Бяла книга за сигурността

Enterprise Security Overview - AI Interview Analyzer

Доставчик:	AI Interview Analyzer Sp. z o.o.
Адрес:	ul. Jedrusik 6/53, 01-748 Warszawa, Poland
NIP:	5253079974
REGON:	54402118500000
Класификация:	PUBLIC
Дата:	24.06.2026

Contents

Бяла книга за сигурността

Доставчик: AI Interview Analyzer Sp. z o.o., Warszawa, Poland

Аудитория: Екипи по корпоративна сигурност, ИТ и обществени поръчки

Класификация: Публично

1. Резюме за ръководството

AI Interview Analyzer е корпоративна платформа за наемане на служители, която записва интервюта с изричното съгласие на кандидата, транскрибира и структурира съдържанието им и предоставя подкрепа за оценяване, основана на доказателства, за специалистите по подбор. Тъй като платформата обработва лични данни на кандидати и подпомага процеси по наемане, сигурността и поверителността се разглеждат като основни проектни ограничения, а не като функции, добавени впоследствие.

Тази бяла книга описва в конкретни и проверими термини как защитаваме данните на клиентите и кандидатите. Тя е написана за хората, които извършват оценка на доставчици: инженери по сигурност, ИТ администратори, длъжностни лица по защита на данните и специалисти по обществени поръчки. Всяка стойност в този документ е извлечена директно от нашите собствени инженерни системи, а не от маркетингови материали.

Основното послание е просто: **ние не просто твърдим, че платформата е сигурна, а непрекъснато проверяваме, че това е така.** Нашата кодова база съдържа **3,171 автоматизирани теста**, включително специализиран пакет за сигурност, който упражнява автентикация, оторизация, изолация между организации, защити срещу инжекции и изтриване на данни. Освен това изпълняваме възпроизводим механизъм за penetration testing срещу работещи внедрявания и изготвяме писмени одитни доклади. В рамките на седем вътрешни одита по сигурността през март и април 2026 отчетохме **zero critical findings**, като най-скорошният ни одит приключи с оценка **PASS**. (Формалната сертификация на тези контроли от трета страна е част от нашата пътна карта; вижте Раздел 18.)

Характеристика на сигурността	Обобщение
Хостинг	Microsoft Azure, само региони в ЕС
Мрежов модел	Частни крайни точки, мрежова сегментация по модела default-deny, без публична база данни
Криптиране	AES-256 в покой, TLS 1.2 или по-висока версия при пренос
Идентичност	Краткоживеещи подписани токени, bcrypt хеширане на пароли, поддръжка на SSO
Контрол на достъпа	Role-based access control със строга изолация по организация
Тайни	Централизиран vault за тайни с достъп чрез managed identity
Поверителност	Изрично съгласие, конфигурируемо съхранение, изтриване като единична единица
Отговорен AI	Само подпомагане на решения, човек винаги участва в процеса
Осигуреност	3,171 автоматизирани теста плюс периодични penetration tests и одити

1.1 Как да четете този документ

Раздели 3 до 11 описват контролите, които защитават данните: архитектура, мрежа, идентичност, приложение, защита на данните, поверителност и жизнен цикъл на сигурната разработка. Раздели 12 и 13 разглеждат нашата отличителна програма за непрекъснато тестване и историята на одитите ни. Раздели 14 до 17 обхващат операциите, моделирането на заплахи, управлението на уязвимости и съпоставянето с изискванията за съответствие. Приложенията предоставят каталог на контролите, ЧЗВ за проверяващи и речник, който екип по сигурност може да използва директно по време на

оценка.

2. Обхват и подход на документа

2.1 Какво обхваща този документ

Тази бяла книга обхваща архитектурата по сигурност и практиките на услугата AI Interview Analyzer: хостинг средата, мрежовия дизайн, управлението на идентичностите и достъпа, контролите на ниво приложение, защитата на данните, поверителността и съответствието с регулаторните изисквания, жизнения цикъл на сигурната разработка и нашата програма за непрекъснато тестване на сигурността.

2.2 Какво я прави проверима

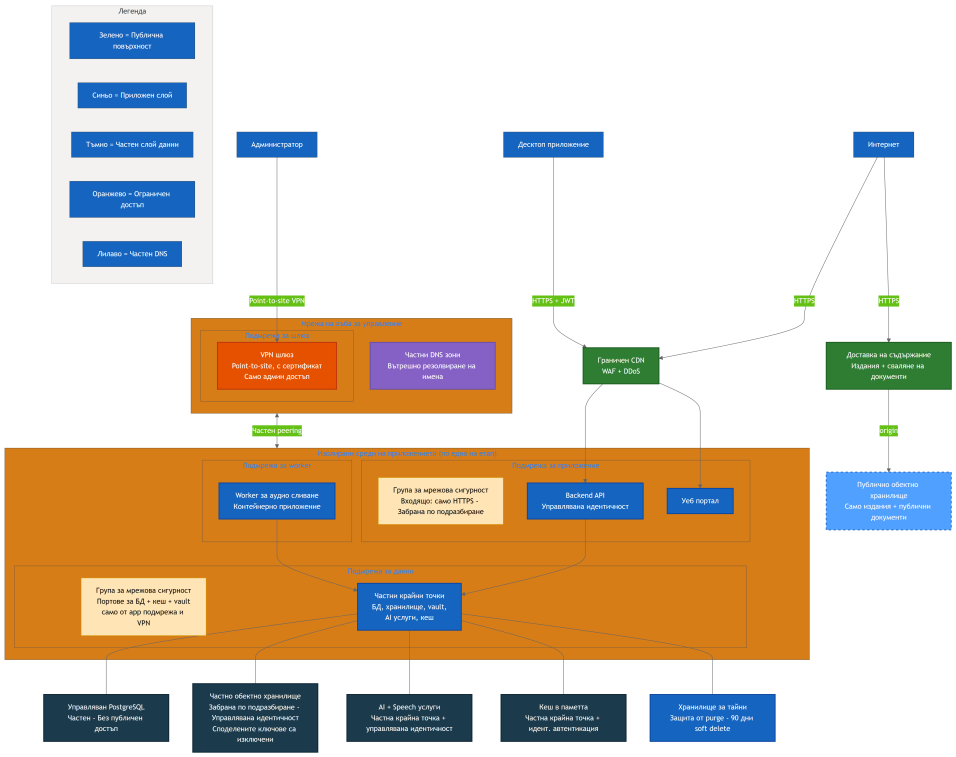
Твърденията на доставчиците относно сигурността са лесни за написване и трудни за доверяване. Затова сме обвързали всяко основно твърдение в този документ с нещо конкретно и измеримо в нашите инженерни системи: контрол, реализиран в код, тест, който доказва, че контролът работи, инфраструктурна дефиниция, която го налага, или одитен доклад, който документира извършена проверка. Когато даден контрол е част от бъдещата ни пътна карта, а не е внедрен днес, ние го заявяваме изрично. Предпочитаме да твърдим по-малко и да ни се вярва, отколкото да твърдим повече и да бъдем уличени.

2.3 Споделена отговорност

Платформата се предоставя като софтуер като услуга. Ние управляваме инфраструктурата, приложението, AI конвейера и обработката на данни. Клиентът носи отговорност за управлението на собствените си потребителски акаунти и роли, конфигурирането на прозорците за съхранение на данни така, че да съответстват на вътрешната му политика, и за гарантиране, че съгласието на кандидатите се получава чрез работния поток за съгласие, предоставен от платформата. Раздел 14 описва това разделение по-подробно.

