сенсус

ЦИФРОВОЙ МАРКЕТПЛЕЙС

Kazan Digital Week 2026 · аналитика по направлениям

Кибербезопасность и цифровая устойчивость: ИИ ускорил атаки, а защита по-прежнему начинается с гигиены

Что на Kazan Digital Week 2026 говорили об атаках через подрядчиков, уязвимостях в эпоху ИИ, агентных SOC, защите ИИ-агентов, приказе ФСТЭК № 117 и замене Microsoft AD. Цифры, практики и позиции сторон.

Цифровой маркетплейс · platforms.su Аналитика: Сенсус · sensus.com.ru 11.11.2026

Содержание

Ключевые показатели

Кратко

1. Устойчивость – задача первого лица, а не службы ИБ

2. Деньги: как считать ущерб и почему бюджеты уходят не туда

3. Гигиена: откуда на самом деле приходят атаки

4. Подрядчики и цепочки поставок: доступ, который компании открывают сами

5. Уязвимости в эпоху ИИ: окно до эксплуатации схлопнулось

6. Восстановление: исходить из того, что взлом случится

7. Мониторинг и реагирование: данные, SIEM и автоматизация

8. Agentic SOC: автономии нет, спор идёт об экономике и людях

9. ИИ-агенты в компании: доступы, контроль и теневой ИИ

10. Регулирование и ответственность: приказ ФСТЭК № 117 и «человек в контуре»

11. Импортозамещение и технологический суверенитет: от «замещать всё» к управляемости

12. Исследования вузов: интерпретируемый ИИ, АСУ ТП и биометрия

Выводы

Источники

Ключевые показатели

около 30%

Доля атак через доступ подрядчиков, прошлый год

Против 15% в 2024 году · Маслов Дмитрий, BI.ZONE

около 42 дней

Среднее время присутствия злоумышленника в сети

Статистика за прошлый год · Маслов Дмитрий, BI.ZONE

5 дней

Срок от выхода патча до эксплуатации уязвимости

Было 63 дня · Буравцов Александр, СБК

271 из 423

Уязвимостей Mozilla, найденных с помощью ИИ

Статистика Mozilla Foundation · Буравцов Александр, СБК

25,5 часа

Выпуск исправления после публикации уязвимости

Сборки под Windows, Linux и контейнеры · Якупов Лев, TrueConf

около 44%

Начальных атак через общедоступные приложения

Ещё 25–26% – через учётные данные · Кулашова Анна, Лаборатория Касперского

88%

Успешных атак из-за слабой аутентификации

Официальная аналитика · Груздев Сергей, Аладдин

около 98%

Заявок bug bounty, сгенерированных с помощью ИИ

Программа вендора · Кулашова Анна, Лаборатория Касперского

около 25%

Событий в SOC, которые разбирает ИИ

Финальный вердикт за людьми · Кулашова Анна, Лаборатория Касперского

91%

Компаний, применяющих ИИ-агентов

Исследование EY · Шепелявый Дмитрий, Innostage

85%

Компаний, где агенты действуют без человека

Исследование EY · Шепелявый Дмитрий, Innostage

в 10–40 раз

Удорожание внутренних ИИ-моделей к внешним

Закрытый контур с GPU и защитой · Соловьев Святослав, Альфа-Банк

50%

Доля рынка ИБ у пяти крупнейших вендоров

Оценка · Груздев Сергей, Аладдин

80–90%

Компаний, оставшихся на Microsoft AD

Оценка · Добрушский Сергей, Сайберпик

около 30%

Рынка, перешедшего на отечественный мониторинг

При 15 системах на рынке · Ганюшкин Николай, Монк Диджитал Лаб

2 раза в год

Тестирование планов непрерывности

Каждый план · Лутфиев Рамиль, Газпромнефть ИТО

Кратко

Кибербезопасности и цифровой устойчивости на Kazan Digital Week 2026 были посвящены восемь сессий: семь дискуссий трека «Кибербезопасность нового времени» и партнёрской программы и научная секция об интеллектуальных системах защиты. Выступали Innostage, BI.ZONE, «Лаборатория Касперского», ИнфоТеКС, «Аладдин», РЕД СОФТ, Альфа-Банк, Домклик, ПСБ, «Газпромнефть ИТО», Госкорпорация по ОрВД, вузы Казани, Уфы и Москвы. Главные выводы такие. Доля атак через доступ подрядчиков за год выросла примерно с 15% до 30% (Маслов Дмитрий, BI.ZONE). Срок от выхода патча до эксплуатации уязвимости сократился с 63 до 5 дней (Буравцов Александр, СБК). При этом 88% успешных атак происходят из-за слабой идентификации и аутентификации (Груздев Сергей, «Аладдин»), а 85% компаний уже допускают действия ИИ-агентов без участия человека (Шепелявый Дмитрий, Innostage, со ссылкой на EY).

1. Устойчивость – задача первого лица, а не службы ИБ

Участники сходились в том, что устойчивость бизнеса нельзя обеспечить силами одной службы ИБ: она возникает только там, где ИТ, безопасность и владельцы процессов работают по общим метрикам, а ответственность поднята до уровня руководства.

1.1 Три вида устойчивости и разрыв между командами

Рустем Хабибуллин (Innostage) разделил киберустойчивость (данные и инфраструктура), функциональную устойчивость (функции систем для пользователей) и операционную (бизнес целиком). По его словам, функциональная устойчивость – связующее звено, которое заказчики чаще всего упускают. При сбое ИТ перезагружает серверы, ИБ изолирует контуры, а бизнес теряет клиентов – команды действуют как в басне Крылова. Innostage свела метрики ИТ, ИБ и бизнеса в единые дашборды и в одном из проектов смогла связать деградацию бизнес-функции со сбоем СУБД, который вызвала DDoS-атака. Задачу Хабибуллин сформулировал так: построить отказоустойчивый процесс на заведомо неустойчивой инфраструктуре.

Дмитрий Шепелявый (Innostage) описал взгляд генерального директора: ему неважна причина остановки, важно, чтобы бизнес не падал, быстро восстанавливался и не разорился на мерах защиты. Компания всегда находится в треугольнике «реальная защищённость – требования регуляторов – бюджет». Типичный конфликт он описал так: ИТ и ИБ перекладывают ответственность друг на друга и на финансистов, главный инженер не пускает средства защиты в АСУ ТП, после чего вторжение идёт именно туда. Владимир Дрюков (Innostage) добавил, что дистанция от ИТ-сбоя до кибератаки, способной обрушить бизнес, очень короткая, поэтому ИТ, ИБ и технологические процессы нужно рассматривать вместе.

1.2 Кто отвечает

Указ Президента № 250 предписывает назначить ответственного за информационную безопасность; на практике это обычно заместитель генерального директора, часто финансовый (Шепелявый Дмитрий, Innostage). Евгений Хасин (BI.ZONE) считает ключевой идеей указа связь ИТ и ИБ ради устойчивости бизнеса. Но, по его наблюдению, знания ИТ и ИБ в одном лице свести в большинстве случаев не удалось, а ответственного за ИБ нередко делали ответственным за всю безопасность, включая пожарную. ИБ долго «поднималась снизу»: руководство не понимало языка безопасников, а его показатели не были увязаны с задачами защиты.

Николай Ганюшкин («Монк Диджитал Лаб») сформулировал правило: чем выше точка ответственности за операционную надёжность, тем больше вероятность, что проект взлетит; крупные вендоры советуют вводить должность вице-президента по операционной устойчивости. Андрей Лагоденко (Домклик) признал, что в его компании ИТ и ИБ пока не объединены: служба ИБ делится лишь частью дашбордов, поскольку требования конфиденциальности сохраняются.

