

# ソフトバンクの国産LLM「Sarashina」の最新動向と知財業務への適用可能性に関する包括的分析

Gemini 3.1 pro

## 1. 序論: 生成AIの独自進化と「デジタル主権」が要請される新たな時代背景

2020年代前半における生成人工知能(AI)の黎明期は、主に海外の巨大テクノロジー企業が開発した汎用大規模言語モデル(LLM)によって市場が牽引されてきた。これらのモデルは圧倒的なパラメータ規模と多言語対応能力を誇る一方で、エンタープライズ領域における本格的な社会実装が進行するにつれ、企業は新たな構造的課題に直面することとなった。それは、自社のコアコンピタンスに関わる機密データの保護、各国の複雑化するデータ保護法規制への準拠、そして日本特有の高度な専門用語や商習慣への深い理解という壁である。

こうした転換期において、ソフトバンク株式会社の子会社であるSB Intuitions株式会社が開発を主導する国産LLM「Sarashina(さらしな)」は、日本市場における生成AIのパラダイムを根本から変革し得る中核的なプラットフォームとして進化を遂げている。Sarashinaは、単に日本語が流暢なAIを構築することを目的としたものではない。その根底には、国家や企業のデータが外国の法規制や技術的ブラックボックスに依存することなく、自律的に管理・運用されるべきであるという「ソブリンクラウド(主権型クラウド)」および「デジタル主権」の確立という極めて戦略的なビジョンが存在している<sup>1</sup>。

本稿では、2026年最新の状況における「Sarashina」モデルの技術的アーキテクチャ、マルチモーダル化の進展、およびそれを支えるインフラストラクチャの現状を網羅的かつ詳細に分析する。その上で、企業活動において最も高度な機密性と専門的かつ論理的な言語処理能力が要求される「知的財産(知財)業務」に対するSarashinaの適用可能性について、技術的、運用的、戦略的な観点から多角的な検証を行う。知財領域は、特許明細書という特殊な言語空間と、未公開技術という絶対的な機密性が交差する領域であり、Sarashinaが志向する「国産・高セキュリティ・日本語特化」という特性がシナジーを生み出す有力な選択肢となり得る。

## 2. 「Sarashina」モデルの技術的系譜と最新アーキテクチャの全容

SB Intuitionsが開発するSarashinaの技術的特性は、海外のオープンソースモデルをベースに日本語データで追加学習(継続事前学習)を行うアプローチを意図的に排し、モデルの学習を初期段階からすべて独自に行う「フルスクラッチ開発(ゼロからの構築)」の哲学を貫徹している点にある<sup>1</sup>。このアプローチは、計算リソースの観点からは多大な投資を要するものの、学習データやアルゴリズムの内部構造を完全に掌握できるため、バグや予期せぬ挙動が発生した際の内部メカニズムの解析が容易となる。結果として、産業用途において不可欠な「説明可能性(Explainability)」を担保し、エンタープライズ顧客に対して高い安心感と導入判断材料となる重要な要素を提供している<sup>1</sup>。

## 2.1 パラメータ規模の拡大と基盤モデルの精緻化

Sarashinaモデルは、初期世代の「Sarashina1」シリーズから始まり、その開発過程で得られた知見を基に大幅なアーキテクチャのブラッシュアップと性能向上を果たした「Sarashina2」シリーズへと進化を遂げた<sup>3</sup>。2026年現在では、さらに派生モデルである「Sarashina2.1」および「Sarashina2.2」シリーズへとエコシステムが細分化・拡張されている<sup>5</sup>。

