

生成AI技術を用いた製品のセキュリティ・リスク評価・ガバナンス体制

2025年11月4日

株式会社 Preferred Networks
福田 昌昭 / Masaaki Fukuda

Preferred Networks (PFN) 会社概要

ミッション： 現実世界を計算可能にする

<https://www.preferred.jp>

設立 2014年3月26日

本社 東京都千代田区

代表取締役 西川徹（最高経営責任者）
岡野原大輔（最高技術責任者）

従業員数 約350名（2025年2月）

事業内容 AIチップ、計算基盤、生成AI基盤モデルなどのAI関連技術を活用したソリューション・製品の開発・販売および研究開発

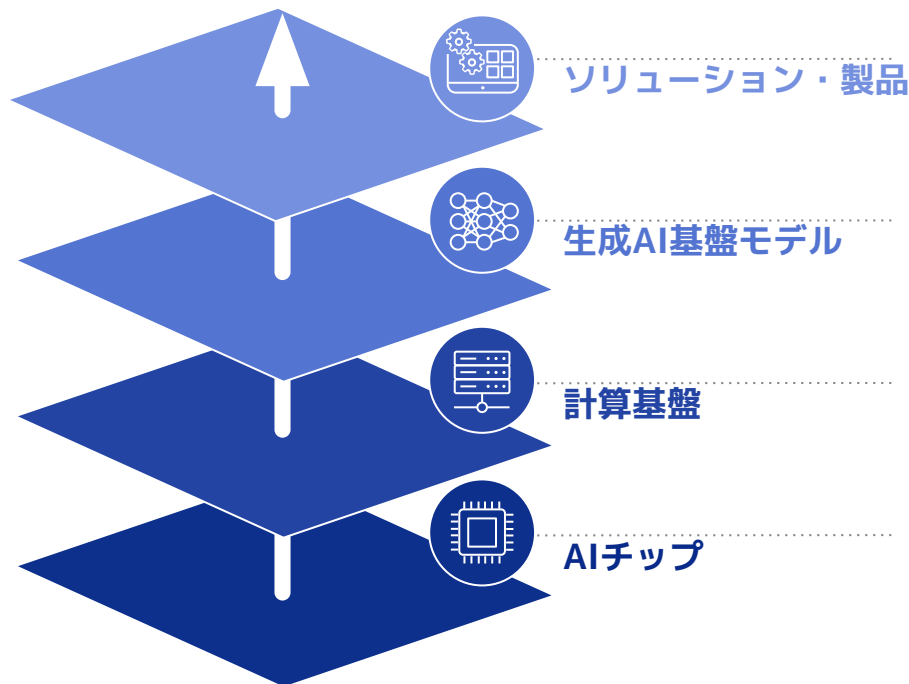
主要子会社 Matlantis株式会社（2021年6月設立、2025年7月Preferred Computational Chemistryから社名変更）
株式会社Preferred Robotics（2021年11月設立）
株式会社Preferred Computing Infrastructure（2025年1月設立）

出資企業
(五十音順) SBIグループ NTT株式会社 ENEOSイノベーションパートナーズ合同会社 株式会社講談社
信越化学工業株式会社 SUMISEI INNOVATION FUND 積水ハウス投資事業有限責任組合 中外製薬株式会社
TBSイノベーション・パートナーズ3号投資事業組合 TEL Venture Capital, Inc. 東映アニメーション株式会社
トヨタ自動車株式会社 株式会社日本政策投資銀行 株式会社博報堂DYホールディングス 株式会社日立製作所
ファナック株式会社 株式会社みずほ銀行 三井住友信託銀行株式会社 三井物産株式会社 三菱商事株式会社
三菱UFJ信託銀行株式会社 株式会社ワコム 他



PFNの事業: AI技術のバリューチェーンを垂直統合

PFNは、チップ、計算基盤、生成AI基盤モデル、ソリューション・製品まで、AI技術のバリューチェーンを垂直統合し、ソフトウェアとハードウェアを高度に融合することで、競争力の高い技術の開発および産業応用を進めています。



様々な産業・消費者向けのソリューション・製品



PLaMo

PLaMo Prime (国産LLM)
PLaMo Lite (エッジ向けSLM)

PFP

物質のエネルギー計算モデル



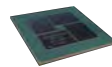
GPUクラスタ



MN-3
(MN-Core™
クラスタ)

