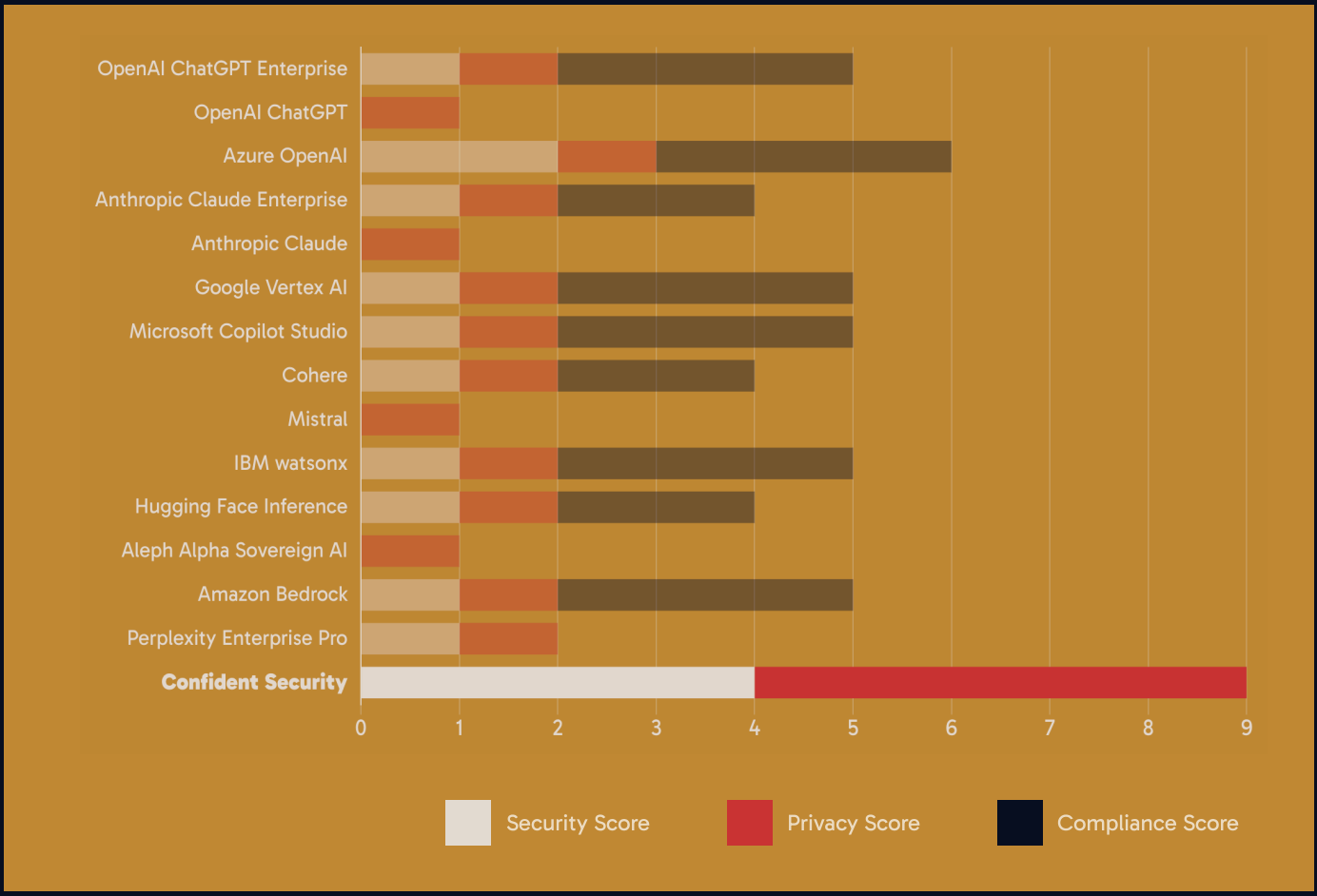
CONFIDENT SECURITY

The Private AI Scorecard



Last Updated: April 24, 2025

Disclaimer: Based on publicly available data.

For corrections or updates, email public links to checklist@confident.security.

Private & Secure AI, Delivered.

Join the waitlist

What makes an AI solution private?

Each AI solution is graded against the private AI checklist, making it easy to determine what solutions may fit your needs. The checklist includes key security and privacy guarantees, along with important compliance certifications.

