



【Python】AI学習用コーパス（テキストデータ）の自動検品・検証プログラム開発

完全在宅（フリーランス）でのプログラミング業務

募集職種

採用企業名

株式会社フランシール

求人ID

1609982

業種

ソフトウェア

会社の種類

中小企業 (従業員300名以下)

外国人の割合

外国人 少数

雇用形態

フリーランス

勤務地

東京都 23区, 豊島区

最寄駅

山手線駅

給与

経験考慮の上、応相談

時給

出来高払い

更新日

2026年09月01日 12:34

応募必要条件

職務経験

1年以上

キャリアレベル

中途経験者レベル

英語レベル

ビジネス会話レベル (英語使用比率: 50%程度)

日本語レベル

流暢

最終学歴

大学卒：学士号

現在のビザ

日本での就労許可が必要です

募集要項

AI（大規模言語モデルなど）の学習に使用する「コーパスデータ（主にJSON / CSV / 正解ラベル付きテキスト）」の品質を自動でチェックするスクリプト・プログラムの開発をお願いします。

【具体的なチェック項目】

- データのフォーマット・構造チェック（ネストの崩れ、文字コード、空データの検出）
 - テキストの基本バリデーション（禁止文字の混入、文字数制限の超過など）
 - アノテーション（タグ付け）の整合性チェック（開始タグと終了タグの不一致など）
 - ※要件定義から一緒に相談しながら進めさせていただける方を歓迎します。
-

スキル・資格**【必須スキル・経験（Must）】**

- **Python**を用いたデータ処理、テキスト処理のの実務経験（2〜3年以上）
- **正規表現（RegEx）**を用いた高度な文字列パターンの抽出・照合スキル
- **pandas**などのデータ解析ライブラリの使用経験
- Git / GitHub を用いたソースコード管理

【歓迎スキル・経験（Want）】

- **自然言語処理（NLP）**に関する知見、ライブラリ（GiNZA, Spacy, MeCabなど）の使用経験
- AIコーパス、アノテーションデータの取り扱い経験
- 大量データ（数万〜数百万件）を高速に処理するプログラムの最適化経験

【求める人物像】

- 単なる指示通りではなく、データの不備を見抜くための仕様と一緒に考えてくれる方
 - 細かい条件分岐や例外処理をロジカルに組み立てられる方
 - 報告・連絡・相談がスムーズで、納期を守れる方
-

会社説明