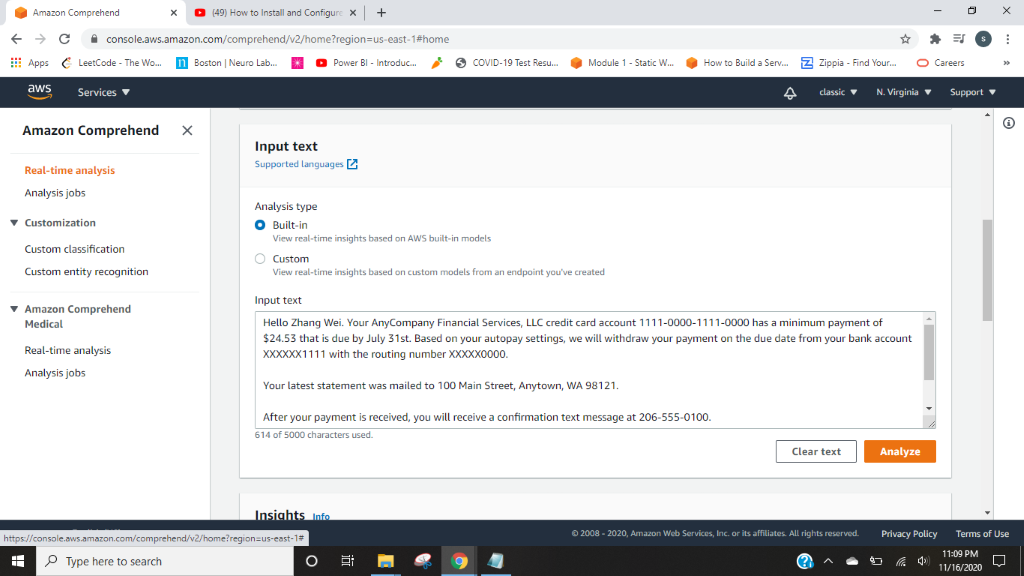
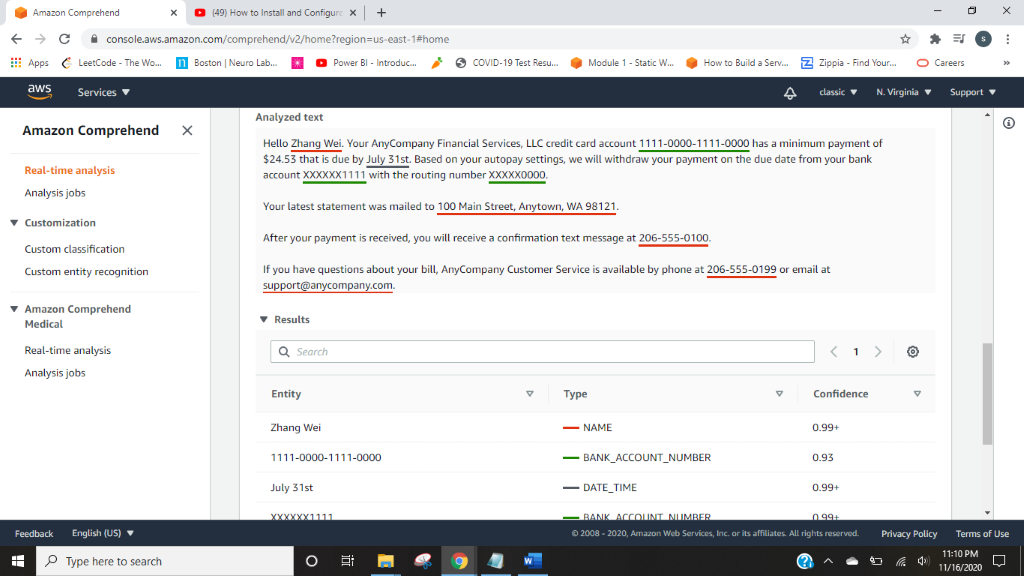
# Detecting and redacting PII using Amazon Comprehend

