

Cierpa & Co. (シェルパ) は、自然言語処理を用いてサステナビリティ情報を可視化するSaaSを開発・提供してきました。現在はその基盤技術を拡張し、多様なビジネス文書から情報を抽出し構造化する新たなVLM開発にも取り組んでいます。

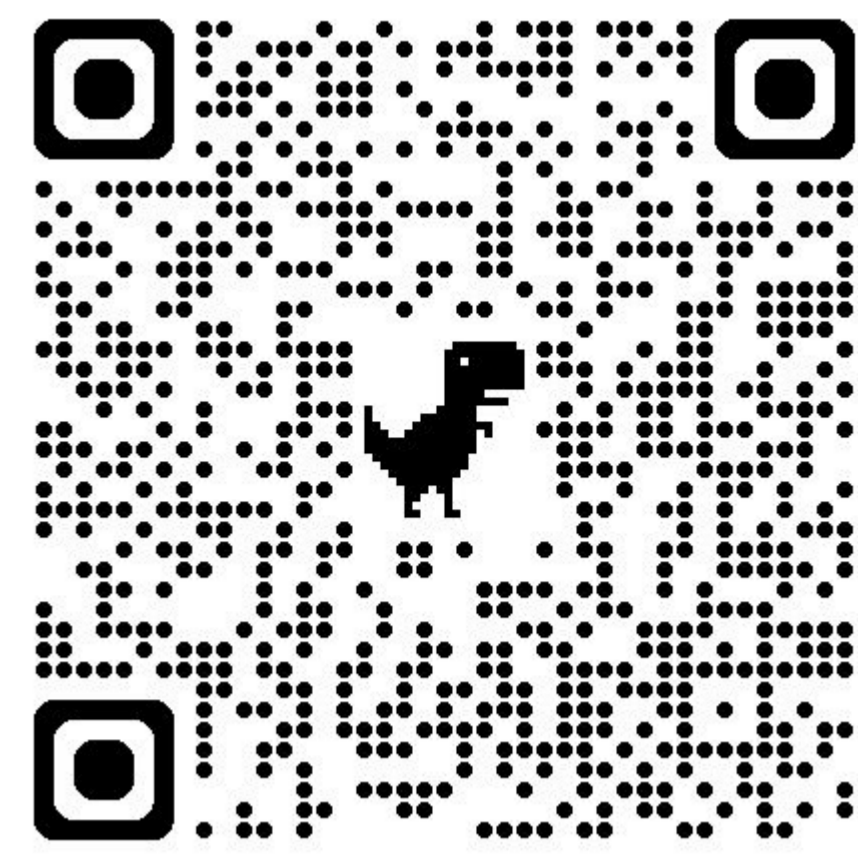
特定のドメインに縛られない多様なドキュメント解析と、その社会実装に興味のある研究者・エンジニアをお待ちしています。

エンジニア・インターン積極採用中

こんな方を求めています！

- NLPやマルチモーダルに深い関心がある方
- 最先端の論文を読み解き、新しいアイデアを積極的に提案できる方
- 社会課題の解決に直結する文書解析や情報抽出の難題に取り組みたい方

採用ページはこちら



弊社プロダクト SmartESG

ESG情報開示支援クラウド「SmartESG」にNLP関連技術を組み合わせて、組織内外に蓄積された非財務情報の収集と分析を実現。ESG経営を通じた企業の価値創出を支援しています。

