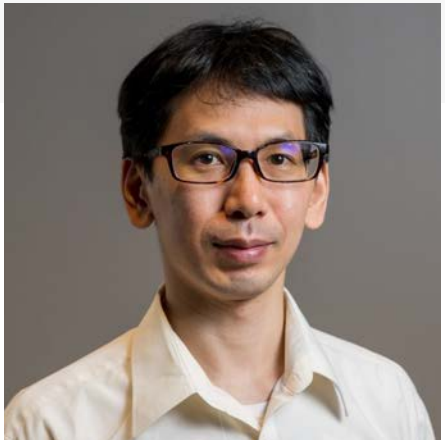
LLMの安全性シンポジウム 2025
@国立情報学研究所 大規模言語モデル研究開発センター

Preferred NetworksのAIガバナンスへの取り組み

2025年7月23日

株式会社Preferred Networks
大野健太 <oono@preferred.jp>

大野健太（おおのけんた）



株式会社Preferred Networks
　AIプロダクツ&ソリューションズ事業部
　エンジニアリングマネージャー兼
　AIガバナンス推進責任者

博士（情報理工学）
日本メディカルAI学会　評議員
デジタル庁・データセキュリティワーキンググループ　委員

Preferred Networks

Preferred Networks (PFN) 会社概要

ミッション： 現実世界を計算可能にする

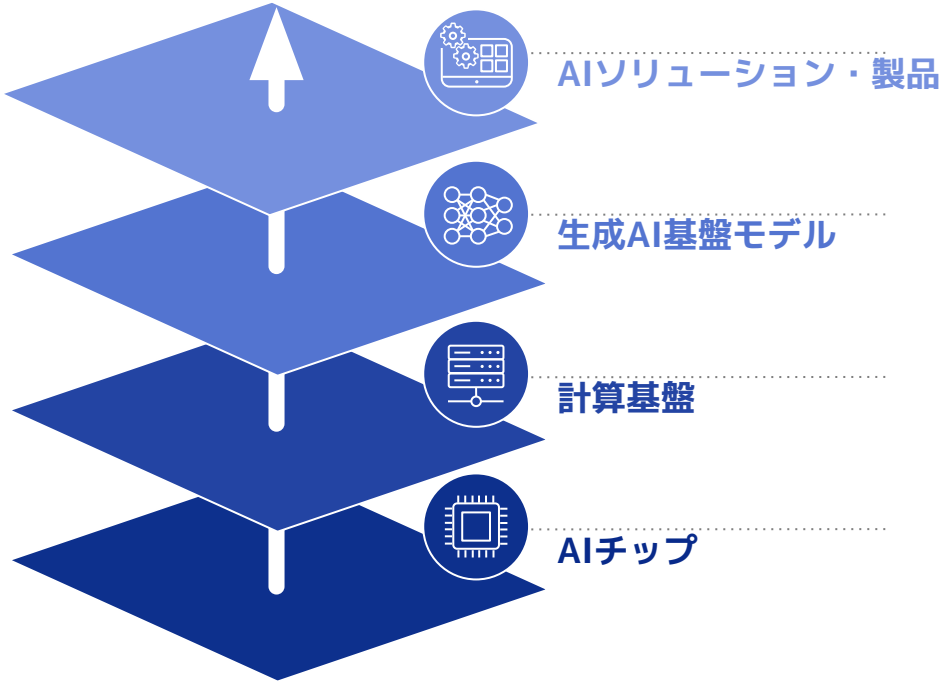
https://www.preferred.jp

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　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

Preferred Networks

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。



様々な産業向けのAIソリューション・製品

PreferredAI　MiseMise　MATLANTIS

PFN 3D Scan　Preferred Networks Visual Inspection　kachaka

PLaMo

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

PFP

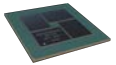
物質のエネルギー計算モデル

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

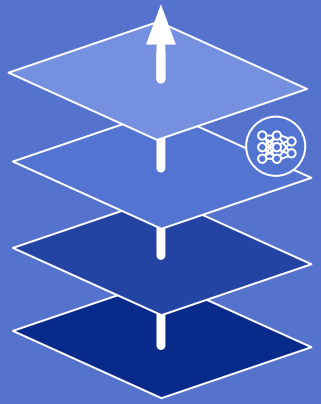
MN-Core 2

生成AI（推論）向け
MN-Core L1000
（2027年提供予定）

MN-Core
次世代

Preferred Networks

生成AI・
基盤モデル



PLaMo™

PFNグループの生成AI基盤モデル

PLaMoは、PFNのグループ会社Preferred Elements (PFE)がフルスクラッチで開発する純国産の生成AI基盤モデルです。高品質な学習データを用い、既存の大規模言語モデルをベースに用いずに全く新規に開発。主要な日本語ベンチマークで世界最高レベルの精度を示しています。