PFCP

MN-Core™ 2を
計算資源とした
クラウドサービス



MN-Core™



MN-Core™ 2



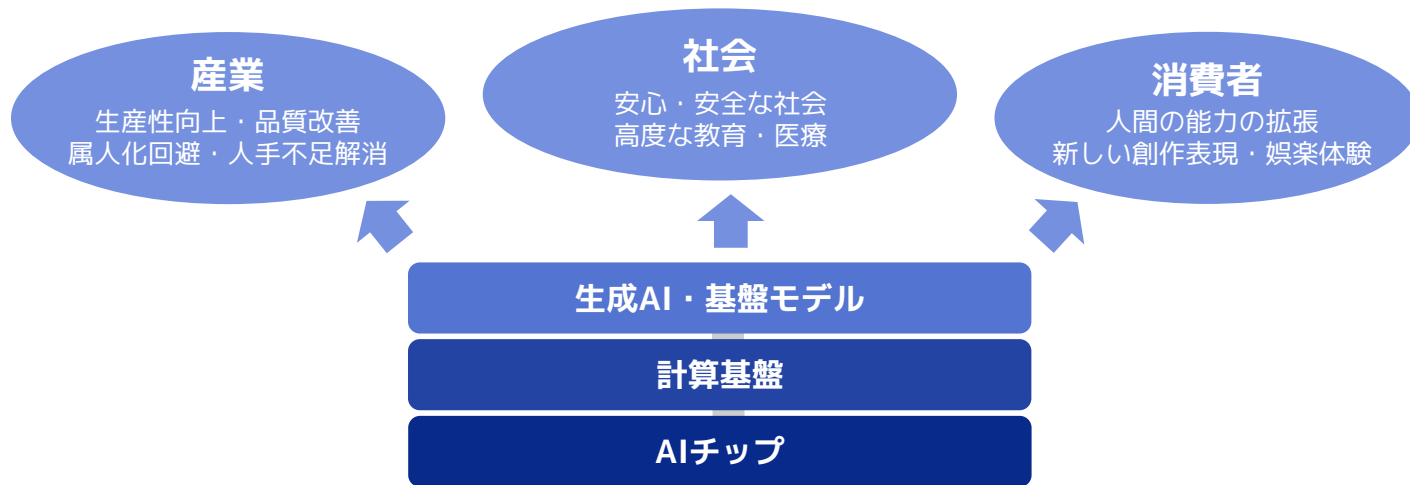
生成AI (推論) 向け
MN-Core L1000
(2026年提供予定)



MN-Core
次世代

PFNの事業: AI技術の水平展開

PFNは、AI技術のバリューチェーンを垂直統合し、産業、コンシューマー、社会に向けて様々な領域でソリューション・製品を水平展開しています。



- 生成AIの活用が、業務効率化に大きく貢献することが明らかになっている

年間 **400** 時間/ 人の
工数削減

日清食品HD*¹

年間 **18.6** 万時間の
工数削減

パナソニックコネクト*²

2026年までに
80%以上の企業が、
生成AIに対応した
アプリケーションを
本番環境に展開するように
ガーートナー社調べ*³

*1

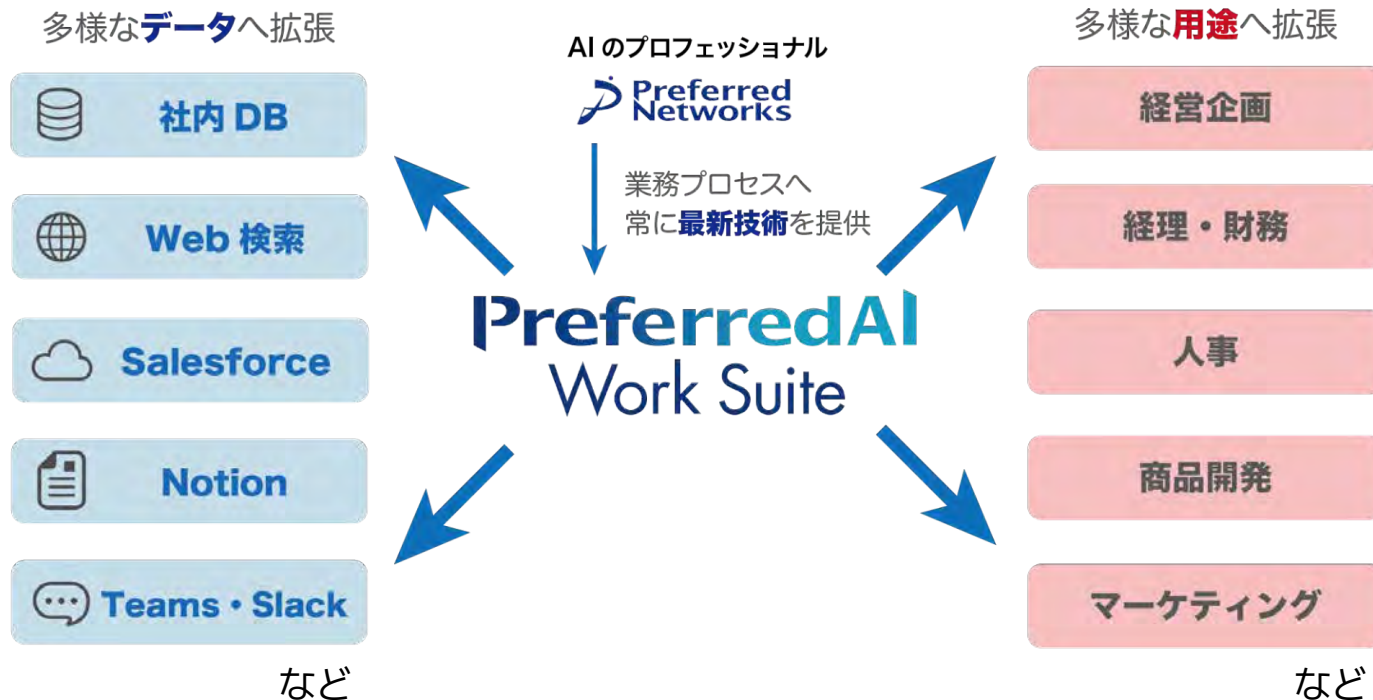
<https://xtech.nikkei.com/atcl/nxt/column/18/02689/010500001/>