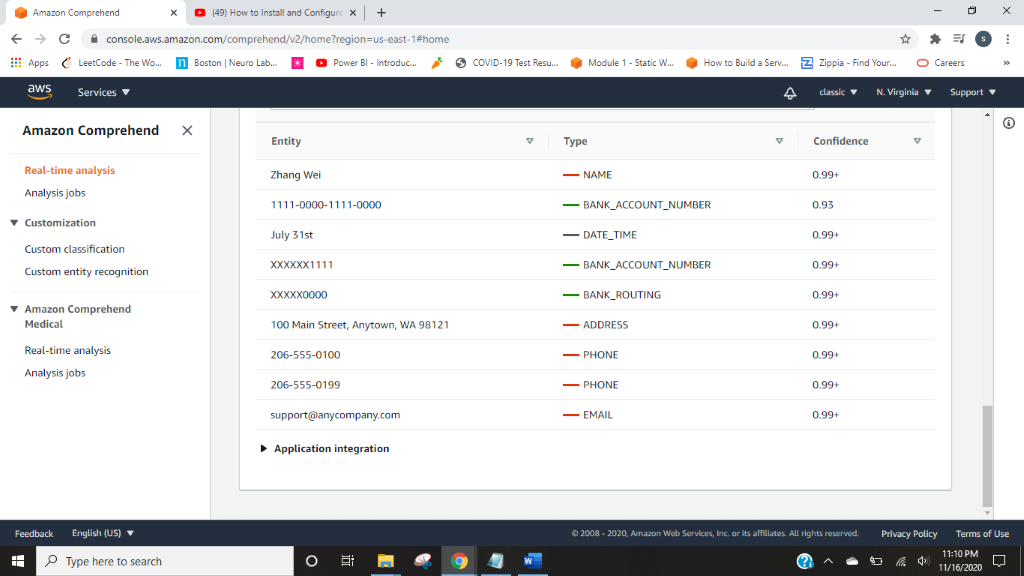
* Amazon Comprehend is a natural language processing (NLP) service that uses machine learning (ML) to find insights and relationships like people, places, sentiments, and topics in unstructured text.
* These ML capabilities of Amazon Comprehend can be used to detect and redact personally identifiable information (PII) in customer emails, support tickets, product reviews, social media, and more.
* Amazon Comprehend automatically identifies PII,When you analyze text using Amazon Comprehend real-time analysis.

# Analyzing text on the Amazon Comprehend console

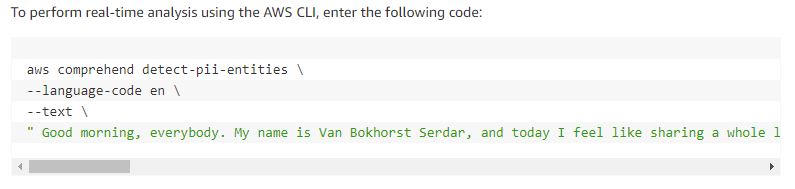


* The PII tab shows color-coded text to indicate different PII entity types, such as name, email, address, phone, and others.
* Each entry shows the PII entity, its type, and the level of confidence Amazon Comprehend has in this analysis.

