы пользователей по юриспруденции, расшифровывать юридические термины и искать статьи из всех кодексов Российской Федерации. В мае 2023 г. разработчики выпустили бета-версию LawAl. Предполагается, что он может взять на себя основную часть рутинной работы юриста: поиск и анализ статей и т.д.
- Компания «Диагностика-М» создала и запустила в производство досмотровую технику для проверки багажа и грузов с использованием ИИ. Уровень локализации разработки – 100%. Интроскоп сможет самостоятельно идентифицировать опасные и запрещённые предметы, что упростит работу операторов досмотровой техники и повысит уровень общественной безопасности.

1.2.5. Мониторинг технологического развития в России: ИКТ

Коммуникационные технологии

- Компания «Швабе» (ГК «Ростех») сообщила о создании нового типа оптоволокна. Основное отличие инновационного оптоволокна от существующих аналогов – увеличенный до 100 мкм диаметр сердцевины (вместо типовых 50 и 62.5 мкм) и специальный градиентный профиль показателя преломления, благодаря которым обеспечивается высокая надежность сети передачи данных без потери пропускной способности. Такое оптоволокно может использоваться, например, в кабельных системах, работающих в агрессивных условиях окружающей среды.

Машинное обучение

- Исследователи Лаборатории искусственного интеллекта Сбербанка при поддержке учёных факультета компьютерных наук ВШЭ смогли увеличить скорость работы градиентного бустинга — одного из самых эффективных алгоритмов для решения задач машинного обучения. Анализ данных с целью прогнозирования часто сводится к классификации (отнесение объекта к определенной группе) или регрессии (прогнозирование числового значения). В случае, если количество классов или размерность регрессии очень велики, применяется градиентный бустинг. Он создает прогноз в виде совокупности слабых моделей, из которых в итоге получается одна эффективная модель. Таким образом, предложенный разработчиками подход позволит быстрее решать задачи классификации и регрессии в машинном обучении.

1.2.6. Мониторинг технологического развития в России: новые материалы и нанотехнологии

Новые материалы

- НПО «Технология» им. А.Г. Ромашина (г. Обнинск, ГК «Ростех») разработало теплопроводный углепластик нового поколения. Изделия, выполненные из этого материала, в три раза легче мировых аналогов и выдерживают нагрев до 400°С. Такая теплостойкость делает материал подходящим для изготовления радиаторов систем терморегулирования космических аппаратов – они будут значительно легче алюминиевых аналогов при более высокой прочности и лучшей теплопроводности.
- Учёные Сибирского федерального университета (СФУ, г. Красноярск) разработали метод усовершенствования стали Гадфильда за счёт увеличения её вязкости. Сталь Гадфильда – первая легированная и с высоким содержанием углерода и марганца сталь, которая применяется в самых сложных условиях (траки гусениц, рельсовые крестовины и т.д.), так как имеет высокую износостойкость при ударах, сильном давлении и перепадах температуры. Новый метод будет использоваться при обучении металлургов и литейщиков в профильном институте СФУ, а практические решения уже внедряются на одном из предприятий Красноярского края при производстве конусных броней для дробилок.

3D-печать

- В Заполярье с помощью 3D-принтера построено первое здание. Конструкция из пенобетона появилась в городе Лабытнанги Ямalo-Ненецкого автономного округа. При его строительстве применяли 3D-принтер российского производства, который позволил создать не только стены, но также мебель и элементы декора. Технология позволяет строить здания в 1,5 раза быстрее и на 30% дешевле, чем при традиционном строительстве.