Environment
Society
Governance

1 企業のESG情報を自動収集

企業の公開情報を自動的にデータベースに蓄積



- 有価証券報告書
- 企業サイト
- 統合報告書
- ESGデータブック
- CG報告書 など

2 情報抽出・分析

NLP関連技術を活用し、PDFやWebページに記載された情報を抽出・分析



3 レポートの作成

評価機関から送付されるレポートを取り込み、競合や業界平均とスコアを比較できるほか、改善すべき点や重要項目の特定を容易に



非定型なレイアウトの統合報告書の文書理解タスク

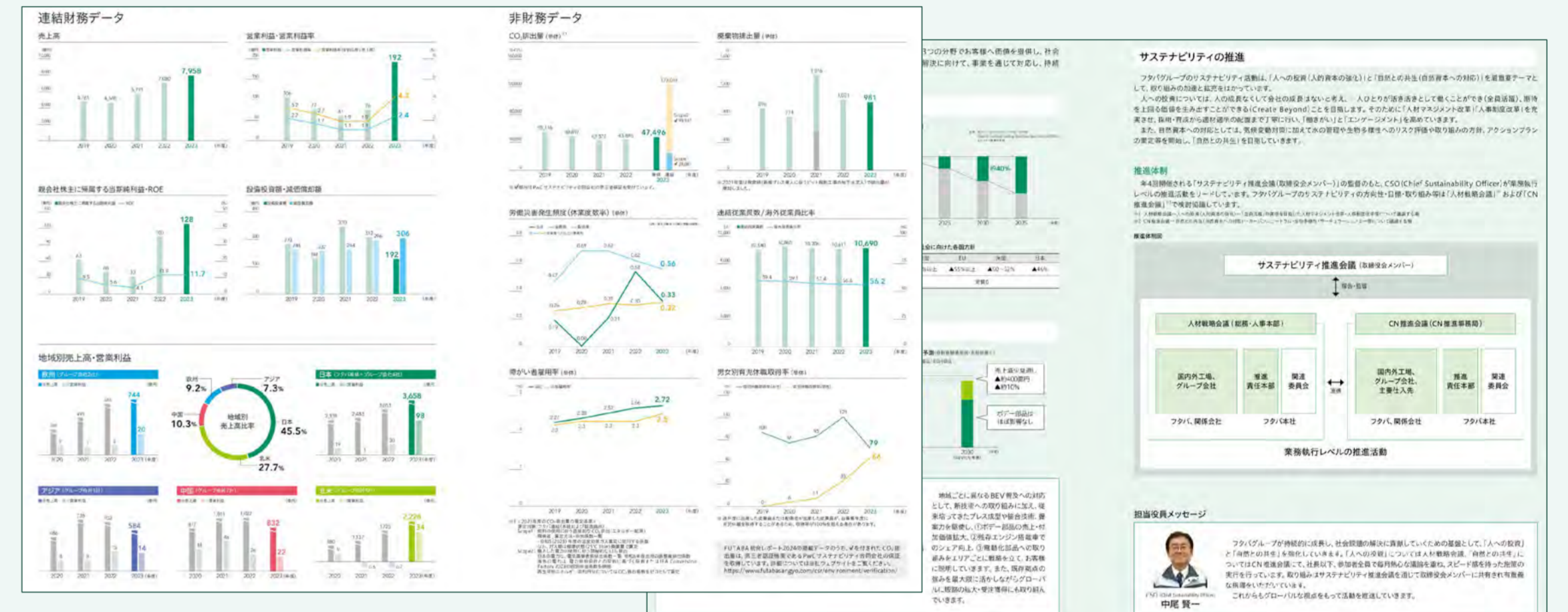
従業員の労働環境は大丈夫？

取締役はどのように選出してる？

温室効果ガス排出量は？

下請け企業に目を向けてる？

児童労働・強制労働を防止してる？



[FUTABA統合レポート2024より]

