

Resources / Assignments (/COMP9321/22T1/resources/73584)
/ Week 5 (/COMP9321/22T1/resources/73586) / Assignment-2

Assignment-2

[Specification](#)[Make Submission](#)[Check Submission](#)[Collect Submission](#)

REST API for Actor/Actress

In this assignment, you are asked to develop a Flask-Restx data service that allows a client to read and store information about actors/actresses, and allows the consumers to access the data through a REST API.

The assignment is based on The TV Maze API, which provides a detailed list of TV shows and people. You can explore the TVmaze API using the following links

- The source URL: (<http://api.worldbank.org/v2/>) <http://api.tvmaze.com/> (<http://api.tvmaze.com/>)
- (<http://api.worldbank.org/v2/>) Documentations on API Call Structure: <https://www.tvmaze.com/api> (<https://www.tvmaze.com/api>)

***Disclaimer: We are using an extremal API provided by TV Maze (<https://www.tvmaze.com/>). We want the students to interact with real-life web services to add to the learning experience and hence we are not responsible nor liable for the wording or inclusion/exclusion of entities and descriptions within the API. This is a "building your REST API" exercise and should be treated as such.

In this assignment, you are going to use the information provided in this API and add a few functionalities as listed below:

Assignment Specification

Question-1: Add a new Actor (3 marks)

This operation can be considered as an on-demand 'import' operation to get the details of an actor to store them in your application. The service will download the JSON data for the given actor/actress (by its name) ; You must use **SQLite** for storing the data (the name of the database should be **YOURZID.db**) locally after importing the actor .

You can use the following to query the API: <https://api.tvmaze.com/search/people?q=> (<https://api.tvmaze.com/search/people?q=>)

For example, you can check the following query: <https://api.tvmaze.com/search/people?q=Brad%20Pitt> (<https://api.tvmaze.com/search/people?q=Brad%20Pitt>)

To Add an actor, your API accepts a query parameter called "name".:

- **name** : name of an actor/actress

After importing the collection, the service should return a response containing at least the following information:

- **id** : a unique integer identifier automatically generated (this might be different than actor ID in tvmaze API)
- **last-update** : the time the collection is stored in the database
- **_links** : the URL with which the imported collection can be retrieved (as shown in below example)

Important : For this assignment , you are asked to access the given Web content programmatically. Some Web hosts do not allow their web pages to be accessed programmatically, some hosts may block your IP if you access their pages too many times. During the implementation, download a few test cases and access the content locally - try not to perform too many HTTP requests programmatically. Check the documentation of the tvmaze API to get insights about their rate-limiting policy.

Example:

```
POST /actors?name=bard pitt
```

An example response [This is not what you store in DB, it is the call's response]

: 201 Created

```
{
  "id" : 123,
  "last-update": "2021-04-08-12:34:40",
  "_links": {
    "self": {
      "href": "http://[HOST_NAME]:[PORT]/actors/123"
    }
  }
}
```

- You must import an actor (e.g, Brad Pitt) if only the given name fully matches ("brad pitt", "brad-pitt", "Brad pitt" all match each other but they should not match a name like "bard pit"). Be noted that the TVMaze API provides a fuzzy search and hence tolerates typos, capital/small, dashes...etc. and at the same time provides more results than the exact name. You should only import the matching one (ignoring cases, and any character except the English alphabet, **SPACE**, and numbers) . If there are a few actors with the same exact name, you need to only import one of them (any)
- Hint: before sending the name to tvmaze API, replace all non-English characters (e.g., , - . _ , ?) with Space and then send it to the API; as such "brad pit-t" will be converted to "brad pit t" and this will not match "brad pitt". Do not make this very complicated; we are not expecting you handle all exceptions, etc. Keep it as simple as mentioned here in the spec.**
- What and how you store the data in DB is up to you, but take a look at the rest of the questions to know what attributes you need to keep for each Actor. At least you need to store: "name", "country", "birthday", "deathday", "gender", and a list of his/her "shows". When you are importing an actor, you should also store all his/her "shows". For this, you need to explore "Tvmaze" API to see how you can automatically get such a list.
- You should never change an ID of a resource (do not update ids)
- You should replace [HOST_NAME]:[PORT] with correct values
- A brief description about _links here: <https://developer.wordpress.org/rest-api/using-the-rest-api/linking-and-embedding> (<https://developer.wordpress.org/rest-api/using-the-rest-api/linking-and-embedding/>) /

Question 2 - Retrieve an Actor (1.5 marks)

This operation retrieves an actor by their ID (the ID that is generated by your application) . The response of this operation will show an actor's details. Please see the provided response example below to see what attributes should be included in the response. "_links" gives the links for previous, next, and current resources if they exist . The next and previous resources are based on the sequential ID generated by your application.