3. Преглед на архитектурата по сигурност

Платформата е изградена като малък брой взаимодействащи си услуги, а не като единен монолит. Десктоп приложение и уеб портал действат като клиенти. Централен backend API управлява цялото съхранение, автентикацията, фактурирането, AI конвейера, съгласието, електронната поща, обработката на файлове и таблата за управление. Worker за обединяване на аудио обработва записите асинхронно. Всички чувствителни състояния се намират зад backend API; клиентите никога не комуникират директно с базата данни, хранилището или AI услугите.



Диаграмата по-горе показва производствената топология, като имената на ресурсите умишлено са обобщени. В нея се виждат три принципа:

- **Без директно излагане на услуги за данни.** Базата данни, частното object storage, AI услугите и кешът имат деактивиран публичен мрежов достъп и са достъпни само чрез частни крайни точки в рамките на изолирана виртуална мрежа. Vault за тайни се достига от приложението през частна крайна точка и е допълнително защитен чрез автентикация с платформена идентичност и политики за достъп с минимални привилегии, така че всеки достъп изисква валидна, оторизирана идентичност независимо от мрежовия път.
- **Отделена публична повърхност.** Единственото публично object storage съдържа файлове за изтегляне на версии и публични документи. То никога не съдържа данни на кандидати. Трафикът към приложението, насочен към клиентите, преминава през edge слой, който предоставя web application firewall, защита срещу distributed-denial-of-service и content delivery.
- **Административният достъп е контролиран.** Операторите достигат вътрешните ресурси само чрез point-to-site VPN с удостоверяване на базата на сертификати към мрежа management hub, а не през публичния интернет.

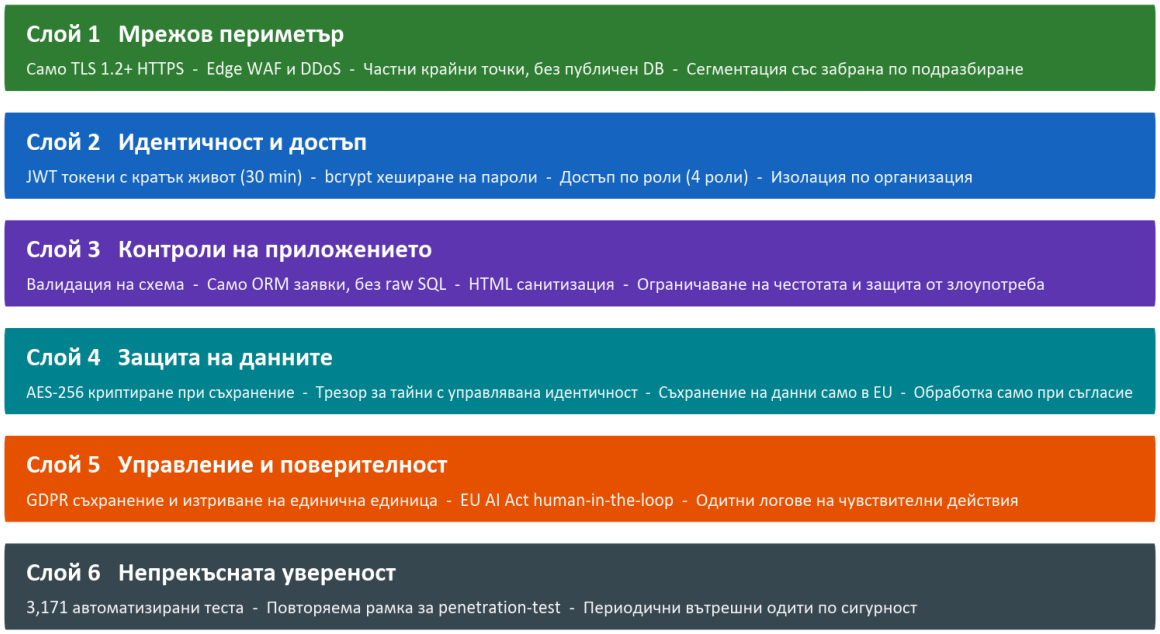
Всеки етап на внедряване (development и production) е напълно изолирана среда със собствена мрежа, storage accounts, база данни и тайни. Производствени данни на клиенти никога не присъстват в по-ниски среди. Споделен management hub съдържа само VPN gateway и private DNS, свързани частно с всяка среда.

---

4. Многослойна защита

На нито един отделен контрол не се разчита да спре всяка атака. Платформата наслагва независими контроли, така че отказът на който и да е слой да не води до излагане на данни. Слоевете по-долу са всеки поотделно реализирани и, както е описано в Раздел 12, индивидуално тествани.

Многослоен модел за сигурност: независими контроли на всяко ниво



Слой	Представителни контроли
Мрежов edge	Транспорт само чрез TLS, edge WAF и DDoS защита, частни крайни точки, сегментация по модела default-deny
Идентичност и достъп	Краткоживеещи подписани токени, bcrypt хеширане, role-based access control, изолация по организация
Приложение	Валидация по схема за всички входни данни, достъп до данни само чрез ORM, кодиране на изхода, rate limiting
Защита на данните	Криптиране в покой, vault за тайни с managed identity, съхранение на данни в ЕС, обработка, контролирана чрез съгласие
Управление и поверителност	Конфигурируемо съхранение, изтриване като единична единица, AI с човек в процеса, audit logging
Непрекъсната осигуреност	Автоматизиран пакет тестове, възпроизводими penetration tests, периодични вътрешни одити по сигурността

Останалата част от този документ разглежда всеки слой поотделно и след това описва как непрекъснато доказваме, че слоевете издържат.

5. Мрежова сигурност

5.1 Частно по подразбиране

Слоят с данни е частен по конструкция. Управляваната PostgreSQL база данни има деактивиран публичен мрежов достъп и е достъпна само чрез частна крайна точка. Частното object storage е конфигурирано по подразбиране да отказва мрежов достъп, изцяло деактивира shared access keys и е достъпно само чрез managed identity от подмрежата на приложението. Кешът, AI услугите и vault за тайни по същия начин се достигат чрез частни крайни точки с private DNS резолюция.

На практика това означава, че не съществува internet-facing connection string към базата данни и няма публичен storage URL за аудио на кандидати: базата данни и частното storage имат публичния мрежов достъп изцяло деактивиран. Vault за тайни се достига от приложението през частна крайна точка и е защитен чрез автентикация с платформена идентичност и политики за достъп с минимални привилегии, като на идентичностите на приложението е предоставен read-only достъп само до нужните им тайни, така че тайните не могат да бъдат извлечени без валидна, оторизирана идентичност. Повърхността за атака, до която външен противник изобщо може да се докосне, е ограничена до HTTPS крайните точки на приложението зад edge слоя.

5.2 Мрежова сегментация

Всяка среда е разделена на отделни подмрежи за слоя на приложението, слоя на данните и асинхронния worker. Всяка подмрежа се управлява от network security group, чието последно правило отказва целия входящ трафик. Подмрежата на приложението приема само входящ HTTPS. Подмрежата на данните приема само конкретните портове за база данни, кеш и vault и то само от подмрежата на приложението или административния VPN. Това означава, че дори нападател, който по някакъв начин достигне слоя на приложението, не може свободно да се придвижва към слоя на данните; разрешени са само пътищата, които приложението легитимно използва.

5.3 Edge слоят

Публичният трафик към приложението е фронтиран от edge слой, предоставящ web application firewall, DDoS защита и content delivery network. Изтеглянията на версии и документи се обслужват от специализиран публичен storage account чрез content-delivery front door, напълно отделен от частното storage, което съхранява данни на кандидати. Двата storage слоя никога не се смесват: неправилна конфигурация в публичния слой не може да изложи частни данни на кандидати, тъй като това са различни акаунти с различни мрежови правила.

