クラウド利活用におけるガバナンスとリスク対策

Governance and Risk Countermeasures in Cloud Utilization

伊 藤 直 行

要 約 パブリッククラウドの利用拡大に伴い，IT部門が自社の情報システム基盤として活用するだけでなく，事業部門がDXビジネスの付加価値訴求を目的としたサービス基盤として利用する企業が増えている．IT部門の要員数が必ずしも十分ではない企業では，パブリッククラウドを利用する上でのセキュリティ対策の標準ルール策定を含めたクラウドガバナンスや運用体制の整備が，本番開始後の後追いでの対応になる傾向がある．クラウドガバナンス不全の状態に陥らないようにするには，クラウド利用に伴うリスクを評価し，それらのリスク対応案を検討・コントロールすることで，クラウドガバナンス体制を強化することが重要である．また，OWASPやNISTが公開しているフレームワークは，生成AIのセキュリティリスクへの有益な対策ガイダンスとして，技術的なリスク項目の認識と対策の評価に有効である．

Abstract As the use of public clouds expands, an increasing number of companies are using them not only as their own information system platforms in IT departments, but also as service platforms in business departments for the purpose of promoting added value in DX business. For companies that do not necessarily have a sufficient number of IT department personnel, the maintenance of cloud governance and operational systems, including the formulation of standard rules for security measures when using public clouds, tend to be a follow-up response after the start of production. In order to avoid falling into a state of cloud governance failure, it is important to strengthen the cloud governance system by evaluating the risks associated with cloud usage, and considering and controlling plans to deal with those risks. In addition, the frameworks published by OWASP and NIST are effective in recognizing technical risk items and evaluating countermeasures as useful guidance against security risks of generated AI.

1. はじめに

新型コロナウイルス感染症の世界的な拡大により，リモートワークの活用が急速に広がった．こうした「ニューノーマル」による働き方の変化に伴い，パブリッククラウドの利用が加速している．2023年7月に総務省が発表した「令和5年 情報通信に関する現状報告（令和5年版情報通信白書）[1]」では，2021年から2026年までの国内企業のクラウドサービスの利用状況において，コロナ禍後もリモートワーク需要喚起等の継続要因により，クラウド利用が今後一層加速することが予想されている．

パブリッククラウドの利用拡大に伴い，企業における利活用シーンも様々な状況になりつつあり，IT部門が自社の情報システム基盤として活用するだけでなく，事業部門がデジタルトランスフォーメーション（以降，DX）ビジネスの付加価値訴求を目的としたサービス基盤として利用するケースが見受けられる．パブリッククラウドの導入時の敷居の低さ，構築期間の短さがその普及要因として考えられるが，こうした状況はパブリッククラウドの乱用・乱立を

引き起こし，企業として管理できないシャドー IT の温床に繋がる可能性がある．

本稿では，Microsoft Azure（以降，Azure）や Amazon Web Services（以降，AWS）に代表されるパブリッククラウドを利活用する際の IT ガバナンスの考え方，ガバナンス不全が企業に与える影響や問題点，及びそうした問題点を解決するためのコントロールについて，クラウド業界におけるクラウド導入フレームワークというトレンドを踏まえながら解説する．また，利用者側の立場に立った客観的な視点として，ENISA（European Network and Information Security Agency[*1]），NIST（National Institute of Standards and Technology[*2]），OWASP（Open Web Application Security Project[*3]）等に代表されるクラウド利活用のリスク評価のフレームワークの活用方法や，クラウド業界での最新のトピックスとして生成 AI 利用時のリスク対策の考え方について解説する．まず 2 章でクラウドガバナンスと国内企業での実情を述べ，3 章でパブリッククラウドベンダーが提供するフレームワークとクラウドガバナンスの考え方を述べた後，4 章でクラウドリスク評価の進め方とフレームワーク活用方法を述べる．また，広域の意味でのパブリッククラウドサービスにおいて，生成 AI のクラウドサービスには，従来の IaaS，PaaS，SaaS とは異なる，特有のリスク対策があることから，5 章では，生成 AI のリスク対策に関するガイドラインとフレームワーク活用方法を述べる．

2. クラウドガバナンスの必要性

世界的に SDGs への社会的な注目度が高まりつつある状況を受けて，国内でも SDGs を達成することを意識した企業経営を導入・推進する企業が増加してきている．その企業の投資判断基準として，今後，ESG 指標[*4] が重要な位置付けになってくる可能性が高い．ESG の三つの指標の中でも，働き方改革を始めとしたイノベーションの視点と，コーポレートガバナンス強化としての内部統制の視点が，企業価値を向上・可視化する重要なポイントとなり，それぞれに対するバランス感覚を持った経営手腕が国内企業の経営層に求められている．

本稿では，パブリッククラウドやプライベートクラウドといったクラウドを対象にした IT ガバナンス（2.1 節）を「クラウドガバナンス」と定義付け，金融業界での事例や，ガバナンス不全に至る企業の背景，ガバナンス不全による企業への影響・リスクの説明を交えながら，クラウドガバナンスの必要性を解説する．

2.1 DX の拡大に伴う IT ガバナンスの必要性

本節では，一般的な IT ガバナンスの考え方として金融業界の事例を紹介する．金融庁は，主な業態（メガバンク，地方銀行，大手生損保）のいくつかの金融機関との対話や有識者に対するヒアリングを重ねて，2023 年 6 月に「金融機関の IT ガバナンスに関する対話のための論点・プラクティス整理　第 2 版[2]」で IT ガバナンスの概念を整理している．有識者や金融機関との議論の中で，金融庁は，IT ガバナンスを「経営者がリーダーシップを発揮し，IT と経営戦略を連携させ，企業価値の創出を実現するための仕組み」と定義している．それは IT ガバナンスの代表的なフレームワークである COBIT[*5] による考え方と齟齬はないものの，企業価値を創出するビジネス・業務の変革的な動きとして DX を経営戦略及び IT 戦略に取り込むことを前提としている点が特徴的である．こうした考え方は金融業界に特化したものではなく，様々な業界・業種において，DX ビジネスの拡大を下支えするクラウドを活用する上で，IT ガバナンス，ひいてはクラウドガバナンスが非常に重要な概念であることを意味している．

