Explainable AI for the molecular subclassification of DLBCL

Tobias Schmidt[1], Michael Altenbuchinger[2] Wolfram Gronwald[3], Rainer Spang[1]

[1] Department of statistical Bioinformatics, University of Regensburg, Germany
[2] Institute of Medical Bioinformatics, University Medical Center Göttingen, Germany
[3] Institute of functional Genomics, University of Regensburg, Germany

Introduction

Diffuse large B-cell lymphoma (DLBCL)

- Neoplasm of medium or large **B lymphoid cells** with a diffuse growth pattern and nuclei at least twice the size of normal lymphocyte nuclei
- Incidence rate: 5.4 / 100.000
- Makes up for **1.2% of all cancers**
- 5-year relative survival: **63.9%** (2011-17)
- Current standard-therapy: R-CHOP
- ≈ 20% of DLBCLs classifiable based on morphology or biological/clinical issues
- ≈ 80% NOS → classified based on molecular features
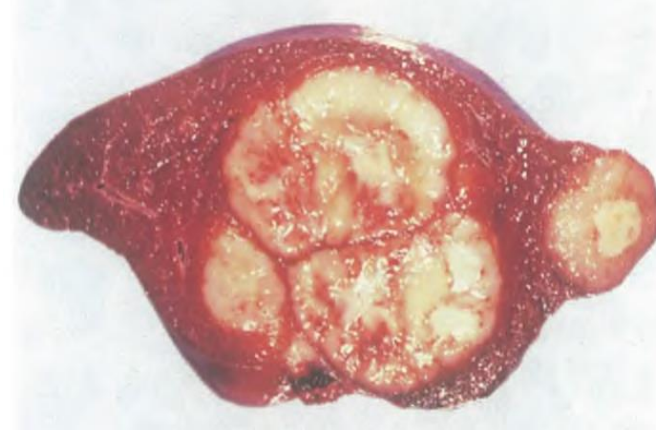
Source: WHO Classification of Tumours and Haematopoietic Tissues by S. Swerdlow et al.

Omics data can be used to classify patients

- 4 Microarray (Rosenwald02, Ship02, Wright03, Lenz08) [12-15]
- 5 Nanostring nCounter (Masque-Soler13, Scott14, Ennishi19, Staiger20, Nordmo20) [7-11]
- 1 SWATH-MS (Renders20) [16]
- 3 WGS/WES/CNA/SNPA/CRISPR-Screening (Reddy17, Chapuy18, Schmitz18) [17-19]
- 8 IHC (Hans04, Muris06, Natkunam08, Nyman09, Choi09, Meyer11) [1-6]

But: only a few are used in clinical practice! [20]