5.4 Административен достъп

Няма публична административна крайна точка към частната мрежа. Операторите се свързват чрез point-to-site VPN gateway, който използва удостоверяване на базата на сертификати. Административният достъп до базата данни и кеша е възможен само отвътре на този тунел, тъй като тези услуги имат деактивиран публичен мрежов достъп. Това изцяло извежда ежедневните операции извън публичния интернет.

---

6. Управление на идентичност и достъп

6.1 Автентикация

Потребителските сесии се установяват с подписан access token, валиден за thirty minutes, в комбинация с отделен непрозрачен refresh token от страна на сървъра. Access token-ите се проверяват при всяка заявка, а потребителят се валидира повторно спрямо базата данни (включително проверка за активен акаунт), вместо да се разчита единствено на съдържанието на токена. Изходът от системата незабавно отменя refresh сесията на сървъра, така че откраднат refresh token не може да надживее logout.

Паролите никога не се съхраняват в открит текст. Те се хешират с bcrypt с уникална salt стойност за всяка парола. За организации, които предпочитат single sign-on, платформата поддържа OAuth вход с Microsoft и Google, като в този случай изобщо не се съхранява парола.

Собствеността върху email адреса се проверява чрез еднократна, ограничена във времето връзка за потвърждение, преди саморегистриран акаунт да се счита за потвърден, а повторните изпращания на email за потвърждение са ограничени чрез rate limiting, за да се предотврати злоупотреба.

6.2 Role-Based Access Control

Оторизацията се налага чрез модел с четири роли с нарастващи привилегии: interviewer, hiring manager, recruiter и administrator. Достъпът до привилегировани операции се налага чрез зависимости от страна на сървъра, които проверяват както ролята, така и статуса на верификация на извикващия. Тези проверки на роли защитават значително повече от сто отделни API операции.

Роля	Типични възможности
Interviewer	Провежда възложени интервюта; вижда само интервютата, възложени на него
Hiring manager	Управлява подбори, които притежава или в които е член
Recruiter	Пълно управление на подбори и кандидати в рамките на организацията
Administrator	Настройки на организацията, фактуриране, управление на потребители и API ключове

Отвъд грубите проверки по роли, платформата прилага правила за видимост на ниво данни. Hiring manager-ите виждат само подборите, които са създали или в които членуват; interviewer-ите виждат само интервютата, които са им възложени. Следователно привилегиите се налагат както на ниво „какво действие“, така и на ниво „кои записи“.

6.3 Изолация по организация

Платформата е multi-tenant и изолацията между tenants се разглежда като първокласен контрол по сигурността. Всяка автентикирана идентичност носи идентификатор на организация, а заявките за данни са ограничени до тази организация. Когато потребител поиска запис, който принадлежи на друга организация, платформата връща отговор „not found“, вместо да разкрива, че записът съществува. Вътрешните идентификатори в базата данни никога не се излагат по канала; API представя display идентификатори и ги пренасочва при всяка заявка, което премахва често срещан клас атаки за enumeration между tenants.

Това не е само проектно намерение. Както е описано в Раздел 12, нашият автоматизиран пакет изпълнява голяма матрица между организации, която опитва да достигне данните на една организация с идентификационните данни на друга и потвърждава, че всеки подобен опит се проваля.

6.4 Програмен достъп

За интеграции организациите по допустими планове могат да издават API ключове. Ключовете използват разпознаваем префикс, носят 128 bits of entropy и се съхраняват само като hash; суровият ключ се показва еднократно при създаване и

никога повече. Всеки ключ носи изричен permission scope (read, write или ATS integration), може да бъде ограничен до конкретни source networks, може да бъде отменен незабавно и подлежи на per-key rate limits, произтичащи от тарифния план на организацията. Верификацията на ключовете използва timing-safe comparison, за да се избегне изтичане на информация чрез времето за отговор.

7. Сигурност на приложението

Приложението е написано така, че да премахва цели категории уязвимости, а не да ги коригира поотделно случай по случай.

• **Инжекции.** Целият достъп до базата данни преминава през object-relational mapper с parameterized queries. Кодова база не съдържа raw SQL, форматиран като низове. Това структурно елиминира SQL injection.

• **Валидация на входа.** Всяко тяло на заявка се валидира спрямо строга схема, преди да достигне бизнес логиката. Прекомерно големи payload-и се отхвърлят, а list endpoint-ите са paginated, за да се ограничи използването на ресурси.

• **Кодиране на изхода и cross-site scripting.** Подаденият от потребителя и генерираният от AI текст се третират като недоверени. Когато съдържанието трябва да бъде визуализирано като HTML, то преминава през sanitizer с allow-list към момента на запис, а специализиран пакет тестове потвърждава, че script тагове, event handlers и javascript URL-и се премахват.

• **Mass assignment.** Операциите за актуализация използват изрични схеми, които изключват привилегировани полета като role, organization и credit balance, така че клиентът да не може да ескалира привилегии чрез публикуване на допълнителни полета.

• **Rate limiting.** Крайните точки за автентикация и тези, склонни към злоупотреба, са ограничени чрез устойчив limiter, базиран на база данни, който преживява рестартирания и работи коректно в множество инстанции на приложението. Login, registration, password reset и verification resend имат собствени лимити. Резолюцията на IP адреса на клиента е защитена срещу spoofing на forwarding headers.

• **Webhooks.** Входящите webhooks от доставчици на плащания и електронна поща се проверяват спрямо signatures на доставчика върху raw request body преди обработка.

• **Качване на файлове.** Upload-ите са ограничени по размер, валидирани, съхранявани под генерирани идентификатори вместо имена, подадени от потребителя, и ограничени за всяка заявка и за всяка организация.

• **Security headers.** В production отговорите носят strict transport security, опции за content-type и frame, referrer policy и рестриктивна permissions policy, и потискат server и framework banners.

8. Защита на данните

8.1 Криптиране

Всички данни са криптирани в покой чрез AES-256 посредством платформените слоеве за криптиране на storage и база данни в Azure. Целият мрежов трафик се обслужва изключително през HTTPS с TLS 1.2 или по-висока версия; некриптираният HTTP се пренасочва към HTTPS на всяко ниво. В production API и уеб порталът излъчват strict transport security headers заедно с набор от headers за втвърдяване и потискат банерите за версия на сървъра и framework-а.

8.2 Управление на тайни

Тайните на приложението се съхраняват в централизиран vault за тайни с активирана purge protection и ninety-day soft-delete window. Приложенията се автентикират към ресурсите в Azure чрез system-assigned managed identities, а не чрез дългоживеещи ключове; например частното storage има shared access keys, изцяло деактивирани, така че достъпът е възможен само чрез role assignments на база идентичност, ограничени до конкретния ресурс. Политиките за достъп до vault предоставят на принципалите на приложението read-only достъп до конкретните тайни, от които се нуждаят, съгласно принципа на минималните привилегии.

8.3 Местонахождение на данните

Всички данни на клиенти и кандидати се съхраняват и обработват в рамките на Европейския съюз. Хостингът на приложението, базата данни, storage, кешът и тайните се намират в West Europe, а AI обработката се изпълнява в региони на ЕС. Доставчикът на AI не използва клиентски данни за обучение на своите модели.