2.2　クラウドガバナンスに係る国内企業の背景

国内企業では，全社の事業部門のビジネス収益に貢献するサービス基盤として，プライベートクラウド（オンプレミス基盤）を導入し，その活用を推進しているケースがある．そうした企業では，プライベートクラウドの利用者である事業部門において，以下の点が課題となることがある．

・DX ビジネス用の先進的なサービス（ビッグデータ，AI/IoT，ブロックチェーン，ローコード/ノーコード開発基盤，量子コンピューティング等）が利用できない
・自社内にシステム関連機器（サーバ，ストレージ，ネットワーク等）を設置するため，調達や導入・設置に時間がかかり，タイムリーに利用できない

こうした課題を背景に，IT 部門主導でクラウドガバナンス体制を全社に浸透させる前に，事業部門が，DX ビジネスから求められるスピード感に対応するためにパブリッククラウドを先行利用するケースが見受けられる．これらの企業では IT 部門の要員数は必ずしも十分ではない事情もあり，パブリッククラウドを利用する上でのセキュリティ対策の標準ルール策定を含めたクラウドガバナンスや運用体制の整備は，本番開始後の「後追い」での対応になる傾向がある．全社的にクラウドガバナンスを効かせ，成熟度の高い利活用ができている企業もあるが，国内の多くの企業においては，クラウドガバナンスの整備が追い付かずに，クラウドガバナンス不全になるリスクを抱えているのが実情であろう．

2.3　クラウドガバナンス不全の問題点

パブリッククラウドを活用したシステムがガバナンス不全を引き起こす主な問題点や，それぞれの原因，想定される影響とリスクを表 1 に列挙する．不正アクセス等による情報漏洩が発生した場合，高額な賠償金や対応コストの支払いによって企業経営に多大な影響を及ぼす．表 1 に列挙した問題点は，本番運用前に検討・対処しておくことで，対応策を事前に施せるものが大半を占めており，システムのクラウド移行の計画フェーズや上流工程（要件定義，基本設計）といった前工程でこれらを検討することが重要である．

表 1　クラウドガバナンス不全による主な問題点とそれぞれの原因と想定影響・リスク

No.	クラウドガバナンス不全による主な問題点	主な原因	想定される影響とリスク
1	不適切なロール/権限付与	・運用現場での職務分掌ルールの不在 ・定期的なアカウント・権限の管理プロセス不在	・機密情報に対する不正アクセスによる情報漏洩リスク
2	不適切なクラウドサービス利用	・クラウドリソース把握の管理プロセス不在 ・クラウドサービスの適切な利用方法の理解不足 ・クラウドサービスプロバイダーとの責任範疇の理解不足 ・IT リソース増設時の適切な承認プロセス不在 ・組織横断での一元的なコスト把握プロセス不在	・事前想定コスト超過による IT 予算の圧迫

3	リソース乱用によるシャドー IT 発生	・適切なクラウドリソース管理，構成管理プロセス不在 ・全社横断でのシステム利用方針ポリシー，IT 管理規程不在	・シャドー IT 経由での情報漏洩リスク
4	システム間でのノウハウ共有不足	・組織間でのノウハウ共有プロセス不在 ・組織間のセクショナリズム ・全社横断でのシステム利用方針ポリシー，IT 管理規程不在	・非効率な IT 投資 ・非効率なシステム構築プロジェクト運営 ・サービス品質の低下（不均一化）

3. パブリッククラウドが提供するガバナンスフレームワーク

Azure や AWS といったパブリッククラウドベンダーは，オンプレミスからパブリッククラウドへの移行を推進している企業向けに「クラウド導入フレームワーク（CAF）」を公開している．これらは，企業のクラウド導入を成功に導くために開発された，戦略策定から運用までの様々なフェーズ毎の各種方法論やベストプラクティス，ツールの集合体であり，パブリッククラウド利用者の成功体験に基づいて整理された良い手本となっている．

Microsoft 社の「Cloud Adoption Framework for Azure（以降，CAF for Azure）[3]」は，顧客のクラウド導入ライフサイクル全体に対応するように設計されたフレームワークである．このフレームワークでは，クラウド導入の短期的および長期的な目標を達成するためのベストプラクティスやツールが提供されている．

一方，AWS 社の「AWS Cloud Adoption Framework（以降，AWS CAF）[4]」は，パブリッククラウドへ確実に移行するためのガイダンスという位置付けになっている．AWS CAF は同社から公開されている「An Overview of the AWS Cloud Adoption Framework Version 2[5]」で整理されている通り，「パースペクティブ（視点）」と呼ばれる六つの重点分野毎に編成されている．これらのパースペクティブは，機能的に関連するステークホルダーが所有または管理する明確な責任範囲をカバーしており，一般的に，BUSINESS，PEOPLE，GOVERNANCE のパースペクティブではビジネス遂行能力に焦点を当て，PLATFORM，SECURITY，OPERATIONS のパースペクティブでは技術的能力に焦点を当てている．

3.1 CAF for Azure におけるクラウドガバナンスの考え方

パブリッククラウドの活用におけるビジネスプロセスまたは技術プラットフォームの変更には様々なリスクを伴うが，企業においてクラウドガバナンスを推進するチームは，こうしたリスクを軽減し，クラウドの導入やイノベーションの作業をできるだけ中断させないことが求められる．図 1 は，CAF for Azure におけるクラウドガバナンスの主要スコープとして，企業におけるポリシーの策定と統制すべき五つの規範，及び関連する Azure の機能を示している．

これらの規範は，クラウドサービス基盤の枠を超えた企業ポリシーの自動化と実施の適切なレベルに関する意思決定の指針となり，コスト管理，セキュリティベースライン，リソースの整合性，ID ベースライン，展開の加速の五つで構成されている．従来型のアプローチでは，ガバナンス不全による事故や改ざんを防止することを重視し，承認ワークフローによる管理でビジネススピードを犠牲にしていた．五つの規範が特徴的なのは，組み込み型のポリシーをベースとして管理することでガバナンスとビジネスへの俊敏性の両立を目指している点である．

図 1　CAF for Azure におけるクラウドガバナンスの主要スコープと関連する Azure の機能[6]

