Building an Image Database with Searchable Metadata

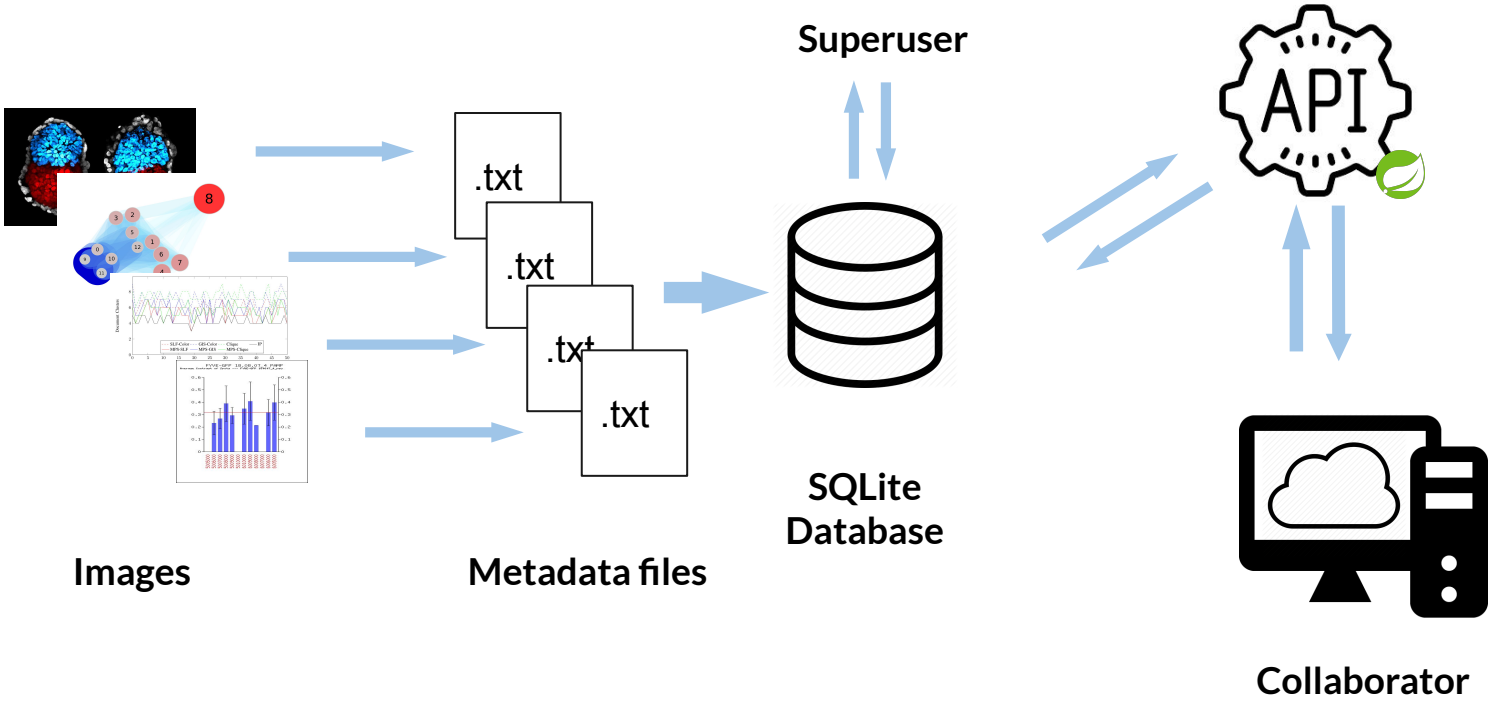
Programming Project 5
Group 3 - Descartes

LSI | b-it | Rheinische Friedrich-Wilhelms-Universität Bonn
Programming Lab II
Instructors: Dr. Jens Dörpinghaus, Dr. Sebastian Schaaf
WS 2019/2020

Shreya Kapoor
Sophia Krix
Gemma van der Voort

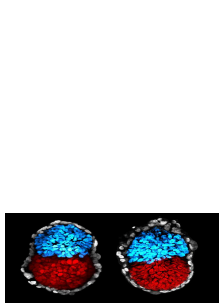
How to store and retrieve scientific image data?