8.4 Жизненият цикъл на едно интервю

Най-ясният начин да се разберат контролите за защита на данните е да се проследи едно интервю от началото до края. Съгласието се събира и записва, преди каквото и да е да бъде обработено. Upload-ът е криптиран при пренос. Транскрипцията и анализът се изпълняват в центрове за данни в ЕС. Резултатите се записват в криптирано storage. След това всеки запис се управлява от единен часовник за съхранение, който завършва с регистрирано каскадно изтриване. Във всеки момент права на кандидата като оттегляне, изтриване, достъп или преносимост могат да прекъснат този поток.

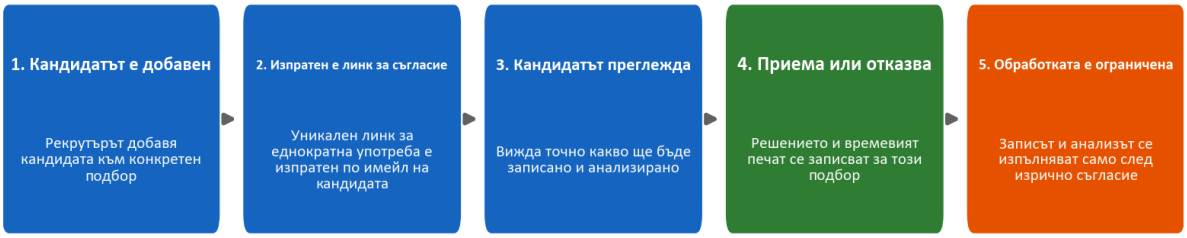
9. Поверителност по дизайн и GDPR

Поверителността е вградена в модела на данните и работния поток, а не е добавена впоследствие само чрез политика.

9.1 Съгласие

Нито едно интервю не се записва или анализира без изричното съгласие на кандидата. Когато кандидат бъде добавен към подбор, платформата изпраща по email уникална, еднократна връзка за съгласие. Кандидатът преглежда какво ще се случи и или приема, или отказва. Състоянието на съгласието, включително времето на отговора, се записва към този конкретен подбор, така че съгласието винаги е обвързано с конкретен процес по наемане, а не е предоставено глобално.

Съгласие на кандидата: изрично и записано преди всяка обработка

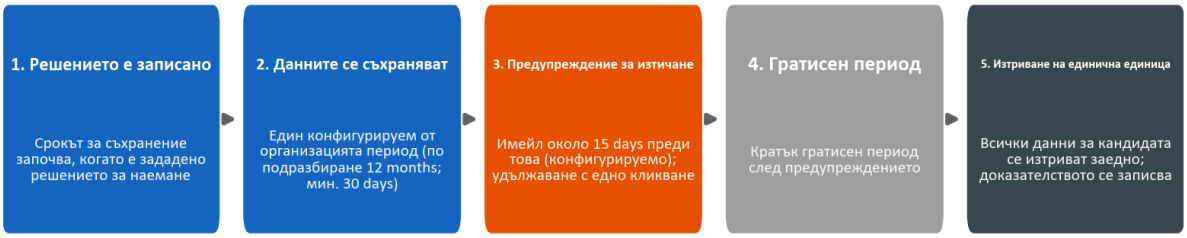


9.2 Съхранение и изтриване

Съхранението на данни е конфигурируемо за всяка организация, със стойност по подразбиране от twelve months и конфигурируем минимум от thirty days, и може да бъде променяно за всеки кандидат. За данните на кандидата има единен часовник за съхранение, а не отделен таймер за всеки артефакт. Часовникът започва, когато бъде записано решение за наемане. Преди изтичането на данните платформата изпраща предупреждение (по подразбиране около fifteen days предварително) и предлага удължаване с едно щракване. Когато данните бъдат изтрити, те се изтриват като единична единица: записът за кандидата, интервютата, транскриптите, аудиозаписите, документите и сравненията се премахват заедно, а изтриването се записва в audit log. Не остава частичен или осиротял остатък.

Жизненият цикъл по-долу показва този единен часовник и как той се свежда до едно каскадно изтриване с регистрирано доказателство за изтриване.

Съхранение на данни: един срок за кандидат, изтриване на единична единица



9.3 Права на субектите на данни и подизпълнители

Платформата поддържа правата на субектите на данни, изисквани съгласно GDPR, включително достъп, изтриване, преносимост, възражение и обяснение. Обработката се извършва съгласно DPA, която клиентите приемат при

регистрация и която е версионирана за всяка организация. Нашите подизпълнители и техните роли, всички в рамките на ЕС или при подходящи гаранции, са разкрити в това споразумение, а клиентите получават предварително уведомление за всяка промяна. Раздел 17 съдържа регистъра на подизпълнителите и съпоставянето на съответствието член по член.

10. Отговорен AI и EU AI Act

Платформата попада в категорията high-risk на EU AI Act, тъй като подпомага решения за заетост, и ние третираме тази класификация сериозно.

Определящото правило на продукта е, че **AI е средство за подпомагане на решения, а не вземащ решения.** Системата никога не приема или отхвърля автоматично кандидат. Тя транскрибира реч, структурира въпроси и отговори, оценява отговорите спрямо критерии, определени от специалиста по подбор, и изготвя проект на обратна връзка, като човек преглежда всеки резултат, преди той да бъде използван. Това държи човека твърдо в процеса.

Също толкова важно е какво AI не прави. Той не оценява личност, „културна съвместимост“, емоционално състояние, тон на гласа, акцент, пол, възраст, етническа принадлежност, външен вид или език на тялото. Оценяването е обвързано с доказателства от транскрипта и с критерии, дефинирани от специалиста по подбор, а имената на кандидатите са изключени от входа за оценяване, за да се намали пристрастността. Ние публикуваме карта за прозрачност, потребителска документация и декларация за съответствие, описващи системата, нейните ограничения и нейните предпазни мерки.

Контрол за отговорен AI	Как работи
Човек в процеса	Всеки резултат и всяка обратна връзка се преглеждат от специалист по подбор преди употреба
Без автоматизирани решения	Системата никога не приема или отхвърля автоматично кандидат
Оценяване, основано на доказателства	Оценките се позовават на подкрепящи доказателства от транскрипта
Дизайн срещу пристрастия	Имената са изключени от оценяването; оценява се съдържание, а не стил
Ограничения на обхвата	Никога не се оценяват личност, емоции, акцент и защитени характеристики
Безопасност на обратната връзка към кандидата	Частната обратна връзка към кандидата преминава през guardrail за безопасност за генериране и валидация

Тези ограничения не са само заявени в документация; те са кодирани в слоя с prompt-и за AI и се упражняват от специализирана програма за тестване на AI safety, описана в Раздел 12.3.

11. Жизнен цикъл на сигурната разработка

Сигурността се налага в начина, по който изграждаме и доставяме софтуер, а не само в работещата система.

• **Разделяне на средите.** Development и production са напълно отделени, всяка със собствена инфраструктура, storage accounts, база данни, тайни и поддомейни. Няма споделено състояние.

• **Infrastructure as code.** Цялата cloud среда е дефинирана като код и се преглежда като код, което прави състоянието на сигурността одитируемо и възпроизводимо. Проверяващ може да прочете точно кои портове са отворени, кои ресурси са частни и кои идентичности какви разрешения имат.

• **Фиксирани, контролирани внедрявания.** Всяка стъпка в continuous-integration pipeline-а е фиксирана към точна, неизменяема версия. Production внедряванията са базирани на tag, изпълняват се само чрез защитения production pipeline и са поставени зад задължително одобрение. Автоматизираният пакет тестове се изпълнява като release gate: внедряване не може да бъде публикувано, ако тестовете се провалят.