1.3 Как это устроено у тех, кто дошёл до практики

Рамиль Лутфиев («Газпромнефть ИТО») рассказал, что культура непрерывности бизнеса в группе «Газпром нефть» внедряется по распоряжению Александра Дюкова 2017 года. Система работает 7 лет, 5-й год подряд группа ресертифицируется по международному стандарту непрерывности бизнеса, а каждый план аварийного восстановления и непрерывности тестируется 2 раза в год. В прошлом году принята стратегия технологической устойчивости на трёх постулатах: непрерывность деятельности, киберустойчивость и импортозамещение. Компания рассчитывает индексы технологической устойчивости и надёжности и использует их в инвестиционных решениях при закупках и интеграциях. Внедрение начинается с анализа воздействия на бизнес – картирования процессов и оценки финансового, юридического и имиджевого ущерба, а владелец процесса и ответственные за активацию плана закрепляются под подпись.

Александр Кравцов (Межгосударственный банк) описал «альбом процессов»: у каждого процесса есть бизнес-владелец, ИТ-владелец и администраторы ИБ, а SLA включены в показатели эффективности подразделений. Процессы делятся на mission critical, business critical и офисные, и путать метрики процесса с метриками систем нельзя. Проблема в том, что каждый владелец считает свой процесс критичным и требует RTO и RPO в 5–10 минут, не понимая, что это прежде всего деньги. В финансовом секторе дополнительным рычагом служат жёсткие требования Банка России.

В крупной цифровой компании такой подход не масштабируется: при 1 000 постоянно меняющихся сервисов договориться о SLA с каждой командой нереально, поэтому Домклик использует чек-листы бизнес-критичности на этапе проектирования и еженедельные разборы всех инцидентов (Лагоденко Андрей, Домклик).

2. Деньги: как считать ущерб и почему бюджеты уходят не туда

Спикеры критиковали оба популярных способа обосновать расходы на защиту – завышенный расчёт ущерба и покупку «того, что на слуху» – и предлагали опираться на реальные векторы атак и правило Парето.

2.1 Оценка ущерба

Шепелявый назвал расчёт эффекта по формуле «час простоя × упущенная прибыль» компрометирующим всю экономику безопасности: он даёт нереалистичные триллионы. Профессиональный подход – считать реальную упущенную прибыль с учётом компенсирующих и организационных мер и перехода на пониженный режим; тогда вложения в устойчивость можно обосновать или, наоборот, показать, что они не нужны. Буквальное прочтение требований регуляторов тоже ведёт к абсурду: по словам Шепелявого, тогда «вместо закупки доменной печи надо закупать систему безопасности». Он привёл случай владельца холдинга, чьи системы стали объектами КИИ: безопасники предложили бюджет, поглощающий всю инвестиционную программу.

Хасин считает, что общей методики оценки ущерба быть не может: задача решается для конкретных предприятий и процессов, «как задача трёх тел» – общего решения нет, частные есть. Неготовым заказчикам Шепелявый предлагает начинать с оценки эффективности защиты «крупными мазками» – по экспертным знаниям, опросникам и журналам. Хасин также обратил внимание, что новая редакция приказа ФСТЭК № 117 строит защиту вокруг целей безопасности в политике ИБ: если формулировать их вместе с первым лицом и финансистами, безопасность перестаёт быть бумажной.

2.2 Сколько тратить и на что

Универсальной нормы нет. Часто называемые 10% от ИТ-бюджета в некоторых компаниях действительно работают, но рост ИТ-бюджета, например покупка GPU-кластеров, не должен автоматически увеличивать бюджет ИБ (Лагоденко Андрей, Домклик). Кравцов считает, что 10% близки к реальности на старте ИИ-проекта, когда закладываются переиспользуемые меры защиты, а позже прямая зависимость исчезает.

Сергей Груздев («Аладдин») назвал рынок ИБ «кривым зеркалом»: организации тратят деньги обратно пропорционально тому, куда их реально бьют, – «мы тратим деньги не туда, куда нас бьют». По его оценке, 5 вендоров («Лаборатория Касперского», Positive Technologies, UserGate, ИнфоТеКС и «Код безопасности») забирают 50% российского рынка ИБ, ещё 10% приходится на SOC и киберучения, на весь остальной рынок остаётся 40%. В экспресс-опросе на конференции по identity 63–65% участников ответили, что условный миллион потратили бы в первую очередь на NGFW. Сергей Добрушский («Сайберпик») объяснил это тем, что бюджет проще получить на популярное: с 2022 года произошёл перекос в сторону сетевой безопасности. Александр Махновский («Аванпост») добавил, что

многоуровневая сегментация – такая же гигиена, как многофакторная аутентификация, только дороже, отсюда и перекос. Груздев также упрекнул рынок в том, что он «продаёт страхи»: руководителю ИБ проще показать дашборд ночных атак и получить деньги.

Выход Груздев видит в «эффективной безопасности» по правилу Парето: 20% усилий должны давать 80% результата, а немногие оставшиеся бюджеты – уходить на топ-5 реальных проблем. Илья Лорич (Innostage) дополнил: золотой пули нет, «эту пулю нужно отливать для каждой организации самостоятельно», последовательно проходя уровни гигиены, реагирования и устойчивости.

2.3 Киберстрахование

Рынок киберстрахования в России, по оценке Шепелявого, пока не откалиброван: полисы, выписанные крупным клиентам, были «лотереей», поскольку в стране нет сервисов актуарной статистики по ущербам и видам инцидентов. Он прогнозирует, что через 2–3 года страховая премия будет зависеть от измеримого уровня устойчивости клиента. Хасин напомнил, что полис не освобождает от мер защиты: уголовная и административная «ответственность не страхуется».

3. Гигиена: откуда на самом деле приходят атаки

Самый согласованный вывод направления: успешные атаки почти всегда используют базовые изъяны – слабую аутентификацию, незакрытые уязвимости на периметре и плоские сети, и закрыть их дешевле, чем покупать сложные системы.

3.1 Статистика первичного проникновения

По данным «Лаборатории Касперского», около 44% начальных атак начинаются с уязвимостей в общедоступных приложениях и ещё 25–26% – с легитимных учётных данных, полученных через фишинг (Кулашова Анна, «Лаборатория Касперского»). Груздев со ссылкой на официальную аналитику привёл цифру 88%: столько успешных атак происходит из-за неправильной или слабой подсистемы идентификации и аутентификации. Махновский по международной аналитике назвал диапазон 85–95% первичных проникновений через слабую аутентификацию. Груздев также отметил, что в текущем полугодии на первое место среди источников успешных атак вышли атаки через уязвимости, а по имеющейся статистике чуть ли не в каждой 3-й компании есть скрытое присутствие посторонних, которые готовят атаку. Светлана Старовойт (ИнфоТеКС) назвала самым частым вектором веб-ресурсы, поскольку они доступны снаружи.

Рустем Ясовеев (Innostage) перечислил типовые причины громких инцидентов, выявленные расследованиями: плоская сеть, неконтролируемый состав сервисов на периметре и одинаковые пароли для разных учётных записей. Круглосуточный SOC, по его словам, не делает компанию защищённой, если удалённый доступ организован без должной аутентификации.