*2 <https://www.businessinsider.jp/post-268299>

PreferredAI Work Suite

Work Suite 機能紹介

PreferredAI Work Suiteは業務効率化のためのAIプラットフォーム。
常に最新のAIを業務プロセスへ活かすことが可能です。



PreferredAI Work Suiteは、
ストレージ・ノート・チャット・ワークフローで業務を支えます。



主な機能

ストレージ

ファイルのアップロード

ノート・チャット

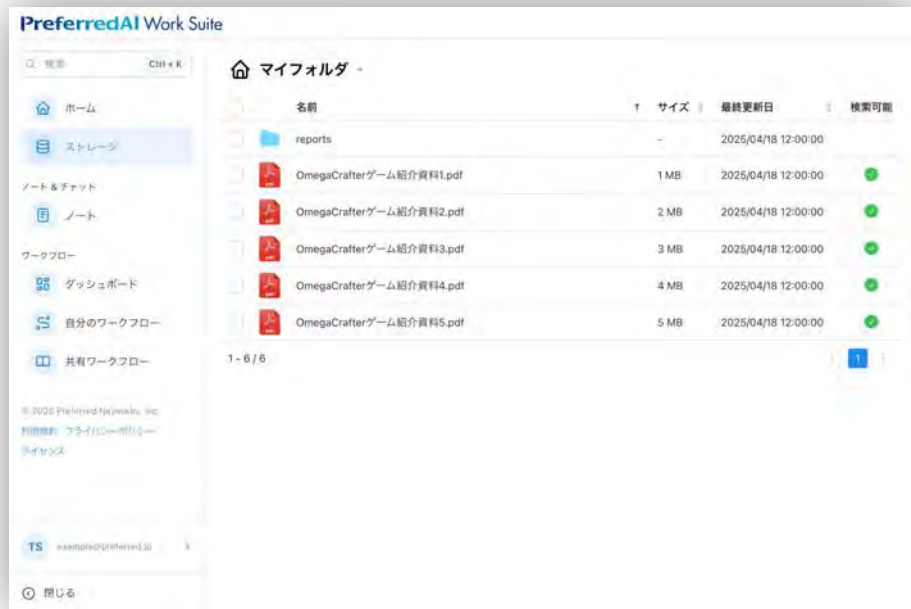
データ検索・文章作成

ワークフロー

業務の自動化

ストレージ

ファイルをアップロードするだけの簡単活用。外部データの連携も可能。



主な機能

ストレージ

ファイルのアップロード

ノート・チャット

データ検索・文章作成

ワークフロー

業務の自動化

ノート・チャット

チャットで質問し、RAGで回答。AIで文章推敲や翻訳も。



主な機能

ストレージ

ファイルのアップロード

ノート・チャット

データ検索・文章作成

ワークフロー

業務の自動化

例えば以下のような用途ですぐに活用することが可能です。

- ✔ アップロードされた大量の社内データを参照し、
AIで叩き台を作りながら既存プロダクトの新しい使い方を考える。
- ✔ 有価証券報告書など、表データを持つ大量のPDFファイルを検索し、
企業についてのまとめを作成する。
- ✔ ミーティングの録音から文字起こしを自動作成し、
AIで要約しながらまとめを作成する。
- ✔ 長大なPDFファイルをアップロードし、
内容について質問しながら理解を深める。

