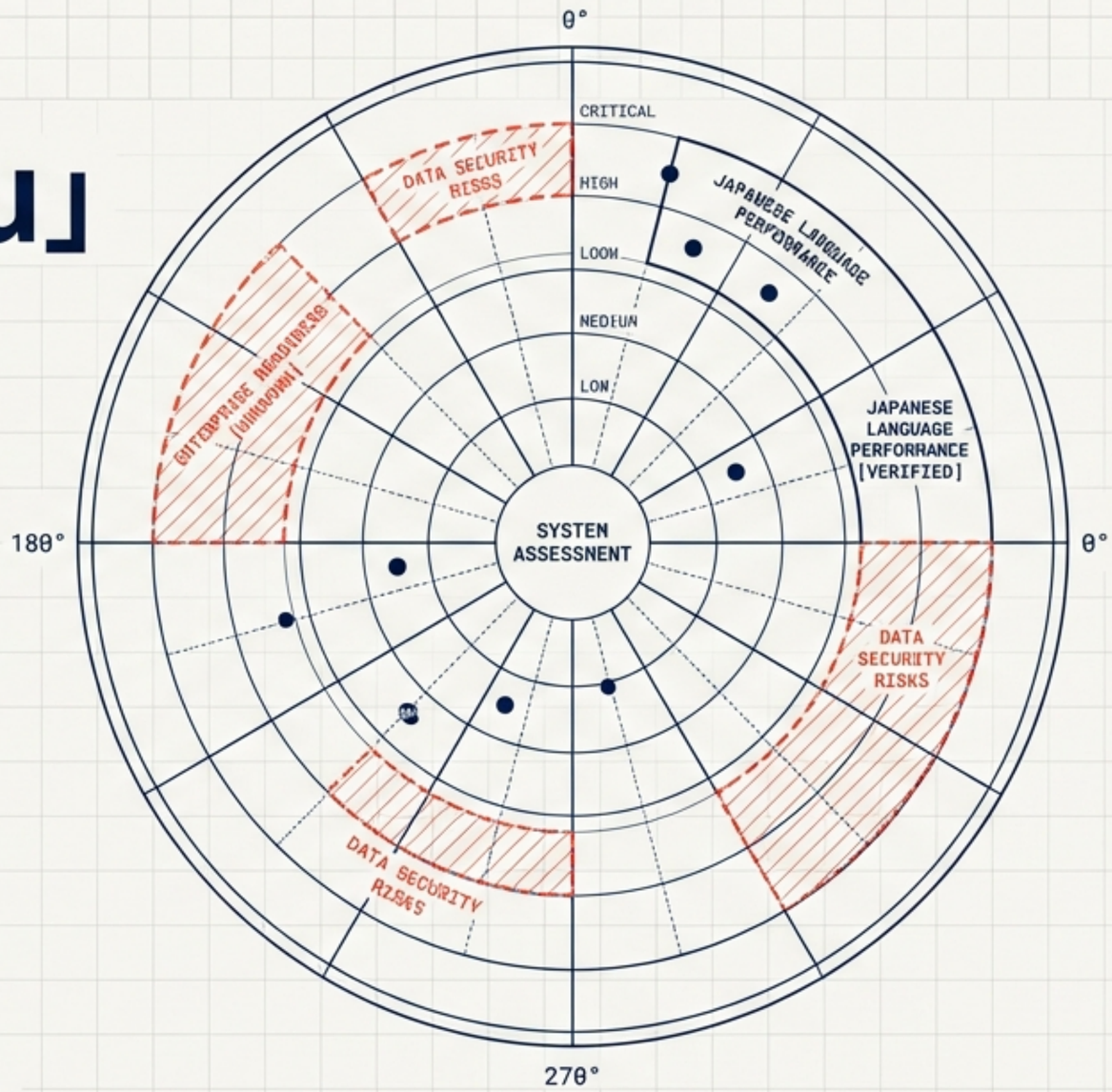
STRATEGIC INTELLIGENCE BRIEFING

Sakana AI「Namazu」&「Sakana Chat」公開調査報告

日本特化AIの実力と、企業導入における「光と影」の戦略的評価



公開日: 2026年3月24日 | 評価対象: Namazu（α版）/ Sakana Chat

Executive Summary

公開の事実 (THE FACT)

公開日: 2026年3月24日一般公開（約1,000名のβテスト経由）
対象: 日本特化LLM「Namazu (α版)」と検索統合チャット「Sakana Chat」
条件: 無料・登録不要（国内IP限定）