Security Guarantees

- Prompts and metadata are never stored or logged beyond the processing of the request, for debugging, service management, or any other purpose
- Prompts and metadata are guaranteed to never be used in training AI models
- Prompts and metadata are guaranteed to never be shared with a third party
- The system provides provable guarantees about the exact model, model version, and operating environment for each and every request

Privacy Guarantees

- It is guaranteed that no person can access unencrypted user data
- All interactions are provably guaranteed to be anonymous – no one can determine which user is associated with any individual interaction
- A client can verify that its traffic is not being modified or tampered with via cryptographic controls
- The provider does not have access to inspect any private computation or user interactions
- Prompts and metadata cannot be decrypted if the operator's LLM environment does not match what is claimed or expected

Compliance Certifications

- SOC2: a security framework that helps companies demonstrate that they've implemented controls to protect customer data
- GDPR: The General Data Protection Regulation, abbreviated GDPR, is a European Union regulation on information privacy in the European Union and the European Economic Area
- CSA STAR: The Security, Trust, Assurance, and Risk (STAR) Registry is a publicly accessible registry that documents the security and privacy controls provided by popular cloud computing offerings

Note: These assessments rely exclusively on publicly available information and may not reflect custom terms negotiated on a per-customer basis

CONFIDENT SECURITY

OpenAI ChatGPT Enterprise

5/12
AI Scorecard

View Product Page

Security Guarantees

- 🚫 Prompts and metadata may be stored or logged
- ✅ “We do not train on your business data or conversations”
- 🚫 Prompts and metadata may be shared with subprocessors
- 🚫 OpenAI does not provide a cryptographic attestation mechanism or verifiable hash for each model version at request time

Privacy Guarantees

- 🚫 Data is encrypted at rest and in transit, but is not guaranteed to not be accessed by administrators
- 🚫 ChatGPT Enterprise tracks usage by organization and may tie sessions to user accounts for security and analytics -- there is no guarantee of anonymity
- ✅ Data in transit is encrypted and can be verified via TLS certificates to ensure data integrity and authenticity
- 🚫 Personnel can access logs or data in special circumstances (e.g., to investigate abuse or significant service issues)
- 🚫 There is no confidential computing or enclave-based attestation system that restricts decryption only to a specifically attested environment

Compliance Guarantees

- ✅ ChatGPT Enterprise is SOC 2 Type 2 compliant
- ✅ ChatGPT Enterprise is STAR Level 1 compliant
- ✅ ChatGPT Enterprise supports GDPR compliance through data processing terms and controls

References

ChatGPT Enterprise · Enterprise Privacy at OpenAI · OpenAI Subprocessors · OpenAI Security & Privacy · STAR Registry

CONFIDENT SECURITY

OpenAI ChatGPT

1/12
AI Scorecard

View Product Page

Security Guarantees

- 🚫 Prompts and metadata may be stored or logged
- 🚫 User data is used to help train and improve OpenAI's models
- 🚫 Prompts and metadata may be shared with subprocessors
- 🚫 OpenAI does not provide a cryptographic attestation mechanism or verifiable hash for each model version at request time

Privacy Guarantees

- 🚫 Data is encrypted at rest and in transit, but is not guaranteed to not be accessed by administrators
- 🚫 Usage can be correlated to specific customers or users for billing, monitoring, or audit logs
- ✅ Data in transit is encrypted and can be verified via TLS certificates to ensure data integrity and authenticity
- 🚫 OpenAI or its automated systems can review content to detect violations of usage policies
- 🚫 There is no confidential computing or enclave-based attestation system that restricts decryption only to a specifically attested environment

Compliance Guarantees

- 🚫 ChatGPT is not confirmed to be SOC 2 Type 2 compliant
- 🚫 ChatGPT is not confirmed to be STAR Level 1 compliant
- 🚫 ChatGPT is not confirmed to be GDPR compliant

References

OpenAI ChatGPT | OpenAI Privacy Policy | OpenAI Security & Privacy | OpenAI Policies | OpenAI Subprocessors