• **Хигиена на зависимостите.** Автоматизираното наблюдение на зависимостите предлага актуализации всяка седмица за backend, desktop, web, инфраструктура и pipeline дефиниции, а одитите на зависимостите са част от нашия периодичен преглед на сигурността.

• **Подписани артефакти.** Инсталаторите за desktop са code-signed, така че клиентите да могат да проверят, че инсталираният от тях софтуер действително идва от нас.

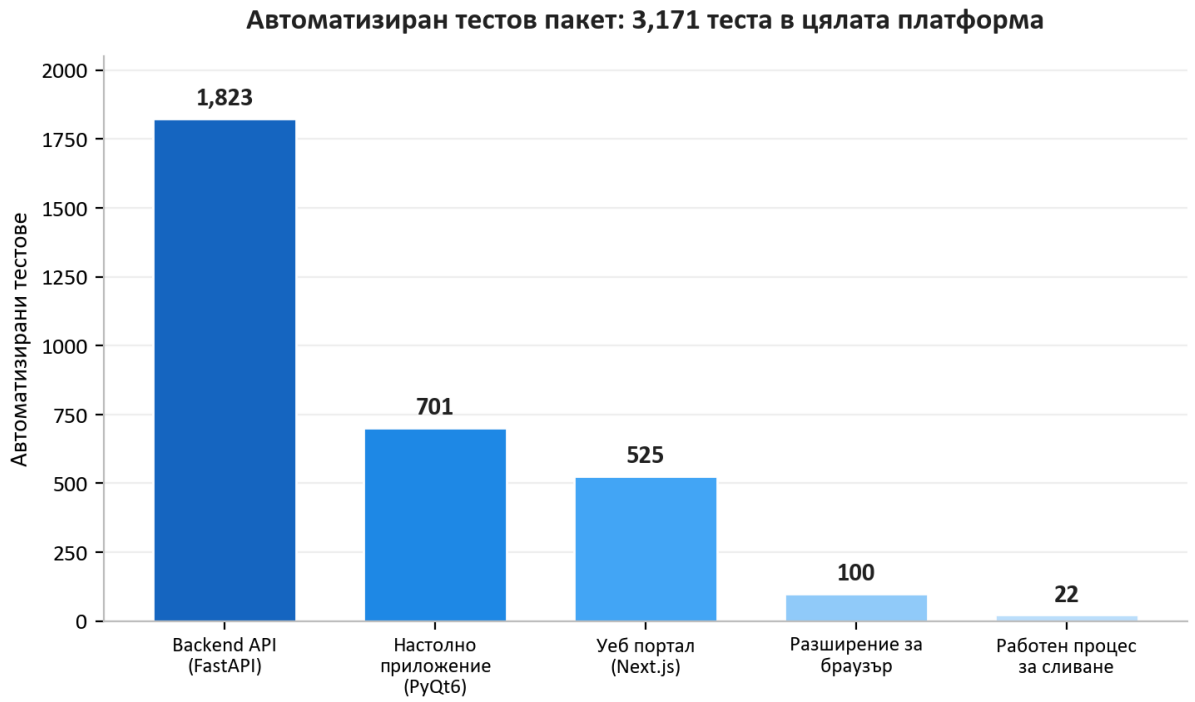
• **Дисциплина при тайните.** Тайните се намират във vault и в защитените pipeline secrets, никога в изходния код.

12. Непрекъснато тестване на сигурността

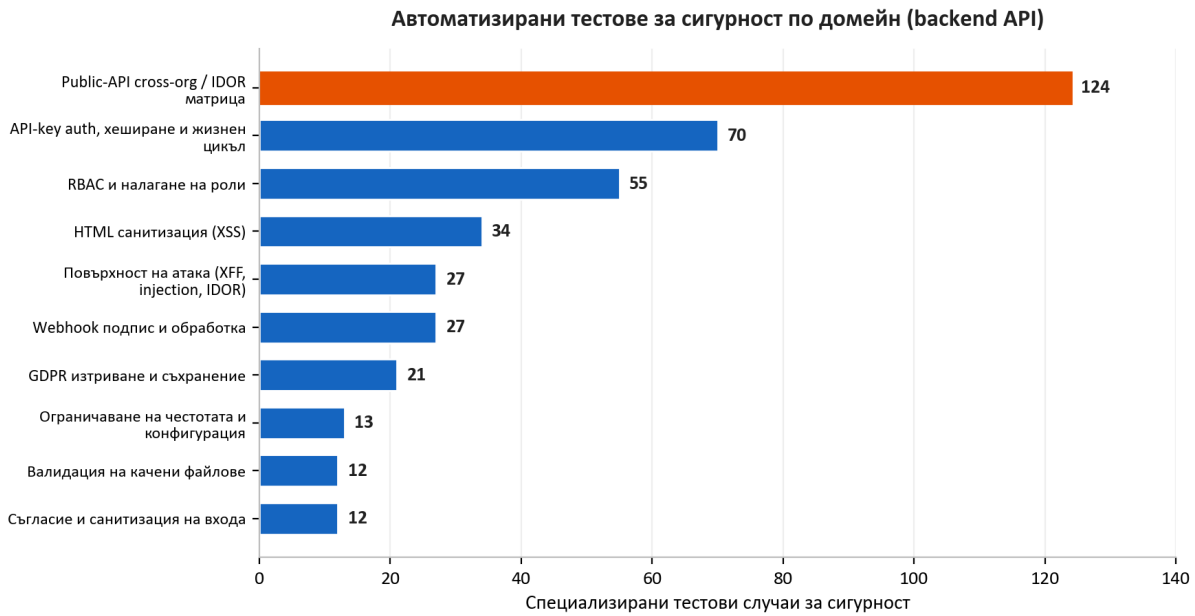
Това е сърцевината на нашата история за осигуреност и частта, която повечето доставчици не могат да покажат. Ние третираме сигурността като нещо, което трябва да се измерва непрекъснато чрез изпълними проверки, а не като нещо, което се заявява еднократно.

12.1 Автоматизираният пакет тестове

Платформата е покрита от **3,171 автоматизирани теста**, обхващащи backend API, desktop приложението, уеб портала, browser extension и audio merge worker.



Това не са само функционални тестове. Съществен, специализиран пакет за сигурност упражнява контролите, описани по-рано в този документ. Графиката по-долу разбива тестовете, специфични за сигурността в backend API, по домейни.



Наред с много други, този пакет включва голяма матрица за публичния API, която изпълнява всяка крайна точка като легитимен потребител, като собствен API ключ на организацията и като API ключ на конкурентна организация, като потвърждава, че всеки опит между организации е блокиран. Той включва десетки adversarial тестове за attack surface за spoofing на forwarding headers, header injection и изтичане на идентификатори, фокусиран пакет за HTML sanitization за cross-site scripting, тестове за налагане на роли за пълния модел от роли и тестове, които доказват, че данните на кандидата действително се изтриват като единица. Тъй като тези тестове се изпълняват като release gate, регресия, която отслабва който и да е от тези контроли, би спряла публикуването, вместо да достигне до клиентите.

