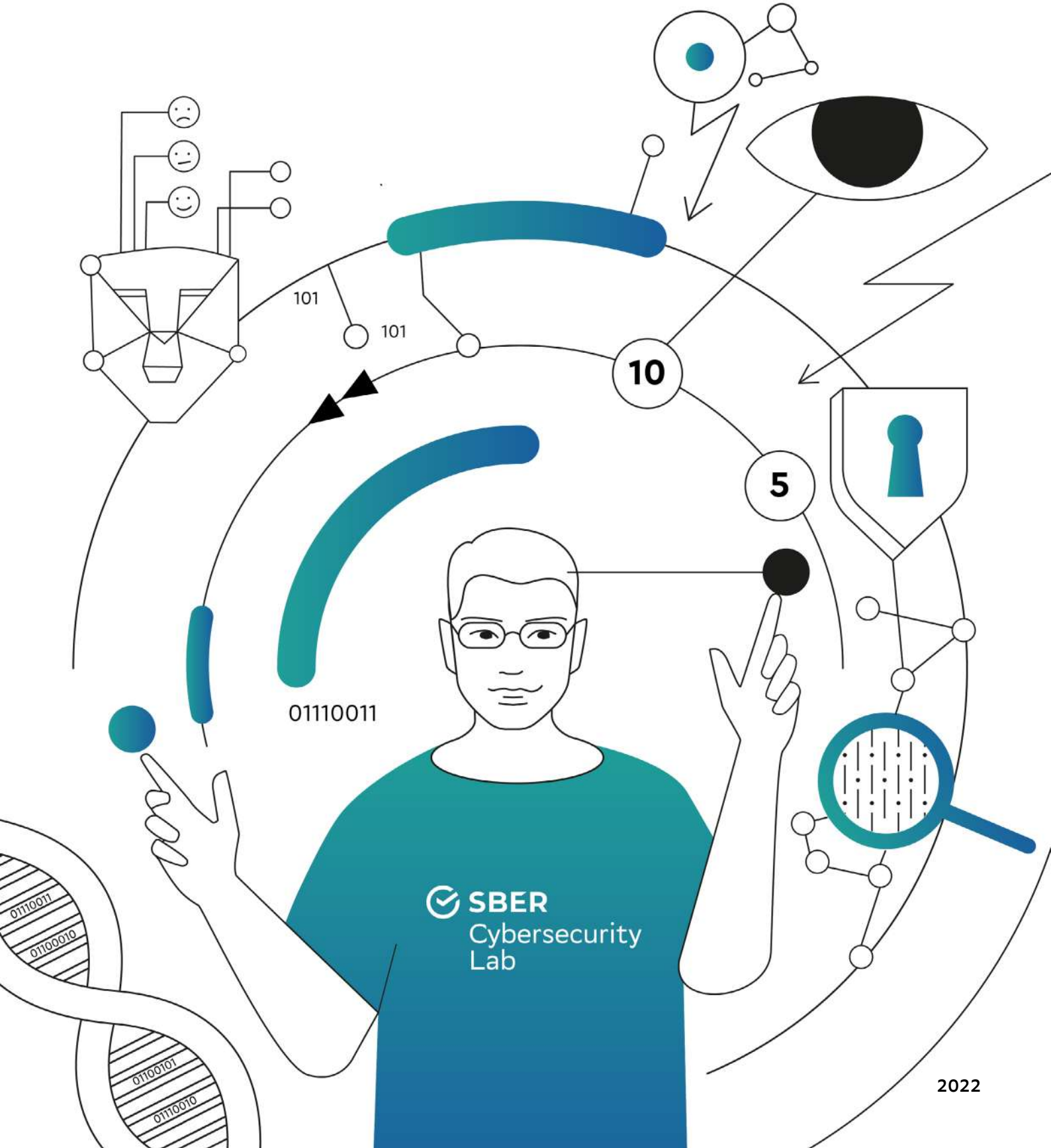
АНАЛИТИЧЕСКИЙ ОТЧЕТ

# Прогноз влияния перспективных технологий на ландшафт угроз кибербезопасности




2022

**СТАНИСЛАВ КУЗНЕЦОВ,**
Заместитель Председателя Правления
ПАО Сбербанк

Высокотехнологичное будущее расширяет наши возможности, открывая при этом дорогу одновременно и для хорошего, и для плохого. Лаборатория кибербезопасности Сбера работает над самыми серьёзными вызовами цифровизации: прогнозирует риски, изучает методы противостояния угрозам, создаёт уникальные технологии кибербезопасности.

Рад представить вашему вниманию исследование нашей Лаборатории, которое показывает, как может измениться ландшафт киберугроз под влиянием новых технологий. Этот аналитический отчёт станет хорошим подспорьем для тех профессионалов, кто не только ведёт борьбу с киберпреступностью в текущем моменте, но и заблаговременно готовится к отражению новых угроз.

**СЕРГЕЙ ЛЕБЕДЬ,**
Вице-президент, директор
департамента кибербезопасности
ПАО Сбербанк

Стремительное развитие технологий изменило нашу жизнь до неузнаваемости и открыло нам новые возможности для общения, работы и развлечений. Но вместе с развитием технологий развиваются и эволюционируют угрозы кибербезопасности.

Любые новые технологии, даже созданные с самыми благими намерениями, очень быстро становятся грозным оружием в руках злоумышленников в киберпространстве, охотящихся за нашими деньгами и данными. Поэтому сейчас для специалистов в области кибербезопасности крайне важно уделять внимание не только текущим угрозам, но и угрозам завтрашнего дня, с которыми нам только предстоит столкнуться.

# Введение

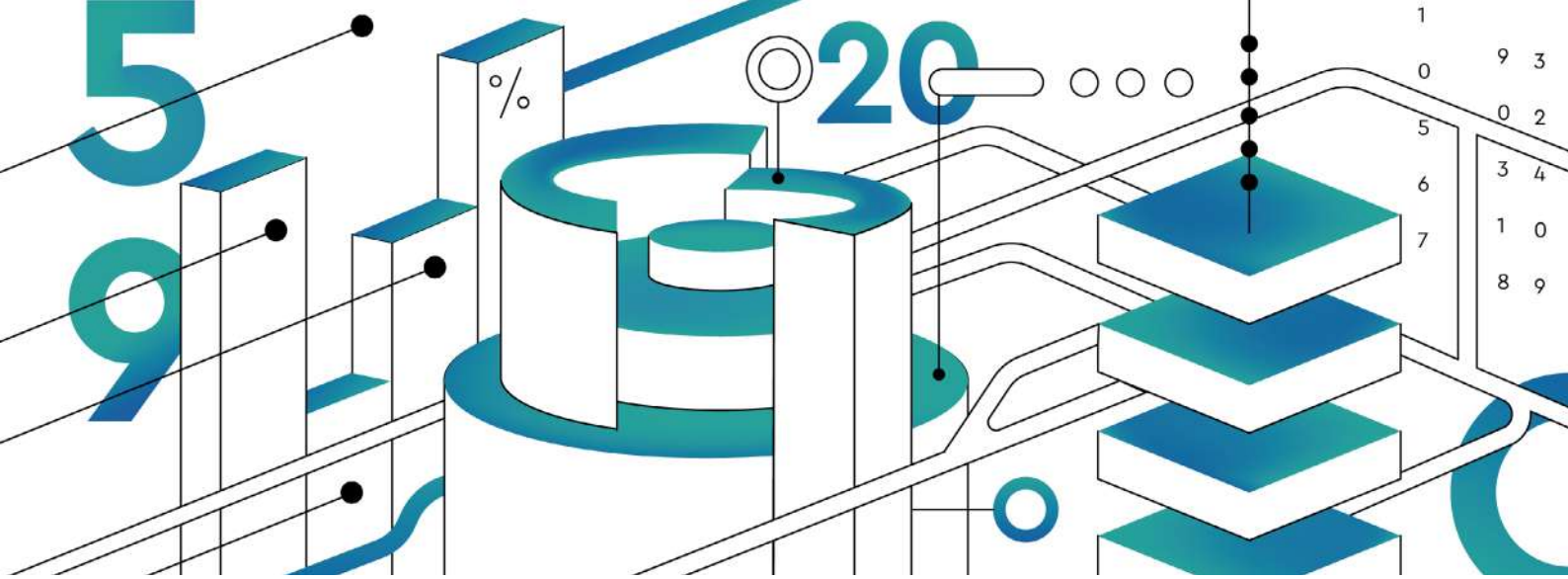
## Появление и развитие новых угроз кибербезопасности неразрывно связано с развитием технологий.

Развивающиеся технологии изменяют ландшафт угроз, так как создают новые возможности для атак, или же находят применения как средства атаки или защиты. Поэтому данный анализ рассматривает новые и развивающиеся технологии как основу для прогноза возможных изменений в ландшафте киберугроз.

В качестве источников для прогноза развития технологий были использованы исследования и прогнозы развития технологий из научных публикаций, публикации Gartner, Institute for The Future и других.

Каждая из развивающихся технологий была проанализирована и оценена с точки зрения степени влияния на ландшафт киберугроз и приведен временной промежуток, в который прогнозируется достижение данной технологией зрелости для широкого промышленного применения.

Цель данного анализа – предоставить структурам кибербезопасности в различных компаниях прогноз относительно сроков возникновения новых или существенной трансформации существующих угроз для своевременной выработки методов противодействия им.



# Методология

В отчете приводится анализ развивающихся технологий с точки зрения их влияния на ландшафт угроз кибербезопасность.

**Оценка влияния на ландшафт угроз имеет 3 градации:**

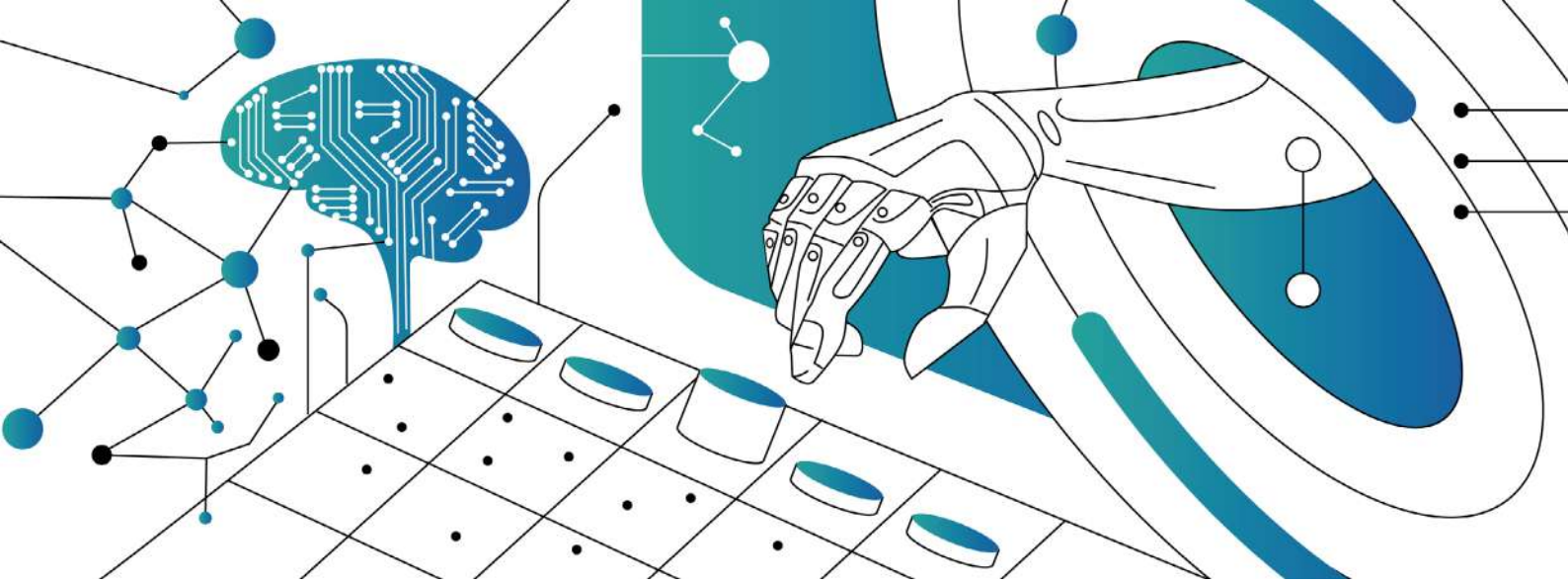
1 **Высокое** – технология существенно влияет на ландшафт угроз, создает возможности для реализации принципиально новых атак, создает новые поверхности атаки или существенно изменяет существующие. Также влияние оценивается, как высокое, если технология может быть использована для защиты от киберугроз и имеет существенное преимущество по сравнению с уже используемыми методами защиты.

2 **Среднее** – технология влияет на ландшафт угроз, не создает возможностей для принципиально новых атак или методов защиты, но при этом позволяет значимо повысить эффективность существующих методов атаки или защиты.

3 **Низкое** – технология практически не влияет на ландшафт угроз, не порождает новых угроз или методов защиты. Обеспечивает лишь небольшой рост эффективности атак или методов защиты от них.

Помимо этого, рассмотренные технологии отнесены к одному из трех временных промежутков, в зависимости от того, когда прогнозируется достижение пика зрелости технологии, делающее возможным ее широкое промышленное применение в различных сферах.

При этом важно отметить, что прогнозирование сроков возникновения тех или иных технологий не является целью данного отчета. В этом вопросе отчет опирается на прогнозы известных аналитических агентств, таких как Gartner, а также на прогнозы в научных и аналитических публикациях.



# Прогноз влияния перспективных технологий на ландшафт угроз кибербезопасности

График, иллюстрирующий прогноз с учетом описанных ранее параметров, представлен далее.