3. 2　AWS CAF におけるクラウドガバナンスの考え方

AWS CAF におけるガバナンスパースペクティブ[7]では，IT ガバナンスと組織ガバナンスを統合し，IT ガバナンスのベストプラクティスの特定と実装，およびテクノロジーを使用したビジネスプロセスのサポートに関するガイダンスを提供している．表 2 にガバナンスパースペクティブの主要な七つの管理機能を列挙する．

表 2　AWS CAF のガバナンスパースペクティブにおける管理機能

No.	管理機能	管理機能の概要
1	プログラムおよびプロジェクト管理	相互依存するクラウド施策を柔軟かつ調整的に実行する．
2	ベネフィット管理	クラウド投資から確実にビジネス利益が生まれ，持続するようにする．
3	リスク管理	クラウドを活用してリスクプロファイルを低減する．
4	クラウドの財務管理	クラウド利用料を計画，測定，および最適化する．IT システム，サービス，およびソフトウェアに必要なライセンスを調達し，配布し，管理する方法も定義する．
5	アプリケーションポートフォリオ管理	ビジネス戦略を支えるアプリケーションポートフォリオを管理し，最適化する．
6	データガバナンス	データに対する権限とコントロールを行使してステークホルダーの期待に応える．
7	データキュレーション	データカタログのデータプロダクトのインベントリを整理する．

特徴的なのは，クラウドの財務管理において，ライセンス管理についても言及しており，誤ったライセンス利用による損害請求リスクを意識している点である．パブリッククラウドを利用するサービス形態（IaaS，PaaS，SaaS）やライセンス持ち込み（BYOL[*6]）の有無に応じて，オンプレミスでのライセンス管理とは異なるケースがあり，明示的な管理を利用者側に促している．

4. クラウドリスク

一般的な定義として，クラウドガバナンスは企業の経営層がクラウド戦略の一環として実施すべき上位概念であり，その構成要素をなすのがクラウドリスク評価である．クラウドガバナンスは，クラウド戦略に基づいたクラウドの全社的な利活用方針を策定する際の様々なルール作りの起点となり，クラウド利用に伴う企業リスクを最小限にコントロールするための概念となる．

2.3 節で述べたようなクラウドガバナンス不全の状態に陥らないためには，クラウド利用に伴うリスクを評価し，それらのリスク対応案を検討・コントロールすることで，クラウドガバナンス体制を強化することが重要である．本章ではクラウド利活用に伴うリスク評価の進め方とリスク評価フレームワークの活用方法を解説する．

4.1 クラウドリスク評価の進め方

クラウドを利用する企業は，クラウドベンダー選定と絡めてリスク評価とその対策立案を実施する．BIPROGY 株式会社（以降，BIPROGY）は，客観的な第三者（利用者）視点に立った汎用的なリスクフレームワークを活用し，選定候補の複数のクラウドを同じ評価軸に基づいて客観的に相対評価することが有効であると考えている．リスク項目の評価軸はクラウドに特有の技術的な観点の他，法務，契約，コストと多岐にわたることが多いため，企業として特に重視すべき評価軸に重み付けを加えて，メリハリのある評価結果とすることが肝要である．

リスク評価も含めたクラウドベンダーの選定評価を実施した後，3 章で述べた CAF を参考にしながら，選定したクラウドの特色に沿ったリスク対策やガバナンスの施策を進める．企業内でリスク評価のフレームワークや CAF を使いこなせない場合は，専門用語の知識や進め方の実績という観点で，IT ベンダーの支援も含めて進め方を検討すると良い．

4.2 クラウドリスク評価のフレームワーク

本節ではクラウドリスク評価のフレームワークの有効性について述べる．対象となるフレームワークの一覧は表 3 の通りである．これらのフレームワークを採用する一般的な有効性は以下 1）～ 3）の 3 点である．

1）クラウドやセキュリティ，リスクマネジメントの有識者による客観的な意見が集約された評価項目であり，サービス提供主体であるクラウドベンダーの視点とは一線を画した内容であること．
2）発行元の団体と企業が連携して記載内容の検証を行っており，必要に応じて継続的な改訂が行われている．クラウドの日進月歩の変化に則した評価項目が，利用者側でタイムリーに利活用できること．
3）クラウド上のシステムに対する外部からのシステム監査において，リスク評価項目の妥当性及び合理性を，監査主体（監査法人等）とスムーズに認識共有できること．

BIPROGY ではこの中で特に 1）を重要視している．本節では，リスク項目の客観性と網羅性の広さという観点で No.1 の ENISA と No.2 の JASA を取り上げて，それぞれの詳細を説明する．

表3　主なクラウドリスク評価のフレームワーク

No.	基準・ガイドライン	制定・発行元	内　　容
1	Cloud Computing: Benefits, risks and recommendations for information security[8]	ENISA（European Network and Information Security Agency）	欧州の公共機関である ENISA が発行したクラウドコンピューティングに関するリスク評価のガイドライン（邦題：情報セキュリティに関わる利点，リスクおよび推奨事項[9]）
2	クラウドサービスにおけるリスクと管理策に関する有識者による検討結果 2011 年度版[10]	JASA（日本セキュリティ監査協会）	ENISA のクラウドリスク項目をベースにして作成されたクラウドリスク評価のフレームワーク
3	Guidance for Critical Areas of Focus in Cloud Computing	CSA（Cloud Security Alliance）	クラウドのサイバーセキュリティに特化したリスク評価フレームワーク
4	FISC 安全対策基準　追補版 v.9	金融情報システムセンター（FISC）	金融事業者向けの FISC 安全対策基準のクラウド版として，利用者と事業者のそれぞれが取るべきセキュリティとデータセンター対策の管理基準がまとめられている．

4. 2. 1　ENISA「Cloud Computing: Benefits, risks and recommendations for information security」

グローバルなクラウドリスク評価のフレームワークとして ENISA のガイドラインを解説する．2011 年に IPA によって日本語に翻訳されており[9]，想定読者はクラウドの利用企業である．

ENISA では全部で 35 個のクラウドリスクを以下の四つのカテゴリに分類している．

また，それぞれのリスクの頻度（最高・高・中・低・最低・NA（非該当）の 6 段階）と影響度（最高・高・中・低の 4 段階）から，リスクの高さ（高・中の 2 段階）で評価している．企業がこのリスクフレームワークを活用する場合は，リスクの高さが「高」の九つのリスク（R1-3，R9-10，R21-23，R26）と共に，自社の事情に合わせて重要視すべきリスクを追加選定し，リスク項目毎にリスク対策を検討するアプローチが有効である．