No software implementations for end users

Increasing model complexity leads to loss of interpretability

Methods

- Collection of model specifications and corresponding datasets
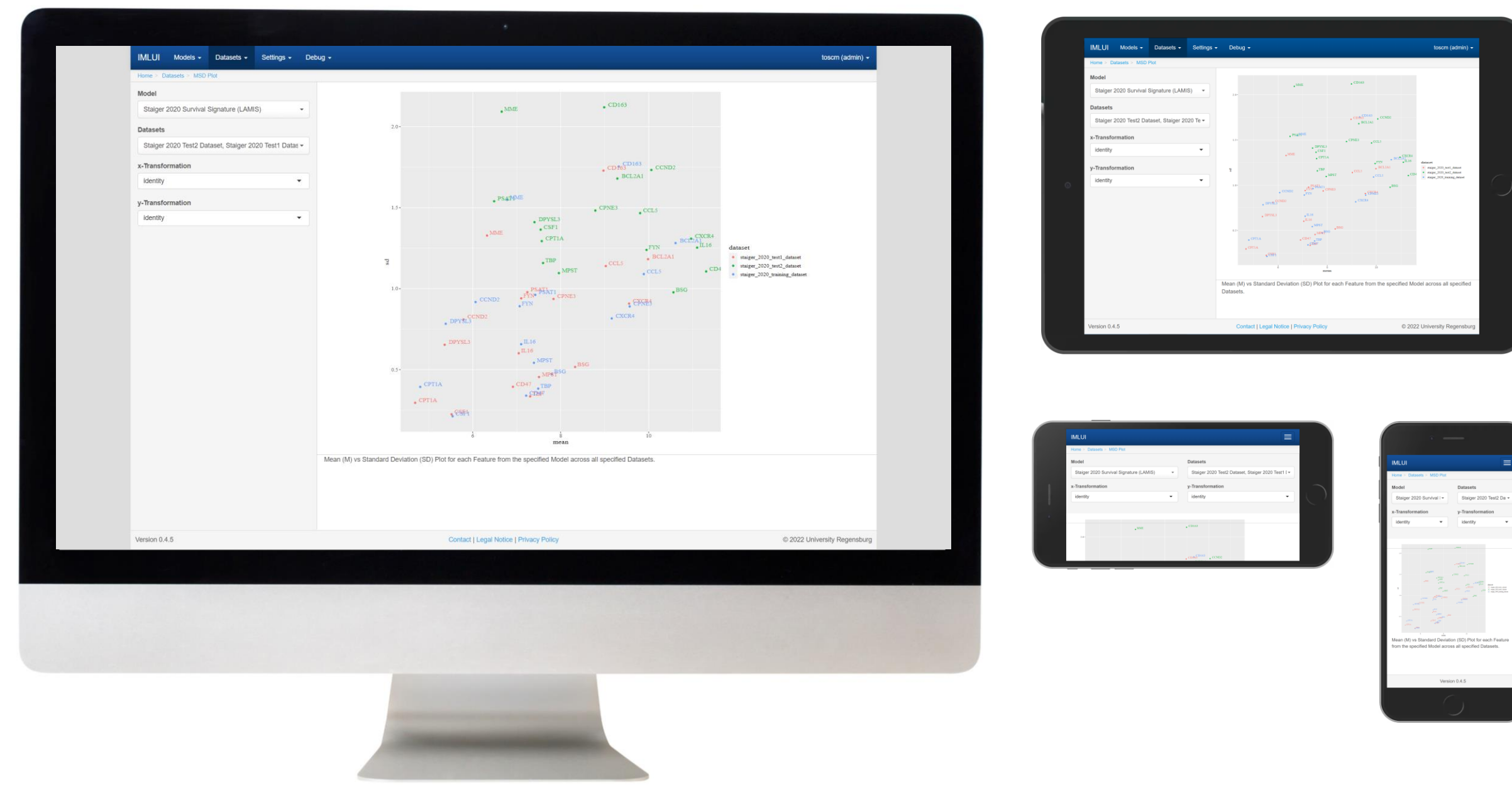- Implementation of visualizations to enhance interpretability
- Implementation of R package *imlui*
- Usage of *imlui* to host a publicly accessible webservice

Results

### Web App with responsive design

Easy Installation and Configuration procedure

1. Install *R*
2. Install the *R* package *imlui* by entering the following commands in a *R* session:
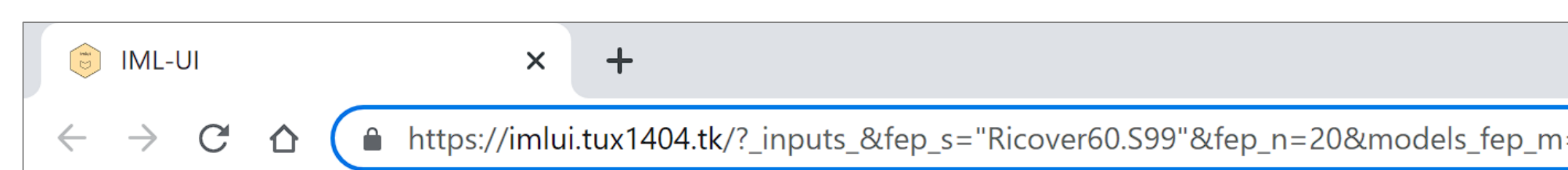
```
install.packages("devtools")
devtools::install_github(repo = "toscm/imlui")
```