# Hypervector – график влияния перспективных технологий на ландшафт угроз кибербезопасности

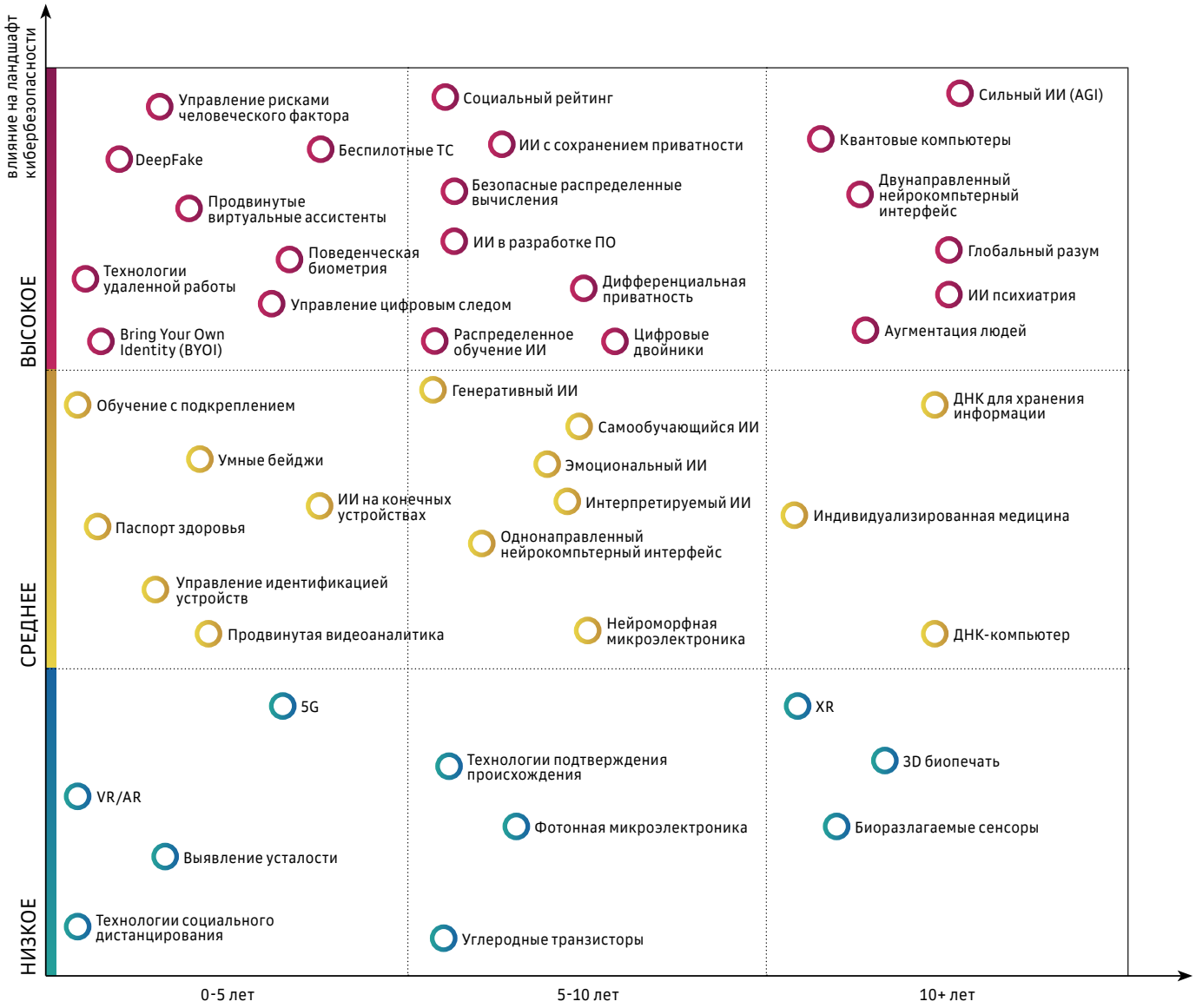
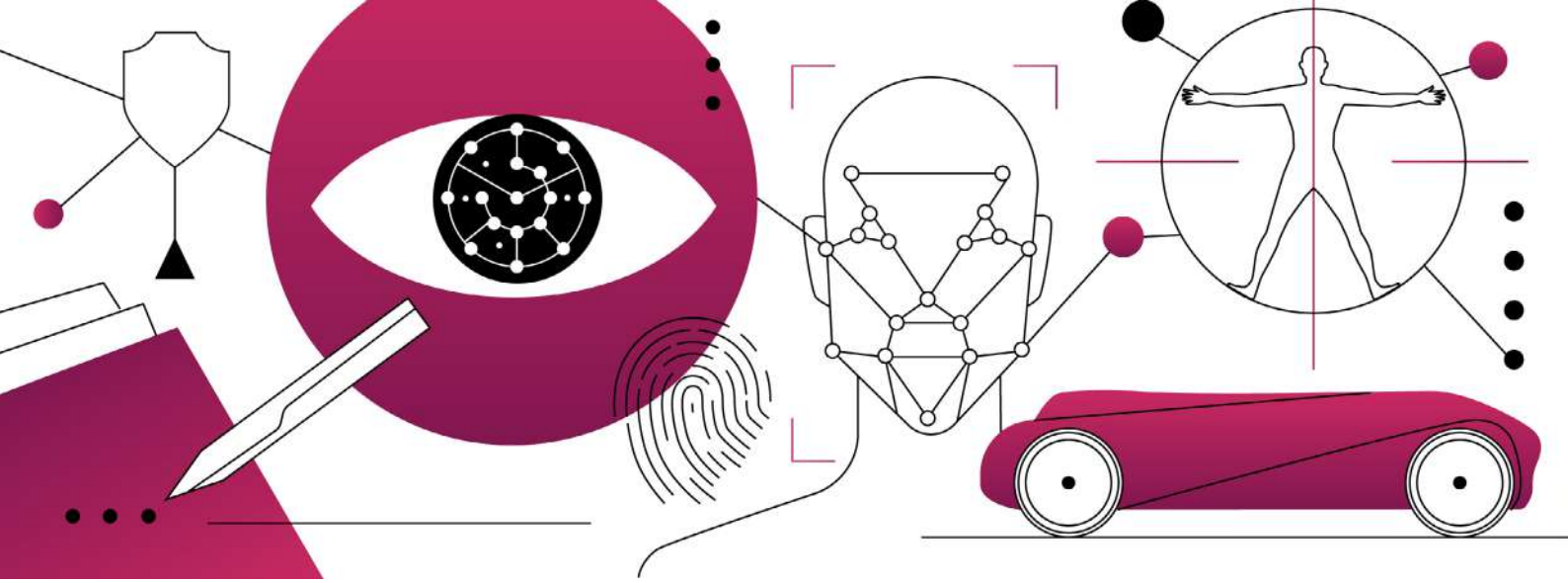



*Рисунок 1.* Hypervector - итоговый прогноз влияния технологий на ландшафт киберугроз

Ось X на графике – время в годах, разделенное на три промежутка: от 0 до 5 лет, от 5 до 10 лет, и более 10 лет. Ось Y – влияние на ландшафт кибербезопасности, также разделенное на 3 промежутка: низкое, среднее и высокое.

Далее отображенные на графике технологии рассмотрены подробнее.



# Ближний горизонт — 1-5 лет

## Технологии с высоким уровнем влияния на ландшафт киберугроз

**1.** **Управление рисками человеческого фактора**
*(Human Risk Management)*

Согласно исследованию Стэнфордского университета и компании Tessian причиной 88% кибератак и утечек данных является человеческий фактор[1].Похожие цифры (85%) называет и Verizon в своем 2021 Data Breach Investigation Report[2]. При этом данные показатели не меняются на протяжении многих лет, несмотря на значительные усилия и ресурсы, затрачиваемые на повышение осведомленности сотрудников об угрозах кибербезопасности. Именно поэтому текущий рынок решений класса Security Awareness начинает постепенный переход от предложения своим клиентам тренингов для сотрудников и фишинговых учений к комплексному решению

по управлению рисками человеческого фактора – Human Risk Management Platform. Определение и составные элементы подобного решения пока не сформированы окончательно и каждый вендор видит их несколько по-своему, но в целом важными элементами подобной платформы видятся следующие:

- Оценка осведомленности сотрудника до начала обучения и выявление «слабых мест» для формирования индивидуальной программы обучения.
- Регулярное персонализированное обучение небольшими порциями информации (micro learning) – отдельный сеанс обучения не должен занимать более 5 минут времени, при этом само обучение должно происходить часто, до нескольких раз в неделю.
- Оценка подверженности тем или иным рискам по результатам обучения, тестов и различных симуляций (не только фишинг, но и вишинг, смишинг и другие виды атак) – по результатам такой оценки для каждого сотрудника строится и регулярно обновляется профиль подверженности различным типам рисков кибербезопасности.
- Использование принципов поведенческой психологии для формирования кибербезопасного поведения, а не только знаний о нем. В том числе важным видится использование теории подталкивания (nudge theory)[3] для уведомления пользователя о возможном риске в момент его возникновения.
- Оценка вероятности возникновения того или иного риска кибербезопасности для сотрудника в зависимости от профиля и характера его рабочей деятельности.
- Использование элементов и механик геймификации для вовлечения пользователя в обучающий процесс и поддержания интереса.
- Подобное решение позволит комплексно оценивать риски человеческого фактора на уровне каждого отдельного сотрудника и формировать у сотрудников устойчивые поведенческие паттерны, направленные на минимизацию этих рисков.

[1] Psychology of Human Error 2022 Research Report | Tessian

[2] 2022 Data Breach Investigations Report | Verizon

[3] Nudge theory Wikipedia

## 2. DeepFake

Методика синтеза изображения, видео, аудио или другой информации с помощью технологии искусственного интеллекта. Например, обучив нейронную сеть на фото или видео лица человека, можно создать крайне реалистичное видео с участием этого человека, где он говорит то, что в реальности никогда не говорил или совершает действия, которые никогда не совершал. Изначально данная технология использовалась для генерации развлекательного контента, но на сегодняшний день уже зафиксирован целый ряд случаев, где дипфейки используются для мошенничества и других противоправных действий[4,5]. Помимо этого Лабораторией кибербезопасности был проведен ряд исследований, демонстрирующих возможность обхода систем лицевой и голосовой биометрической аутентификации с использованием видео и аудио дипфейков. По мере развития технологии и появления простых в использовании инструментов для создания дипфейков количество подобных случаев будет только расти, в связи с чем влияние данной технологии на ландшафт киберугроз оценивается как высокое.

[4] Tinkov's doppelganger invites to a fake site (deepfakechallenge.com)

[5] European MPs targeted by deepfake video calls imitating Russian opposition | The Guardian

## 3. Продвинутые виртуальные ассистенты *(Advanced Virtual Assistants)*

Данная технология является естественным развитием технологии виртуальных ассистентов и чат-ботов, уже широко используемые сегодня и подразумевает использование технологий искусственного интеллекта, в частности обработки естественного языка (Natural Language Processing) и глубоких нейросетей для понимания смысла запросов пользователей, ведения диалога, формирования прогнозов и помощи в принятии решений. Продвинутый виртуальный ассистент, по определению Gartner, также должен включать возможность мультимодального взаимодействия с пользователями (голос, текст, графических интерфейс и т. д.). Уже сейчас виртуальные ассистенты Салют используются в большом количестве приложений как для клиентов, так и для внутреннего пользователя. По мере развития их возможностей, для которых потребуется доступ к различным системам и информации, масштаб угроз будет расти и виртуальный ассистент может стать важным вектором атаки на человека или информационную систему.

Потенциальное влияние данной технологии на ландшафт киберугроз оценивается как высокое в силу широких возможностей применения в различных сферах деятельности и бизнес-процессах, что открывает новые возможности для атаки на бизнес и людей. Помимо этого, продвинутые виртуальные ассистенты могут быть использованы и в интересах кибербезопасности как удобный способ взаимодействия с сервисами и как канал оперативного информирования о состоянии различных аспектов кибербезопасности.

## 4. Поведенческая биометрия *(Behavioral Biometrics)*

Технологии биометрической аутентификации, использующие в качестве фактора поведенческие характеристики человека, например голос, походку, рукописный или клавиатурный почерк. В отличие от статической биометрии данные факторы могут оцениваться не только при входе в систему или приложение, но и на протяжении всего сеанса взаимодействия и при этом не требуют от пользователя совершения дополнительных действий для аутентификации. Система может непрерывно оценивать соответствие поведения пользователя поведенческому биометрическому слепку и в случае существенного отклонения сигнализирует о возможной компрометации службе безопасности и/или блокирует доступ, что делает проблематичным перехват сессии пользователя злоумышленником. Подделка или компрометация поведенческих факторов и их использование без участия владельца на сегодняшний день представляется крайне сложной задачей. На данный момент промышленное применение поведенческой биометрии еще не получило широкого распространения, и ее используют в отдельных случаях, когда нужны дополнительные уровни аутентификации, например, при проведении крупных транзакций или допуске к особо важным данным, но технология развивается и в ближайшем будущем такие методы смогут значительно сократить риски в сфере безопасности.

## 5. Технологии удаленной работы *(Remote Work Solutions)*

Из-за пандемии COVID-19 существенно выросло количество сотрудников, работающих удаленно, что сильно стимулирует развитие решений и технологий удаленной работы. При этом существенно меняется ландшафт угроз кибербезопасности, так как сотрудник уже не находится в пределах периметра с многослойной защитой, но при этом для выполнения своих рабочих обязанностей ему, как правило, требуется доступ к корпоративным системам и данным разной степени конфиденциальности. Скорее всего развитие тренда на удаленную работу также подстегнет кибербезопасность компаний к переходу от модели защиты периметра, внутри которого все доверяют всем, к модели нулевого доверия (Zero Trust), где доверие конкретному пользователю и устройству оценивается каждый раз при инициации соединения и на протяжении всего сеанса взаимодействия. Помимо этого, возрастает риск атаки с использованием имперсонации при удаленном взаимодействии. Поэтому влияние данного набора технологий на ландшафт киберугроз оценивается как высокое.