ワークフロー

繰り返される業務をAIにより自動化。定期的な実行、サービス連携も可能。



主な機能

ストレージ

ファイルのアップロード

ノート・チャット

データ検索・文章作成

ワークフロー

業務の自動化

ワークフローの主な機能（カスタマイズ不要で利用可能なもの）

基本機能に加え、複雑な処理を必要とする機能については別途カスタマイズにて追加することも可能です。

LLM

LLMプロンプト

モデルを選択可能

データ抽出

ファイル・文章からデータを自動抽出

ファイル出力

LLMでファイルを作成

外部連携

外部サービス連携

連携サービス例

- Salesforce
- SharePoint

※カスタマイズ対応可

HTTPリクエスト

URL・APIを参照させる



データ参照

ファイル・Web検索

アップロード済ファイルや Webを検索

高度な機能

条件分岐

処理を分岐させる

日々の業務を支える お役立ち機能

定期実行

特定の時刻に毎日実行

まとめて実行

ワークフローを
複数の条件で一括実行

共有・コピー

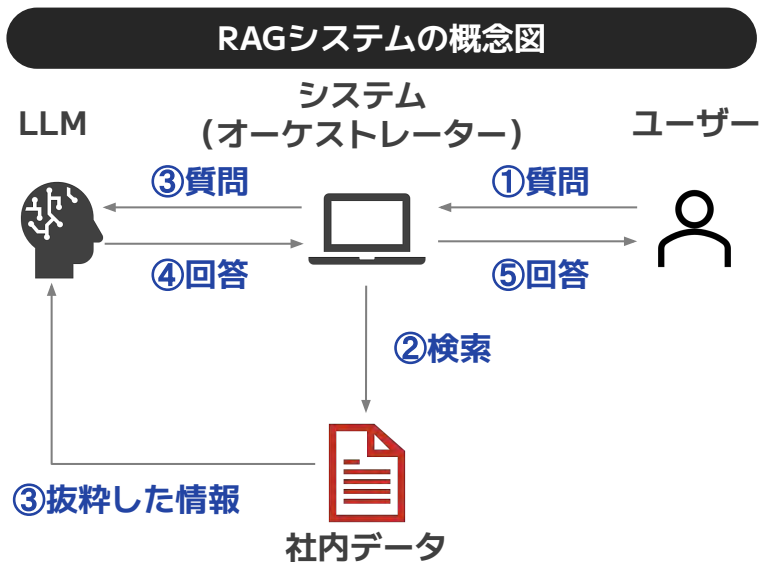
他のメンバーに
ワークフローを共有

API実行

外部からの呼び出し

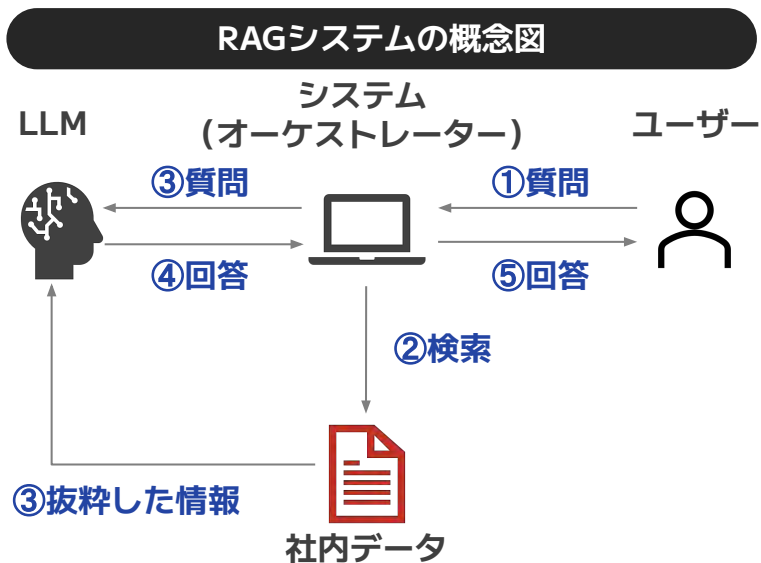
RAGを用いたシステムのセキュリティ課題

システムの性能を客観的に測るための指標を定義。定義した評価指標のスコアが最大になるように、各種パラメータを調整する。



最大化を目指す指標の例	概念図との対応
Q&Aベンチマーク(回答の正解率)	⑤
回答に対するユーザー評価	⑤
検索ベンチマーク(検索のヒット率)	②

そのために調整するパラメータの例	
プロンプトの文面パターン	①③
検索方法(BM25, Embedding, ハイブリッドなど)	②
LLMに渡す検索上位件数	③



- セキュリティの課題と対応例

- 社内データの情報漏洩リスク

- 外部LLMを利用する場合、プロンプトや検索結果に社外秘情報が含まれるリスク

- ルールベース、LLMベースのフィルタを挿入
- LLM提供元とのデータ利用ポリシーを契約面で明確化

- アクセス制御・権限管理の不備

- 検索結果として、見てはいけない文書をAI経由で情報を取得してしまうリスク

- ユーザー、ロールベースでのアクセスコントロールの実装、検索範囲の制限
- 監査ログの保存、トレーサビリティを確保

- LLMプロンプトインジェクション攻撃

- 悪意のあるユーザーが意図しない動作や内部情報を出力するように誘導する

- プロンプトサニタイザーの導入
- ユーザー入力と社内データを明確に分離して管理する
- ルールベース、LLMベースのフィルタを挿入
- 自由入力の制限、安全なテンプレート入力など

PreferredAI Work Suiteは、安心してAIの活用を目指し、セキュリティ確保に努めています。以下の方針に基づき、安全で信頼性の高いAI運用を推進しています。

1. ガイドラインの遵守

本プロダクトは、国や業界団体が定めるAI事業者ガイドラインに基づき、情報管理とセキュリティ対策を適切に実施し、安全性の確保に取り組んでいます。

2. 責任の分担・共有

AIの安全な活用には、AI利用者だけではなく、AI開発者・AI提供者にも責任が生じると考えています。本プロダクトは、利用者とともに、安心してAIを活用できる環境づくりを進めていきます。

3. 柔軟な対応

AIの活用目的や求められる成果はお客様によってさまざまです。

本プロダクトは、利用目的や要件に応じて柔軟に対応し、最適なAI活用を支援します。

4. データの安全管理

本プロダクトは、インターフェースおよびデータを適切に管理する仕組みを整えています。

入力内容や出力結果の取り扱いを明確にし、利用者がアクセス制御やデータ保護を自ら管理できる仕組みを提供することで、安心してAIを活用できる環境の実現を目指しています。

