



prefix_to_name_dict

prefix_to_name_dict

Increment by 1 the value of i
(i.e.: $i += 1$)

Is the just
processed **row** the last
element of **csv_reader** to
be processed?

yes

Execute **write_to_csv** (**publisher_data**, **prefix_to_name_dict**, **correct_dois_data**, **incorrect_dois_data**), whose function is that of creating cache files ("**correct_dois.csv**", "**incorrect_dois.csv**", "**publisher_data.csv**", "**prefix_name.json**") where the processed data are stored, in order to limit the risk of losing already retrieved information in the case some problem occurs during the code run.

Empty the lists
correct_dois_data and
incorrect_dois_data
by re-assigning to these
variables two empty
lists

End

Call the function **create_output(publisher_data)**, aimed at storing all the collected data in an output file.

- In particular, this function adds to the **output_dict** dictionary two main key-value pairs: "**citations**", a **dictionary containing two lists** (i.e.: "**valid**" and "**invalid**", where the citational data are separately stored), and "**publishers**", a **list of dictionaries** (each of which represents a publisher with the numbers of valid and invalid citations it received and addressed).
- In addition to that, it counts the **number of citational data which were validated** during the process, and adds to the **output_dict** structure another key-value pair, where the key is "**total_num_of_valid_citations**" and the value is the aforementioned number.
- Further, **for each dictionary in "publishers"** another key-value pair is stored, where the key is the string "**prefix_list**" and the value is the **list of prefixes associated with the publisher in question**.
- In the end, the **cache files are deleted** and **output_dict** data are used to fill the final JSON file (i.e.: "**output.json**"), which is then returned.