

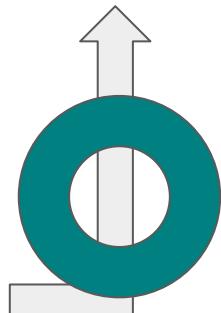
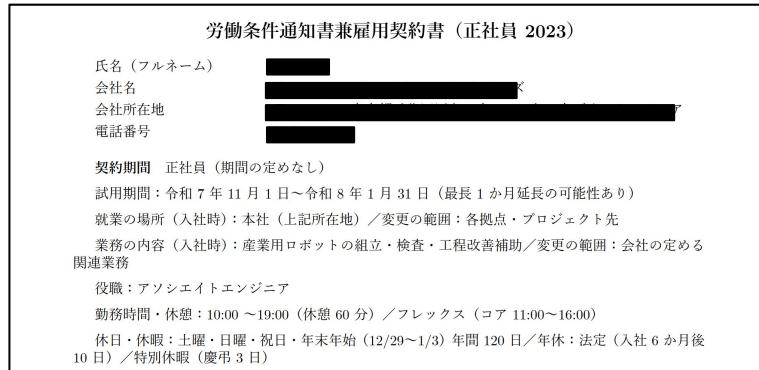
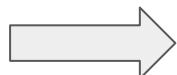
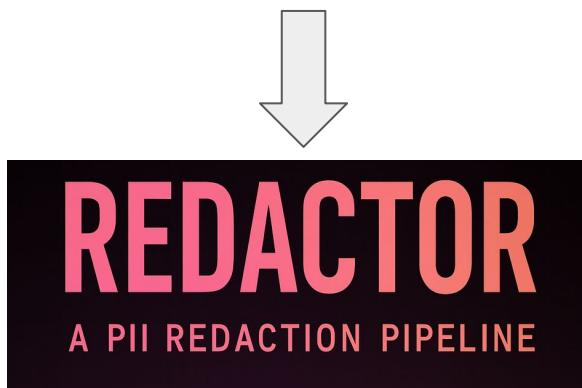
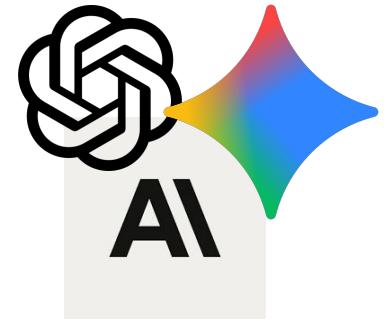
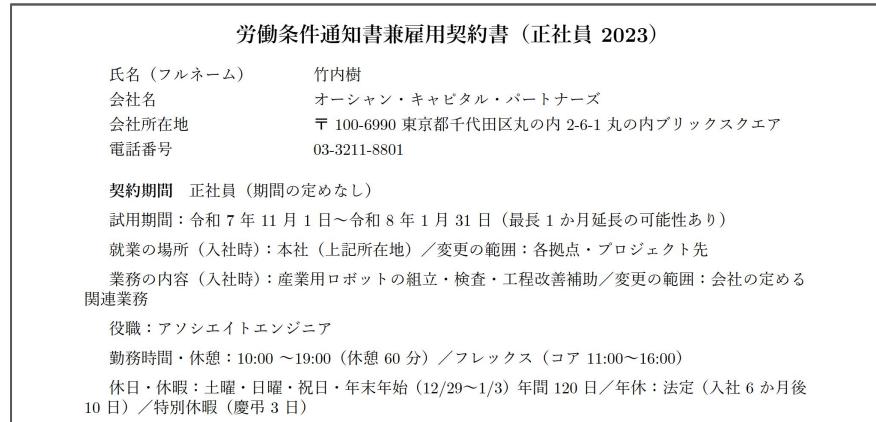
LIQUID AI & W&B TOKYO HACKATHON 2025