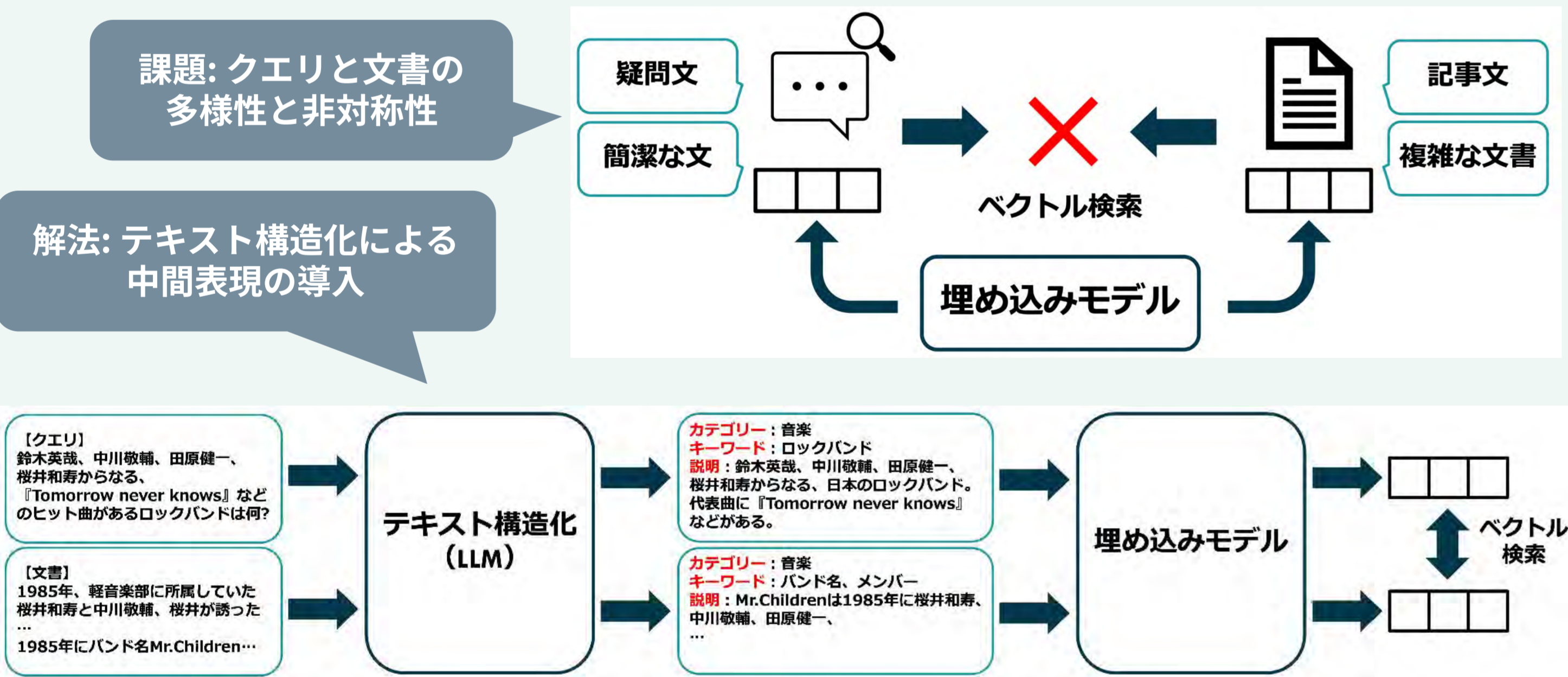
研究活動

NLP2025参加記

ベクトル検索におけるテキスト構造化の効果分析

梶川 怜恩^{1,2} 神田 峻介¹ 赤部 晃一¹ 小田 悠介^{1,3} [NLP2025]

¹⁾シェルパ・アンド・カンパニー株式会社 ²⁾愛媛大学大学院理工学研究科 ³⁾国立情報学研究所 大規模言語モデル研究開発センター



ESGデータセット構築における

Human-in-the-Loop全自動パイプラインの構築

赤部 晃一¹ 櫻田 幸子¹ 小田 悠介^{1,2} [YANS2025]

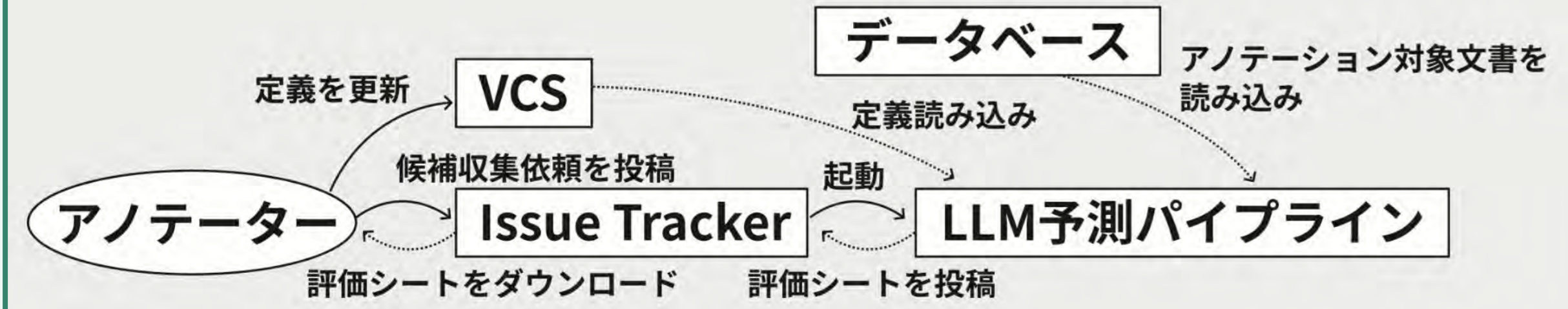
¹⁾シェルパ・アンド・カンパニー株式会社 ²⁾国立情報学研究所 大規模言語モデル研究開発センター

