

Accelerating Policy Review with AI

A PBAC Document Case Study

WHAT'S THE RESEARCH QUESTION?

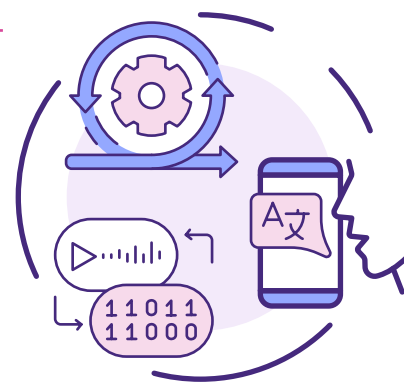
To what extent can **AI/Natural Language Processing (NLP) techniques** be applied to facilitate the policy document review process by **accurately** identifying key information to **improve overall efficiency**? Using the Pharmaceutical Benefits Advisory Committee (PBAC) Public Summary Documents (PSD) as a case study.

Note: PBAC is the body responsible for making recommendations for listing of new medicines in Australia. Each PSD captures the decision made on a drug submission, complete with detailed records from PBAC meetings on the drug's economic and clinical evidence, and other factors.

WHAT IS NATURAL LANGUAGE PROCESSING?

NLP is a branch of Artificial Intelligence (AI) that helps computers understand human language. Text is messy. Unlike numbers, which have clear values, language has nuance—one word can have many different meanings, and different documents use language in unique ways.

NLP techniques **convert words into numerical representations** that a computer can process. It can analyse the structure of sentences to understand how words are related to each other, identifying parts of speech.



RESULTS

NLP tools enabled **faster reviews**, processing hundreds of documents in minutes, and **reduced manual workload**, allowing analysts to **focus on high-value decisions**.

	 MANUAL REVIEW Previous practice	 CHATGPT 4o Versatile AI for General Understanding & Summarization	 NLP TOOLS Targeted & Rule-Based for Research-Specific Needs
Review Speed	Time intensive, lengthy and tedious.	Quick to give answers, but limited to uploading 10 files at a time, which can be time-consuming.	Fast and efficient, in minutes.
Accuracy for Key Details	High but vulnerable to human errors.	Unreliable. Good for summaries but may miss precise details. Likely to fabricate information due to 'completion' function.	High for extracting explicit information, low to middle for summarising implicit and less quantifiable factors.
Scalability	Low. Reviewing more documents requires proportional increases in human effort.	Middle. Handles large text inputs consistently, but limited by input size and requires repeated uploads for many documents.	High. Once the rules and models are set up, they can process a large number of documents efficiently.
Complexity Handling	High. Human reviewers can understand nuanced language and context.	Middle. Understands overall context and provides summarise, but may struggle with highly detailed or nuanced requirements.	Low to High. The effectiveness of NLP in handling complexity can vary based on the programmer's skills.
Privacy concerns	Low. Human processing, minimal privacy risk.	Middle to High. Data uploaded externally to ChatGPT servers.	Low. Locally processed data, minimal privacy risk.