モデルの根幹をなす基盤アーキテクチャにおいて、特筆すべきは「Sarashina2-70B」の存在である。このモデルは、強固なLlama2アーキテクチャとRoPE (Rotary Position Embedding) を採用して構築された700億パラメータの大規模言語モデルである<sup>6</sup>。事前学習においては、CCNetやHojiCharといった高度なクリーニング技術で処理された1兆トークン規模の日本語Common Crawlデータと、書籍データ(books3)を除外した英語のSlimPajamaコーパスを含む、計2.1兆トークンという膨大なデータセットが投入されている<sup>6</sup>。さらに、言語処理の効率性を決定づける語彙(ボキャブラリー)サイズは102,400トークンという広範なカバレッジを誇り、8,192トークンのコンテキスト長をサポートしている<sup>6</sup>。また、日本語処理に最適化するため、Unigram言語モデルとバイトフォールバック(byte-fallback)機能を備えたSentencePieceトークナイザを採用している。これにより、日本語特有の事前の単語分割(形態素解析など)を必要とせず、生の文章を直接的かつ高効率に処理することが可能となり、処理遅延の低減と文脈理解の向上に大きく寄与している<sup>6</sup>。

## 2.2 Sparse Upcycling技術によるMoEアーキテクチャの到達点

モデルの大規模化に伴う推論コストの増大というジレンマを克服するため、SB Intuitionsは最新のアーキテクチャ技術を導入している。その集大成が、4,650億(465B)パラメータを超える巨大なMixture-of-Experts(MoE)モデル「Sarashina2-8x70B」である<sup>7</sup>。

このモデルは、既存の高精度なDense(密)モデルであるSarashina2-70Bのチェックポイントから初期化を行い、そこに「Sparse Upcycling(スパース・アップサイクリング)」と呼ばれる高度な手法を適用することで効率的に構築されている<sup>7</sup>。スパースモデルは、モデル全体のパラメータサイズと計算コストを切り離すことを可能にする。すなわち、465Bという莫大なパラメータを持ちながらも、特定の推論タスクにおいては関連性の高い一部のエキスパートネットワークのみが活性化されるため、推論時の計算負荷を劇的に抑えつつ、表現力と精度を飛躍的に高めることができる。

推論には16基のH100またはA100(80GB)GPUを使用し、BF16(Bfloat16)テンソル型で動作するなど、極めて高度なハードウェア要件が設定されている<sup>7</sup>。日本語と英語のウェブコーパスをミックスして学習されたこのバイリンガルモデルは、専門性の高いタスクにおいて、計算リソースを最適化しながら海外の最先端モデルに匹敵する、あるいはそれを凌駕する推論能力を発揮する基盤となっている<sup>1</sup>。

## 2.3 マルチモーダル化と特定タスク向けモデルのエコシステム

言語生成能力の高度化と並行して、Sarashinaのエコシステムは視覚や音声といったマルチモーダル領域へと急速に外延を広げている<sup>1</sup>。これにより、テキストデータだけでなく、企業内に眠る多様な非構造化データの統合的解析が可能となっている。

- 視覚・画像処理(VLM)とOCR: 「Sarashina2-Vision」シリーズは、日本国内のベンチマークにおいて最高レベルの性能を記録した画像理解エンジンである<sup>1</sup>。その後継となる「Sarashina2.2-Vision-3B」は、40億(4B)パラメータのコンパクトなサイズでありながら、ユー

ザーが入力した画像と質問に基づき、視覚情報内から回答を抽出する推論能力を備えている<sup>3</sup>。また、日本語と英語の解析に特化して調整されたエンドツーエンドのOCR(光学文字認識)モデル「Sarashina2.2-OCR(3B)」は、複雑なレイアウトを持つ帳票やスキャン文書からのテキスト抽出を実現する<sup>10</sup>。

- テキスト音声合成(TTS):「Sarashina2.2-TTS」は、0.8Bパラメータの大規模言語モデルをベースにしたテキスト音声合成システムである。日本語と英語の両言語に対応し、ゼロショットの音声生成をサポートしている。これにより、特定の事前学習なしに、多様な発話スタイルで自然な音声を出力することが可能であり、コンタクトセンターの自動化などに威力を発揮する<sup>5</sup>。
- ベクトル化・エンベディング: RAG(検索拡張生成: Retrieval-Augmented Generation)システムの構築に不可欠なテキスト埋め込みモデルとして「Sarashina-Embedding-v2-1B」が公開されている<sup>3</sup>。この10億(1B)パラメータのモデルは、コントラスト学習を用いて構築されており、ユーザーがクエリに指示文を付与することで、個別のタスクに合わせてテキストベクトルを動的に調整できる特性を持つ。JMTEBベンチマークの検索(Retrieval)、意味的類似度(STS)、およびリランキングタスクにおいて国内トップレベルのスコアを記録しており、社内情報検索の要となる<sup>3</sup>。