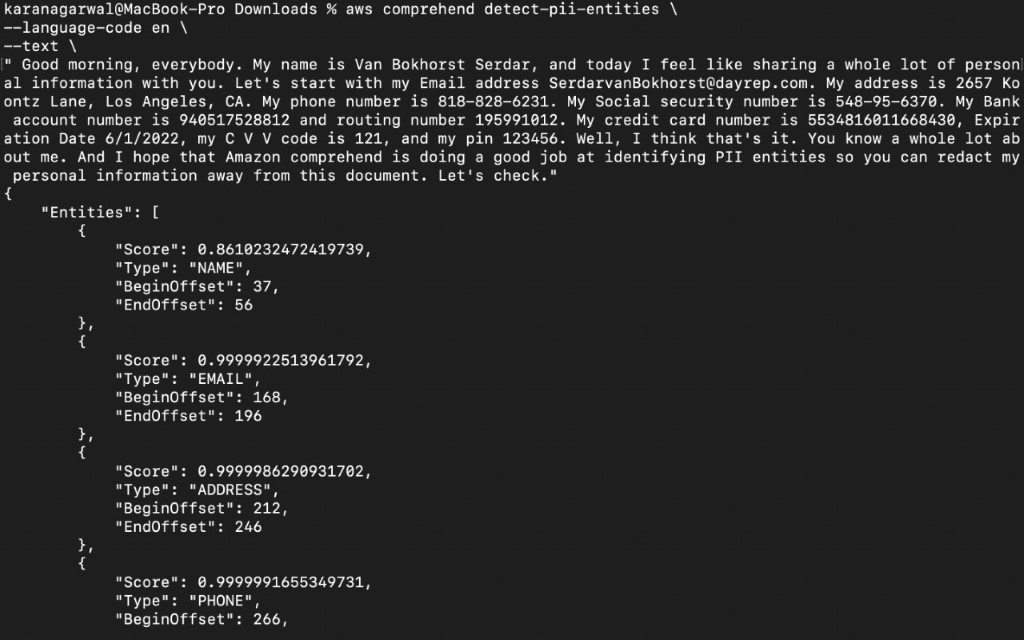


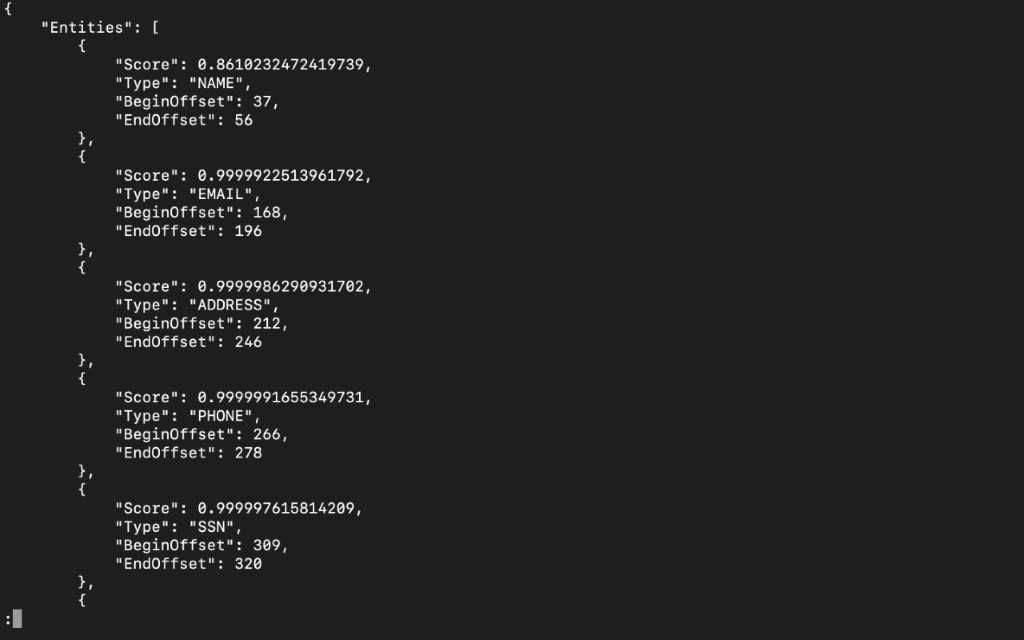


# Analyzing text via the AWS CLI



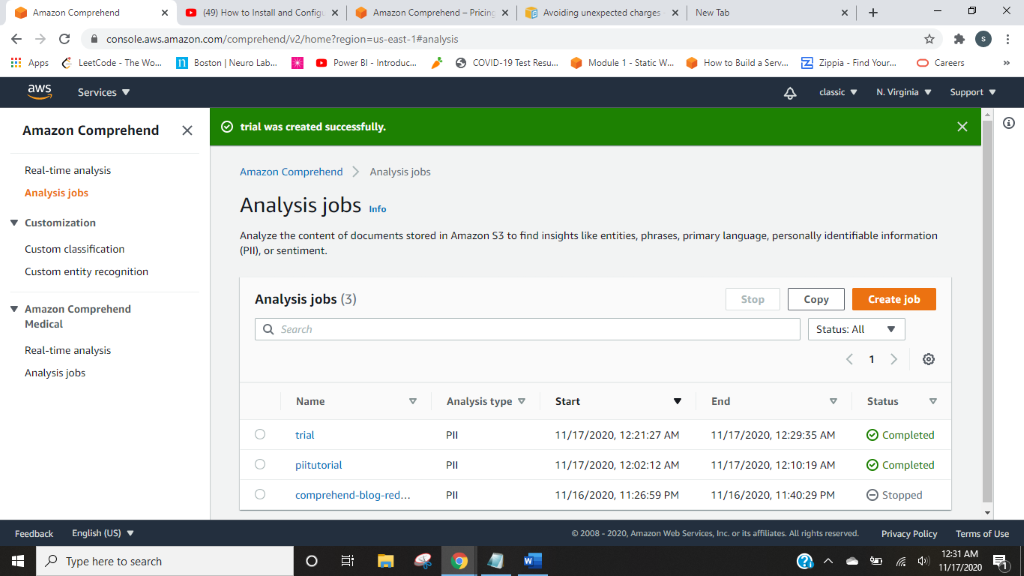
* For each entity, the service returns the type of PII, confidence score metric, BeginOffset, and EndOffset.
* The offsets help you locate PII entities in your documents for document processing to redact it at the secure storage.