3.2 Пароли и второй фактор

Груздев предложил «забыть про пароли»: в эпоху ИИ все утёкшие парольные фразы давно собраны моделями, а в мерах защиты приказа ФСТЭК № 117 пароли оставлены только для одного случая. Код по SMS или почте он не считает вторым фактором: «это не двухфакторная, это двухэтапная. Просто ключ распилен на 2 части». Методы атак унифицированы, и компанию с внедрённой двухфакторной аутентификацией атакующие обходят стороной; тратить деньги на киберучения и киберполигоны, пока «ключ лежит под ковриком», по его мнению, бессмысленно. Лорич уточнил систему координат: защита строится на трёх уровнях – гигиена, реагирование, устойчивость; двухфакторная аутентификация относится к гигиене, киберучения – к реагированию.

3.3 Инвентаризация и сегментация

Кравцов начинает с инвентаризации и реестра внешних публикаций. В Межгосударственном банке после отключения сервиса без владельца осталась публикация, и со временем на том же IP-адресе оказался доступный извне центр управления vSphere, который быстро привлёк атакующих; злоумышленник низкой квалификации сбросил пароль администратора и этим себя выдал. Публикации, о которых никто не знает, Кравцов советует отключать и смотреть, кто придёт с претензией. В МГТУ им. Баумана, где работает более 1 000 информационных систем, веб-ресурсы паспортизируют, а ресурс без ответственного отключают до его назначения (Полянский Алексей, МГТУ им. Баумана).

Сегментации, по словам Кравцова, боятся потому, что сервисы в плоских сетях не описаны и при переносе в сегменты начинают рушиться процессы; метрикой успеха он считает прозрачность меры для бизнеса. Эльман Бейбутов (UserGate) видит корень проблемы в разделённой ответственности: ИТ отвечает за доступность, ИБ – за ограничения, поэтому сеть и её безопасность нужно вести в одном домене у архитектора. Практический совет Бейбутова – выгрузить access-листы межсетевого экрана во внутреннюю языковую модель, попросить оптимизировать правила и вместе со статистикой их работы за 90 дней получить свод описания сервисов и доступов.

Хасин и Шепелявый согласились, что новому руководителю стоит начинать с самых дешёвых мер с наибольшим эффектом: инвентаризации, харденинга, патчей и сегментации. Дрюков предложил простой тест зрелости ИТ-команды – спросить, когда она в последний раз проверяла восстановление из резервной копии.

4. Подрядчики и цепочки поставок: доступ, который компании открывают сами

Значительная часть инцидентов связана не с уязвимостями, а с легитимными доступами – каналами для удалённых сотрудников и контрагентов, которые компании открывают сами (Ясовеев Рустем, Innostage). Эта доля растёт быстрее остальных.

4.1 Масштаб угрозы

Дмитрий Маслов (BI.ZONE) привёл динамику: в 2024 году через доступ подрядчиков начиналось порядка 15% атак, в прошлом году – порядка 30%, итоги текущего года ещё не подведены. По статистике BI.ZONE, среднее время нахождения злоумышленников в инфраструктуре за прошлый год составило порядка 42 дней, и эта длительность растёт. Груздев отметил, что атаки через цепочки поставок растут «линейно или экспоненциально» и бьют по контрагентам – как правило, небольшим и хуже защищённым компаниям. Анна Кулашова («Лаборатория Касперского») отнесла атаки через цепочки поставок и доверительные отношения к тройке главных угроз.

4.2 От входа в сеть к доступу к приложению

VPN с многофакторной аутентификацией управляет только входом и не показывает, куда подрядчик движется дальше и с какими системами работает, поэтому доступ нужно рассматривать на уровне конкретных приложений (Маслов Дмитрий, BI.ZONE). Аудит обычно выявляет общие учётные записи и записи с неконтролируемым временем доступа; их исключают, всех заводят в единую точку входа со вторым фактором, а учётные записи выдают по запросу на ограниченный срок, с логированием и видеофиксацией действий. Zero Trust, по словам Маслова, внедряется поэтапно с оценкой критичности учётных записей.

Илья Борисов (Билайн) считает, что к подрядчику должны применяться как минимум такие же требования, как к своей команде, а обычно жёстче. Новая проблема – ИИ-агенты: запущенные подрядчиком агенты продолжают работать после его ухода, создают других агентов и инфраструктуру, и блокировка учётной записи подрядчика больше ничего не решает. Справиться с этим без столь же быстрых автоматизированных систем не получится, а каждый доступ любого субъекта – агента, системы или человека – должен быть «всегда проверен, всегда минимизирован, всегда отвечать какой-то бизнес-задаче». Привычные индикаторы тоже перестали работать: активность администраторов в нерабочее время больше не признак атаки, нужна поведенческая аналитика.

Груздев исходит из того, что владелец системы не контролирует среду, из которой подключается подрядчик, и должен считать её скомпрометированной. По требованиям приказа № 117 доступ из недоверенной среды запрещён; для удалённой работы есть специализированные средства, сертифицированные ФСТЭК по приказу № 32, которые создают доверенную среду и защищённый канал даже на заражённом компьютере.

4.3 Когда подрядчиком становится вендор

Николай Ганюшкин напомнил об атаке через систему мониторинга SolarWinds, затронувшей десятки тысяч компаний, а Денис Назаренко («Сайберлифа») предупредил, что централизованная система управления устойчивостью хранит карты сетей и сведения о критичных процессах и становится приоритетной целью; размещать её в плоской сети нельзя. Махновский назвал два недопустимых для вендора события: закладки в исходном коде с атакой на клиентов через цепочку поставок и утечку данных клиентов. В GitLab, по

его словам, за последние 2 месяца выявлено 10 уязвимостей уровня 9 и выше и 2 уязвимости уровня 10, поэтому репозиторий с основным кодом нужно убирать внутрь периметра с многофакторной аутентификацией.

5. Уязвимости в эпоху ИИ: окно до эксплуатации схлопнулось

ИИ одновременно ускорил поиск уязвимостей и их эксплуатацию, и главным узким местом стала скорость установки обновлений у заказчиков.

5.1 Что изменилось

Александр Буравцов (СБК, CommuniGate Pro) связал резкий рост числа обновлений у крупных вендоров с тем, что ИИ позволил находить больше уязвимостей: по статистике Mozilla Foundation, 271 из 423 найденных и исправленных в её продуктах проблем выявлены с помощью ИИ. Срок от выхода патча до использования уязвимости в реальных атаках раньше составлял 63 дня, пару лет назад сократился до 5 дней, а теперь уязвимости эксплуатируют ещё до официального исправления: в одном из недавних случаев Amazon зафиксировал попытки эксплуатации за 4 дня до патча. В эксперименте Anthropic модель по описанию исправленных уязвимостей без исходного кода воспроизвела 18 из 21 и для 6 создала готовые эксплойты.

Кулашова и Игорь Аникин (КНИТУ-КАИ) оценили сдвиг сдержаннее: новых тактик и техник у атакующих не появилось, ИИ не меняет ни вектор, ни сценарий атак, но даёт колоссальное преимущество в скорости и масштабе. Аникин привёл пример: бесплатная модель Qwen на сильных соревнованиях CTF решает ряд задач, особенно по криптографии, за 5–10 минут. Старовойт добавила, что порог входа резко снизился: раньше злоумышленнику нужно было знать инструменты Kali Linux, теперь атака задаётся промптом на обычном языке, а эксплойты и вредоносное ПО меняют ИИ-агенты. Лев Якупов (TrueConf) считает иначе: ИИ находит уязвимости «совершенно другого уровня» – сложные переполнения буфера, которые человек найти не в состоянии.