# Sarashinaモデル・エコシステムの技術的機能マトリクス

カテゴリ	代表モデル	パラメータ規模	主要機能・特徴
テキスト生成	Sarashina2-70B	70B	Llama2アーキテクチャ、102.4kボキャブラリー
MoE / 高性能	Sarashina2-8x70B	465B	Sparse upcycling手法、日英バイリンガルウェブコーパス
ビジョン / マルチモーダル	Sarashina2.2-Vision-3B	4B	Image-to-Text
OCR	Sarashina2.2-OCR	3B	日英ドキュメント解析
エンベディング	Sarashina-embedding-v2-1B	1B	RAGおよびSTS向けに最適化
音声	Sarashina2.2-TTS	0.8B	ゼロショット音声生成

SB Intuitionsが提供するSarashinaモデル群は、用途や推論リソースに応じて最適化されており、テキスト生成から視覚・音声処理まで広範なエンタープライズニーズに対応する。

Data sources: [SB Intuitions](#), [Featherless.ai](#), [PromptLayer](#), [GitHub \(SB Intuitions\)](#), [Interfaze.ai](#), [SoftBank](#)

これらのモデル群は事前学習済みモデルとして提供されるだけでなく、ソフトバンクが提供する「Sarashina API」を通じて、実務的な形態（チャット補完API、エンベディングAPIなど）として企業システムに統合される<sup>1</sup>。API経由でアクセス可能なモデルには、軽量で高速な応答を実現する「Sarashina2 mini」や、差別的発言や違法行為を助長する有害な入出力を検知・ブロックし、企業としてのレピュテーションリスクを低減するセキュリティフィルター「Sarashina2 guard」などが含まれており、運用環境が整えられている<sup>1</sup>。

## 3. インフラストラクチャ基盤と「デジタル主権」を担保するソブリクラウド構想

Sarashinaモデルがいかに高性能であっても、それをエンタープライズの現場、とりわけ機密性の高い業務で運用するためには、強固な計算基盤と法的に保護されたデータ運用体制が不可欠である。ソフトバンクは国内AI計算基盤の拡充を進めており、2025年7月時点で合計1万基超のGPU規模に達したと公表している。そして、このハードウェア基盤の上で、データの安全性と自律性を高める「ソ

ブリッククラウド(Sovereign Cloud)」構想を推進している。

### 3.1 デジタル主権を構成する4つの柱

「主権(Sovereign)」とは、国家が自律的に統治を行う権利を持つように、データやクラウドシステムの運用権限を自国の管理下に置くことを意味する。世界の潮流を見ると、欧州ではGDPR(一般データ保護規則)によるデータ移転制限やGAIA-Xによる独自のクラウド標準化が進み、米国ではCLOUD法(Clarifying Lawful Overseas Use of Data Act)により、自国企業が管理するデータであれば物理的所在地を問わず政府がアクセスを要求できる体制が敷かれている。また中国では、サイバーセキュリティ法等により国内データの国内保存が原則化されている<sup>1</sup>。

このような地政学的なデータ囲い込みが加速する中、日本政府も「DFFT(Data Free Flow with Trust:信頼性のある自由なデータ流通)」を提唱しつつ、機密性の高いデータの国内保存を推奨している。ソフトバンクはCloud PF Type Aにおいて、主にデータ主権・運用主権・技術主権を重視している<sup>14</sup>。なお、本稿では知財業務への適用観点から、システム主権も含めて整理する。