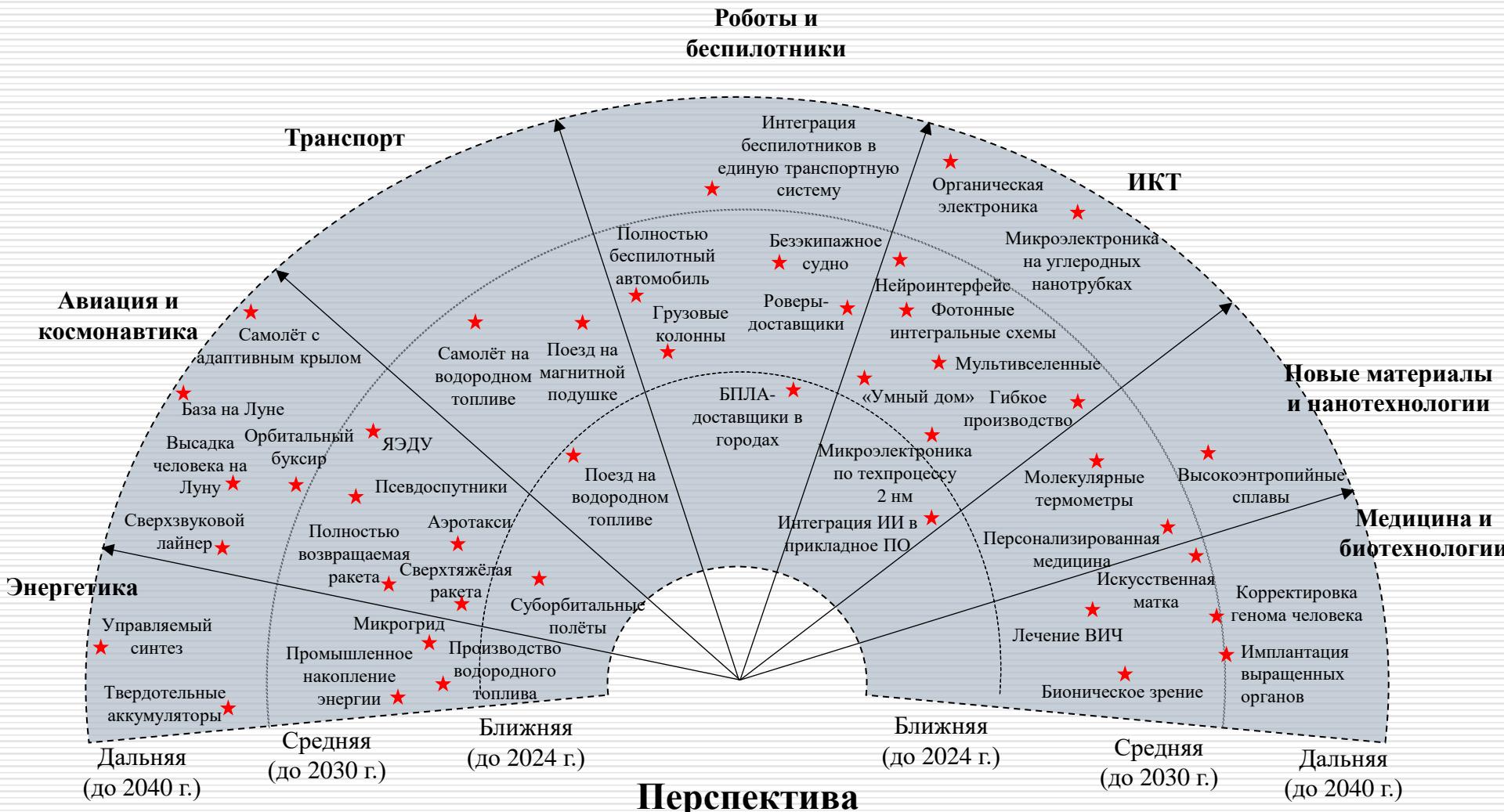
1.2.7. Мониторинг технологического развития в России: медицина и биотехнологии

Медицинские технологии

- Ученые Института автоматики и электрометрии (ИАиЭ) Сибирского отделения РАН совместно с коллегами из Новосибирского государственного университета разработали наносенсоры, работающие в терагерцовом диапазоне частот, которые позволяют распознавать на ранних этапах заболевания маркеры глиобластомы агрессивной опухоли головного мозга, которая плохо поддается ранней диагностике другими методами.
- В Сеченовском университете разработали искусственную роговицу на основе коллагена. Искусственная роговица получена при помощи инновационного метода электрофоретического осаждения. Она обладает высокой светопропускающей способностью (до 90%). При этом ученые могут регулировать ее толщину и набухаемость в зависимости от патологии пациента. Искусственная роговица может использоваться для замещения роговицы при травмах, а также при кератоконусе - заболевании глаза, при котором роговица принимает коническую форму, что приводит к ухудшению зрения.
- На Уральском оптико-механическом заводе им. Э.С. Яламова (ГК «Ростех») разработали портативный аппарат искусственной вентиляции лёгких АИВЛ-Д для транспортировки пациентов и применения в домашних условиях. Параметры прибора можно настроить для проведения дыхательной поддержки как взрослых, так и детей. Масса аппарата всего 4 кг. Встроенный аккумулятор позволяет работать до трёх часов без подзарядки. Аппарат оснащён влаго- и пылеустойчивым корпусом и бесшумен в работе. Таким образом, возможно его использование не только в стационарах, но также в домашних условиях.
- ГК Ростех получила регистрационное удостоверение на аппарат экстренного восстановления кровообращения, который поможет спасать жизни пациентов при внезапной остановке сердца и при острой дыхательной недостаточности.