Обратная сторона – поток мусора. По словам Кулашовой, Google как минимум временно закрыл программу bug bounty из-за потока сгенерированных ИИ заявок и ложных срабатываний, а в программе «Лаборатории Касперского» около 98% поданных заявок сгенерированы с помощью ИИ.

5.2 Уязвимость закрыта, когда патч установлен

Буравцов предложил считать уязвимость закрытой не в момент выпуска исправления, а когда оно установлено в инфраструктуре заказчика. Заказчики часто не готовы к внеплановым срочным обновлениям: запускаются стандартные процедуры поиска ответственных и тестовой среды, и время уходит. Поэтому срочные обновления стоит репетировать в рамках плановых, а признак активной эксплуатации считать достаточным поводом для внепланового обновления. От вендоров он ждёт понятных бюллетеней:

насколько срочно обновляться, какие версии затронуты, что делать, если быстро обновиться нельзя, и как проверить результат. Он же предостерёг от инфляции тревог: «если очень долго кричать „волки-волки“, то настоящего волка можно пропустить».

Якупов рассказал, что TrueConf несколько лет назад выпускал обновления безопасности раз в квартал, с прошлого года – раз в месяц, а только в октябре вышло 2 таких обновления. После публикации одной из недавних критических уязвимостей компания выпустила исправленный релиз через 25,5 часа, пересобрав его под Windows, разные дистрибутивы Linux, контейнеры и среды виртуализации. Его вывод: «Лучший способ защиты – это иметь обновленную систему», и обновиться важнее, чем сохранить работоспособность для 100% пользователей. Отсутствие продукта в базах уязвимостей, по его словам, значит лишь, что продукт менее распространён, а вынос сервисов в облако прозрачности не добавляет. Якупов ожидает, что нынешний вал найденных ИИ уязвимостей пройдёт пик. Кулашова напомнила, что на технологических процессах вендор не может обновляться с частотой антивирусного движка и должен соответствовать регламентам предприятия, хотя ИТ и ОТ срастаются.

5.3 Open source и вайб-кодинг

Груздев указал на системный риск импортозамещения: российский софт последних 5 лет построен на open source, а ИИ-модели теперь находят в нём десятки тысяч уязвимостей, под которые автоматически создаются эксплойты. Второй по значимости проблемой он назвал вайб-кодинг: на выставке он видел предложения сгенерировать CRM за 3 часа. Сгенерированный код при этом профессионален, но содержит архитектурные уязвимости и захардкоженные пароли. Эльдар Зиганшин («Атомдата-Иннополис») подтвердил, что клиенты ЦОД размещают непроверенный софт, в том числе написанный вайб-кодингом, и оператор вынужден отдельно защищать свою инфраструктуру. Кулашова назвала безопасный цикл разработки по рекомендациям ФСТЭК обязательным для разработчиков.

6. Восстановление: исходить из того, что взлом случится

Многие заказчики уже исходят из того, что атака рано или поздно пройдёт, и вкладываются в быстрое восстановление и живые резервные копии (Дрюков Владимир, Innostage). Спор шёл о балансе между защитой и восстановлением и о том, где хранить копии.

6.1 Баланс защиты и восстановления

Ясовеев описал «треугольник киберустойчивости» Innostage с вершинами «предотвращение», «обнаружение и реагирование» и «восстановление». Перекос в защиту даёт отражение большинства атак, но успешные приводят к неделям простоя и потере данных; обратный перекос – резервная площадка и учения без защиты – даёт слишком много инцидентов, а восстановленные данные могут содержать вредонос. Модель зрелости Innostage включает 3 уровня, от гигиенического до непрерывного, на котором компания

способна оперативно восстановить критичные процессы. Андрей Щербаков (Innostage) сформулировал крайний вариант: если взлом неизбежен, нужно строить «ЦОД судного дня», который поднимется при атаке.

Шепелявый подчеркнул, что «бэкап – это одна из конструкций правильных, но далеко не достаточных»: вкладываться нужно в обучение людей, ЗИП, телеметрию, управляемость и финансовую оценку. Хасин предупредил, что восстановление без устранения причин ведёт к повторным падениям из-за тех же уязвимостей, а некоторые технологические процессы, например доменную печь, быстро перезапустить нельзя. Маслов добавил, что в резервных копиях могут остаться инструменты и учётные записи злоумышленников.

6.2 Где хранить резервные копии

Ясовеев напомнил о зависимости от сервисных партнёров: простои провайдеров защиты, в частности Curator, на время сделали часть бизнес-сервисов недоступными при полностью работающей внутренней инфраструктуре, а при инциденте у японского облачного провайдера быстро восстановились только те, кто вынес резервные копии из облака к себе. Эльдар Зиганшин, напротив, считает, что хранение копий в облачном хранилище решает проблему доступа шифровальщика к бэкапам. Добрушский советует делать копии на инфраструктуре, непохожей на боевую: другая ОС, другой сегмент; в известных ему случаях шифровальщики, атаковавшие Windows-инфраструктуру, не доходили до резервного копирования на Linux. Груздев назвал частую ошибку: бэкап базы 1С лежит в соседнем каталоге, и шифровальщик находит его сразу после базы.

6.3 Опыт критической инфраструктуры

Никита Самойленко (Госкорпорация по ОрВД) описал многоуровневое резервирование: резервные копии корпоративного сегмента защищены так, что даже администратор с полными правами не может их удалить – только по заявке производителю оборудования; при отказе основной системы поднимается холодный резерв с преднастроенными конфигурациями, у диспетчера есть горячий резерв – соседний компьютер, а на случай полного отказа есть независимые запасные пункты управления. Катастрофоустойчивый кластер строится на 3 взаимозаменяемых ЦОДах; для корпоративных сервисов создаётся собственный кластер без облаков. Планы предусматривают наихудший сценарий – переход управления воздушным движением в ручной режим, а в учения теперь включён сценарий атаки с использованием ИИ. Наихудшим на учениях оказалась не авария всей системы, а подмена с помощью ИИ данных о положении самолётов. Первым шагом для любой компании Самойленко считает учения по наихудшему для неё сценарию.

Алексей Полянский (МГТУ им. Баумана) рассказал, что в 2024 году университет вошёл в число вузов, которым пришлось восстанавливать инфраструктуру с нуля. Утверждённый руководством годовой план по ИБ за 2 года дал межсетевые экраны, бюджеты, отдел ИБ и пройденные аудиты ФСБ и ФСТЭК, хотя каждый аудит – это месяцы работы групп специалистов. Полянский предложил заказчикам и вендорам обмениваться планами

восстановления. Ирек Вафин (ИНТ) привёл пример из энергетики: при отключении заражённого энергоблока питание переключают на другую подстанцию, перерыв для потребителей составляет около 5 минут, а распространение вредоноса локализовано.

7. Мониторинг и реагирование: данные, SIEM и автоматизация

Спикеры согласились, что атаки развиваются быстрее, чем человек принимает решение, но автоматическое реагирование допускают только по проверенным сценариям, а главным препятствием считают не технологии, а качество данных.

7.1 «Ад данных» и бесконечный SIEM

Назаренко назвал частым заблуждением промышленных заказчиков фразу «у нас всё мониторится»: мониторится всё подряд без приоритизации и связи с процессами. Проблема, по его словам, не в метриках, а в «аде данных»: данных много, метрик мало, им не доверяют, а на производстве самая достоверная информация иногда хранится в Excel. Ганюшкин поддержал: «хаос на входе дает хаос на выходе». Одна компания собирала терабайты логов в день в расчёте, что ИИ потом найдёт аномалии, но без нормализации данные не помещались даже в контекстное окно. Хабибуллин советует начинать с 3–5 фокусных метрик: на одном проекте управленческий дашборд с сотнями метрик постоянно горел красно-жёлтым. Ганюшкин добавил, что «если инженер сидит и смотрит дашборды, это плохой результат», а долгое отсутствие плохих событий само по себе должно настораживать.