公式の主張 (THE CLAIM)

海外モデルの基礎能力を維持しつつ、日本の文化・社会文脈に合わせて事後学習。
政治・歴史のトピックで中立性・正確性を改善（回答拒否率72%→ほぼ0%へ）。

実務上の障壁 (THE RISK)

オプトアウト不可: 入力データは学習に利用される可能性あり。
未公開のブラックボックス: 技術詳細、学習データ出典、モデルウェイトは現時点で未公開（後日発表予定）。

現時点の結論 (THE VERDICT)

教育・個人の情報収集には「推奨（Go）」。
機密情報を扱う企業・業務利用には「原則待機（No-Go）」。

戦略的意図：なぜ今、日本特化なのか？

Massive Open Models (US/China)

＋

Targeted Post-Training (Japan)

＝

Sovereign Cultural Alignment

1. 開発コストの回避

巨大モデルのゼロからの事前学習競争を避け、既存の高性能オープンウェイトモデルを基盤として活用。

2. イデオロギーの書き換え

開発元（米・中）の価値観や情報統制による「バイアス」を、独自の事後学習（Post-training）で上書きする。

3.「ソブリンAI」の軽量な実現

各国の文化・安全保障要件に適合させる技術実証の第一弾として日本市場を選択。

技術仕様：「Namazu」モデルの系譜

DeepSeek-V3.1-Terminus
MIT License | 671B / 37B Active

Llama-3.1-Namazu-405B
Llama 3.1 Community | 405B

Namazu-gpt-oss-120B
Apache 2.0 Base | 117B / 5.1B Active

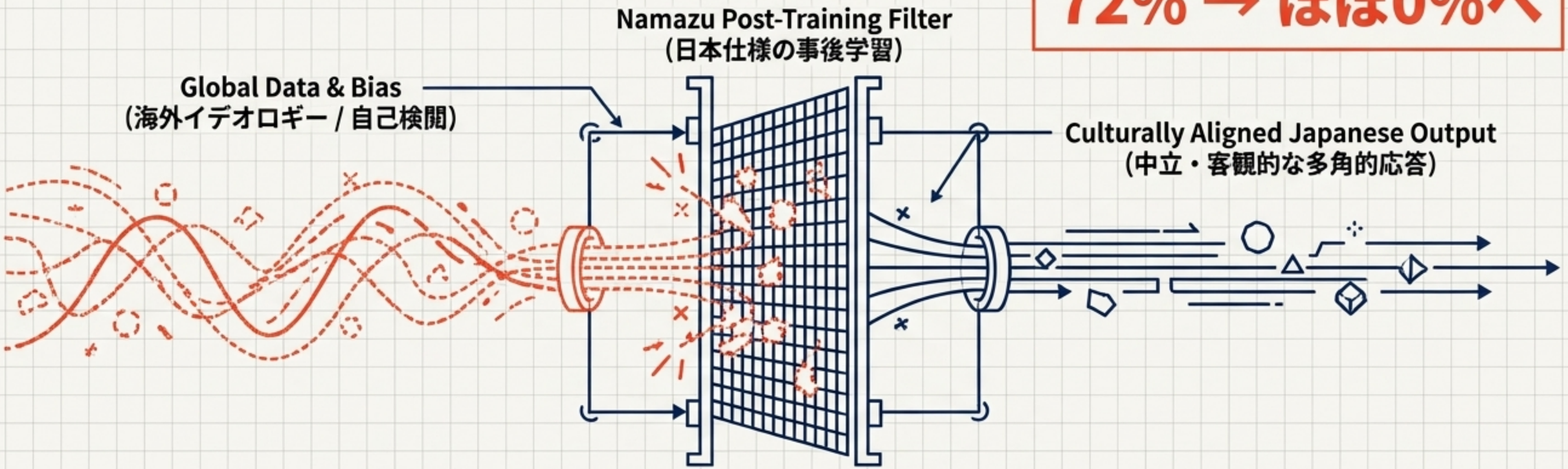
Namazu Alpha: 未公開の領域 (Post-Training)

- **独自データセット**
  (出典、収集方法、合成データ比率：未公開)
- **日本語最適化手法**
  (トークナイザ改修、SFT/RLHF設計：未公開)
- **モデルウェイト＆アーキテクチャ**
  (推論モードの差異等：準備中)