3. Optional: adjust database credentials in file *imlui_config.yml*
4. Use the following command to serve the imlui web app via port 8080:

```
imlui::serve_web_app()
```

Supports authentication through 3rd party Identity providers
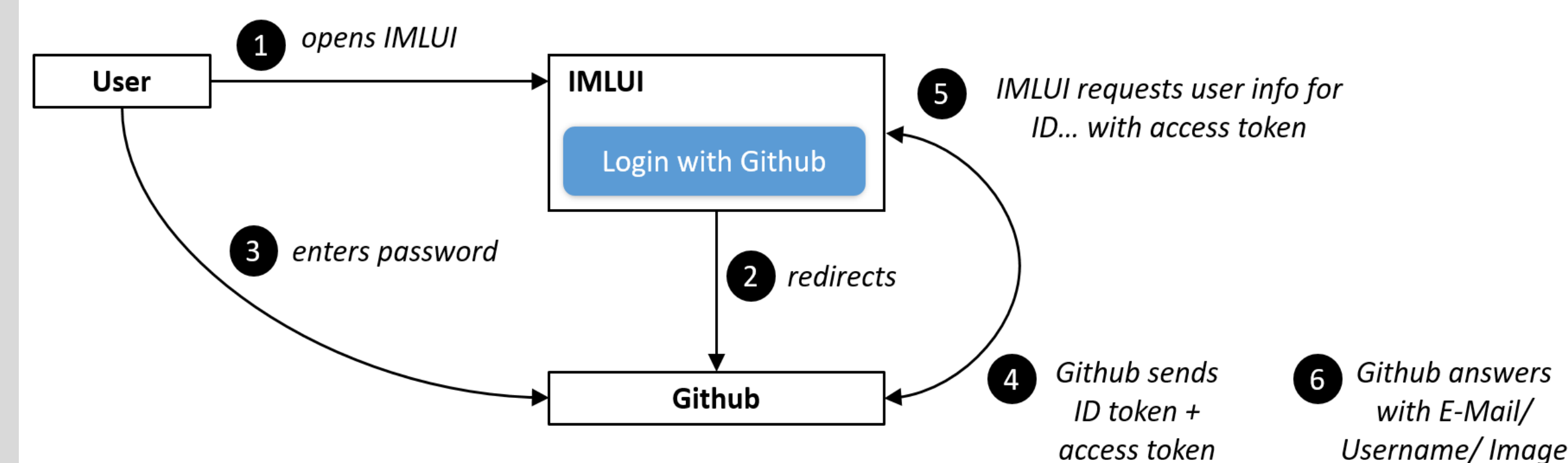
Possibility to bookmark the exact state of each page


Eight Models and Seven Datasets currently included

Models	Datasets
Seifert 2021 COO Signature	Staiger 2020 Test2 Dataset
Seifert 2021 Survival Signature (TRIC)	Staiger 2020 Test1 Dataset
Nordmo 2020 OS Signature	Staiger 2020 Training Dataset
Nordmo 2020 PFS Signature	Nordmo 2020 Training Dataset
Reinders 2020 COO Signature	Reinders 2020 Test Dataset
Staiger 2020 Survival Signature (LAMIS)	Reinders 2020 Training Dataset
Age-Adjusted IPI	
IPI	

