

Dear candidate,

Thank you for taking the time to carry out this assessment task.

Task: Build a web crawler that gets data from a secure forum

You need to create a crawler that scans, downloads and saves into files.

Fetch as many documents as possible from this URL:

<https://foreternia.com/community/announcement-forum>

Requirements:

- Python 3.9+
- Use concurrency / threading (up to 2 connections at the same time)
- Use [Typer](#) in order to create a simple CLI.

Data structure:

- The files will be saved in the ./data/ directory
- Each unique thread will have a filename prefix (chosen by you).
- Each post will be a JSON file with the page link, title, published time and content.

Coding quality:

- Make sure the code is as good as you would require from your team.
- Use type hints, document the functions with at least a single sentence.
- Use classes and reusable code.

Notes:

- You need to register for free in order to access this forum.
- You can pause for 3 seconds between each document retrieval in order not to get blocked.
- You can use any module / library out there.
- In case assumptions were made, corners were cut - please note what was assumed and what happened.
- You have 48 hours to complete this task.

Thank you for your time and effort for doing this task.

If you need any assistance or have any questions, here is my contact information:

054-553-9018 - idan@webz.io

Good luck,

Idan Beer Hagai