評価の制約：ベースモデルの仕様は判明しているが、事後学習の詳細な手法と学習データが未公開のため、現時点での外部監査は不可能。

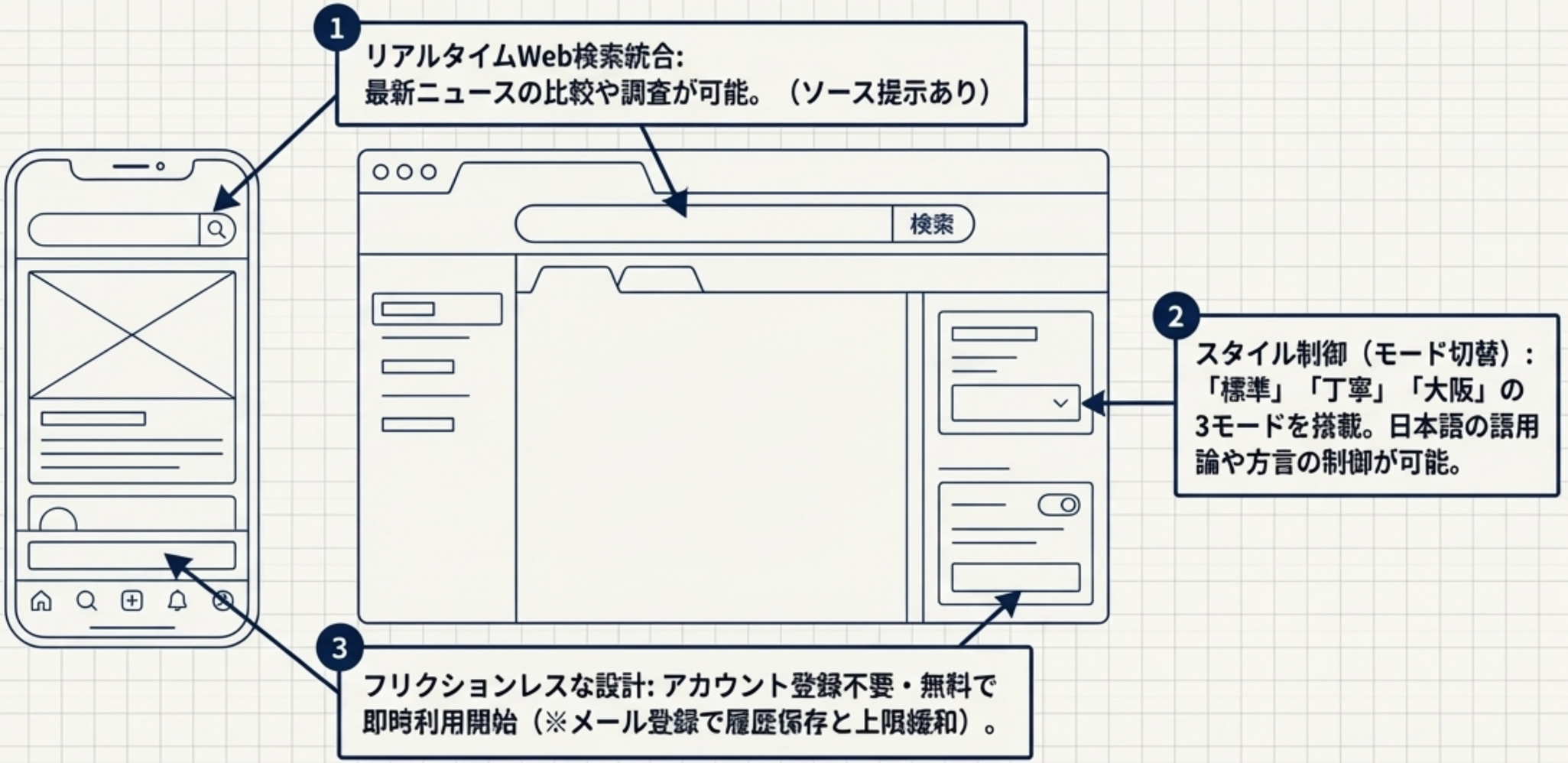
コア・バリュー：「日本仕様」へのバイアス是正

回答拒否率：
72% → ほぼ0%へ



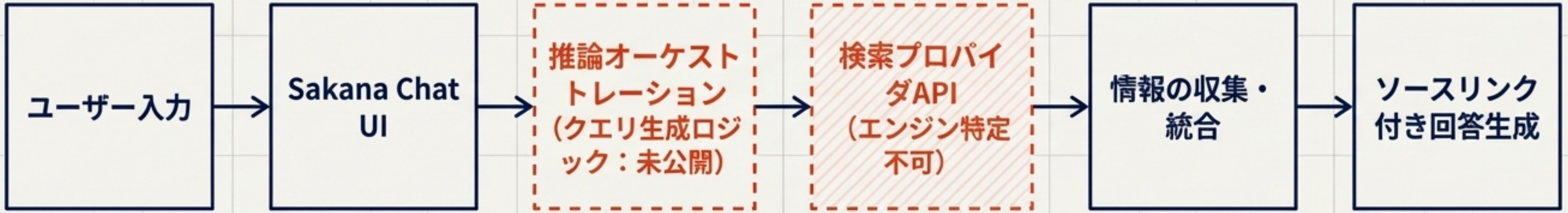
The Problem (課題)	The Solution (解決策)	The Result (結果)
海外のオープンモデルは、開発国の政治的・文化的な安全性基準に過剰に縛られ、日本の文脈において無害な質問でも回答を拒否（自己検閲）する傾向がある。	政治・歴史・外交などのトピックにおいて、日本の社会的文脈に沿った独自データで事後学習を実施。	中立的かつ客観的事実に即した多角的な応答が可能に。基礎能力（推論・コーディング）を低下させずにこれを実現したと公式は主張している。

プロダクト分析：Sakana ChatのUXと主要機能



ユーザー体験の焦点：単なる「精度の高い日本語」ではなく、レスポンスの速さと「大阪モード」のようなUI上の遊び心が、アーリーアダプターの高評価に直結している。