1）ポリシーと組織関連リスク（表 4：R1 ～ R7）
2）クラウドの技術関連リスク（表 5：R8 ～ R20）
3）クラウドの法的なリスク（表 6：R21 ～ R24）
4）クラウドに特化していないリスク（表 7：R25 ～ R35）

表4　ENISA によるクラウドのポリシーと組織関連のリスク（R1-R7）

No.	リスク項目	頻　度	影響度	リスク
R1	ロックイン	高	中	高
R2	ガバナンスの喪失	最高	最高	高
R3	コンプライアンスの課題	最高	高	高
R4	他の共同利用者の行為による信頼の喪失	低	高	中
R5	クラウドサービスの終了または障害	NA	最高	中
R6	クラウドプロバイダの買収	NA	中	中
R7	サプライチェーンにおける障害	低	中	中

表5　ENISA によるクラウドの技術的なリスク（R8-R20）

No.	リスク項目	頻　度	影響度	リスク
R8	リソースの枯渇（リソース割当の過不足）	中/低	低/中	中
R9	隔離の失敗	低/中	最高	高
R10	クラウドプロバイダ従事者の不正－特権の悪用	中	最高	高
R11	管理用インタフェースの悪用（操作，インフラストラクチャアクセス）	中	最高	中
R12	データ転送途上における攻撃	中	高	中
R13	データ漏えい（アップロード時，ダウンロード時，クラウド間転送）	中	高	中
R14	セキュリティが確保されていない，または不完全なデータ削除	中	最高	中
R15	DDoS 攻撃（分散サービス運用妨害攻撃）	中/低	高/最高	中
R16	EDoS 攻撃（経済的な損失を狙ったサービス運用妨害攻撃）	低	高	中
R17	暗号鍵の喪失	低	高	中
R18	不正な探査またはスキャンの実施	中	中	中
R19	サービスエンジンの侵害	低	最高	中
R20	利用者側の強化手順と，クラウド環境との間に生じる矛盾	低	中	中

表6　ENISA によるクラウドの法的なリスク（R21-R24）

No.	リスク項目	頻　度	影響度	リスク
R21	証拠提出命令と電子的証拠開示	高	中	高
R22	司法権の違いから来るリスク	最高	高	高
R23	データ保護に関するリスク	高	高	高
R24	ライセンスに関するリスク	中	中	中

表7　ENISA によるクラウドに特化していないリスク（R25-R35）

No.	リスク項目	頻　度	影響度	リスク
R25	ネットワークの途絶	低	最高	中
R26	ネットワークの管理（ネットワークの混雑，接続ミス，最適でない使用）	中	最高	高
R27	ネットワークトラフィックの改変	低	高	中
R28	特権の（勝手な）拡大	低	高	中
R29	ソーシャルエンジニアリング攻撃（なりすまし）	中	高	中
R30	運用ログの喪失または改ざん	低	中	中
R31	セキュリティログの喪失または改ざん（フォレンジック捜査の操作）	低	中	中
R32	バックアップの喪失，盗難	低	高	中
R33	構内への無権限アクセス（装置その他の設備への物理的アクセスを含む）	最低	高	中
R34	コンピュータ設備の盗難	最低	高	中
R35	自然災害	最低	高	中

4. 2. 2　JASA「クラウドサービスにおけるリスクと管理策に関する有識者による結果検討 2011 年度版」

本項では日本セキュリティ監査協会（Japan Information Security Audit Association：以降，JASA）のクラウドリスク項目（クラウドサービスにおけるリスクと管理策に関する有識者による結果検討 2011 年度版[10]）について解説する．基本的には前項の ENISA のリスク項目に基づいているものの，JASA はセキュリティ監査の団体であることから，クラウド情報セキュリティの管理基準としてクラウドサービス事業者の視点で読み替えを行っている．事業者視点の対策として考慮すべきでない項目（ENISA のリスク項目 No. R3, R5, R6, R20, R25-35）を除外して，リスクの重大性に応じて高（H），中（M），低（L）の 3 段階に分類し，さらに JASA として重要と考えられる以下の 2 項目を追加している．前項と同様，自社の事情に合わせて重要視すべきリスクを追加選定し，リスク項目毎にリスク対策を検討するアプローチが有効である．

・リソース・インフラの高集約によるインシデントの影響の拡大
・仮想/物理の設計・運用の不整合

5.　生成 AI リスク

本章では，クラウド業界の最新のトピックスとして，世界中で利用が加速している生成 AI のリスク対策について解説する．これからクラウド上の生成 AI サービスを活用し，企業内外に向けて様々な利活用シーンを検討しようとしている読者に対し，海外発の AI リスク対策のガイダンスやフレームワークの概要を紹介する．リスク評価の進め方を検討する際の有益な情報として参考になれば幸いである．

5. 1　生成 AI のリスク対策ガイドラインやフレームワーク

日本国内での OpenAI を始めとした生成 AI の利活用は，コロナ禍後の経済的回復を重視すべく日本国政府が後押ししており，統制や法規制等のブレーキを踏まずに，急速な利活用のアクセルを踏んでいる状況が続いている．一方，世界各国では，既に AI リスクが顕在化しており，米国では画像生成 AI の開発元への訴訟，韓国では企業の機密情報の漏洩等が起きている．そうした状況を踏まえ，EU を始めとした欧米諸国では，ChatGPT の利用規制が施行されており（イタリア），統一規制の検討や，公的機関によるリスク評価の包括的なフレームワークガイダンスの整備・普及も進んでいる．AI に対するガバナンスやリスクマネジメントについては G7 先進国の中でも日本は非常に遅れている状況である．

昨今，対話型生成 AI のコア機能である大規模言語モデル（Large Language Models：以降，LLM[*7]）への取り組みが世界各国で急速に進んでいる中で，日本国内では 2023 年 5 月に，日本ディープラーニング協会（Japan Deep Learning Association：以降，JDLA）から生成 AI の利用ガイドライン[11]がリリースされたものの，本稿執筆時点（2023 年 9 月末）でリスク評価にフォーカスしたフレームワークは国内の政府機関や公共機関含めて，まだ公開されていない．一方，海外においても，サイバーセキュリティ対策を観点とした政府機関や公共機関からの対策ガイダンスやフレームワークがドラフト版として徐々に公開されつつあるものの，一般的な利用者にまでは十分浸透していないのが現状である．本節では，生成 AI のセキュリティ