主権の分類	定義とソフトバンク・Sarashinaにおける実現メカニズム	企業・国家にとっての戦略的意義
データ主権 (Data Sovereignty)	データの物理的な保管場所、アクセス権限、保護スキームを自国(日本)の法制度に基づいて厳密に管理する。Sarashina APIに入力されたデータはAIの再学習に利用されることはなく、国内データセンターから外部に持ち出されることはない。	外国法・域外移転・外部委託に伴うリスクを大幅に低減し得る。
システム主権 (System Sovereignty)	クラウドやITシステムの重要コンポーネントを国内で構築・維持する。金融機関や政府機関などの重要インフラにおいて、海外ハイパースケーラーへの過度なアーキテクチャ依存を避ける。	外国製ソフトウェアの仕様変更やサービス停止に伴うビジネスの断絶を防ぎ、システムの設計・拡張に対する決定権を自国企業が保持する。
運用主権 (Operational Sovereignty)	クラウドシステムやデータセンターの運用管理、障害対応、監視体制をすべて自国内の組織(ソフトバンク等)で完結させる。	有事や災害時において、システムメンテナンスやシャットダウンの権限を国内で掌握でき、地政学的リスクに伴うサービス停止を防ぐ。

<p>技術主権 (Technology Sovereignty)</p>	<p>国産AI基盤であるSarashinaをゼロからスクラッチ開発することで、海外ベンダーへのコア技術の依存から脱却し、継続的な研究開発と人材育成を国内で推進する。</p>	<p>AI、半導体、クラウドネイティブ技術における国家および企業の長期的な自律性と競争力を維持・発展させる。</p>
--	--	--

### 3.2 「Cloud PF Type A」とSarashinaの統合による安全域の創出

Cloud PF Type AIは2026年4月に東日本データセンターで提供開始され、同基盤上でSarashinaを活用した生成AIサービスが2026年6月から順次提供される。このサービスは、ソフトバンクが保有するデータセンター内で運用され、顧客が暗号鍵を管理することで高度なアクセス制御を実現する。知財や営業秘密といった極めてセンシティブなデータを扱う企業にとって、このインフラは「Sovereign IP AI Workflow(主権型知財AIワークフロー)」とも呼ぶべきセキュアな閉域環境を提供する。具体的には、研究開発部門の技術者や知財部門が入力した未公開の発明情報が、国内の「Cloud PF Type A」データセンターというトラストゾーン内へ送信される。この環境内で、SarashinaのマルチモーダルLLMや、特許公報データベースをベクトル化したRAGシステムが情報の処理・分析を行い、その結果が知財部門へ還元される。この一連のアーキテクチャは、生成AIの業務適用において企業が抱いていた主要な懸念を軽減する重要な要素となる。

## 4. 社会実装の最前線：公共インフラから産業特化型マルチエージェントへの飛躍

Sarashinaの技術的完成度とソブリンクラウドの信頼性は、日本社会のインフラや高度な産業領域におけるプロジェクトにおいて実証されつつある。

### 4.1 デジタル庁「G-Biz Portal」による公共セクターでの採用

2026年3月、SB Intuitionsはデジタル庁が運営する事業者向けポータルサイト「G-Biz Portal」のデータベース拡充プロジェクトにおいて、技術協力者として「Sarashina API」を提供した<sup>12</sup>。行政情報という正確性と公平性が厳しく問われる領域において、Sarashinaが基盤技術として採用された事実は、同モデルの出力精度と日本語理解能力が社会的な水準に達していることを証明するマイルストーンである。

### 4.2 中外製薬との共同研究が示す「自律型AIエージェント」のパラダイム

さらに専門的な領域での適用例として、2025年1月に合意された中外製薬株式会社、ソフトバンク、SB Intuitionsの3社による共同研究が挙げられる<sup>13</sup>。このプロジェクトは、新薬開発プロセスにおいて莫大なコストと時間を要する「臨床開発業務(治験)」を生成AIによって革新することを目的としている<sup>13</sup>。

この取り組みの核心は、汎用言語モデルの導入にとどまらず、「製薬産業特化型のLLM」と、人間の指示を待たずに自律的に業務を遂行する「AIエージェント」の共同開発にある。プロジェクトは以下の3つのフェーズで段階的に推進されるよう設計されている<sup>13</sup>。