アーキテクチャ図：Web検索統合の裏側



確定情報（KNOWN ELEMENTS）

- リアルタイム検索を実行し、情報を「収集・統合」して回答を生成する。
- 生成結果には、出典元へのソースリンクが提示される。

未公開情報（UNKNOWN ELEMENTS）

- **検索プロバイダ:** 裏側でどの検索エンジンを使用しているか。
- **クエリ処理:** 検索キーワードの生成ロジックや、再ランキングの仕組み。
- **安全制御:** 検索結果のフィルタリング基準。

「第三者サービスが含まれる」と規約にあるが、検索インデックスの依存先がブラックボックスであるため、網羅性や偏りの検証が困難。

データとプライバシー：国内保管の利点と「学習利用」の罠

データ保管 (Data Storage) - 評価：ポジティブ

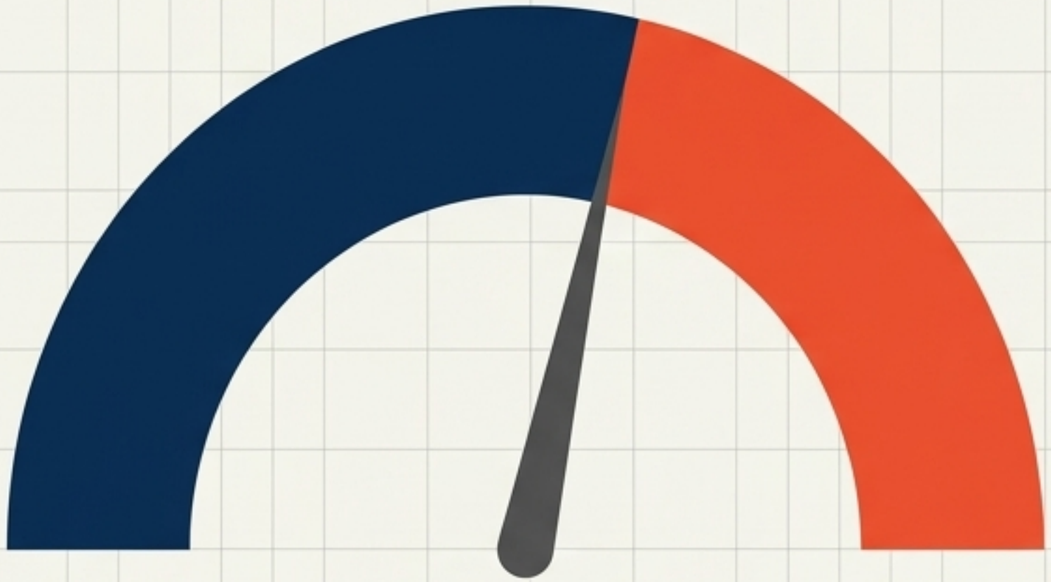
- 会話履歴やアカウント情報は日本国内のGoogle Cloud（Firestore）に保管。
- 原則として国外移転は行わない（データ・ローカライゼーションの観点では高評価）。

