



LIG201

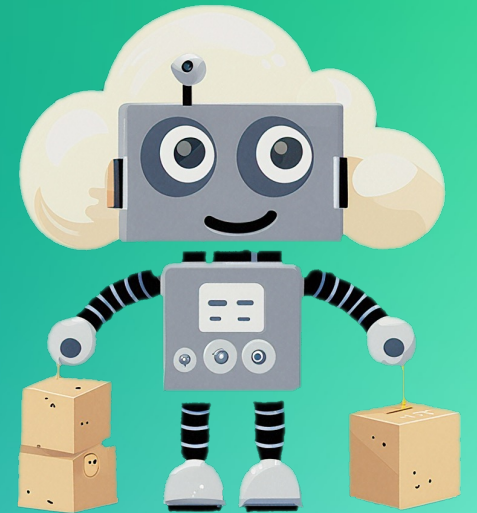
Beyond OCR: Leveraging Vision LLMs for Contract Analysis at SIXT

Ilya Tarygin, PhD

Staff Data Scientist
SIXT

Nikita Kozodoi, PhD

Senior Applied Scientist
Amazon Web Services



GenAI to Analyze Fleet Purchase Contracts at SIXT

Business context:

- 400+ fleet purchase contracts p.a.
- 6+ languages, various document format/scan quality
- Various information to be extracted from the contracts to operate the fleet:
 - *What tire brand should be on the vehicle, when we return it?*
 - *What are the charges in case we don't return vehicle on time?*

Project wins:

- Some use cases are live
- Expansion beyond fleet contracts
- **New innovative approach to recognize the text in the contracts proposed: Vision LLM + OCR**



Common Challenges in PDF Processing

Word Template for Full Length Paper

1st Author
1st author's institute affiliation
Telephone number
E-mail address

2nd Author
2nd author's institute affiliation
Telephone number
E-mail address

3rd Author
3rd author's institute affiliation
Telephone number
E-mail address

Abstract

The abstract is to be in fully-justified italicized text, at the top of the left-hand column as it is here, below the author information. Use the word "Abstract" as the title, in 12-point Times, boldface type, centered relative to the column, initially capitalized. The abstract is to be in 10-point, single-spaced type, and may be up to 3 in. (7.62 cm) long. Leave two blank lines after the abstract, then begin the main text. All manuscripts must be in English.

1. Introduction

These guidelines include complete descriptions of the fons, spacing, and related information for producing your proceedings manuscripts. Please follow them and if you have any questions, direct them to the production editor in charge of your proceedings at the IEEE Computer Society Press: Phone (714) 821-8380 or Fax (714) 761-1784.

2. Formatting Your Paper

All printed material, including text, illustrations, and charts, must be kept within a print area of 6-7/8 inches (17.5 cm) wide by 8-7/8 inches (22.54 cm) high. Do not write or print anything outside the print area. All text must be in a two-column format. Columns are to be 3-1/4 inches (8.25 cm) wide, with a 5/16 inch (0.8 cm) space between them.

Text must be fully justified. A format sheet with the margins and placement guides is available in Word files as <format.doc>. It contains lines and boxes showing the margins and print areas. If you hold it and your printed page up to the light, you can easily check your margins to see if your print area fits within the space allowed.

3. Main Title

The main title (on the first page) should begin 1-3/8 inches (3.49 cm) from the top edge of the page, centered, and in Times 14-point, boldface type. Capitalize the first letter of nouns, pronouns, verbs, adjectives, and adverbs;

SIXT

aws

Design and build accessible PDF tables

Sample tables

Table 1

Column header (TH)	Column header (TH)	Column header (TH)
Row header (TH)	Data cell (TD)	Data cell (TD)
Row header(TH)	Data cell (TD)	Data cell (TD)

Table 2: example of footnotes referenced from within a table


Expenditure by function £ million		2009/10	2010/11 ¹
Policy functions	Financial	22.5	30.57
	Information ²	10.2	14.8
	Contingency	2.6	1.2
Remunerated functions	Agency services ³	44.7	35.91
	Payments	22.41	19.88
	Banking	22.90	44.23
	Other	12.69	10.32

(1) Provisional total as of publication date.
(2) Costs associated with on-going information programmes.
(3) From the management accounts, net of recoveries, including interest charges.