2. Технологические тренды в мире

2.1.1. Мировой технологический радар



2.2.1. Мониторинг технологического развития в мире: энергетика

Биомеханическая энергия

- Инженеры Чанчуньского технологического университета (Китай) разработали гибкий прибор, который за счет получения и накопления биомеханической энергии (энергия движения человека). Устройство состоит из суперконденсатора и трибоэлектрического наногенератора. Суперконденсатор отвечает за хранение энергии и представляет собой «сэндвич» из никель-ванадиевого катода, нанесенного на углеродную ткань, цинк-содержащего электролита и углеродного анода. Генератор состоит из никель-ванадиевого композита на углеродной ткани и полимерного слоя из полидиметилсилоксана. Принцип работы наногенератора основан на трибоэлектрическом эффекте — природном явлении генерации разницы потенциалов при трении двух материалов. При растяжении или сжатии устройства при контакте с кожей происходит перераспределение зарядов между слоями и появление электрического тока. Суперконденсатор и наногенератор прибора соединены с помощью выпрямителя, который преобразует переменный ток в постоянный. За счет гибкости и легкого веса наногенератор можно закрепить на теле, а суперконденсатор надеть на запястье как браслет и использовать для заряда портативной электроники. Так, с его помощью уже удалось зарядить электронные наручные часы.

Устойчивое авиационное топливо

- Компания Honeywell International представила технологию производства низкоуглеродного авиационного топлива из водорода и углекислого газа. Технология позволяет использовать водород, производимый в электролизёрах из возобновляемых источников энергии и воды, и углекислый газ, откачиваемый из промышленных дымовых труб, для создания менее углеродистого метанола, который затем превращается в устойчивое авиационное топливо (SAF). По данным разработчиков, технология может сократить выбросы парниковых газов на 88% по сравнению с традиционным реактивным топливом, получаемым из нефти.

2.2.2. Мониторинг технологического развития в мире: авиация и космонавтика

Космическая техника

- Компания Virgin Galactic осуществила первый полностью коммерческий суборбитальный полёт ракетоплана SpaceShipTwo, который запускается с двухфюзеляжного самолёта-носителя. В ходе 14-минутного полёта была достигнута высота 85 км.
- Американское управление по аeronавтике и исследованию космического пространства NASA сообщило о двух новых рекордах вертолёта Ingenuity, установленных в ходе полётов на Марсе — по скорости и высоте. Вертолёт массой 1,8 кг достиг максимальной скорости 23,3 км/ч и максимальной высоты 16 метров. Предыдущие рекорды составляли 21,6 км/ч и 14 м соответственно. Новые рекорды расширяют возможности исследования Марса, а также пополняют научно-техническую базу для создания новых летательных аппаратов для будущих марсианских миссий.

Гиперзвук

- Компания Stratolaunch (США) произвела успешный запуск испытательного аппарата TA-0 со своего самолета-носителя Roc. TA-0 должен будет развивать скорость более 5 Маха. После завершения полета он самостоятельно приземляется на взлетно-посадочную полосу. Успешный эксперимент означает шаг вперед в создании системы гиперзвукового полета.

Авиадвигатель

- General Electric разработала авиадвигатель с открытым ротором, сочетающий в себе топливную экономичность турбовинтового двигателя со скоростью и производительностью турбовентиляторного двигателя. На двигателе без воздуховода в задней части установлены два винта особой формы. Считается, что такая конструкция позволит повысить экономичность. General Electric подписала соглашение с Airbus о проведение лётных испытаний прототипа.

2.2.3. Мониторинг технологического развития в мире: транспорт

Автономный транспорт

- В Японии (префектура Фукуи) запустили линию беспилотного транспорта. Семиместный электромобиль, по компоновке напоминающий гольф-кар, может преодолевать двухкилометровый маршрут за 10 минут, максимальная скорость — 12 км/ч. Линия стала первой в стране линией общественного транспорта с использованием 4 уровня автономного транспорта. Данный уровень предполагает фиксированный маршрут движения и не требует контроля со стороны водителя или удаленного оператора.

