



Generative AI at scale: Serverless workflows for enterprise-ready apps

[Pre-requisite] Enable foundation model access in Amazon Bedrock

[Pre-requisite] Configuring the front-end application

► Playground

▼ Use cases

► Building a RAG pipeline

▼ Document extraction and summarization

▼ Intelligent document processing with Generative AI

Generative AI Setup

Module 1 - GenAI Introduction for IDP

**Module 2 - Document Classification and Summarization**

Module 3 - Data Extraction

► Scaling with serverless

▼ **AWS account access**

[Open AWS console \(us-west-2\)](#)

[Get AWS CLI credentials](#)

Exit event

[Event dashboard](#) > [Use cases](#) > [Document extraction and summarization](#) > [Intelligent document processing with Generative](#)

## Module 2 - Document Classification and Summarization

### Important

Make sure you have performed the steps described in the **Module 1 - GenAI Introduction for IDP** section before beginning this module.

## Document Classification

A common first step in Intelligent Document Processing (IDP) workflows is document classification. In this step, a document or page is classified according to its content so that it can be routed to the correct downstream process. Previous solutions to this problem involved assembling a collection of sample documents of various types and training a classification model with a set number of types.

With Generative AI, we can skip the pre-training step and perform zero-shot classification by leveraging Large Language Models (LLMs) which are trained on a large corpus of text, so they can understand the context and meaning of words within the broader text. This allows them to make more nuanced and accurate classifications by providing just a name and brief description of the task. This lab will show you how to use Amazon Textract to extract all the text in a document and pass it to the downstream LLMs for classification.

This lab will also highlight the multi-modal capabilities of LLMs. These models can take both text and images as input, including instructions in a prompt.

To get started with the lab, navigate to the gen-ai folder in the file menu, then into the Bedrock folder. Open the notebook titled `02-idp-genai-introduction.ipynb`.

Once the kernel starts, follow along with the instructions in the notebook through the end of **1. Document Classification**.

## Document Summarization

As companies digitize their documents and knowledge libraries, it is often necessary to summarize long documents into a few sentences for showing in search results and cataloging tools. Writing concise and accurate summaries is time-consuming, but with GenAI, it can be automated with minimal effort.

This lab will show you how you can use Amazon Textract with Bedrock to extract text from long documents and quickly summarize it.

To get started with the lab, continue working in the notebook you have open, starting with **2. Document**

© 2008 - 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved. [Privacy policy](#) [Terms of use](#) [Cookie preferences](#)

---

## Conclusion

In this notebook, we have shown how you can use Amazon Textract with GenAI to classify and summarize text from a scanned document. We have also shown how you can do this in a single step by using GenAI for both the OCR and the classification or summarization.

In the next notebook, we will look at data extraction from forms and tables.

**Let's move on to Module 3**

---

[Previous](#)[Next](#)