

Airbyte JSON

This covers how to load any source from Airbyte into a local JSON file that can be read in as a document

Prereqs: Have docker desktop installed

Steps:

1. Clone Airbyte from GitHub - `git clone https://github.com/airbytehq/airbyte.git`
2. Switch into Airbyte directory - `cd airbyte`
3. Start Airbyte - `docker compose up`
4. In your browser, just visit `http://localhost:8000`. You will be asked for a username and password. By default, that's username `airbyte` and password `password`.
5. Setup any source you wish.
6. Set destination as Local JSON, with specified destination path - lets say `/json_data`. Set up manual sync.
7. Run the connection!
8. To see what files are create, you can navigate to: `file:///tmp/airbyte_local`
9. Find your data and copy path. That path should be saved in the file variable below. It should start with `/tmp/airbyte_local`

```
from langchain.document_loaders import AirbyteJSONLoader
```

```
!ls /tmp/airbyte_local/json_data/
```

```
_airbyte_raw_pokemon.jsonl
```

```
loader =  
AirbyteJSONLoader('/tmp/airbyte_local/json_data/_airbyte_raw_pokemon.jsonl')
```

[Skip to main content](#)

```
data = loader.load()
```

```
print(data[0].page_content[:500])
```

```
abilities:
ability:
name: blaze
url: https://pokeapi.co/api/v2/ability/66/

is_hidden: False
slot: 1

ability:
name: solar-power
url: https://pokeapi.co/api/v2/ability/94/

is_hidden: True
slot: 3

base_experience: 267
forms:
name: charizard
url: https://pokeapi.co/api/v2/pokemon-form/6/

game_indices:
game_index: 180
version:
name: red
url: https://pokeapi.co/api/v2/version/1/

game_index: 180
version:
name: blue
url: https://pokeapi.co/api/v2/version/2/

game_index: 180
version:
n
```