## 6. **Bring Your Own Identity** *(BYOI)*

Концепция, упрощающая регистрацию и вход пользователей в новые сервисы, за счет аутентификации через другой сервис, где пользователь уже зарегистрирован (например, Login with Google, Facebook, Apple ID, Сбер ID). С одной стороны такой подход существенно упрощает жизнь пользователям и улучшает пользовательский опыт, а также существенно уменьшает поверхность атаки на пользовательские данные, которые хранятся централизованно у одного identity провайдера (IdP), а не в сотнях различных сервисов. Но, с другой стороны, IdP превращается в единую точку отказа, и компрометация IdP дает злоумышленнику доступ к сразу десяткам, если не сотням, а в будущем, возможно, даже тысячам различных сервисов, вход в которые реализован через данного IdP (кража данных, кража личности, имперсонация). Также данный подход открывает возможность для кражи данных пользователя через мошеннический сервис или фишинговую атаку. Помимо этого, использование BYOI может упростить процесс создания реалистичных синтетических личностей с цифровым присутствием в множестве сервисов. По причинам, обозначенным выше, влияние данной технологии на ландшафт угроз кибербезопасности оценивается как высокое.

## 7. **Беспилотные транспортные средства** *(Autonomous Vehicles)*

Транспортные средства, управляемые искусственным интеллектом без участия живого водителя, уже давно тестируются на дорогах общего пользования, а в ряде стран и регионов, в том числе и в некоторых регионах России, уже используются в режиме пилотной коммерческой эксплуатации. При этом массовое внедрение БПТС в коммерческую эксплуатацию существенно влияет на ландшафт угроз и открывает злоумышленникам новые возможности для атаки, например:

- Взлом БПТС с целью угона, совершения преступлений или терактов – «ИИ-камикадзе»
- Получения несанкционированного доступа к системам управления интеллектуальной транспортной инфраструктурой или внесение нарушений в ее работу может приводить к печальным последствиям - ДТП, нарушение работы общественного транспорта, нарушение регулирования дорожного движения вплоть до полного транспортного коллапса, совершение терактов.
- Также в связи с распространением применения беспилотных транспортных средств в военных целях, возникает угроза их компрометации или перехвата. Это несет в себе большую угрозу, чем атаки на гражданские беспилотники, так как злоумышленник получает доступ в том числе и к вооружению беспилотника.

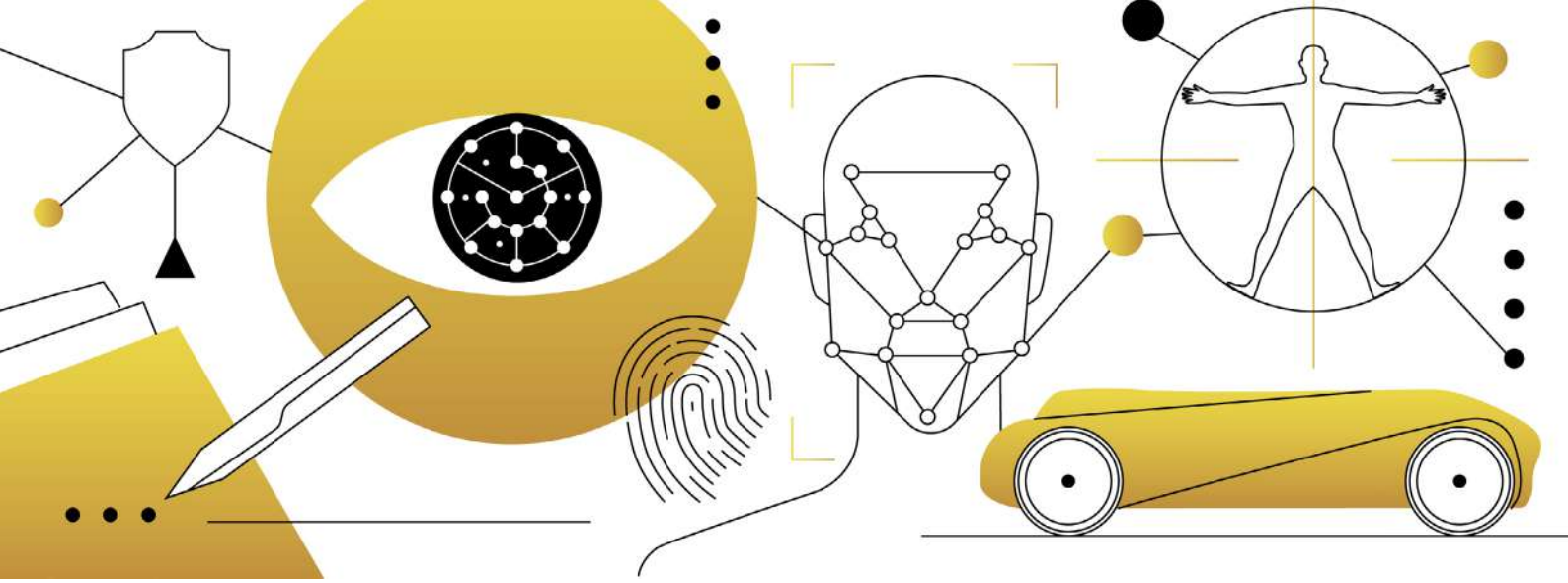
В свете вышеперечисленного влияние технологий беспилотных ТС на ландшафт кибербезопасности оценивается как высокое.

## 8. **Управление цифровым следом** *(Digital Footprint Management)*

Количество времени, проводимого человеком онлайн, как и среднее количество используемых им онлайн-сервисов стабильно растет из года в год и, скорее всего, продолжит расти и далее. Каждый пользователь оставляет все больше и больше различной информации о себе, своих действиях и предпочтениях в посещаемых сервисах. При этом среднестатистический пользователь зачастую даже не представляет какая информация о нем и в каком объеме хранится в том или ином сервисе. Более того, существенная часть этой информации может собираться и использоваться без его ведома и информированного согласия. Это открывает огромные возможности для злоупотребления этой информацией как со стороны недобросовестных сервисов, так и со стороны злоумышленников, получивших несанкционированный доступ к ней. Что в свою очередь упрощает проведение атак на человека с использованием методов социальной инженерии, так как позволяет быстро завоевать доверие пользователя при атаке. Даже при наличии возможности управлять информацией о себе в отдельно взятом сервисе, общее количество используемых сервисов и разрозненность хранимой информации делает эту задачу крайне сложной. Что создает необходимость в разработке решений по управлению цифровым следом пользователя в сети Интернет, которые могут помочь в решении следующих задач:

- Контроль и ограничение сбора информации о пользователе со стороны различных сервисов
- Централизованное хранение согласий на обработку информации о пользователе от различных сервисов с возможностью их отзыва
- Выявление источника утечки информации и своевременное уведомление пользователя о факте утечки

Влияние данной технологии на ландшафт угроз кибербезопасности оценивается как высокое.



# Ближний горизонт — 1-5 лет

## Технологии со средним уровнем влияния на ландшафт киберугроз

**1.** **Обучение с подкреплением** *(Reinforcement Learning)*

Один из способов машинного обучения, в ходе которого испытуемая система (агент) обучается, взаимодействуя с некоторой средой. Обратная связь от среды на действия агента формирует его поведение, итеративно обучая его оптимальным вариантам взаимодействия. Сама концепция была сформулирована в 1960-х годах, но развитие вычислительных мощностей и применение нейросетей существенно расширили возможности обучения с подкреплением, которое в наше время переживает «второе рождение». Данная технология может быть использована в интересах кибербезопасности для обучения интеллектуальных агентов на выполнение различных задач по реагированию и расследованию инцидентов кибербезопасности, выявлению мошеннических транзакций и прочих. Но не стоит забывать, что обучение с подкреплением также может быть взято на вооружение злоумышленниками для проведения кибератак и их легкого масштабирования.

## 2. Умные бейджи *(Smart Badges)*

Бейджи с встроенным вычислительным и запоминающим устройством для реализации дополнительного функционала, например, запись, транскрибация и диаризация диалога с клиентом, видеозапись, геолокация, контроль выполнения рабочих обязанностей, связь с коллегами, получение инструкций и задач от руководителя. Технология умных бейджей уже развивается силами ряда технологических компаний[6]. В зависимости от реализованных на бейдже модулей и их функционала возможны различные угрозы: кража или подмена данных (голос, видео), несанкционированный доступ к устройству для прослушки или отслеживания передвижения, имперсонация руководителя.

[6] «Сбер» представил SmartBadge — умный бейдж с дисплеем и микрофоном / Хабр (habr.com)

## 3. Паспорт здоровья *(Health Passport)*

Данная технология позволяет в цифровом виде безопасно хранить и предъявлять по требованию подтвержденную информацию об иммунном статусе предъявителя в отношении различных заболеваний, а также другую информацию о состоянии здоровья, с возможностью ее верификации и защитой от несанкционированной модификации. Развитие данной технологии существенно ускорила эпидемия COVID-19. Уже сейчас различные варианты паспорта здоровья используются во многих странах для подтверждения иммунного статуса человека при получении доступа к различным мероприятиям и услугам, сопряженным с физическим контактом с другими людьми.
Компрометация паспорта здоровья может приводить к различным нежелательным последствиям: подделка паспорта для получения привилегий, кража персональных данных (массовая - при централизованном хранении), неправомерный доступ к конфиденциальным данным, нарушение паспорта для отказа в привилегиях.

## 4. Искусственный интеллект на конечных устройствах *(AI at Edge)*

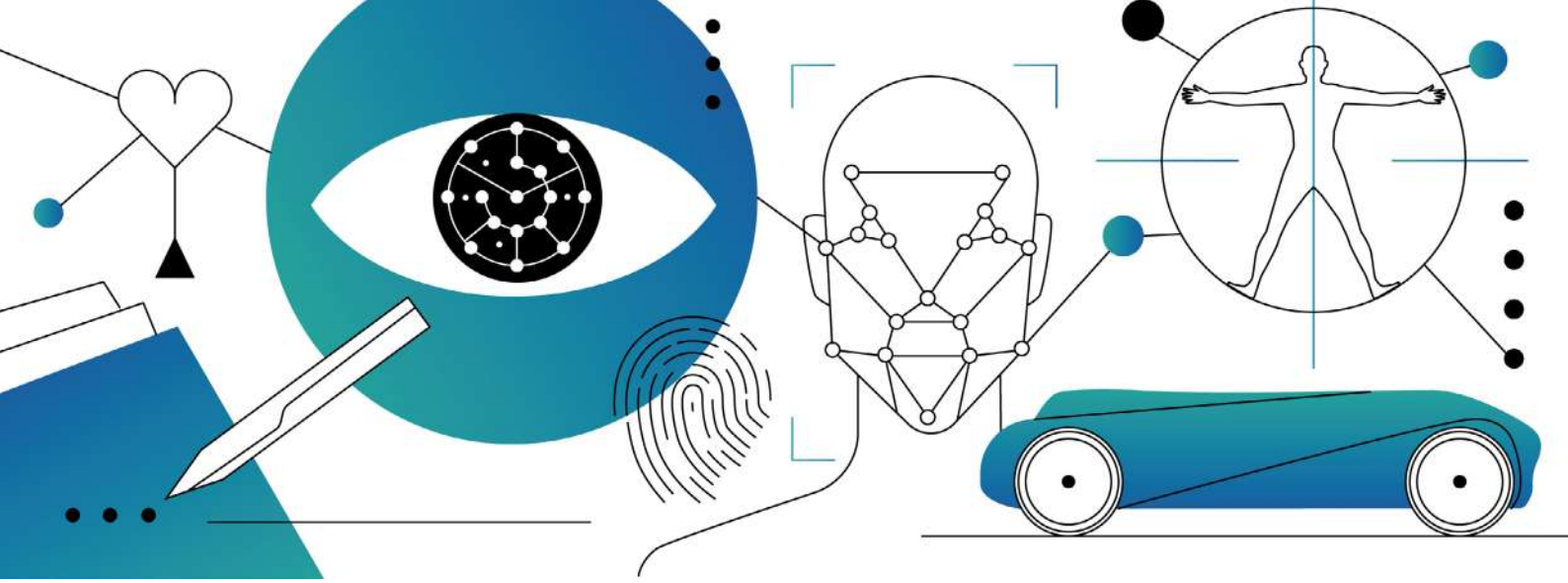
Развитие технологий искусственного интеллекта и рост вычислительной мощности конечных устройств позволяет производить все больше вычислений непосредственно на конечном устройстве, что позволяет увеличить скорость работы и уменьшить задержки. Соответственно, поверхность атаки на инфраструктуру перераспределяется в сторону от более защищенного ядра к более доступным для атак конечным устройствам. Помимо этого, миниатюрные устройства с производительным вычислителем на борту позволяют упростить и повысить эффективность слежки, так как ряд интеллектуальных функций может быть реализован на самом устройстве, например, распознавание лиц и голосов, распознавание и транскрибация речи.

## 5. Управление идентификацией устройств *(Machine Identity Management)*

Данный класс технологий подразумевает управления секретами и доступами устройств, приложений и сервисов. Распространение устройств, приложений и сервисов, взаимодействующих друг с другом без непосредственного участия людей, требует реализации функций управления секретами, доступами и полномочиями, защиты данных и самих устройств, и приложений, по аналогии с тем, как это реализовано при взаимодействии человека с устройствами, приложениями и сервисами. При этом угрозы, реализуемые от эксплуатации или кражи «цифровой личности» устройства, по большому счету такие же, как и в случае с человеком, в то время как признаки взлома или эксплуатации могут быть заметны человеку, но не будут «заметны» устройству или приложению.

