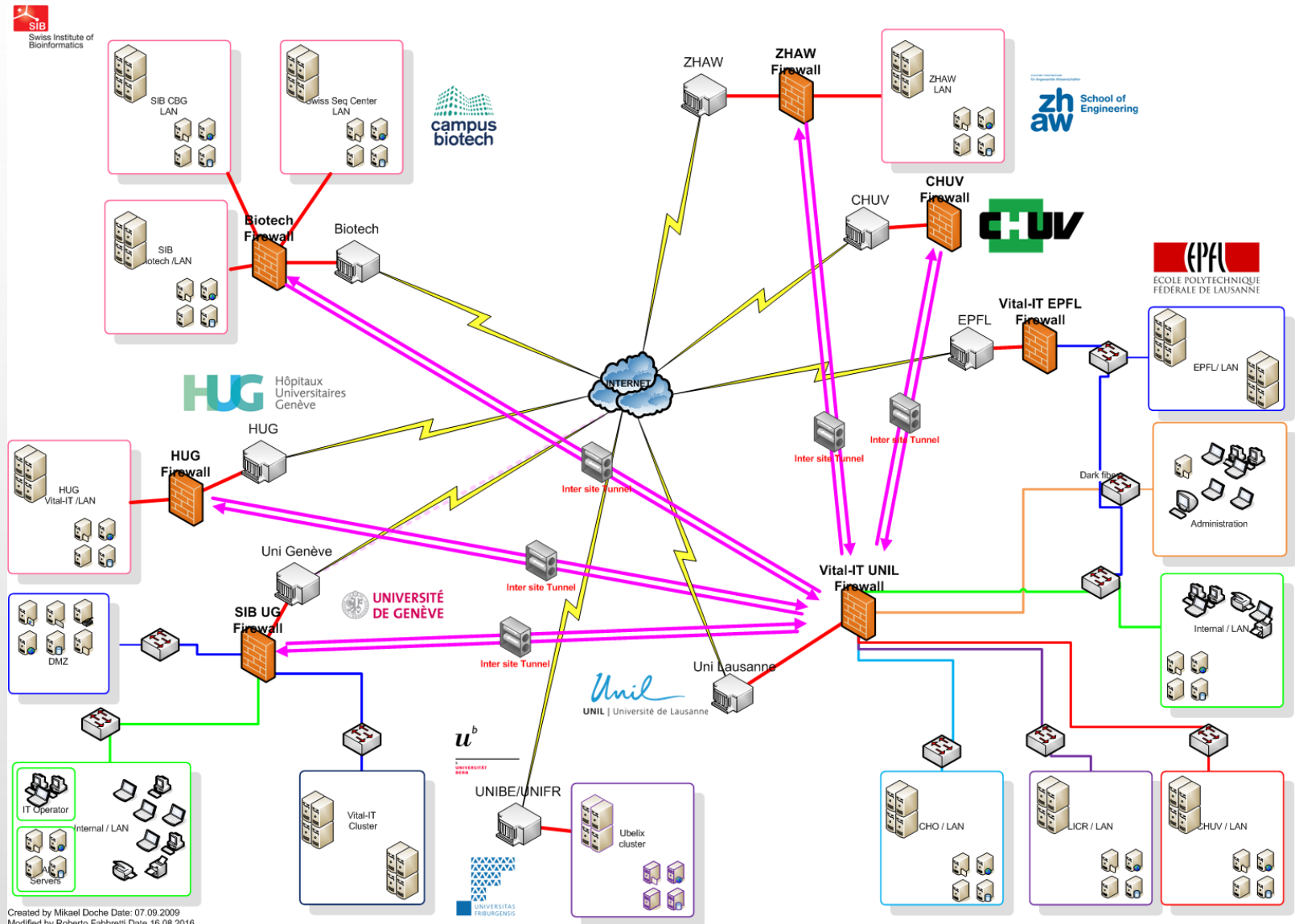


A Data Science's Perspective of the Use of Containers.

Bioinformatics x Digital humanities

Vital-IT, lead by **I. Xenarios**, Competency Centre in Bioinformatics and Computational Biology that provides **infrastructure, support and technological R&D** for life science and clinical research.

Vital-IT Embassies



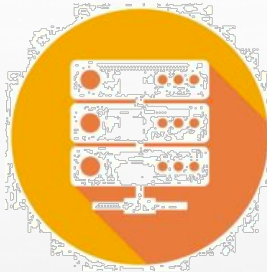
Vital-IT Key Numbers



77 scientists

12 seniors

1 COO



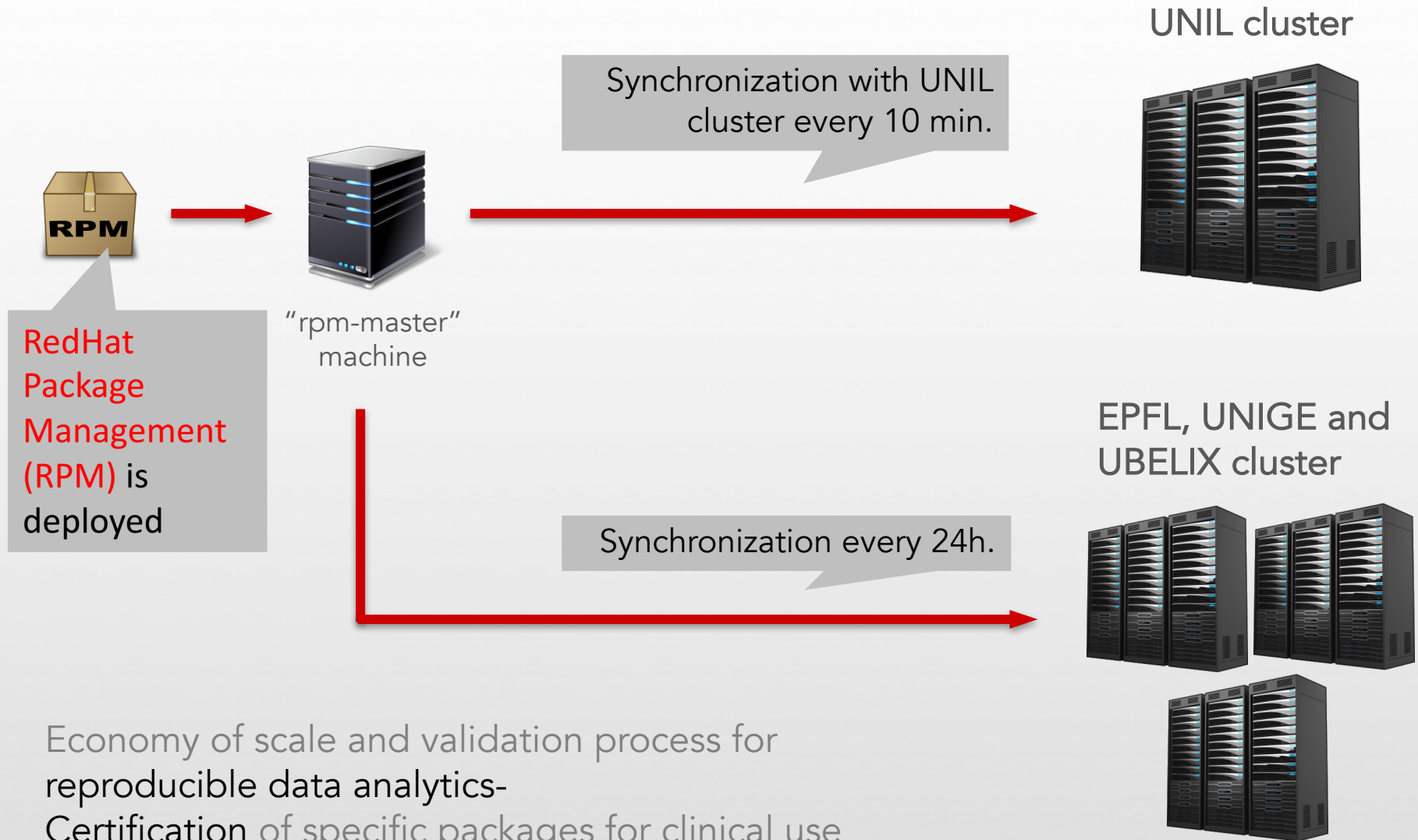
14 Life-Science
tech. platforms
and "core"
facilities entrust
Vital-IT with their
data



