**Introduction**

The objective of this assignment is to find the top contributors for the Conference on Computer Vision and Pattern Recognition. Authors’ names are scraped from the 2021, 2022, and 2023 web pages. For each author, the program will record both the total amount of submissions and the submissions per year. At the end, the data of the three authors who contributed the most is displayed in an Excel spreadsheet.

**Table of Functions**

|  |  |
| --- | --- |
| **Function Prototypes** | **Purpose** |
| getYear(url)  compile pattern searching for 4 digit numbers  search for pattern in url  return matches in url | To get the year of the webpage, to count papers by year |
| getNames(url, year)  download website using requests.get  creating a BeautifulSoup object  search for author names  return list of html elements | Get list of html elements holding the author’s names |
| countNames(names, authors, year)  for each html element  get ‘value’ tag from element (author’s name)  if authors name is a key in dictionary  tally counter in corresponding Author object value  else  create new Author object dictionary element  tally counter in corresponding Author object value | Takes a dictionary and uses the list of html elements to 1) create a new ‘name’ : Author item or tallies Authoro object of existing item if key ‘name’ is already there |
| topContributors(authors)  if less than 3 authors:  exit()  elif length == 3  return authors  else:  create list of dictionary values  sort list (key = lambda x:x.sum(), reverse = True)  return list[0], list[1], list [2] | Takes dictionaries of ‘name’ : Author items, makes a list of only the Author object values, sorts by total contributions, return list of top three contributors |
| sendExcel(topThree)  open new workbook  create sheet object  set up row label cells  for x in range 3 (one column for each author)  print author[0].name  print author[0].contributions[2021]  print author[0].contributions[2022]  print author[0].contributions[2023]  save workbook | Displays top three contributors in excel |
| Main()  Create dictionaries to hold authors  Create list of urls  For url in urls  Call getYear  Call getNames  Call countNames  Call topContributors  Call sendExcel | Bring together all the functions to create the program’s entire functionality |

**Classes**

|  |  |
| --- | --- |
| **Prototype** | **Purpose** |
| Class Author  Def \_\_init\_\_(self, name)    Self.name = name  Self.contributions = {‘2021’: 0, etc.}  Def tally(self, year)  Self.contributions[year] += 1  Def sum(self)  Return sum(self.contributions.values()) | Create a unique profile for each author, no repeats, store data, and allow for operations such as tallying contributions and summing all contributions |

**Program Snapshots**

Not much to show for program execution, nothing is printed in the console

At the end, the Excel spreadsheet looks like this

A screenshot of a computer

Description automatically generated

**Group Members & Their Roles**

* Riley (implemented data handling and sorting functionalities)
* Kylie (implemented web scraping methods and the Author class)