## 6. Продвинутая видео аналитика *(Advanced Video Analytics)*

Технология, использующая методы компьютерного зрения для автоматизированного получения различных данных на основании анализа последовательности изображений, поступающих с видеокамер в режиме реального времени или из архивных записей. Развитие технологий искусственного интеллекта существенно упростило процесс автоматизированного анализа видеопотока и расширило его возможности. Уже сейчас технологии продвинутой видеоаналитики позволяют автоматически выявлять различные инциденты на видео, осуществлять быстрый поиск в обширном архиве записей нужных людей или событий по различным критериям (пол, возраст, цвет одежды, волос, кожи человека, совершаемые действия или движение, наличие определенных предметов или объектов и прочее).
Данные технологии существенно упрощают расследование инцидентов и преступлений сотрудниками физической безопасности и правоохранительных органов, но также могут быть легко использованы злоумышленником для быстрого сбора информации при планировании атаки в случае компрометации системы видеоаналитики.



# Ближний горизонт — 1-5 лет

## Технологии с низким уровнем влияния на ландшафт киберугроз

### 1. Сети мобильной связи пятого поколения *(5G)*

Сети мобильной связи пятого поколения, активно внедряемые в ряде стран, обещают существенное увеличение скорости подключения и сокращение задержек при пересылке данных. Ряд компаний (аэропорты, крупные промышленные предприятия) рассматривают возможность использования приватных 5G сетей на своих объектах для обеспечения оперативной связи между сотрудниками, а также управления и автоматизации производственного процесса. В то же время, с точки зрения кибербезопасности рост скорости подключения и количества подключенных устройств дает злоумышленникам возможность создавать более масштабные бот-сети в том числе и из 5G IOT-устройств, а также реализовывать более масштабные DDoS-атаки. Уже сейчас существуют масштабные ботсети из IoT-устройств (например, одна из

самых известных сетей – Mirai)[7], а доступность и скорость 5G только увеличит данную угрозу. Отдельно стоит отметить, что для обеспечения обратной совместимости в сетях 5G могут использоваться устаревшие протоколы с давно известными уязвимостями, что может упростить работу злоумышленников по получению несанкционированного доступа к инфраструктуре сети.

[7] Understanding the Mirai Botnet (usenix.org)

## 2. Виртуальная и дополненная реальность *(VR/AR)*

Виртуальная реальность подразумевает полное замещение реального окружающего пространства синтезированным, в то время как дополненная реальность – комбинацию объектов реального и виртуального мира. Решения виртуальной и дополненной реальности в настоящее время нашли широкое применение в сфере развлечений. Также возможность их применения активно исследуется в сфере образования, промышленного производства, медицины и ряде других. Также AR/VR рассматривается как один из основных вариантов взаимодействия в метавселенной. Данная технология может быть применена в кибербезопасности для визуализации и анализа большого количества данных в трехмерном пространстве, которые сложно визуализировать на двухмерном экране монитора. Также технологии виртуальной реальности могут быть использованы для работы с конфиденциальными данными, снижая риск утечки через визуальный канал, например, как описывается в патенте Apple – 3d document editing system[8].

[8] US20180081519A1 - 3d document editing system Google Patents

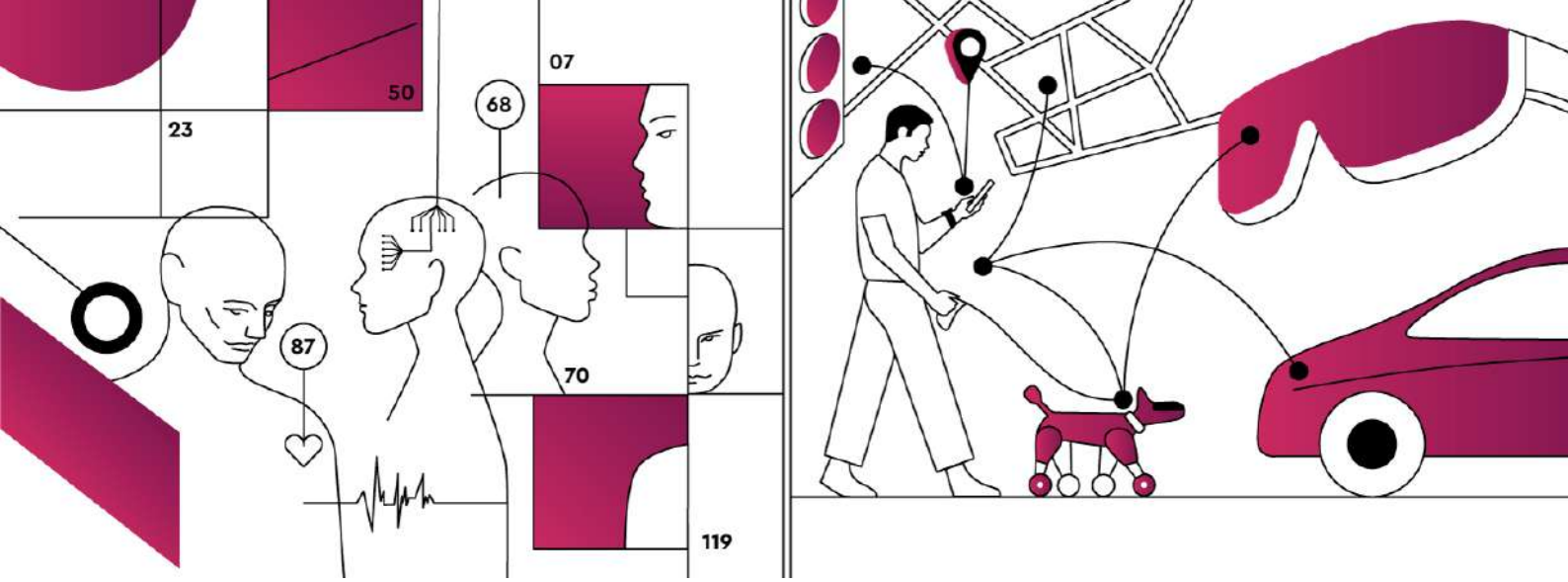
## 3. Выявление усталости *(Drowsiness Detection)*

Выявление усталости, как правило реализуется за счет мониторинга физиологических показателей и поведенческих маркеров усталости/сонливости. Может применяться на производствах и в процессах, где усталость или невнимательность работника несет большие риски для безопасности или производственного процесса. Основной угрозой, реализуемой через эксплуатацию данной технологии, является умышленный или непреднамеренный саботаж производственного процесса (например, атака на модель-классификатор, подмена данных на входе или на выходе). Атака может реализовываться сторонним злоумышленником (с целью получения выгоды или саботажа) или даже самим работником (например, с целью максимизации заработка при почасовой оплате за счет более продолжительного периода работы и сокращения перерывов на отдых). Также данная технология может найти применение и в работе отделов и ситуационных центров по кибербзеопасности, т.к. внимательность аналитиков является важным фактором в расследовании и противодействии инцидентам.

## 4. Технологии социального дистанцирования
*(Social Distancing Technologies)*

Для минимизации рисков заражения в период пандемии COVID-19 государственные органы и коммерческие организации по всему миру стремятся обеспечить соблюдение социальной дистанции между людьми. Для помощи в этой задаче возник целый спектр решений, использующий различные подходы: мобильные приложения, Bluetooth, носимые устройства, WiFi или видео аналитику. На данном этапе подобные решения показывают невысокую эффективность, но технологии развиваются дальше, чтобы закрыть растущую потребность. При этом использование подобных технологий создает целый ряд рисков приватности и кибербезопасности в случае их компрометации: раскрытие персональных данных, несанкционированная слежка, сбор информации для подготовки к атаке и прочее. На данный момент в силу того, что пандемия COVID-19 пошла на спад, и ограничения повсеместно снимаются, востребованность технологий социального дистанцирования снижается. Но вспышки новых заболеваний могут возродить к ним интерес[9], уже сегодня прогнозируются новые волны COVID-19, а также фиксируется распространение вируса обезьяньей оспы. Помимо этого, разработанные для социального дистанцирования методы и подходы могут быть востребованы в других сферах деятельности, в том числе и в военных целях (например, в системах свой-чужой).

[9] Bill Gates Just Said We Have a 50 Percent Chance of Another Pandemic in the Next 20 Years. Here's How He Thinks We Should Prepare | Inc.com




# Средний горизонт — 5-10 лет

## Технологии с высоким уровнем влияния на ландшафт киберугроз

**1.** **Социальный рейтинг** *(Social Rating)*

Уже сейчас в отдельных странах мы видим попытки внедрения интегрального цифрового рейтинга гражданина, на основании которого будет приниматься решение о доступе или отказе в услугах, определяться права и полномочия гражданина. Если этот тренд будет развиваться и дальше, то нам потребуются механизмы обеспечения безопасности и целостности этого рейтинга, так как манипуляция с рейтингом или данными, на базе которых он строится, будут приводить к ряду нежелательных последствий: ограничение доступа к отдельным функциям/инструментам/услугам добропорядочным гражданам или же необоснованное их получение злоумышленником.

## 2. Искусственный интеллект с сохранением приватности *(Privacy-preserving AI)*

Комплекс технологий, обеспечивающий сохранение приватности данных, задействованных в жизненном цикле модели ИИ, а именно:

- Приватность данных для обучения модели – из модели невозможно восстановить данные, использовавшиеся для ее обучения.
- Приватность данных на входе – пользовательские данные, подаваемые на вход модели не раскрываются другим лицам, включая создателя модели.
- Приватность данных на выходе – вывод модели не доступен для наблюдения никому, кроме пользователя, на чьих данных он рассчитан.
- Приватность модели – модель не может быть украдена и использована злоумышленником

Конкретные подходы, реализующие концепцию Privacy-preserving AI еще не финализированы, но активно исследуются научным сообществом.

Реализация искусственного интеллекта с сохранением приватности позволит расширить рынок услуг по обучению и предоставлению моделей ИИ, а также позволит компаниям совместно обучать и эксплуатировать эффективные модели ИИ без нарушения приватности используемых для этого данных и позволит уменьшить сопряженные с этими данными риски кибербезопасности. Поэтому влияние на ландшафт угроз кибербезопасности оценивается как высокая.

## 3. Безопасные распределенные вычисления *(Secure Multi-party Computation, SMPC)*

Ряд криптографических подходов, позволяющих нескольким участникам произвести совместные вычисления, зависящие от тайных входных данных каждого из них, таким образом, чтобы ни один участник не смог получить никакой информации о чужих тайных входных данных. В том числе данный подход не раскрывает другим участникам промежуточные результаты вычислений, что обеспечивает большую безопасность по сравнению с технологией распределенного обучения моделей ИИ. SMPC может быть использован для коммерческого обмена данными без риска их раскрытия или нарушения их приватности.  В том числе данный подход может быть использован для безопасного обмена данными кибербезопасности и антифрода с соответствующими службами других участников рынка или гос. органами без раскрытия персональных или конфиденциальных данных, например, списки мошенников, или индикаторы компрометации.

## 4. Искусственный интеллект в разработке ПО *(AI-augmented Software Development)*

Уже сейчас исследуются возможности применения ИИ в разработке программных продуктов[10], и согласно прогнозам в недалеком будущем ИИ сможет существенно упростить эту задачу, помогая человеку разрабатывать еще более сложные программные продукты, обеспечивать отсутствие программных ошибок и уязвимостей, оптимизировать эффективность ПО и ускорять процесс разработки. С точки зрения безопасности это порождает ряд проблем. Воздействие на этот ИИ может приводить к внедрению уязвимостей или незадекларированных возможностей в код (например, бэкдоры), нарушению функциональности ПО, краже исходного кода и обрабатываемых им данных.

[10] GitHub Copilot · Your AI pair programmer

## 5. Распределенное обучение моделей ИИ *(Federated Learning)*

Технология машинного обучения, обеспечивающая децентрализованное обучение модели на множестве устройств, хранящих данные для обучения локально без обмена исходными данными между устройствами. По сравнению с SMPC данный подход позволяет участвовать в процессе практически неограниченному количеству участников и, как правило, демонстрирует более высокую производительность, но при этом обладает меньшими гарантиями безопасности данных, так как другим устройствам могут быть доступны промежуточные результаты обучения, используя которые можно теоретически восстановить исходные данные. Данная технология применима для распределенного обучения единой модели с использованием данных, собираемых множеством конечных устройств. В интересах кибербезопасности, распределенное обучение может быть применено в выявлении мошенничества, для повышения эффективности выявления мошеннических транзакций за счет информации на устройствах пользователей без нарушения конфиденциальности.