Автомобильный транспорт

- Компания Weishi Energy (Китай) представила автомобильную систему хранения жидкого водорода Jupiter, призванную обеспечить запас хода для водородных автомобилей более чем в 1000 км. Система прошла основные тесты (падение, вибрация и т.д.) и готова к серийному производству. Система позволяет выдавать пиковую удельную мощность 4 кВт/л, подходящую для тяжелых грузовиков и перевозок на дальние расстояния.

Гибрид мультикоптера и автомобиля

- Китайская компания Guangzhou Automobile Group (GAC) представила гибрид мультикоптера и автомобиля GOVE, объединяющий самодвижущуюся четырехколесную платформу и отделяющуюся от нее пассажирскую кабину с 6 вертикальными электромоторами. При необходимости дрон и шасси смогут разделяться для перехода в режим полета. По задумке разработчиков, шасси можно будет управлять дистанционно и отправлять его подзаряжаться на станцию.

2.2.4. Мониторинг технологического развития в мире: работы и беспилотники

БПЛА

- Учёные из Имперского колледжа Лондона (Великобритания) и исследовательского института Empa (Швейцария) разработали жаропрочный дрон, способный выдерживать температуру до 200 °С в течение 10 минут. Термостойкость обусловлена, главным образом, изолирующим слоем аэрогеля, который образует оболочку вокруг электроники. Внешняя оболочка из алюминия дополнительно отражает тепло. Дрон может использоваться для работы в агрессивных температурных средах, например, в пожарах.

Работы

- Инженеры из Федеральной политехнической школы Лозанны (Швейцария) разработали робота, состоящего из треугольных модулей. Это позволяет роботу трансформироваться практически в любую форму наподобие оригами. В частности, разработчики собрали из нескольких базовых модулей манипулятор, подвижную гусеницу и четырехногого робота. Таким образом такой робот может использоваться для выполнения широкого спектра задач, заменяя собой большое количество специализированных устройств.
- Ученые из Осакского университета (Япония) разработали новый тип шагающего робота-многоноски, обладающего высокой устойчивостью к поломкам отдельных конечностей и низкими требованиями к вычислительным мощностям для расчёта движений. Робот состоит из 6 сегментов, у каждого из которых 2 ноги с гибкими суставами.
- В Университет штата Вашингтон (США) разработали роботизированную пчелу, способную летать во всех направлениях и даже поворачивать. Это первый прототип робота-пчелы, который стабильно летает во всех направлениях – в том числе под углом относительно вертикальной оси. Возможная область применения - опыление растений, поисковые работы в ограниченном пространстве и т.д.

2.2.5. Мониторинг технологического развития в мире: ИКТ

Искусственный интеллект

- Учёные из Уханьского университета (Китай) протестировали систему искусственного интеллекта для автономного управления спутников дистанционного зондирования земли. В ходе 24-часового эксперимента спутником Qimingxing 1 были просканированы несколько объектов. Эксперимент дал новые представления о возможностях автономных космических аппаратов.

Квантовые технологии

- Компания IBM (США) заявила о прогрессе в работе квантовых компьютеров: впервые система, состоящая из 127 кубитов продемонстрировала возможность давать более точные результаты в сложных вычислениях, чем традиционные суперкомпьютеры. Достичь такого результата помог предложенный разработчиками специальный метод для исправления ошибок в квантовых вычислениях.

Передача данных

- В Федеральном институте технологий (ETH Zürich, Швейцария) совместно с рядом космических компаний разработали технологию передачи данных с помощью лазера по воздуху со скоростью несколько Тбит/с. В ходе эксперимента с помощью лазеров был передан сигнал между альпийской горной вершиной Юнгфраугох и Бернской обсерваторией (расстояние около 53 км). Несмотря на турбулентность воздуха, которая создает помехи, удалось достичь чистой средней скорости передачи данных 0,94 Тбит/с. Преимущество технологии в том, что лазерные системы могут использовать ближний инфракрасный диапазон с длинами волн в несколько микрометров, что примерно в 10 000 раз короче радиоволн, что позволяет передавать больше информации в единицу времени. Технология может использоваться в системах, требующих беспроводной передачи больших объемов данных, например, в спутниковой связи.