当社のAIガバナンス体制

AIガバナンス体制概要

PFNは、政府のガイドラインに沿ったAIガバナンスの体制を構築し、経営陣のコミットメントのもと、ガバナンスとイノベーションを両立する開発手法を追求しています。また、政府のAI戦略や国際的枠組みへも積極的に貢献しています。

AIガバナンス推進体制

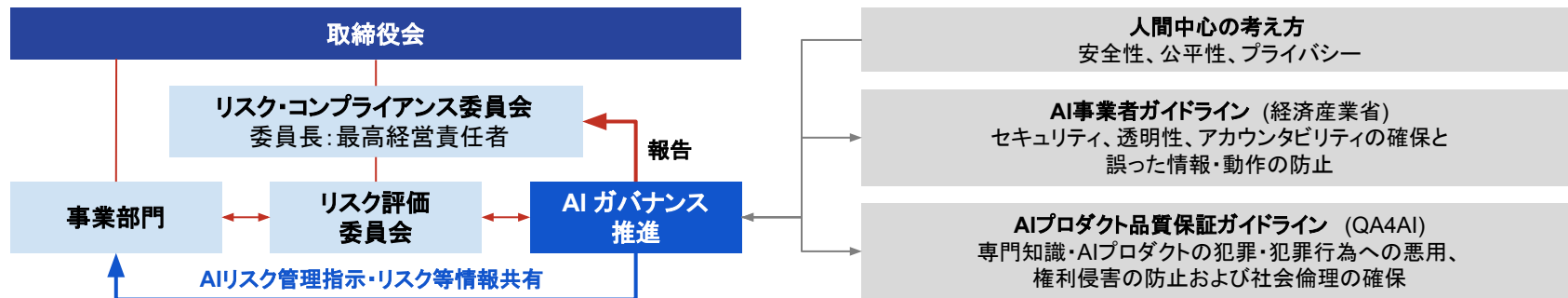
- 経産省のAI事業者ガイドライン・業界団体のAIプロダクト品質保証ガイドラインに沿ったガバナンス体制
- AIガバナンスとイノベーション促進を両立する開発手法を提示

経営陣のコミットメント

- 経営陣が経営ガバナンスの一環としてAIガバナンスを実施
- 最高経営責任者が委員長を務めるリスクコンプライアンス委員会がAIリスクマネジメントを統括

政府・国際的枠組への貢献

- 総務省・経産省AI事業者ガイドライン、内閣府AI時代の知的財産権検討会に委員として参加
- G7が主導する広島AIプロセス*に協力



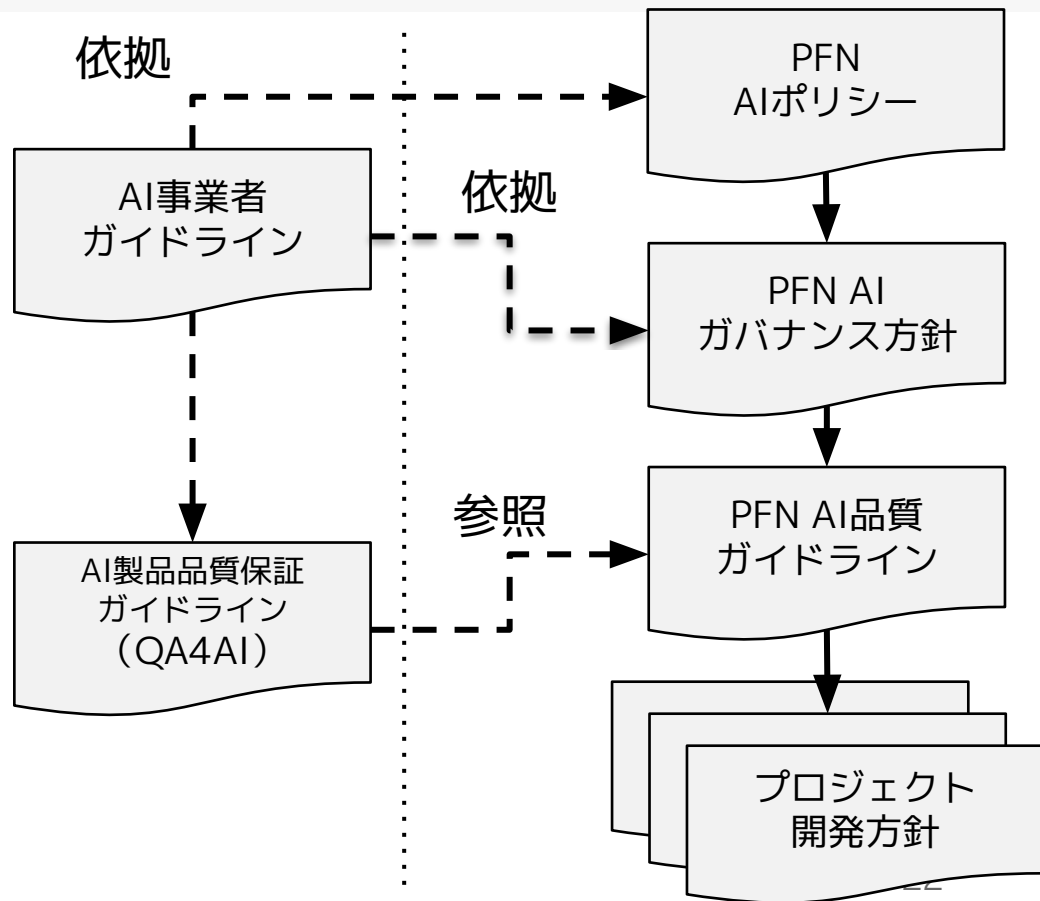
*広島AIプロセス: 2023年にG7首脳に承認された、安心して信頼できる高度な AIシステムの普及を目的とした指針と行動規範からなる国際的政策枠組み。
PFNのG7 Hiroshima AI Process (HAIP) Transparency Reportはこちら: <https://transparency.oecd.ai/reports/a86f4925-5cd5-4af7-b4f6-1b1f0984419e>