リスクへの有益な対策ガイダンスとして，OWASP と NIST から公開されたフレームワークの活用方法を解説する．

5. 1. 1　OWASP「OWASP Top 10 for LLM」「OWASP Top 10 for Large Language Model Applications」

OWASP Top 10 とは，OWASP によって発表された Web アプリケーションのセキュリティ上の最も重要な 10 個のリスクのリストである．このリストは，開発者，セキュリティ専門家，組織にとって，Web アプリケーションのセキュリティ強化のためのガイドラインとして広く普及しており，WAF 等のクラウド上のセキュリティ対策サービス（Azure WAF，AWS WAF 等）へも標準実装されている．

OWASP は，LLM を利用するアプリケーションで発見された 10 個の重大な脆弱性とセキュリティガイドラインに関するレポートとして「OWASP Top10 for LLM[12]」と「OWASP Top 10 for Large Language Model Applications[13]」を公開した．このレポートは，LLM テクノロジーを活用したアプリケーションとプラグインの設計と構築を担当する開発者，データ・サイエンティスト，セキュリティ専門家を主な想定読者としている．OWASP Top 10 for LLM は，500 人近くの専門家からなる国際チームの集合的な専門知識に基づいて作成されており，こられのチームが潜在的な脆弱性について話し合った結果，最も重大な以下の 10 個の脆弱性に絞り込んでいる．これから生成 AI システムを外部向けに公開する企業にとっては，これら 10 個の重大な脆弱性への対策状況をセキュリティ監査部門のチェック項目にすることで，セキュリティリスクの予防的統制に役立てることができるだろう．それぞれの脆弱性の詳細については，OWASP Top10 for LLM の URL[12]を参照されたい．

・Prompt Injection（プロンプト・インジェクション）
・Insecure Output Handling（安全でない出力処理）
・Training Data Poisoning（トレーニングデータの汚染）
・Model Denial of Service（モデルのサービス（DoS）攻撃）
・Supply Chain Vulnerabilities（サプライチェーンの脆弱性）
・Sensitive Information Disclosure（機密情報の開示）
・Insecure Plugin Design（安全でないプラグイン設計）
・Excessive Agency（過剰な主体性）
・Overreliance（過度の依存）
・Model Theft（モデルの盗難）

5. 1. 2　NIST「AI リスクマネジメントフレームワーク（Artificial Intelligence Risk Management Framework（AI RMF 1.0））」

AI RMF（NIST-AI 100-1[14]）は NIST が 2023 年 1 月，AI の信頼性を高めるためのガイダンスとして公表した．一般的な AI リスクの管理手法と，生成 AI の利用組織の成熟度を把握する上で有効なフレームワークである．AI RMF は 2 部構成で，第 1 部は信頼できる AI システムに求められる特性として，AI に関するリスクの考え方や信頼できる AI システムの特徴を説明している．セキュリティとレジリエンスの特性では，同じNISTのサイバーセキュリティ

フレームワークやリスク管理フレームワークにも言及している．第2部では，AIシステムのリスクに対処するための実務の進め方として，組織がAIシステムのリスクに対処するための四つの機能（ガバナンス，マップ，測定，管理）を説明している．NISTはAI RMFの解説と実践方法について記載されたプレイブックも同時公開している．主な想定読者は，AIシステムの設計，開発，展開，評価，利用を行う者であり，AIのライフサイクル全体にわたってリスク管理の取り組みを推進する関連者である．生成AIに特化した内容ではないものの，AIの利用者としてこれからAIシステム全般のリスクマネジメントを進める組織にとって，非常に有益なフレームワークとして活用できる．

5. 1. 3　NIST「敵対的機械学習：攻撃と防御の分類と用語集（NIST AI 100-2e2023)」

NISTは2023年3月8日付けで，ドラフトガイダンス（敵対的機械学習：攻撃と防御の分類と用語集：NIST AI 100-2e2023[15]）を公表した．

OWASP Top 10 for LLMがChatGPTで採用されているLLMを対象にしたサイバーセキュリティ対策のガイドラインであるのに対して，NIST AI 100-2e2023は，現在，画像生成AI分野における敵対的生成ネットワーク（Generative Adversarial Network：以降，GAN[*8]）を始めとした敵対的機械学習（Adversarial Machine Learning：以降，AML）に対するサイバーセキュリティ対策の一環として，NISTがドラフト版を公開し，AMLにおける攻撃手法と防御策を分類し，関連する用語の定義を整理した報告書である．将来的には，AIシステムのセキュリティを評価・管理するための標準化や実践ガイドを提供することを目的としている．現在，パブリックコメントを公募中であり，想定読者はAMLの提供者である開発者やSIerであり，利用者側ではない．本稿執筆時点（2023年9月末）では，ChatGPTを始めとした対話型生成AIは報告対象になっていないものの，AMLのような画像生成AIを提供する側（クラウドベンダー，モデル開発者，モデルを応用したアプリケーションやSaaSベンダー等）が取るべきサイバーセキュリティ対策に関する情報を整理した初のガイドラインである．

NIST AI 100-2e2023は，クラウド上のAMLを利用したPaaSやSaaSを利用する際に，責任共有モデルに則って責任範囲を明確化している．利用者側で講じるべき対策については言及されていないものの，パブリッククラウドベンダーが提供している，AMLを始めとしたクラウド上の画像生成AIサービスの上で，SaaSを提供するCSP（クラウドサービスプロバイダー）にとっては，技術的なリスク項目の認識と対策を評価するための有効な参考情報となるだろう．

5. 2　生成AIのリスク対策

本節ではChatGPTを対象にした対話型生成AIに対する技術的対策例を報告する．いくつかの対策は，BIPROGYの対話型生成AIサービスである「Azure OpenAI Service スターターセット Plus」に実際に組み込まれており，安全・安心なChatGPTシステムとしてBIPROGYの顧客に利用されている．

5. 2. 1　一般的なリスク対策方針

本項ではクラウドの対話型生成 AI サービス（Azure OpenAI Service 等）を利用し，社内の業務効率改善や，社外の顧客向けの商用サービスとして利用する際の一般的な対策方針について説明する．以下 1）～ 4）の四つの観点を軸として方針検討を進める．