CONFIDENT SECURITY

Azure OpenAI

6/12
AI Scorecard

View Product Page

Security Guarantees

- 🚫 Prompts and metadata are retained for up to 30 days for abuse-detection and service health
- ✅ Customer data is not used to train, retrain, or improve Azure OpenAI or OpenAI foundation models
- ✅ Azure OpenAI does not share your data with third parties without your permission
- 🚫 Azure OpenAI does not provide a cryptographic attestation mechanism or verifiable hash for each model version at request time

Privacy Guarantees

- 🚫 Data is encrypted at rest and in transit, but is not guaranteed to not be accessed by administrators
- 🚫 Logs may contain identifiers for abuse investigation; anonymity isn't guaranteed
- ✅ Data in transit is encrypted and can be verified via TLS certificates to ensure data integrity and authenticity
- 🚫 Personnel can access logs or data in special circumstances (e.g., to investigate abuse or significant service issues)
- 🚫 There is no confidential computing or enclave-based attestation system that restricts decryption only to a specifically attested environment

Compliance Guarantees

- ✅ Azure services, including Azure OpenAI, are covered by SOC 2 Type 2 reports
- ✅ Azure maintains CSA STAR Level 1 self-assessment.
- ✅ Azure OpenAI adheres to GDPR, per Azure Trust Center guidance

References

Azure OpenAI Service Azure Data, Privacy, and Security Protecting Azure AI Customer Data Azure AI Customer Data FAQ

CONFIDENT SECURITY

Anthropic Claude for Enterprise

4/12
AI Scorecard

View Product Page

Security Guarantees

- 🚫 Prompts and metadata may be stored or logged for security, debugging, or abuse detection
- ✅ Enterprise data is not used to train or improve Anthropic's foundation models
- 🚫 Prompts and metadata may be shared with subprocessors
- 🚫 Anthropic does not advertise a cryptographic attestation mechanism

Privacy Guarantees

- 🚫 Data is encrypted at rest and in transit, but is not guaranteed to not be accessed by administrators
- 🚫 Anthropic Claude for Enterprise associates prompts with user accounts for service improvement, policy enforcement, and billing
- ✅ Data in transit is encrypted and can be verified via TLS certificates to ensure data integrity and authenticity
- 🚫 Anthropic or its automated systems can review content to detect violations of usage policies
- 🚫 There is no confidential computing or enclave-based attestation system that restricts decryption only to a specifically attested environment

Compliance Guarantees

- ✅ Anthropic Claude for Enterprise is stated to be SOC 2 Type 2 compliant
- 🚫 Anthropic Claude for Enterprise is not confirmed to be STAR Level 1 compliant
- ✅ Anthropic Claude for Enterprise offers GDPR compliance for EU customers

References

Anthropic Claude for Enterprise | Anthropic Privacy Policy | Anthropic Commercial Terms | Anthropic Subprocessors

CONFIDENT SECURITY

Anthropic Claude

1/12
AI Scorecard

View Product Page

Security Guarantees

- 🚫 Prompts and metadata may be stored or logged for security, debugging, or abuse detection
- 🚫 User data is used to help train and improve Anthropic's models
- 🚫 Prompts and metadata may be shared with subprocessors
- 🚫 Anthropic does not advertise a cryptographic attestation mechanism

Privacy Guarantees

- 🚫 Data is encrypted at rest and in transit, but is not guaranteed to not be accessed by administrators
- 🚫 Anthropic associates prompts with user accounts for service improvement, policy enforcement, and billing
- ✅ Data in transit is encrypted and can be verified via TLS certificates to ensure data integrity and authenticity
- 🚫 Anthropic or its automated systems can review content to detect violations of usage policies
- 🚫 There is no confidential computing or enclave-based attestation system that restricts decryption only to a specifically attested environment

Compliance Guarantees

- 🚫 Anthropic Claude is not stated to be SOC 2 Type 2 compliant
- 🚫 Anthropic Claude is not stated to be STAR Level 1 compliant
- 🚫 Anthropic Claude is not stated to be GDPR compliant

References

Anthropic Claude | Anthropic Privacy Policy | Anthropic Commercial Terms | Anthropic Subprocessors

CONFIDENT SECURITY

Google Vertex AI

5/12
AI Scorecard

View Product Page

Security Guarantees

- 🚫 Prompts and metadata may be stored or logged for security, debugging, or abuse detection
- ✅ User data is not used to train models unless a customer explicitly opts in
- 🚫 Prompts and metadata may be shared with subprocessors
- 🚫 Google does not advertise a cryptographic attestation mechanism

