

Kurahi セキュリティ概要・チェックシート回答 v1

製品：Kurahi 2.2.0（エージェントランタイム）・作成：KODA SōKEN チーム・日付：2026年10月25日

すべての回答は、Kurahiのドキュメントまたはソースツリーに基づいており、根拠となる文書を示しています。パスは公開リポジトリ github.com/koda-tokyo/kurahi のものです。同じ文書は <https://kurahi.kodasoken.com/docs/> でも公開しています（主に英語。セキュリティ概要、プライバシー、エンタープライズ ID、添付ファイルの取り込み、エンタープライズプロファイル、データガバナンス、オートパイロット、OWASP Agentic Top 10 は日本語版あり）。「該当なし（セルフホスト）」は、Kurahiがお客様のインフラで動作するため存在しないサービスに関する回答です。「未対応」は、現時点の不足をそのまま示しています。（KODA SōKENにて確認予定）の項目は、ご契約手続きの中で記入します。

A. 会社・サービス

#	質問	回答	根拠
A1	委託先・供給者の商号	KODA SōKEN（東京）。契約主体は書面契約に記載します。	会社概要ページ kurahi.kodasoken.com/ja/company
A2	所在地・法人番号・適格請求書発行事業者登録番号	（KODA SōKENにて確認予定）	—
A3	セキュリティ窓口	security@kodasoken.com	SECURITY.md
A4	提供形態	セルフホスト型ソフトウェア。アイデンティティごとに一つのゲートウェイプロセスが、お客様のマシンまたはサーバー上でOSのサービス管理機能のもとで動作します。評価構成に弊社運用のクラウド構成要素はありません。	docs/security/SECURITY_OVERVIEW.md §1
A5	供給者がお客様データを保管するインフラを運用しているか	該当なし（セルフホスト）。ゲートウェイは実行中にKurahiや弊社へ接続しません。	docs/security/SECURITY_OVERVIEW.md §2 、 docs/security/privacy.md
A6	再委託先（サブプロセッサー）	既定ではなし。データは、お客様が設定したモデルプロバイダー、チャンネル、業務コネクターにのみ送られます。	docs/security/privacy.md 、 docs/capabilities.md
A7	保険（サイバー保険・賠償責任保険）	（KODA SōKENにて確認予定）	—

B. 認証・第三者保証

#	質問	回答	根拠
B1	ISO/IEC 27001認証	未対応。取得していません。	トラストセンター kurahi.kodasoken.com/ja/trust
B2	SOC 2報告書	未対応。保有していません。	トラストセンター
B3	第三者による侵入テスト	未対応。第三者の侵入テスト報告書はありません。	—
B4	適合性テスト等	ランタイム継続性の適合性テスト（C0～C9）をCIで実行し、署名付きレポートを出力します。状態・ファイル・権限・会話記録の保持を確認するもので、モデル間で推論が同等であることは確認しません。	docs/capabilities.md 、 docs/continuity.md

#	質問	回答	根拠
B5	AI事業者ガイドライン	共通の指針のいくつかを支援できるよう設計しています。準拠や認証は主張しません。	AIガバナンスページ kurahi.kodasoken.com/ja/ai-governance
B6	OWASP Top 10 for Agentic Applications	KODA SōKEN チームによる自己評価（2026年10月24日時点）：対応済み4、一部対応6、未対応0。認証でも第三者監査でもありません。「対応済み」はリスクがなくなったことを意味しません。	docs/security/owasp-agentic.md