Pipeline


Superuser
API
.txt
.txt
.txt
.txt
SQLite
Database
Images
Metadata files
Collaborator

1) Mapping between images and metadata

Generation of structured metadata files



Images

.png
.jpg
.jpeg

Associated

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4584022/figure/Fig1/

Metadata files (if exist)

Input Metadata

Author : xx
Title : xx
Database : xx
Infographic : xx
https://www.xx.com

Metadata files (generated)

1) Mapping between images and metadata

Example Terminal Input

```
(base) eduroam0097:task01 sophiakrix$ java -cp target/task01-0.0.1-SNAPSHOT.jar de.
bit.pl02.pp5.task01.CommandLineInterface -d PP5 -ip d41586-018-07663-9_16315876.jpg
 -m -im "Nicolas Rivron","Debate ethics of embryo models from stem cells",Nature,1
-o -p
```

-d or --directory

-ip or --inputfile

-m or --meta

-im or --inputmeta

-o or --overwrite

-p or --print

1) Mapping between images and metadata

Entering meta information with the -im parameter

```
* Author - String
* Title - String
* Database_name - String
* Infographic_number - Integers in range 1 to 4 i.e. (1,2,3,4)
    1. Implies image of a cell/tissue
    2. Implies image of a biological cartoon
    3. Implies that the image is a graph
    4. Implies the type of the image doesn't fit into the above classification
```

All the metadata entered after the parameter -im shall be separated by commas. If any values are unknown just add a dummy string.

1) Mapping between images and metadata

Example Output for Terminal and Text File

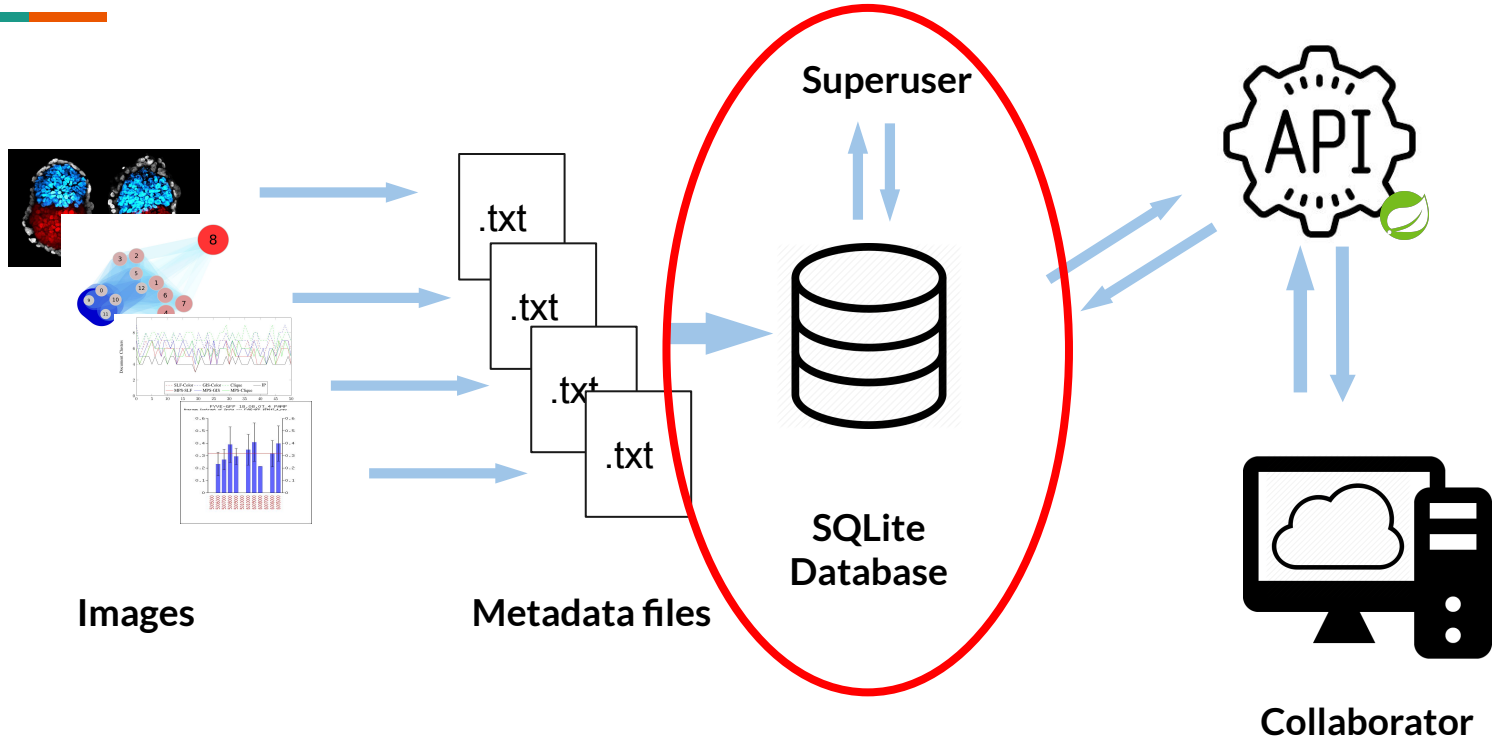
```
************************************************************
File name: d41586-018-07663-9_16315876.jpg
Reading file in directory, path is: PP5/d41586-018-07663-9_1
6315876.meta
-------------------- MESSAGE --------------------
Metafile exists, will deal with existing metafile
-------------------------------------------------

Contents of the current file
=================================================
https://www.nature.com/articles/d41586-018-07663-9
Author:Nicolas Rivron
Title:Debate ethics of embryo models from stem cells
Database: Nature
Infographic:1
=================================================
```

