

Amazon Textract vs Google Document AI vs Azure Document Intelligence.

Ian Bruwer

School of Enterprise Computing and Digital Transformation, TU Dublin, Ireland

X00142242@myTUDublin.ie

Introduction

This paper conducts a comparison of Azure AI Document Intelligence, Amazon Textract, and Google Document AI by evaluating price, ease of use, performance, and accuracy. It looks at the mechanisms by which data is collected, from designing a task appropriate form with which to gather hand writing samples, but also discussed the need for redaction of the data being sent for analysis. After sending the data for analysis through the various providers the responses are analysed by comparing actual and expected results using the Ratcliff-Obershelp similarity technique. This measure of accuracy is used to measure how accurately each of the sample documents is captured, but also to compare various sections of the sampled forms to ascertain if certain processors perform better with certain form elements. Ultimately, the paper concludes with a recommendation that if one is already utilising AWS or Azure workloads, then one can comfortably stick with using their respective document processing solutions. However, if starting a greenfield project, overall pricing, ease of use, performance, and accuracy mean that Azure AI Document Intelligence should be your first choice.

Approach

1. Data collection:

A form was design to collect handwriting samples to represent subset of the types of input fields contained in the Irish Naturalisation Process, consisting of various input types. i.e. Text boxes, Check boxes, Tables. The forms were designed to provide expected text such that the research did not need to be concerned with data protection legislation.

2. Data Preparation:

72 Forms were scanned to PDF, with each form also undergoing redaction to ensure that results were not poisoned due to form containing both typed and written versions of the responses.

3. Model Training:

Where appropriate Neural and Template based AI models were trained for each provider; 10 documents used for unredacted documents, 15 documents used for redacted documents.

4. Document Processing:

Primary objective was to compare Generic Pre-trained models for each provider, however, Neural and Template custom models were included in order to ascertain if future research was warranted regarding custom trained models or if generic pre-trained models were adequate for most processing tasks. Documents were sent for processing using REST APIs provided by each provider. This allowed the evaluation of both the performance of the respective APIs, but also for analysis of responses. Though the responses from the providers varied, none were overtly more complicated to process. All results were parsed into consistent structure and were compared with the expected responses using Ratcliff-Obershelp similarity score.

Performance

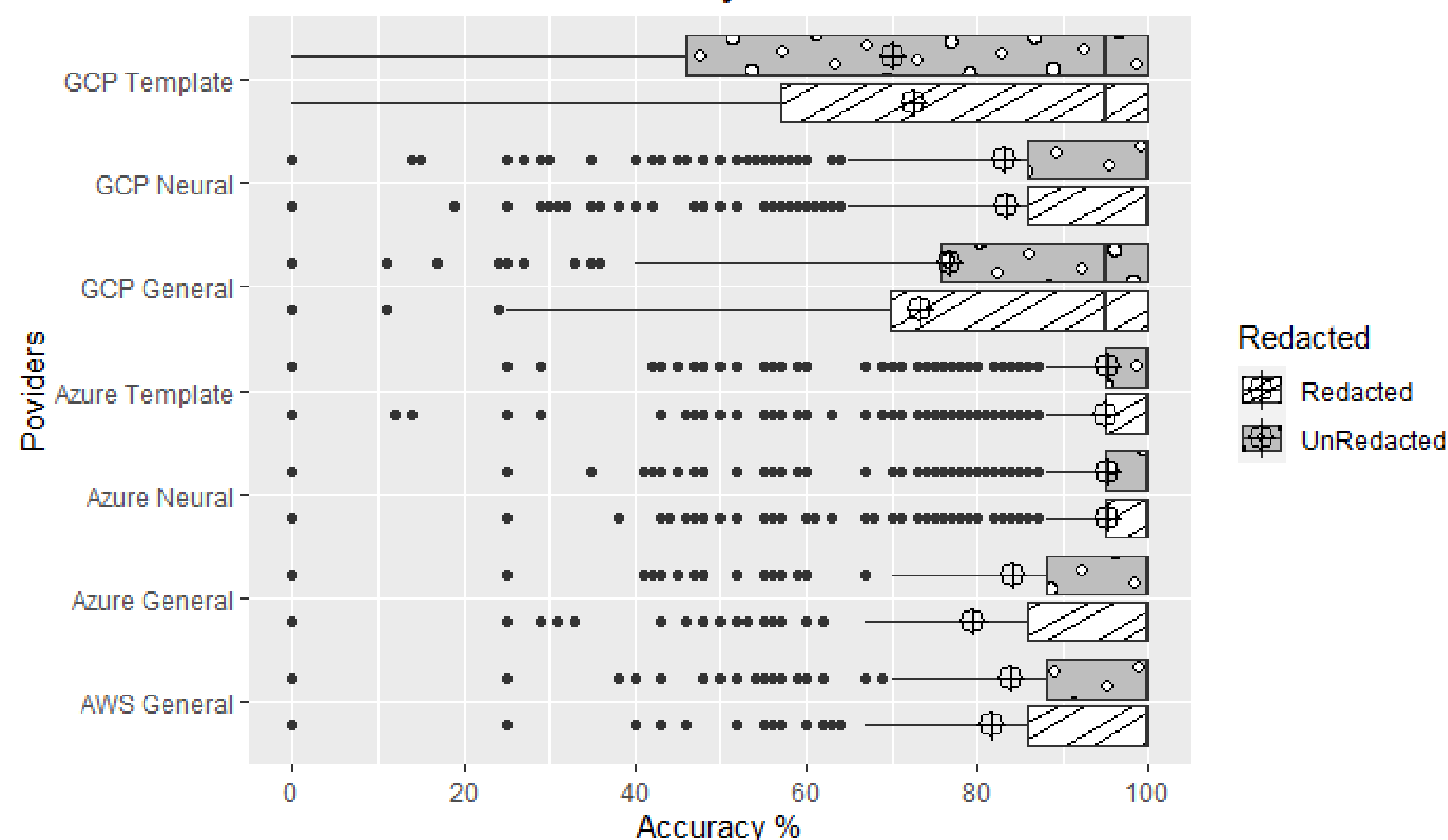
Performance metrics collected for the various providers indicated that Azure is the most performant and consistent processor across all models for both redacted and unredacted documents. Google showed the widest variance in performance between processing of redacted and unredacted documents. AWS and Google processors showed relatively equal processing times for unredacted documents.