C. データ保護・プライバシー

#	質問	回答	根拠
C1	お客様環境の外に出るデータ	設定したプロバイダーへのモデルリクエスト、有効にしたチャネルのメッセージ、バージョン情報のみの更新確認（npmインストール時）、LAN内のmDNS通知、インストーラーのダウンロード。それぞれ無効化できます。業務コネクタ（Microsoft 365、Gmail、Entra ID、kintone、freee、マネーフォワード、SmartHR、LINE WORKS、Box）は既定で無効で、管理者が有効にしたものだけが各サービスとデータをやり取りします。エンタープライズ ID を使う場合、サインインは組織の ID プロバイダーを経由します。	docs/security/privacy.md、docs/security/SECURITY_OVERVIEW.md §2、docs/capabilities.md、docs/security/identity.md
C2	テレメトリ・アクセス解析・クラッシュレポート	一切送信しません。	docs/security/privacy.md
C3	データの所在地	お客様が選択：ローカルモデル、日本国内ホストのモデル（既定プロファイル）、外部プロバイダー、エアギャップ（別途ライセンス）。	docs/deployment.md
C4	データの保存場所	お客様が管理する一つのディレクトリ（KOLLO_HOME）。プログラムファイル以外は他の場所に書き込みません。	docs/security/SECURITY_OVERVIEW.md §1、§8
C5	保存期間	セッション・会話記録は削除まで、ログは日次ローテーション、監査チェーンは月次ファイル、サポートバンドルは作成時のみ。データ区分（会話記録、ワークアイテム、証跡、監査、添付ファイル、出力、メモリ、コネクタのキャッシュ）ごとの保存ルール、訴訟ホールド、署名付きの eDiscovery エクスポートを設定できます。保存ルールが指定されるまで各区分は無期限に保持され、削除された証跡には署名付きのトゥームストーンが残ります。	docs/security/SECURITY_OVERVIEW.md §8、docs/security/data-governance.md
C6	データの削除	kurahi uninstall --all --yes で設定・シークレット・セッション・監査・ログを削除。--dry-run で削除対象を一覧表示。	docs/security/SECURITY_OVERVIEW.md §9
C7	AIモデルに渡す個人情報の最小化	コンテキストガードが、ツールの結果に含まれる個人情報をモデルに渡る前にプレースホルダーへ置き換えます。制約：画像は検査せず、検出はパターンにより、利用者がチャットに入力した文章はそのまま送られます。	docs/context-guard.md
C8	サポートのために供給者がデータにアクセスするか	お客様が kurahi support で作成したサポートバンドル（秘匿処理済み）を送付された場合のみ。	docs/security/SECURITY_OVERVIEW.md §8

#	質問	回答	根拠
C9	日本の個人識別番号（マイナンバー）	Kurahi の外へ送信・書き込みをするツールの実行前に、最終的な引数のすべての文字列で日本の個人識別情報を検査します。既定では、チェックデジットが正しい個人番号を含む呼び出しは拒否され、監査ログと証跡ではマスクされます。	docs/security/data-governance.md

D. アクセス制御・認証

#	質問	回答	根拠
D1	既定のネットワーク公開範囲	ループバックのみ。リモートアクセスには明示的な gateway .bind 設定と、接続ごとのトークンが必要です。	docs/security/SECURITY_OVERVIEW.md §3
D2	API呼び出しの認証	ヘルスチェックを除き、すべての /api/* ルートと WebSocket メソッドに認証情報が必要です。認証試行は送信元ごとに毎分20回に制限されます。	docs/security/SECURITY_OVERVIEW.md §3
D3	権限モデル	端末ペアリングのスコープ（observe < attach < admin < owner）、短命のアクセストークン、ローテーションするリフレッシュトークン、失効、再同意（既定90日）。エンタープライズ ID では、人は部門・チームの中でロール（org-admin、security-admin、auditor、department-manager、approver、operator、member、viewer）を持ち、IdP のグループから対応付けられます。ゲートウェイの admin スコープを与えるロールはありません。	docs/security/SECURITY_OVERVIEW.md §4、docs/security/identity.md
D4	シングルサインオン（SAML/OIDC）	標準搭載：人のための OIDC（Entra ID、Google、Okta、Keycloak など）または SAML 2.0 のサインイン、SCIM 2.0 プロビジョニング、部門・チーム・ロール。既定で無効。Free を含むすべてのプランに含まれます。ssoRequired: true では、人のトークンログインは理由の記載が必要な監査付きブレイクグラスになります。	docs/security/identity.md、docs/capabilities.md、docs/releases/2.2.0.md
D5	エージェント間の認証情報の分離	auth.isolation: "explicit" では、各エージェントは付与された認証プロファイルだけを参照し、メインエージェントからのコピーや統合はありません。	docs/gateway/security/index.md（Credential isolation between agents）
D6	多要素認証	該当なし（セルフホスト）。端末ペアリングはEd25519の鍵所持証明を用います。ホストへのログインはお客様の管理です。エンタープライズ ID では、人は組織の ID プロバイダーでサインインします。	docs/security/SECURITY_OVERVIEW.md §4、docs/security/identity.md

E. シークレット・暗号

#	質問	回答	根拠
E1	シークレットの保管	KOLO_HOME 配下の認証プロファイルストアと暗号化されたシークレットストア。ゲートウェイトークンはファイルモード0600。	docs/security/SECURITY_OVERVIEW.md §4、§5
E2	ログ・コマンドラインへのシークレットの露出	表示せず、コマンドライン引数で受け付けず、ログでは秘匿します。	docs/security/SECURITY_OVERVIEW.md §5
E3	保存データの暗号化		

