

CASE STUDY

HEALTHCARE INDUSTRY

AI-enabled platform ensuring privacy-compliant health data use

Project: Datarwe NLP De-identification & Classification Initiative



KEY OUTCOME 1

Improved healthcare insights through AI models



KEY OUTCOME 2

Automated analysis of complex clinical text



KEY OUTCOME 3

Enhanced data accuracy and decision support

BACKGROUND

Datarwe, in partnership with Queensland Health, sought to enhance their ICU data platform by ensuring patient data could be de-identified, securely shared and effectively classified to support both operational clinical use and secondary applications such as research and clinical trials.

A major component of this initiative involved working with unstructured clinical text — progress notes, radiology reports, and other narrative documentation — to unlock insights while ensuring strict compliance with privacy regulations (e.g., HIPAA, local Public Health Acts).

To facilitate secondary uses of this data, it was essential to accurately remove personally identifiable information (PII) from free-text content. Achieving this at scale, reliably, and in a privacy-preserving manner required sophisticated natural language processing (NLP) and robust data governance.

WHY KJR WAS ENGAGED

KJR as Datarwe joint venture partner brought key experience to this task due to its unique dual expertise in:

- Artificial Intelligence and NLP engineering, with experience in real-world healthcare AI applications.
- Data assurance, privacy, and governance, ensuring sensitive data handling aligns with both legal standards and ethical best practices.

The combination of deep technical skill and a robust assurance framework made KJR an ideal partner to build trusted AI systems in sensitive healthcare settings.

CHALLENGE

The central challenge lay in reliably identifying and redacting personally identifiable information (PII) like names, locations, and dates, embedded in free-text medical records. Off-the-shelf tools like AWS Medical Comprehend struggled with the nuances of Australian place names, medical terminology, and localised linguistic patterns, resulting in unacceptable de-identification accuracy (~80%).

Furthermore, the project needed to manage data governance risks while enabling machine learning and generative AI to support clinical and research use cases.

Key technical and operational hurdles included:

- Recognising and handling diverse PII types within natural language.
- Adapting models to regional data idiosyncrasies.
- Ensuring human-in-the-loop processes preserved patient confidentiality.
- Enabling downstream AI use, such as disease phenotype classification and billing code summarisation, without re-identification risks.

CASE STUDY

HEALTHCARE INDUSTRY

AI-enabled platform ensuring privacy-compliant health data use

Project: Datarwe NLP De-identification & Classification Initiative

SOLUTION

KJR applied its Validation Driven Machine Learning (VDML) approach to build a tailored NLP pipeline that balanced accuracy, privacy, and scalability:

1. Baseline Assessment: Evaluated AWS Comprehend against local ICU data, revealing gaps in handling regional terms and clinical nuances.
2. Model Fine-tuning: Customised a BERT-based model using manually labeled local data, achieving >99% accuracy in detecting PII.
3. Hybrid Filtering: Augmented the model with rule-based filters using regular expressions to capture local place names and anomalies.
4. Privacy-Preserving Validation: Deployed segmented views to human annotators, restricting PII context and minimising privacy risks.
5. Custom Interfaces: Developed lightweight web tools to streamline labeling, validation, and model testing.
6. LLM Integration: Used large language models (e.g., Claude) for classification tasks such as phenotype detection (e.g., vasospasm suspicion), guided by prompt engineering.
7. Functional Testing: Built a functional test framework to verify model output accuracy and mitigate hallucination risks in generative tasks.

For classification tasks, including disease phenotyping (e.g., vasospasm), generative models like ChatGPT were leveraged using prompt engineering, with outcomes validated via functional equivalence tests and human cross-checking.

DELIVERABLES

KJR was key in producing the following:

- AI Model Development: Fine-tuned transformer models for de-identification and classification.
- Privacy Engineering: Designed secure data validation processes to minimise PII exposure.
- Tooling and Automation: Created custom interfaces for annotation and testing workflows.
- Testing and QA: Built and maintained the functional testing framework for both NER, Classification and LLM tasks.
- Governance Advisory: Provided strategic input on data protection, ethics, and regulatory compliance.
- A scalable, reusable pipeline for de-identification and classification across new datasets.
- Comprehensive documentation of the VDML process for future reuse.

CASE STUDY

HEALTHCARE INDUSTRY

AI-enabled platform ensuring privacy-compliant health data use

Project: Datarwe NLP De-identification & Classification Initiative

KEY OUTCOMES

- High Accuracy: Achieved >99% precision and recall in PII detection.
- Cost Reduction: Reduced de-identification costs by up to 10,000x compared to manual redaction.
- Rapid Data Provisioning: Enabled on-demand generation of de-identified datasets for research and operational use.
- AI Enablement: Provided a foundation for phenotype classification, summarisation, and billing code generation from medical records.
- Robust Governance: Ensured data governance best practices through clean-room protocols and controlled access.
- Scalable Application: Supported distribution shift resilience through retraining and adaptive validation methods.

VALUE TO CLIENT

For Datarwe and Queensland Health, the project delivered a trustworthy, scalable AI infrastructure enabling:

- Rapid, ethical, compliant sharing of ICU data at scale
- Reduced reliance on manual labour for data privacy assurance.
- Accelerated deployment of AI in clinical contexts.
- New pathways for research funding and collaboration through data access.
- Demonstrated leadership in privacy-centric AI applications in healthcare.

KJR's unique expertise in both AI engineering and data assurance/governance was instrumental in achieving these outcomes, solidifying the client's capability to handle sensitive medical data in a modern, AI-driven environment.

TOOLS & TECHNOLOGIES

The Datarwe NLP project leveraged a variety of tools and technologies across the AI development, data engineering, privacy assurance, and validation workflows. Key tools and technologies used:

Category	Key Tools / Technologies
NLP Models	BERT, AWS Comprehend Medical, Claude, OpenAI, (LLMs)
De-identification	Fine-tuned BERT, Regular Expressions
Infrastructure & Pipelines	AWS (cloud compute, storage, orchestration)
Validation & Annotation	Custom labeling UIs, segmentation tools, VDML framework
Governance & Privacy	Clean room protocols, segmented views, access control policies
Testing & Evaluation	Functional test harnesses, human-in-the-loop validation, performance benchmarking frameworks