

LIQUID AI & W&B TOKYO HACKATHON 2025

REDACTOR

A PII REDACTION PIPELINE

VALERIO · PRZEMEK · GANGA

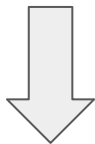
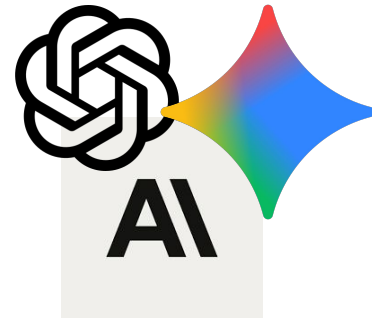
Remove PII from Japanese contracts before LLMs

労働条件通知書兼雇用契約書（正社員 2023）

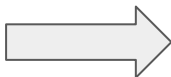
氏名（フルネーム） 竹内樹
会社名 オーシャン・キャピタル・パートナーズ
会社所在地 〒 100-6990 東京都千代田区丸の内 2-6-1 丸の内ブリックスクエア
電話番号 03-3211-8801

契約期間 正社員（期間の定めなし）
試用期間：令和 7 年 11 月 1 日～令和 8 年 1 月 31 日（最長 1 か月延長の可能性あり）
就業の場所（入社時）：本社（上記所在地）／変更の範囲：各拠点・プロジェクト先
業務の内容（入社時）：産業用ロボットの組立・検査・工程改善補助／変更の範囲：会社の定める関連業務

役職：アソシエイトエンジニア
勤務時間・休憩：10:00～19:00（休憩 60 分）／フレックス（コア 11:00～16:00）
休日・休暇：土曜・日曜・祝日・年末年始（12/29～1/3）年間 120 日／年休：法定（入社 6 か月後 10 日）／特別休暇（慶弔 3 日）



REDACTOR
A PII REDACTION PIPELINE

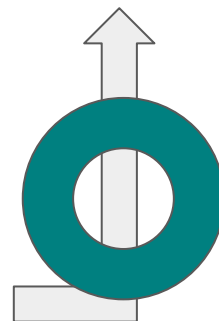


労働条件通知書兼雇用契約書（正社員 2023）

氏名（フルネーム） [REDACTED]
会社名 [REDACTED]
会社所在地 [REDACTED]
電話番号 [REDACTED]

契約期間 正社員（期間の定めなし）
試用期間：令和 7 年 11 月 1 日～令和 8 年 1 月 31 日（最長 1 か月延長の可能性あり）
就業の場所（入社時）：本社（上記所在地）／変更の範囲：各拠点・プロジェクト先
業務の内容（入社時）：産業用ロボットの組立・検査・工程改善補助／変更の範囲：会社の定める関連業務

役職：アソシエイトエンジニア
勤務時間・休憩：10:00～19:00（休憩 60 分）／フレックス（コア 11:00～16:00）
休日・休暇：土曜・日曜・祝日・年末年始（12/29～1/3）年間 120 日／年休：法定（入社 6 か月後 10 日）／特別休暇（慶弔 3 日）