正例の割合が約1%と非常にimbalancedな二値分類データセット構築のための効率的な方法の提案。自動パイプラインを構築することで技術者の関与なしに構築可能。

関連ブログ



Phase 1: LLMによる予測結果の検証によるデータ作成



Phase 2: データベースの各事例に対して交差検証で予測結果を算出し、確信度が高い事例から順にアノテーション(非該当例が15件続いたら作業終了)



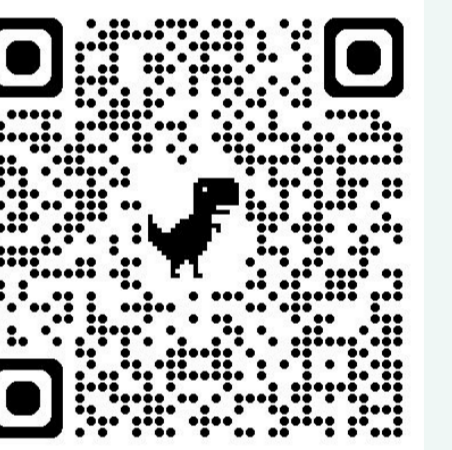
ESG-QA: 日本語ESG文書を対象とした

NLP2026投稿論文

マルチモーダル質問応答ベンチマークの構築

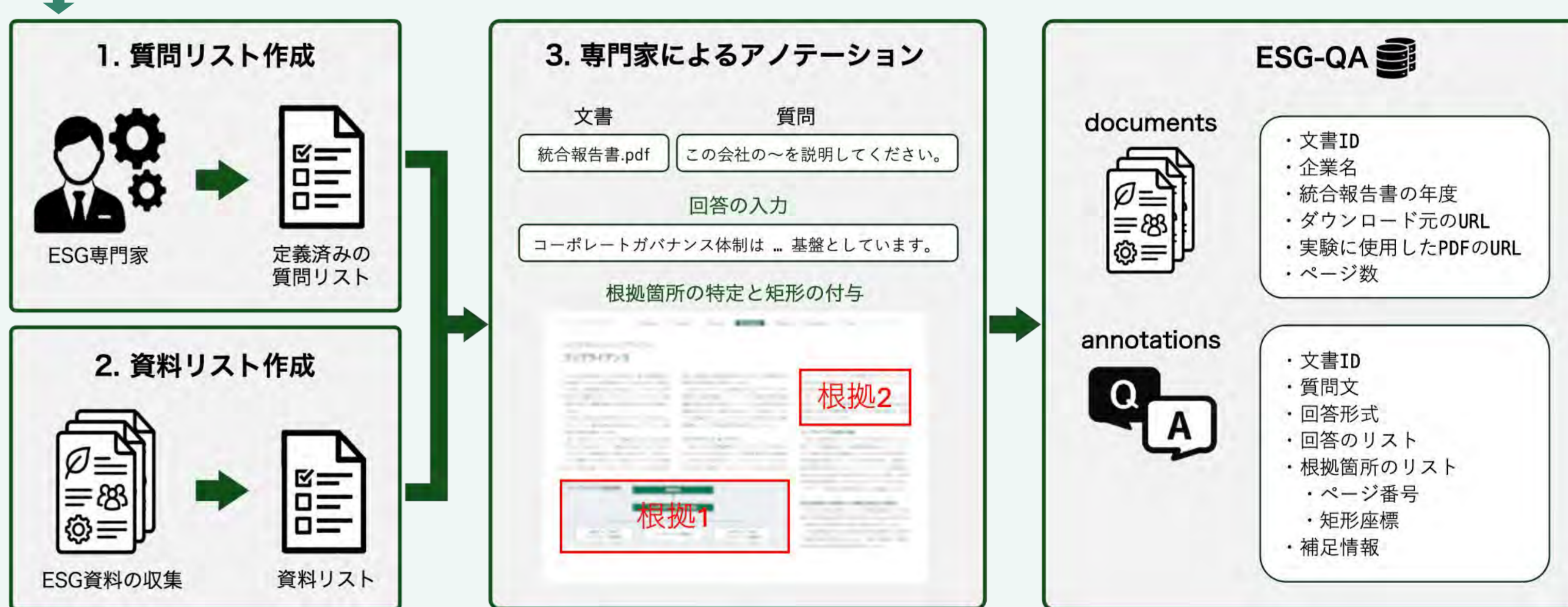
佐多 亮明^{1,2} 赤部 晃一¹ 神田 峻介¹ 小田 悠介^{1,3} [NLP2026 Q4-10]