#	質問	回答	根拠
		シークレットストアのみ。ホストのディスク全体暗号化を推奨します。Kurahiはそれ以外のディスク暗号化を追加しません。	docs/security/SECURITY_OVERVIEW.md §8
E4	通信の暗号化	リモートクライアントはHTTPS/WSS（例：tailnet経由）を推奨。リモート wss:// ではTLSフィンガープリントの固定が可能。既定のループバックでは通信はホスト内に留まります。	docs/gateway/security/index.md
E5	電子署名	ポリシー文書、証跡バンドル、AIインベントリはEd25519で署名されます。	docs/security/SECURITY_OVERVIEW.md §7a、docs/evidence.md、docs/inventory.md

F. エージェントの統制・AI固有のリスク

#	質問	回答	根拠
F1	エージェントの操作の制限方法	ゲートウェイごとの署名付きポリシー：ツールの許可・拒否、プロバイダー、モデル、チャンネル、予算、パス、外部通信とMCPの許可リスト。すべてのツール呼び出し、モデルリクエスト、チャンネルメッセージ、ジョブ開始時に適用。	docs/policy.md、docs/security/SECURITY_OVERVIEW.md §7a
F2	重要な操作の人による承認	ツールの種類ごとに、ターミナルまたはチャットチャンネルでの承認を必須にできます。承認待ちはポリシーのタイムアウトで区切られます。	docs/policy.md、docs/capabilities.md
F3	緊急停止	kurahi policy kill --agent <id> --all。実行中のジョブは中止され、停止状態は解除まで再起動後も維持されます。	docs/security/SECURITY_OVERVIEW.md §7a
F4	外部エージェントCLI	アダプターごとに、実行前に強制される事項、実行後に検知のみの事項、統制外の事項を互換性契約で明記。	docs/agents/external-adapters.md
F5	重要ファイルの保護	snapshots.requiredにより、復元用スナップショットを取得できない場合は保護対象ファイルの編集を拒否。サンドボックス内の編集とシェルコマンドは対象外。	docs/tools/snapshots.md、docs/capabilities.md
F6	AIシステムの台帳	kurahi inventory がエージェント、モデル、ツール、ポリシールール、MCPサーバー、チャンネル、証跡を一覧化。署名付きで書き出し、検証可能。	docs/inventory.md
F7	ランサムウェア	検知のみ：おとりファイルと大量暗号化の検知で全エージェントを停止し、監査ログに記録。検知前に暗号化されたファイルは損なわれ、	docs/shield.md、docs/

#	質問	回答	根拠
		停止するのはエージェントでありマルウェアではなく、マシン外への複製は保持しません。	capabilities.md
F8	プロンプトインジェクション	外部からのコンテンツ（ウェブページと検索結果、リモートエージェント、外部フック、業務コネクタの結果）は、信頼できないコンテンツのマーカーで囲んでモデルに渡します。そうした入力後のメモリ書き込みは隔離され、重要な操作にはポリシーの判断か人の承認が必要です。マーカーはプロンプトインジェクションのリスクを減らしますが、防ぎきるものではありません。チャネル利用者のメッセージと MCP の結果はマーカーで囲まれません。評価は一部対応（ASI01）。	docs/security/owasp-agentic.md
F9	受信ファイルとマルウェア	メールの添付ファイルは隔離し、静的なファイルポリシー（実行ファイル、スクリプト、マクロ、暗号化アーカイブは拒否）で確認し、ClamAV または Microsoft Defender でスキャンしてから Project に保存します。テキストは別の制限付きプロセスで抽出します。スキャンは失敗時に拒否（フェイルクローズ）します：タイムアウト、異常終了、判定のない出力は決してクリーン扱いにならず、既定（requiresCan）では未スキャンのファイルは拒否されます。残るリスク：未知（ゼロデイ）のマルウェア。	docs/security/attachment-intake.md、docs/capabilities.md
F10	業務環境でのコンシューマー向けチャネル・プラグイン	エンタープライズパッケージにはコンシューマー向けチャネルのプラグインとコンシューマー向けログインを含めません。エンタープライズプロファイルでは、除外されたプラグインを誰かがコピーして入れてもランタイムが拒否し、拒否を監査ログに記録します。コンシューマー向けチャネルのコードの一部はコアに残りますが不活性です：チャネルプラグインが登録したときだけ動作し、そのプラグインは同梱されていません。	docs/security/enterprise-profile.md、docs/releases/2.2.0.md
F11	自律的な操作（オートパイロット）	ルーティンは人が設定した自律レベル（L0 観察～L3 ポリシー内で外部に実行）で動作します。判断するモデルはツールを持たず、エンジンはレベルが許す操作だけを実行します。外部への操作には、宛先・テンプレート・1日の上限・有効期限を限定した、人による常設承認が必要です。行動ウォッチドッグがパターンを外れたルーティンを一時停止し、グローバル停止ですべてのルーティンを止められます。	docs/autopilot.md、docs/releases/2.2.0.md

G. ログ・監視・証跡

#	質問	回答	根拠
G1	監査ログ	認証、ペアリング、アタッチ・デタッチ、管理操作、ツール実行をハッシュチェーンの監査ログに追記。kurahi audit verify でチェーンを検証。	docs/security/SECURITY_OVERVIEW.md §6
G2	改ざん耐性	改ざんは検知できますが、追記専用ストレージへ転送しない限り防止はできません。	docs/security/SECURITY_OVERVIEW.md §6
G3	SIEMへの監査ログ連携	別途ライセンス（エンタープライズ機能）。	docs/capabilities.md

#	質問	回答	根拠
G4	エージェントの作業の証跡	ジョブごとの署名付き・ハッシュチェーン化された証跡バンドル。ku rahi evidence verify でオフライン検証。オートパイロットの実行 ごとにも署名付きの証跡バンドルを作成し、そのルーティンの前回の 実行と連結します。署名付き証跡が示すのは記録の完全性と来歴であ り、判断が正しかったことではありません。	docs/ evidence. md、docs/ autopilot .md
G5	設定の完全性	起動時の整合性ゲートが保護対象のポリシーファイルを毎回再検証。 enforceモードでは不一致時に起動しません。	docs/ security/ SECURITY_ OVERVIEW. md §7

H. セキュアな開発・サプライチェーン

#	質問	回答	根拠
H1	バイナリのコード署名	未対応。現在のビルドはコード署名・公証をして いません。	docs/capabilities.md
H2	リリースの完全性	各リリースに SHA256SUMS.txt。インストーラー は不一致のダウンロードを拒否。	docs/security/ SECURITY_OVERVIEW.md §10
H3	SBOM・来歴証明	各リリースにCycloneDX形式のSBOMとビルドの 来歴証明 (gh attestation verify)。	docs/security/ SECURITY_OVERVIEW.md §10、d ocs/PACKAGING.md
H4	依存関係の管理	pnpm-lock.yaml で固定、CIでシークレッツ キャナーを実行、Dependabotを設定済み。	docs/security/ SECURITY_OVERVIEW.md §10、 .github/dependabot.yml
H5	テスト	約9,700件の自動テスト（ユニット、ゲートウェ イ、エンドツーエンド、拡張機能）と、CIでの適 合性チェック。	Kurahiホームページの数値、 .github/workflows/ci.yml

I. 脆弱性・インシデント管理

#	質問	回答	根拠
I1	脆弱性の報告窓口	security@kodasoken.com へ非公開で報告。	SECURITY.md
I2	対応目標	受領連絡3営業日以内、トリアージ10営業日以内、確認された 高・重大な問題の修正または緩和30日以内、協調開示は報告か ら最長90日。	SECURITY.md
I3	お客様へのインシデント通 知	(KODA SōKENにて確認予定) (サポート条件で定めます)。	—
I4	サービスの事業継続	該当なし（セルフホスト）。ソフトウェアは弊社がなくても動 作し続け、KOLO_HOME のバックアップはお客様が管理します。	docs/security/ SECURITY_OVERV IEW.md §1、§8

J. ライセンス・評価

#	質問	回答	根拠
J1	ライセンス形態		docs/ security/

#	質問	回答	根拠
		オープンコア：MIT（プロトコル、SDK、プラグインAPI、CLI）、FSL-1.1-Apache-2.0（ゲートウェイランタイム）、プロプライエタリ（アイデンティティエンジン、エンタープライズ機能）。	SECURITY_OVE RVIEW.md §12
J2	評価条件	90日間の評価ライセンス。現在は法務レビュー待ちの草案です。	docs/ EVALUATION_L ICENSE.md