2.2.6. Мониторинг технологического развития в мире: новые материалы и нанотехнологии

Новые материалы

- В Национальной лаборатории Ок-Риджа Министерства энергетики США изобрели покрытие, которое значительно уменьшает трение. В основе разработки —углеродные нанотрубки, выращенные на поверхности стали. В процессе трения из таких трубок образуется стружка. Вместе со смазкой стружка образует богатую графеном пленку, которая снижает трение почти до нуля: анализ показал, что использование такой смазочной поверхности снижает коэффициент трения до 0,001. При этом свойства сохранялись при испытаниях более чем на 500 000 циклов трения. Технология может применяться в несущих системах с движущимися частями, от приводов транспортных средств до ветряных и гидроэлектрических турбин.

3D-печать

- Ученые из Университета штата Вашингтон (США) разработали способ 3D-печати двух типов стали в одном круговом слое с использованием двух сварочных аппаратов. Полученный биметаллический материал оказался на 33-42% прочнее каждого металла по отдельности за счет давления, возникающего между металлами в процессе совместного охлаждения. Современные методы 3D-печати несколькими металлами требуют остановки сварочной установки и замены металлических проволок. Новый метод устраняет эту паузу и помещает два или более металла в один и тот же слой, пока металлы еще горячие, что позволяет обеспечить одновременное охлаждение материалов и, соответственно, повысить прочность полученного материала.
- Компания Stakeholder Foods (Израиль) впервые в мире при помощи 3D-принтера создала аналог мяса рыбы. 3D-печать осуществлялась с помощью биочернил, в составе которых содержались рыбы клетки. Развитие данной технологии расширяет возможности по производству альтернативных пищевых белков.

2.2.7. Мониторинг технологического развития в мире: медицина и биотехнологии

Клеточные технологии

- Ученые из Орхусского университета (Дания) разработали новый метод управления формой ДНК. На основе взаимодействия Хугстена к двум нитям ДНК присоединяется третья, образуя тройную спираль. Полученная триплекс нить способна изгибать или «сворачивать» двухцепочечную ДНК. Таким образом ученые смогли достичь нового уровня контроля над двухцепочечной ДНК. Метод может найти применение в генных технологиях.
- Ученые Базельского университета (Швейцария) определили механизм, который запускает процесс «самоуничтожения» клеток. Получив команду самоликвидации, два белка ниндзюрин-1 объединяются в нитевидные структуры и «протыкают» мембрану. Постепенно клеточная мембрана расщепляется по частям и клетка разрушается. Более глубокое понимание процесса гибели клеток может помочь в разработке лекарств от болезней, связанных с клеточными патологиями, такими как рак и болезнь Паркинсона.
- Американские молекулярные биологи создали двойную генную терапию на базе геномного редактора CRISPR/Cas9, которая одновременно очищает зараженные клетки от большинства следов ВИЧ и меняет ДНК здоровых иммунных клеток таким образом, что те становятся неуязвимыми для атак вируса, сообщила пресс-служба университета Темпл.

Медицинские технологии

- Ученые из Кембриджского университета (Великобритания) создали из стволовых клеток модель человеческого эмбриона. Она представляет собой организованную трехмерную структуру, полученную из плuriпотентных стволовых клеток, которые воспроизводят некоторые процессы развития, происходящие у ранних человеческих эмбрионов. Такая модель позволит воспроизвести развитие эмбриона на второй неделе беременности, когда эмбрион имплантируется в матку и когда есть риск потери беременности, а также проанаблюдать развитие органов и специализированных клеток, выявить причины ранней потери беременности и тд.

тема номера:

3. Обработка естественного языка

3.1. Введение

Терминология

- Обработка естественного языка (англ. - *Natural Language Processing, NLP*) находится на пересечении лингвистики с искусственным интеллектом и машинным обучением.
- В NLP входят понимание естественного языка (англ. - *Natural Language Understanding, NLU*) и генерация естественного языка (англ. - *Natural Language Generation, NLG*), которые во многих случаях неразрывно связаны друг с другом.

История

- NLP восходит к исследованиям в области компьютерной лингвистики 1950-70х годов, в которых предпринимались попытки на основе грамматической структуры языка генерировать связные предложения или выполнять команды пользователя, сформулированные в виде предложений естественного языка.
- Более современные методы NLP опираются на адаптированные для обработки языка методы глубокого обучения (англ. - *Deep learning*), в которых текст оцифровывается и обрабатывается нейронными сетями.
- Такой подход позволил получить впечатляющие результаты в обработке и генерации текстов на естественном языке: распознавание речи, «умный» поиск ответов на вопросы, поддержка разговора с человеком на произвольную тему, извлечение нужной информации из текста, качественный перевод и т.п.
- К настоящему времени создано большое количество прикладных сервисов, основанных на технологиях NLP и обладающих разнообразным функционалом

3.2. Примеры сервисов

Распознавание речи

- Предполагает способность программы идентифицировать слова в речи человека и преобразовать их в машиночитаемый формат.
- Технологией владеет большое количество ИТ компаний.