Table 3: "film credits" style layout

Main character	Daniel Radcliffe
Sidekick 1	Rupert Grint
Sidekick 2	Emma Watson
Lovable ogre	Robbie Coltrane
Professor	Maggie Smith
Headmaster	Richard Harris

When you open a ticket, assign it to yourself, fix the issue and close it



Produced with the help of the PDFAnnotator.com

Pictures

FORM - 1102(6B)

APPLICATION FOR EXPLORATION PERMIT

This is an example document used for the purposes to demonstrate how a .PDF can undergo advanced document information extraction using modern technologies. It is not intended for use as a business document in any sense. All contents are fictitious in entirety and do not relate to any specific business(es), their clients, legal agreements or otherwise.

PERMIT DETAILS

Permit Applicant Name	Cortina Minerals Group
Term of Permit	5 Years
Minerals	<input checked="" type="checkbox"/> Coal <input type="checkbox"/> Bauxite <input type="checkbox"/> Nickel <input type="checkbox"/> Tin <input type="checkbox"/> Copper <input type="checkbox"/> Gold <input type="checkbox"/> Other:
Block List	2205656, 2204578, 2201397

PERMIT APPLICANT DETAILS

Company Name	Cortina Minerals Group
ACN/ARBN	37 391 734 800
Address	101 Dempsey Square
Town/City	Brisbane City
State	Queensland
Country	Australia
Phone	+6182266774
Email	heritage@cortinaminerals.com.au

DETAILS OF AREA

I agree that no activities to cause major ground disturbance	<input checked="" type="checkbox"/> Yes <input type="checkbox"/> No
I agree to and accept to the heritage protection conditions	<input checked="" type="checkbox"/> Yes <input type="checkbox"/> No
Undertaken a land availability search	<input checked="" type="checkbox"/> Yes <input type="checkbox"/> No
Will the permit area exceed 80 sub-blocks	<input type="checkbox"/> Yes <input checked="" type="checkbox"/> No

DECLARATION

I understand that this application must comply with the Minerals and Resource Guidelines for this State/Territory and my/our obligations.

Print Name	Karen Greyson
Position	Managing Director
Signature	
Date	04/11/2023

Form – 1102(6B)

Forms

3

Text-Picture Mixture: Amazon Textract

Produced with a Trial Version of PDF Annotator – www.PDFAnnotator.com

When you open a ticket,
assign it to yourself,
fix the issue and close it



Produced with a Trial Version of PDF Annotator
– www.PDFAnnotator.com

Text-Picture Mixture: Amazon Bedrock (Claude 3.5)

Produced with a Trial Version of PDF Annotator - www.PDFAnnotator.com

When you open a ticket,
assign it to yourself,
fix the issue and close it



<section>NO TITLE</section>

When you open a ticket,
assign it to yourself,
fix the issue and close it

<figure>The image shows two men in suits standing in front of a yellow curtain and an American flag. One man is smiling broadly while the other has a more serious expression. The scene appears to be in an official setting, possibly the White House.</figure>

Parsing Tables: Amazon Textract

Design and build accessible PDF tables Sample tables

Table 1

Column header (TH)	Column header (TH)	Column header (TH)
Row header (TH)	Data cell (TD)	Data cell (TD)
Row header(TH)	Data cell (TD)	Data cell (TD)

Table 2: example of footnotes referenced from within a table

Expenditure by function £ million		2009/10	2010/11 ¹
Policy functions	Financial	22.5	30.57
	Information ²	10.2	14.8
	Contingency	2.6	1.2
Remunerated functions	Agency services ³	44.7	35.91
	Payments	22.41	19.88
	Banking	22.90	44.23
	Other	12.69	10.32

Sample tables

Table 1

Column header (TH)

Column header (TH)

Column header (TH)

Row header (TH)

Data cell (TD)

Data cell (TD)

Row header(TH)

Data cell (TD)

Data cell (TD)

Table 2: example of footnotes referenced from

Expenditure by function million

* not using TABLES and LAYOUT

Parsing Tables: Amazon Bedrock (Claude 3.5)

Design and build accessible PDF tables

Sample tables

Table 1

Column header (TH)	Column header (TH)	Column header (TH)
Row header (TH)	Data cell (TD)	Data cell (TD)
Row header(TH)	Data cell (TD)	Data cell (TD)

Table 2: example of footnotes referenced from within a table

Expenditure by function £ million		2009/10	2010/11 ¹
Policy functions	Financial	22.5	30.57
	Information ²	10.2	14.8
	Contingency	2.6	1.2
Remunerated functions	Agency services ³	44.7	35.91
	Payments	22.41	19.88
	Banking	22.90	44.23
	Other	12.69	10.32