12.2 Live Penetration Testing

Автоматизираните unit тестове доказват, че контролите се държат коректно в изолация. За да докажем, че те издържат заедно в реално внедряване, поддържаме възпроизводима методология за penetration testing, която изпълнява реални attack scripts срещу работеща среда. Тя е организирана в шест фази:

Фаза	Фокус	Примери за това, което се упражнява
1. Статичен анализ	Изходен код	Тайни, модели на инжектиране, опасни функции, липсваща auth, небезопасен HTML
2. Преглед на архитектурата	Инфраструктура	Частни крайни точки, сегментация, TLS, конфигурация на тайните
3. Анализ на векторите за атака	Source control и cloud	Защита на branch-ове, обхват на идентичностите, публично излагане
4. Live penetration testing	Работеща среда	Неавтентикирано сондиране, достъп между организации, инжекции, подправяне на токени, SSRF, bursts срещу rate limits
5. Enterprise scoring	Зрелост	Шестнадесет категории сигурност, оценени спрямо enterprise baseline
6. Зависимости и supply chain	Риск от трети страни	Одит на dependency CVE, фиксирани pipeline actions, цялост на lock file

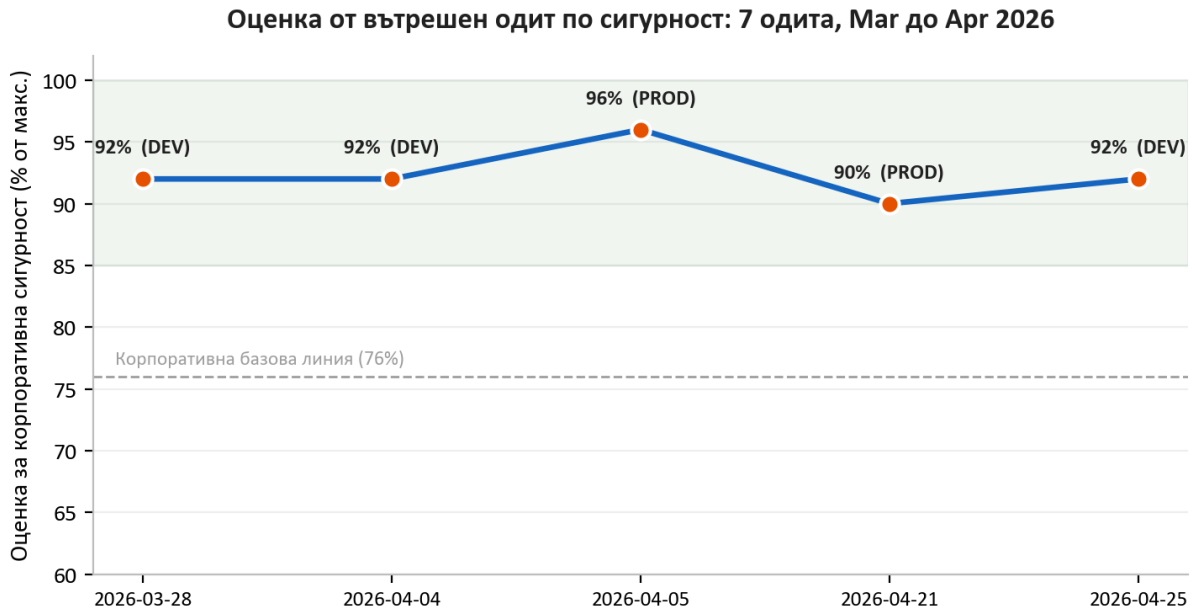
Фаза 4 представлява реално adversarial тестване срещу внедрена система, а не контролен списък. Тя сондира защитени крайни точки без идентификационни данни и потвърждава, че отказват достъп; регистрира две организации и се опитва да достигне записи на едната организация с акаунта на другата; инжектира payload-и за cross-site scripting и server-side template и потвърждава, че са неутрализирани; подправя authentication токени и потвърждава, че са отхвърлени; опитва server-side request forgery срещу cloud metadata endpoints; и изпраща bursts към authentication endpoint-ите, за да потвърди, че rate limiting действително се задейства в live средата, а не само на теория.

12.3 Тестване на безопасността на обратната връзка към кандидата

Тъй като платформата може да генерира частна развиваща обратна връзка за кандидати, изпълняваме отделна adversarial програма за безопасност срещу тази функционалност. Тя умишлено подава на системата груби и враждебни бележки от специалисти по подбор и потвърждава, че изходът, насочен към кандидата, никога не съдържа вулгарен език, никога не разкрива или приписва самоличност или лично мнение на специалист по подбор и никога не прилага осъдителни личностни етикети. Това защитава както кандидата, който следва да получава конструктивна и уважителна обратна връзка, така и клиента, чието вътрешно мнение никога не следва да изтича навън.

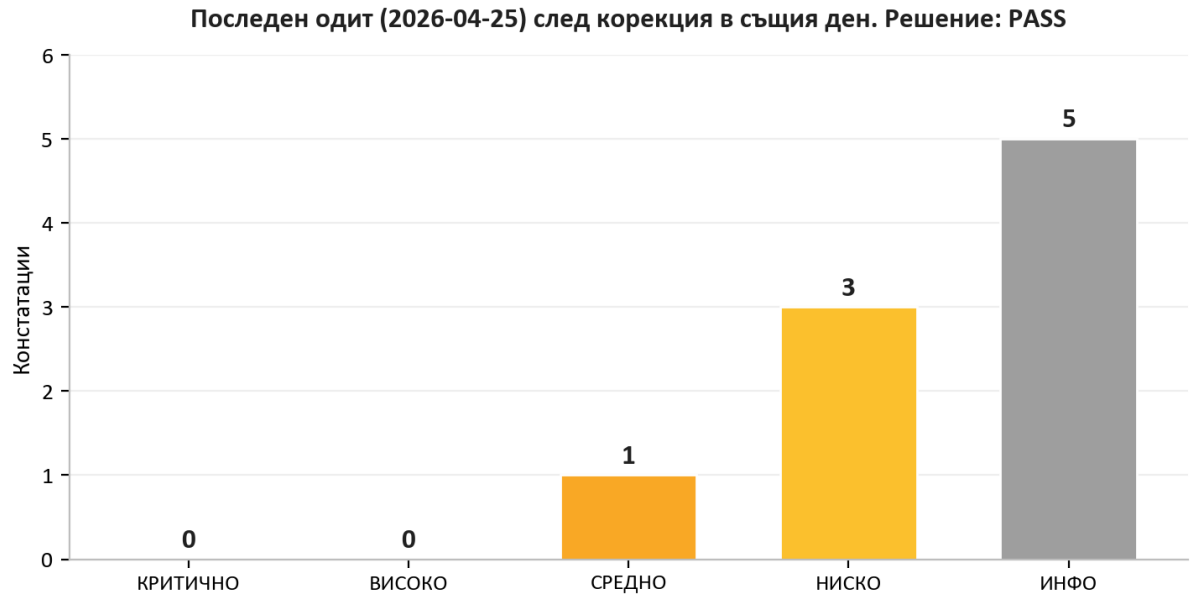
13. Резултати от одитите по сигурността

Провеждаме периодични одити по сигурността, използвайки структурирана, възпроизводима методология за penetration testing, и документираме всеки от тях като датиран доклад с констатации, оценени по тежест, доказателства и коригиращи действия. Това са вътрешни одити, изпълнявани чрез нашия собствен процес по сигурност; формалната сертификация на същите контроли от трета страна е част от нашата пътна карта. Между края на март и края на април 2026 завършихме **seven such audits** в development и production.



Резултатът, който е най-важен за потенциален клиент, е последователността: **в рамките на всички седем одита имаше zero critical findings.** В редките случаи, когато се появеше проблем с по-висока тежест, той беше коригиран бързо, често в същия ден, и проверен повторно. Скалата за оценяване беше умишлено затегната през този период (максимално възможният резултат беше повишен, тъй като добавяхме повече категории за оценяване), поради което линията на нормализирания резултат остава висока, дори когато летвата се повишаваше.