Privacy Guarantees

- 🚫 Data is encrypted at rest and in transit, but is not guaranteed to not be accessed by administrators
- 🚫 Vertex AI usage is tied to a Google Cloud project and IAM credentials for billing and monitoring
- ✅ Data in transit is encrypted and can be verified via TLS certificates to ensure data integrity and authenticity
- 🚫 Google may inspect logs or user content for reasons such as abuse, debugging, or legal compliance
- 🚫 There is no confidential computing or enclave-based attestation system that restricts decryption only to a specifically attested environment

Compliance Guarantees

- ✅ Google Cloud is SOC 2 Type 2 certified, and Vertex AI falls under the broader scope
- ✅ Google Cloud is CSA STAR Level 2 certified
- ✅ Google Cloud states it provides tools and contractual assurances to help customers comply with GDPR

References

Google Vertex AI | Google Cloud Privacy Notice | Google Vertex AI docs | Google Cloud Subprocessors

CONFIDENT SECURITY

Microsoft Copilot Studio

5/12
AI Scorecard

View Product Page

Security Guarantees

- 🚫 Prompts and metadata may be stored or logged for security, debugging, or abuse detection
- ✅ User data is not used to train models
- 🚫 Prompts and metadata may be shared with subprocessors
- 🚫 Microsoft does not advertise a cryptographic attestation mechanism

Privacy Guarantees

- 🚫 Data is encrypted at rest and in transit, but is not guaranteed to not be accessed by administrators
- 🚫 Copilot Studio usage is tied to Azure AD credentials for audit trails, billing, and governance
- ✅ Data in transit is encrypted and can be verified via TLS certificates to ensure data integrity and authenticity
- 🚫 Microsoft may inspect logs or user content for reasons such as abuse, debugging, or legal compliance
- 🚫 There is no confidential computing or enclave-based attestation system that restricts decryption only to a specifically attested environment

Compliance Guarantees

- ✅ Microsoft Azure is SOC 2 Type 2 certified, and Copilot Studio falls under the broader scope
- ✅ Microsoft Azure is CSA STAR Level 2 certified
- ✅ Microsoft Azure states it provides tools and contractual assurances to help customers comply with GDPR

References

Microsoft Copilot Studio Data, Privacy, and Security for Microsoft 365 Copilot Microsoft Commercial Support Subprocessors

CONFIDENT SECURITY

Cohere

4/12
AI Scorecard

View Product Page

Security Guarantees

- 🚫 Prompts and metadata may be stored or logged for security, debugging, or abuse detection
- ✅ User data is not used to train models
- 🚫 Prompts and metadata may be shared with subprocessors
- 🚫 Cohere does not advertise a cryptographic attestation mechanism

Privacy Guarantees

- 🚫 Data is encrypted at rest and in transit, but is not guaranteed to not be accessed by administrators
- 🚫 Interactions are associated with API keys for purposes like billing and analytics
- ✅ Data in transit is encrypted and can be verified via TLS certificates to ensure data integrity and authenticity
- 🚫 Cohere may inspect logs or user content for reasons such as abuse, debugging, or legal compliance
- 🚫 There is no confidential computing or enclave-based attestation system that restricts decryption only to a specifically attested environment

Compliance Guarantees

- ✅ Cohere has publicly stated that it has achieved SOC 2 Type 2 compliance for its cloud offerings
- 🚫 Cohere is not stated to be STAR Level 1 compliant
- ✅ Cohere offers tools and contractual terms to help customers meet GDPR requirements

References

Cohere

Cohere Privacy Policy

Cohere Security

Cohere FAQ

CONFIDENT SECURITY

Mistral

1/12
AI Scorecard

View Product Page

Security Guarantees

- 🚫 Prompts and metadata may be stored or logged for security, debugging, or abuse detection
- 🚫 User data may be used to train models
- 🚫 Prompts and metadata may be shared with subprocessors
- 🚫 Mistral does not advertise a cryptographic attestation mechanism