The interface should look as like below:

```
GET /actors/{id}
```

Example Response returns: 200 OK

```
{
  "id": 124,
  "last-update": "2022-03-08-12:34:40",
  "name": "Some One",
  "country": "Australia",
  "birthday": "22-05-1987",
  "deathday": null,
  "shows": [
    "show 1",
    "show 2",
    "show 3"
  ],
  "_links": {
    "self": {
      "href": "http://[HOST_NAME]:[PORT]/actors/124"
    },
    "previous": {
      "href": "http://[HOST_NAME]:[PORT]/actors/123"
    },
    "next": {
      "href": "http://[HOST_NAME]:[PORT]/actors/125"
    }
  }
}
```

Question 3- Deleting an Actor (1.5 marks)

This operation deletes an existing actor from the database. The interface should look like as below:

```
DELETE /actors/{id}
```

Returns: 200 OK

```
{
  "message" : "The actor with id 134 was removed from the database!",
  "id": 134
}
```

Question 4 - Update an Actor (2 marks)

This operation partially updates the details of a given Actor.

The interface should look like the example below:

```
PATCH /actors/{id}
{
  "name": "Some One",
  "country": "Australia",
  "birthday": "22-05-1987",
  "deathday": null
}
```

The above payload is just an example; it can contain any of the actors' attributes. Take a look at the example response in Question 2 to know the existing attributes.

Returns: 200 OK

```
{
  "id" : 123,
  "last-update": "2021-04-08-12:34:50",
  "_links": {
    "self": {
      "href": "http://[HOST_NAME]:[PORT]/actors/123"
    }
  }
}
```

Question 5 - Retrieve the list of available Actors (4 marks)

This operation retrieves all available actors. The interface should be:

```
GET /actors?order=<CSV-FORMATED-VALUE> & page=1 & size=10 & filter=<CSV-FORMATED-VALUE>
```

All four parameters are optional with default values being "order=+id", "page=1", and "size=10", "filter=id,name". "page" and "size" are used for pagination; "size" shows the number of actors per page. "order" is a comma-separated string value to sort the list based on the given criteria. The string consists of two parts: the first part is a special character '+' or '-' where '+' indicates ordering ascendingly, and '-' indicates ordering descendingly. The second part is an attribute name which is one of {id, name, country, birthday, deathday, last-update}. Here are some sample values of "order" :

+name,+id order by "name ascending" and "id ascending"
In this case, sorting by "name" has priority over "id". This is similar to SQL order by clause :
" name ASC, id ASC "

- country order by "country descending"

"filter" is also another comma-separated value (id, name, country, birthday, deathday, last-update, shows), and shows what attribute should be shown for each actor accordingly. Take a look at the following example:

```
GET /actors?order=+id&page=1&size=10&filter=id,name
```

All four parameters are optional with default values being "order=+id", "page=1", and "size=10", "filter=id,name"

Returns: 200 OK

```
{
  "page": 1,
  "page-size": 10,
  "actors": [
    {
      "id" : 1,
      "name" : "Brad Pitt"
    },
    {
      "id" : 2,
      "name" : "Angelina Jolie"
    },
    ...
  ],
  "_links": {
    "self": {
      "href": "http://[HOST_NAME]:[PORT]/actors?order=id&page=1&size=10&filter=id",
    },
    "next": {
      "href": "http://[HOST_NAME]:[PORT]/actors?order=id&page=2&size=10&filter=id",
    }
  }
}
```

Question 6 - get the statistics of the existing Actors (3 marks)

This operation accepts a parameter called "format" which can be either "json" or "image". Depending on the format your operation should provide the following information: In case when the format is an image, your operation should return an image (can be in any image format) and the image illustrates the requested information in a visualization (apply all your knowledge when creating the visualization such as choosing appropriate visualization type and making sure that it is human-readable, clear, and informative).