Лагоденко назвал внедрение SIEM бесконечным проектом с постоянными OPEX и CAPEX, а подход «соберём всё в Data Lake, потом решим» упирается в лицензии, ротацию логов и стоимость хранения. Кравцов был резче: SIEM, в который данные только стекаются, защищает «максимум от регулятора». Хабибуллин подытожил: «мониторинг – это не проект, мониторинг – это процесс».

Ганюшкин оценил рынок: в России около 15 систем мониторинга, но на отечественные решения перешло лишь около 30% рынка. Назаренко назвал Zabbix «условно бесплатным»: работу вендора по обновлениям выполняет своя команда. Ганюшкин добавил, что самосборная платформа держится на двух-трёх инженерах, и если их переманят, компания остаётся с ней один на один. Кравцов предупредил, что сократить штат при переходе на коробочный мониторинг, как правило, не удаётся.

7.2 Что можно отдать машине

Дмитрий Семидоцкий (Security Vision) предложил автоматизировать в первую очередь понятные сценарии, а сценарии с неопределённостью и принятием риска оставлять человеку; условие автоматизации – логирование и возможность отката автоматически выполненных действий. Вафин перечислил, что можно передать ИИ: детекцию аномалий, автотриаж, обогащение данных и некритичные действия – изоляцию хоста, блокировку

учётных записей, сбор форензики, черновики отчётов. Остановку техпроцесса, отключение критичных сервисов и изоляцию сегмента АСУ ТП он считает решениями исключительно для человека.

Самойленко допускает автоматическое реагирование только по подтверждённым плейбукам и считает преждевременной автоматическую блокировку на IPS: легальное приложение может сгенерировать трафик под сигнатуру. Кравцов назвал перевод правил из режима обучения в режим блокировки сложным решением, которое каждая компания принимает по своему уровню допустимого риска. Бейбутов предложил различать нарушение политики и инцидент: многие SOC поставили на поток отчёты о нарушениях, а о работе шифровальщика узнают постфактум. Обращение хоста к известному командному серверу, по его мнению, – уже инцидент, и хост желательно изолировать автоматически. Для этого SOC нужны коннекторы к бизнес-системам, например к системе командировок, а плановые работы сетевиков должны попадать в SOC, чтобы обслуживание не принимали за атаку. В Межгосударственном банке ИБ заранее знает о регламентных работах (Кравцов Александр, Межгосударственный банк), а в Госкорпорации по ОрВД все технические работы согласуют 3 подразделения, включая владельца бизнес-процесса (Самойленко Никита, Госкорпорация по ОрВД).

Бейбутов также советует заранее разделить системы по критичности и закрепить мандаты на быстрые решения: если все вопросы адресовать гендиректору, быстрой реакции не будет. Вафин добавил, что сотрудник не последний рубеж: если он перешёл по фишинговой ссылке, заражение должны остановить эшелонированная защита и дежурная смена. Когда у заказчика коммерческого SOC нет ответственного за реагирование, выполнение рекомендаций растягивается на сутки и даже недели (Старовойт Светлана, ИнфоТеКС).

8. Agentic SOC: автономии нет, спор идёт об экономике и людях

Дискуссия об агентных SOC показала, что практики относятся к автономии скептически: ИИ уже помогает разбирать поток событий, но решения остаются за человеком, а экономия от замены первой линии не доказана.

8.1 Где ИИ уже работает

В SOC «Лаборатории Касперского» около 25% событий разбирает ИИ, остальное уходит на дополнительный анализ; ИИ стоит на первой линии как второе мнение, а финальный вердикт выносят люди, потому что агентная модель «недоверенная по определению» (Кулашова Анна, «Лаборатория Касперского»). Старовойт рассказала, что в центре мониторинга ИнфоТеКС ИИ-агент для триажа с объяснениями и рекомендациями существенно сократил время на ложные срабатывания, которые раньше занимали почти половину времени специалистов; первый продукт ИнфоТеКС с машинным обучением для анализа событий вышел 12 лет назад. Игорь Залевский (Innostage) описал случай, когда агент разобрал бинарный файл вредоноса, извлёк индикаторы компрометации и написал

отчёт за 4 минуты на домашней видеокарте. Лагоденко отметил, что за последний год число срабатываний в SOC значительно выросло почти у всех компаний, в том числе из-за использования ИИ сотрудниками.

8.2 Спор об автономии

Залевский определил agentic SOC как момент, когда ИИ снимает линейную зависимость штата от нагрузки: плюс 10 источников и 200 правил больше не требуют найма людей на первую, вторую линии и detection engineering. Сергей Исхаков (ПСБ) назвал ИИ «продвинутой автоматизацией» и заявил, что передача первой линии агентам – самообман: карточек в SOC стало слишком много, большинство ложноположительные, а людей не увольняют, а переводят на вторую линию, формально выполнив KPI. Его итог: «к автономии мы не приблизились ни на йоту, это факт». Залевский согласился, что полной автономии в SOC не будет никогда.

Антон Байтала (ТрансТелеКом) считает, что для рутинных операций ИИ не обязателен: классическое машинное обучение давно себя зарекомендовало и стоит в разы дешевле, а блокировку хоста можно прописать в плейбуке SOAR. Ценность агентов – в вариативности: обогатить контекст и предложить несколько гипотез. Байтала настаивает, что решение по инциденту принимает человек, а если компания доверяет агенту действия в инфраструктуре, решение и ответственность нужно зафиксировать на бумаге: неясно, кто понесёт административную ответственность, если агент не выполнит операцию и НКЦКИ не будет уведомлён вовремя. Алёна Егорова (ЦПЛ) спросила, кто будет сопровождать навыки агентов, и напомнила, что агенты работают поверх имеющейся автоматизации, а та – поверх процессов, которых во многих компаниях нет.

8.3 Экономика и инфраструктура

Залевский указал на ловушку: лицензии SIEM привязаны к объёму данных, а ИИ требует как можно больше данных, поэтому экономия может обнулиться ростом платы за SIEM; выход – общее с ИТ озеро данных или Security Data Lake перед SIEM. Экономию от замены первой линии он считает неочевидной: придётся платить либо за GPU, либо за токены. Егорова отметила, что модель с LLM, SOAR, людьми и агентами по карману далеко не каждой компании, а один из участников дискуссии заметил, что маленькой компании дешевле нанять двух специалистов.

ТрансТелеКом продвигает локальные модели: первую модель, развёрнутую без настроек, собственные специалисты «завалили» за 4 часа; российские, американские и китайские модели на схожих данных дали совершенно разные результаты, а следующим шагом компания видит малые модели, обученные только на задачах кибербезопасности (Байтала Антон, ТрансТелеКом). По подсчётам на SOC-форуме, в России разрабатывается 20 решений класса AI firewall, но единого понимания этого класса нет, как раньше с XDR (Залевский Игорь, Innostage); собственный AI firewall ПСБ уже готов к внутренним пилотам (Исхаков Сергей, ПСБ). Залевский допускает отправку аккуратно очищенных логов зарубежным моделям, признавая, что многие с ним не согласны; Исхаков считает это

неприемлемым для in-house SOC крупного бизнеса без маскирования, тем более что в прикладных логах отечественных импортозамещающих систем встречаются пароли в открытом виде.

8.4 Люди

Самый жёсткий прогноз дал Залевский: через 3 года 30% специалистов отрасли останутся без работы. Исхаков и Байтала заявили, что увольнять никого не планируют; Байтала добавил, что привлекать новых людей всё тяжелее, а вузы отстают: практиканты приходят через 2–4 года. Начинающие специалисты, по словам Исхакова, не исчезнут, а ключевым навыком станет умение создавать и тренировать агентов. Байтала предупредил, что специалист, умеющий только общаться с моделями, априори им верит, и тогда human in the loop теряет смысл. Залевский видит и плюс: ИИ может снять монотонную работу с логами, из-за которой у сотрудников SOC самое сильное выгорание.

9. ИИ-агенты в компании: доступы, контроль и теневой ИИ

ИИ-агенты внедряются быстрее, чем выстраивается их защита, и спикеры предложили относиться к агенту как к новому сотруднику с минимальными правами, внешним контролем и кнопкой экстренного отключения.

9.1 Масштаб и экономика внедрения

Шепелявый со ссылкой на исследование EY 2026 года привёл цифры: ИИ-агентов применяют 91% компаний, 85% допускают их действия без участия человека, почти половина ради скорости отказалась от собственных процедур, а около 40% сообщили регуляторам о существенных инцидентах с действиями агентов.

Святослав Соловьев (Альфа-Банк) рассказал, что ИИ-агенты в валютном контроле за полгода достигли точности 95%, что позволило оптимизировать десятки сотрудников. ИИ-интерфейс в мобильном приложении раскатан более чем на 100 000 клиентов, до конца года их будет около 1 млн. Внутренние модели в закрытом контуре со своими GPU и средствами защиты в 10–40 раз дороже внешних, поэтому данные категорируют: самые чувствительные обрабатывают внутри, остальные – на внешних моделях через шлюзы с маскированием. Рабочая группа при Минцифры России с участием крупных организаций строит экономическую модель стоимости инференса и до конца года должна опубликовать «тепловую карту» зрелости ИИ-агентов по областям бизнеса. При этом прохождение кибербезопасности для нового ИИ-решения в банке занимает около 3 месяцев.

9.2 Агент как субъект доступа

Бейбутов предупредил, что агент с доступом почти ко всем бизнес-системам становится суперадминистратором, и компании должны понимать, кто им управляет и к чему он реально имеет доступ. Антон Башарин (Swordfish Security) предложил относиться к агенту как к новому сотруднику: сначала чтение, поиск и суммаризация, а частые подтверждения

его действий – сигнал расширить полномочия через белый список операций. Кравцов назвал порог: если в режиме обучения пользователи одобряют около 95% действий агента, функцию можно отдать ему; доступы выдаются в минимально необходимом объёме. Али Шабанов (ЦИТ РТ) выдаёт агентам временные токены под каждую задачу по принципу «запрещено всё, что не разрешено»; сотрудники при этом даже просят передать агенту свою электронную подпись. Альфа-Банк переходит для агентов от ролевой модели доступа к модели на основе задач и строит реестр всех агентов с регрессионными тестами, контролем дрифта и kill switch (Соловьев Святослав, Альфа-Банк).

Айдар Фатыхов (Innostage) перечислил меры допуска агента в продуктив: песочница с ограниченным доступом к ресурсам, централизованный шлюз для всех сетевых обращений агента, включая модели и MCP-инструменты, гардрейлы против промпт-инъекций и джейлбрейков, квоты (агент может зациклиться и сам устроить DoS), человек в цикле для необратимых изменений и наблюдаемость: «Мы не можем контролировать, не можем защищать то, чего мы не мониторим». Кравцов не доверяет встроенным в агента сторожевым механизмам: при компрометации модели они бесполезны, нужна внешняя поведенческая аналитика и kill switch. Лагоденко добавил, что часть операционных сотрудников нужно сохранить, иначе после отключения агента резервный сценарий выполнять будет некому.

Башарин предупредил, что агенты «убегали» из песочниц даже у грамотных команд, а атаки на модели эволюционировали от промпт-инъекций к постепенной «раскачке контекста». Соловьев указал на цепочку поставок: непроверенные навыки из сторонних библиотек могут содержать трояны, а прокси-провайдеры в третьих странах предлагают трафик в 40 раз дешевле официального без гарантий, что модель не подменят. Кулашова попросила руководителей не поручать агентам разбор почты: письмо может содержать промпт-инъекцию, ведь «агент – это не самый умный, но очень активный, никогда не спящий, постоянно трудящийся подрядчик».

9.3 Теневой ИИ

Запрет не работает, сходились спикеры: даже при изоляции контура пользователь сфотографирует экран и отправит снимок во внешнюю модель (Лагоденко Андрей, Домклик). Валентина Степанова («ИнфоЛинк») сформулировала то же иначе: организация может вложить 1 000 000 рублей в защиту, но сотрудник одним запросом к нейросети сольёт конфиденциальные данные, а «если сотрудникам полностью запретить использовать внешнюю нейросеть, это так не работает». Исхаков советует находить таких пользователей и предлагать альтернативу, а не ругаться.

Альтернативы описали несколько компаний. Домклик разрешает работу с мощными внешними моделями через шлюз с обфускацией и гардрейлами, а клиентские данные обрабатывает только на выданной модели (Лагоденко Андрей, Домклик). Межгосударственный банк опросил пользователей, составил альбом процессов и перевёл процессы с чувствительными данными на внутренние модели; драйвером внедрения ИИ там стала служба ИБ (Кравцов Александр, Межгосударственный банк). «ИнфоЛинк» запустил локальную транскрибацию совещаний на сервере с видеокартой на 16 ГБ, а

локальные агенты маскируют чувствительные данные перед отправкой во внешнюю модель и выполняют обратную подстановку примерно за 2 секунды (Степанова Валентина, «ИнфоЛинк»).

Сергей Кравченко (ICL) добавил, что ИТ-компании часто внедряют ИИ, но не защищают его, а коробочных решений для защиты ИИ нет; организациям в зоне жёсткого регулирования он советует только суверенные национальные модели.

10. Регулирование и ответственность: приказ ФСТЭК № 117 и «человек в контуре»

Приказ ФСТЭК № 117 стал центральным нормативным сюжетом направления: он переводит безопасность ИИ из хороших практик в обязательные требования, хотя ряд ключевых понятий ещё не определён. В вопросе ответственности за действия ИИ позиции спикеров совпали – отвечает человек.

10.1 Что уже требует регулятор

Фатыхов напомнил, что приказ № 117 вступил в силу 1 марта 2026 года, и к нему уже опубликован проект изменений, который вводит агентские системы как отдельный объект защиты. В банке данных угроз ФСТЭК появился раздел по защите ИИ – открытая базовая модель угроз; методика анализа защищённости описывает аудит систем на базе ИИ, включая полную инвентаризацию моделей, агентов и инструментов (AI-BOM). Методические рекомендации требуют контролировать промпты и ответы моделей, подключать ИИ-инфраструктуру к мониторингу ИБ и выводить большие языковые модели в отдельный сетевой контур со шлюзами безопасности. Аникин считает, что приказ перевёл защиту систем с ИИ в нормативные требования, включая контроль запросов и ответов агентов и оценку достоверности решений.