Переводчики

- Сервисы для перевода развиваются несколько десятилетий. Современные версии позволяют переводить как текст, так и аудио информацию, в т.ч. в режиме реального времени с нескольких десятков языков.
- Примеры сервисов: Google Переводчик, IBM Watson Language Translator (США), DeepL Переводчик (Германия), AliTranx (КНР), Яндекс.Переводчик (Россия) и др.

Автоматический анализ текстов

- Предполагает автоматическое сканирование текста и извлечение из него нужных сущностей: фактов, ответов на вопросы и т.п.
- Microsoft text analytics позволяет извлекать нужные сведения из неструктурированного текста, анализировать тональность, выделять ключевые слова и именованные сущности.
- IBM Watson Explorer — набор моделей на основе IBM Watson для когнитивной аналитики контента. В частности, Watson Explorer Content Miner позволяет удалить посторонние результаты и сфокусировать внимание пользователя на релевантной информации. Watson Explorer предоставляет несколько способов визуальной организации и навигации по неструктурированным данным. Watson Explorer использует возможности машинного обучения для выявления закономерностей в данных.

3.2. Примеры сервисов

Голосовые помощники (*англ. - Virtual assistant*)

- Представляют собой программные агенты (т.е. программы, предназначенные для упрощения взаимодействия пользователя с другими программами), которые могут на основе команд пользователя, сформулированных на естественном языке, выполнять для него полезные действия с использованием данных других программ: искать ответы на вопросы, подбирать товары и услуги, бронировать билеты и т.п.
- Голосовые помощники сочетают в себе технологии распознавания голоса, обработки естественного языка и взаимодействия с другими сервисами.
- Первым голосовым помощником стал Siri, выпущенный в 2011 году и вскоре выкупленный компанией Apple. Он был разработан Международным центром искусственного интеллекта SRI, который до этого на протяжении 40 лет вёл проект CALO — Cognitive Agent that Learns and Organizes Program (рус. — Программа «Познавательный агент, который учится и организует»), финансируемый DARPA.
- Широкое распространение получили Voice Mate (LG, 2012 г.), S Voice и его развитие - Bixby (Samsung, 2012 и 2017 гг. соответственно), Amazon Alexa (2014 г.), Cortana (Microsoft, 2014 г.), Google Assistant (2016 г.) и др.
- Российскими компаниями созданы голосовые помощники Алиса (Яндекс, 2017 г.), Маруся (ВКонтакте, 2019 г.), Салют (Сбер, 2020 г.).

3.2. Примеры сервисов

Анализ тональности (англ. - *sentiment analysis*)

- Предполагает машинное распознавание эмоций и отношения автора к предмету.
- Позволяет обогатить ключевые слова контекстом и более точно сформулировать смысл текста.

Анализ тональности применяется в ряде областей:

- Маркетинг, связи с общественностью (PR), таргетированная реклама: на основе отзывов выявляются сильные и слабые стороны товаров и услуг компаний, выявляются интересы и предпочтения пользователей сеансов, запросы потребителей и т.п.
- Политология: исследуются политические взгляды населения, отношение различных социальных групп к различным событиям и персонам, в т.ч. с учётом отдельных аспектов и характеристик.
- Психология, медицина: выявляются признаки депрессии и иных расстройств.
- Известны по меньшей мере сотни моделей анализа настроения, опубликованных для использования на языке Python в качестве открытого исходного кода.
- Для русского языка адаптирована библиотека для обработки языка Dostoevsky, (использует язык Python).
- Анализ настроений интегрирован в платформы Microsoft text analytics , IBM Watson и т.д.
- В психологической науке используют сервис LIWC, выделяющий в тексте шаблоны, отражающие эмоции, стили мышления и социальные проблемы. На сегодняшний день ведётся разработка русскоязычного словаря для LIWC.
- Российскими учеными описан подход по использованию бинарной классификации с моделью обучения LSTN-сети, который может составить психологический портрет человека по записи его речи в повседневной жизни.