- Actors break down by an attribute determined by the "by" parameter (a comma-separated value); this parameter can be any of the following Actors' attributes: "country" (showing the percentage of actors per country), "birthday", "gender", and "life_status" (represents what percentages of actors are alive). You may decide to use different types of charts based on the given attribute.
- Total Number of actors
- Total Number of actors updated in the last 24 hours
- UPDATE: Both parameters are required

The interface should look as below when the format is JSON:

```
GET /actors/statistics?format=json&by=country,gender
```

Returns: 200 OK

```
{
  "total": 1241,
  "total-updated": 24,
  "by-country" : { "Australia": 60.7, "USA": 19.2, ... },
  "by-gender" : { "Male": 48.6, "Femail": 51.4}
}
```

Notice:

- You can only use the libraries (latest versions) used in the Labs
- TVMAZE API is only used in Question 1 for importing a new actor. The rest of the operations rely on the existing actors in your local database
- There is no template code for this assignment, your submission should stick to the assignment specifications.
- Your submission will be marked manually.
- You should adhere to the best design guidelines for REST API mentioned in the lecture (e.g., appropriate responses based on JSON format with proper status codes, full API documentation: the generated Swagger should be fully self-explanatory with operation summary, parameter descriptions, default values).
- You should consider cases such as invalid titles or any invalid attempts to use the endpoint (e.g. If the input title doesn 't exist in the data source, return error 404)
- You should return appropriate responses (JSON) in case of already imported collections
- Your code must be implemented in **flask-restx** and automatically generate swagger doc for testing the endpoints.
- Your code must be executable in CSE machines
- Your code must not require installing any software (python libraries are fine)
- Your code must be written in Python 3.5 or newer versions.
- Your operations (API methods) must return appropriate responses in JSON format, and appropriate HTTP response code! e.g., **500 Internal Server Error** is an inappropriate response!
- Make sure you are using the right datatypes in the database and in your API methods (e.g., not string for years '2017')
- Some of the responses of some operations contain "_links", depending on the response this should include links to "self", "next", and "previous" resources.

FAQ:

- **What is the date-time format?**
You can use any format (human-readable), but it should be the case for all date/time values, including birthday, last updated
- **How can I get an actor's SHOWs?**
This is part of the assignment, and you should explore the API and find a way to link actors and shows
- **What libraries can I use?**
You can use all libraries imported in any labs, plus SQLAlchemy
- **Can API users update ID, links, and Last-update time?**
No, these attributes are not editable by the user. An ID should be never changed, links and last-update fields are not part of an actor's attribute.

Submission:

- The Deadline is **Monday the 4th of April 2022 20:00**

- One and only one Python script file named "YOUR_ZID .py" which contains your code
- **How I can submit my assignment?**
Go to the assignment page click on the "Make Submission" tab; pick your files which must be named "YOUR_ZID.py". Make sure that the files are not empty, and submit the files together.
- **Can I submit my file after the deadline?**
Yes, you can. But 5% of your assignment will be deducted as a late penalty per day. No late submission is accepted after 5 days.

Plagiarism

This is an *individual assignment*. The work you submit must be your own work. Submission of work partially or completely derived from any other person or jointly written with any other person is not permitted. The penalties for such offense may include negative marks, automatic failure of the course, and possibly another academic discipline. Assignment submissions will be examined manually.

Do not provide or show your assignment work to any other person - apart from the teaching staff of this course. If you knowingly provide or show your assignment work to another person for any reason, and work derived from it is submitted, you may be penalized, even if the work was submitted without your knowledge or consent. Pay attention that is **also your duty to protect your code artifacts**. If you are using an online solution to store your code artifacts (e.g., GitHub) then make sure to keep the repository private and do not share access to anyone.

Reminder: Plagiarism is defined as (<https://student.unsw.edu.au/plagiarism>) using the words or ideas of others and presenting them as your own. UNSW and CSE treat plagiarism as academic misconduct, which means that it carries penalties as severe as being excluded from further study at UNSW. There are several online sources to help you understand what plagiarism is and how it is dealt with at UNSW:

- Plagiarism and Academic Integrity (<https://student.unsw.edu.au/plagiarism>)
- UNSW Plagiarism Procedure (<https://www.gs.unsw.edu.au/policy/documents/plagiarismprocedure.pdf>)

Make sure that you read and understand this. Ignorance is not accepted as an excuse for plagiarism. In particular, you are also responsible for ensuring that your assignment files are not accessible by anyone but you by setting the correct permissions in your CSE directory and code repository, if using one (e.g., Github and similar). Note also that plagiarism includes paying or asking another person to do a piece of work for you and then submitting it as your own work.

UNSW has an ongoing commitment to fostering a culture of learning informed by academic integrity. All UNSW staff and students have a responsibility to adhere to this principle of academic integrity. Plagiarism undermines academic integrity and is not tolerated at UNSW.