d41586-018-07663-9_16315876.meta

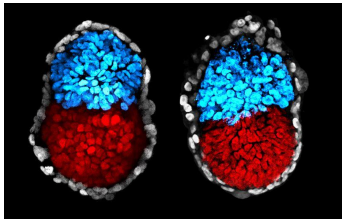
```
https://www.nature.com/articles/d41586-018-07663-9
Author:Nicolas Rivron
Title:Debate ethics of embryo models from stem cells
Database: Nature
Infographic:1
```

Pipeline


Superuser
API
.txt
.txt
.txt
.txt
SQLite
Database
Images
Metadata files
Collaborator

2) Database from metadata files and images

Storing image files and metadata files into a database



Author : xx
Title : xx
https://www.example.com

ID	AUTHOR	TITLE	LINK	PICTURE
1				
2				
...				

Images

Metadata files (generated)

SQLite Database

2) Database from metadata files and images

Example Terminal Input

```
(base) eduroam0097:~ sophiakrix$ java -cp ~/gitlab/group-03-descartes/ProgrammingProject05
/Task02/task02/target/task02-0.0.1-SNAPSHOT.jar de.bit.pl02.pp5.task02.CommandLineInterfac
e -d ~/gitlab/group-03-descartes/ProgrammingProject05/Task01/task01/PP5 -n Desktop,trialyy
 -git "Debate ethics of embryo models from stem cells",Desktop
```

-d or --directory

-git or --getImagebyTitle

-n or --name

2) Database from metadata files and images

Querying the database for metadata

-gma or **--getMetabyAuthor**: Enter the name of the author of which you want to retrieve the metadata and the outputpath where to save it at

```
-gma author_name,output_path
```

-gmt or **--getMetabyTitle**: Enter the name of the title of which you want to retrieve the metadata and the outputpath where to save it at

```
-gmt title_name,output_path
```