1）対策分類の観点：

a）セキュアな対策（情報漏洩対策）

対話型生成 AI はチャットの入力となるプロンプト情報に企業の機密情報や個人情報が含まれることによって，情報漏洩のリスクが潜在化している．技術的な観点の対策としては，以下の情報漏洩対策が中心になる．

・会話履歴をクラウド上に保存しない設定（オプトアウト申請）
・アクセス履歴（ログ）の取得
・特権管理者に対する最小特権の付与
・多要素認証（特権アカウント盗難時の不正ログイン防止）
・外部からのアクセス遮断（ファイアウォール）
・DDoS 攻撃対策
・通信暗号化（TLS 化）

図 2 に，上記の対策を講じることで，安全・安心なセキュリティとガバナンスに配慮した社内向けの ChatGPT 環境のシステム構成例を掲載する．

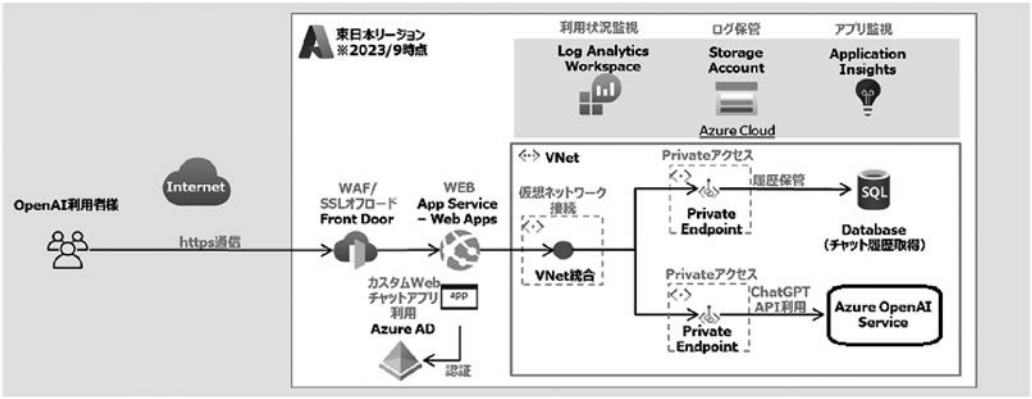

図 2　セキュリティとガバナンンスを考慮した社内 ChatGPT 環境の構成例

b）ハルシネーション（hallucination）対策

対話型生成 AI は「一見するともっともらしい嘘をつく（ハルシネーション）」ことがあり，出力結果に誤情報が含まれている可能性を排除することができない．出力結果をそのまま引用し社外との契約処理や顧客との情報共有として使用する場合は，情報の正確性を再確認する等のチェックが不可欠である．更に踏み込んで，商用サービスのコア機能として利用する場合は，正確性・真正性が担保されている社内情報に基づいた仕組みや，生

成 AI 情報の正確性を担保する事前チェックの仕組み・プロセスをサービスに組み込むことを検討しなければならない.

ｃ）人種差別・ヘイト対策

対話型生成 AI の学習データは欧米諸国を始めとした英語圏のインターネット情報に基づいており，LLM のモデル学習の際にアジア圏や中東・アフリカ圏の情報不足により人種差別やヘイトに繋がる可能性が指摘されている．商用サービスのコア機能として利用する場合は，生成 AI からの出力情報の正確性・真正性を担保しながら，人種差別・ヘイトに繋がるような情報が含まれていないか事前チェックする仕組み・プロセスを商用サービスに組み込むことを検討しなければならない.

ｄ）法的規制対策

生成 AI が関係する法律としては，個人情報保護法，著作権法，独占禁止法，特許法，不正競争防止法等が該当する．個人情報保護法の観点としては LLM や GAN 等の学習モデルに個人の機微な情報，企業の機密情報が学習データとして取り込まれることにより，社外への情報漏洩リスクが考えられる．著作権法の観点としては，生成 AI が生成した画像・音楽等が著作権侵害で訴訟されるリスクがあることを利用者は認識しなければならない．特に社外向けの商用サービスとして利用を検討している場合は，クラウドベンダー側の利用規約（利用約款）を自社の法務部門と共によく確認した上で，こうしたリスクを企業としてどのように担保できるかを検討しなければならない.

ｅ）契約違反対策

Microsoft 社は，以下に該当するコンテンツ生成を，Azure OpenAI Service の利用者の禁止事項として定めている.

・子どもの性的虐待や搾取，いじめ，暴力
・兵器製造，軍事，戦争
・アダルトコンテンツ
・政治的キャンペーン，ロビー活動
・法務実務，金融取引に関するアドバイス提供
・医療従事者でない者による医療行為の提供等

利用者は，これらのいずれかに違反した場合はサービス自体の利用を停止されるリスクがあることを認識しなければならない．社内利用する場合は，利用者である社員に対して利用規則や利用ガイドラインを周知徹底する等のガバナンス対策が不可欠である．併せて，違反した場合の罰則規定を人事規則として定め，社内周知することにより，利用者に対する牽制効果も含める.

2）利用者・利用範囲の観点：

社内での業務改善等の社員向けの限定利用か，商用サービス等での顧客向けの社外利用かという利用者・利用範囲の違いによって，対策レベルの深さが異なってくる．後者の場合は，Google で発生した株価下落の事例のような企業影響を考慮しながら，影響を極小化するための対応策を事前に検討しなければならない．

3）対策レベルの手段の観点：

システムによる自動化か，または統制・ルール等の人による手動プロセスを基本とするのかという手段の違いによって，対策にかかるコストの規模や，対策影響が波及する社内関連部署のスコープに違いが生じてくる．これらの観点は経営層に生成 AI システムの予算化稟議を上申する際の重要な判断基準となることから，適切な経営判断を仰ぐためにも，対応方針案に応じた複数の選択肢を提示することが肝要である．

4）データ保護の観点：

ChatGPT への入力が想定される情報の種類や機密度レベルを整理し，会社の機密情報等の入力を禁止する情報を特定する．禁止された情報を特定した後，禁止則を利用者に対してどのように周知するか，またはシステム的に入力制限やフィルタリングの機能をどのように実装するか，入力制限が難しい場合には入力情報をバックエンドで監視する機能を実装できないかどうか等の検討を進める．