¹⁾シェルパ・アンド・カンパニー株式会社 ²⁾筑波大学大学院 システム情報工学研究群 ³⁾国立情報学研究所 大規模言語モデル研究開発センター



ESGenius (He+, '25) やMMESGBench (Zhang+, '25) などの既存ベンチマークの作成方法は実務に則していない。実務では質問がある上で文書をもとに回答するが、既存ベンチマークでは文書をもとに質問と回答を作成している。

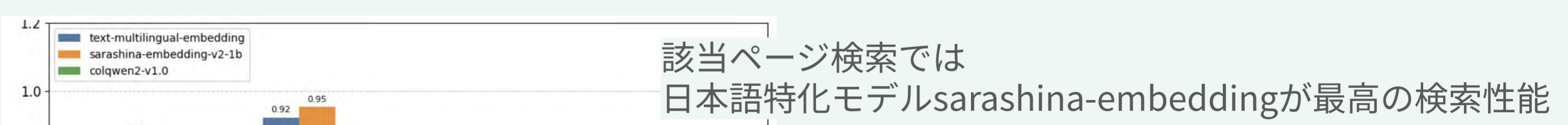
ESG-QAでは文書に依らず**質問起点でデータを作成した。**



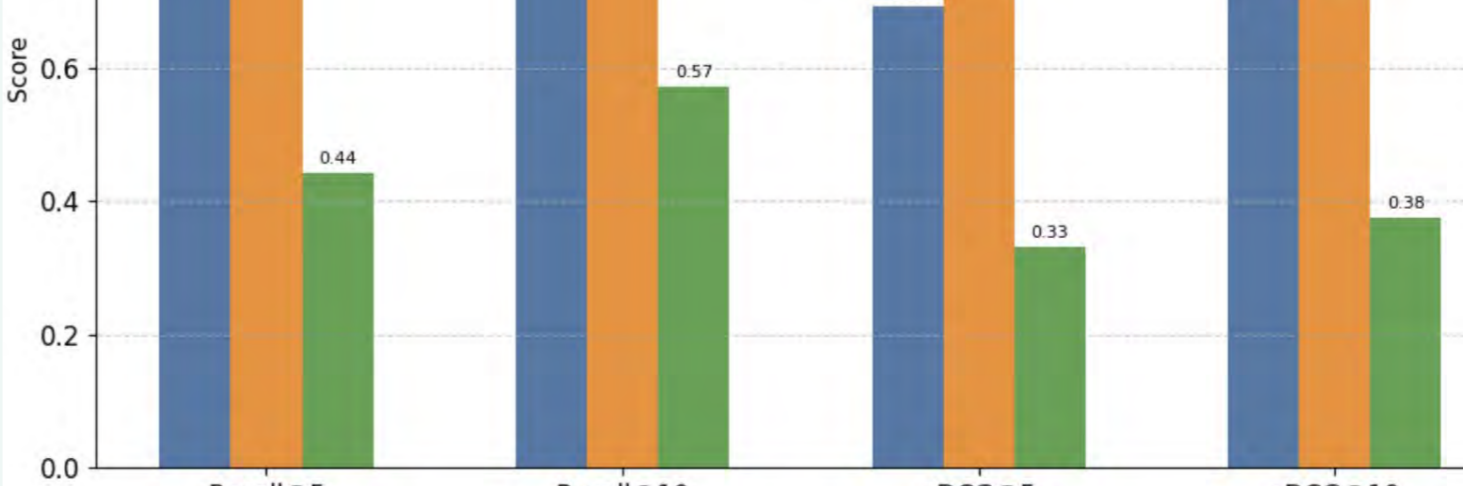
• 質問は実務上重要な質問20件。Yes/No, 事実抽出, 数量, 自由記述の4形式。