データ利用 (Data Usage) - 評価：重大なリスク

- **オプトアウトの不在:** 入力データ（プロンプト等）はAIモデルの学習・改善に利用される可能性があり、現時点で拒否設定が存在しない。
- **消去義務の免責:** アカウント削除は可能だが、「学習済みモデルの重み」や「一時キャッシュ」についてはSakana側に削除義務がないと規約に明記されている。

法人向けアラート：個人情報・機密情報の入力は規約違反かつ漏洩リスクに直結する。

市場の反応：期待と懸念のセンチメント分析



57%
ポジティブ

43%
ネガティブ

データソース：Yahoo!リアルタイム検索（自動まとめ：120件のポスト）

ポジティブな反応（Positive Drivers）

- Speed: 無料かつ圧倒的な体感レスポンス速度。
- Naturalness: 極めて自然な日本語表現。
- UX：「大阪弁モード」の面白さと、ソースリンクが辿れる実用性。

懸念とネガティブな反応（Concerns）

- Privacy: 学習利用のオプトアウト機能がないことへの強い警戒。
- Access: 地域制限（日本国内IPのみ）により海外拠点・VPN 経由で使えない不満。
- Quality: 複雑な文脈での論理破綻（ハルシネーション）やUIの導線課題。

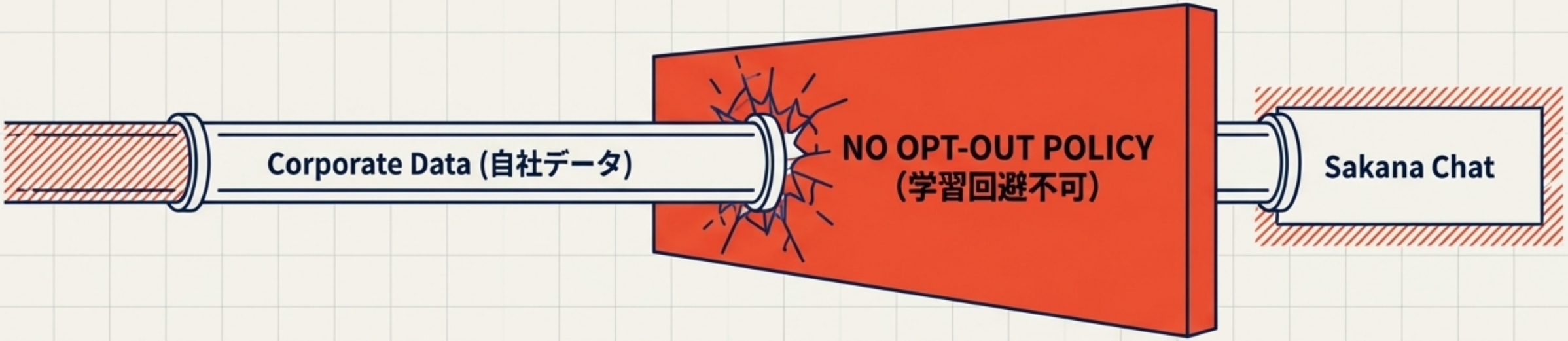
定性ソース：はてなブックマーク（262 users / 39コメント），LinkedIn専門家レビュー

競合比較：「AI検索」ランドスケープにおける独自の立ち位置

サービス	検索・ソース提示	提供形態	最大の違い（思想）
Sakana Chat	**リアルタイム統合/リンクあり**	**無料/国内限定**	**「日本文化への事後学習」によるモデル自体のローカライズ**
ChatGPT Search (OpenAI)	インライン引用/Sourcesパネル	グローバル	GPTエコシステム内でのシームレスな統合体験
Perplexity	リアルタイム検索/出典リンク	グローバル	「Answer Engine（回答エンジン）」としての特化UX
Gemini (Google)	Google Search grounding	API/クラウド	RAG/根拠付け（Grounding）の信頼性重視
Copilot Search (Microsoft)	要約＋探索導線	Bing統合	検索UI中心の「次の探索」の提示