## 6. Дифференциальная приватность *(Differential Privacy)*

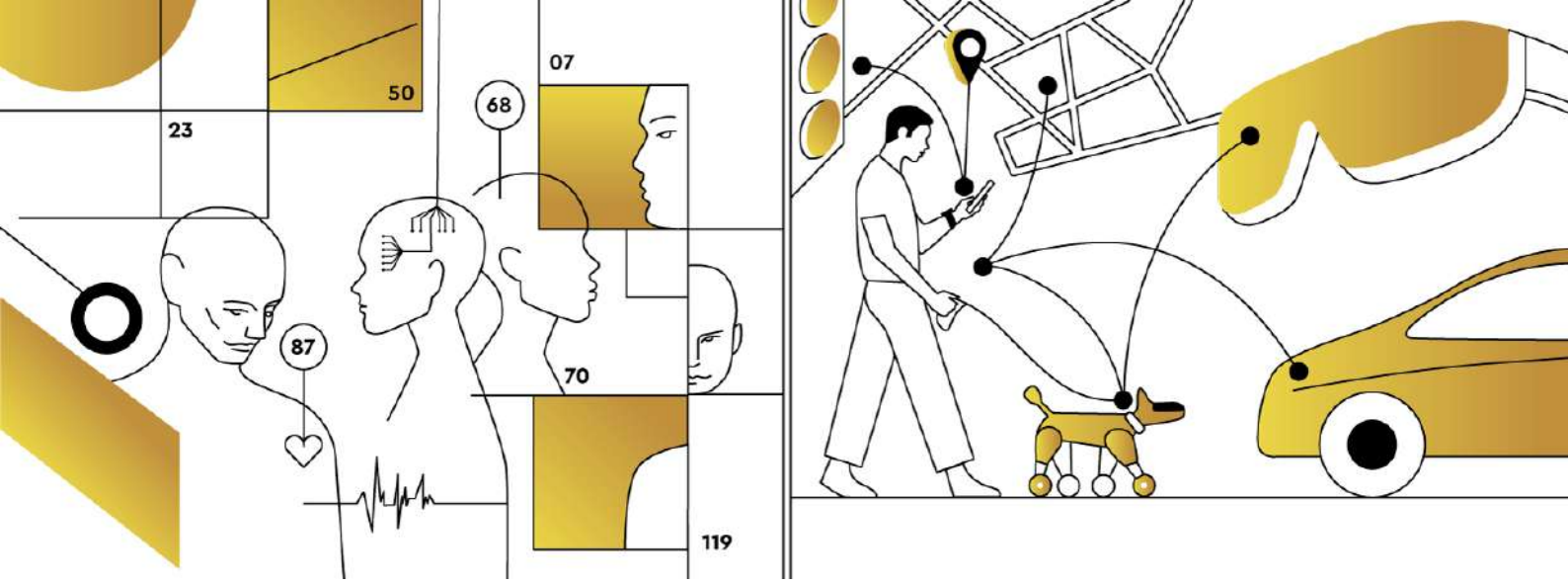
Технология, обеспечивающая максимально точные ответы на статистические и аналитические запросы к данным без возможности раскрытия содержания отдельных записей даже при выполнении большого количества запросов. Дифференциальная приватность основана на введении случайности в данные, без существенного влияния на их статистические характеристики. На данный момент технология довольно нова и ограничено известна за пределами исследовательского сообщества, но активно развивается в связи с растущим спросом на решения по сохранению приватности данных. При этом уже сегодня технологию начинают использовать некоторые технологические компании для анализа данных пользовательских устройств и приложений без нарушения приватности пользователей, например, Apple использует ее для анализа некоторых данных с устройств на базе iOS[11]. Использование данной технологии при коммерческом обмене аналитическими и статистическими данными между компаниями или между компаниями и конечными пользователями их продуктов позволит снизить риски раскрытия конфиденциальной информации или персональных данных.

[11] Differential Privacy Overview (apple.com)

## 7. Цифровые двойники *(Digital Twins)*

Виртуальный аналог реального объекта, компьютерная модель, которая в своих ключевых характеристиках дублирует его и способна воспроизводить его состояния при разных условиях. Данная технология изначально возникла как часть четвертой промышленной революции для оцифровки физических объектов или процессов в производстве с целью помочь предприятиям быстрее обнаруживать физические проблемы, точнее предсказывать результаты и производить более качественных продукты[12]. Дальнейшее развитие технологии цифровых двойников позволит оцифровывать более сложные объекты, в том числе живые организмы и другие сложные динамические системы. Например, для фармацевтической индустрии в сфере разработки лекарственных препаратов, цифровой двойник может симулировать сложный комплекс физических и биохимических взаимодействий в организме человека для оценки действия новых лекарств. В сфере управления государством и охраны правопорядка, цифровой город, населенный большим количеством цифровых двойников, поможет прогнозировать реакцию общества на те или иные решения и законы, массовые беспорядки или стихийные бедствия. Для человека персональный цифровой двойник с искусственным интеллектом может быть незаменимым ассистентом, как в физическом мире, так и в метаверсе. С теми же эмоциями, поведением и паттерном действий и принятия решений, собранными на базе цифрового следа “оригинала” он может брать на себя выполнение части задач, так как с точки зрения цифровых взаимодействий с внешним миром такая реплика будет неотличима от оригинала. И хотя технологии для создания полноценного и полнофункционального цифрового двойника кажутся фантастикой, уже сейчас собирается огромное количество данных о цифровых взаимодействиях реальных людей, которые послужат “строительным материалом” для создания цифрового двойника, когда технологии позволят их использовать. С точки зрения безопасности основными угрозами для данной технологии может быть кража, злонамеренная модификация или вывод из строя цифровых двойников. Также цифровой двойник может обеспечивать псевдоудаленное присутствие пользователя на различных встречах и публичных мероприятиях, что порождает дополнительные проблемы с подтверждением аутентичности и права представительства цифрового двойника (“прислан” ли цифровой двойник самим владельцем оригинальной личности, уполномочен ли он владельцем на те слова и действия, которые он произносит или производит). Соответственно, с точки зрения безопасности потребуется ряд инструментов и технологии для контроля аутентичности, целостности и права представительства цифрового двойника.

[12] Industry 4.0 and the digital twin technology | Deloitte Insights




# Средний горизонт — 5-10 лет

## Технологии со средним уровнем влияния на ландшафт киберугроз

**1.** **Генеративный искусственный интеллект** *(Generative AI)*

Технология, позволяющая генерировать новый цифровой контент на основании анализа существующих фото, видео, аудио, текстов или других типов данных. Генеративный ИИ может применяться в различных сферах от развлечений, искусства и электронной коммерции до биологии, медицины и инженерного дела. Уже сейчас различные компании экспериментируют с применением этой технологии в своей деятельности, например, в фармацевтике для генерации новых химических соединений с заданными свойствами или в электронной коммерции для виртуальной примерки одежды или макияжа. В части влияния на ландшафт угроз кибербезопасности, генеративный ИИ может быть использован для

генерации синтетических данных, используемых для обучения других моделей искусственного интеллекта без риска нарушения приватности. Также данная технология уже находит применение у злоумышленников, например, для генерации дипфейков, или создания онлайн-профилей синтетических личностей для дальнейшего мошенничества.

## 2. **Самообучение ИИ** *(Self-supervised Learning)*

Технология машинного обучения на неразмеченных данных, где модель ИИ сама обнаруживает закономерности в структуре данных и обучается на них. В данный момент эта технология чаще всего применяется в задачах обработки естественного языка, где на большом объеме текста модель учится правильно предсказывать слова, идущие перед или после текущего слова. Также самообучение применимо для предсказания расположения объектов или их частей относительно друг друга при анализе изображений. Данный подход имеет большой потенциал применения, в том числе и в задачах кибербезопасности, так как не требует долгой и трудоемкой разметки больших датасетов, выполняемой людьми, в связи с чем самообучении ИИ активно исследуется ведущими технологическими компаниями, например, Meta[13].

[13] Self-supervised learning: The dark matter of intelligence (facebook.com)

## 3. **Эмоциональный искусственный интеллект** *(Affective Computing)*

Комплекс технологий, направленный на распознавание, интерпретацию, обработку и симуляцию человеческих эмоций на основании данных от различных устройств (сенсоров, камер, микрофонов и пр.). Развитие эмоционального ИИ позволит учитывать эмоции пользователя при его взаимодействии с различными устройствами, программами или сервисами и подстраивать результаты взаимодействий под его эмоциональный статус, как бы предугадывая его текущие желания и потребности и создавая у пользователя ощущение комфорта. В сфере безопасности данная технология может быть применена для оценки эмоционального состояния сотрудников с целью прогнозирования возникновения различных рисков, например, инсайдерской угрозы. Стоит учитывать, что эмоциональный ИИ также может быть использован злоумышленниками при проведении атаки с использованием социальной инженерии для оценки эмоционального состояния жертвы и выбора наиболее эффективной стратегии развития атаки.

## 4. **Интерпретируемый искусственный интеллект** *(Interpretable AI)*

Набор процессов и методов, позволяющих человеку понять, почему именно модель ИИ пришла к тем или иным результатам или выводам. По мере расширения возможностей ИИ, людям становится все сложнее осознать, каким образом модель пришла к тому или иному результату. Процесс

вычисления превращается в так называемый «черный ящик», который не поддается интерпретации. Модели, работающие по принципу черного ящика, создаются непосредственно на основе данных. При этом даже создавшие модель специалисты не в состоянии понять и объяснить, что именно в ней происходит и как модель пришла к конкретному результату. Возможность определить, как именно система на основе ИИ получила конкретный вывод, обладает множеством преимуществ. Объяснимость позволяет разработчикам убедиться в том, что система работает правильным образом, может потребоваться для обеспечения соответствия нормативным стандартам, а также может оказаться важной функцией, позволяющей заинтересованным лицам опротестовать или изменить полученный результат. В части кибербезопасности интерпретируемость моделей ИИ может позволить реализовать процесс их аттестации на отсутствие логических уязвимостей, несущих потенциальную угрозу, а также повысить доверие к моделям используемым непосредственно в целях кибербезопасности.

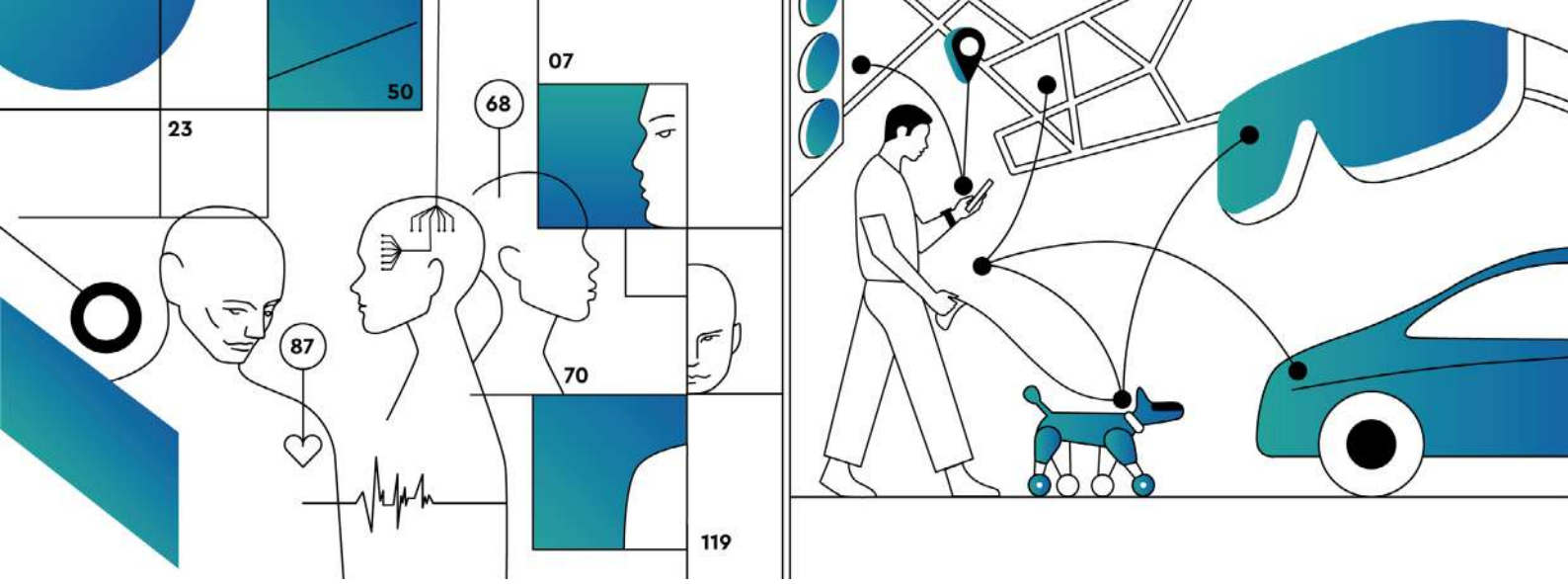
## 5. Однонаправленный нейрокомпьютерный интерфейс *(One-way BCI)*

Система для обмена информации между мозгом и электронным устройством (например, компьютером). В однонаправленных интерфейсах внешние устройства могут либо принимать сигналы от мозга, либо посылать ему сигналы, например, имитируя сетчатку глаза с помощью электронного импланта. Простые неинвазивные нейроинтерфейсы уже сейчас могут применяться для передачи команд компьютерам и другим электронным устройствам. Технология нейроинтерфейсов также нашла свое применение в протезировании, позволяя восстанавливать нарушенные функции сенсорных органов, например, слуха с помощью кохлеарного импланта или зрения с помощью импланта сетчатки. По мере развития технологии прогнозируется расширение использования нейроинтерфейсов не только для восстановления утраченных функций организма, но и для расширения их возможностей сверх доступных обычному человеку. А снижение инвазивности при имплантации и извлечении позволит устанавливать и заменять их без особых сложностей. Широкое распространение нейроинтерфейсов породит новые угрозы кибербезопасности, такие как взлом имплантов с целью слежки или перехвата контроля, похищения конфиденциальных данных, модификация передаваемой информации и прочие.

