

Teaching with CLARIN

Learning by CLARIN

Michał Gawor
Twan Goosen

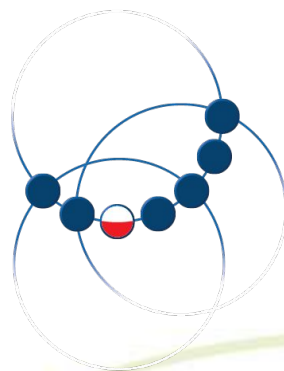




Resources



NLP services



Virtual workbook

Jupyter Notebook - a virtual workbook

- web-based application for creating and sharing computational documents
- +40 programming languages, including **Python**, **R**, **Julia**, and **Scala**
- designed with single instance in mind

XML basics

XML (Extensible Markup Language) is a markup language that does not do anything on its own. Its purpose is to improve machine readability of data. Below you can see an exemplary XML file that we will use in the following cells to demonstrate field value access.

```
[1]: !cat "./example_xml.xml"
```

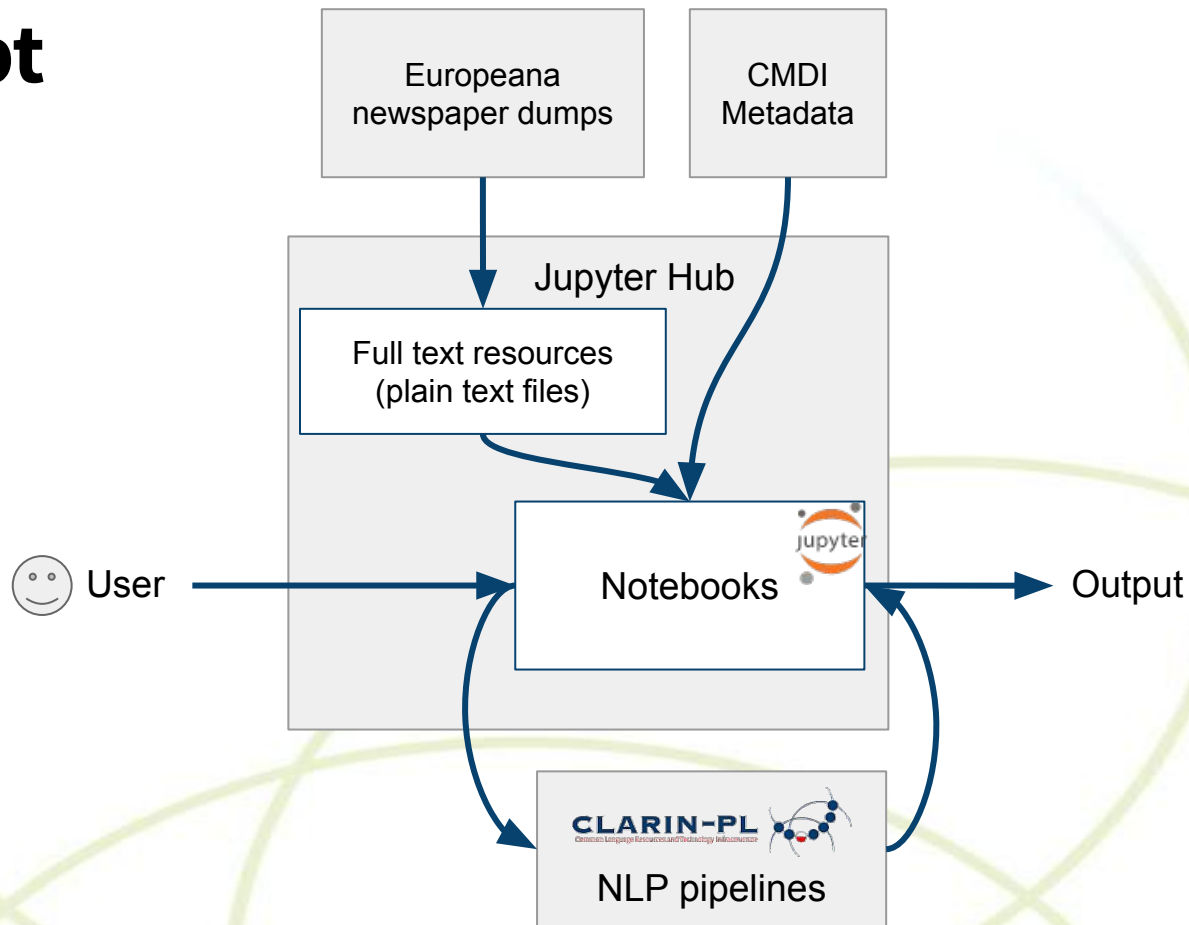
```
<studentsList>
  <student id="1">
    <firstName>John</firstName>
    <lastName>Snow</lastName>
    <scores>
      <homework1>80</homework1>
      <homework2>70</homework2>
      <project>85</project>
    </scores>
  </student>
  <student id="2">
    <firstName>Jozef</firstName>
    <lastName>Szwejk</lastName>
    <scores>
      <homework1>70</homework1>
      <homework2>65</homework2>
      <project>35</project>
    </scores>
  </student>
  <student id="3">
    <firstName>Franek</firstName>
    <lastName>Dolas</lastName>
    <scores>
      <homework1>65</homework1>
      <homework2>80</homework2>
      <project>45</project>
    </scores>
  </student>
</studentsList>
```