リリース時にリスク評価(ERM)を実施

中分類	小分類	項番	リスク項目
1. 技術領域	a 技術領域(規制、コンプライアンス、その他)	1	医療関係・BIO関係
		2	金融関係
		3	化学
		4	交通
		5	規制のある領域(輸出入規制を除く)
		6	その他
	b 人命等に関わる事故の可能性	7	アクチュエーター なし/小規模/大規模
		8	事故の可能性 なし/小規模/大規模
	c 環境汚染	9	通常時に懸念がある
		10	事故時に懸念がある
	d 倫理面と社会的な影響(人権・公平性)	11	社会的バイアスのかかった出力を行う可能性がある(犯罪予測、内定判断、保険、Tay、動物判定 google など)
		12	
	a データ・情報・	1	個人情報扱う(氏名、住所等、画像、医療情報、行動履歴)
		2	広義の個人情報扱う(Cookie など: GDPR)
		3	機微情報を扱う
		4	インターネットサービスを提供する

→ 大規模言語モデルの開発に伴う、AIガバナンス体制強化の必要性を認識

規程・ガイドライン体系



PFNのAIに関する基本方針

AIに関する適切な企業活動
のための体制・規程

プロジェクトが守るべき
ルールや手続き

ガイドラインの対応可否・
選択理由などを記述

AIポリシー: PFNのAIに関する基本方針

構成

- PFNの企業理念・事業目的
 - METIのガイドラインの「基本理念」の文脈で記載
 1. 人間の尊厳が尊重される社会(共通指針 1)
 2. 多様な背景を持つ人々が多様な幸せを追及できる社会
 3. 持続可能な社会
 - 公正な競争(共通指針 9)、イノベーション(共通指針 10)に関しても記載
- AIなどへの懸念に対する基本姿勢 (Responsibility)
 - 透明性(共通指針 6)、アカウントビリティ(共通指針 7)
- ガバナンス体制
 - 安全性(共通指針 2)、公平性(共通指針 3)、プライバシー(共通指針 4)、セキュリティ(共通指針 5)
- 教育、社会への還元
 - 教育・リテラシー(共通項目 8)

PFN AIポリシー

AI Policy

PFNは現実世界を計算可能にすることで、これまで解決が困難であった課題解決を目指します。人々のより豊かな生活の実現のために、生産性の向上だけでなく、人間の能力や創造性を拡大するとともに、一人ひとりの関与の促進をサポートし、これまでの人々の賢知を受け継いだ上でこれを拡張し、尊厳を持って豊かな生活の実現に寄与することを目的とします。

AI技術自体を支える技術領域の革新に寄り添い、AIを活用するだけでなくこれらのAIを支えるための高度な計算力の構築により持続可能な社会の実現を目指します。そして、これらの取り組みは、企業や個人の選択性を確保し、公正な競争を実現することにつながるものと考えています。

一方で、新技術は、機密が漏洩されれば様々な場面で使われるようになります。そのような技術の適正な開発と活用のため、私たちは以下の方針を定めます。

1. 透明性 (正しく伝える)

私たちは、その技術自体が何をするか、何ができるか、何ができないかを可能な限り明確にします。それがどのような仕組みで動くかを、論文その他の出版、ブログ、ドキュメント等を通して、様々なレベルのステークホルダーが理解できるよう、丁寧に説明します。

2. リスク (適切に怖がる)

私たちは、その技術を使うことによって、社会の様々な場面で起さるリスクを、私たちの想像力が及ぶ限り考えます。リスクには、定量化できないもの、想定しにくいものがあり、また、ある時点での社会適応ではリスクではなかったものが、その後リスクと考えられるようになることもあります。新技術は真に便益とリスクを持っています。私たちは、社会との対話を通じて、そのようなリスクに前向きにいく努力を積み重ねます。

3. インテグリティ (見たくないものを、見る)