Visualizations to enhance interpretability of predictions and/or validate new datasets
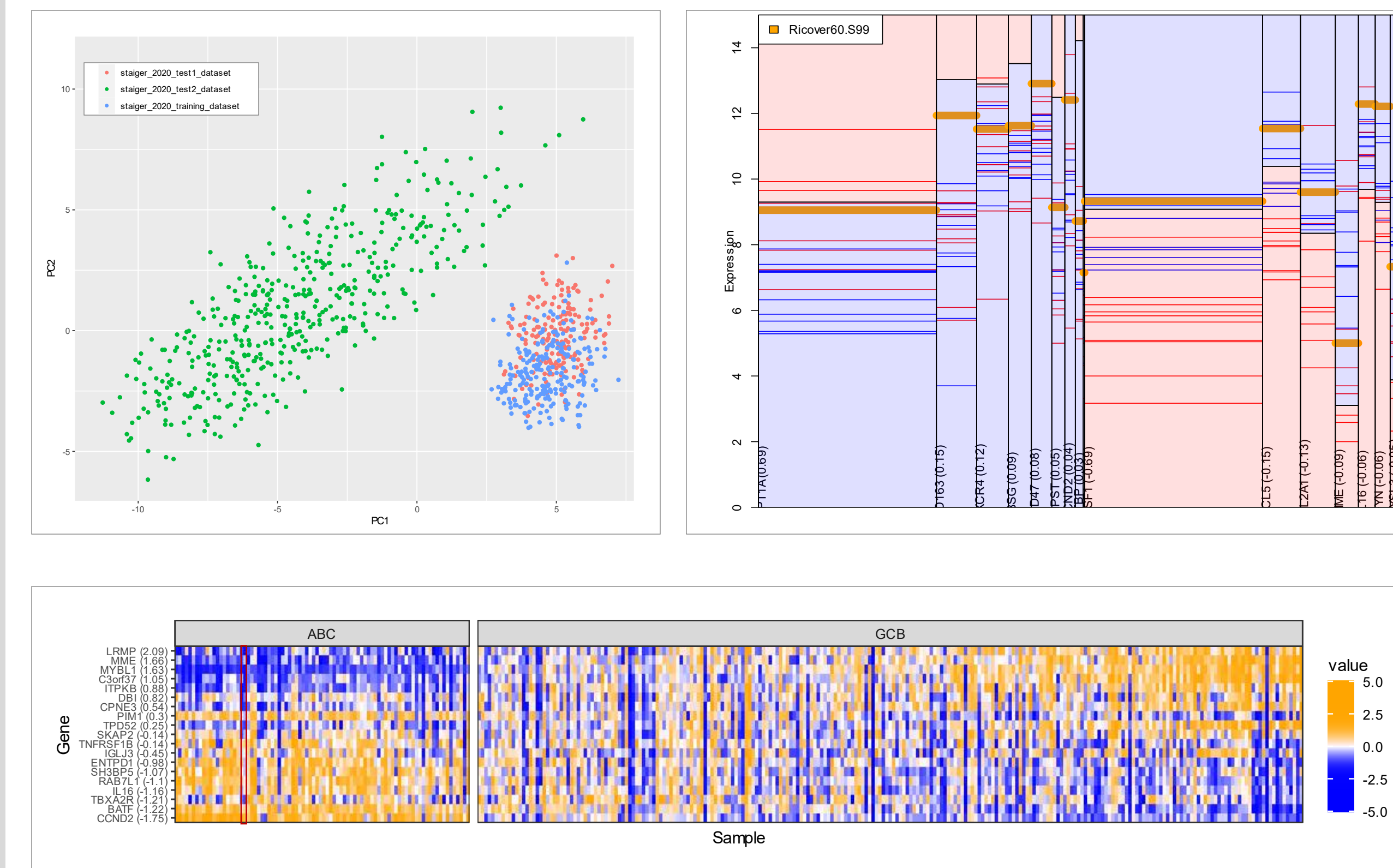
Fine grained permission system for models, datasets and methods

group	model	permission
Spang	seifert_2021_coo	select
Public	nordmo_2020_os	download
Public	reinders_2020_coo	download
Public	staiger_2020_lamis	download
...	...	...

group	dataset	permission
spang	seifert_2021_test	select
spang	staiger_2020_test2	select
public	dummy_staiger_2020_test2	download
public	dummy_nordmo_2020_train	download
...	...	...

user	model	permission
toscm	seifert_2021_coo	download
...	...	...

user	dataset	permission
toscm	seifert_2021_test	download
...	...	...