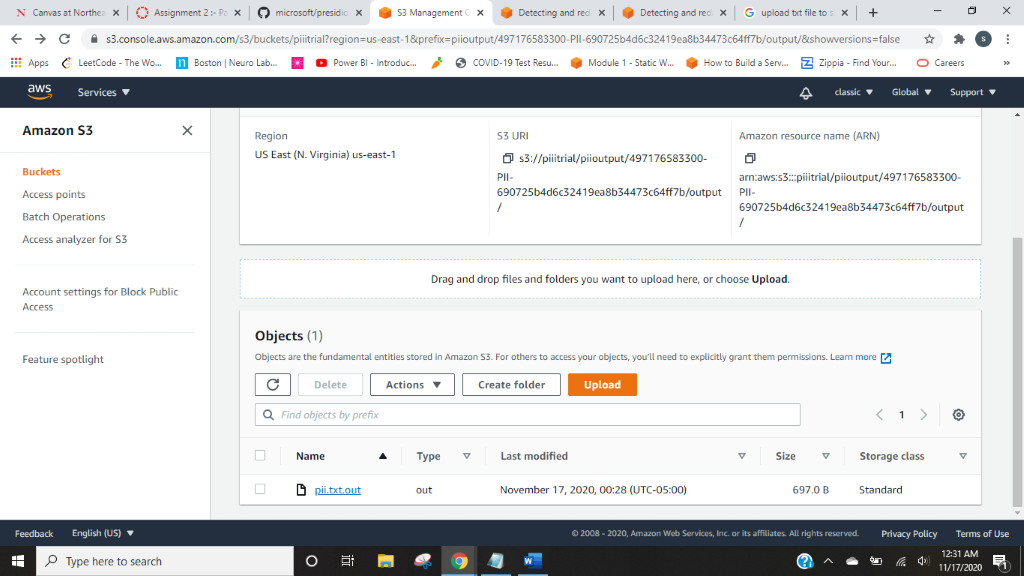


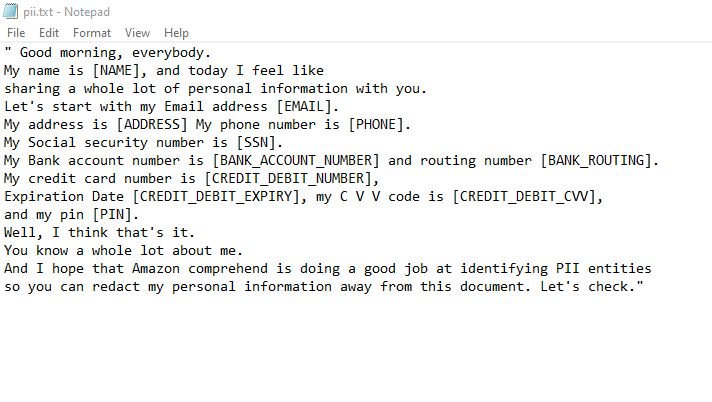


# Asynchronous PII redaction batch processing on the Amazon Comprehend console

* You can choose redaction mode Replace with PII entity to replace PII entities with PII entity type, or choose to mask PII entity with redaction mode Replace with character and replace the characters in PII entities with a character of your choice (!, #, $, %, &, \*, or @).
* To analyze and redact the documents, we need to set up an Amazon s3 bucket to store the document.
* The results of the analysis are returned in an S3 bucket.
* Then create an Amazon Comprehend Analysis Job by entering the Input and output the path to the respective folder in Amazon S3 and creating a role with required permissions to access the input and output S3 buckets for the job that is created and propagated.







# Asynchronous PII redaction batch processing via the AWS CLI

To perform the PII redaction job using the AWS CLI, enter the following code:

