Reproducible Self-Publication via PythonTEX

Introduction and Reference Slides

[github.com/TheChymera/RepSeP]

Horea Christian

[@TheChymera]

Institute for Biomedical Engineering, ETH and University of Zürich

August 31, 2018

Publish From Code, Openly.

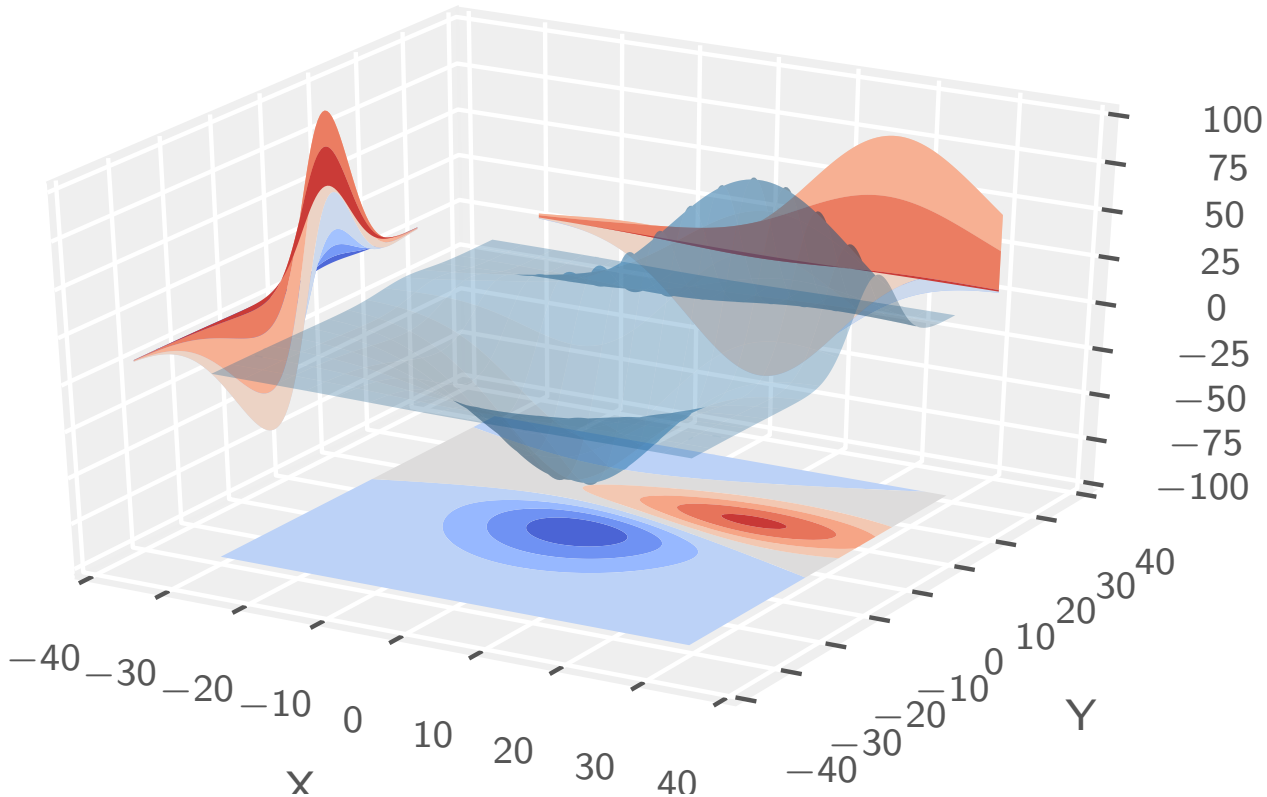
- Transparency $\longrightarrow$ verifiability.
- Reproducibility $\longrightarrow$ hackability.
- Version management support:
  - `diff` -ability.
  - `blame` -ability.

Publish in a Distributed Model, Free.