- **商用版PLaMo Primeを**2024年12月にリリース
- **エッジ向け小規模言語モデル PLaMo Lite を提供中**
- 金融・医療などの特定タスクに強いモデルも順次リリース予定

世界最高クラスの日本語性能

- 主要な日本語ベンチマーク*でGPT-４を超える精度を記録
- 現在も精度向上中
- 日英の翻訳性能にも優れる

*日本語LLMの性能を測るため標準的に使われるベンチマーク「Jaster」を使用

純国産フルスクラッチモデル

- 独自のアーキテクチャと学習データでゼロから事前学習･指示学習を実施
- 他社モデルベースではないため社外ライセンスの縛りや開発上で不明瞭な点がない

APIで簡単に導入可能

- API経由でクラウドを介して提供、簡単なコードの書き換えで実装可能
- モデル自体をお客様のクローズドなオンプレミス環境に実装することも可能

詳細: https://plamo.preferredai.jp/

Preferred Networks

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

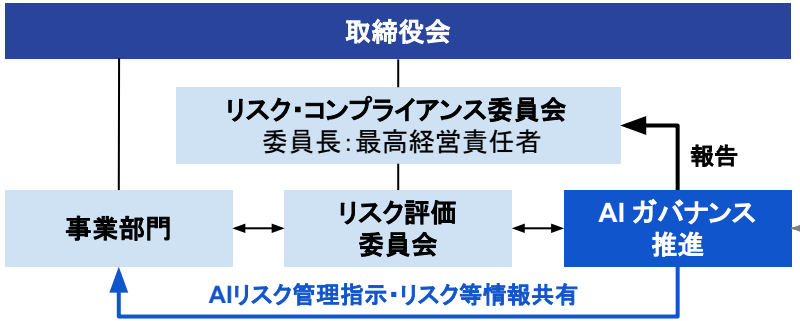
- 経産省のAI事業者ガイドライン・業界団体のAIプロダクト品質保証ガイドラインに沿ったガバナンス体制
- AIガバナンスとイノベーション促進を両立する開発手法を提示

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



人間中心の考え方
安全性、公平性、プライバシー

AI事業者ガイドライン（経済産業省）
セキュリティ、透明性、アカウンタビリティの確保と
誤った情報・動作の防止

AIプロダクト品質保証ガイドライン（QA4AI）
専門知識・AIプロダクトの犯罪・犯罪行為への悪用、
権利侵害の防止および社会倫理の確保

*広島AIプロセス：2023年にG7首脳に承認された、安心で信頼できる高度な 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