1. フェーズ1(特定タスク向けプロトタイプの開発): 治験申請に必要な厳格なフォーマットの文書

の自動生成、疾患データや業界規制・社内標準作業手順書(SOP)の自動検索、そして治験計画に規定されていない安全性シグナルを検出するための「探索的データ解析」を行うAIエージェントを開発する。

2. フェーズ2(プロセスの全域拡張): エージェントの適用範囲を、臨床開発プロセスの特定タスクから全体のワークフローへと拡大する。
3. フェーズ3(マルチエージェント協調システムの構築): 単一のAIではなく、専門機能を持った複数のAIエージェントが相互に連携し、人間の専門家と協調しながら業務を実行する「マルチエージェントシステム」を構築する。

この事例は、生成AIの活用が能動的なエージェント・フェーズへと進化していることを示している<sup>1</sup>。

## 5. 知財業務 (Intellectual Property Operations) への適用可能性の検討

これまでに詳述したSarashinaの技術的特性(フルスクラッチ開発による日本語処理、機密性、マルチモーダル対応、エージェント連携の概念)を俯瞰し、知財業務への活用が可能かどうかを検討する。

知財業務は「機密保持」「特許法に基づく論理と特殊な日本語用語の解釈」「数千万件に及ぶ先行技術文献の処理」、そして「図面等を含むマルチモーダルな情報の理解」が要求される領域である。Sarashinaがこれらの課題に対していかにアプローチし得るか、具体的な知財ワークフローに沿って詳述する。

### 5.1 【機密性の確保】発明提案書および営業秘密の漏洩リスクの大幅な低減

知財業務において致命的なインシデントは、未公開の特許情報や発明提案書(営業秘密)が外部に漏洩することである。出願前の発明情報が第三者に知得された場合、新規性が喪失し、特許を取得する権利を失うリスクがあるため、パブリックな海外製生成AIサービスに発明のコア技術を入力することは多くの企業で禁止されている。

**Sarashinaによる解決策:** ソフトバンクの「Cloud PF Type A」上で稼働するSarashinaを利用することで、データ処理プロセスが国内のデータセンター内で完結する<sup>1</sup>。さらに「API経由で入力されたデータはモデルの学習には使用しない」というポリシーが適用されている<sup>1</sup>。これにより、研究開発部門や知財担当者は、データ漏洩のセキュリティリスクを管理しながら、未公開のアーキテクチャなどを入力してAIによる支援を受ける環境を構築しやすくなる。

### 5.2 【日本語の言語特性への適応】特許文書ドラフト作成と実務への適用

特許明細書、とりわけ権利範囲を画定する「特許請求の範囲(クレーム)」は、独自の論理構造と厳密な法的・技術的語彙で構成される。「前記」「～からなる群から選択される」といった用語は日常的な日本語とは異なり、汎用LLMでは特許庁(JPO)の実務に基づく解釈を正確に捉えきれないケースがある。

**Sarashinaによる解決策:**

Sarashinaの日本語処理能力を基盤に、JPO審査基準、過去の意見書・補正書、特許公報、社内出願実績をRAGや追加評価データとして組み合わせることで、知財業務に適した支援システムを構築できる可能性がある。

- 発明提案から明細書の自動起案: 発明提案書から明細書ドラフトを作成する補助に活用できる余地がある。ただし、権利範囲、サポート要件、明確性、禁反言リスクへの対応について

は、弁理士・知財担当者によるレビューを前提とする。

- 特許翻訳の一次案作成: 原文の意図や背景を考慮した翻訳機能<sup>1</sup>を活用し、パテントファミリー等向けの特許翻訳の一次案作成に利用できる。これも各国実務への適合性については専門家による確認が必要である。

### 5.3 【RAGの活用】ベクトル検索による先行技術調査・クリアランス調査の効率化

知財業務において多大な人的リソースを消費するのが、自社発明の特許性を確認する「先行技術調査」や、他社の特許権を侵害していないかを確認する「クリアランス調査(FTO調査)」である。従来の手法では、複雑な論理検索式を作成し、膨大な公報を人間が目視でスクリーニングしていた。