<table>

Column header (TH)	Column header (TH)	Column header (TH)
Row header (TH)	Data call (TD)	Data call (TD)
Row header(TH)	Data call (TD)	Data call (TD)

</table>

<table_summary>This table demonstrates the basic structure of an accessible PDF table. It includes column headers (TH) and row headers (TH), along with data calls (TD) to illustrate the proper use of table elements for accessibility.</table_summary>

<table>

Expenditure by function £ million	2009/10	2010/11 ¹
Policy functions		
Financial	22.5	30.57
Information ²	10.2	14.8
Contingency	2.6	1.2
Remunerated functions		
Agency services ³	44.7	35.91
Payments	22.41	19.88
Banking	22.90	44.23
Other	12.69	10.32

Parsing Forms: Amazon Textract

FORM - 1102(6B)

APPLICATION FOR EXPLORATION PERMIT

This is an example document used for the purposes to demonstrate how a .PDF can undergo advanced document information extraction using modern technologies. It is not intended for use as a business document in any sense. All contents are fictitious in entirety and do not relate to any specific business(es), their clients, legal agreements or otherwise.

PERMIT DETAILS

Permit Applicant Name	Cortina Minerals Group
Term of Permit	5 Years
Minerals	<div><div><input checked="" type="checkbox"/> Coal</div><div><input type="checkbox"/> Nickel</div><div><input type="checkbox"/> Copper</div><div><input type="checkbox"/> Other:</div></div> <div><div><input type="checkbox"/> Bauxite</div><div><input type="checkbox"/> Tin</div><div><input type="checkbox"/> Gold</div></div>

PERMIT DETAILS

Permit Applicant Name

Term of Permit

Minerals

Block List

| Cortina Minerals Group

5 Years

X Coal

0 Nickel

[Copper

0 Other:

2205656, 2204578, 2201397

0 Bauxite

0 Tin

0 Gold

* not using TABLES and LAYOUT

Parsing Forms: Amazon Bedrock (Claude 3.5)

FORM - 1102(6B)

APPLICATION FOR EXPLORATION PERMIT

This is an example document used for the purposes to demonstrate how a .PDF can undergo advanced document information extraction using modern technologies. It is not intended for use as a business document in any sense. All contents are fictitious in entirety and do not relate to any specific business(es), their clients, legal agreements or otherwise.

PERMIT DETAILS

Permit Applicant Name	Cortina Minerals Group
Term of Permit	5 Years
Minerals	<input checked="" type="checkbox"/> Coal <input type="checkbox"/> Bauxite <input type="checkbox"/> Nickel <input type="checkbox"/> Tin <input type="checkbox"/> Copper <input type="checkbox"/> Gold <input type="checkbox"/> Other:
Block List	2205656, 2204578, 2201397

<section>FORM - 1102(6B)</section>

APPLICATION FOR EXPLORATION PERMIT

This is an example document used for the purposes to demonstrate how a .PDF can undergo advanced document information extraction using modern technologies.

<section>PERMIT DETAILS</section>

<table>

Permit Applicant Name	Cortina Minerals Group
Term of Permit	5 Years
Minerals	<input checked="" type="checkbox"/> Coal <input type="checkbox"/> Bauxite <input checked="" type="checkbox"/> Nickel <input type="checkbox"/> Tin <input type="checkbox"/> Copper <input type="checkbox"/> Gold <input type="checkbox"/> Other:
Block List	2205656, 2204578, 2201397

</table>

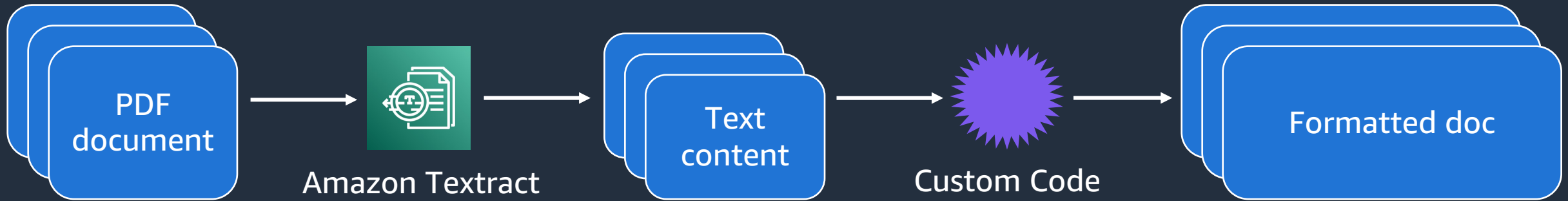
<table_summary>

This table contains permit details for an exploration permit application. It includes the applicant name, permit term, minerals of interest (with checkboxes), and a block list of numbers.