REDACTOR

A PII REDACTION PIPELINE

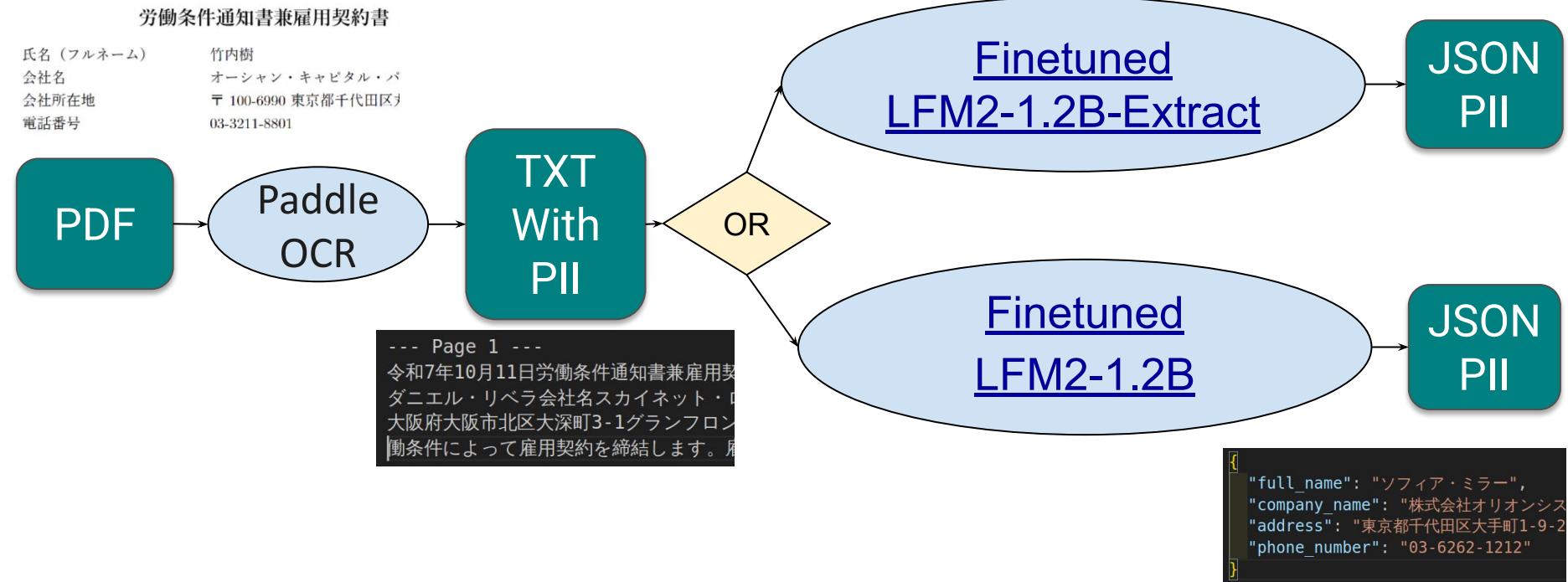
VALERIO · PRZEMEK · GANGA

Remove PII from Japanese contracts before LLMs

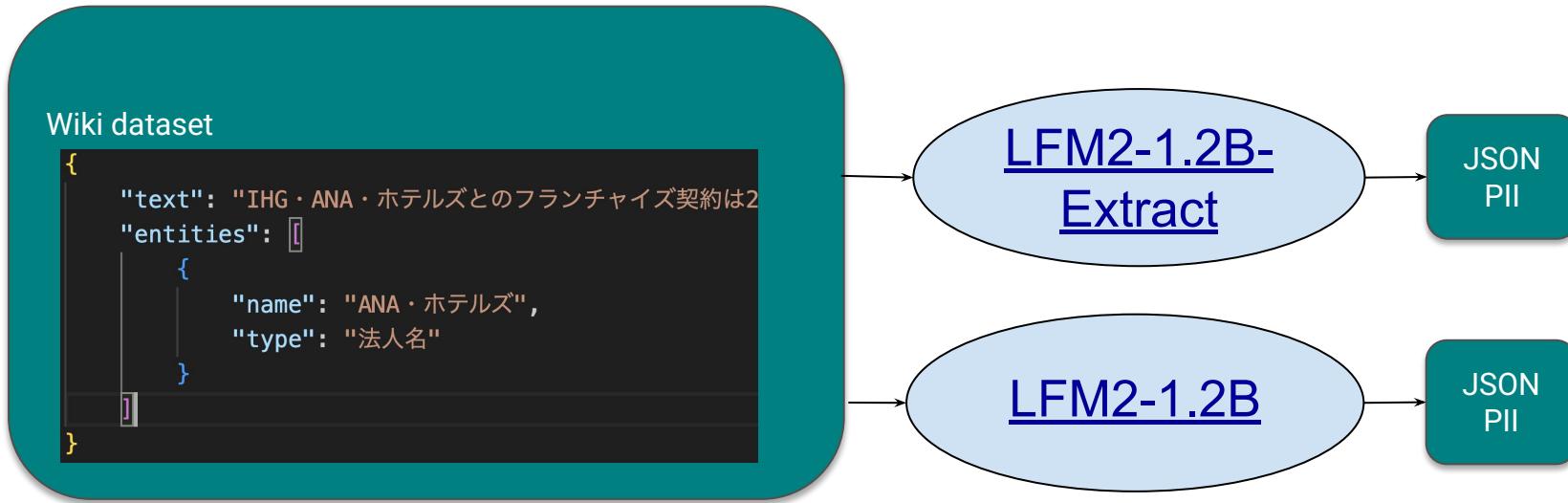


Secure , Onprem, content preserved without PII

How Redactor Works



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
<u>Wiki Dataset</u>	0 . 43		0 . 91	
Our 64 Generated Contracts	0 . 03		0 . 58	

(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](https://twitter.com/kainoj)

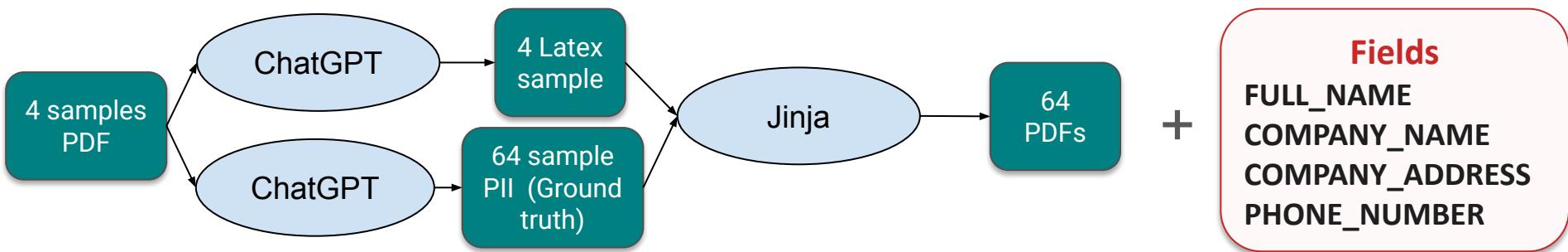


Valerio Salvucci
[@valeriosalvucci](https://twitter.com/valeriosalvucci)



Ganga
[@gangadhara691](https://twitter.com/gangadhara691)

Dataset Preparation



参考様式第1～6号

雇用条件書

年月日

殿

特定技能専属機関名 _____
所在地 _____
電話番号 _____
代表者 役職・氏名 _____

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 大手町ファニシャルシティ
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得る（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