Fast response times due to process wide caching of models and datasets

Further Information

Gitub Repository:
https://github.com/spang-lab/imlui

Online Manual:
https://github.com/spang-lab/imlui/wiki
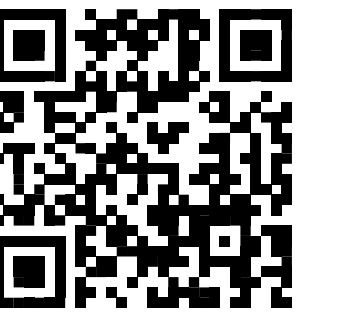
Alpha Version:
https://imlui.tux1404.tk
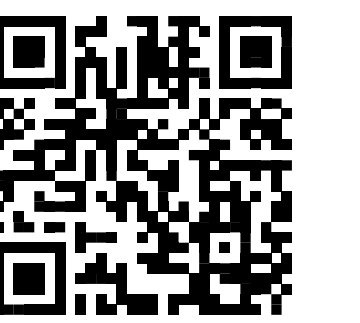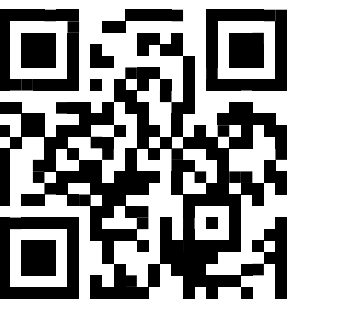
References

1) Hans et al., Blood, 103.1, Jan. 2004, 10.1182/blood-2003-05-1545
2) Muris et al., J Pathol, 208.5, Apr. 2006, 10.1002/path.1924
3) Natkunam et al., J. Clin. Oncol., 26.3, Jan. 2008, 10.1200/JCO.2007.13.0690
4) Nyman et al., Mod. Pathol., 22.8, Aug. 2009, 10.1038/modpathol.2009.73
5) Choi et al., Clin. Cancer Res., 15.17, Sept. 200, 10.1158/1078- 0432.CCR-09-0113
6) Meyer et al., J. Clin. Oncol., 29.2, Jan. 2011, 10.1200/JCO.2010.30.0368
7) Masqué-Soler et al., Blood, 122.11, Sept. 201, 10.1182/blood-2013-06-508937.
8) Scott et al., Blood, 123.8, Feb. 2014, 10.1182/blood-2013-11-536433.
9) Ennishi et al., J. Clin. Oncol., 37.3, Jan. 2019, 10.1200/JCO.18.01583.
10) Staiger et al., J. Clin. Oncol., 35.22, Aug. 2017, 10.1200/JCO.2016.70.3660.
11) Nordmo et al., Leuk. Lymphoma, 62.5, Dec. 2020, 10.1080/10428194.2020.1861268.
12) Rosenwald et al., NEJM, 346.25, June 2002, 10.1056/NEJMoa012914.
13) Shipp et al., Nat. Med., 8.1, Jan. 2002, 10.1038/nm0102-68.
14) Lenz et al., NEJM, 359.22, Nov. 2008, 10.1056/NEJMoa0802885.
15) Wright et al., PNAS, 100.17, Aug. 2003, 10.1073/pnas.1732008100
16) Reinders et al., Sci. Rep., 10.1, May 2020, 10.1038/s41598-020-64212-z.
17) Reddy et al., Cell, 171.2, Oct. 2017, 10.1016/j.cell.2017.09.027.
18) Chapuy et al., Nat. Med., 24.5, May 2018, 10.1038/s41591-018-0016-8.
19) Schmitz et al., NEJM, 378.15, Apr. 2018, 10.1056/NEJMoa1801445.
20) Liu et al., Am. J. Hematol., 94.5, May 2019, 10.1002/ajh.25460