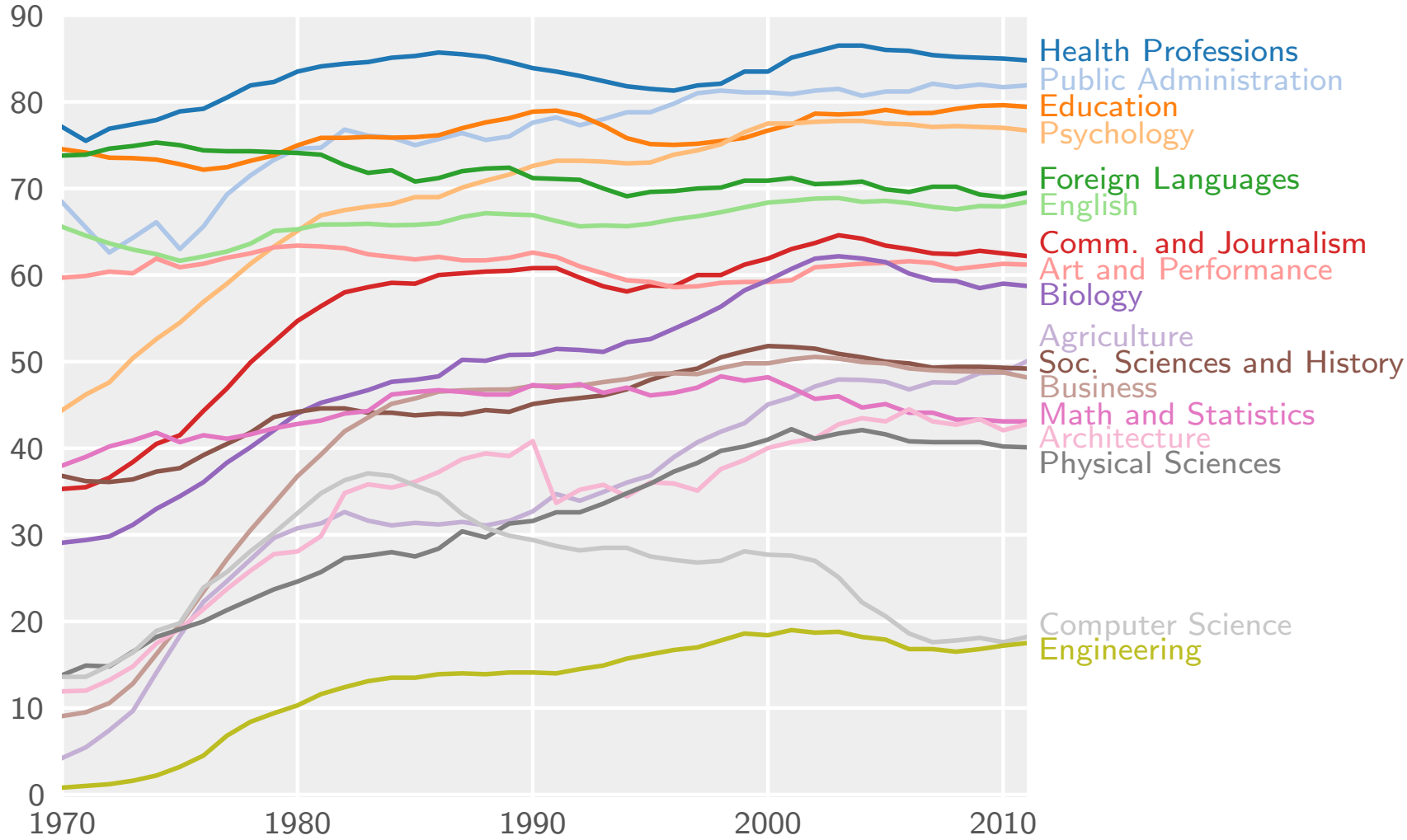
- No entry barrier $\longrightarrow$ citizen science.
- No institutional bias $\longrightarrow$ free science.
- *Less* publication bias $\longrightarrow$ honest science.
- "Direct Market Access".

Publish, in a Presentable Format.

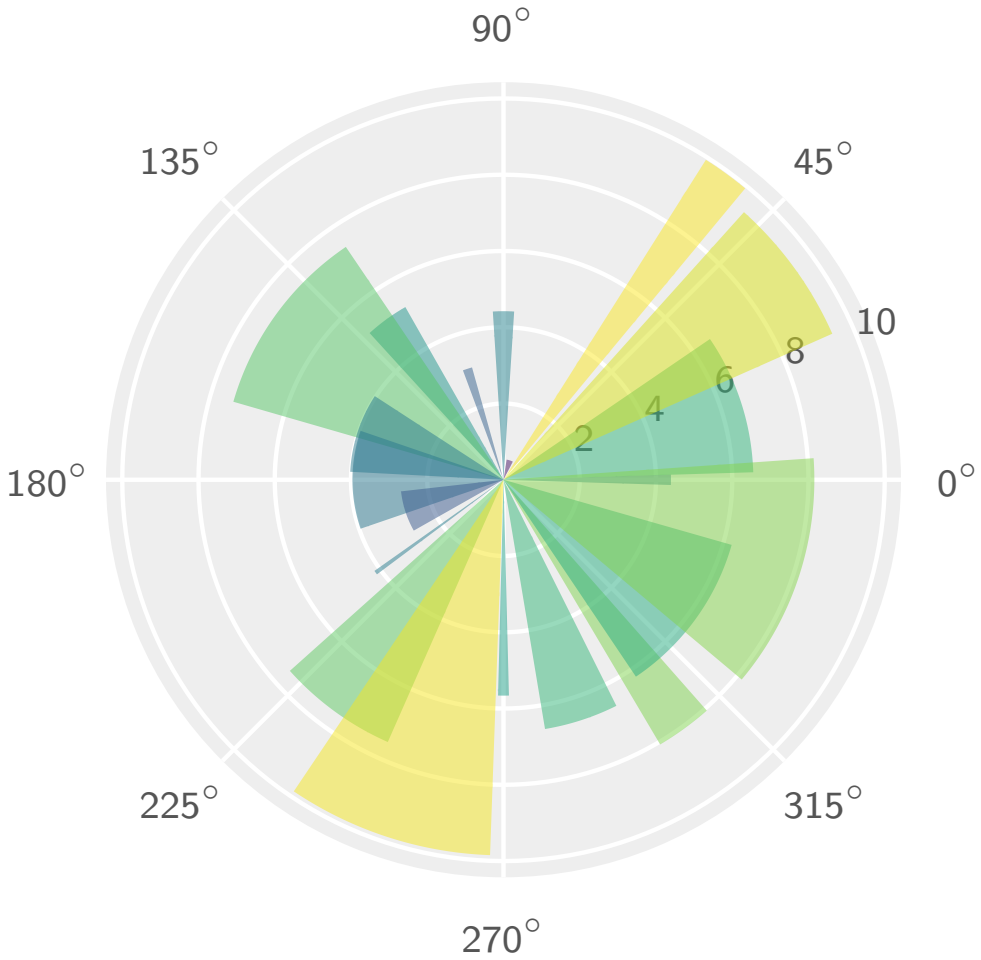
- Article.
- Poster.
- Slides.
- "Notebooks" integrate poorly with both presentation and development.