結論：機能面での差別化要素は薄いが、「思想（脱・海外バイアス）」と「運用（国内保管）」に特異性がある。

リスクプロファイル①：企業利用の壁「オプトアウト不在」の実務影響



1. 個人情報の取り扱いリスク

ユーザーの規約上で「個人情報入力は禁止」と明記。日本の個人情報保護委員会のガイダンスに照らしても、入力制御なしでの業務利用は極めてハイリスク。

2. 機密情報の蒸留懸念

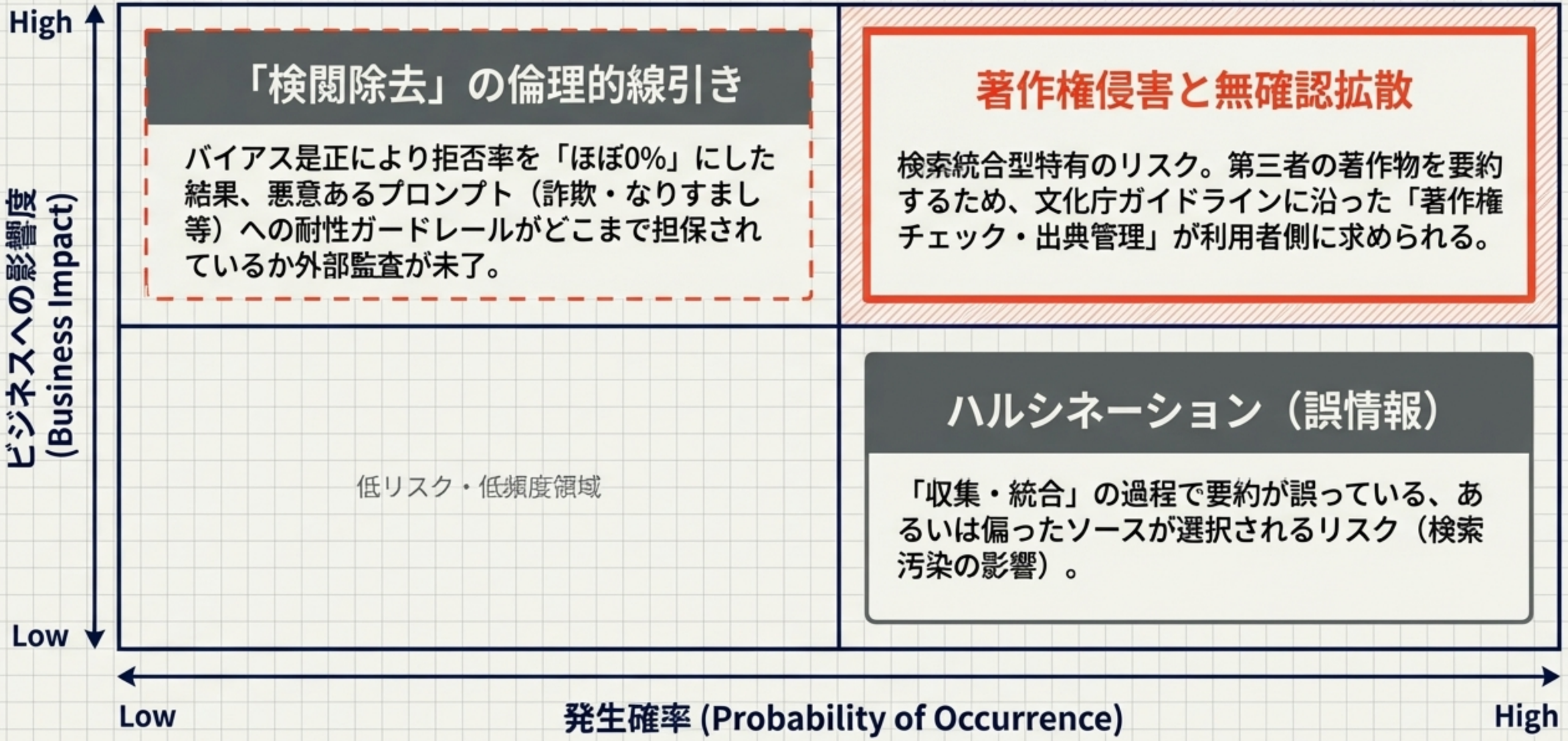
入力データが学習に回るため、自社の営業秘密や非公開コードがモデルの重み　（Weights）に吸収され、他者の回答として出力されるリスクを排除できない。