Privacy Guarantees

- 🚫 Data is encrypted at rest and in transit, but is not guaranteed to not be accessed by administrators
- 🚫 Interactions are associated with API keys for purposes like billing and analytics
- ✅ Data in transit is encrypted and can be verified via TLS certificates to ensure data integrity and authenticity
- 🚫 Mistral may inspect logs or user content for reasons such as abuse, debugging, or legal compliance
- 🚫 There is no confidential computing or enclave-based attestation system that restricts decryption only to a specifically attested environment

Compliance Guarantees

- 🚫 Mistral is not stated to be SOC 2 Type 2 compliant
- 🚫 Mistral is not stated to be STAR Level 1 compliant
- 🚫 Mistral is not stated to be GDPR compliant

References

Mistral | Mistral ToS | Mistral Data Processing | Mistral Privacy Policy

CONFIDENT SECURITY

IBM watsonx

5/12
AI Scorecard

View Product Page

Security Guarantees

- 🚫 Prompts and metadata may be stored or logged for security, debugging, or abuse detection
- ✅ Customer data will not be used to train IBM foundation models without explicit permission
- 🚫 IBM does not explicitly guarantee that prompts and metadata are never shared with third parties
- 🚫 IBM does not advertise a cryptographic attestation mechanism

Privacy Guarantees

- 🚫 Data is encrypted at rest and in transit, but is not guaranteed to not be accessed by administrators
- 🚫 Interactions are associated with API keys for purposes like billing and analytics
- ✅ Data in transit is encrypted and can be verified via TLS certificates to ensure data integrity and authenticity
- 🚫 IBM may inspect logs or user content for reasons such as abuse, debugging, or legal compliance
- 🚫 IBM does not have confidential computing offerings

Compliance Guarantees

- ✅ IBM Cloud is SOC 2 Type 2 certified, and watsonx falls under the broader scope
- ✅ IBM Cloud is CSA STAR Level 1 certified
- ✅ IBM Cloud states it provides tools and contractual assurances to help customers comply with GDPR

References

IBM WatsonX · IBM Cloud Security · IBM Privacy Statement · IBM Compliance · STAR Registry

CONFIDENT SECURITY

Hugging Face Inference

4/12
AI Scorecard

View Product Page

Security Guarantees

- 🚫 Prompts and metadata may be stored or logged for security, debugging, or abuse detection
- ✅ Hugging Face states that customer data and prompts used through Inference Endpoints are not used to train or retrain shared models
- 🚫 Hugging Face does not explicitly guarantee that prompts and metadata are never shared with third parties
- 🚫 Hugging Face does not advertise a cryptographic attestation mechanism

Privacy Guarantees

- 🚫 Data is encrypted at rest and in transit, but is not guaranteed to not be accessed by administrators
- 🚫 Interactions are associated with API keys for purposes like billing and analytics
- ✅ Data in transit is encrypted and can be verified via TLS certificates to ensure data integrity and authenticity
- 🚫 Hugging Face may inspect logs or user content for reasons such as abuse, debugging, or legal compliance
- 🚫 Hugging Face does not have confidential computing offerings

Compliance Guarantees

- ✅ Hugging Face has publicly stated that it has achieved SOC 2 Type 2 compliance
- 🚫 Hugging Face Inference Endpoints is not stated to be STAR Level 1 compliant
- ✅ Hugging Face states it provides tools and contractual assurances to help customers comply with GDPR

References

Hugging Face Enterprise Endpoints

Hugging Face Security

Hugging Face FAQ

CONFIDENT SECURITY

Aleph Alpha Sovereign AI

1/12
AI Scorecard

View Product Page

Security Guarantees

- 🚫 Prompts and metadata may be stored or logged for security, debugging, or abuse detection
- 🚫 Aleph Alpha does not explicitly state that customer data and prompts are not used to train models
- 🚫 Aleph Alpha does not explicitly guarantee that prompts and metadata are never shared with third parties
- 🚫 Aleph Alpha does not advertise a cryptographic attestation mechanism

Privacy Guarantees

- 🚫 Data is encrypted at rest and in transit, but is not guaranteed to not be accessed by administrators
- 🚫 Interactions are associated with API keys for purposes like billing and analytics
- ✅ Data in transit is encrypted and can be verified via TLS certificates to ensure data integrity and authenticity
- 🚫 Aleph Alpha may inspect logs or user content for reasons such as abuse, debugging, or legal compliance
- 🚫 Aleph Alpha does not have confidential computing offerings