JupyterHub - a virtual classroom

Multi-user version of the Notebook

- pluggable authentication (PAM, OAuth, custom)
- centralized deployment
- container friendly (docker, Kubernetes)
- code meets data

Concept



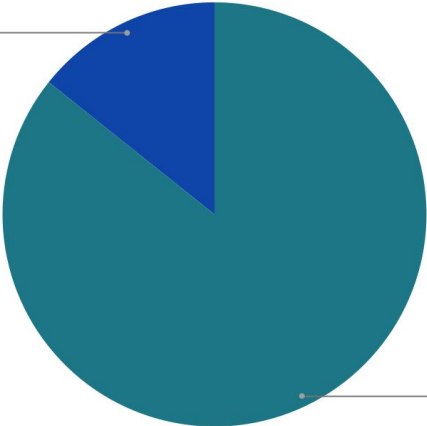
The workshop

- online setup via Zoom, 35 min. tutorial, 35 min. exercises, 10 min. break
- prepared JupyterHub
- 16 participants, 2 break-out rooms for exercises
- each receives credentials to own Notebook
- persistent teaching materials accessible online
- one-click in Binder to launch

Feedback

To what extent are you satisfied with this training?

Completely satisfied
14,3%



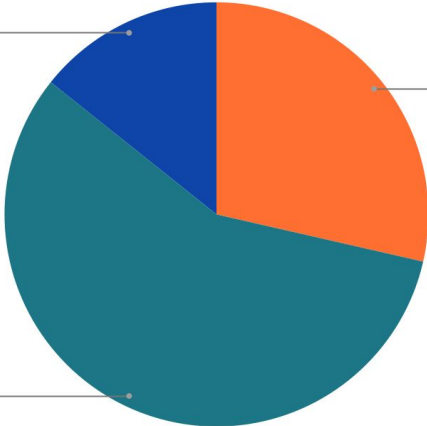
How difficult did you find this workshop to be?

Very difficult
14,3%

Neither difficult nor
28,6%

Satisfied
85,7%

Difficult
57,1%



Lessons learned - internal evaluation

4 main aspects:

- communication
- educational content
- infrastructure
- admission and registration