2) Database from metadata files and images

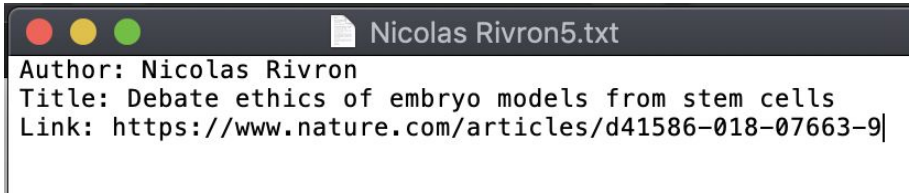
Querying the database for images

Terminal Output

```
Connecting to the database:
jdbc:sqlite:Desktop/trialyy.db
The database currently contains 16 elements
Files from the particular folder /Users/sophiakrix/gitlab/grou
p-03-descartes/ProgrammingProject05/Task01/task01/PP5have alre
ady been added to the database named trialyy
------------------------------------------------------------

Author: Nicolas Rivron
Title: Debate ethics of embryo models from stem cells
Link: https://www.nature.com/articles/d41586-018-07663-9
Author: Nicolas Rivron
Title: Debate ethics of embryo models from stem cells
Link: https://www.nature.com/articles/d41586-018-07663-9
```

File Output



2) Database from metadata files and images

Querying the database for images

-gia or **--getImagebyAuthor**: Enter the name of the author from which you want the image and the outputpath where to save it at

```
-gia author_name,output_path
```

-git or **--getImagebyTitle**: Enter the name of the title from which you want the image and the outputpath where to save it at

```
-git title_name,output_path
```

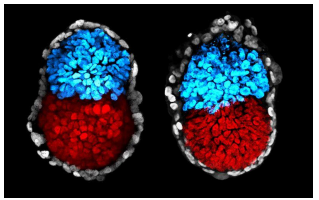
2) Database from metadata files and images

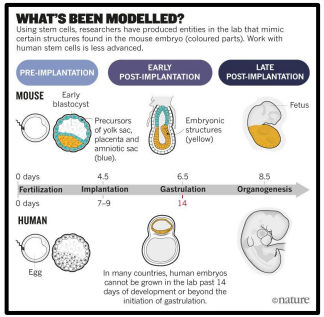
Querying the database for images

Terminal Output

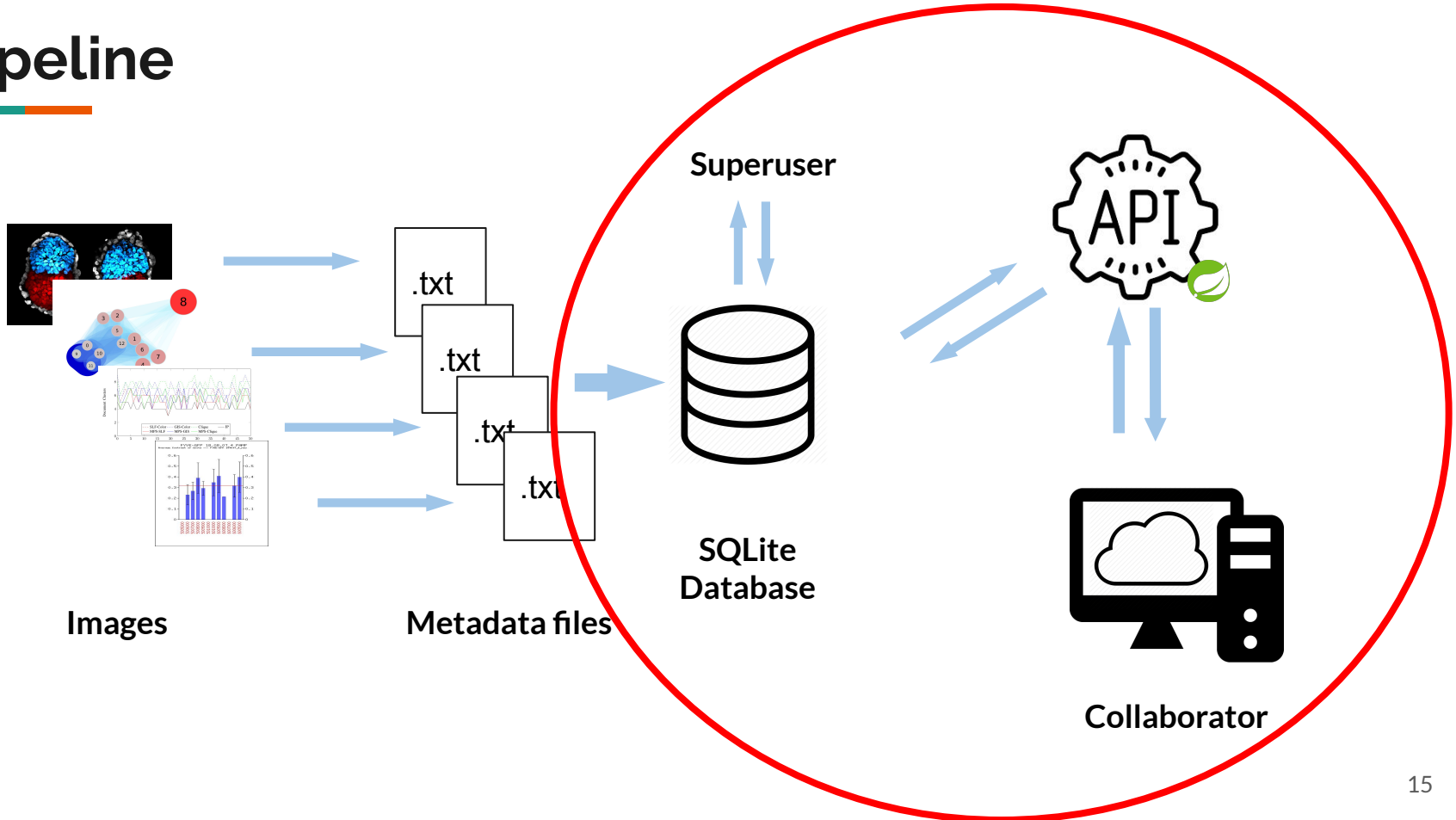
```
Connecting to the database:
jdbc:sqlite:Desktop/trialyy.db
The database currently contains 0 elements
Executing Store method
executed the query:SELECT * FROM IMAGES WHERE TITLE='Debate ethics of embryo model
s from stem cells';
saved image at: Desktop/Nicolas Rivron5.png
saved image at: Desktop/Nicolas Rivron11.png
```

Image Output





Pipeline


Superuser
API
.txt
.txt
.txt
.txt
SQLite
Database
Images
Metadata files
Collaborator

3) RESTful service for collaborators

Design Considerations

1. Superuser
    a. flexible set-up
    b. limiting options for collaborators
    c. privacy
2. Collaborator
    a. easy to access
    b. easy to use

3) RESTful service for collaborators

/store method

```
gemma@maanvis:~/Downloads$ curl -F 'author=Nicolas Rivron' -F 'title=Deba
te ethics of embryo models from stem cells' -F link=https://www.nature.co
m/articles/d41586-018-07663-9 -F file=@embryomodelsdebate.png http://loca
lhost:8080/trialyy/store
{"fileName":"You have succesfully uploaded file: embryomodelsdebate.png"}
```

curl command line tool

query

response

3) RESTful service for collaborators

/get method

```
gemma@maanvis:~$ curl http://localhost:8080/trialyy/get?author=Nicolas%20Riv
ron
[{"id":5,"author":"Nicolas Rivron","title":"Debate ethics of embryo models f
rom stem cells","link":"https://www.nature.com/articles/d41586-018-07663-9"}
,{"id":11,"author":"Nicolas Rivron","title":"Debate ethics of embryo models
from stem cells","link":"https://www.nature.com/articles/d41586-018-07663-9"
}]gemma@maanvis:~$
```