A 3D plot.



Percentage of Bachelor's degrees conferred to women in the U.S.A. by major



A radar plot.

Cage ID	Period [days]	Sucrose Bottle Position	Sucrose Concentration	Sucrose Preference Ratio	Treatment
32	0 to 2	left	1	0.700000	Fluoxetine
32	2 to 4	right	1	0.729032	Fluoxetine
32	4 to 6	left	1	0.788820	Fluoxetine
33	0 to 2	left	1	0.756098	Fluoxetine
33	2 to 4	right	1	0.595041	Fluoxetine
33	4 to 6	left	1	0.669492	Fluoxetine
34	0 to 2	left	1	0.656934	Control
34	2 to 4	right	1	0.645455	Control
34	4 to 6	left	1	0.574074	Control
35	0 to 2	left	1	0.687500	Control
35	2 to 4	right	1	0.672566	Control
35	4 to 6	left	1	0.742424	Control
36	0 to 2	left	1	0.701571	Control
36	2 to 4	right	1	0.647541	Control
36	4 to 6	left	1	0.725191	Control

Sucrose preference table.

Sometimes Less is More

$$F_{4,26} = 7,\ p = 5 \times 10^{-4}$$

Typesetting the Previous Radar Plot

```
\py{
    pytex_fig('scripts/radar.py',
        label='radar',
        caption='A radar plot.',
        )
    }
```

Typesetting the Previous Table

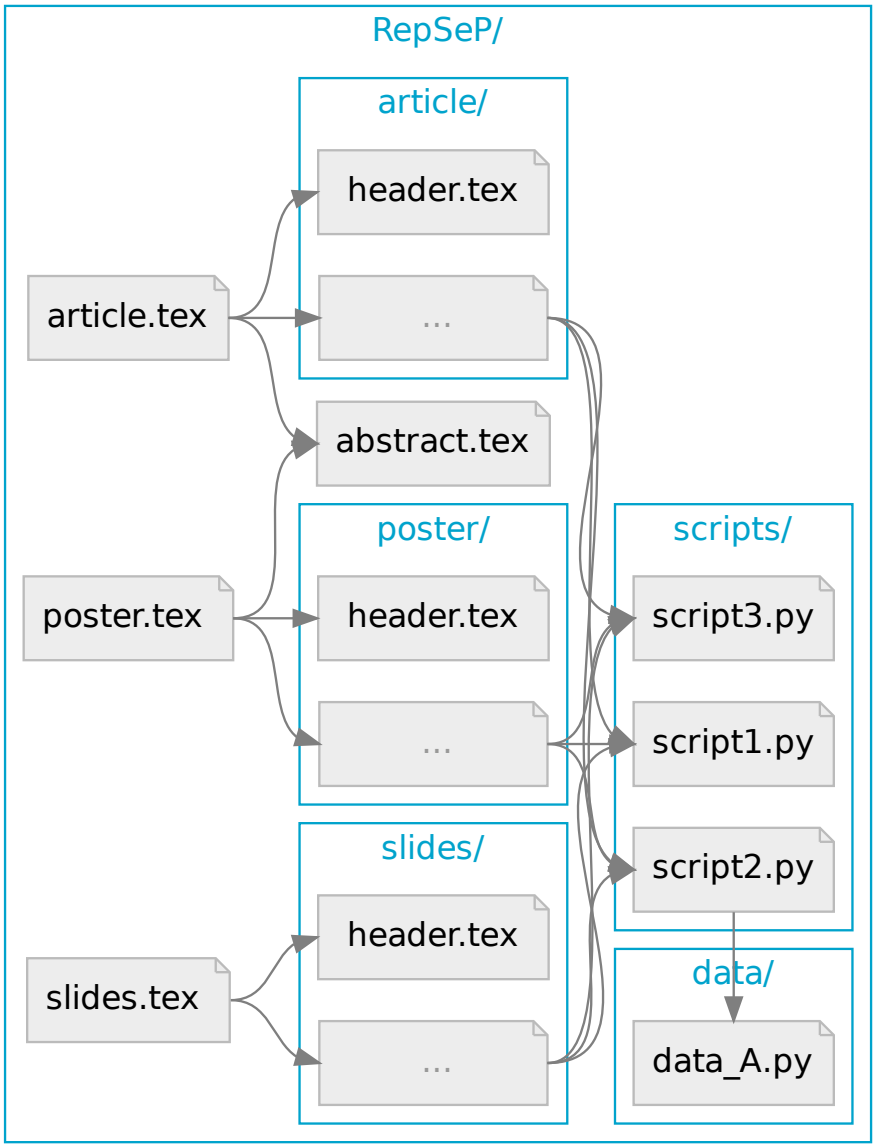
```
\py{
    pytex_tab('scripts/small_table.py',
        label='sp',
        caption='Sucrose preference table.',
        options='\\tiny',
        data='data/sucrosepreference.csv',
        )
    }
```

