Egnyte Crawler Documentation

Introduction

In this document, we will cover how to navigate, create and use the *Egnyte Crawler* in the Admin Interface.

Permissions [IMPORTANT]

Before going further in this guide, be sure you use the *admin account* that is associated with your Egnyte space, service accounts do *not have the necessary permissions* to create API keys, which are required for this crawler to work.
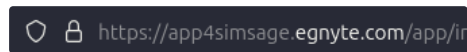
Prerequisites

For our crawler, we require the following credentials:

- **Username** - the email address associated with the *admin* account we want to *crawl* data from
- **Password** - the password of your *admin* account (respective of email)
- **Domain** - the domain of your Egnyte instance (e.g., app4simsage)
  - Please note - do *not* include 'https://' or '.engyte.com' within this field
- **Client ID** - the oath client id from Egnyte (see https://developers.egnyte.com/apps/mykeys)
- **Client Secret** - the oath secret associated with the ID
- **Root path** - The path you wish to crawl from (inclusive)

> 📄 **Grant type** does not need to be specified - is fixed to type **password** *always.*
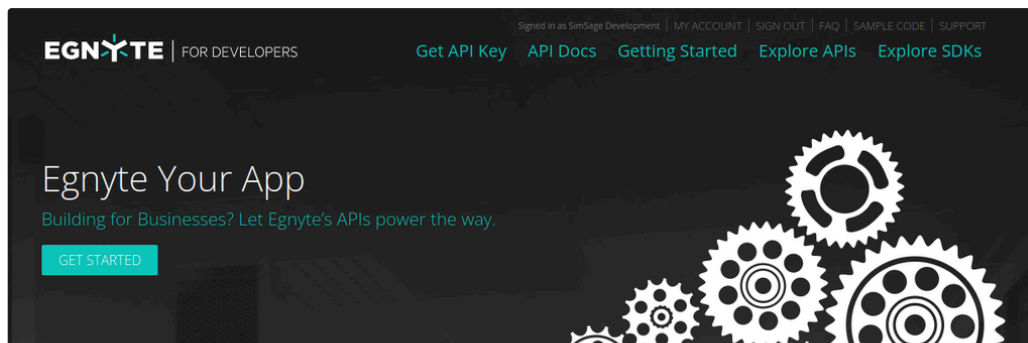
Domain

The domain we want is the one you use to login, alternatively, you can find this in the URL itself:
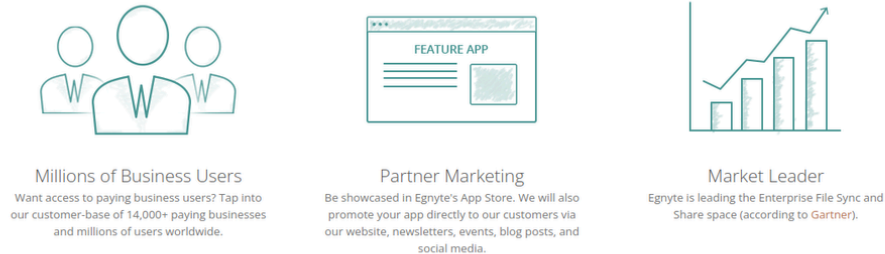

 Please remember - all we want here is the specific domain name (e.g., *app4simsage)*

Client Secret & Client ID

In order to obtain these credentials, you must first login to the *mykeys* page provided by Egnyte ([Egnyte for Developers](#)). Once you have logged in, you should be greeted with the following page or something similar:
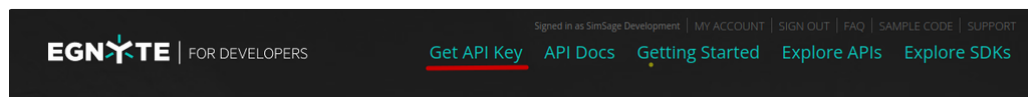

'Egnyte for Developers' *homepage*

From here, navigate to the **Get API Key** page:

'Get API Key' in Navigation Bar

Egnyte Connect API

If You Have Already Requested for Keys

Once you have been directed to the 'Get API Keys' page, you should see the following (if you have applied for API Keys):

The above shows the Client ID and Secrete we wish to use, note these for the crawler.

> 📄 We want the Keys from the **Connect API**, not the Protect API

If You Need to Apply

If you cannot see connect/protect keys, you should have the option to *request* API credentials via the application tab:

From here you can apply for keys, and Egnyte should be in contact to confirm.

> ⚠️ **NOTE:** SimSage recommend contacting Egnyte to increase the User (Access Token) Rate Limits.
>
> As a default these limits are set to 2 calls per second & 1,000 calls per day. Depending on your environment size SimSage recommends adjusting these limits to 1,000 calls per second & 10,000,000 calls per day for the initial data load.
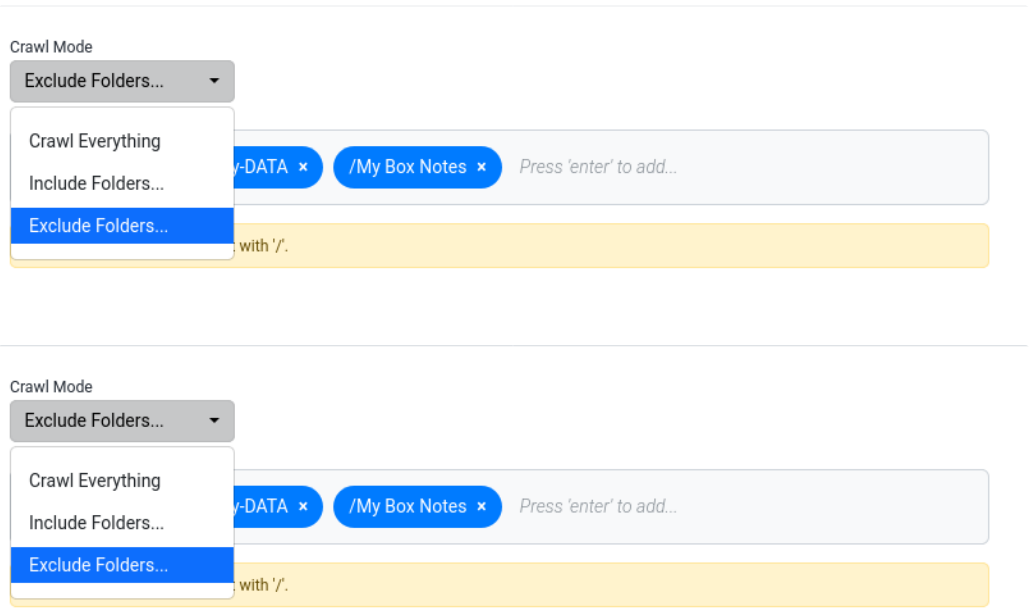
Folder Filtering

Once you have successfully set-up your box account we can begin crawling through specific folders, you have *three* options available to you:

- Include Folders
- Exclude Folders (beginning from root)
- Crawl Everything

> 🗒 Folders need to be referenced by their *absolute path (starting with* `/` *)*

To select on of these three options, select from the drop down shown below:

Crawl Mode Drop-down

if the option selected is to include or exclude folders, a text box will appear prompting you to specify the *paths* of these folders.

> 🗒

By default - all *sub-directories* are also resolved when including/excluding folders

Things to Note

Unfortunately Egnyte does not support crawling for *shortcuts*, there for any files referenced as a shortcut will be excluded from a given crawl of a specific directory.

Further Reading

Egnyte for Developers - Overview