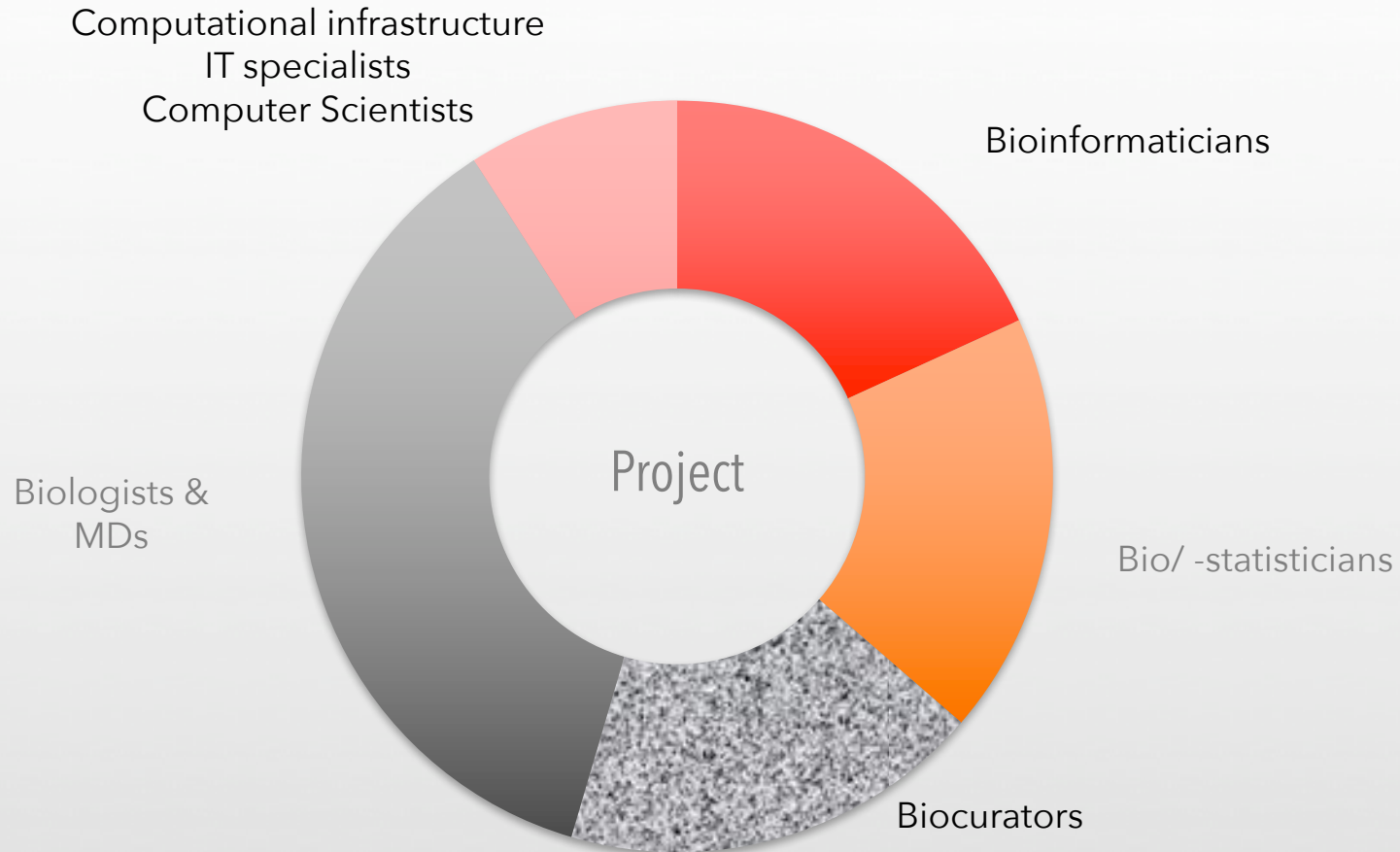
2'500 packaged
and
verified/certified
softwares
maintained as
RPM release

Maintaining and deploying “Fedora-like” Scientific packages

Enabling reproducible science



Industry-like approach to research projects



Various projects – diversity of competencies

Wide range of
Issues

Solving collaborators/customers Issues

Portability of pipeline/tools
with/wo complex setup

, Reproducibility

, Interoperability and Safety

& Others.

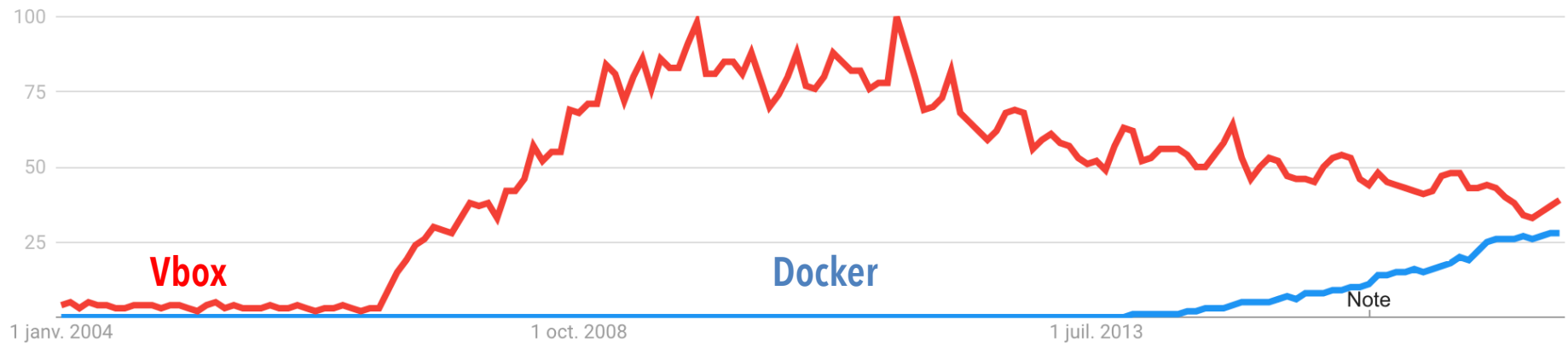
“

PACKAGE THE APPLICATION INTO A
STANDARDIZED UNIT FOR SOFTWARE
DEVELOPMENT

Source : <https://www.docker.com/what-docker>



Docker trends

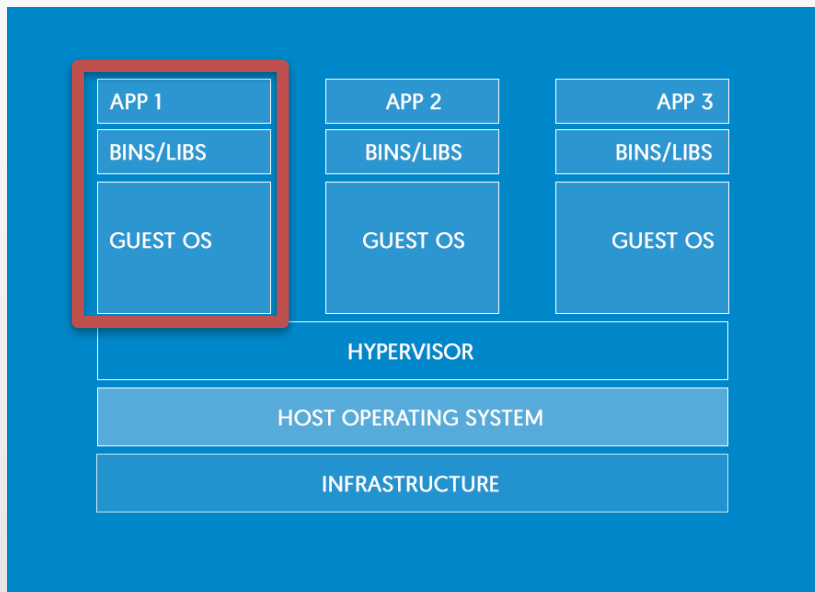


Source : google trends

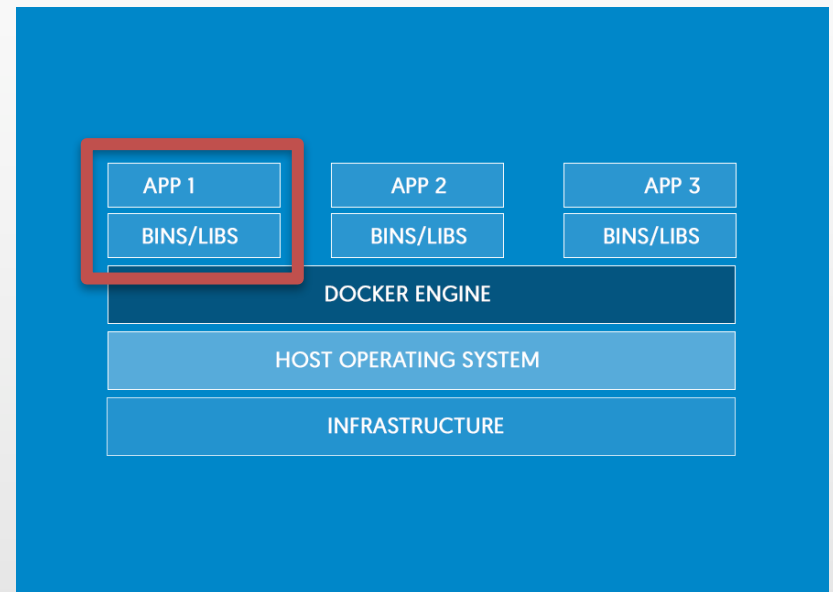
Focus on Application

Docker vs classic Vbox

Virtual machine



Docker Container



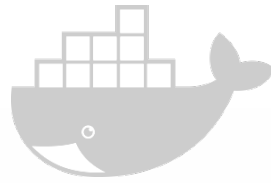
Key advantages 

Lightweight

,Versatile

~~, & "Secure by default".~~

Drawbacks



Security breach

, MPI not supported

& HPC is not the target segment.

Singularity

Solving collaborators/customers Issues

Portability of pipeline/tools
with/wo complex setup

, Reproducibility

, Interoperability and Safety

& Others.

Portability

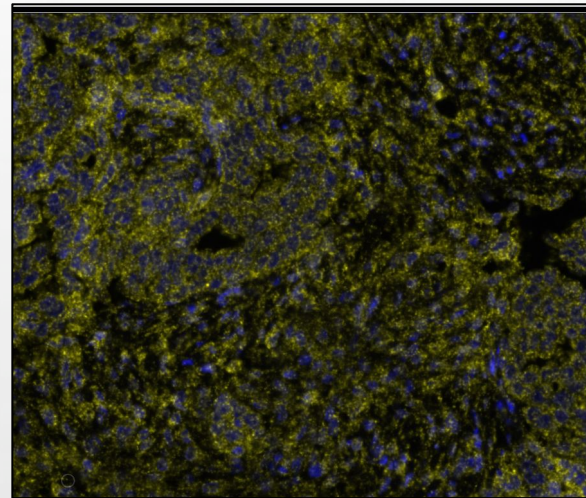
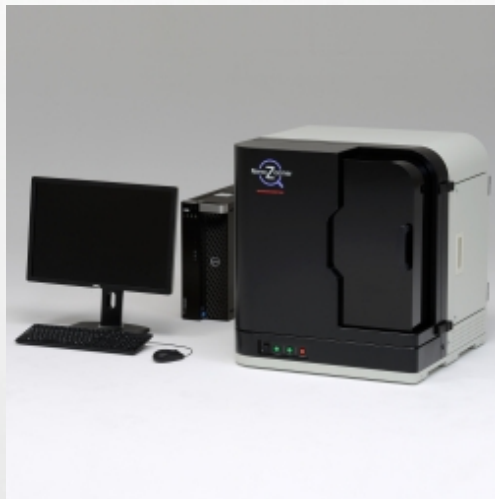
Center of Experimental Pathology Collaboration

With O. Martin, R. Engler & N. Guex

Collaboration : Rusakiewicz S.(CHUV) & Danenberg E.
(CHUV)

NDP.READ

SDK used to access image data in proprietary format (produced by the Hamamatsu NanoZoomer digital slide scanner).



Issue : NDP.Read SDK only for specific Windows version and Linux.

NDP.READ

Dockerfile

```
FROM debian:jessie
MAINTAINER Martial Sankar "martial.sankar@sib.swiss"
RUN apt-get update
RUN apt-get install -y build-essential
RUN apt-get install -y libc6-dev
RUN apt-get install -y libc6-dev-i386
RUN apt-get install -y gcc-multilib g++-multilib
RUN apt-get install -y vim
RUN apt-get clean
RUN mkdir -p software
COPY NDP.read_2.0.4 software
WORKDIR /software/sample/src/
RUN ["make", "x64"]
ENTRYPOINT ["/software/sample/src/NDPread2Sample"]
```

```
$ docker run -it raw/ndpreadraw
```


NDP.READ

Dockerfile

```
FROM debian:jessie
MAINTAINER Martial Sankar "martial.sankar@sib.swiss"
RUN apt-get update
RUN apt-get install -y build-essential
RUN apt-get install -y libc6-dev
RUN apt-get install -y libc6-dev-i386
RUN apt-get install -y gcc-multilib g++-multilib
RUN apt-get install -y vim
RUN apt-get clean
RUN mkdir -p software
COPY NDP.read_2.0.4 software
WORKDIR /software/sample/src/
RUN ["make", "x64"]
ENTRYPOINT ["/software/sample/src/NDPread2Sample"]
```

```
$ docker run -it raw/ndpreadraw
```

Cross-platform,
Bypassing installation procedure.

SourceData

Making Figures Discoverable

With A. Chasapi, J. Delafontaine, L. Gotz, I. Crespo, I. Xenarios & R. Liechti

Collaboration : T. Lemberger (EMBO), E. B. Garagorri (EMBO), N. George (EMBO), S. El-Gebali (EMBO), E. Boutet (Swiss-Prot), N. Nospikel (Swiss-Prot)



Robert Bosch **Stiftung**

Collaborative effort

nature publishing group



HighWire

WILEY



EMBOpress

Figure centered- data curation

Curation by experts & validation by
authors

Data are stored in semantic graph
using neo4j.

Mining the source data graph database

Building sd-graph from scratch using the SourceData API

Install neo4j according to the instructions provided at <http://neo4j.com>. IMPORTANT: To be able to run the commands below with the neo4j-tool, please download the [TAR/ZIP distributions](#). The scripts below were tested under neo4j community edition 2.2 and 3.1.4.

To enable the use of the `neo4j-shell` tool, uncomment this line in `neo4j/conf/neo4j.conf` :

```
# Enable a remote shell server which Neo4j Shell clients can log in to.
#dbms.shell.enabled=true
```

If you run into trouble due to insufficient memory for Java, you may have to edit `neo4j/conf/neo4j.conf` and increase the Java heap size:

```
# Java Heap Size: by default the Java heap size is dynamically
# calculated based on available system resources.
# Uncomment these lines to set specific initial and maximum
# heap size.
#dbms.memory.heap.initial_size=512m
#dbms.memory.heap.max_size=512m
```

Install the [Neo4j Python REST Client](#) with

```
pip install neo4jrestclient
```

or

```
easy_install neo4jrestclient
```

Don't forget to launch neo4j

```
neo4j start
```

Before populating the database, set some constraints:

```
neo4j-shell -file SD-constraints.cql
```

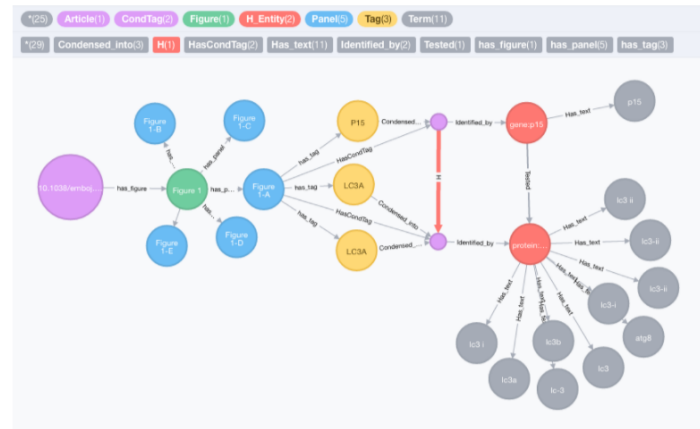
Download the content of the SourceData database through the SourceData API and populate the neo4j database (Warning: this will take a while, so be patient...):

```
python sdneo.py --password <your_password_to_your_neo4j_instance> PUBLICSEARCH
```

Next, build the relationships to create the sd-graph model:

```
neo4j-shell -file SD-processing.cql
```

This will create the following model, linking papers, figures, panels, tags and biological entities:



As a last step, genes and proteins have to be mapped to each other. First, go in the neo4j client in your browser and run this CYPHER command to extract all the Uniprot identifiers:

```
MATCH (t:Tag)
WHERE t.type = "protein" AND t.ext_id <> ""
WITH split(t.ext_id,"/") AS ids
UNWIND ids as id
RETURN DISTINCT id
```

Save the results as csv file (for example to a `export.csv`). Go to <http://www.uniprot.org/uploadlists/> and upload this file to generate a UniProtKB AC/ID to GeneID (Entrez Gene) mapping `protein2gene.tab` (select the "mapping table" format).