Typesetting the Previous Inline Statistic

```
\py{
    pytex_printonly('scripts/drs_activityANOVA.py')
    }
```

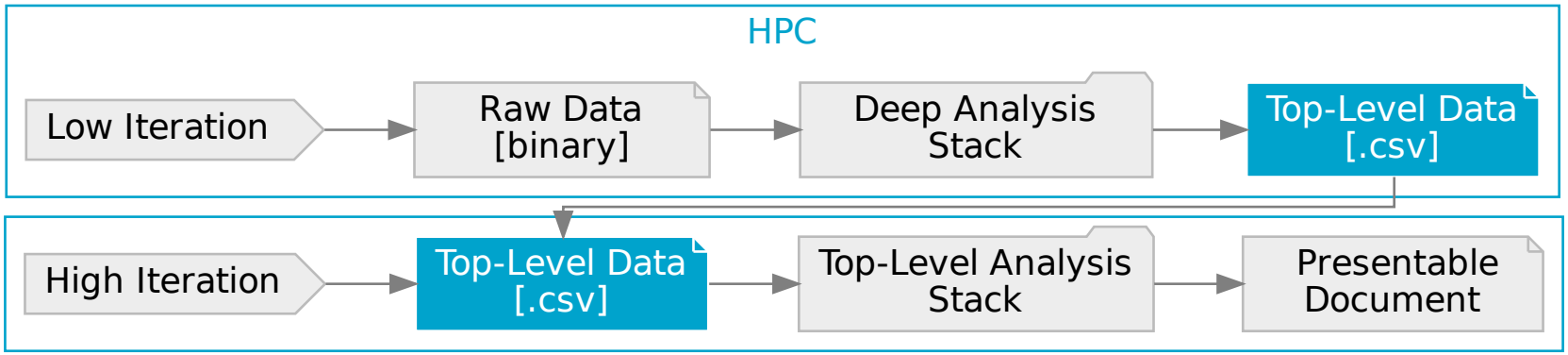
The Framework Topology

- Reduced information duplication.
- Continuous development support.



The Workflow

- Asynchronous offloading for time-consuming analysis.
- Separate packaging for deep analysis stacks.



Time is ripe for code-document integration, cf.:

- Sayako
- Antonio
- et al.

Minimal First-Level Dependencies

```
DEPEND=""
RDEPEND="
        dev-python/matplotlib[${PYTHON_USEDEP}]
        dev-python/numpy[${PYTHON_USEDEP}]
        dev-python/pandas[${PYTHON_USEDEP}]
        >=dev-tex/pythontex-0.16[${PYTHON_USEDEP}]
        dev-texlive/texlive-latex
        "
```

Following the Package Manager Standard (PMS):

- Because dependency graphs should never be managed ad hoc.

Co-Author the Reference Implementation

- The `article.tex` reference document is still in early draft.
- You can contribute, fork, and publish it however you want.

Gain the Best Exposure for Your most Underexposed Work

- Pay-for-Paywall vs. “Direct Market Access”.

Gain the Best Exposure for Your most Underexposed Work

- Pay-for-Paywall vs. “Direct Market Access”.