Шепелявый описал глобальный контекст: в 2026 году регулирование ИИ смещается от алгоритмов к действиям агентов – правам доступа, порогам автономии и участию человека. ЕС движется к обязательным регламентам, США – к децентрализации, Китай – к разрешительному порядку с регистрацией и сертификацией моделей; Россия, по его мнению, ближе к китайской модели, хотя пока регистрирует и сертифицирует только медицинские изделия с ИИ. Груздев прогнозирует, что требования приказа № 117 станут базовыми не только для ГИС, но и для КИИ, значимых объектов КИИ и ИСПДн; приказ также вынес носители за периметр, поэтому данные на ноутбуках и флешках должны быть зашифрованы. Груздев напомнил, что повторная утечка большого объёма персональных данных может обернуться оборотным штрафом 1–3%, а требования разных регуляторов – ФСТЭК, ФСБ, Роскомнадзора, Банка России – часто конфликтуют, и тогда действует самое строгое.

10.2 Чего пока нет

Обновляемый приказ требует применять доверенные ИИ-модели, но критерии доверенности не утверждены (Шепелявый Дмитрий, Innostage). Башарин уточнил, что ФСТЭК отказалась от термина «доверенный ИИ» в самом приказе, а в готовящихся изменениях посвящённый ему пункт убран; ГОСТ 2020 года о доверенных системах ИИ написан до генеративных моделей и требует объяснимости, которую для них реализовать очень тяжело. В Минцифры России обсуждают 3 базовых критерия доверенного ИИ – безопасность, функциональность и надёжность – и формулирование недопустимых действий ИИ. Требования к ИИ на объектах КИИ содержатся в пунктах 60 и 61 приказа № 117, и пункт 60 будет переписан; Альфа-Банк пока не использует большие языковые модели на объектах КИИ и обсуждает это с Банком России (Соловьев Святослав, Альфа-Банк). Кравченко напомнил, что юридически система ИИ сейчас – такая же информационная система, как любая другая, и, по его словам, закон о безопасном использовании систем ИИ, пока существующий как законопроект, вступит в силу в 2027 году, часть статей – в 2028-м.

10.3 Кто отвечает за действия агента

На вопрос, кто отвечает за ошибку агента, Башарин ответил: «сядет, конечно, человек» – тот, кто делегировал агенту полномочия, ведь «всегда курок нажимает человек». Кравцов сравнил агента с Metasploit: решение применить инструмент принимает человек. Соловьев предложил заменить слово «агент» на «виртуальный сотрудник»: тогда за утечку отвечает директор по кибербезопасности, за финансовые потери – владелец бизнес-линии. Лагоденко считает, что привлечь к ответственности разработчиков моделей вряд ли получится, тем более что большинство из них нерезиденты, поэтому возможный ущерб от ошибок агента бизнес должен заранее закладывать в бюджет операционных рисков. Шепелявый предложил страховому рынку страховать действия агентов по аналогии с гражданской и профессиональной ответственностью и ожидает, что риски ИИ войдут в банковские нормативы непрерывности.

Шепелявый назвал передачу ИИ ответственности за устойчивость на уровне совета директоров фантастикой: регулирование оставляет окончательное решение за человеком, а если действие защитного ИИ может привести к недопустимому событию, решать должен человек. Хасин напомнил, что регулирование ответственности за беспилотный транспорт не сформировано за 9 лет, поэтому к 2035 году вопрос ответственности ИИ вряд ли будет решён. Фатыхов отметил, что в США уже разрабатывают законопроекты о компаниях без сотрудников, где работают только агенты. Одним условием безопасной ИИ-трансформации Соловьев назвал «human in the loop для всех критических решений», Кравцов – управляемость: «Очень важно, внедрив такую агентную систему, не потерять управление».

11. Импортозамещение и технологический суверенитет: от «замещать всё» к управляемости

Спор об импортозамещении в ИБ шёл не о том, нужно ли оно, а о том, что считать результатом: число замещённых продуктов или реальную защищённость от внешнего отключения и атак.

11.1 «Красная кнопка» и цена инерции

Груздев предложил остановиться спустя 4–5 лет бурного импортозамещения: заместить «всё и вся» не хватит ни сил, ни времени, ни денег, а цель – технологический суверенитет, то есть отсутствие за океаном «красной кнопки», блокирующей всё построенное. Ошибкой он считает знак равенства, при котором «российское значит безопасное и российское значит из реестра». Груздев связывает риск с американскими документами PPD-20 и CLOUD Act, которые, по его словам, обязывают вендоров закладывать сбор и передачу телеметрии, – и с отзывом сертификатов, который в начале СВО нарушил цепочку доверия в российском интернете. Отключение контроллера домена или центра сертификации из-за рубежа, по его оценке, может положить инфраструктуру в течение суток, поскольку билет Kerberos живёт 24 часа. Заместить Microsoft AD и Certificate Services нужно прежде всего санкционным компаниям, включая системообразующие банки, и это, по Груздеву, «вопрос времени».

Махновский возразил: централизованной «красной кнопки» для отключения KDC в каждой организации нет, а реальная глобальная угроза – возможность ассоциации GSMA отозвать сертификаты мобильных операторов. Он спросил, адекватно ли сейчас бежать заменять Microsoft с деградацией функциональности и непонятными последствиями для ИБ, и напомнил, что в Samba, на базе которой российские производители делают службы каталогов, за 2026 год найдено больше критичных уязвимостей, чем за 5 предыдущих лет, а сведений о них в БДУ ФСТЭК нет. Рустам Рустамов (РЕД СОФТ) ответил, что продукты Microsoft могут перестать обновляться, а спор нужно решать практикой; по его словам, службой каталогов РЕД СОФТ пользуются крупные банки и нефтегазовые компании. Сергей Добрушский оценил, что 80–90% компаний остаются на Microsoft AD, и назвал неочевидный плюс перехода: менее изученная злоумышленниками область.

11.2 Open source: опора или риск

Зиганшин заметил, что при импортозамещении все побежали в open source, а «open source у нас это то, за что никто не отвечает»; импортозамещение – это платные решения на его основе. Рустамов объяснил, зачем open source нужен: уйти от возможности выключить системы из-за рубежа, ведь остановка обновлений – тоже способ давления. Но «голый» open source опасен: нужна команда, которая сопровождает и патчит его независимо от основного ядра. Разработчик ОС оказывается «между молотом и наковальней»: прикладное ПО требует расширять пакетную базу, а любая несовместимость останавливает работу. На опубликованную в день сессии уязвимость ядра у РЕД СОФТ уже был готов патч. Быстро отказаться от иностранных enterprise-решений невозможно, поэтому российские продукты

строят с расчётом на совместную работу с ними. Зиганшин оценил, что ЦОД «Атомдата-Иннополис» на 80% российский по составу, и признал, что по удобству российские офисные пакеты во многом пока не дотягивают.

11.3 Сертификат или bug bounty

Махновский считает, что называть сертифицированный софт защищённым – профанация. Пентест при сертификации одного из решений «Аванпоста» выявил 5 IDOR низкого уровня, а за первые полгода публичной программы bug bounty нашли десяток уязвимостей уровня 8,5 и выше; в публичной программе, по его словам, вендор отвечает своими деньгами. Груздев предложил компромисс: безопасность должна подтверждать третья доверенная сторона – сначала сертификация в независимой лаборатории, затем публичный bug bounty; требования ФСТЭК к сертификации за последние годы сильно выросли и не дают вендорам расслабляться. Один из участников дискуссии добавил, что сертификат получают на версию продукта одного года, а поставляются версии следующих лет, безопасность которых неизвестна; Лорич указал, что сертификация долгая и редкая. РЕД СОФТ сейчас проходит публичную программу bug bounty для своей СУБД (Рустамов Рустам, РЕД СОФТ).