curl command line tool

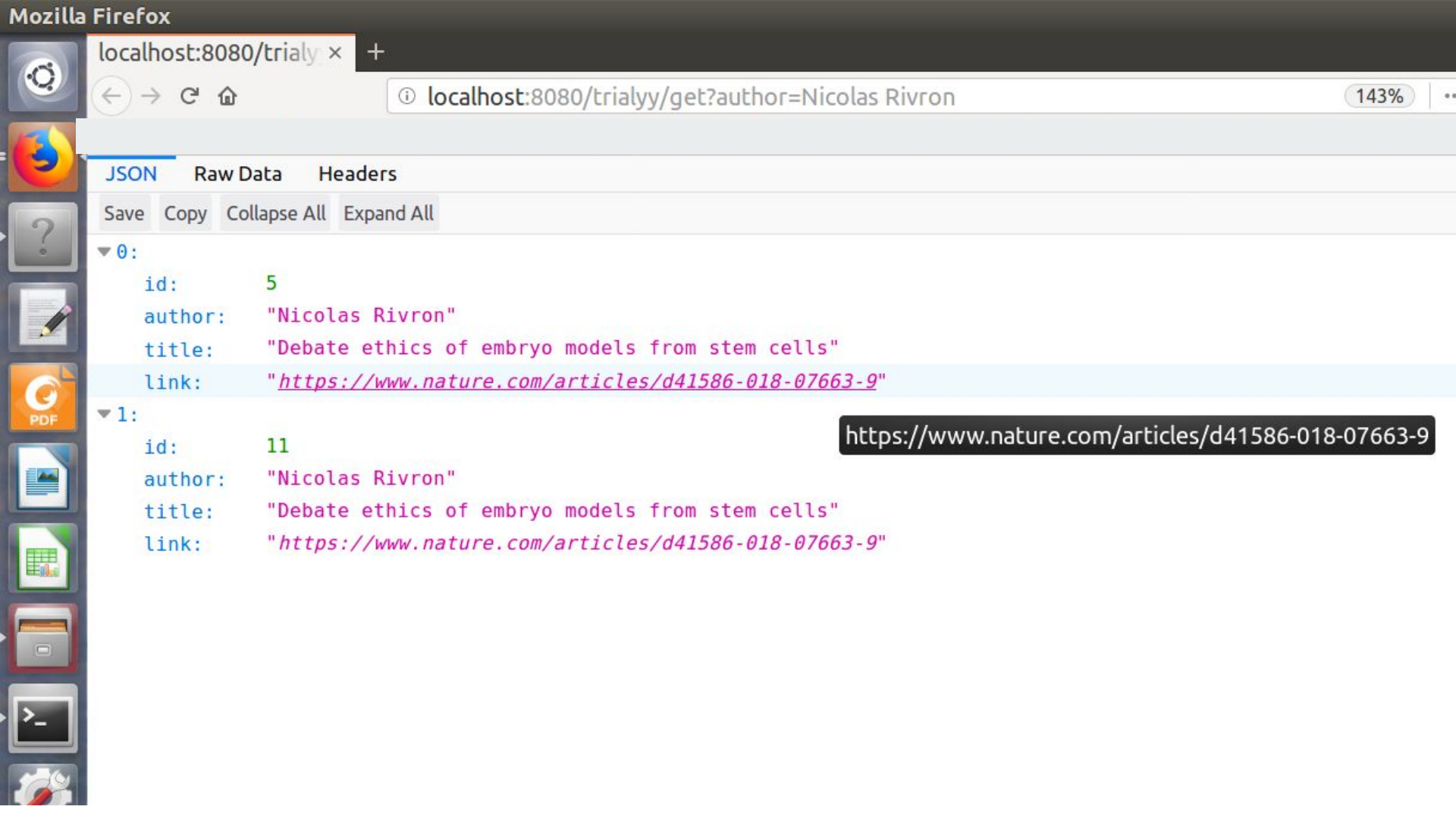
query

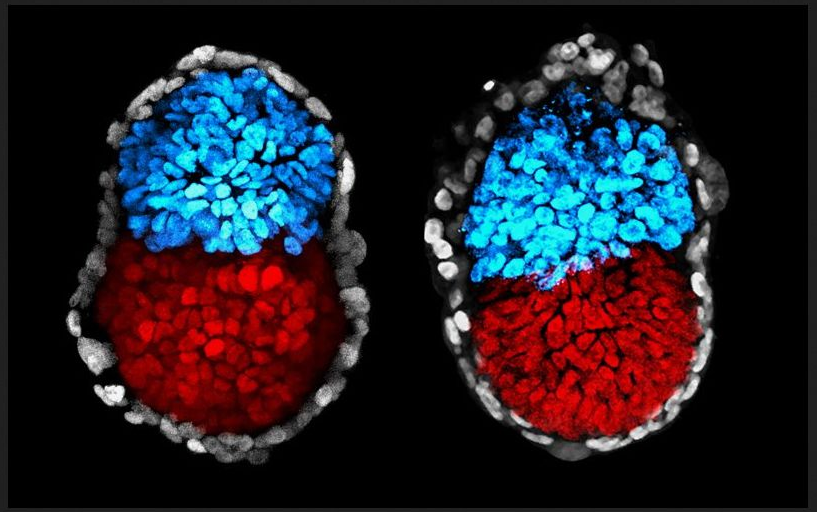
response

3) RESTful service for collaborators

/get method

```
gemma@maanvis:~/Downloads$ curl http://localhost:8080/trialyy/get?id=11 -o e
mbryomodelsdebate.png
  % Total    % Received % Xferd  Average Speed   Time    Time     Time  Curr
ent
                                 Dload  Upload   Total   Spent    Left  Spee
d
  0     0    0     0    0     0      0      0 --:--:-- --:--:-- --:--:--
100  461k  100  461k    0     0  1002k      0 --:--:-- --:--:-- --:--:-- 100
2k
```


Mozilla Firefox
localhost:8080/trialy
localhost:8080/trialyy/get?author=Nicolas Rivron
143%
JSON
Raw Data
Headers
Save
Copy
Collapse All
Expand All
0:
id: 5
author: "Nicolas Rivron"
title: "Debate ethics of embryo models from stem cells"
link: "https://www.nature.com/articles/d41586-018-07663-9"
1:
id: 11
author: "Nicolas Rivron"
title: "Debate ethics of embryo models from stem cells"
link: "https://www.nature.com/articles/d41586-018-07663-9"
https://www.nature.com/articles/d41586-018-07663-9



4) Conclusion

- Application that allows storing, querying and retrieving of images and corresponding metadata
- Interact with SQLite database
- Easily accessible via RESTful API

5) Possible Improvements

- Addition of more metadata information
  - ontology controlled keywords
- Enhanced search with composite of meta parameters
- Image class assignment with group 6
- Allow partial search
- Front-end for API store method

5) Further information

Please see our:

- gitlab page: https://gitlab-sysprog.informatik.uni-bonn.de/ProgrammingLab2/winterterm-2019-20/group-03-descartes/tree/master (with permissions)
- README files and Javadocs

5) Acknowledgements

Dr. Jens Dorpinghaus

Dr. Sebastian Schaaf

2) Database from metadata files and images

SQLite Database Format

- self-contained
- serverless
- zero-configuration
- transactional
- portable

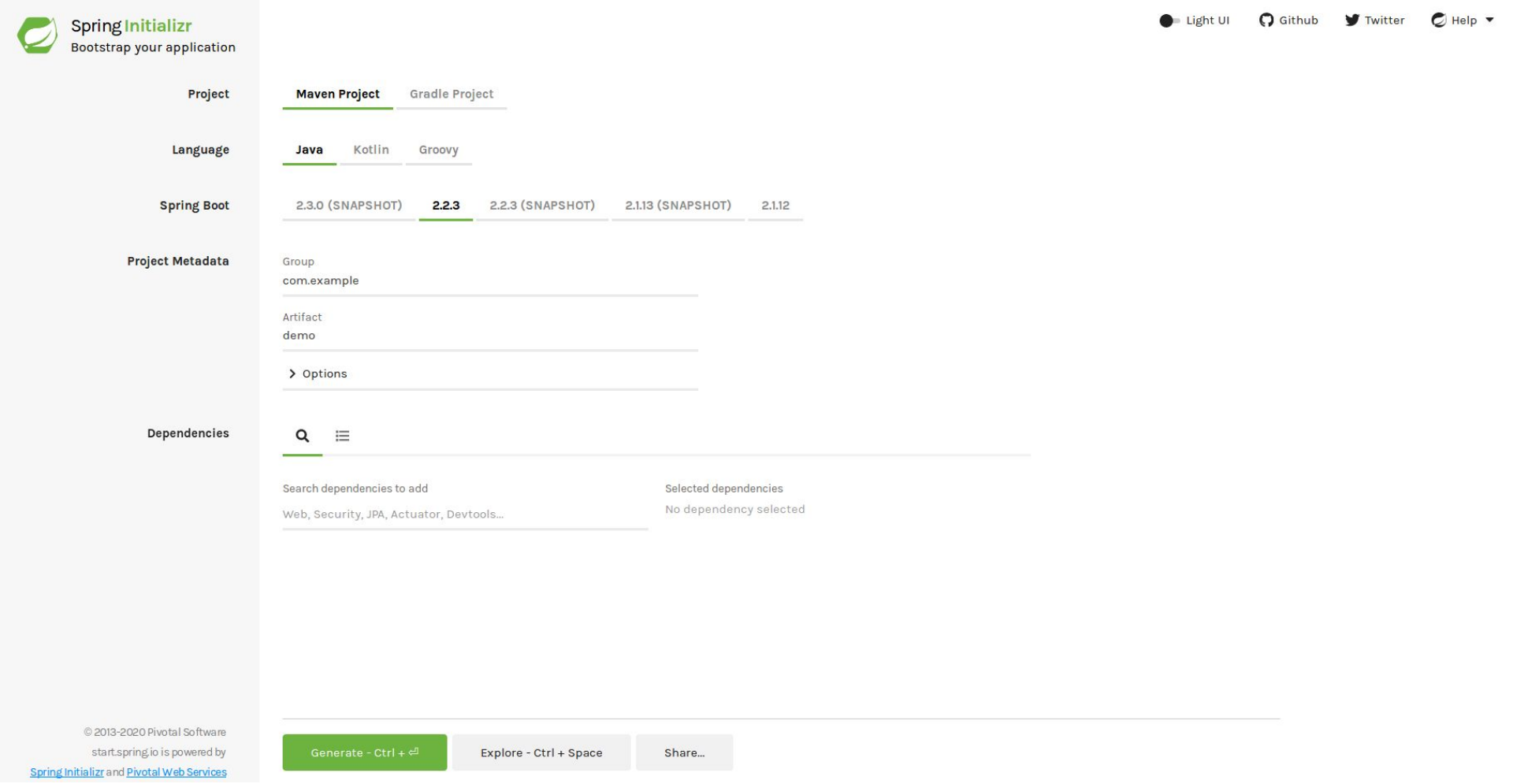
3) RESTful service for collaborators

Tool: Spring boot

Accelerated application development for Java

- build a variety of applications
- automatic configuration with Spring initializr
- curated dependencies
- good starter guides and documentation

"Spring.io." https://spring.io/. Accessed 17 Jan. 2020.



"Spring Initializr." https://start.spring.io/. Accessed 17 Jan. 2020.