3. 削除権の限界

「当社がコンテンツに基づき学習を行ったモデルの重み等は削除義務を負わない」。事後的なデータ消去要求が技術的・法務的に通用しない。

実務影響：エンタープライズ向けの管理機能（チーム版、ログ監査、データ学習除外設定）が実装されるまで、機密情報を扱う業務での全社導入は不可能。

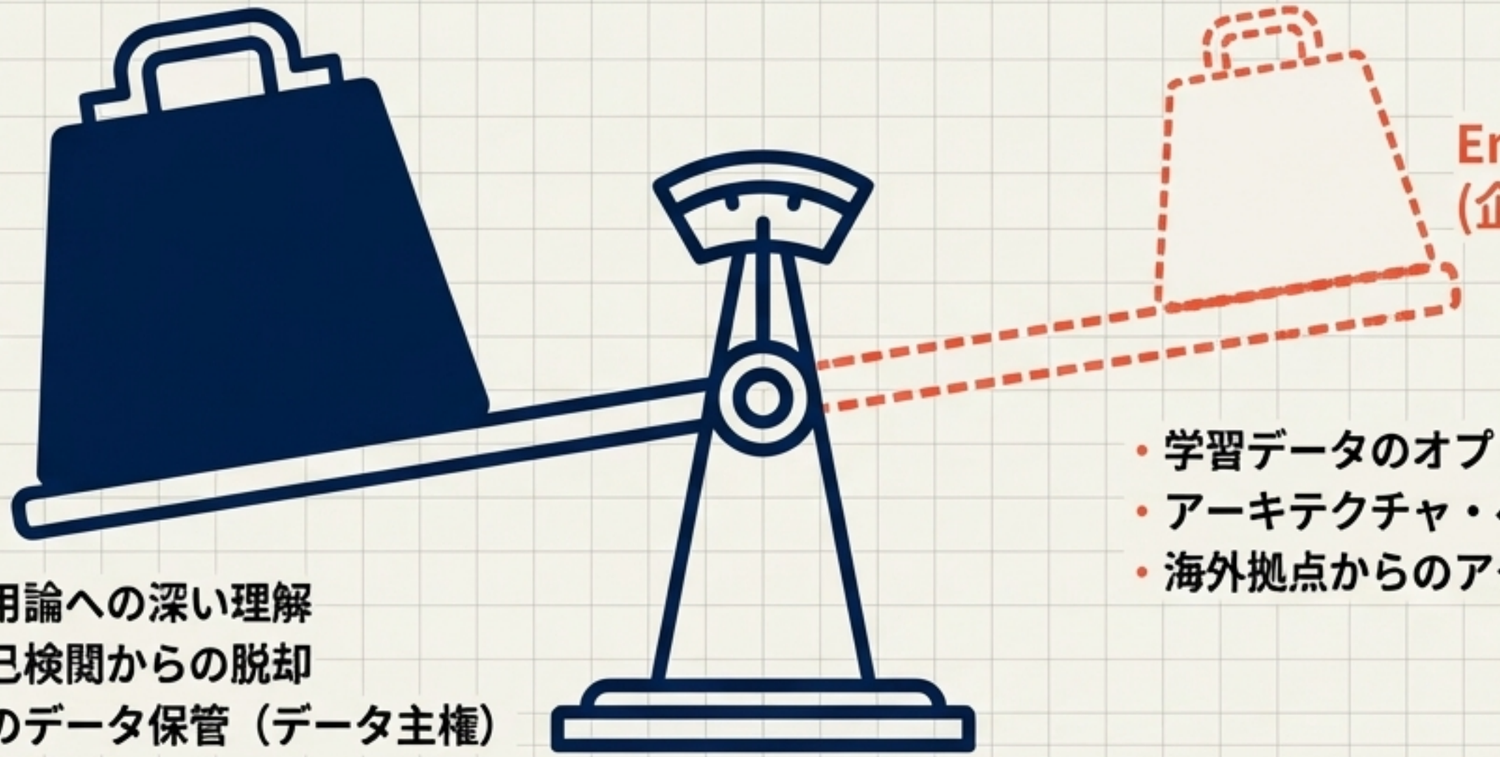
リスクプロファイル②：法務・倫理の境界線



シンセシス：ローカルな最適化 vs エンタープライズ要件

Cultural Fit
(文化的適合: 高)