Resource created 18 days ago (Friday 11 March 2022, 03:20:24 PM), last modified about 2 hours ago (Tuesday 29 March 2022, 09:18:06 AM).

Comments

 [Q \(/COMP9321/22T1/forums/search?forum_choice=resource/73777\)](/COMP9321/22T1/forums/search?forum_choice=resource/73777)

 [\(/COMP9321/22T1/forums/resource/73777\)](/COMP9321/22T1/forums/resource/73777)

 Add a comment



Haoyu Wang (/users/z5330735) about 2 hours from now (Tue Mar 29 2022 13:02:06 GMT+0800 (China Standard Time))

Hi, can I use the dicttoxml package for the ass2?

Reply



Mohammadali Yaghoubzadehfard (/users/z5138589) about 2 hours from now (Tue Mar 29 2022 13:28:44 GMT+0800 (China Standard Time))

No

Reply



Jiagan Wang (/users/z5306506) about 12 hours ago (Mon Mar 28 2022 23:24:42 GMT+0800 (China Standard Time))

Hi teacher, I have a silly question: Both of Q1 and Q5 has used /actors with different query string. The RequestParser could only deal with one of them and the other one will be overlapped. So is it correct to use parser for both questions? Or should I use the parser only for Q5 and use @api.expect for Q1? If so, will the example for Q1 (POST /actors?name=bard pitt) be still working with Q5 implemented?

Reply



Mohammadali Yaghoubzadehfard (/users/z5138589) about 2 hours ago (Tue Mar 29 2022 09:23:35 GMT+0800 (China Standard Time))

The difference is the HTTP VERB: POST vs GET
So they are not the same

Reply



Jiagan Wang (/users/z5306506) about an hour from now (Tue Mar 29 2022 12:13:47 GMT+0800 (China Standard Time)), last modified about an hour from now (Tue Mar 29 2022 12:30:02 GMT+0800 (China Standard Time))

So for the examples of Q1 and Q5, "http://127.0.0.1:5000/actors?order=+id&page=1&size=10&filter=name" and "http://127.0.0.1:5000/actors?name=bard pitt",

are these 2 URLs expected to work well simultaneously in browser? In URLs, HTTP VERB could not be selected so they will get overlapped in my case.

In other words, will the query parameter be forced to use in both questions? If not, will it be the similar as lab5.py, in which GET is implemented with parser() and POST implemented with @api.expect(book_model, validate=True) instead?

Hope for your reply!

Reply



Mohammadali Yaghoubzadehfard (/users/z5138589) about 2 hours from now (Tue Mar 29 2022 13:29:27 GMT+0800 (China Standard Time)), last modified about 2 hours from now (Tue Mar 29 2022 13:30:04 GMT+0800 (China Standard Time))

No, post requests cannot be made by the browser like that

Yes for the 2nd question.

Reply



Jiagan Wang (/users/z5306506) about 3 hours from now (Tue Mar 29 2022 14:05:39 GMT+0800 (China Standard Time))

Alright, so I will use @api.expect for POST and parser for GET. I got confused for the example you give in Q1 (POST /actors?name=bard pitt) and now I'm clear. Thanks for advice!

Reply



Yifan He (/users/z5173587) [about 14 hours ago \(Mon Mar 28 2022 21:10:52 GMT+0800 \(China Standard Time\)\)](#)

I have a weird question, is the testing going to last more than 1 day? Because otherwise, for question 6, "total" should always be equal to "total-updated" unless the testing lasts for more than 1 day. Or I understand the spec wrong?

Thanks.

Reply



Mohammadali Yaghoubzadehfard (/users/z5138589) [about 2 hours ago \(Tue Mar 29 2022 09:21:49 GMT+0800 \(China Standard Time\)\)](#)

total means the total number of actors in your API

total-updated: how many actors have been updated in the last 24 hours

Reply



Ka Cheung (/users/z5258869) [about 15 hours ago \(Mon Mar 28 2022 20:15:48 GMT+0800 \(China Standard Time\)\)](#)

Hi are we allowed to use the "io" library to encode the image from q6 into a base64 string? If not, what method are we supposed to use to send the image back in the response?

Reply



May Altulyan (/users/z5131400) [about 13 hours ago \(Mon Mar 28 2022 22:07:31 GMT+0800 \(China Standard Time\)\)](#)

yes

Reply



Sijie Liu (/users/z5329422) [about 15 hours ago \(Mon Mar 28 2022 19:55:05 GMT+0800 \(China Standard Time\)\)](#)

Hi, I wonder if all the input can be considered valid?

For example, if I give a hint the input should be a number, do I still need to judge whether the input is a number?