3.2. Примеры сервисов

Медицина и фармацевтика

- В медицине технологии NLP применяются для упрощения заполнения врачом электронных медицинских карт (ЭМК), круглосуточного фиксирования жалоб пациентов (технология «окружающего клинического интеллекта – *ambient clinical intelligence – ACI*) и обобщения неструктурированной информации для последующего применения специалистами.
- Например, приложение Nuance Communications (США, Канада) позволяет медицинскому персоналу надиктовывать текст для ЭМК, сокращая время заполнения на 45%, а голосовой помощник для пациентов сокращает расходы на круглосуточную поддержку на 25%.
- В России группа компаний ЦРТ разработала программу для голосового заполнения медицинской документации Voice2Med, которая адаптирована для распознавания медицинских терминов и сокращает время заполнения ЭМК в 2-3 раза.
- Помимо сбора данных и заполнения ЭМК, NLP помогает преобразовывать неструктурированные данные в структурированный формат, извлекать и нормализовывать (приводить к словарной форме) понятия, определять их контекст, что делает возможным использование собранных данных для научных исследований.
- Например, платформы Linguamatics NLP (США, Великобритания) или Amazon Comprehend Medical извлекают информацию медицинского характера из неструктурированных данных ЭМК, научных публикаций, отчётов о клинических испытаниях лекарств и т.д., что позволяет исследователям и разработчикам медицинских препаратов получать информацию в удобной форме по интересующим их вопросам.
- В России компания К-Скай (Резидент Сколково) разработала платформу прогнозной аналитики Webiomed, которая на основе данных миллионов пациентов выявляет новые предпосылки или признаки распространённых заболеваний, а также предлагает лечащему врачу возможные диагнозы на основе клинической картины конкретного пациента.

3.2. Примеры сервисов

Чат-боты

- В 2022-23 годах получили широкий резонанс чат-боты, способные поддерживать разговор на свободную тему и объяснять пользователю интересующую его информацию на доступном для него языке.
- ChatGPT — разработанный компанией OpenAI на основе большой языковой модели.
- Bard — чат-бот с искусственным интеллектом, разработанный компанией Google на основе языковой модели LaMDA.
- GigaChat - чат-бот, обученный Сбером (Россия) на русскоязычных текстах. По заявлению разработчиков, бот способен генерировать код на языке Python по запросу пользователя, составлять письма и рассказы по тезисам пользователя и т.д.

AI Copywriting Software

- Rytr, Nichesss, Jasper, Hypotenuse AI, SEO Content Machine и другие сервисы упрощают создание текста нужной стилистики по тезисам автора или наоборот помогают выделить ключевые тезисы, краткое содержание, а также сгенерировать вопросы по тексту.
- Некоторые из таких сервисов могут использовать данные поиска гугл или иных сервисов.
- Большинство программных продуктов для копирайта адаптированы для наполнения контентом сайтов.
- Остается не реализованной интеграция такого рода программных продуктов в стандартные текстовые редакторы широкого профиля.

3.3. Перспективные технологии

Извлечение сущностей из текста

- В 2019-23 годах американский институт развития IARPA, финансирующий развитие новых технологий в интересах разведывательного сообщества США, финансировал проект по извлечению сущностей из неструктурированных текстов под кодовым названием BETTER.
- В 2023 году IARPA запустил проект REASON (англ. – Rapid Explanation, Analysis and Sourcing Online), направленный на разработку системы ИИ, которая будет помогать аналитикам в поиске доказательств, сильных и слабых сторон своих суждений на основе всех доступных данных.

Межкультурная коммуникация

- До сих пор переводчики на основе технологий ИИ позволяют переводить точный смысл текста с одного языка на другой.
- Одним из возможных направлений развития NLP является перевод с языка одной культуры на язык другой культуры, т.е. перевод с учётом культурного контекста вместо перевода буквального смысла.
- Помимо неизменного культурного контекста могут меняться эмоции участников диалога, которые по-разному проявляют представители различных культур.
- Созданию такой системы посвящён проект CCU (англ. - *Computational Cultural Understanding*), финансируемый DARPA в интересах Министерства обороны США.
- CCU стремится создавать технологии обработки естественного языка, которые распознают, адаптируются и рекомендуют, как действовать в рамках эмоциональных, социальных и культурных норм, которые различаются в разных обществах, языках и групповых связях.
- Система ИИ создаётся с учётом достижений психологии, социологии и других гуманитарных наук.