Compliance Guarantees

- 🚫 Aleph Alpha is not stated to be SOC 2 Type 2 compliant
- 🚫 Aleph Alpha is not stated to be STAR Level 1 compliant
- 🚫 Aleph Alpha is not stated to be GDPR compliant

References

Aleph Alpha: Sovereign AI Solutions

CONFIDENT SECURITY

Amazon Bedrock

5/12
AI Scorecard

View Product Page

Security Guarantees

- 🚫 Prompts and metadata may be stored or logged for security, debugging, or abuse detection
- ✅ Amazon has publicly committed that prompts and data sent to Amazon Bedrock are not used to train or improve foundation models without customer consent
- 🚫 Amazon does not explicitly guarantee that prompts and metadata are never shared with third parties
- 🚫 Amazon does not advertise a cryptographic attestation mechanism

Privacy Guarantees

- 🚫 Data is encrypted at rest and in transit, but is not guaranteed to not be accessed by administrators
- 🚫 Amazon logs requests and ties them to an AWS account for billing, auditing, and usage monitoring
- ✅ Data in transit is encrypted and can be verified via TLS certificates to ensure data integrity and authenticity
- 🚫 Amazon may inspect logs or user content for reasons such as abuse, debugging, or legal compliance
- 🚫 Amazon does not use their Nitro Enclaves to support Amazon Bedrock to provide a confidential computing offering

Compliance Guarantees

- ✅ Amazon AWS has SOC 2 Type 2 certification, and Amazon Bedrock likely falls under the broader scope
- ✅ Amazon AWS is CSA STAR Level 2 certified and Amazon Bedrock likely falls under the broader scope
- ✅ Amazon states it provides tools and contractual assurances to help customers comply with GDPR

References

Amazon Bedrock | Amazon Bedrock FAQ | AWS STAR Registry

CONFIDENT SECURITY

Perplexity Enterprise Pro

2/12
AI Scorecard

View Product Page

Security Guarantees

- 🚫 Prompts and metadata may be stored or logged for security, debugging, or abuse detection
- ✅ Perplexity Enterprise says that user prompts and queries are never used for training
- 🚫 Perplexity does not explicitly guarantee that prompts and metadata are never shared with third parties
- 🚫 Perplexity does not advertise cryptographic attestation or environment verification

Privacy Guarantees

- 🚫 Data is encrypted at rest and in transit, but is not guaranteed to not be accessed by administrators
- 🚫 Interactions are associated with API keys for purposes like billing and analytics
- ✅ Data in transit is encrypted and can be verified via TLS certificates to ensure data integrity and authenticity
- 🚫 Perplexity may inspect logs or user content for reasons such as abuse, debugging, or legal compliance
- 🚫 Perplexity does not have confidential computing offerings

Compliance Guarantees

- 🚫 Perplexity is not stated to be SOC 2 Type 2 compliant
- 🚫 Perplexity is not stated to be STAR Level 1 compliant
- 🚫 Perplexity is not stated to be GDPR compliant

References

Perplexity Enterprise Pro

Perplexity Privacy Policy

Perplexity Terms of Service

Confident Security

9/12
AI Scorecard

View Product Page

Security Guarantees

- ✅ Prompts and metadata are guaranteed to never be stored or logged
- ✅ User data is not used to train models
- ✅ Prompts and metadata are never shared with subprocessors
- ✅ Confident Security provides a cryptographic attestation mechanism for each model version at request time

Privacy Guarantees

- ✅ Data is encrypted at rest and in transit, and is guaranteed to not be accessed by administrators
- ✅ All interactions are provably guaranteed to be anonymous – no one can determine which user is associated with any individual interaction
- ✅ Data in transit is encrypted and can be verified via TLS certificates to ensure data integrity and authenticity
- ✅ Confident Security does not have access to inspect any private computation or user interactions
- ✅ Prompts and metadata cannot be decrypted if the operator's LLM environment does not match what is claimed or expected

Compliance Guarantees

- 🚫 Confident Security is not yet SOC 2 Type 2 compliant
- 🚫 Confident Security is not yet STAR Level 1 compliant
- 🚫 Confident Security is not yet GDPR compliant

References

Confident Security