また，生成 AI のクラウドサービスを利用する場合，クラウドの海外リージョンでしかサービスがリリースされておらず，海外リージョンにデータが保存される可能性がある．企業のセキュリティ・ポリシーによっては海外でのデータ保存が許されない場合があり，国内リージョンでのみ生成 AI サービスが利用できるクラウドを選定することが肝要である．なお，Azure OpenAI Service は，本稿執筆時点（2023 年 9 月末）では日本リージョン（東日本のみ）で利用できる．

5. 2. 2　社外利用時のリスク対策

企業の商用サービスとして社外利用する際の対策レベルは数段上がる．まだ事例は少ないが，いくつかの国内外の企業で ChatGPT を活用したエンドユーザ向けの商用サービスを展開しており，こうしたサービス事業者では，技術的な観点の他に，前項「5. 2. 1 一般的なリスク対策方針」の 1）で説明した b）c）d）の観点の対策も同時に講じなければならない．また，利用者であるエンドユーザが意図的な悪意を持っていない場合でも，以下 4 項目の情報が入出力情報に含まれることで，サービス事業者の企業経営を揺るがす事態になりかねない．サービス事業者は，入力出情報の事前チェックや監視，利用ルールに抵触する情報検出時のアラート通知等の仕組みや，プロセスを AI システムに組み込むことを検討しなければならない．

・ヘイト，人種差別，性（LGBTQ）差別等
・機微な情報（病気，遺伝情報，出生情報等）
・特定政治団体，宗教，暴力，戦争，個人攻撃
・嘘の情報（社会的な通念との乖離の有無）

これらを防ぐ対策としては，システムによる入力・出力情報の事前チェック機能（フィルタリング）を自社サービスの仕組みに組み込む方式が考えられる．出力情報の事前チェック機能は，米国 Robust Intelligence（ロバストインテリジェンス）社のサービス[16]として既に存在している．当該サービスを社外向けの ChatGPT システムに組み込んだ場合のシステム構成を図 3 に例示する．社外のエンドユーザからの入力情報については，内容のチェック機能を独自実装するか，対話型生成 AI サービスの提供元で監視する手段を取ることができる．対話型生成 AI サービスからの出力情報については，内部連携機能として Robust Intelligence 社のサービスを呼び出すことによって内容をチェックすることができる．これにより，エンドユーザに出力情報を含めた回答結果を返す前に，回答結果の正確性等の「機能・品質」の確認，人種差別等の偏見が含まれていないかどうかの「公平性」の確認や，プロンプト・インジェクション*9等の「セキュリティ」リスクの事前確認ができるようになる．今後，同様なサービスを展開する SaaS ベンダーの登場と共に，クラウドベンダー自らがそのような安全・安心な機能を実装した生成 AI サービスを開発しリリースしてくる可能性も高い．

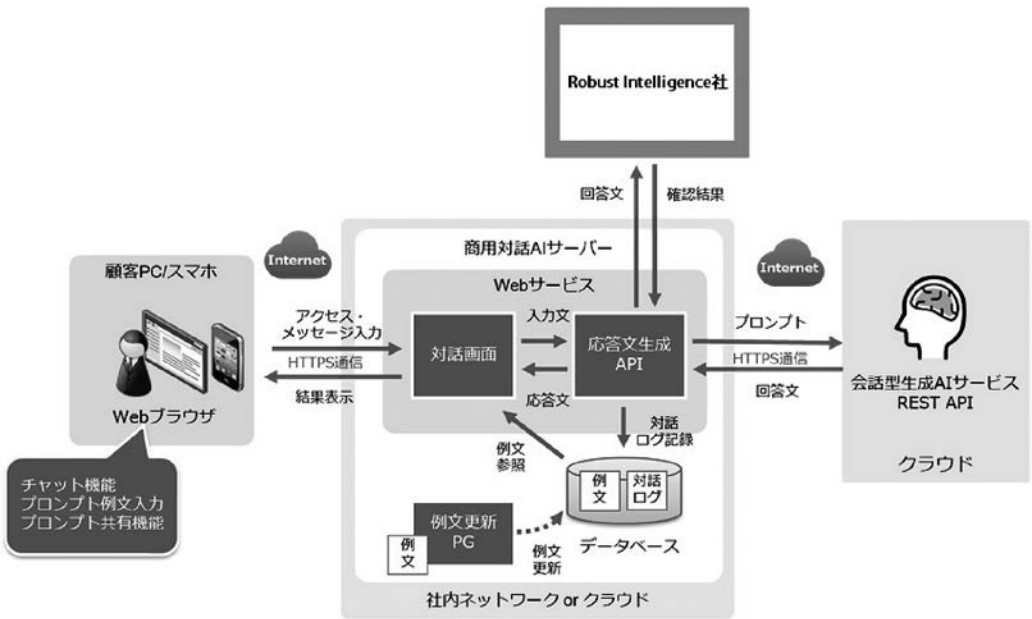

図 3　米国 Robust Intelligence 社のサービスを活用した社外 ChatGPT 環境

6. お わ り に

今後，DX ビジネスの拡大を目指す企業にとって，クラウドガバナンスは，IT ガバナンス，ひいてはコーポレートガバナンスの観点において，非常に重要な要素となることは自明であり，その整備が急務となってくるであろう．そのためには，今回紹介したガバナンスフレームワークを積極的に活用しながら，クラウドガバナンスを整備し，浸透させていくことが肝要である．クラウドガバナンスは，DX ビジネスの成長性とのバランスが重要である．IT 利用に対する統制をきつく締めすぎると，DX ビジネスの成長スピードを阻害する要因になりかねないが，ビジネススピードを優先するあまり統制を緩めすぎると，不正アクセスの温床や IT 予算の垂れ流しに繋がりかねない．

現在，国内企業では，欧米企業と比較するとリスクマネジメント等のガバナンス面への対応の遅れが指摘されている状況である．今後，OpenAI や敵対的生成ネットワーク（GAN）等の生成 AI テクノロジーを応用したクラウド上の商用サービスが普及していく可能性が高まる中で，企業経営への影響を考慮したガバナンスの徹底とリスクマネジメントへの対応がより一層求められる局面になってきている．本稿で紹介したフレームワークを有効活用し，生成 AI を始めとしたクラウド利用に伴う企業リスクの極小化に努めることが肝要である．