Thank you.

Reply



May Altulyan (/users/z5131400) [about 13 hours ago \(Mon Mar 28 2022 22:11:39 GMT+0800 \(China Standard Time\)\)](#)

you should consider the input not only a hint

Reply



Kinto Wan (/users/z5353871) about 17 hours ago (Mon Mar 28 2022 18:40:00 GMT+0800 (China Standard Time))

Q6: Are we allowed to import io to plot the graph ?

Reply



May Altulyan (/users/z5131400) about 13 hours ago (Mon Mar 28 2022 22:11:47 GMT+0800 (China Standard Time))

yes

Reply



Luke Junsuo Chen (/users/z5264602) about 20 hours ago (Mon Mar 28 2022 15:07:32 GMT+0800 (China Standard Time))

Hi, I'm a little confused with what to import for Q1. Exactly what characters should we be replacing with a space before we pass it to the tvmaze api since both "brad_pitt" and "brad-pitt" don't return any results and we are expected to match b-r-a-d-p-i-t-t when bradpitt (removing all '-') and b r a d p i t t (replacing '-' with a space) don't return anything. How are we meant to know where to insert a space? Also should something like 'b!@#\$rad p!@#!@#itt' given in the original question match "Brad Pitt" or do we only ignore non-alphanumeric characters if they're at the start/end of the string (e.g. the example given in a previous question is that "Brad pitt***" should match)?

Reply



Mohammadali Yaghoubzadehfard (/users/z5138589) about 2 hours ago (Tue Mar 29 2022 09:19:16 GMT+0800 (China Standard Time))

You do not need to handle such cases, just replace non-english characters with space, and if it does not match do not import.

Do not import: 'b!@#\$rad p!@#!@#itt', "b-r-a-d-p-i-t-t"

Reply



Tam Lam (/users/z5208582) about 21 hours ago (Mon Mar 28 2022 14:36:46 GMT+0800 (China Standard Time))

Hi, I just wanted to clarify something with question 5. If we order birthday by ascending, that means that the least recent birthday comes first right? Similar thing with last-update, if we order by ascending, that means that the least recently updated record comes first right?

Reply



Mohammadali Yaghoubzadehfard (/users/z5138589) about 20 hours ago (Mon Mar 28 2022 15:07:31 GMT+0800 (China Standard Time))

Hi,

When it comes to dates/time, ascending order would mean that the oldest ones come first and the most recent ones last.

Reply



Jiajin Chen (/users/z5337562) a day ago (Mon Mar 28 2022 09:23:54 GMT+0800 (China Standard Time))

it's mentioned that we need import b-r-a-d-p-l-t before, but what if someone's names are: Lee Le and Le Ele?

Reply



Mohammadali Yaghoubzadehfard (/users/z5138589) about 20 hours ago (Mon Mar 28 2022 14:57:15 GMT+0800 (China Standard Time))

Do as the spec, and if there is multiple matches, only import one

Reply



Jiajin Chen (/users/z5337562) about 20 hours ago (Mon Mar 28 2022 15:01:00 GMT+0800 (China Standard Time)), last modified about 20 hours ago (Mon Mar 28 2022 15:05:10 GMT+0800 (China Standard Time))

yah but if we just randomly import one how can we achieve the accuracy of the database? Don't people want to import and get the specific data they want instead of random one?

Reply



Mohammadali Yaghoubzadehfard (/users/z5138589) about 2 hours from now (Tue Mar 29 2022 13:32:04 GMT+0800 (China Standard Time))

For simplicity I updated the spec, to let you keep the spaces, and ignore them if they do not match.

We are not expecting students build a perfect solution for matching; even a lower case matching will be sufficient for this assignment.

Reply



Yang Cao (/users/z3343513) a day ago (Mon Mar 28 2022 00:55:42 GMT+0800 (China Standard Time))

Hi, I am using the wildcard feature from Flask-RESTX to declare one of my models. It gives the correct format in the API response while the doc on Swagger page shows an incorrect model declaration. It seems to be an open issue from Flask-RESTX according to their GitHub - <https://github.com/python-restx/flask-restx/issues...> (<https://github.com/python-restx/flask-restx/issues/57>)

Just want to confirm will this cause any mark lose? Thanks.

Reply



Mohammadali Yaghoubzadehfard (/users/z5138589) a day ago (Mon Mar 28 2022 08:03:10 GMT+0800 (China Standard Time))

Your documentation will be marked, and if it does not make sense you will lose mark

Reply

Load More Comments