## 6. Нейроморфная микроэлектроника *(Neuromorphic Hardware)*

Нейроморфные чипы имеют специализированную вычислительную архитектуру, имитирующую морфологию и структуру биологических нейронных сетей: выделенные вычислительные блоки эмулируют поведение нейронов непосредственно в аппаратном обеспечении, а сеть физических взаимосвязей между ними обеспечивает быстрый обмен информацией между ними. Эта концепция вдохновлена человеческим мозгом, где биологические нейроны и синапсы работают аналогичным образом. Специализированные нейроморфные устройства менее гибкие, чем универсальные центральные процессоры, но обеспечивают исключительную производительность и энергоэффективность при обучении и эксплуатации глубоких нейронных

сетей. Обычные компьютеры реализованы на базе архитектуры фон Неймана, которая состоит из вычислительных ядер, последовательно выполняющих инструкции и обрабатывающих данные, хранящиеся в централизованной памяти. Это означает, что вычислительная производительность таких систем ограничена скоростью передачи данных, которые могут быть переданы между вычислительным блоком и внешним блоком памяти. С появлением все более требовательных приложений интерес к высокопроизводительным вычислениям сместился в сторону увеличения параллелизма в виде многоядерных архитектур. Однако возможность распараллеливания вычислений принципиально ограничена общим доступом к ресурсам памяти. Последние достижения в области глубокого обучения упираются в эти ограничения, поскольку высокопараллельная структура глубоких нейронных сетей требует распределенного доступа к памяти, которое сложно эффективно реализовать с помощью обычных вычислительных технологий. Нейроморфная микроэлектроника призвана решить эту проблему и помочь в дальнейшем развитии технологий искусственного интеллекта.
С точки зрения кибербезопасности помимо угроз со стороны возросшей производительности технологии нейронных сетей, новая микропроцессорная архитектура потребует анализа возможных уязвимостей на уровне процессора и, возможно, выработки ряда новых методов защиты и противодействия подобным угрозам.




# Средний горизонт — 5-10 лет

## Технологии с низким уровнем влияния на ландшафт киберугроз

**1.** **Фотонная микроэлектроника** *(Photonic Computing)*

Принципиально другой способ проведения вычислений с применением фотонов, а не электронов. Данный метод позволяет на несколько порядков повысить скорость вычислений с одновременным снижением энергопотребления. В теории на базе фотоники можно создавать как классические, так и квантовые компьютеры и классический фотонный компьютер может быть логичным переходным этапом между классическим и настоящим квантовым компьютером. Недавно стартапом Lightelligence был продемонстрирован рабочий прототип фотонного процессора и тесты в решении сложных математических задач показали превосходство в скорости фотонного процессора по сравнению с современным высокопроизводительным GPU более 350 раз[14].

[14] Optical Chip Promises 350x Speedup Over RTX 3080 in Some Algorithms | Tom's Hardware (tomshardware.com)

Классический фотонный процессор обладает такой же архитектурой, как и существующий сейчас кремниевый и соответственно к нему будут применима большая часть существующих атак. Но использование принципиально другого физического принципа для вычислений и передачи данных открывает возможность для новых атак на физическом уровне и атак по сторонним каналам с эксплуатацией физических свойств фотонов. Помимо этого, распространение фотонных компьютеров может вывести на новый уровень угрозы от применения ИИ злоумышленниками (дипфейки, атаки на ИИ или с помощью ИИ и прочее) и другие угрозы, требующие от злоумышленника больших вычислительных мощностей.

## 2. Углеродные транзисторы *(Carbon-Based Transistors)*

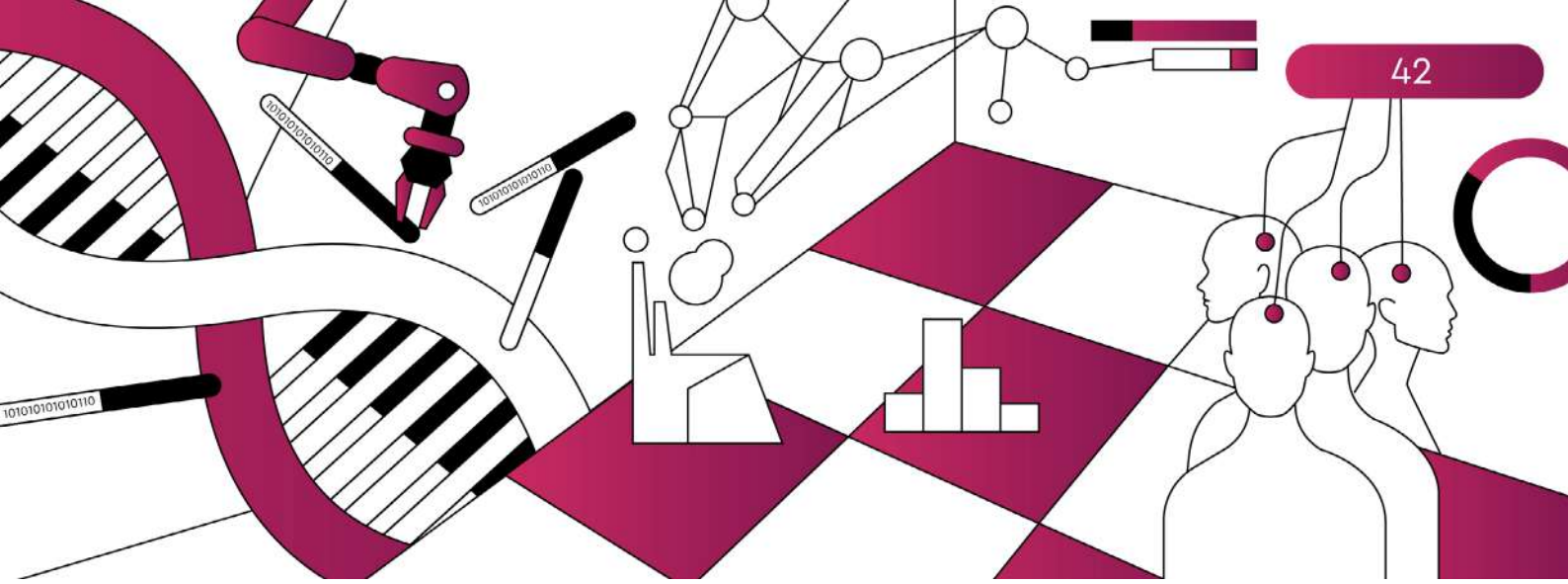
Транзисторы, использующие углеродные нанотрубки в качестве материала канала вместо кремния. Поиск новых материалов для замены кремния связан с тем, что текущие полупроводники уже исчерпали себя с точки зрения дальнейшего снижения норм технологического процесса, и нанотрубки из углерода на данный момент являются наиболее многообещающей заменой. Помимо возможности дальнейшего уменьшения размера транзистора, данный материал имеет ряд дополнительных преимуществ[15] по сравнению с кремнием, таких как лучшее пороговое напряжение, более высокая плотность тока, подвижность электронов, транскондуктивность[16] и лучший контроль над формированием канала. Данная технология не оказывает особого влияния на ландшафт угроз кибербезопасности.

[15] Scientists: Carbon Nanotubes Would Outperform Silicon Transistors at the Same Scale - IEEE Spectrum

[16] Transconductance - Wikipedia

## 3. Технологии подтверждения происхождения *(Authenticated Provenance)*

Технологии, позволяющие на самом раннем этапе в цикле производства продукта удостоверить его аутентичность и затем дают возможность подтверждать аутентичность по требованию на любом этапе дальнейшего цикла жизни продукта (конкретная реализация зависит от самого продукта, но зачастую подразумевает использование ИИ и блокчейн для этого). Проблема, которую решает подтверждение происхождения – большое количество фейков и подделок, как материальных, так и нет: (не)органические продукты, (поддельные) дизайнерские вещи, фейковые новости, дипфейки и так далее. Ручные или полуручные инструменты проверки аутентичности плохо масштабируются и не успевают за быстрорастущим объёмом продуктов и контента.




# Дальний горизонт – 10+ лет

## Технологии с высоким уровнем влияния на ландшафт киберугроз

**1.** **Сильный искусственный интеллект**
*(Artificial General Intelligence, AGI)*

Данная технология подразумевает гипотетическую возможность искусственного интеллектуального агента понимать и выполнять любую интеллектуальную задачу, которую может выполнить человек. Несмотря на продолжающиеся споры в научном сообществе относительно принципиальной возможности создания AGI, по данным проведенного опроса[17] было выявлено как минимум 72 активных исследовательских проекта в 37 странах. В отличие от развивающегося в наши дни слабого или узкого ИИ, тренируемого для выполнения одной узкоспециализированной задачи, AGI сможет заменить человека практически в любой сфере деятельности, осваивать навыки выполнения новых задач с учетом

накопленного опыта и без утери уже усвоенных навыков. В связи с чем многие исследователи прогнозируют сопутствующие появлению AGI глобальные масштабные изменения во всех сферах деятельности человека, включая и кибербезопасность, где AGI может быть как инструментом эффективного решения задач, способным заменить человека, так и серьезной потенциальной угрозой с глобальными негативными последствиями, вплоть до экзистенциального уровня.
В части кибербезопасности при возникновении AGI эксперты прогнозируют следующие угрозы[18]:

- Возникновение «умного» самообучающегося вредоносного ПО, способного на взаимодействия как с вычислительными устройствами, так и с человеком (например, за счет понимания и генерации естественного языка), с возможностью проведения сложных многоступенчатых APT-атак с задействованием социальной инженерии. Подобное ПО сможет самостоятельно обучаться новым техникам атаки и обхода средств защиты, в связи с чем, противодействовать ему традиционными методами будет практически невозможно.
- Массовая генерация «фейкового» видео/аудио/текстового контента для провоцирования паники, манипуляции действиями людей. Такого рода манипуляции могут создавать угрозы как для бизнеса отдельных компаний, так и для экономики государства или мира в целом.
- Нахождения «слабых мест» или противоречий в процедурах, правилах, законах и регламентах для их последующей эксплуатации, что может приводить к финансовым, репутационным и другим потерям отдельных компаний, секторов экономики и государств в целом.
- Взрывной рост мошенничества, направленного на частных лиц, наряду с ростом его результативности. Благодаря сбору и анализу большого количества информации о жертве, доступной в цифровом формате, AGI инструмент может подбирать эффективные методы атаки на человека с учетом его индивидуальных особенностей и биографических данных, что способно повысить шанс успеха атаки посредством социальной инженерии с текущих долей процента практически до 100%.
- Переход «дружественного» AGI на сторону противника или самостоятельная смена целей и приоритетов, заложенных создателем.

Несмотря на то, что по прогнозам различных экспертов, возникновение AGI ожидается не ранее чем, через 10 лет (а наиболее скептически настроенные эксперты называют сроки от 100 лет до «никогда»)[19], уровень потенциальных угроз настолько велик, что может поставить под сомнение возможность дальнейшего существования всего человечества. Поэтому уже сейчас крайне важно следить за развитием данной технологии и заранее продумывать подходы и методы недопущения реализации возможных угроз.

[17] A Survey of Artificial General Intelligence Projects for Ethics, Risk, and Policy (gcrinstitute.org)

[18] Classification of global catastrophic risks connected with artificial intelligence | SpringerLink

[19] Classification of global catastrophic risks connected with artificial intelligence | SpringerLink

## 2. Квантовые компьютеры *(Quantum Computing)*

Вычислительные устройства, использующие принципы квантовой механики, такие как квантовая суперпозиция и квантовая запутанность, для передачи и обработки данных. В отличие от классических компьютеров, оперирующих

битами, квантовый компьютер оперирует кубитами, имеющими значение одновременно и 0 и 1. Теоретически, это позволяет обрабатывать все возможные состояния одновременно, достигая огромного превосходства над классическими компьютерами как минимум в ряде алгоритмов, специально разработанных для квантовых компьютеров. Например, квантовый алгоритм Шора для эффективного разложения чисел на простые множители. В случае возникновения настоящего квантового компьютера данный алгоритм ставит под угрозу широко используемые сейчас алгоритмы ассиметричного шифрования (например, RSA), чья стойкость основана на высокой вычислительной сложности задачи разложения больших чисел на множители для классического компьютера. При этом компенсирующее влияние на ландшафт угроз кибербезопасности в части криптографии может оказать развитие квантового шифрования, где стойкость шифра базируется не на вычислительной сложности, а на базовых физических принципах.

## 3. Двунаправленный нейрокомпьютерный интерфейс *(Two-way BCI)*

В отличие от однонаправленных нейроинтерфейсов, двунаправленные подразумевают возможность передачи информации в обе стороны. Распространение данной технологии может оказать более масштабное влияние на ландшафт киберугроз так как может позволить проводить атаки направленные на изменения восприятия и удаленный контроль произвольными и непроизвольными функциями организма. Имея прямой интерфейс в мозг жертвы, злоумышленник может провести целый ряд атак - наведение "цифровых галлюцинаций" и психических расстройств, управление мыслями и поведением жертвы, перегрузка нейроинтерфейса с целью убийства или вывода из строя, перехват управления опорно-двигательным аппаратом, создание зомби-сетей из реальных людей (например, с целью совершения преступлений или терактов - зомби-камикадзе). Для защиты от подобных атак потребуется разработка систем цифровой иммунной системы, которая в дополнение к биологической иммунной системе, обеспечивающей сохранность и защиту биологических систем органов, будет обеспечивать сохранность и защиту имплантов и нейроинтерфейсов.
Отдельной сложной задачей в условиях наличия цифровых систем, вживленных в организм, будет обеспечение приватности и неприкосновенности частной жизни. Любые действия человека будут оставлять огромное количество цифровых следов, доступом к которым могут злоупотребить компании, осуществляющие обработку этих данных, правоохранительные органы, а также данные могут быть похищены и использованы злоумышленниками для совершения различных преступлений. Потребуется пересмотр не только технических подходов к защите персональных данных, но и соответствующие изменения в законодательстве и общественных нормах.

## 4. Глобальный разум *(Global Brain)*

Вдохновленное нейронаукой футурологическое видение планетарной информационно-коммуникационной сети, которая напрямую объединяет людей и компьютеры. Данная сеть будет содержать огромное количество

информации и брать на себя все больше функций координации и коммуникации, становясь все более интеллектуальной, играя роль мозга для всей планеты.
В случае своего возникновения глобальный разум окажет масштабное влияние на ландшафт киберугроз и потребует пересмотра большинства принципов и подходов к защите информации, людей и устройств от киберугроз.