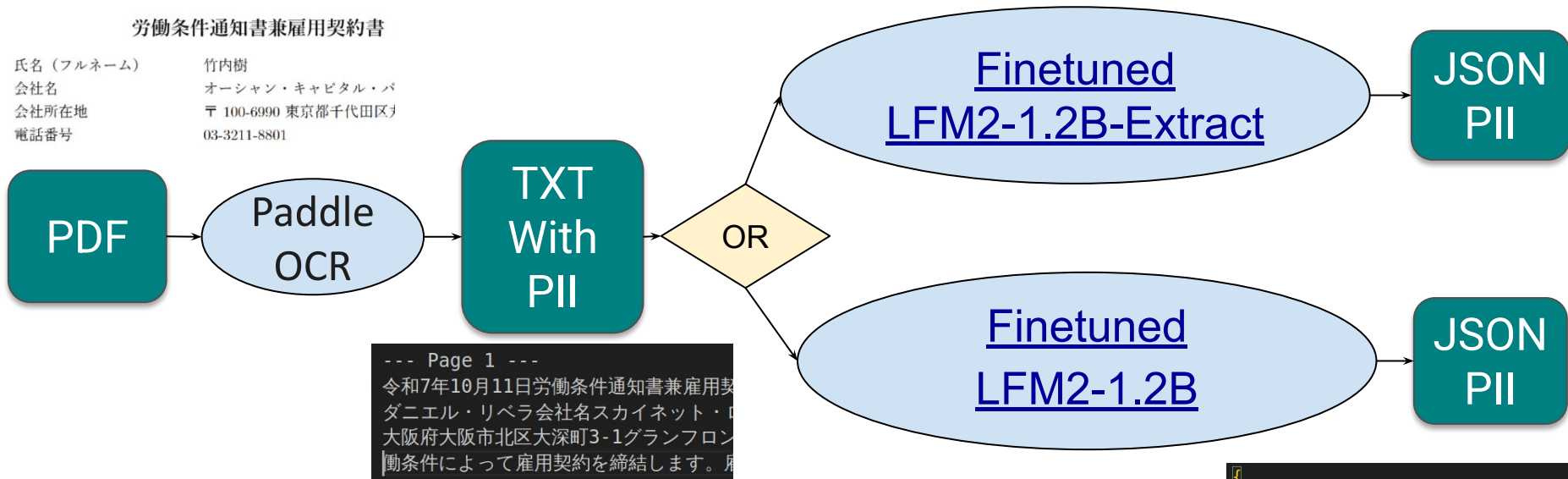


Secure , Onprem, content preserved without PII

How Redactor Works

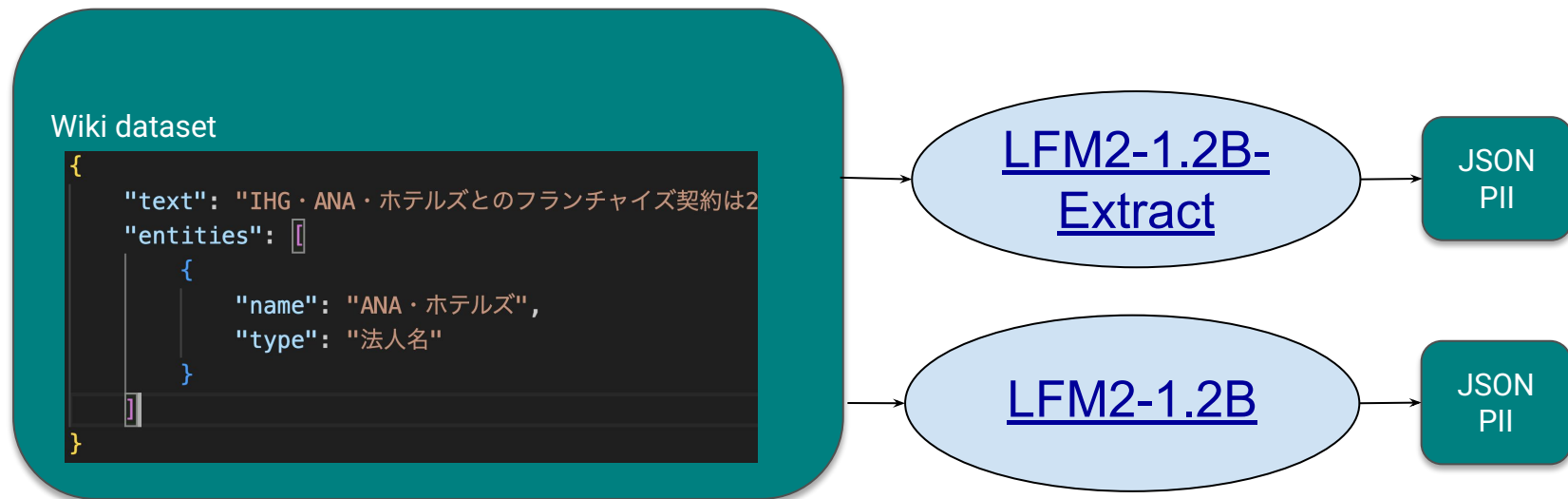
労働条件通知書兼雇用契約書

氏名（フルネーム）	竹内樹
会社名	オーシャン・キャピタル・パ
会社所在地	〒 100-6990 東京都千代田区
電話番号	03-3211-8801



```
{  
  "full_name": "ソフィア・ミラー",  
  "company_name": "株式会社オリオンシス",  
  "address": "東京都千代田区大手町1-9-2",  
  "phone_number": "03-6262-1212"  
}
```

How we Built Redactor



Dataset completely different
than test dataset

SFT+LORA fine tuning

Results

	LiquidAI-LFM2-1.2B		LiquidAI-LFM2-1.2B-Extract	
	Before Finetune	After Finetune	Before Finetune	After Finetune
Wiki Dataset	0 . 4 3		0 . 9 1	
Our 64 Generated Contracts	0 . 0 3		0 . 5 8	

(exact match)

Hugging Face Checkpoints