</table_summary>

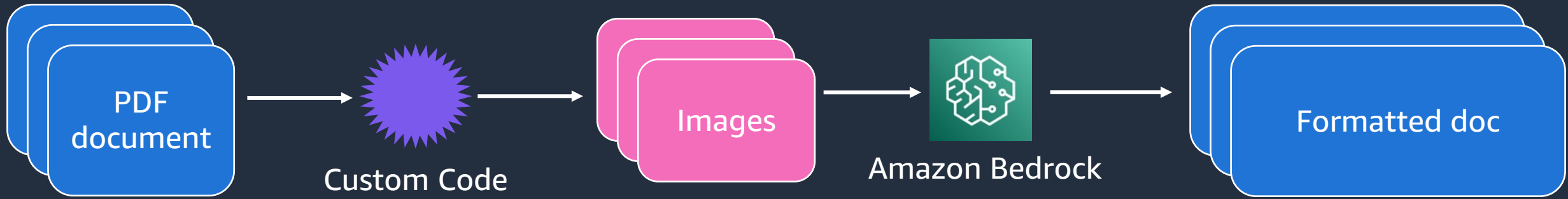
Solution: Hybrid OCR Approach

Traditional OCR



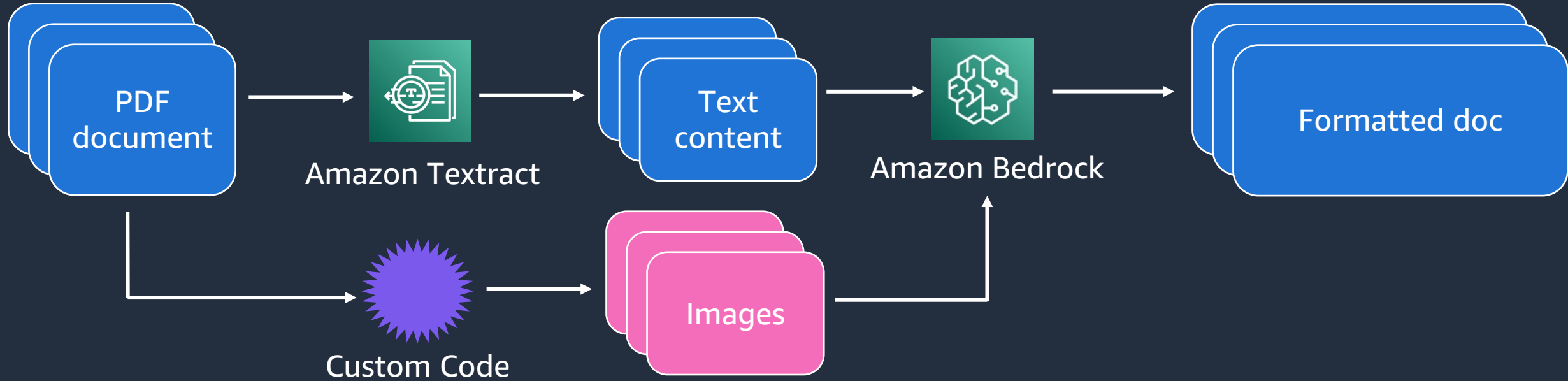
- PDF documents are processed using an **OCR service**
- OCR output is parsed into a **well-formatted text document** (e.g. Markdown)