Нашият най-скорошен одит, на 25 April 2026, илюстрира как процесът работи на практика. Бяха идентифицирани два проблема с по-висока тежест, и двата бяха отстранени и повторно проверени в същия ден, а одитът приключи с оценка **PASS** без оставащи проблеми, готови за експлоатация, в рамките на текущия модел на заплахи.



Одит	Среда	Critical	Оценка
2026-03-28	Development	0	Готово за production
2026-04-04	Development	0	Готово за enterprise
2026-04-05	Production	0	Готово за enterprise
2026-04-20	Development	0	Готово за production, бележки
2026-04-20	Development	0	Pass с бележки
2026-04-21	Production	0	Безопасно, без експлоатируеми констатации
2026-04-25	Development	0	Pass

Моделът в тези одити е най-честното доказателство, което можем да предложим: проблемите се откриват, защото ги търсим активно, и се затварят бързо, защото процесът е създаден да ги затваря. Доставчик, който никога не докладва констатация, обикновено е доставчик, който не търси.

---

14. Оперативна устойчивост и споделена отговорност

14.1 Наблюдение и регистриране

Telemetry на приложението и платформата постъпва в централизиран workspace за log analytics и услуга за наблюдение на приложения, което ни дава видимост върху наличността и поведението. Чувствителни действия като изтриване на данни, приемане на правни споразумения и извиквания към AI се записват в специализирани audit таблици, така че да има траен запис кой какво е направил с важни данни.

14.2 Резервни копия и възстановяване

Управляваната база данни съхранява автоматизирани резервни копия, а частното storage е защитено чрез soft-delete retention както за blobs, така и за containers, така че случайно или злонамерено изтриване може да бъде възстановено в рамките на прозореца за съхранение. Критичната инфраструктура е снабдена с deletion locks, за да се предотврати случайно премахване на производствени ресурси.

14.3 Обобщение на споделената отговорност

Област	AI Interview Analyzer	Клиент
Инфраструктура, мрежа, patching	Да	-
Сигурност на приложението и AI конвейер	Да	-
Криптиране, тайни, местонахождение на данните	Да	-
Администриране на потребители и роли	Предоставя контролите	Управлява потребителите и ролите
Конфигуриране на политика за съхранение	Предоставя контролите	Определя прозореца за съхранение
Съгласие на кандидата	Предоставя работния поток	Гарантира, че се използва
Силни идентификационни данни на крайните потребители и SSO	Поддържа SSO и политика	Налага вътрешната политика

15. Модел на заплахите и съпоставяне с OWASP

Проектираме срещу конкретен набор от противници: външен нападател без идентификационни данни, любопитен или злонамерен автентикиран потребител от една организация, който се опитва да достигне данните на друга организация, компрометирана зависимост и вътрешна грешка. Таблицата по-долу съпоставя широко използваните категории риск от OWASP Top 10 със специфичните контроли, които ги адресират в тази платформа, всеки от които се упражнява от тестването, описано в Раздел 12.

OWASP риск	Как платформата го смекчава
Нарушен контрол на достъпа	Role-based access control на всяка привилегирована крайна точка; ограничаване по организация; „not found“ при достъп между организации; пренасочване на идентификатори; матрица от тестове между организации
Криптографски откази	TLS 1.2+ при пренос; AES-256 в покой; bcrypt хеширане на пароли; тайни в управляван vault
Инжекции	Само ORM parameterized queries; строга валидация по схема; HTML sanitization при запис
Несигурен дизайн	Слоеста многослойна защита; моделиране на заплахи и преглед на архитектурата при всеки одит
Неправилна конфигурация на сигурността	Infrastructure as code; мрежови групи по модела default-deny; security headers; деактивирани shared storage keys; API schema не е изложена в production
Уязвими компоненти	Седмично автоматизирано наблюдение на зависимостите; одити на dependency CVE при периодичен преглед
Неуспехи в идентификацията и автентикацията	Краткоживеещи токени; login с rate limiting; email verification; поддръжка на SSO; без plaintext пароли
Неуспехи в целостта на софтуера и данните	Фиксирани, неизменяеми pipeline стъпки; подписани desktop инсталатори; webhook signature verification; production deploys, контролирани чрез tag
Неуспехи в регистрирането и наблюдението на сигурността	Централизирана telemetry; специализирани audit таблици за чувствителни действия
Server-side request forgery	Изходящите извиквания са ограничени до доверени крайни точки; SSRF probe-ове в механизма за penetration testing

Това съпоставяне е гръбнакът на нашия аргумент за осигуреност: за всеки добре познат клас атака има именуван контрол, а за всеки именуван контрол има тест.

16. Управление на уязвимости и отговорно разкриване

Сигурността никога не е завършена, затова изпълняваме непрекъснат цикъл на откриване и коригиране.

• **Откриване.** Уязвимостите се извеждат от четири източника: автоматизирания пакет тестове, периодичните penetration-test одити, автоматизираното наблюдение на зависимостите и сигнали от клиенти или изследователи.

• **Триаж.** На всяка констатация се присвоява тежест (critical, high, medium, low или informational) с доказателства и собственик на коригиращите действия, точно както е записано в нашите одитни доклади.

• **Цели за коригиране.** Констатациите с critical и high тежест се приоритизират за незабавно коригиране; в историята на нашите одити констатациите с по-висока тежест обикновено са били отстранявани и повторно проверявани в същия ден. Констатациите със medium и по-ниска тежест се планират в обичайния цикъл на поддръжка.

• **Проверка.** Поправките се тестват повторно, а когато е относимо, се изпълнява live проверка срещу внедрената среда, за да се потвърди, че проблемът е действително затворен, а не само затворен в кода.

• **Разкриване.** Притеснения, свързани със сигурността, могат да ни бъдат докладвани директно. Ние потвърждаваме получаването на сигналите, разследваме ги и държим подателя информиран до разрешаването им.

17. Съпоставяне със съответствието

17.1 GDPR

Област по GDPR	Реализация в платформата
Правно основание (Art. 6)	Изрично съгласие на кандидата, събрано преди обработка
Минимизиране на данните и ограничение на съхранението (Art. 5)	Обработват се само данни, свързани с интервюто; конфигурируемо съхранение с автоматично изтриване
Право на изтриване (Art. 17)	Изтриване като единична единица на всички данни на кандидата с регистрирано доказателство за изтриване
Права на субектите на данни (Art. 15 to 20)	Поддържат се достъп, изтриване, преносимост и възражение
Задължения на обработващия (Art. 28)	DPA се приема при регистрация и е версионирано за всяка организация
Сигурност на обработването (Art. 32)	Криптиране, контрол на достъпа, изолация и непрекъснато тестване, както е описано в този документ
Прозрачност относно подизпълнителите	Разкрити в DPA с предварително уведомление за промяна

17.2 EU AI Act

Платформата се третира като high-risk AI system, подпомагаща решения за заетост, и ние поддържаме документация, съобразена с регулацията, включително карта за прозрачност, потребителска документация и декларация за съответствие. Основните предпазни мерки, човешкият надзор, прозрачността, оценяването, основано на доказателства, и строгите ограничения на обхвата на това, което AI оценява, са описани в Раздел 10. Продължаваме да развиваме нашата формална документация за съответствие с напредването на графика за прилагане на регулацията.

17.3 Сертификации на хостинга

Платформата работи изцяло върху Microsoft Azure, чиито центрове за данни притежават независими сертификати, включително ISO 27001 и SOC 2. Тези сертификати обхващат физическите и платформените слоеве под нашето приложение; контролите на ниво приложение са тези, описани в целия този документ.