Lessons learned - Communication

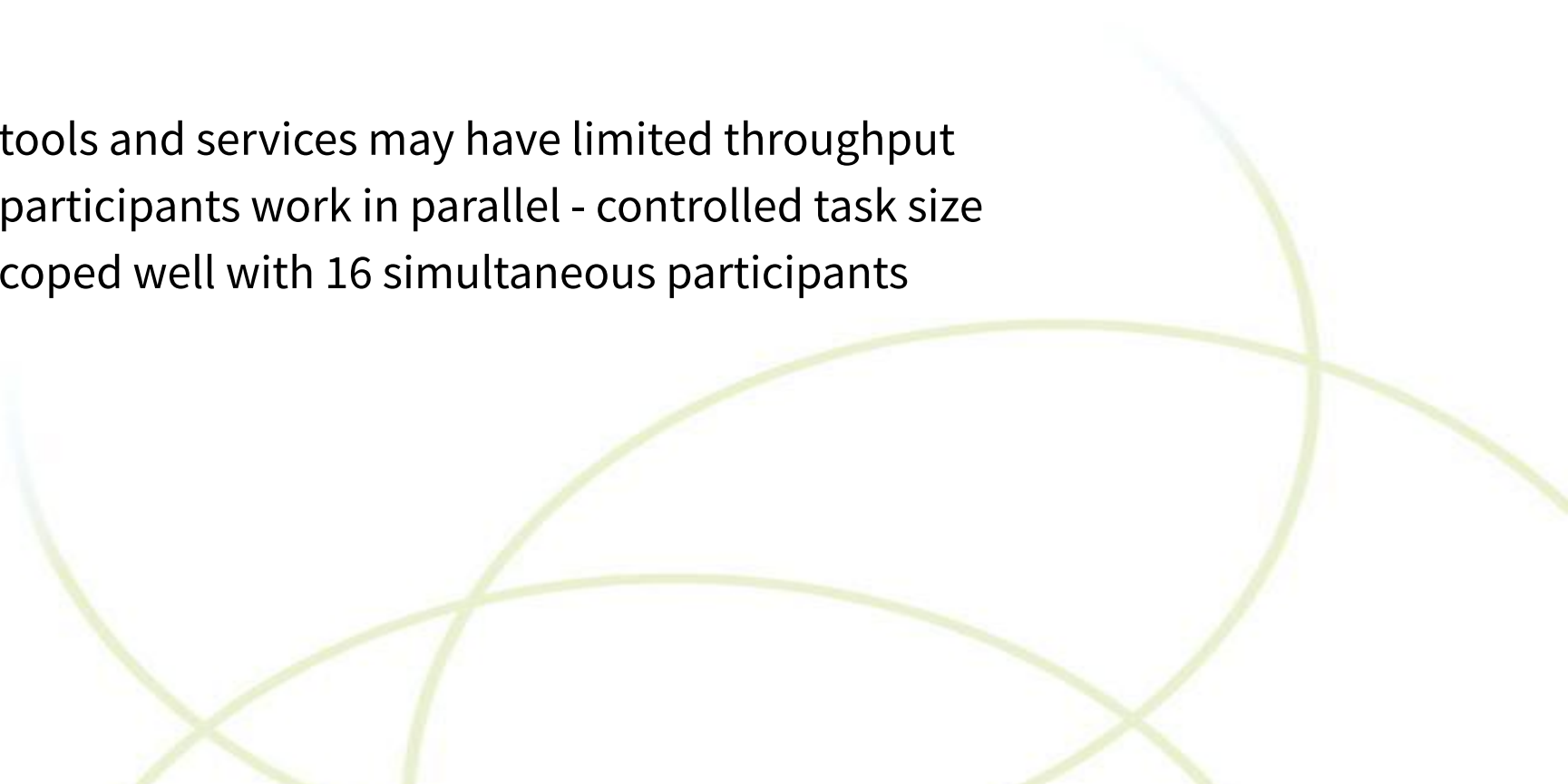
- alignment with audience is a key
 - required skill set
 - recommended skill set
 - end goal
- virtual classroom
 - proactively check on participants

Lessons learned - Educational content

Provide:

- pre-tutorial materials for self-assessment
- links to recommended reading/watching material for recommended skills

Lessons learned - Infrastructure

- tools and services may have limited throughput
 - participants work in parallel - controlled task size
 - coped well with 16 simultaneous participants
- 

Lessons learned - Admission and registration

- high interest with ~300 views of EventBrite page
- change in target audience
- registration for waiting list, last minute fair participation selection at random (day before event)
- at least two users (one via [twitter](#)) reported expecting an invitation but not receiving one

Reusability

- profile target audience
- evaluate material in terms of target audience
- provide a clear and fair description of required and/or recommended pre-existing knowledge and skills
- same (though less critical) for learning goals

Open questions

Could a two part tutorial (with ~1 day in between sessions) be a better approach?

How to encourage/invoke interaction between participants in a virtual setting?

How to arrange long(er) term support for notebooks?

We have our stall at Bazaar!

Our notebook (out-of-the-box with Binder):

<https://github.com/clarin-eric/europeana-newspapers-notebooks>

Screencast:

<https://www.youtube.com/watch?v=FQcunA185N0>

More CLARIN notebooks:

<https://www.clarin.eu/notebooks>