Выбор продукта, по словам Груздева, начинается не с каталога, а с выяснения классов систем, контуров и регуляторов. Лорич предупредил о заблуждении «технология равно процесс»: антивирусная защита – это процесс, а купленный продукт без человека, ведущего политики, превращается в «диски на полке». Рустамов и Добрушский призвали сразу планировать эксплуатацию, обучение и техподдержку и учитывать совокупную стоимость владения. Андрей Финохин (Айтисфера) сформулировал общий принцип: «Ты сначала процесс положи как-то на бумагу, а потом уже оцифровывай, иначе будет оцифрованный бардак».

12. Исследования вузов: интерпретируемый ИИ, АСУ ТП и биометрия

Научная секция показала, что вузы работают над теми же проблемами, что и практики, – непрозрачностью моделей, оценкой рисков и защитой данных, – и предлагают для них прикладные инструменты.

12.1 Интерпретируемость и оценка рисков

Аникин назвал главной проблемой моделей машинного обучения в кибербезопасности «чёрный ящик», что противоречит требованию регулятора о доверии к решениям ИИ. Кафедра КНИТУ-КАИ развивает гибридные нейро-нечёткие системы, где интерпретируемые базы правил строятся по данным, а генетические алгоритмы сокращают число правил без потери точности. Большинство показателей ущерба в методиках категорирования КИИ и моделирования угроз качественные – например, жизнь и здоровье трудно выразить в рублях, – поэтому их оценивают методами нечёткой логики и анализа иерархий. Нечёткая модель кафедры применялась для оценки опасности уязвимостей как аналог CVSS: размытие результата показывает степень доверия к оценке. Генератор гаммы

для потоковых шифров на нечёткой логике, по словам Аникина, после настройки превзошёл все 16 генераторов из набора Diehard. Аникин также отметил, что значительная часть научных работ этого года посвящена сертифицируемой робастности систем ИИ.

Лина Сафиуллина (КНИТУ) указала, что риски ИБ в АСУ ТП оценивают раз в квартал или полгода, а в перерабатывающей промышленности периодичность обновления оценок может достигать 10 лет, хотя угрозы меняются непрерывно. Она предложила архитектуру киберустойчивости АСУ ТП на трёх опорах: динамическая оценка рисков на нечёткой когнитивной карте, веса которой корректируются по телеметрии ПЛК, цифровой двойник для проверки защитных мер (одно ошибочное правило межсетевого экрана на живом производстве может остановить линию) и автоматическая оркестрация мер. По её прогнозу, это сократит время реакции с минут до секунд; прототип проверят на испытательном стенде КНИТУ и в пилотах с промышленными партнёрами.

В докладе исследователей Уфимского университета науки и технологий описана система для SOC, которая на графе знаний связывает БДУ ФСТЭК, MITRE ATT&CK и отчёты кибербразведки и формирует ответ аналитику с помощью языковой модели; по оценке авторов, типизированный граф с указанием происхождения данных снижает галлюцинации лучше, чем RAG.

12.2 Биометрия

Эллина Анисимова (КНИТУ-КАИ) рассказала о 4 бакалаврских работах по нейросетевой биометрии, доведённых до программных комплексов. Система проверки рукописной подписи на сиамской нейросети показала вероятность ложного допуска 3,3% и ложного отклонения 6,2%, проверка живости лица обнаруживает атаки предъявления с точностью 96%, аутентификация по радужной оболочке – с точностью 95,4% и ложным допуском 2,1%. Сервер хранит только хэши и биометрические векторы, а не изображения. Вероника Гончаренко (КНИТУ-КАИ) сообщила, что адаптированная модель аутентификации по походке на эталонном датасете достигла ориентировочной точности 98,94%, но на реальных видео качество ниже из-за сложности выделения силуэтов; классический параметрический метод даёт равный уровень ошибок 11%. Аутентификация по походке может работать вторым фактором в СКУД: после прикладывания карты камера подтверждает, что идёт её владелец. Анисимова назвала главными вызовами устойчивость к состязательным атакам, снижение ложных срабатываний и соблюдение 152-ФЗ, а перспективами – мультимодальную проверку и федеративное обучение без передачи данных на сервер.

Пётр Филяк (РЭУ им. Плеханова) обратил внимание на массовое производство генеративным ИИ фейковых новостей и видео, воздействующих прежде всего на молодёжь. Дрюков на другой сессии признался, что сам, будучи опытным безопасником, 3 месяца назад принял за настоящие сгенерированные ролик и профиль в соцсети.

Выводы

- Доступ подрядчиков стал одним из главных каналов атак: их доля выросла примерно с 15% в 2024 году до 30% в прошлом году, а злоумышленник находится в инфраструктуре в среднем около 42 дней (Маслов Дмитрий, BI.ZONE). Ответ спикеров – Zero Trust на уровне приложений, временные учётные записи и учёт ИИ-агентов, которые продолжают работать после ухода подрядчика (Борисов Илья, Билайн).
- Окно между патчем и атакой почти исчезло: срок до эксплуатации сократился с 63 до 5 дней, а 271 из 423 уязвимостей в продуктах Mozilla найдены с помощью ИИ (Буравцов Александр, СБК). Узким местом стала скорость установки обновлений у заказчиков, а не их выпуск у вендоров.
- Эффективнее всего вкладываться в гигиену: 88% успешных атак связаны со слабой идентификацией и аутентификацией (Груздев Сергей, «Аладдин»), около 44% начинаются с уязвимостей в общедоступных приложениях (Кулашова Анна, «Лаборатория Касперского»), тогда как 63–65% опрошенных специалистов первым делом купили бы NGFW.
- Устойчивость работает там, где она поднята до первого лица и измеряется: в «Газпром нефти» система непрерывности действует 7 лет, а каждый план тестируется 2 раза в год (Лутфиев Рамиль, «Газпромнефть ИТО»).
- ИИ-агенты внедряются быстрее защиты: 91% компаний их применяют, 85% допускают их действия без человека (Шепелявый Дмитрий, Innostage, со ссылкой на EY). Спикеры предлагают минимальные права, временные токены под задачу, внешний мониторинг и kill switch, а ответственность за действия агента оставляют за человеком.
- Автономного SOC нет: даже в SOC «Лаборатории Касперского» ИИ разбирает около 25% событий (Кулашова Анна, «Лаборатория Касперского»), а экономия от замены первой линии упирается в стоимость GPU, токенов и лицензий SIEM (Залевский Игорь, Innostage).
- Импортозамещение в ИБ упирается в службы каталогов: 80–90% компаний остаются на Microsoft AD (Добрушский Сергей, «Сайберпик»), а спор о том, что надёжнее, спикеры предлагают решать сертификацией и публичным bug bounty, а не принадлежностью к реестру.

Источники

- Сессия «Экономически эффективная цифровая устойчивость бизнеса», Kazan Digital Week 2026
- Сессия «Управление устойчивостью бизнес-процессов», Kazan Digital Week 2026
- Сессия «Актуальные подходы обеспечения киберустойчивости», Kazan Digital Week 2026
- Сессия «Практическая безопасность внедрения ИИ», Kazan Digital Week 2026
- Сессия «Agentic SOC. Насколько мы приблизились к автономии», Kazan Digital Week 2026

- Сессия «Киберриски в эпоху ИИ: что угрожает бизнесу и как выстроить надёжную защиту», Kazan Digital Week 2026
- Сессия «Управляемость в эпоху импортозамещения», Kazan Digital Week 2026
- Сессия «Интеллектуальные системы защиты информации», Kazan Digital Week 2026