Move the `protein2gene.tab` into the `neo4j/import/` directory (this is set in `neo4j.conf` as the default directory for importing files) and build the protein-to-gene mapping with

```
neo4j-shell -file SD-protein2gene.cql
```

Et voilà!

The database is now ready and the examples provided in `SD-examples.cql` can then be run directly in the neo4j browser client.

PUBLIC REPOSITORY

sourcedata/neo4j-sd-graph ☆

Last pushed: 2 months ago

Repo Info

Tags

Short Description

SourceData is a platform that makes scientific papers discoverable based on their data content.

Full Description

This docker contains a neo4j database of sourcedata.

To run it, type:

```
docker run -d --publish=7474:7474 --publish=7687:7687 sourcedata
```

- Open your web browser at: <http://localhost:7474/browser/>
- If necessary log-in with user: `neo4j` and password: `1234`

More info and analysis examples at: <https://github.com/source-data>

Docker Pull Command



```
docker pull sourcedata/neo4j
```

Owner



sourcedata



<https://hub.docker.com/r/sourcedata/neo4j-sd-graph/>

Reproducibility

SIGNAL INTEGRATION IN PLANTS: UV-B, SHADE AND BIOTIC STRESS

With E. Schmit

Collaboration : C. Fankhauser, E. Farmer & R. Ulm



FONDS NATIONAL SUISSE
SCHWEIZERISCHER NATIONALFONDS
FONDO NAZIONALE SVIZZERO
SWISS NATIONAL SCIENCE FOUNDATION

Specs

To build an interface that would
allow to easily :

Mine data

,Visualize data

,Flexibility.

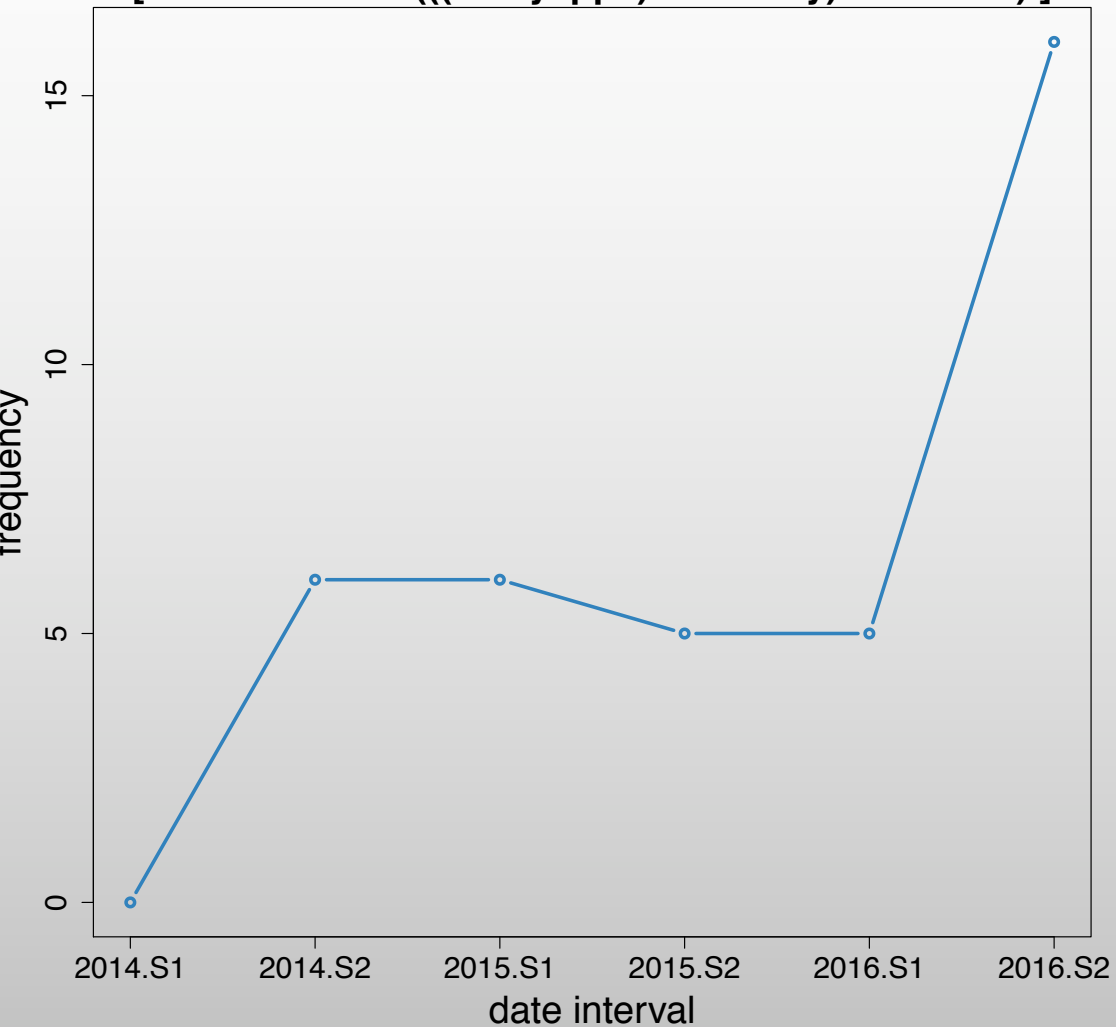
Easy web applications in R



Shiny is an open source R package that provides an elegant and powerful **web framework** for building web applications **using R**. Shiny helps you turn your analyses into interactive web applications **without requiring HTML, CSS, or JavaScript knowledge**.

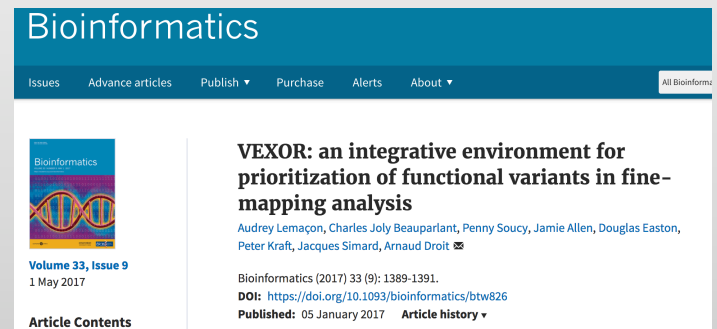
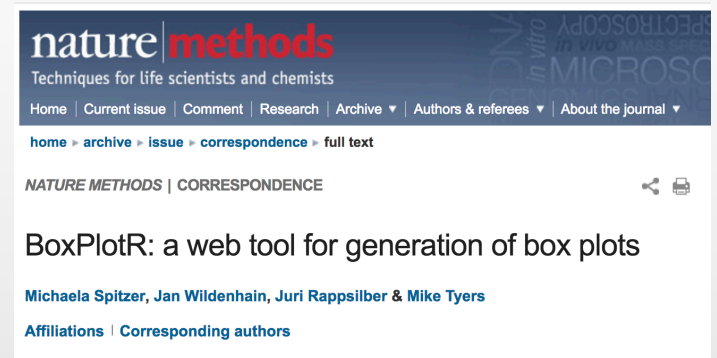
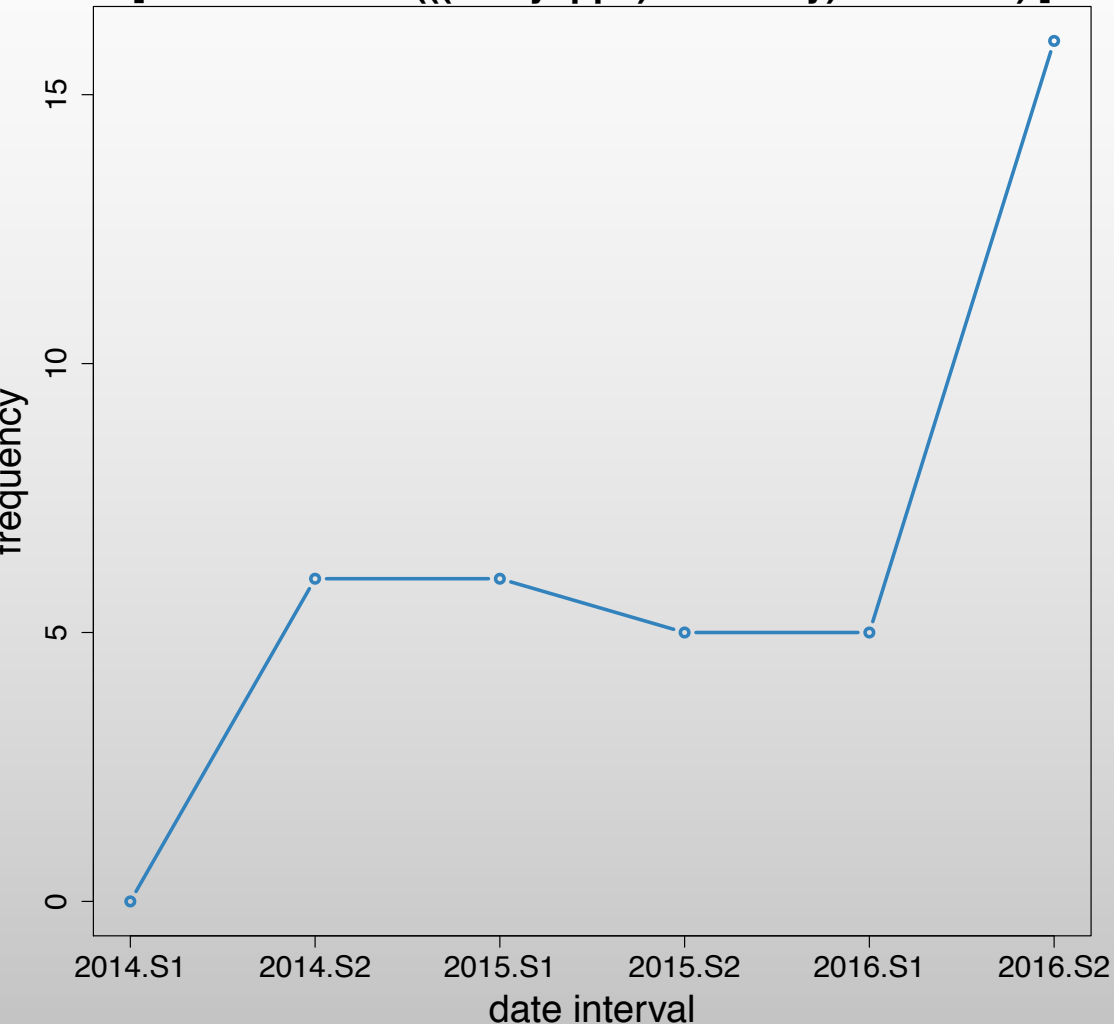
Trendy

Frequency of Shiny-based apps in pubmed
[search terms : (((shinyapps) OR shiny) AND web)]



Trendy

Frequency of Shiny-based apps in pubmed
[search terms : (((shinyapps) OR shiny) AND web)]



Who ?

Target Segment

Who ?

Target Segment

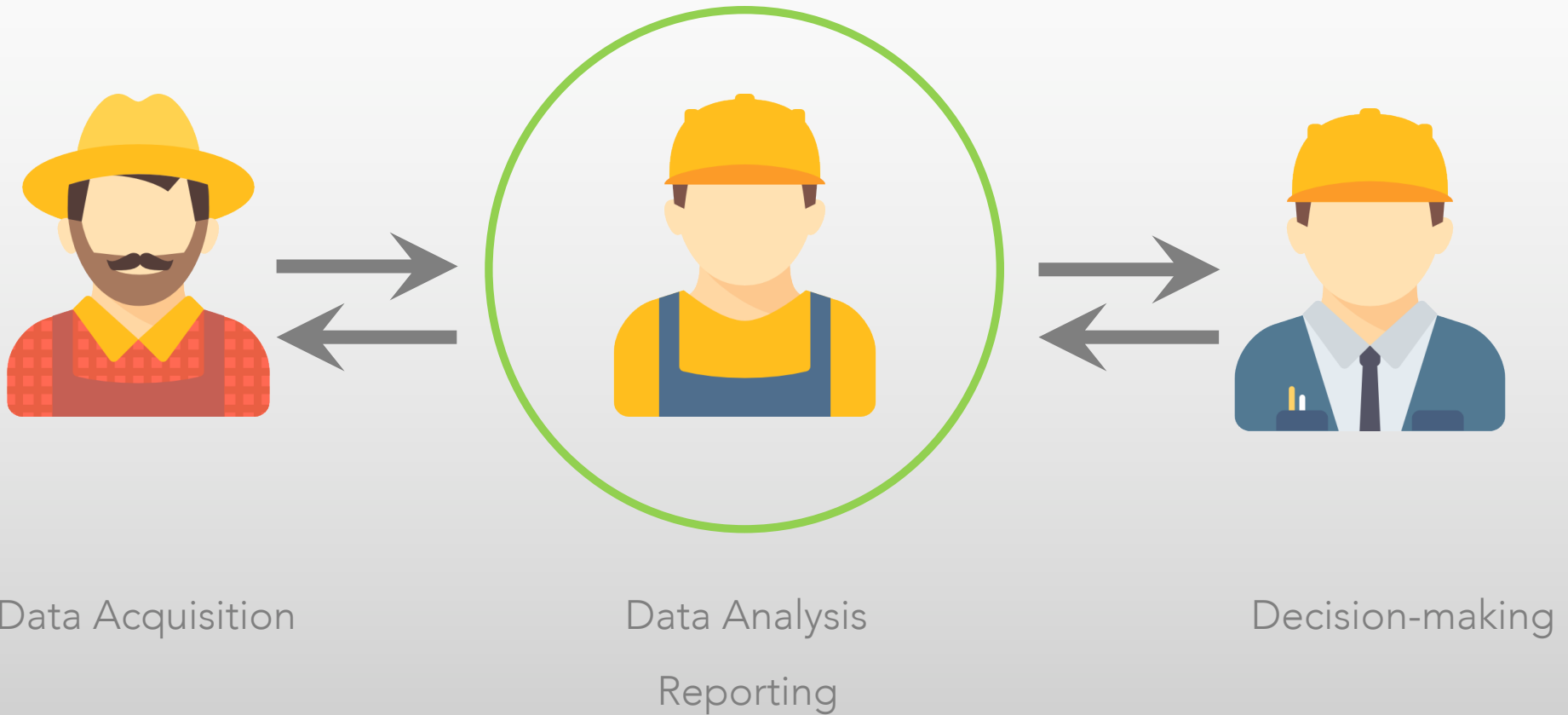


Data Analysis

Reporting

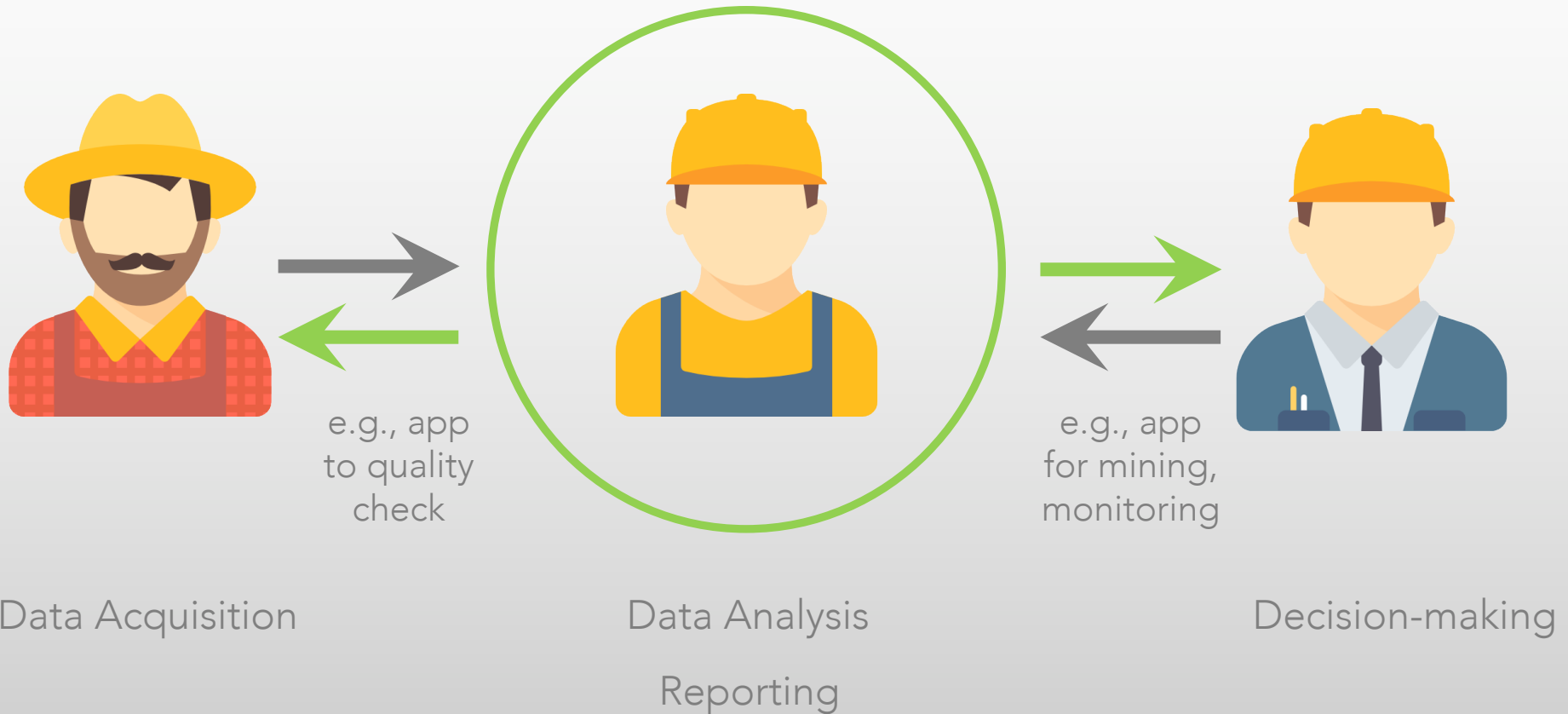
Who ?

Target Segment



Who ?

Target Segment



Why/when ?

Key advantages

- Interactive visualization
- , Reporting
- , Provide dashboard (e.g, QC)
- , Prototyping.

Shinyapps.io pricing

FREE

\$0 /month

New to Shiny? Deploy your applications for FREE.

5 Applications

25 Active Hours

✔ Community Support

❗ RStudio Branding

Sign Up Now

STARTER

\$9 /month
(or \$100/year)

More applications. More active hours!

25 Applications

100 Active Hours

✔ Premium Support

Sign Up Now

BASIC

\$39 /month
(or \$440/year)

Take your users to the next level!

Unlimited Applications

500 Active Hours

✔ Performance Boost

✔ Premium Support

Sign Up Now

STANDARD

\$99 /month
(or \$1,100/year)

Password protection? Authenticate your users!

Unlimited Applications

2,000 Active Hours

✔ Authentication

✔ Performance Boost

✔ Premium Support

Sign Up Now

PROFESSIONAL

\$299 /month
(or \$3,300/year)

Professional has it all! Personalize your domains.

Unlimited Applications

10,000 Active Hours

✔ Authentication

✔ Account Sharing

✔ Performance Boost

✔ Custom Domains

✔ Premium Support

Sign Up Now

All prices are in USD and must be paid with a valid credit card. Taxes may apply.

See our [Terms of Use](#) for more information

Publish Shiny applications on RStudio Shiny Server Pro.

RStudio Shiny Server Pro

\$9,995 per year

BUY NOW

Concurrent Users

20

Upgrade Path

Add Concurrent User packs of
20 for \$4,995 and 150 for \$14,995

Commercial License (not AGPL)



RStudio Support



Deploy Shiny apps and interactive documents to the internet



Controlled access via SSL and LDAP, Active Directory, Google OAuth, PAM, proxied authentication, or passwords



Tune & Scale applications across multiple processes



Metrics & Session Management



Pricing

\$9,995/yr

BUY NOW

ShinyProxy
 About
 Getting Started
 Deploying Apps
 Configuration
 Security
 Usage Statistics
 Downloads
 Troubleshooting
 Support

About

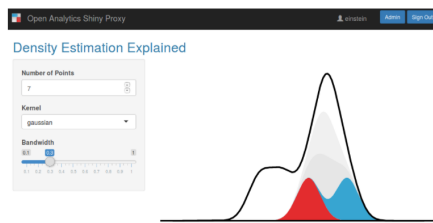


Table of contents

What is ShinyProxy?
 Why use it?
 Open Source
 Java Server Side
 Docker-based technology
 Open Source Shiny Package

What is ShinyProxy?

ShinyProxy is your favourite way to **deploy Shiny apps in an enterprise context**. It has built-in functionality for LDAP authentication and authorization, makes securing Shiny traffic (over TLS) a breeze and has no limits on concurrent usage of a Shiny app.

Why use it?

- you want to **seamlessly deploy Shiny apps** that were developed locally using the Shiny R package
- you need **enterprise features** but want to stay with **open source**
- you trust **Java** on the **server side** for running your Shiny apps
- you want to get all benefits offered by **Docker**-based technology

Open Source

ShinyProxy is 100% open source, released under the [Apache License version 2.0](#) and the sources are on [Github](#).



Java Server Side

ShinyProxy uses time-tested and mature enterprise Java technology bundled nicely as a [Spring boot](#) web application.



Why use it?

Open Source
 Java Server Side
 Docker-based technology
 Open Source Shiny Package

Docker-based technology

When deploying a Shiny application with ShinyProxy, the application is simply bundled as an R package and installed into a Docker image. Every time a user runs an application, a container spins up and serves the application. This has numerous advantages:

- fully isolated 'workspace' per session
- plug and play different docker images (even with different R versions or different Shiny versions)
- control on memory and cpu usage via the Docker API
- monitoring and debugging using standard Docker tooling



Open Source Shiny Package

The ShinyProxy only uses functionality that is available in the open source [R package](#) and does not rely on any server version of Shiny. This allows for a clean separation between the statistical/business logic in R and the management of the applications in ShinyProxy.



ShinyProxy
About
Getting Started
Deploying Apps
Configuration
Security
Usage Statistics
Downloads
Troubleshooting
Support

About

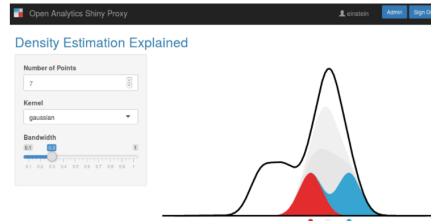


Table of contents

What is ShinyProxy?
Why use it?
Open Source
Java Server Side
Docker-based technology
Open Source Shiny Package

What is ShinyProxy?

ShinyProxy is your favourite way to **deploy Shiny apps in an enterprise context**. It has built-in functionality for LDAP authentication and authorization, makes securing Shiny traffic (over TLS) a breeze and has no limits on concurrent usage of a Shiny app.

Why use it?

- you want to **seamlessly deploy Shiny apps** that were developed locally using the Shiny R package
- you need **enterprise features** but want to stay with **open source**
- you trust **Java** on the **server side** for running your Shiny apps
- you want to get all benefits offered by **Docker**-based technology

Open Source

ShinyProxy is 100% open source, released under the [Apache License version 2.0](#) and the sources are on [Github](#).



Getting Started
Deploying Apps
Configuration
Security
Usage Statistics
Downloads
Troubleshooting
Support

Java Server Side

ShinyProxy uses time-tested and mature enterprise Java technology bundled nicely as a [Spring boot](#) web application.



Why use it?
Open Source
Java Server Side
Docker-based technology
Open Source Shiny Package

Docker-based technology

When deploying a Shiny application with ShinyProxy, the application is simply bundled as an R package and installed into a Docker image. Every time a user runs an application, a container spins up and serves the application. This has numerous advantages:

- fully isolated 'workspace' per session
- plug and play different docker images (even with different R versions or different Shiny versions)
- control on memory and cpu usage via the Docker API
- monitoring and debugging using standard Docker tooling



Open Source Shiny Package

The ShinyProxy only uses functionality that is available in the open source [R package](#) and does not rely on any server version of Shiny. This allows for a clean separation between the statistical/business logic in R and the management of the applications in ShinyProxy.



ShinyProxy

Advantages of a container-based proxy

fully isolated 'workspace' per session

,plug and play different docker images (even with different R versions or different Shiny versions)

,control on memory and CPU usage via the Docker API

& monitoring and debugging using standard Docker tooling.


```
$ docker stats --all --format ...
```

CONTAINER ID	NAME	CPU %	MEM USAGE / LIMIT	MEM %
e10366b5e03caa43b7bfd4097085a343f06d693ce789471da7dde6feb354dff6	eager_goldstine	0.14%	142.6MiB / 1.938GiB	7.18%
7462db9ce9d55b2d724402898a0b2ad88e7f11fed9763f117fa4a475e102b6c0	heuristic_hugle	0.00%	0B / 0B	0.00%
2f845d3f302d76fbaee23bfb6401d2e6d78bf183841e83d69d1e20ba4dd009dd	vigorous_rosalind	0.00%	0B / 0B	0.00%
eee0453a9e27c30208e0e3ab47e20e6ce80d09cae0cb479055caefbc319e5d32	dreamy_lalande	0.09%	123.6MiB / 1.938GiB	6.23%

Please sign in:

User name
Password

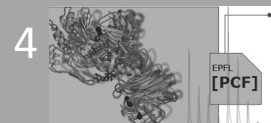
Sign In

Data Interoperability And **safety**

MegaClust IaaS

With R. Liechti, I. Xenarios & N. Guex

Collaboration : A. Harari¹, P. Baumgartner¹,
G. Coukos¹, F. Periklis¹, E. Danenberg¹, P. Meylan², J.
Faget², C. Schwärzler³, Miguel Garcia³, M. Moniatte⁴.



Flow cytometry data analysis

The core of MegaClust implements a **high-performance unsupervised classification algorithm** allowing fast and robust prediction of cell-type population.

No human bias & reproducible



Designed to handle large datasets

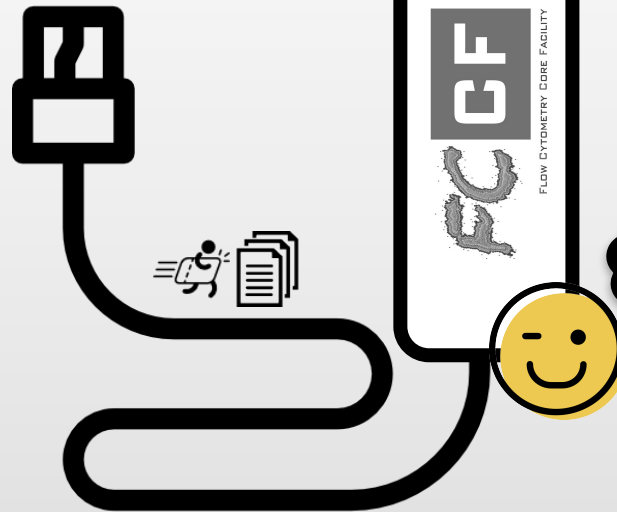
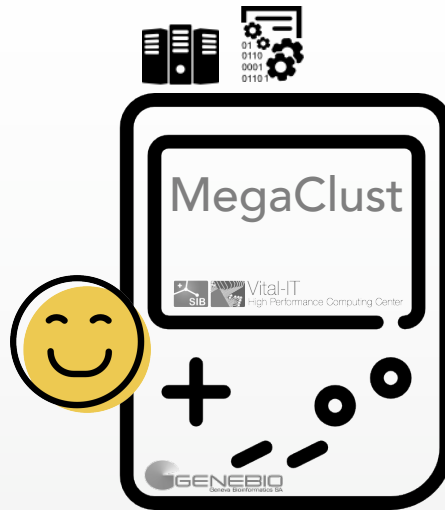


Scale, up to 60 markers (CyTOF)



B2B model

To limit
maintenance cost
of the solution



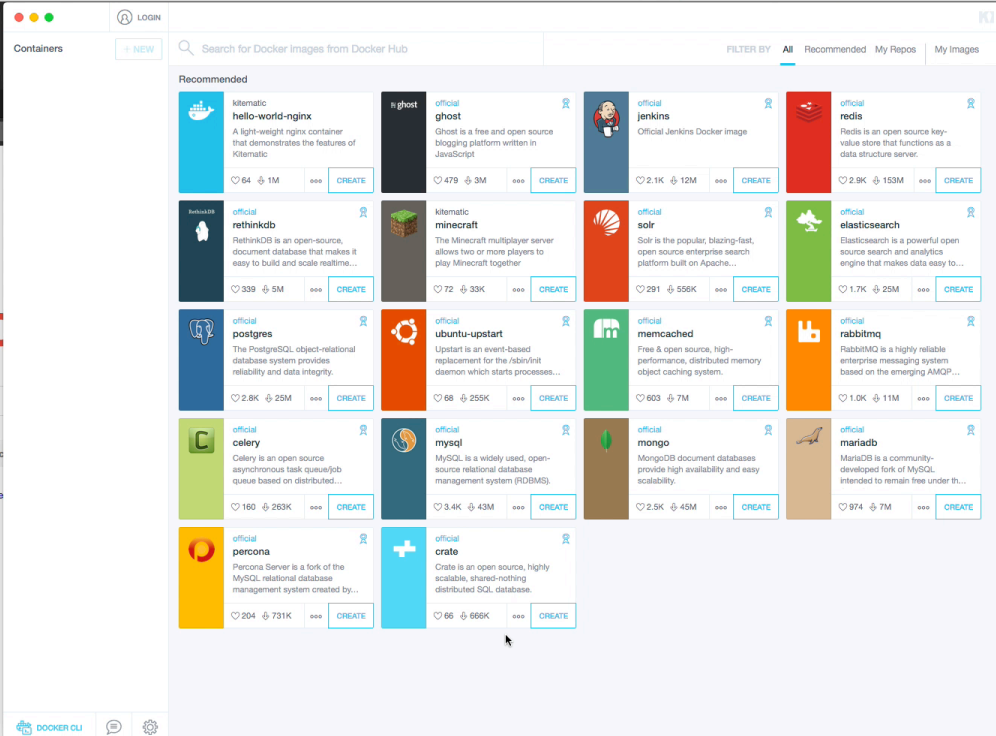
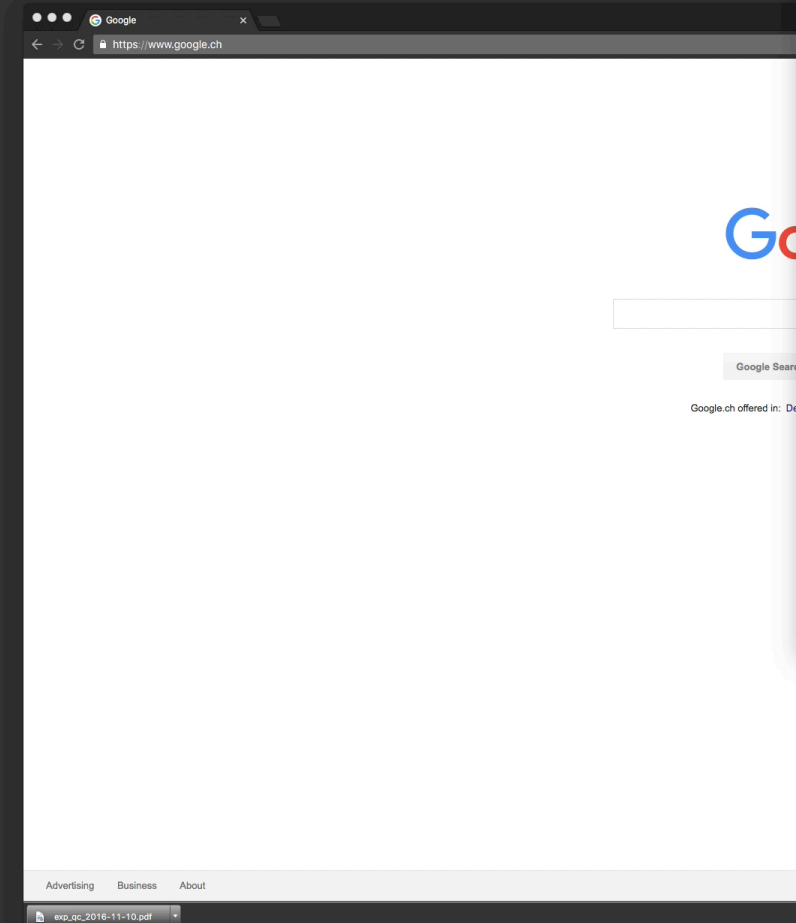
Fast release
Reproducible



To ensure of the quality of
customers' data

NO data transfer to third-party server

Independent of the MegaClust
clustering



Towards a new editorial
model?

Vital-DH@Vital-IT

With A.Chasapi, S. Schulthess I. Xenarios & C.Clivaz



FONDS NATIONAL SUISSE
SCHWEIZERISCHER NATIONALFONDS
FONDO NAZIONALE SVIZZERO
SWISS NATIONAL SCIENCE FOUNDATION

Funded by the
Erasmus+ Programme
of the European Union



Humanities

“

... the study of how people **process and document the human experience.**

Since humans have been able, we have used philosophy, literature, religion, art, music, history and language **to understand and record our world.**

Stanford Humanities | Stanford Universities



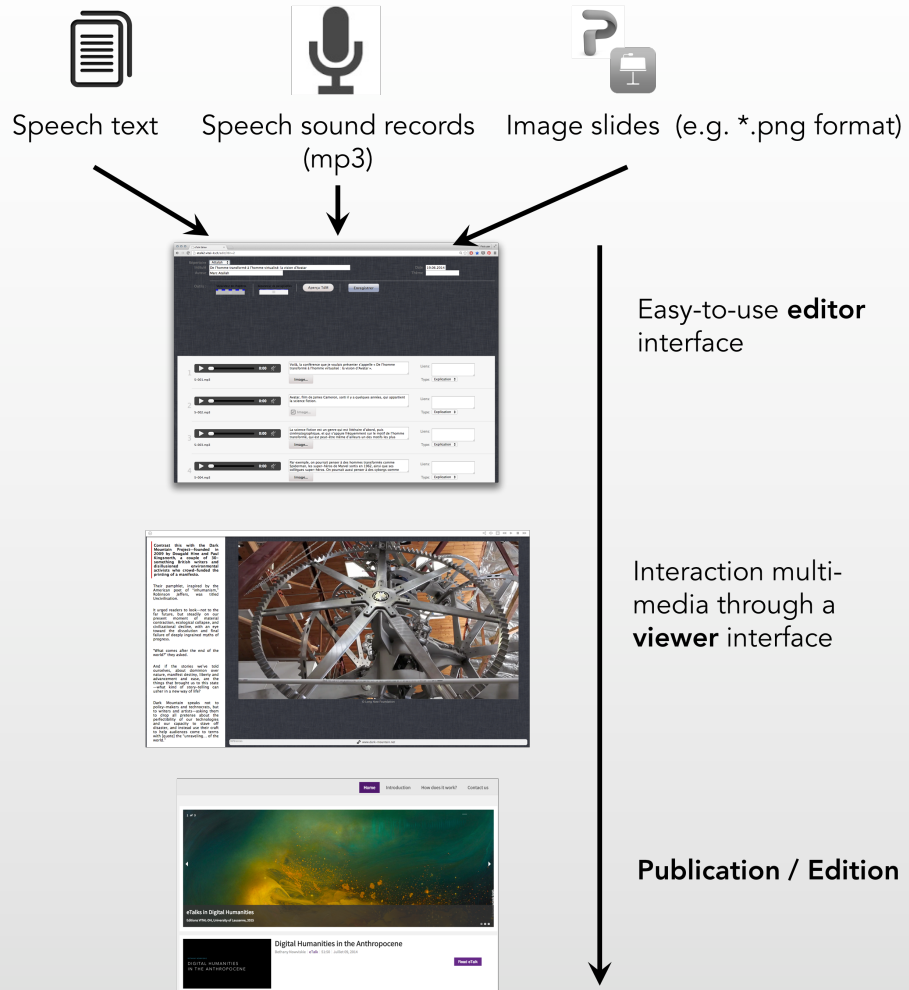
The eTalks

A new digital multimedia editing platform

By Vital-DH

[Find Out More](#)

Text, image and sound together



4

Fully edited publication series,

21

Authors,

24

eTalks

767

Minutes of recorded speeches.


Personalized medicine with F. Schütz

La médecine personnalisée est un sujet dont on parle beaucoup actuellement, en particulier dans les médias.

Si l'on prend simplement l'exemple de Lausanne, il y a de nombreuses activités liées à ce sujet. Qu'elles soient académiques (par exemple à l'Institut Suisse de Bioinformatique) dans des domaines cliniques (par exemple au CHUV), ou dans des sociétés privées.

Et Lausanne n'est qu'un exemple, il y a de telles activités partout en Suisse, que ce soit dans d'autres cantons (par exemple Zurich), ou au niveau de la Confédération qui soutient des initiatives dans ce domaine.

Mais ça peut sembler paradoxal de donner autant d'importance à un sujet tel que la médecine personnalisée. Parce que dans un sens, la médecine a toujours été personnalisée.



**Médecine personnalisée:
Patient vs génétique, risques et
probabilités**

Frédéric Schütz



Swiss Institute of
Bioinformatics

This repository

Search

Pull requests

Issues

Gist

msank / etalk-docker

Unwatch

1

Star

0

Fork

0

Code

Issues 0

Pull requests 0

Projects 0

Wiki

Pulse

Graphs

Settings

No description or website provided. — Edit

2 commits

1 branch

0 releases

0 contributors

Branch: master

New pull request

Create new file

Upload files

Find file

Clone or download

Martial Sankar 0.1

Latest commit b59fa88 on Jul 8

db

0.1

4 months ago

etalk-db

0.1

4 months ago

etalkapp

0.1

4 months ago

README.md

0.1

4 months ago

docker-compose.yml

0.1

4 months ago

README.md

HOW TO CREATE AN ETALK USING THE CONTAINER

This etalk setup is suitable for eduction purpose. It permits to use and test the etalk application on the user's own platform. It uses the docker-compose tool to set-up the application services (etalk php/apache, mysql, phpmyadmin).

FIRST STEP : INSTALL DOCKER and Run the etalk virtual machine (VM)

1. INSTALL docker (free) for your desktop (windows, osx or linux) <https://www.docker.com/products/docker#>
2. Download or clone this repository.
3. Open a terminal in osx or linux or open the "docker quick start terminal" in Windows (that comes along with the installation).
4. Go Inside the directory (use `cd /path/to/directory` more info here [🔗](#))
5. Secondly, build the image with :

```
$ docker-compose build
```

6. Then, run the etalk application with :

```
$ docker-compose up -d
```

7. Open a browser

On osx or linux : go to the url <http://localhost:88> for the **viewer interface** or <http://localhost:88/edit> for the **edit interface**

On windows : go to the url <http://192.168.99.100:88> for the **viewer interface** or <http://192.168.99.100:88/edit> for the **edit interface**

SECOND STEP : MAKE YOU OWN ETALK

To make your own etalk, you can follow the how-to from the etalk *"Make your own etalk"* accessing <http://192.168.99.100:88>

1. Create and name a folder that will contain the mp3 files inside `etalkapp/etalk-master/data/`
2. go to the **edit interface** and start editing your etalk.

© 2016 GitHub, Inc.

[Terms](#) [Privacy](#) [Security](#) [Status](#) [Help](#)

Contact GitHub

API

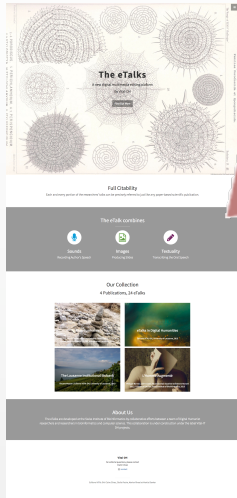
Training

Shop

Blog

About

Towards a Docker-based editorial process (?)



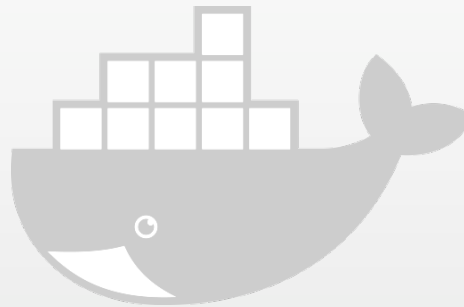
published



App encapsulation

Dev + Edition

Submitted for
editorial process



Pulled / Shipped on request



Author Creation

Summary

Revolution, Evolution ...



Or ...



