

# Kinesis Introduction

The Kinesis connector fetches data from [Amazon Kinesis Data Streams](#).

## On this page

- [Features](#)
- [Content Retrieved](#)
- [Limitations](#)

## Features

Some of the features of the Kinesis connector:

- Support for incremental and full crawling (with limitations, see below)
- Configurable starting point for data retrieval. Starting position can be specified using timestamps or sequence numbers
- Is search engine independent. The content retrieved can be published by Aspire to any search engine
- Runs from any machine that has access to Amazon Kinesis

## Content Retrieved

---

The Kinesis connector publishes all the data available with each record:

- Data (as text)
- Approximate arrival timestamp
- Sequence number
- Shard ID
- Partition key

## Limitations

---

Due to API limitations and the nature of Kinesis Data Streams itself, the Kinesis connector has the following limitations:

- Due to the streaming nature of the data, crawls (whether full or incremental) *run continuously without end*, unless paused/stopped or when the shards that were picked up at the start of the crawl are closed due to a reshard operation.
- Kinesis Data will be fetched as text data only.
- Cannot adapt to [resharding](#). To get all the data you will have to restart the crawl after a reshard operation.
- Cannot keep track of expired (trimmed) messages. This means that the connector is unable to post update or delete operations to publishers.