17.4 Регистър на подизпълнителите

Подизпълнител	Цел	Регион
Microsoft Azure	Хостинг, AI и speech processing, storage, transactional email	EU (West Europe, Sweden Central)
Stripe	Обработка на абонаменти и плащания	EU (Ireland)
Fakturownia	Фактуриране	EU (Poland)
ATS connector (optional)	Интеграция със система за проследяване на кандидати, активира се само при заявка	EU

18. Пътна карта за сигурността

Разглеждаме сигурността като програма за непрекъснато усъвършенстване. Текущите инициативи в нашата пътна карта включват укрепване на възможностите за multi-factor authentication за административни акаунти, разширяване на централизираното audit logging на достъпа до данни, продължаващо затягане на актуалността на зависимостите по редовен график и напредък към формална сертификация от трета страна на контролите, описани в този документ. Нито една от тези инициативи не представлява празнина, която днес излага клиентски данни; всяка е подобрение на вече многослойна позиция.

19. Обобщение

AI Interview Analyzer защитава данните на кандидатите и клиентите чрез многослойна архитектура: мрежа, частна по подразбиране, без публични услуги за данни, силна идентичност и изолация по организация, код на приложението, който елиминира цели класове уязвимости, криптиране и местонахождение на данните в ЕС, и контроли за поверителност, вградени в модела на данните. Това, което отличава платформата, са доказателствата зад тези твърдения. С 3,171 автоматизирани теста, възпроизводима методология за live penetration testing, специализирана програма за AI safety и история от седем вътрешни одита по сигурността с zero critical findings, ние можем да покажем, а не просто да кажем, че платформата е сигурна.

Приложение A: Каталог на контролите по сигурността

Съкратена справка за основните контроли и доказателствата, които подкрепят всеки от тях.

Контрол	Механизъм	Доказателства
Криптиране на транспорта	Само HTTPS, TLS 1.2+, пренасочване на HTTP	Infrastructure as code; одит на архитектурата
Криптиране в покой	Платформено AES-256 криптиране на storage и база данни	Платформена конфигурация; одит на архитектурата
Защита на паролите	bcrypt със salt за всяка парола	Source control; тестове за автентикация
Управление на сесии	30-minute подписани токени, отменим refresh от страна на сървъра	Source control; тестове за автентикация
Оторизация	Контрол на достъпа с четири роли на привилегировани крайни точки	Пакет тестове за налагане на роли
Изолация на tenants	Ограничаване на заявките по организация; 404 при достъп между организации	Матрица от тестове между организации
Сигурност на API ключове	Хеширано съхранение, ограничени разрешения, per-key rate limits	Пакет тестове за API ключове
Защита срещу инжекции	Само ORM parameterized queries	Статичен анализ; тестове за инжекции
Защита срещу cross-site scripting	HTML sanitization при запис	Пакет тестове за HTML sanitization
Rate limiting	Устойчив limiter за auth endpoint-и, базиран на база данни	Тестове за rate limiting; live burst проверки
Цялост на webhook-овете	Проверка на signature на доставчика върху raw body	Пакет тестове за webhooks
Управление на тайни	Управляван vault, purge protection, managed identity	Infrastructure as code; одит на архитектурата
Мрежова изолация	Частни крайни точки; сегментация по модела default-deny	Infrastructure as code; одит на архитектурата
Изтриване на данни	Каскадно изтриване като единична единица с audit log	Пакет GDPR тестове за изтриване
Supply chain	Фиксирани pipeline стъпки; седмично наблюдение на зависимостите	Конфигурация на pipeline; одит на зависимостите

Приложение B: Често задавани въпроси за проверяващи по сигурността

Къде се съхраняват нашите данни? Изцяло в рамките на Европейския съюз, в Microsoft Azure, в West Europe с AI обработка в региони на ЕС. Данните на кандидатите никога не напускат ЕС.

Използват ли се нашите данни за обучение на AI модели? Не. Доставчикът на AI не използва клиентски данни за обучение.

Достъпна ли е базата данни от интернет? Не. Публичният мрежов достъп е деактивиран и базата данни е достъпна само чрез частна крайна точка в рамките на виртуалната мрежа.

Може ли един клиент да вижда данните на друг клиент? Не. Всяка заявка е ограничена до организацията на извикващия, достъпът между организации връща „not found“, а автоматизирана матрица непрекъснато тества тази изолация.

Как се съхраняват паролите? Хеширани с bcrypt и уникална salt стойност за всяка парола. Поддържа се single sign-on с Microsoft и Google, като в този случай не се съхранява парола.

Поддържате ли single sign-on? Да, чрез Microsoft и Google OAuth.

Колко дълго са валидни access token-ите? Thirty minutes, в комбинация с отменима refresh сесия от страна на сървъра, която се инвалидира при logout.

Как се обработва съгласието на кандидата? Всеки кандидат получава уникална, еднократна връзка за съгласие и трябва да приеме преди какъвто и да е запис или анализ. Съгласието се записва спрямо конкретния процес по наемане.

Как се изтриват данните? Като единична единица, обхващаща записа за кандидата, интервютата, транскриптите, аудиото, документите и сравненията, по конфигурируем график за съхранение, с регистрирано доказателство за изтриване. Кандидатите могат също да поискат изтриване директно.

Имате ли DPA? Да, приема се при регистрация и е версионирано за всяка организация, включително регистъра на подизпълнителите.

AI взема ли решения за наемане? Не. Той предоставя само подпомагане на решения; човек преглежда всеки резултат и взема всички решения.

Как доказвате твърденията си за сигурност? Чрез 3,171 автоматизирани теста, включително специализиран пакет за сигурност, възпроизводима шестфазова методология за penetration testing, изпълнявана срещу live среди, програма за тестване на AI safety и периодични писмени одитни доклади.

Какво се случва, когато откриете уязвимост? Присвоява ѝ се тежест с доказателства и собственик, отстранява се по приоритетен график, проверява се повторно, включително чрез live проверки, когато е относимо, и се записва в одитен доклад.

Можем ли да проведем собствен penetration test? Оценки на сигурността могат да бъдат организирани чрез вашия account representative при подходящ обхват и график.

Приложение C: Речник

Термин	Значение
AES-256	Силен симетричен стандарт за криптиране, използван за защита на данни в покой
bcrypt	Специализирана функция за хеширане на пароли със salt за всяка парола
Managed identity	Идентичност, издадена от платформата, която позволява на услуга да се автентикира без съхранявани ключове
Private endpoint	Частен мрежов адрес, който държи cloud услуга извън публичния интернет
Network security group	Набор от правила за разрешаване и отказ, които филтрират мрежовия трафик към подмрежа
RBAC	Role-based access control, предоставящ разрешения според ролята на потребителя
IDOR	Insecure direct object reference, слабост в контрола на достъпа, срещу която платформата се защитава
SSRF	Server-side request forgery, клас атаки, сондиран в нашите penetration tests
Web application firewall	Edge контрол, който филтрира злонамерен уеб трафик
Data processing agreement	Договорът, който урежда как обработващият обработва лични данни от името на администратор

Приложение D: Контакти и контрол на документа

AI Interview Analyzer Sp. z o.o.

ul. Jedrusik 6/53, 01-748 Warszawa, Poland

NIP: 5253079974

За преглед по сигурността, копие от нашата DPA или нашата документация за съответствие с EU AI Act, моля, свържете се с вашия account representative.

Този документ описва състоянието на сигурността на услугата AI Interview Analyzer към датата на генериране, показана в долния колонтитул. Той се предоставя за целите на оценка и не представлява част от какъвто и да е договор. Конкретните договорни ангажименти по сигурността са посочени в приложимото споразумение и DPA.