Preferred Networks

ガバナンスフレームワーク

人間中心

安全性

公平性

プライバシー

セキュリティ

透明性

アカウンタビリティ

誤った情報・動作

専門知識の悪用

犯罪・犯罪行為への悪用

権利および権利侵害

社会倫理

AI事業者ガイドライン（METI / MIC）

AIプロダクト品質保証ガイドライン（QA4AI）

PFN AIポリシー

PFNのAIにかかわる方針
https://preferred.jp/ja/company/aipolicy/

PFN AIガバナンス方針

AIに関わる企業活動の適切さを担保するための方針・規程

PFN AI品質ガイドライン

各プロジェクトが守るべきルール・手続きをまとめたガイドライン

プロジェクト開発方針

ガイドラインに記載された項目の対応の可否や選択理由などをまとめた資料

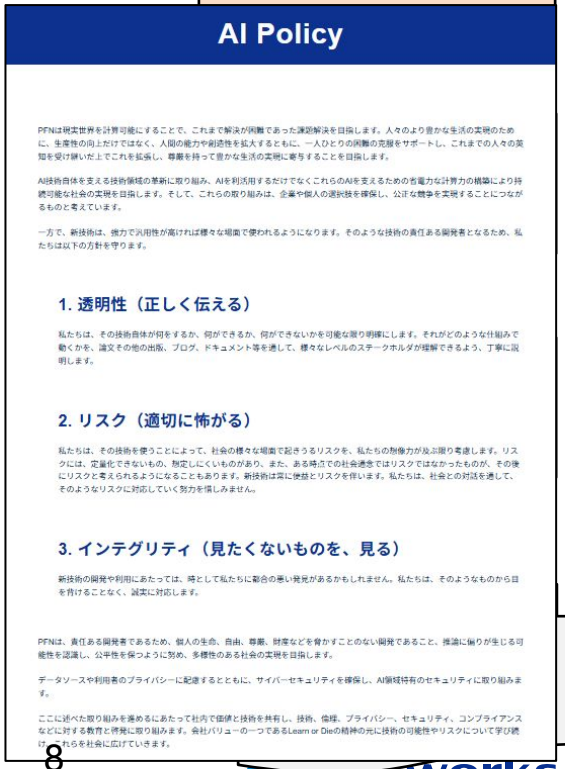
Preferred Networks

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）、アカウンタビリティ（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はなく、人間の能力や創造性を拡大するとともに、一人ひとりの困難の克服をサポートし、これまでの人々の英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指します。

AI技術自体を支える技術領域の革新に取り組み、AIを利活用するだけでなくこれらのAIを支えるための省電力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、企業や個人の選択肢を確保し、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、強力で汎用性が高ければ様々な場面で使われるようになります。そのような技術の責任ある開発者となるため、私たちは以下の方針を守ります。

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瞭にします。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きうるリスクを、私たちの想像力が及ぶ限り考慮します。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、ある時点での社会通念ではリスクではなかったものが、その後にリスクと考えられるようになることもあります。新技術は常に便益とリスクを伴います。私たちは、社会との対話を通して、そのようなリスクに対応していく努力を惜しみません。

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、推論に偏りが生じる可能性を認識し、公平性を保つように努め、多様性のある社会の実現を目指します。

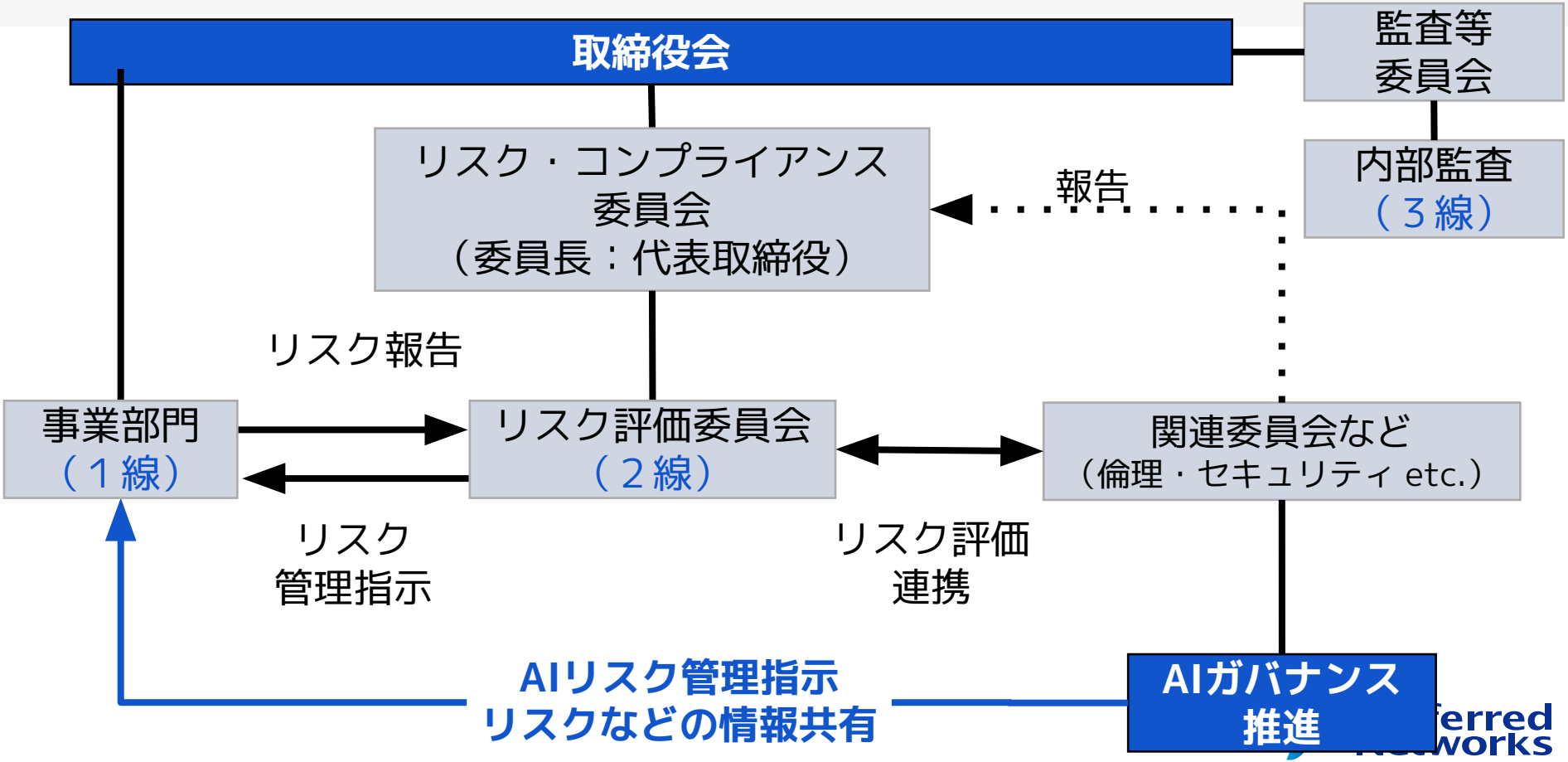
データソースや利用者のプライバシーに配慮するとともに、サイバーセキュリティを確保し、AI領域特有のセキュリティに取り組みます。

ここに述べた取り組みを進めるにあたって社内で個人技術を共有し、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みます。会社バリューの一つであるLearn or Dieの精神の元に技術の可能性やリスクについて学び続け、これらを社会に広げていきます。

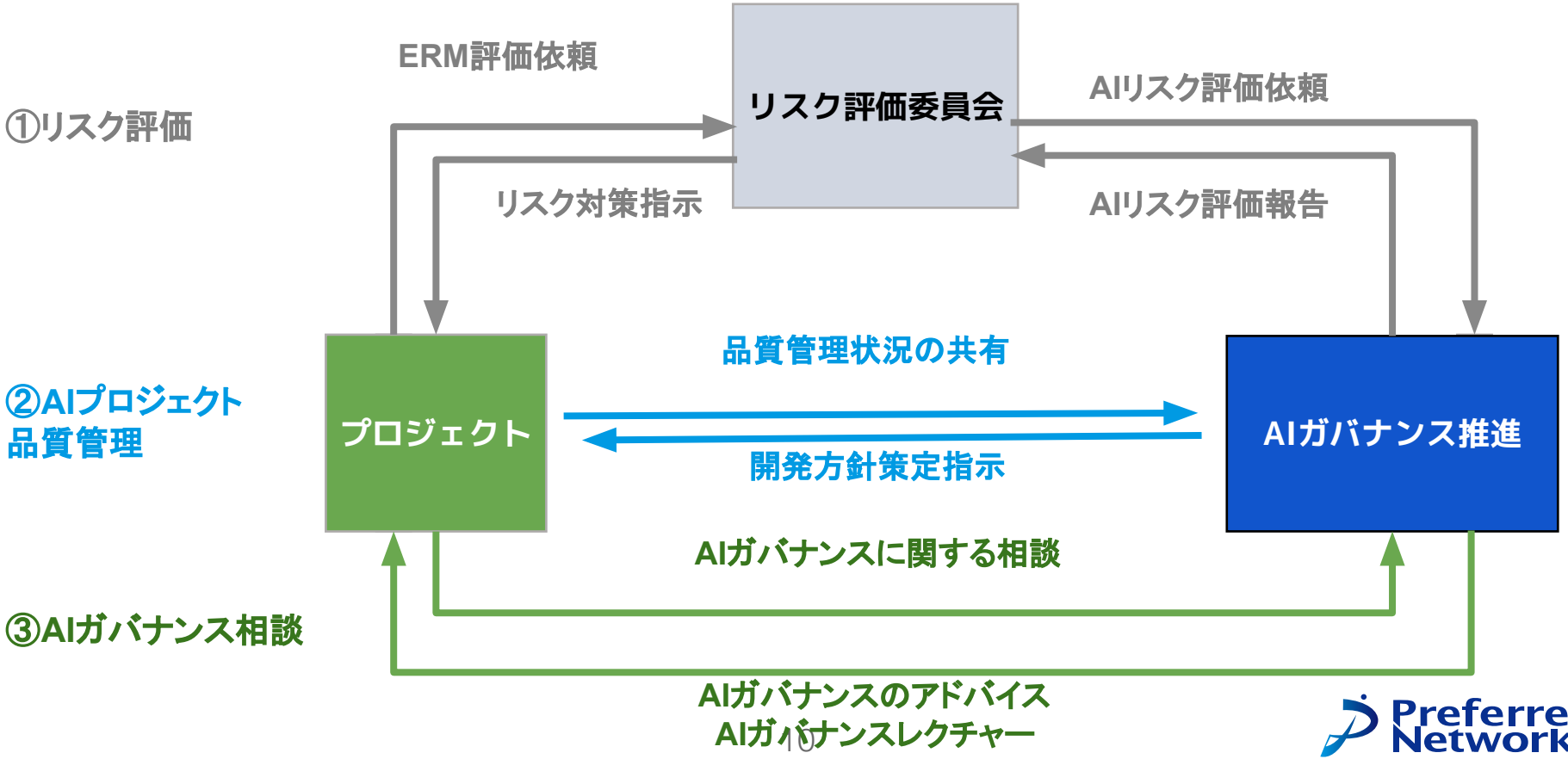
8

https://www.preferred.jp/ja/company/aipolicy/

規程・ガイドライン体系


取締役会
監査等
委員会
リスク・コンプライアンス
委員会
（委員長：代表取締役）
報告
内部監査
（３線）
リスク報告
事業部門
（１線）
リスク評価委員会
（２線）
関連委員会など
（倫理・セキュリティ etc.）
リスク
管理指示
リスク評価
連携
AIリスク管理指示
リスクなどの情報共有
AIガバナンス
推進

AIガバナンス推進


ERM評価依頼
リスク評価委員会
AIリスク評価依頼
①リスク評価
リスク対策指示
AIリスク評価報告
品質管理状況の共有
②AIプロジェクト
品質管理
プロジェクト
AIガバナンス推進
開発方針策定指示
AIガバナンスに関する相談
③AIガバナンス相談
AIガバナンスのアドバイス
AIガバナンスレクチャー

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。

Preferred Networks

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）

Preferred Networks

AI Governance Training

Preferred Networks

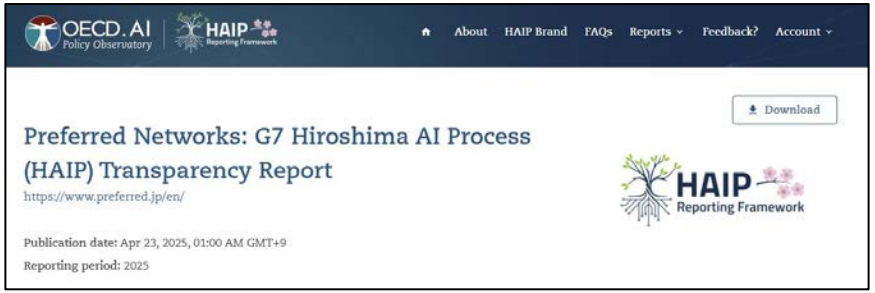
広島AIプロセス報告枠組みへの参加