Accuracy

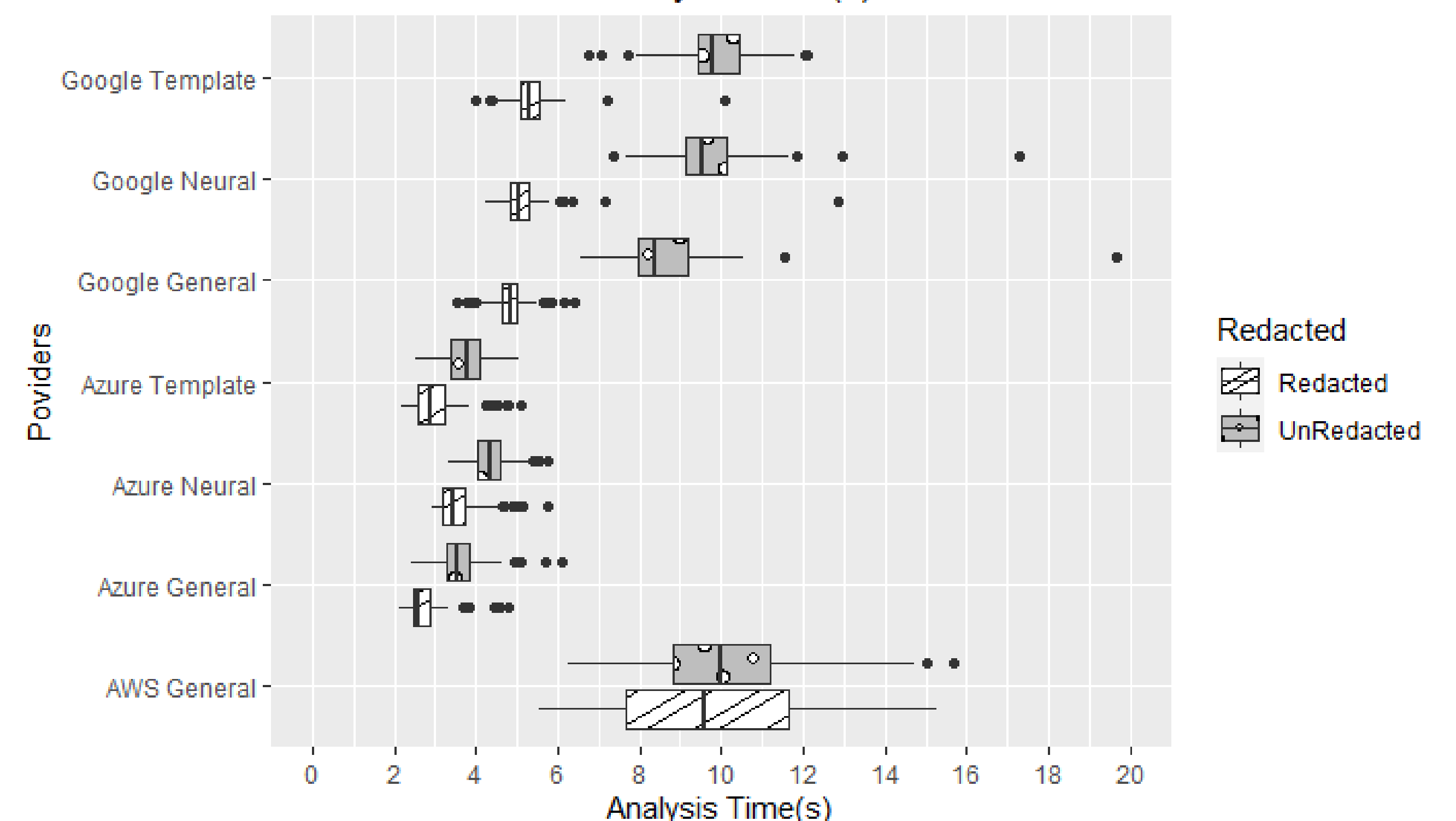
When utilising Pre-trained general models, Azure and AWS proved to be the most accurate of the processors with median accuracy of 100% followed closely by Google. The most accurate processing was delivered by Azure using Neural and Template custom models, with the Google template custom model showed surprising poor performance.

Results

Plot of Providers vs Accuracy



Plot of Providers vs Analysis Time(s)



Conclusions and Future Work

The research shows that if you are looking to perform document analysis, the choice of either Azure AI Document Intelligence or Amazon Textract are suitable option if embedding into an existing Azure or Amazon eco system respectively, however, if only just starting out, the accuracy, cost. and ease of use of the Azure AI Document Intelligence makes it the prime candidate for new endeavours. Future work needs to focus on an in depth comparison of the Neural and Template models provided by the providers in order to ascertain if Google can achieve parity with Azure if provided with enough training data. Generative AI is another area that needs to be investigate in this context as the providers have mechanisms to integrate natural language queries into document analysis.

QR Code for Recording

