Ensembl REST API Workshop

05th to 06th April 2023

helpdesk@ensembl.org

Welcome to the ‘Living Document’ for this Ensembl REST API Workshop. Everyone who is registered for the course has access to edit this document. Please feel free to use this document to ask questions to the Ensembl team throughout the workshop. If you wish to ask questions privately, please do not hesitate to contact the Ensembl HelpDesk.

The ‘Living Document’ is a great way of capturing the knowledge exchanged during the course and saving it for future use by yourselves and those who can’t attend this course. Remember - you don’t have to contribute, but any additions will be welcomed!

The course learning outcomes are:

- Describe the data types accessible through the Ensembl REST API
- Describe the methods in the Ensembl REST API
- Navigate the Ensembl REST API documentation to find available methods
- Write scripts to access data via the Ensembl REST API

Course Overview

Schedule

Wednesday, 05th April 2023, 13:00-17:00 (GMT+3)

- Introduction to the Ensembl REST API
- Accessing GET queries
- Scripting around REST API calls
- Decoding JSON responses
- Using results: fetching specific data points

Thursday, 06th April 2023, 13:00 - 17:00 (GMT+3)

- Content types
- Chaining REST API queries
- Accessing POST queries: bulk REST API calls
- Rate limits
- Closing remarks

Trainers

Aleena Mushtaq, Louisse Paola Mirabueno

Demo, exercises and slides available to download

https://training.ensembl.org/events/2023/2023-04-05-RSGTurkey_API

Feedback survey

https://forms.gle/184KLQvxmhsj6zLe6

We would really appreciate it if you could share your thoughts with us regarding these sessions. We are interested in your opinions, how you feel the experience has benefited you and how it could be improved. If you could find a few minutes to complete a short survey at the end of the last session it would really help us in improving the training we can deliver.

The golden rules for efficient online training

We suggest everyone follows these few simple rules for the course to run as smoothly as possible:

- Mute all your microphones to reduce the background noise - only unmute when asking a question
- If you have questions during the course, either:
    - Unmute yourself and ask question out loud
    - Type your question in this shared reference document. Avoid typing in the Zoom chat as it is difficult to keep track there
- When you are ready to move on from a practical/exercises please click the Raise Hand button Raise Hand so that the tutor knows you are ready and can proceed with the course. Please remember there are more exercises than we have time for. You are not expected to complete all of them, but rather pick and choose the ones most relevant to your work and you are welcome to finish them in your own time.

Resources:

www.ensembl.org - Ensembl genome browser (chordates)
grch37.ensembl.org - Ensembl archive for browsing data associated with the human GRCh37 genome assembly
www.ensemblgenomes.org - Ensembl Genomes genome browser (non-chordates)
www.ebi.ac.uk - EMBL-EBI website
www.ebi.ac.uk/services - EMBL-EBI databases and tools
Train Online - EMBL-EBI e-learning for our databases and tools

Questions

If you have any questions/problems that you would like to share and are applicable to the whole class please write them below. A tutor will answer your question.

Write your question after the last one you can see in this document and write your name.

1. **<İlke>< Can we ask for two or more genes or more species? If so, how do we write on the lookup up or the python request? >**
   <Aleena><You can definitely query for more than one gene at a time, however, you will need to use the POST method for that. Louisse will cover this in tomorrow's session. You will not be able to query for more than one species at a time I am afraid, unless you are looking for orthologues of your gene(s) of interest in other species.>

2. **<Omer Faruk><When I apply the steps in Exercise 3.2 to different phenotype term, e.g. "epilepsy" or "seizure", I encountered the "KeyError: 'Variation'" error. How can I solve this error?>**
   <Louisse><Hi Omer! Could you share your script below, please?>
   <Omer Faruk><

```
import requests, sys, json
from pprint import pprint

def fetch_endpoint(server, request, content_type):

    r = requests.get(server+request, headers={ "Accept" :
content_type})

    if not r.ok:
        r.raise_for_status()
        sys.exit()

    if content_type == 'application/json':
        return r.json()
    else:
        return r.text

print ("Variant\tp-value\tPub-med ID\tRisk allele\tGene")

# define the general URL parameters
server = "http://rest.ensembl.org/"
ext_phen = "/phenotype/term/homo_sapiens/epilepsy?"
con = "application/json"
```

```
# submit the query
get_phen = fetch_endpoint(server, ext_phen, con)

for variant in get_phen:
    id = variant['Variation']
    pv = str(variant['attributes'].get('p_value'))
    pmid = variant['attributes']['external_reference']
    risk = str(variant['attributes'].get('risk_allele'))
    gene = str(variant['attributes'].get('associated_gene'))

    print (id + "\t" + pv + "\t" + pmid + "\t" + risk + "\t" +
gene)
```

\>
<**Louisse**><Thank you! If you look at the [full JSON](), you'll notice that in some dictionaries the key "Variation" isn't present, so you will need to use the get() function (and transform this into a string using str()), so that Python will ignore these null values and move on to the next key.
<**Omer Faruk**><Thank you for answer.>

3. <**Sinem**><How can i apply this method for my data? For example, I want to compare a phenotype between different datasets? >
   <**Tutor**><Answer>

4. <**İlke**><Could you show Ex. 4.1. by modifying the example URL in rest? I tried it as well but I don't think I managed it.>
   <**Louisse**><Answered live>

5. <**Name**><Question>
   <**Tutor**><Answer>

6. <**Name**><Question>
   <**Tutor**><Answer>

7. <**Name**><Question>
   <**Tutor**><Answer>

8. <**Name**><Question>
   <**Tutor**><Answer>

9. <**Name**><Question>
   <**Tutor**><Answer>

10. <**Name**><Question>
    <**Tutor**><Answer>

11. <Name><Question>
    <Tutor><Answer>

12. <Name><Question>
    <Tutor><Answer>

13. <Name><Question>
    <Tutor><Answer>

14. <Name><Question>
    <Tutor><Answer>

15. <Name><Question>
    <Tutor><Answer>

16. <Name><Question>
    <Tutor><Answer>

17. <Name><Question>
    <Tutor><Answer>

18. <Name><Question>
    <Tutor><Answer>

19. <Name><Question>
    <Tutor><Answer>

20. <Name><Question>
    <Tutor><Answer>

21. <Name><Question>
    <Tutor><Answer>

22. <Name><Question>
    <Tutor><Answer>

23. <Name><Question>
    <Tutor><Answer>

24. <Name><Question>
    <Tutor><Answer>

25. <Name><Question>
    <Tutor><Answer>

Other resources

This section of the Living Document provides additional resources that might be useful to you in developing skills and knowledge in the course topic area.

Train online

EMBL-EBI provides an e-learning platform called Train online. Train online provides free courses on Europe's most widely used data resources, created by experts at EMBL-EBI and collaborating institutes. You do not need to have any previous experience of bioinformatics to benefit from this training. We want to help you to become a confident user of our data resources; we are not trying to train you to become a bioinformatician.

Tutorials of interest might include:

- Ensembl: Quick Tour
- Ensembl: Browsing Genomes
- Ensembl Genomes (non-chordates): Quick Tour
- Ensembl REST API

Webinar series

The EMBL-EBI training team also runs regular webinar series featuring the EBI resources. See the Training pages for more information. You can also catch up on any webinars that you might have missed in Train online or on Youtube.