**Sarashina**による解決策: Sarashinaのエンベディングモデル「Sarashina-Embedding-v2-1B」を適用し、特許公報群をベクトルデータ化することで、高度なRAGシステムを構築できる<sup>1</sup>。

- 意味的類似度に基づく検索: キーワードの完全一致に依存せず、「技術の課題解決原理」や「構造の機能的類似性」に基づいた検索が可能になり、従来のキーワード検索では漏れていた特許を抽出できる可能性がある。
- クレーム対比表の生成補助: 抽出した関連特許公報に対して、自社の発明提案内容と引例の構成要件ごとの「一致点」および「相違点」を抽出し、要約されたクレーム対比表のドラフトを生成する。これにより、スクリーニングおよび判断にかかる時間の削減が期待でき、PoC等で検証すべき事項である<sup>1</sup>。

### 5.4 【マルチモーダル機能の適用】図面等を含む特許情報へのアプローチ

特許情報はテキストデータだけで完結せず、「図面」や「ブロック図」が技術を理解する上で重要となる。

**Sarashina**による解決策: SB Intuitionsが展開するエンドツーエンドのOCRモデル「Sarashina2.2-OCR」や、画像理解エンジン「Sarashina2.2-Vision-3B」を知財分析のパイプラインに組み込むことで、視覚情報の解析が期待できる<sup>1</sup>。

- 図面と明細書のクロスリファレンス解析: 特許の図面(PDF画像など)をSarashina2-Visionに入力し、視覚情報とテキスト情報を統合的に解釈し、技術的構造を解説する補助が可能になる<sup>1</sup>。
- 非定型フォーマットからのテキスト抽出: 過去の手書きノートやスキャンされた文献をSarashina2.2-OCRでテキスト化・構造化し、社内データベースへ統合することができる<sup>10</sup>。
- (※なお、化学構造式やDNA配列などは一般のOCR/VLMだけで正確に扱うにはリスクが高いため、化学構造認識、分子式変換、配列照合などの専用ツールと組み合わせる前提でのシステム構築が望ましい。)

### 5.5 【知財AIエージェントの展望】中外製薬モデルの応用

中外製薬との共同研究で示された「自律型AIエージェント」の概念<sup>13</sup>は、知財業務のデジタルトランスフォーメーション(DX)にも応用し得る。知財プロセスは、厳密なガイドラインの遵守、情報の収集、審査官への論理的な文書作成といったステップを持つためである。

臨床開発AIエージェント(中外製薬モデル)の	知財業務におけるAIエージェントへの応用シ
------------------------	-----------------------

役割	ナリオ(期待される用途)
文書作成エージェント 治験申請に必要な関連文書群の自動生成。	明細書起案エージェント 発明提案書から、特許請求の範囲や明細書ドラフトを自動起案(人間によるレビュー前提)。
情報収集エージェント 疾患データ、法規制、社内手順書の検索。	先行技術調査エージェント 特許DBへクエリを発行し、関連度の高い引例を抽出・スコアリングする。
探索的データ解析エージェント 安全性シグナルの検出や治験の妥当性検証。	不適切コンテンツ検知・拒絶理由分析補助 Sarashina2 guardは不適切・有害コンテンツ検知に利用できる。拒絶理由通知書の論点整理などへの応用が考えられる。一方、営業秘密や未公開発明情報の混入検知には、社内DLP、機密語辞書、出願前情報データベース、専用分類モデルを組み合わせる必要がある。

## 6. 結論

SarashinaおよびCloud PF Type AIは、日本語処理、国内運用、RAG、マルチモーダル処理を組み合わせることで、知財業務への活用可能性が高い。特に、発明提案書の整理、先行技術調査の初期スクリーニング、社内特許情報検索、明細書ドラフト作成補助、拒絶理由通知の論点整理などでは有効性が期待できる。

