Engineer Консультування

Кібербезпека ШІ-агентів вашої компанії

Перевірка доступу ШІ-агентів: погоджений перелік, висновки за ризиком, випробування прав і записи перевірок.

З’ясуйте, до чого мають доступ ваші ШІ-агенти, і встановіть контроль. Агент, під’єднаний до пошти, документів чи API, додає права й можливі шляхи помилок, які варто перевірити. Ми описуємо погоджені системи, оцінюємо ризики, впроваджуємо схвалені засоби контролю та випробовуємо обмежені дії. Ви отримуєте карту доступу, висновки й записи перевірок.

Перетворіть послугу на результат

Нехай «Кібербезпека ШІ-агентів вашої компанії» слугує меті бізнесу. Ми можемо пов'язати технічну роботу з користувачами, експлуатацією та рішеннями, які вона має підтримувати. Доступ кожного агента описано й обмежено необхідним, кожну дію записано в журнал, якого агент не може змінити, і кожен захист перевірено на ділі, а не взято на віру.

Переглянути наші послуги Зв'язатися з нами →

Розгляньте відповідні роботи

Використайте роботи на цій сторінці, щоб оцінити наш підхід. Запитайте про компроміси та передумови кожного результату. Запровадила безперервне сканування безпеки — 21 письмових аудитів безпеки поряд з автоматичними

Запровадила безперервне сканування безпеки Зв'язатися з нами →

Перелічено: кожен агент і все, до чого він має доступ

Описуємо агентів у погодженому обсязі: схвалені інтеграції та інші з’єднання, виявлені разом із вашими працівниками. Для кожного карта містить облікові дані, інструменти й доступні дані. Ви отримуєте карту з висновками; відповідальність за її оновлення погоджуємо під час передачі. Межі перевірки та прогалини зазначаємо явно.

Замкнено: кожен агент із найменшими привілеями

Кожен агент отримує власний обліковий запис і найменший привілей, потрібний для його роботи: доступ через ролі (RBAC), а не через спільні акаунти, лише читання за замовчуванням, а запис — тільки через вузькі інструменти з лімітами на частоту й обсяг. Секрети розділено за агентами й хостами, вони ніколи не потрапляють у промпт і змінюються без правок у коді. Зовнішній текст, який читає агент, — лист, вебсторінку чи документ — обробляємо як недовірені дані. Права інструментів і перевірки забезпечують погоджені межі; повідомлення «ігноруй свої правила» не є дозволом на дію. Захист працює за принципом fail-closed: якщо перевірку неможливо виконати, дію відхиляють.

Записано: що агент зробив і для кого

Кожен виклик інструмента записується в журнал: для кого діяв агент, про що його попросили, чого він торкнувся і що отримав у відповідь, — а журнал зберігається там, де сам агент не може його змінити. Сповіщення налаштовуємо за погодженими для системи правилами. Вони можуть охоплювати сплеск читань, роботу поза очікуваними годинами або звернення до даних поза погодженими для агента межами доступу. Коли щось усе ж іде не так, саме журнал перетворює інцидент на чітку хронологію замість здогадок.

Карта, замок і журнал **лишаються лише словами** — доки хтось не спробує їх обійти.

Узгоджені засоби захисту перевіряються на репрезентативних сценаріях доступу, зловживань і відновлення до передачі результатів.

Замок перевіряють, а не довіряють йому

Випробовуємо засоби контролю доступу в межах завдання й записуємо, що перевірено, що відхилено та що потребує уваги. Тести можуть охоплювати заборонені виклики інструментів, ліміти запитів, ізоляцію облікових даних та

спроби ін’єкції інструкцій. Результат — записи перевірки й повторювані тести з описаними межами. Роботу з інцидентом і ротацію облікових даних погоджуємо окремо, коли потрібно локалізувати наявний витік.

Робота, що за цим стоїть

Ми виконували цю роботу на серверах, у системах керування доступом і на платформах. Ми написали 1000 правил RBAC для Linux-серверів і серверів баз даних, керували доступом у M365 і системі єдиного входу (SSO) і вели безперервну програму сканування безпеки з 21 письмовим аудитом. Ми розділили кожен експлуатаційний секрет за хостами, додали односекундну попередню перевірку секретів, знайшли на дверях три замки, усі три відімкнені, і довели, що блокування за спробу вторгнення спрацьовує під час кожного прогону. Ми побудували обмеження частоти запитів за принципом fail-closed і написали правило розмежування доступу після витоку облікових даних. Як ми дбаємо про безпеку власних систем — на сторінці як ми працюємо безпечно.

Як іде співпраця

Усе починається з доступу лише на читання. Ми наносимо на карту агентів у погодженому обсязі, їхні облікові дані й те, до чого вони відкривають доступ, і передаємо вам знахідки, впорядковані за ризиком, — кожну з причиною й виправленням. Потім вносимо виправлення у вашу конфігурацію й репозиторій — через ваше рев’ю і з перевірками, що завершуються помилкою, якщо на охоплених шляхах доступ перевищує погоджені права. Карта, журнал і тести лишаються у вас: саме їх ви передаєте аудиторові чи страховикові. Так само можна розмежувати й дані компанії, які читають агенти. Замовляють цю роботу, як і решту наших послуг, — на умовах, описаних на сторінці скільки це коштує.

Якщо агенти вже мають ключі від ваших систем, розкажіть, від яких саме.

Кожен ваш агент **тримає ключ до чогось** — знайте, який саме і від чого.

Починаємо в режимі лише для читання: карта агентів та облікових даних у погодженому обсязі, з висновками за рівнем ризику.

Попросіть карту агентів Скільки це коштує →

Визначте обсяг до замовлення

Опишіть шлях розгортання, експлуатаційні ризики та чергову команду. Ми можемо визначити зміну з перевірками, шляхом відновлення та чіткою відповідальністю.

Дивитися, скільки це коштує Зв'язатися з нами →

Обговоріть свої вимоги

Розкажіть, що вам потрібно за напрямом «Кібербезпека ШІ-агентів вашої компанії», зокрема про строки та команду, яка підтримуватиме рішення. Обговоримо практичний шлях уперед.

Переглянути наші послуги

welcome@engineer.company +45 50 10 61 84 Копенгаген, Данія

Візьміть цю сторінку із собою

Візитівка

Дбайливо зроблено з любов'ю — D.

CVR 45969703 · Tagensvej 105C, 3. th, 2200 København N

Авторське право © 2025 – дотепер · Engineer ApS · Усі права захищено.

Надруковано з https://advisory.engineer.company/uk/ai-security/