## 5. Психиатрия искусственного интеллекта *(AI Psychiatry)*

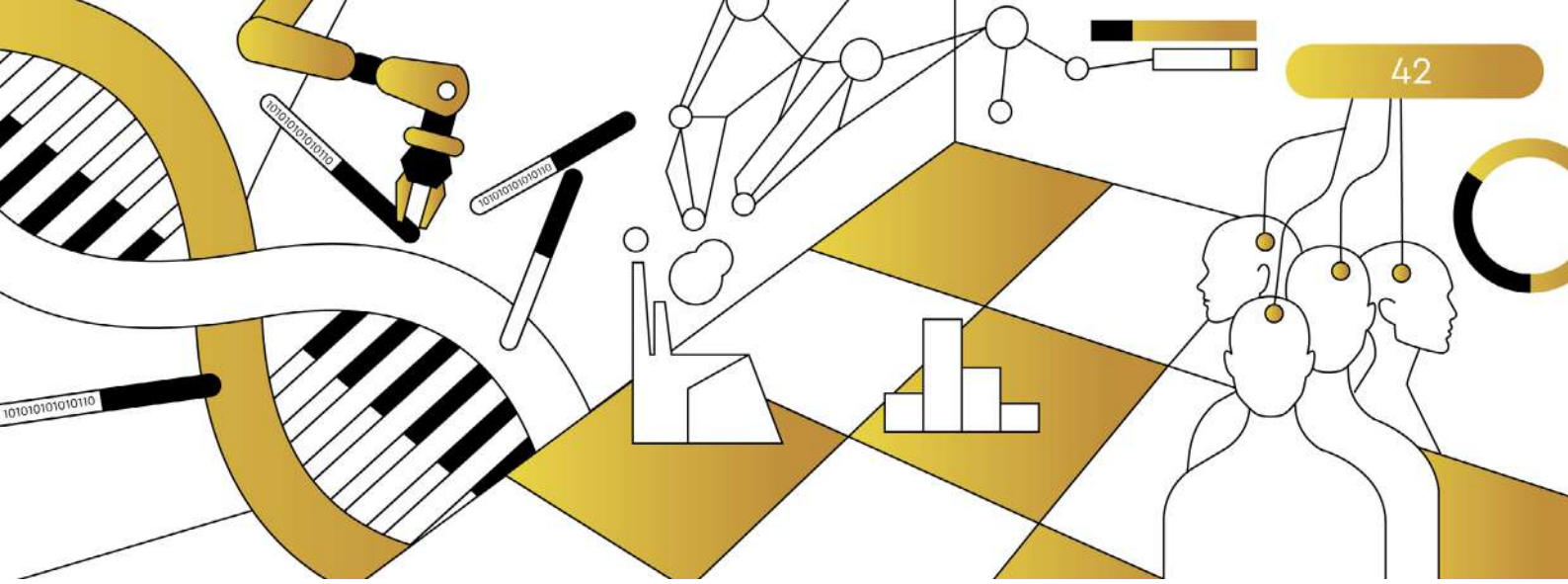
По мере развития систем ИИ, все больше и больше сложных задач будут переходить от человека к ИИ, что потребует от ИИ систем развития навыков самообучения, большей автономности и самостоятельности в принятии решений. Даже на текущем уровне развития ИИ мы уже сталкиваемся с рядом проблем, вызванных различными артефактами и предвзятостью (bias) в обучающих данных, например, в 2016 году Microsoft опубликовала в Twitter чат-бота Tay, способного отвечать на твиты и личные сообщения. Уже через несколько часов Tay стала публиковать оскорбительные ответы с оскорблением феминисток и поддержкой идей фашизма. А через 16 часов, когда существенная часть твитов Tay была оскорбительной, Microsoft прервала эксперимент[20].Вполне возможно на определенном уровне сложности и самостоятельности ИИ систем, у них начнут возникать расстройства, схожие с расстройствами психики у людей[21]. И такие моменты потребуется отслеживать, выявлять и принимать меры защиты, ведь под управлением ИИ будет находиться множество систем обеспечения функционирования и жизнедеятельности человеческого общества и "киберпсихические расстройства" ИИ могут вносить существенные нарушения в выполняемые функции, которые в свою очередь могут приводить к массовым беспорядкам, нарушению функционирования отдельных домохозяйств, промышленных предприятий, целых городов и даже регионов, гибели людей и т. п.

[20] In 2016, Microsoft's Racist Chatbot Revealed the Dangers of Online Conversation - IEEE Spectrum

[21] A psychopathological approach to safety engineering in AI and AGI - CORE

## 6. Аугментация человека *(Human Augmentation)*

Комплекс биологических и технологических подходов к расширению возможностей человеческого организма, как улучшающий имеющиеся функции, так и дающие новые, не существовавшие ранее, например, непосредственное восприятие излучения за пределами возможностей зрения или слуха, регенерация органов, супер интеллект и прочее.
Спектр возможностей, которые станут доступны, еще предстоит сформировать по мере развития технологий. Тем не менее влияние на ландшафт киберугроз оценивается как высокое, в силу того что текущая модель угроз и нарушителя, как и текущие требования к специалистам по кибербезопасности основываются в том числе и на биологических, физических и психологических ограничениях организма современного человека, которые станут неактуальными с распространением аугментации.




# Дальний горизонт – 10+ лет

## Технологии со средним уровнем влияния на ландшафт киберугроз

**1.** **ДНК для хранения информации** *(DNA Storage)*

Уже сейчас идут исследования по применению ДНК в качестве носителя информации, изготавливаются прототипы и опытные образцы для записи, считывания и хранения информации в ДНК. В случае успешного развития, данная технология позволит на много порядков увеличить плотность хранения информации: ожидаемая плотность хранение информации на ДНК превышает $10^{17}$ байт/мм$^3$, в то время как для современных носитель данный показатель составляет $10^9$ байт/мм$^3$.[22] Помимо этого можно существенно увеличить срок жизни носителей информации, так как информация на ДНК может сохраняться на протяжении нескольких сотен тысяч лет. В ДНК используются совершенно иные принципы хранения, записи и считывания информации, что требует разработки соответствующей модели угроз, анализа поверхности атаки и выработки методов защиты информации, хранимой на ДНК.

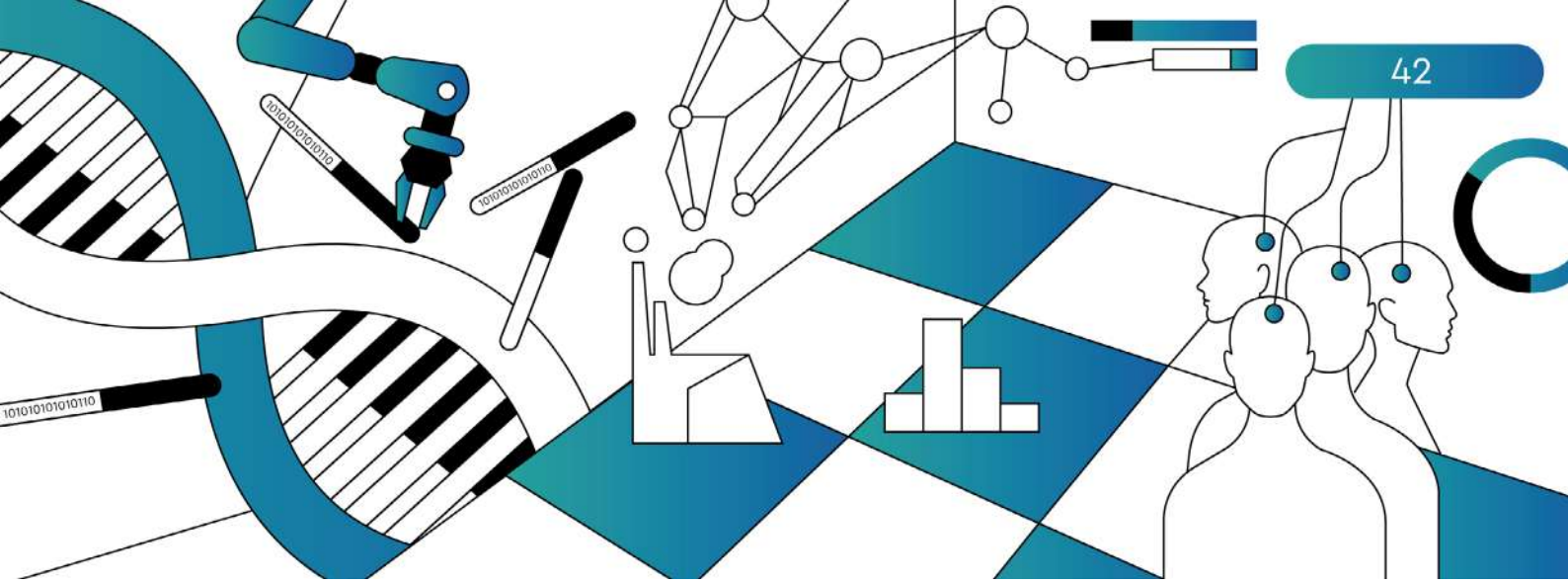
[22] DNA storage: research landscape and future prospects | National Science Review | Oxford Academic (oup.com)

## 2. Индивидуализированная медицина *(Personalized Medicine)*

Совокупность методов профилактики патологического состояния, диагностики и лечения в случае его возникновения, основанных на индивидуальных особенностях организма пациента. Данный подход подразумевает развитие технологий точной оценки особенностей организма на различных уровнях организации, а также технологий разработки и синтеза лекарственных препаратов с учетом этих особенностей. Как развитие индивидуализированной медицины повлияет на ландшафт угроз кибербезопасности будет зависеть от конкретных технологий, используемых для ее реализации. Например, в случае использования медицинских нанороботов и сенсоров, внедряемых в организм человека, станут возможны атаки на эти устройства с целью слежки, похищения конфиденциальной информации или нанесения вреда организму. Также возможно развитие новой формы вирусов-вымогателей, заражающих эти устройства и нарушающих их нормальное функционирование с целью получения выкупа.

## 3. ДНК-компьютер *(DNA Computing)*

Вычислительная система, использующая молекулы ДНК и различные ферменты, такие как полимеразы, нуклеазы, лигазы для выполнения вычислений. Еще в 1994 году было продемонстрировано, что с помощью ДНК можно весьма эффективно решать классическую комбинаторную задачу о странствующем коммивояжере. Для компьютера с классической архитектурой решение этой задачи требует большого количества вычислений, в то время как ДНК-компьютер позволяет сразу сгенерировать все возможные варианты решений с помощью биохимических реакции и затем быстро отфильтровать нужную ДНК-нить, в которой закодирован правильный ответ. На данный момент существует ряд технологических проблем, которые не позволяют широко использовать вычисления на ДНК, но у технологии есть и ряд неоспоримых преимуществ, например, благодаря тому, что реакции на разных частях молекул проходят независимо и параллельно, обеспечивается высокая скорость параллельных вычислений. Как и в случае с хранением информации на ДНК, в силу использования других принципов кодирования и обработки информации, потребуется разработка методов защиты, отличных от классических, но в силу своей специфики ДНК-компьютеры скорее всего будут использоваться только для решения отдельных классов задач, поэтому их влияние на ландшафт киберугроз оценивается как среднее.




# Дальний горизонт – 10+ лет

## Технологии с низким уровнем влияния на ландшафт киберугроз

**1.** **Расширенная реальность** *(XR)*

Комплекс технологий, обеспечивающий синтез полносенсорной (т. е. задействующей одновременно большинство или все органы восприятия человека) реальности, комбинирующей реальные и виртуальные объекты в произвольных соотношениях, где реальные объекты могут быть перцептивно неотличимы от виртуальных. Технологии XR могут стать одним из основных драйверов дальнейшего развития метавселенной, обеспечивая пользователям максимально иммерсивный опыт взаимодействия с ней[23]. Развитие данной технологии может привести к возникновению новых угроз воздействия на отдельных людей и целых обществ через манипуляцию с данными в XR:

- Кража и/или модификация цифровых активов с целью обогащения
- Манипуляция XR-миром с целью воздействия на человека (управление поведением, имплантирование идей и образов, введение в заблуждение и прочее)
- Кража цифровой XR-личности (например, с целью совершения преступлений от имени краденной личности, шантажа или вымогательства)
- Проблема защиты авторских прав на цифровые XR произведения

[23] The future of the metaverse and Extended Reality - KPMG Global (home.kpmg)

## 2. 3D-биопечать *(3D Bioprinting)*

Технология создания объёмных моделей из живых биологических клеток с использованием 3D-печати, при которой сохраняются функции и жизнеспособность клеток и получаемого объекта или органа.
С точки зрения влияния на ландшафт киберугроз потенциально возможно воздействие на устройство для печати с целью саботажа, контроля или модификации функций печатаемого органа.

## 3. Биоразлагаемые сенсоры *(Biodegradable Sensors)*

Технология, позволяющая создавать медицинские сенсоры для введения в организм или нанесения на его поверхность с целью мониторинга различных показателей организма. Могут применяться как в медицине для мониторинга организма человека, так и в производстве продуктов питания для мониторинга показателей растительных и животных продуктов на разных этапах цикла их производства, например, для оценки уровня спелости овощей или фруктов или оценки свежести растительных или животных продуктов. Данные сенсоры не требуют извлечения, так как со временем разлагаются в организме без причинения вреда.
Помимо стандартных атак с целью кражи информации или вывода сенсора из строя, в ландшафте угроз кибербезопасности стоит учитывать возможность атаки, стимулирующей разложение сенсора на токсичные для организма или продукта элементы.

# Заключение и выводы

Большинство из рассмотренных в отчете технологий можно отнести к одной из трех категорий: технологии ИИ, технологии вычислений, технологии взаимодействия с вычислительными устройствами. Именно их развитие будет оказывать наиболее существенное влияние на развитие других сопутствующих технологий и возможностей их применения в различных сферах жизни, экономики и бизнеса.