Vision LLM-Based OCR



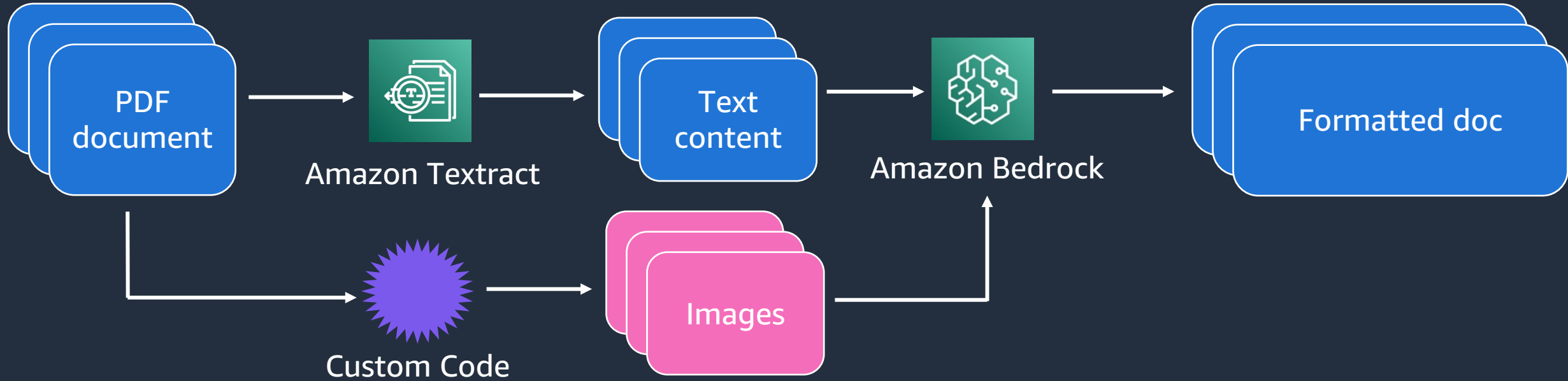
- PDF documents are converted into a **series of images**
- Images are sent to a **vision LLM** (e.g. Anthropic Claude)
- LLM generates a **well-formatted text** version of the document (e.g. Markdown)

Hybrid OCR Approach



- PDF documents are processed using an **OCR service**
- In parallel, PDF documents are converted into a **series of images**
- Both **OCR outputs** and **images** are sent to a vision LLM
- LLM generates a **well-formatted text** version of the document (e.g. Markdown)

Hybrid OCR Approach



Prompt template:

- Use the image as the source of truth, correcting OCR errors.
- Convert content to structured Markdown, preserving format and readability.
- Wrap figures in <figure>, tables in <table_summary>, and sections in <section>.
- Maintain logical grouping without skipping or adding content.
- Output in the document's original language.

Text-Picture Mixture: Hybrid Approach

Produced with a Trial Version of PDF Annotator – www.PDFAnnotator.com

When you open a ticket,
assign it to yourself,
fix the issue and close it



Produced with a Trial Version of PDF Annotator –
www.PDFAnnotator.com

<section>NO TITLE</section>

When you open a ticket,
assign it to yourself,
fix the issue and close it

<figure>The image shows two men in suits standing in front of a yellow curtain and an American flag. One man is smiling broadly while the other has a more serious expression. The scene appears to be in an official setting, possibly the White House.</figure>

Parsing Tables: Hybrid Approach

Design and build accessible PDF tables

Sample tables

Table 1

Column header (TH)	Column header (TH)	Column header (TH)
Row header (TH)	Data cell (TD)	Data cell (TD)
Row header(TH)	Data cell (TD)	Data cell (TD)

Table 2: example of footnotes referenced from within a table

Expenditure by function £ million		2009/10	2010/11 ¹
Policy functions	Financial	22.5	30.57
	Information ²	10.2	14.8
	Contingency	2.6	1.2
Remunerated functions	Agency services ³	44.7	35.91
	Payments	22.41	19.88
	Banking	22.90	44.23
	Other	12.69	10.32

```
<section>Design and build accessible PDF tables</section>

<section>Sample tables</section>

<table>
| Column header (TH) | Column header (TH) | Column header (TH) |
|-----|-----|-----|
| Row header (TH)    | Data cell (TD)      | Data cell (TD)      |
| Row header(TH)     | Data cell (TD)      | Data cell (TD)      |
</table>

<table_summary>This table demonstrates the basic structure of
an accessible PDF table. It includes column headers (TH) and
row headers (TH), along with data cells (TD) to illustrate the
proper use of table elements for accessibility.</table_summary>
```


Parsing Forms: Hybrid Approach

FORM - 1102(6B)

APPLICATION FOR EXPLORATION PERMIT

This is an example document used for the purposes to demonstrate how a .PDF can undergo advanced document information extraction using modern technologies. It is not intended for use as a business document in any sense. All contents are fictitious in entirety and do not relate to any specific business(es), their clients, legal agreements or otherwise.

PERMIT DETAILS

Permit Applicant Name	Cortina Minerals Group
Term of Permit	5 Years
Minerals	<input checked="" type="checkbox"/> Coal <input type="checkbox"/> Bauxite
	<input type="checkbox"/> Nickel <input type="checkbox"/> Tin
	<input type="checkbox"/> Copper <input type="checkbox"/> Gold
	<input type="checkbox"/> Other:
Block List	2205656, 2204578, 2201397

APPLICATION FOR EXPLORATION PERMIT

This is an example document used for the purposes to demonstrate how a .PDF can undergo advanced document information extraction using modern technologies. All contents are fictitious in entirety and do not relate to any specific business(es), their clients, legal agreements or otherwise.