3.4. Выводы

- Продукты, разработанные на основе технологий NLP, принимают на себя часть функций человека по обработке текстовой информации.
- Современные инструменты позволяют упростить часть задач (перевод) или создают возможность в короткий срок обработать очень большие объёмы информации (извлечение сущностей).
- В результате, пользователи NLP-продуктов получают как количественное (способность обработать большие массивы информации), так и качественные (способность сделать то, что не под силу другим) преимущества перед своими коллегами, работающими без таких сервисов.
- Необходимым условием реализации возможностей технологий NLP является их рутинизация через интеграцию в стандартные текстовые редакторы и сервисы поиска информации.
- В случае широкого распространения систем ИИ большое влияние приобретут платформы, предоставляющие услуги ИИ по машинной обработке естественного языка и по организации совместной работы коллективов специалистов.
- В случае дальнейшего успешного развития технологий NLP и их интеграции в стандартные сервисы переход к интеллектуальному труду с использованием NLP-помощников может приобрести масштабы, сопоставимые с переходом от ручного труда к машинному.

Отдельные использованные сокращения

- БПЛА – беспилотный летательный аппарат;
- ВИЧ – вирус иммунодефицита человека;
- ГК – государственная корпорация;
- ДВС – двигатель внутреннего сгорания;
- ИИ – искусственный интеллект;
- ИКТ – информационно-коммуникационные технологии;
- ПО – программное обеспечение;
- ТВЭЛ – тепловыделяющий элемент;
- ЯЭДУ – ядерная энергетическая двигательная установка;
- IARPA – Intelligence Advanced Research Projects Activity (англ., Агентство передовых исследований в сфере разведки)
- NLP – Natural Language Processing (англ., обработка естественного языка)
- SAF – Sustainable Aviation Fuel (англ., устойчивое авиационное топливо).

Использованные источники

Новостные сайты

russiandrone.ru; vedomosti.ru; aviation21.ru; rbc.ru; refnews.ru; nangs.org; roscosmos.ru; ria.ru; rostec.ru; proryv2020.ru; atomic-energy.ru; sk.ru; minenergo.gov.ru; nplus1.ru; robo-hunter.com; robogeek.ru; topwar.ru; tass.ru; techfusion.ru; popmech.ru; rb.ru; web-canape.ru; nag.ru; cnews.ru; kommersant.ru; hi-news.ru, hightech.fm, hightech.plus, i-mash.ru, interfax.ru, mashportal.ru, rg.ru; phys.org; fortune.com; news.tpu.ru; naked-science.ru; scientificrussia.ru; misis.ru; phys.org, mdpi.com, azom.com, nature.com, boeing.com, sciencedaily.com, newatlas.com, smart-energy.com, techinsider.ru, 3dnews.ru, naukatehnika.com и др.

Источники про обработку естественного языка

- 5 Best AI Copywriting Tools in 2023 [Out of the 11 We Tested]. URL: <https://peertopeermarketing.co/ai-copywriting-tools/>
- Computational Cultural Understanding (CCU). URL: <https://www.darpa.mil/program/computational-cultural-understanding>
- Методы обработки естественного языка. URL: <https://developers.sber.ru/help/ml/natural-language-processing-techniques>
- ИИ и Natural Language Processing: большой обзор рынка. Часть 2. URL: <https://vc.ru/future/102380-ii-i-natural-language-processing-bolshoy-obzor-rynka-chast-2>
- Обучаем компьютер чувствам (sentiment analysis по-русски). URL: <https://habr.com/ru/articles/149605/>
- Panicheva P., Litvinova T. Matching LIWC with Russian Thesauri: An Exploratory Study // Artificial Intelligence and Natural Language: 9th Conference, AINL 2020: Helsinki, Finland, October 7—9, 2020: Proceedings / Eds. A. Filchenkov, J. Kauttonen, L. Pivovarova. Cham: Springer, 2020. P. 181—195. DOI:10.1007/978-3-030-59082-6_1
- Татарникова Т. М., Богданов П. Ю. Построение психологического портрета человека с применением технологий обработки естественного языка //Научно-технический вестник информационных технологий, механики и оптики. – 2021. – Т. 21. – №. 1. – С. 85-91.)

Авторы обзора:

Артёменко Владимир, эксперт,

avg@forecast.ru

Волков Роман, ведущий эксперт

rvolkov@forecast.ru