広島AIプロセス・フレンズグループ　パートナーズコミュニティ参加企業・組織一覧（アルファベット順）

（2025年5月時点）

Amazon	Business Software Alliance (BSA)	Center for AI and Digital Policy (CAIDP)	富士通株式会社
Google	Impact AI	国際協力機構 (JICA)	KDDI株式会社
Microsoft Corporation	日本電気株式会社	日本電信電話株式会社	経済協力開発機構 (OECD)
OpenAI LLC	日本オラクル株式会社	Palo Alto Networks	株式会社Preferred Networks
楽天グループ株式会社	SaferAI	Salesforce,Inc.	ソフトバンク株式会社
国連開発計画 (UNDP)	世界銀行 (WB)	世界経済フォーラム (WEF)	

https://www.soumu.go.jp/hiroshimaaiprocess/supporters.html#partners



https://transparency.oecd.ai/reports/a86f4925-5cd5-4af7-b4f6-1b1f0984419e

Preferred Networks

まとめ

- **ビジネスモデル**：AIチップ、コンピューティングインフラ、大規模言語モデル、AIソリューションの垂直統合

- **ガバナンス組織**：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、実働部隊としてAIガバナンス推進を設置

- **ガバナンス体系**：AI事業者ガイドラインとQA4AIに準拠した階層的な方針体系

Preferred Networks

Preferred Networks

Making the real world computable