<section>PERMIT DETAILS</section>

<table>

Permit Applicant Name	Cortina Minerals Group
Term of Permit	5 Years
Minerals	<input checked="" type="checkbox"/> Coal <input type="checkbox"/> Bauxite
	<input checked="" type="checkbox"/> Nickel <input type="checkbox"/> Tin
	<input type="checkbox"/> Copper <input type="checkbox"/> Gold
	<input type="checkbox"/> Other:
Block List	2205656, 2204578, 2201397

</table>

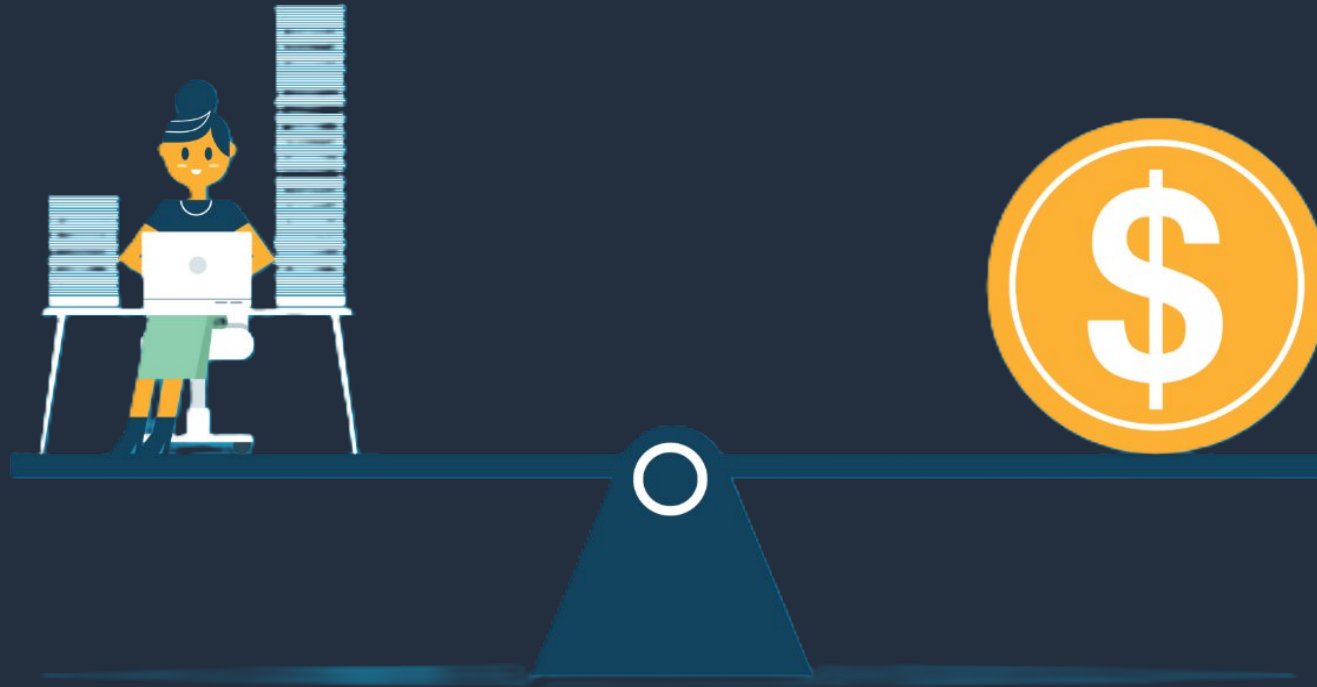
<table_summary>

This table contains permit details for an exploration permit application. It includes the applicant name, permit term, minerals of interest (with checkboxes), and a block list of numbers.

</table_summary>

Conclusions

- Increased the accuracy from 88% to 95% (top-3 retrieval)
- Approach found application in other domains
- Achieved relevant OPEX savings in 25Q1
- Further use-cases with significant potential identified





Thank you!

 [linkedin.com/in/ilya-tarygin](https://www.linkedin.com/in/ilya-tarygin)

Ilya Tarygin, PhD

Staff Data Scientist
SIXT

 [linkedin.com/in/kozodoi](https://www.linkedin.com/in/kozodoi)

Nikita Kozodoi, PhD

Senior Applied Scientist
Amazon Web Services



Please complete the
session survey.