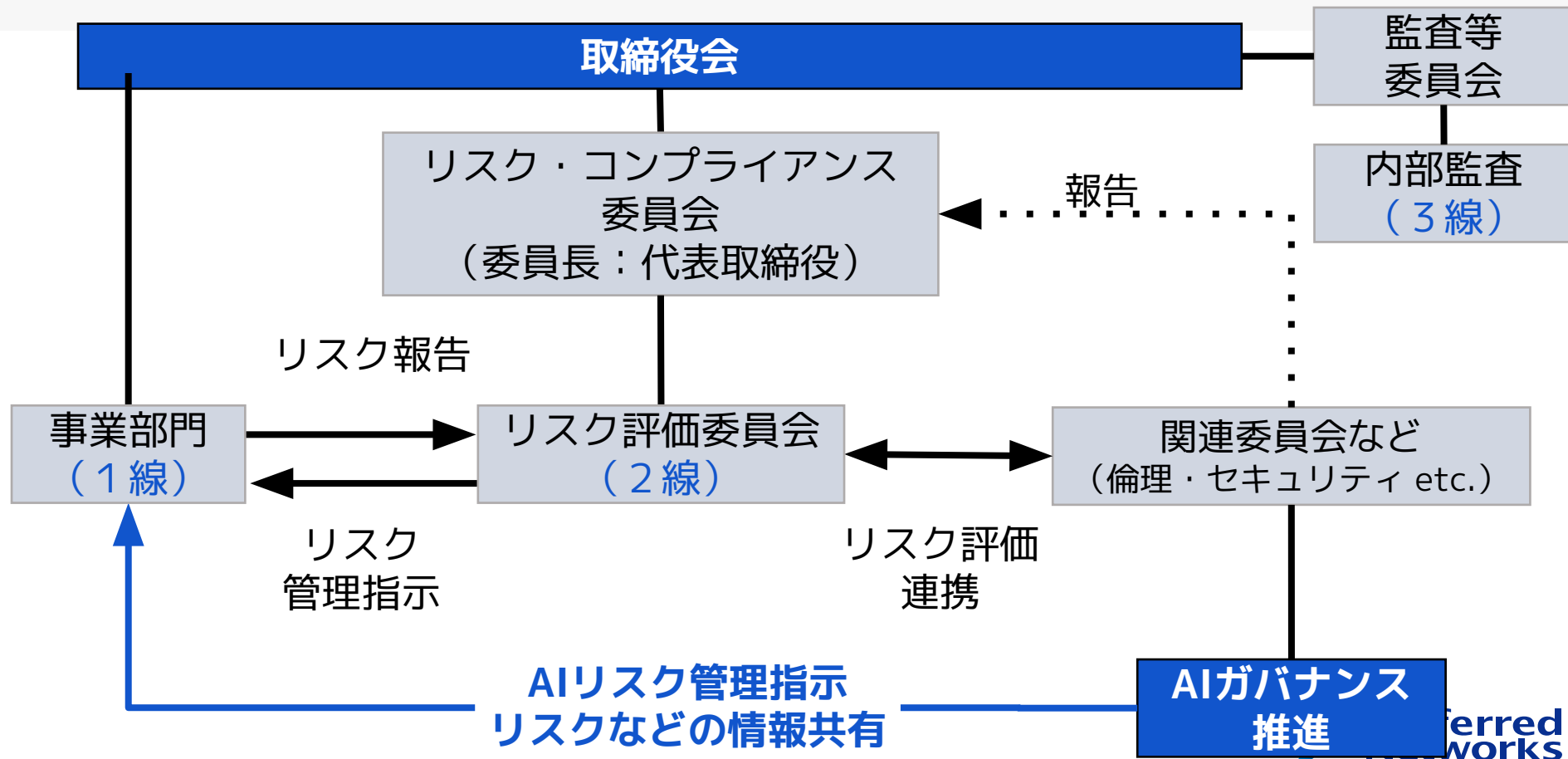
新技術の開発や利用にあたっては、時として私たちに都合の悪い発見があるかもしれません。私たちは、そのようなものから目を背けることなく、誠実に対応します。

PFNは、責任ある開発者であるため、個人の生命、自由、尊厳、財産などを脅かすことのない開発であること、確信に依りながら可能性を認識し、公平性を保つよう努め、多様性のある社会の実現を目指します。