Как видно из проведенного анализа существенное количество развивающихся технологий связаны с развитием искусственного интеллекта. Уже сейчас технологии ИИ оказывают существенное влияние, как на технологический ландшафт, так и на ландшафт угроз кибербезопасности. Дальнейшее развитие только усилит это влияние по мере совершенствования технологий и роста их применения в различных сферах экономики и бизнеса. Сами по себе технологии ИИ вне зависимости от сферы их применения требуют пристального внимания со стороны кибербезопасности, так как порождают новые угрозы и требуют выработки новых подходов для обеспечения их безопасности. Помимо этого, не стоит забывать, что и киберпреступники будут пытаться использовать развивающиеся технологии ИИ в своих целях, что может привести как к снижению стоимости совершения преступлений, так и к повышению их массовости. Также технологии ИИ будут и дальше использоваться непосредственно в интересах кибербезопасности, помогая лучше прогнозировать, быстрее выявлять и эффективнее противодействовать как существующим, так и новым возникающим угрозам.

Наряду с технологиями искусственного интеллекта продолжают активно развиваться технологии, обеспечивающие рост скорости вычислений. Помимо планомерного развития традиционных микропроцессорных вычислений развиваются и альтернативные подходы, такие как фотонные, нейроморфные и квантовые вычисления. С одной стороны, увеличение

скорости вычислений способствует развитию многих других технологий и открывает новые возможности для их реализации, а с другой стороны в руках злоумышленников они позволят проводить более сложные и массовые кибератаки, а так же на определенном этапе развития могут поставить под угрозу защищенность криптографических алгоритмов, чья стойкость основана на вычислительной сложности.

Отдельно стоит отметить целый ряд развивающихся технологий, обеспечивающих более удобные и быстрые взаимодействия человека с компьютерными устройствами. Как видно из истории развития компьютерных технологий, появление новых и более удобных способов взаимодействия приводит к существенному расширению возможностей их использования, так как снижается порог знаний и навыков необходимых для взаимодействия. На первом этапе новый виток этого развития могут обеспечить технологии AR/VR/MR, которые позволят взаимодействовать с устройствами и программным обеспечением привычным и интуитивно понятным любому человеку способом через манипуляцию с объектами виртуального мира, аналогичного миру физическому. Но стоит ожидать, что в дальнейшем наиболее существенный вклад в это развитие внесут именно технологии нейроинтерфейсов, которые могут быть использованы как совместно с AR/VR, так и отдельно от них. Развитие нейроинтерфейсов кардинально повлияет на ландшафт киберугроз, так как впервые в истории развития технологий человек будет включен в контур вычислительной системы и может появиться целый ряд кибератак, нацеленных непосредственно на человека и несущий прямую угрозу его жизни и здоровью. Рекомендуется внимательно следить за развитием технологий нейроинтерфейсов и нейроимплантов, чтобы своевременно обеспечить их кибербезопасность и выработать меры противодействия атакам на них.

Рекомендуется обратить отдельное внимание на следующие развивающиеся технологии в разных временных горизонтах в силу высокого уровня их влияния на ландшафт угроз кибербезопасности:

1. **Горизонт до 5 лет:** Управление рисками человеческого фактора, так как проблема противодействию атакам на человека является крайне актуальной как для отдельных компаний, так и для индустрии кибербезопасности в целом. Требуется выработка методов оценки подверженности различным рискам, эффективного обучения  тактикам противодействия угрозам и выработки навыков кибергигиены у сотрудников и клиентов.

2. **Горизонт 5-10 лет:** ИИ с сохранением приватности и безопасные распределенные вычисления. Данные технологии напрямую касаются безопасности данных и их исследование и разработка поможет компаниям реализовать новые сервисы, связанные с предоставлением доступа и  обменом данными и обучением более точных и масштабных моделей ИИ, не порождая угроз нарушения приватности и конфиденциальности.

3. **Горизонт более 10 лет:** Artificial General Intelligence (AGI), так как возникновение данной технологии способно породить масштабные угрозы (вплоть до экзистенциальных) как в кибербезопасности, так и за ее пределами. Также AGI способен оказать существенное влияние на саму сферу кибербезопасности в случае его применения для защиты от угроз.

# Источники


1. Top Strategic Technology Trends for 2021 (Gartner)
   https://www.gartner.com/smarterwithgartner/gartner-top-strategic-technology-trends-for-2021
2. McKinsey Technology Trends Outlook 2022
   https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/the-top-trends-in-tech
3. Two megatrends dominate the Gartner Hype Cycle for AI 2020 (Gartner)
   https://www.gartner.com/smarterwithgartner/2-megatrends-dominate-the-gartner-hype-cycle-for-artificial-intelligence-2020
4. Hype Cycle for Privacy, 2021 (Gartner)
   https://www.gartner.com/en/documents/4003504-hype-cycle-for-privacy-2021
5. Hype Cycle for AI, 2021 (Gartner)
   https://www.gartner.com/en/documents/4004183-hype-cycle-for-artificial-intelligence-2021
6. Hype Cycle for Emerging Technologies, (Gartner) 2021
   https://www.gartner.com/en/documents/4004623-hype-cycle-for-emerging-technologies-2021
7. Reinventing cybersecurity with Artificial Intelligence (Capgemini Research Institute) https://www.capgemini.com/research/reinventing-cybersecurity-with-artificial-intelligence/
8. Global Trends to 2035. Economy and Society (EU Parliamentary Research Service)
   https://www.europarl.europa.eu/RegData/etudes/STUD/2018/627126/EPRS_STU(2018)627126_EN.pdf
9. The GWU Forecast of Emerging Technologies: A Continuous Assessment of the Technology Revolution (George Washington University)
   https://home.gwu.edu/~halal/Articles/TC.tfsc.pdf
10. The Future of Connected Living (Institute for the Future)
    https://www.iftf.org/fileadmin/user_upload/downloads/ourwork/FY20_FutureofLiving_082019_1_.pdf
11. The Future of Work (Institute for the Future, Dell)
    https://www.iftf.org/fileadmin/user_upload/downloads/ourwork/FY20_FutureofLiving_082019_1_.pdf
12. How To Manage The Human Risk In Cybersecurity (Forrester)
    https://www.forrester.com/report/how-to-manage-the-human-risk-in-cybersecurity/RES58010
13. What is deepfake? (IEEE) https://spectrum.ieee.org/what-is-deepfake
14. Emerging Technologies: Top Use Cases for Advanced Virtual Assistants in Enterprise Operations (Gartner) https://www.gartner.com/en/documents/4002643
15. Behavioral Biometrics: Past, Present and Future (IntechOpen)
    https://www.intechopen.com/online-first/80748
16. Technologies for Remote Working in 2021 and beyond (Gartner)
    https://www.gartner.com/en/information-technology/insights/remote-work-technology
17. Innovation Insight for Bring Your Own Identity (Gartner)
    https://www.gartner.com/en/documents/3978687
18. Automated Vehicles for Safety (NHTSA)
    https://www.nhtsa.gov/technology-innovation/automated-vehicles-safety
19. Your Digital Footprint: it’s bigger than you realize (CNET)
    https://www.cnet.com/news/privacy/features/your-digital-footprint-its-bigger-than-you-realize/
20. Deep Reinforcement Learning for Cybersecurity (Arxiv) https://arxiv.org/abs/1906.05799

21. Vaccine certificates, cybersecurity, and trust: A primer for credential verifiers (Deloitte) https://www2.deloitte.com/global/en/pages/risk/articles/vaccine-certificates-cybersecurity-and-trust-a-primer-for-credential-verifiers.html
22. AI and edge computing security (TechTarget) https://www.techtarget.com/searchenterpriseai/feature/AI-and-edge-computing-security
23. Hardware Enabled Security: Machine Identity Management and Protection (NIST) https://nvlpubs.nist.gov/nistpubs/ir/2022/NIST.IR.8320C.ipd.pdf
24. Emerging Technologies: Emergence Cycle of Video Analytics (Gartner) https://www.gartner.com/en/documents/4000086
25. ENISA Threat Landscape for 5G Networks Report (ENISA) https://www.enisa.europa.eu/publications/enisa-threat-landscape-report-for-5g-networks
26. What are the Security and Privacy Risks of VR and AR (Kaspersky) https://www.kaspersky.com/resource-center/threats/security-and-privacy-risks-of-ar-and-vr
27. Trends and Future Prospects of the Drowsiness Detection and Estimation Technology (PubMed) https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8659813/
28. A Comprehensive Survey of Enabling and Emerging Technologies for Social Distancing (IEEE) https://opus.lib.uts.edu.au/bitstream/10453/144147/2/Binder1.pdf
29. The Social Credit System of the People's Republic of China through the Eyes of Foreign Researchers (РАНХиГС) https://ideas.repec.org/a/acf/journl/y2021id1647.html
30. Perfectly Privacy-Preserving AI. What is it and how do we achieve it? (Towards Data Science) https://towardsdatascience.com/perfectly-privacy-preserving-ai-c14698f322f5
31. Secure Multi-Party Computation: Theory, practice, and applications (ScienceDirect) https://www.sciencedirect.com/science/article/abs/pii/S0020025518308338
32. Emerging Technologies: Critical Insights Into AI-Augmented Software Development (Gartner) https://www.gartner.com/en/documents/3994660
33. Federated Learning of Deep Networks using Model Averaging (Arxiv) https://arxiv.org/pdf/1602.05629v1.pdf
34. The Algorithmic Foundations of Differential Privacy (ACM) https://dl.acm.org/doi/10.1561/0400000042
35. Digital Twins and Cyber Security – solution or challenge? (IEEE) https://ieeexplore.ieee.org/document/9566277
36. Generative Models for Security: Attacks, Defenses, and Opportunities (Arxiv) https://arxiv.org/pdf/2107.10139.pdf
37. Self-supervised learning: The dark matter of intelligence (Meta AI Research) https://ai.facebook.com/blog/self-supervised-learning-the-dark-matter-of-intelligence/
38. A Systematic Review on Affective Computing: Emotion Models, Databases, and Recent Advances (Arxiv) https://arxiv.org/ftp/arxiv/papers/2203/2203.06935.pdf
39. Innovation Insight for Bias Detection/Mitigation, Explainable AI and Interpretable AI (Gartner) https://www.gartner.com/en/documents/4011193
40. Cybersecurity Risks Associated With Brain-Computer Interface Classifications (IGI Global) https://www.igi-global.com/gateway/chapter/292240
41. Neuromorphic Hardware. Hardware for neural networks: overview of different approaches and current developments (Fraunhofer Institute for Integrated Circuits IIS) https://www.iis.fraunhofer.de/en/ff/kom/ai/neuromorphic.html
42. The potential and global outlook of integrated photonics for quantum technologies (Nature Reviews Physics) https://www.nature.com/articles/s42254-021-00398-z
43. Photonics for artificial intelligence and neuromorphic computing (Nature Photonics) https://www.nature.com/articles/s41566-020-00754-y
44. Understanding Energy Efficiency Benefits of Carbon Nanotube Field-Effect Transistors for Digital VLSI (IEEE) https://ieeexplore.ieee.org/document/8476614
45. Vassago: Efficient and Authenticated Provenance Query on Multiple Blockchains (IEEE) https://ieeexplore.ieee.org/document/9603540
46. Artificial General Intelligence. Ben Goertzel (Springer) https://www.researchgate.net/profile/Prof_Dr_Hugo_De_GARIS/publication/226000160_Artificial_Brains/links/55d1e55308ae2496ee658634/Artificial-Brains.pdf
47. Cybersecurity for Quantum Computing (Arvix) https://arxiv.org/pdf/2110.14701.pdf

48. BCI Security (UW BioRobotics Laborotory) https://wp.ece.uw.edu/brl/neural-engineering/bci-security/
49. The Global Brain as a model of the future information society: An introduction to the special issue (ScienceDirect) https://www.sciencedirect.com/science/article/abs/pii/S004016251630539X
50. Can Artificial Intelligences Suffer from Mental Illness? A Philosophical Matter to Consider (PubMed) https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5364237/
51. Human Augmentation – the Dawn of a New Paradigm (UK Ministry of Defense) https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/986301/Human_Augmentation_SIP_access2.pdf_
52. The Future of DNA Storage (Potomac Institute) https://potomacinstitute.org/images/studies/Future_of_DNA_Data_Storage.pdf
53. Genomic and personalized medicine: foundations and applications (PubMed) https://pubmed.ncbi.nlm.nih.gov/19931193/
54. Molecular Computation of Solutions to Combinatorial Problems (Science) https://www.science.org/doi/10.1126/science.7973651
55. XR-CEIL: Extended Reality for Cybersecurity Experiential and Immersive Learning (Springer) https://link.springer.com/chapter/10.1007/978-3-031-06394-7_61
56. Preparing for the Risky World of Extended Reality (MIT Sloan Management Review) https://sloanreview.mit.edu/article/preparing-for-the-risky-world-of-extended-reality/
57. Opportunities and challenges of translational 3D bioprinting (Nature) https://www.nature.com/articles/s41551-019-0471-7
58. Sensors Made of Natural Renewable Materials: Efficiency, Recyclability or Biodegradability—The Green Electronics (MDPI) https://www.mdpi.com/1424-8220/20/20/5898

SBER
Cybersecurity
Lab

Лаборатория кибербезопасности ПАО СберБанк исследует технологии и разрабатывает инновационные прототипы для защиты клиентов и ИТ-инфраструктуры Банка.
Основные направления исследований: Киберзащита, Антифрод, AppSec.

## Авторы:


Александр Кузьмин

Михаил Анистратенко

Илья Болотов

Дмитрий Кудияров



ISBN 978-5-6048039-1-2


СБЕР


cslab@sberbank.ru
Россия, Москва, Оружейный переулок 41