• 28社分の統合報告書 (1,522ページ) , 560事例のQAペア

• 該当ページ検索, 根拠箇所からの回答生成, 根拠箇所の領域検出の3タスク評価可能。



【追加検証】 回答不能な事例を正しく「根拠なし」と識別し、幻覚を抑制する能力を検証

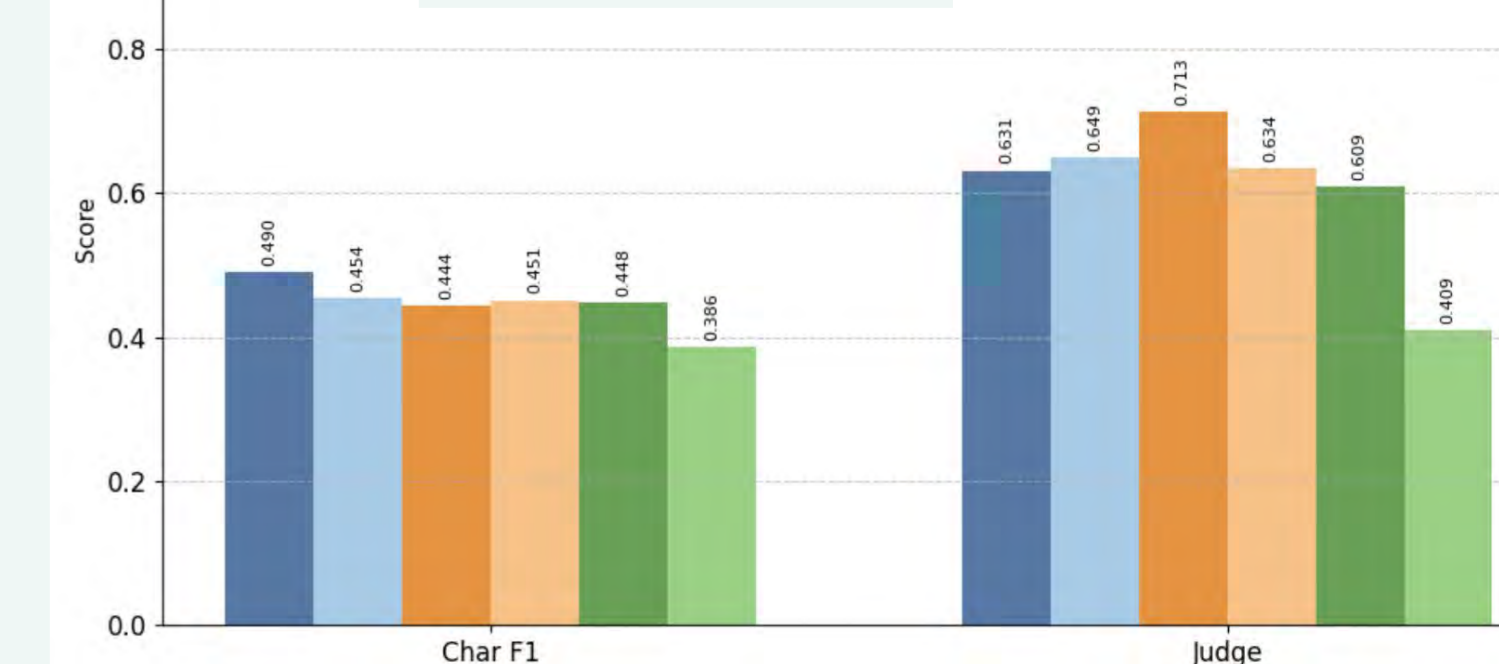


根拠箇所からの回答生成ではgemini-2.5-flash (Text), gpt-5-mini (Image), gpt-5-mini (Text), llm-gp-3.1.8k1.8k (Text) 高い性能を示した

	予測ページ有	予測ページ無
根拠ページ有	245	34
根拠ページ無	196	85

→ 約50% (281/560件) が回答不能な事例

根拠箇所の領域検出ではレイアウト解析モデル+VLMを用いてIoUスコア0.297を示した。



根拠箇所の領域検出ではレイアウト解析モデル+VLMを用いてIoUスコア0.297を示した。