- [kainoj/LiquidAI-LFM2-1.2B-ja-pii-finetuned](#) ([wandb](#))
- [kainoj/LiquidAI-LFM2-1.2B-Extract-ja-pii-finetuned](#) ([wandb](#))
- Source code: [valeriosalvucci/redact](#)



Przemyslaw Joniak
[@kainoj](#)

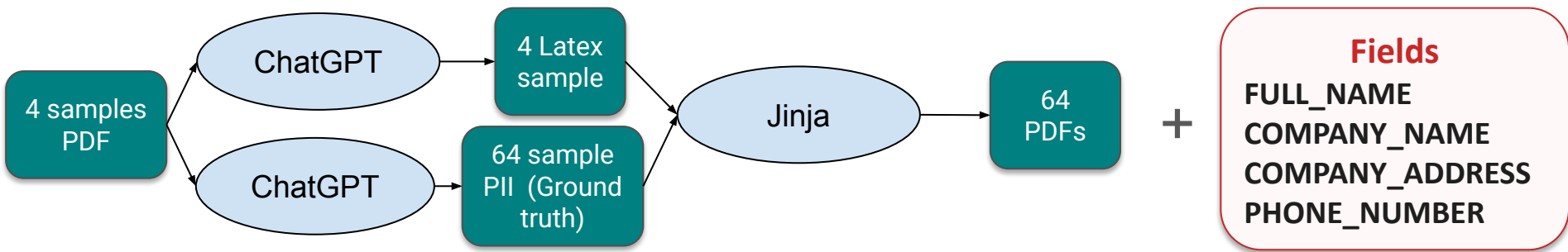


Valerio Salvucci
[@valeriosalvucci](#)



Ganga
[@gangadhara691](#)

Dataset Preparation



参考様式第 1－6 号

雇 用 条 件 書

_____ 殿	年 月 日
特定技能所属機関名 _____	
所在地 _____	
電話番号 _____	
代表者 役職・氏名 _____	印

I. 雇用契約期間

令和 7 年 10 月 11 日

労働条件通知書兼雇用契約書（シンプル）

氏名（フルネーム）	ソフィア・ミラー
会社名	株式会社オリオンシステムズ
会社所在地	〒 100-0004 東京都千代田区大手町 1-9-2 大手町フィナンシャルシティ
電話番号	03-6262-1212

次の労働条件によって雇用契約を締結します。

雇用期間	期間の定めあり（令和 7 年 11 月 1 日～令和 8 年 10 月 31 日）
更新の有無	更新する場合があります（基準：業務量／成績・態度／能力／経営状況／進捗）
試用期間	令和 7 年 11 月 1 日～令和 8 年 1 月 31 日
就業の場所	本社（上記所在地）
仕事の内容	産業機械の組立補助・検査・付随事務