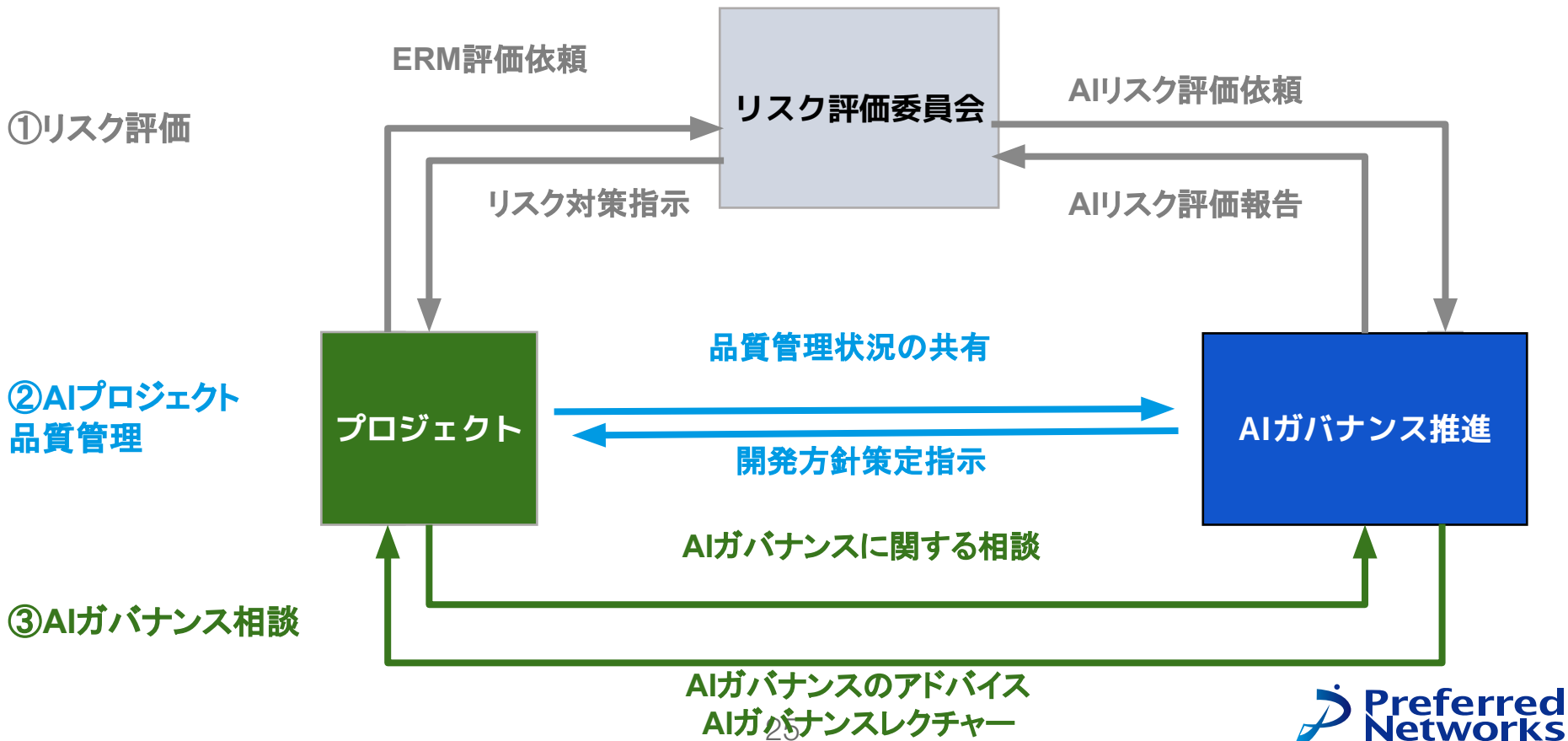
データソースや利用者のプライバシーに配慮するとともに、サイバーセキュリティを確保し、AI開発特有のセキュリティに取り組んでいます。

ここに述べた取り組みを進めるにあたって社内でも知識と技術を共有し、技術、倫理、プライバシー、セキュリティ、コンプライアンスなどに対する教育と啓発に取り組んでいます。会社バリューの一つであるLearn or Dieの精神の下に技術の可能性やリスクについて学び続け、社会に貢献していきます。

規程・ガイドライン体系



AIガバナンス推進



開発方針(抜粋)

【モデルケース】プロダクトA開発方針

XXXX年X月XX日

文責:XXXXX(<メールアドレス>)

更新情報

- 202X年X月XX日:初版作成
- 202X年X月XX日:更新作成
- 202X年X月XX日:更新作成

本ドキュメントについて

PFNのプロジェクトの管理者は、プロジェクトでのAI品質管理方法を定めた開発方針を策定しなければならない。本ドキュメントは、各プロジェクトが開発方針を定めるために利用可能な開発方針テンプレートである。「PFN AI品質ガイドライン」には各プロジェクトがAIシステムの開発・提供を行うにあたって、遵守すべき一般的な内容についてまとめている。必ずしも、上記ガイドラインのすべての項目を満たす必要はないが、遵守事項を採用しない場合は、開発方針にその理由とリスクの担保方法を明確にすることが求められる。

開発方針(抜粋)

1.2. プロジェクト体制・データ管理ポリシー

記載内容

- プロジェクト体制(責任者、開発メンバー、運用体制など)
- データ管理ポリシー

プロジェクト体制が現段階で確定していない場合は、その旨を記載してください。また、データ管理ポリシーを作成していない場合は、その旨記載の上別途セキュリティチームにご相談ください

補足:このプロジェクトに関することを誰に聞けばよいか、どのデータに誰がアクセスすることができるのかを把握できるようにしてください

プロジェクト体制

- 責任者:担当VP xxxxx@
- メンバー: xxxxx@ xxxxx@ xxxxx@ xxxxx@
- 運用体制:データ管理ポリシーに記載

データ管理ポリシー:(資料削除)



Making the real world computable