ただし、特許請求の範囲の設計、権利解釈、FTO判断、出願可否判断、補正・意見書の最終判断は法的・技術的責任を伴うため、AIの出力をそのまま採用するのではなく、弁理士・知財担当者によるレビューを前提とした支援ツールとして位置づけるべきである。

導入にあたっては、特許公報・審査基準・社内出願書類を用いたRAG環境を構築し、検索再現率、クレーム対比精度、ハルシネーション率、機密情報保護、監査ログ、利用権限管理をPoCで検証することが望ましい。

### 引用文献

1. Sarashina API | 法人のお客さま - ソフトバンク, 6月 7, 2026にアクセス、  
<https://www.softbank.jp/business/service/ai/sarashina-api/>
2. ソフトバンクのクラウドサービス「Cloud PF Type A」を基盤にしたノーコードプラットフォームを - システナ, 6月 7, 2026にアクセス、  
<https://www.systema.co.jp/pdf/irnews/20260401.pdf>

3. SB Intuitions、独自の日本語LLMを構築 アカデミアや産業界の研究 ..., 6月 7, 2026にアクセス、[https://www.sbintuitions.co.jp/news/press/20240614\\_01/](https://www.sbintuitions.co.jp/news/press/20240614_01/)
4. SB Intuitions, Building Unique Japanese LLM Released Japanese LLM with 7 Billion, 13 Billion, and 65 Billion Parameters to Contribute to R&D in Academia and Industry | PRESS RELEASE | SB Intuitions Corp., 6月 7, 2026にアクセス、[https://www.sbintuitions.co.jp/en/news/press/20240614\\_01/](https://www.sbintuitions.co.jp/en/news/press/20240614_01/)
5. RESEARCH | SB Intuitions Corp., 6月 7, 2026にアクセス、<https://www.sbintuitions.co.jp/en/sarashinalab/>
6. Run sarashina2-70b API | Serverless Inference | 8K Context | Flat-Rate Pricing - Featherless.ai, 6月 7, 2026にアクセス、<https://featherless.ai/models/sbintuitions/sarashina2-70b>
7. sarashina2-8x70b | PromptLayer Models, 6月 7, 2026にアクセス、<https://www.promptlayer.com/models/sarashina2-8x70b/>
8. sbintuitions/sarashina2-8x70b - Hugging Face, 6月 7, 2026にアクセス、<https://huggingface.co/sbintuitions/sarashina2-8x70b>
9. [Paper Note] Sparse Upcycling: Training Mixture-of-Experts from Dense Checkpoints, Aran Komatsuzaki+, ICLR'23 #1546 - GitHub, 6月 7, 2026にアクセス、[https://github.com/AkihikoWatanabe/paper\\_notes/issues/1546](https://github.com/AkihikoWatanabe/paper_notes/issues/1546)
10. Sarashina2.2 Ocr - Interfaze, 6月 7, 2026にアクセス、<https://interfaze.ai/models/sbintuitionssarashina22-ocr>
11. sbintuitions/sarashina2.2-tts - GitHub, 6月 7, 2026にアクセス、<https://github.com/sbintuitions/sarashina2.2-tts>
12. "Sarashina" Utilized to Support Database Expansion for "G-Biz Portal"| PRESS RELEASE | SB Intuitions Corp., 6月 7, 2026にアクセス、[https://www.sbintuitions.co.jp/en/news/press/20260325\\_01/](https://www.sbintuitions.co.jp/en/news/press/20260325_01/)
13. 2025年01月30日 | 中外製薬、ソフトバンク、SB Intuitionsの3社が ..., 6月 7, 2026にアクセス、[https://www.chugai-pharm.co.jp/news/detail/20250130153000\\_1461.html](https://www.chugai-pharm.co.jp/news/detail/20250130153000_1461.html)
14. Cloud PF Type A - ソブリン性を備えたクラウドサービス | 法人向け - ソフトバンク, 6月 7, 2026にアクセス、<https://www.softbank.jp/business/service/platform/cloud-pf-type-a/>