- 日本の文脈・語用論への深い理解
- 海外モデルの自己検閲からの脱却
- 国内サーバーでのデータ保管（データ主権）

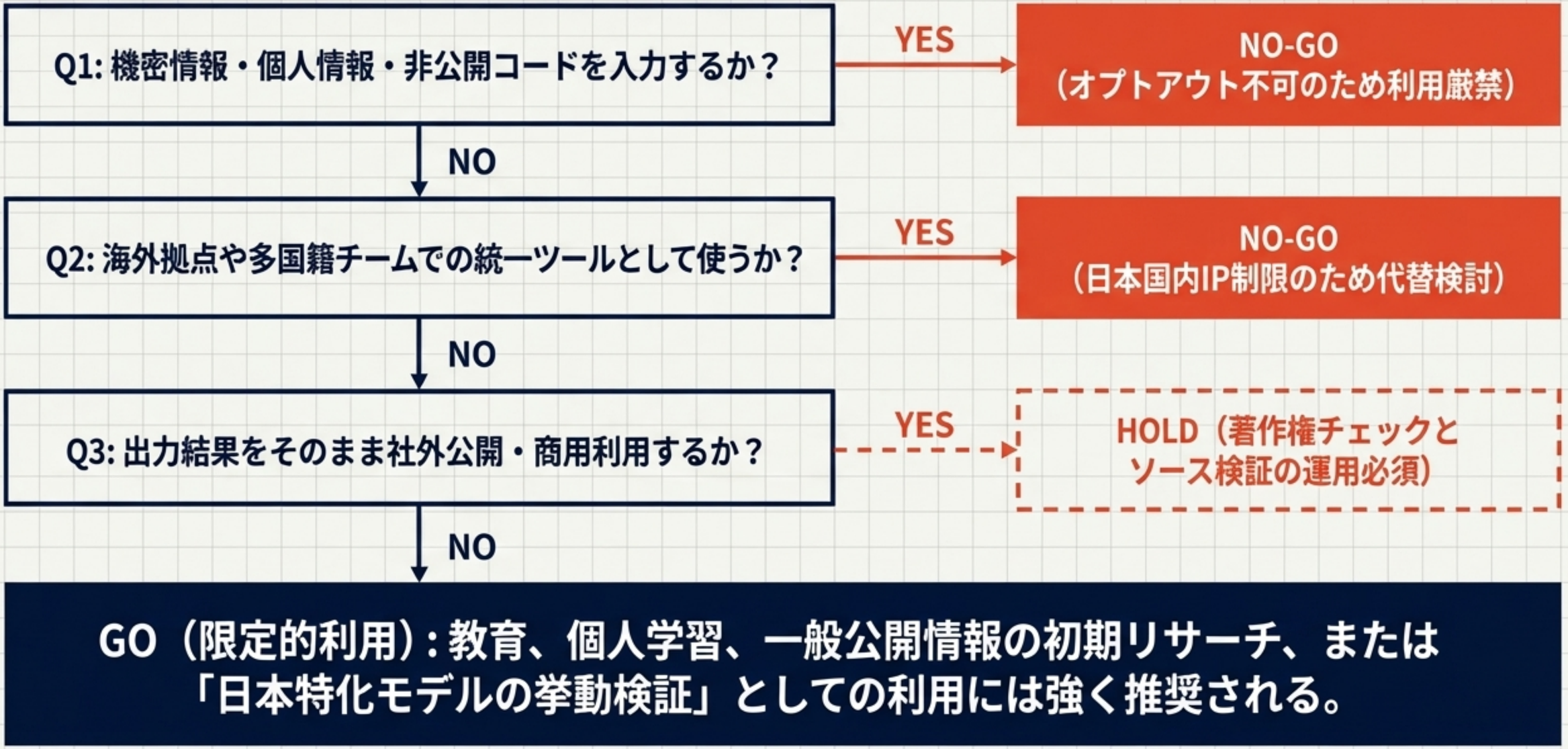


Enterprise Readiness
(企業導入準備度: 低)

- 学習データのオプトアウト機能なし
- アーキテクチャ・ベンチマークの不透明性
- 海外拠点からのアクセス不可

Sakana Chat（α版）は、「日本の文化的コンテキスト」というローカルな最適化には見事に成功している。しかし、「データの秘匿性・統制」というエンタープライズの必須要件にはまだ適応していない。 現段階はあくまで『B2C向けの技術デモンストレーション』として捉えるべきである。

意思決定：組織導入のための「Go / No-Go」診断



今後の展望：短中期のロードマップと「次に来るべき透明性」

短期 (Short-Term)

- **Technical Reportの公開:** ベンチマークの詳細スコア、学習データの出典、安全性評価（レッドチーミング）結果の開示待ち。
- **モデルウェイトの公開:** ローカル環境での推論検証と、詳細なライセンス条件の確定。

中期 (Mid-Term)

- **エンタープライズ統制の提供:** オプトアウト設定、ログ管理、Pro版の登場が、企業本格導入の絶対条件となる。
- **API提供の拡大:** 自社システムへのNamazuモデルの組み込み。

CLOSING STATEMENT: 現時点では「ポテンシャルを示す鋭利なプロトタイプ」である。企業は利用を急ぐのではなく、今後の「透明性」と「統制機能」のアップデートを注視し、導入のトリガーを引く準備をしておくべきである。