なお，BIPROGY では，本稿で紹介したクラウドガバナンスの考え方に準拠したコンサルティングサービスの提供を開始しているとともに，Azure OpenAI Service を利用した社内利用のための安全かつ安心なサービスの提供も開始しており，各業界の顧客から多数のお問い合わせやご相談を受けている．これらの利用を本稿の読者の方々にもご検討いただければ幸いである．

最後に本稿の執筆において協力を頂いた関係各位に感謝の意を表する．

＊ 1　欧州連合（EU）の公的ネットワーク・情報セキュリティ機関．加盟国のセキュリティレベルの向上を目的として 2004 年に設立．
＊ 2　米国商務省配下のアメリカ国立標準技術研究所．情報セキュリティ技術の標準化を行っている政府機関．
＊ 3　セキュアな Web アプリケーション開発を促進する技術・プロセスに関する情報共有と普及啓発を目的としたオープンソース・ソフトウェアコミュニティ．NPO 法人である The OWASP Foundation がその活動を維持している．
＊ 4　ESG（環境（Environment），社会（Social），ガバナンス（Governance））は企業または企業への投資の持続可能性と社会的影響を測定する三つの中心的な要素を指し，これらの基準は，企業の将来の財務状況をより適切に判断するのに役立つ．
＊ 5　COBIT（control objectives for information and related technology）とは，企業・自治体組織の IT ガバナンスの指針として，情報システムコントロール協会（ISACA）や IT ガバナンス協会（ITGI）が提唱する IT ガバナンスの実践規範である．
＊ 6　BYOL（Bring Your Own License）とは，ユーザがオンプレミス用途等で購入していた所有済のソフトウェアライセンスを，クラウドサービス上に持ち込んで利用する方式．
＊ 7　自然言語処理（NLP）で使用される深層学習モデルの一種．大規模なテキストデータを学習し，人間のような自然な言語生成や理解を実装することを目的とする．
＊ 8　2014 年にモントリオール大学（当時）の Ian Goodfellow らによって発表されたニューラルネットワークの教師なし学習手法．現在，画像生成 AI のモデルとして応用されている．
＊ 9　ユーザが悪意を持ったプロンプトを使い LLM の開発者が意図しない情報開示を狙った対話型生成 AI の脆弱性を突くサイバー攻撃．

参考文献

［1］ 令和 5 年度版情報通信白書　第 8 節　データセンター市場及びクラウドサービス市場の動向，総務省，2023 年 7 月
https://www.soumu.go.jp/johotsusintokei/whitepaper/ja/r05/html/nd248200.html
［2］ 金融機関の IT ガバナンスに関する対話のための論点・プラクティス整理　第 2 版，金融庁，2023 年 6 月
https://www.fsa.go.jp/news/r4/sonota/20230630/02.pdf
［3］ Azure 向けの Microsoft Cloud 導入フレームワーク，Microsoft
https://docs.microsoft.com/ja-jp/azure/cloud-adoption-framework/
［4］ An Overview of the AWS Cloud Adoption Framework，AWS
https://docs.aws.amazon.com/whitepapers/latest/overview-aws-cloud-adoption-framework/introduction.html
［5］ An Overview of the AWS Cloud Adoption Framework Version 2，AWS
https://d1.awsstatic.com/whitepapers/aws_cloud_adoption_framework.pdf
［6］ クラウドのガバナンス手法，Microsoft
https://learn.microsoft.com/ja-jp/azure/cloud-adoption-framework/govern/methodology
［7］ AWS クラウド導入フレームワーク，AWS
https://d1.awsstatic.com/whitepapers/ja_JP/aws-cloud-adoption-frameworkja-JP.pdf

［8］ “Cloud Computing: Benefits, risks and recommendations for information security”（クラウドコンピューティング：情報セキュリティに関わる利点，リスクおよび推奨事項），ENISA，2009.11，
https://www.enisa.europa.eu/publications/cloud-computing-risk-assessment/@@download/fullReport
［9］ ENISA クラウドコンピューティング：情報セキュリティに関わる利点，リスクおよび推奨事項の対訳，IPA，2010.5.25，
https://warp.ndl.go.jp/info:ndljp/pid/11630427/www.ipa.go.jp/security/publications/enisa/documents/enisa%20jp-en%20doc.pdf
［10］ クラドサービスにおけるリスクと管理策に関する有識者による検討結果 2011 年度版，JASA，2011 年
https://jcispa.jasa.jp/wp-content/uploads/docs/pdf2012/2012_cloud_doc04.pdf
［11］ JDLA，生成 AI の利用ガイドライン，ISACA，2023.5，
https://www.jdla.org/document/#ai-guideline
［12］ OWASP Top 10 for LLM, OWASP, 2023.8，
https://llmtop10.com/
［13］ OWASP Top 10 for Large Language Model Applications, OWASP, 2023.8，
https://owasp.org/www-project-top-10-for-large-language-model-applications/
［14］ Artificial Intelligence Risk Management Framework (AI RMF 1.0), NIST, 2023.1，
https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf
［15］ Adversarial Machine Learning A Taxonomy and Terminology of Attacks and Mitigations (NIST AI 100-2e2023 ipd), NIST, 2023.3，
https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-2e2023.ipd.pdf
［16］ AI Integrity 実現のための End-to-End のテスト・モニタリング，Robust Intelligence，
https://www.robustintelligence.com/jp

※　上記参考文献に含まれる URL のリンク先は，2023 年 11 月 24 日時点での存在を確認.

執筆者紹介　伊　藤　直　行（Naoyuki Ito）

1991 年日本ユニシス(株)入社．電力事業者向け大規模運用管理システムの開発，ストレージ・バックアップ，内部統制用セキュリティソリューションの適用コンサルティング，自社製クラウドサービス開発プロジェクトを経て，クラウド適用のシニアコンサルタントとして，移行計画策定のコンサルティング，アドバイザリを担当．現在，サポートサービス本部クラウドサービス部に所属．AWS ソリューションアーキテクト，Microsoft Azure 管理者アソシエイツ，公認情報システム監査人（CISA），ISACA 東京支部会員.