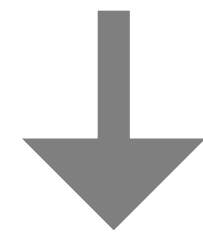# Infrastructures for Reproducible and Transparent Scholarly Communication

Markus Konkol[1], Daniel Nüst[1], Laura Goulier[1]
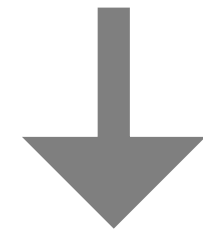[1] Institute for Geoinformatics (ifgi), D-48149 Münster, Germany

ifgi

WWU MÜNSTER

Publishing open reproducible research is sustainable and fair.

⬇

The trend towards open science increases the pressure on authors to provide access to the source code and data underlying the computational results of scientific papers.

⬇

Platforms provide solutions to support the publication of code and data.

What do I need?

Author:
- Supported research areas
- Supported submission formats
- Upload
- Copyright
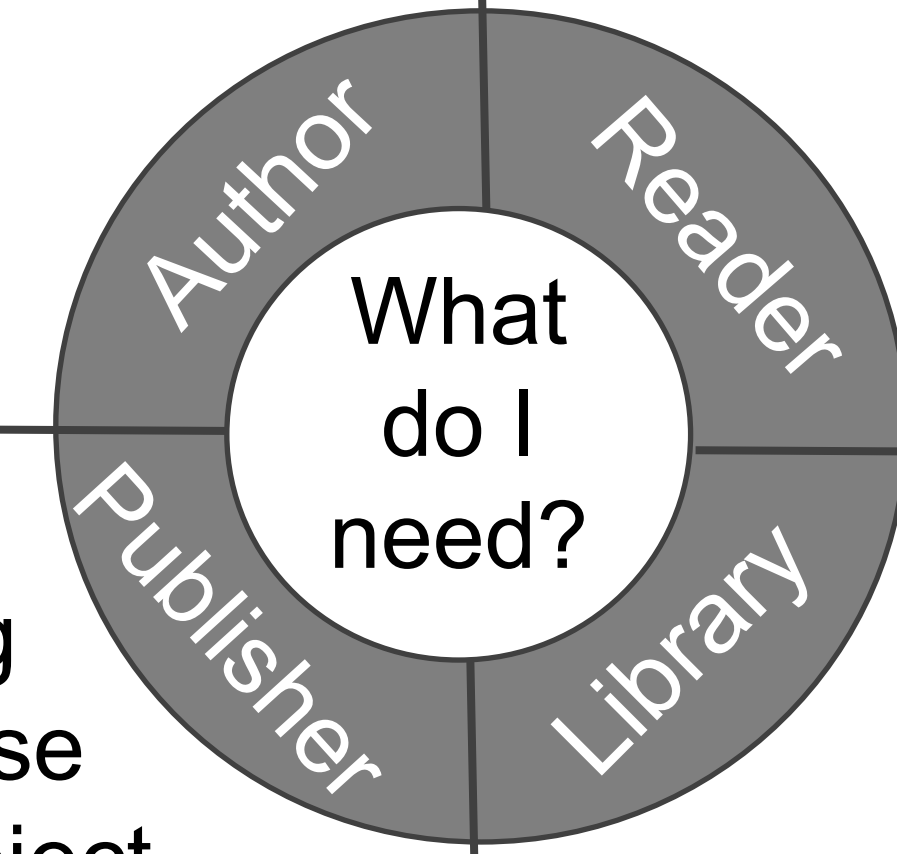
Reader / Support of:
- Searching
- Inspection
- Download
- Execution
- Manipulation

Publisher:
- Platform hosting
- Choice of License
- Stage of the project
- Funding

Library:
- Storing
- Modification/ Deletion after publication
- Sharing

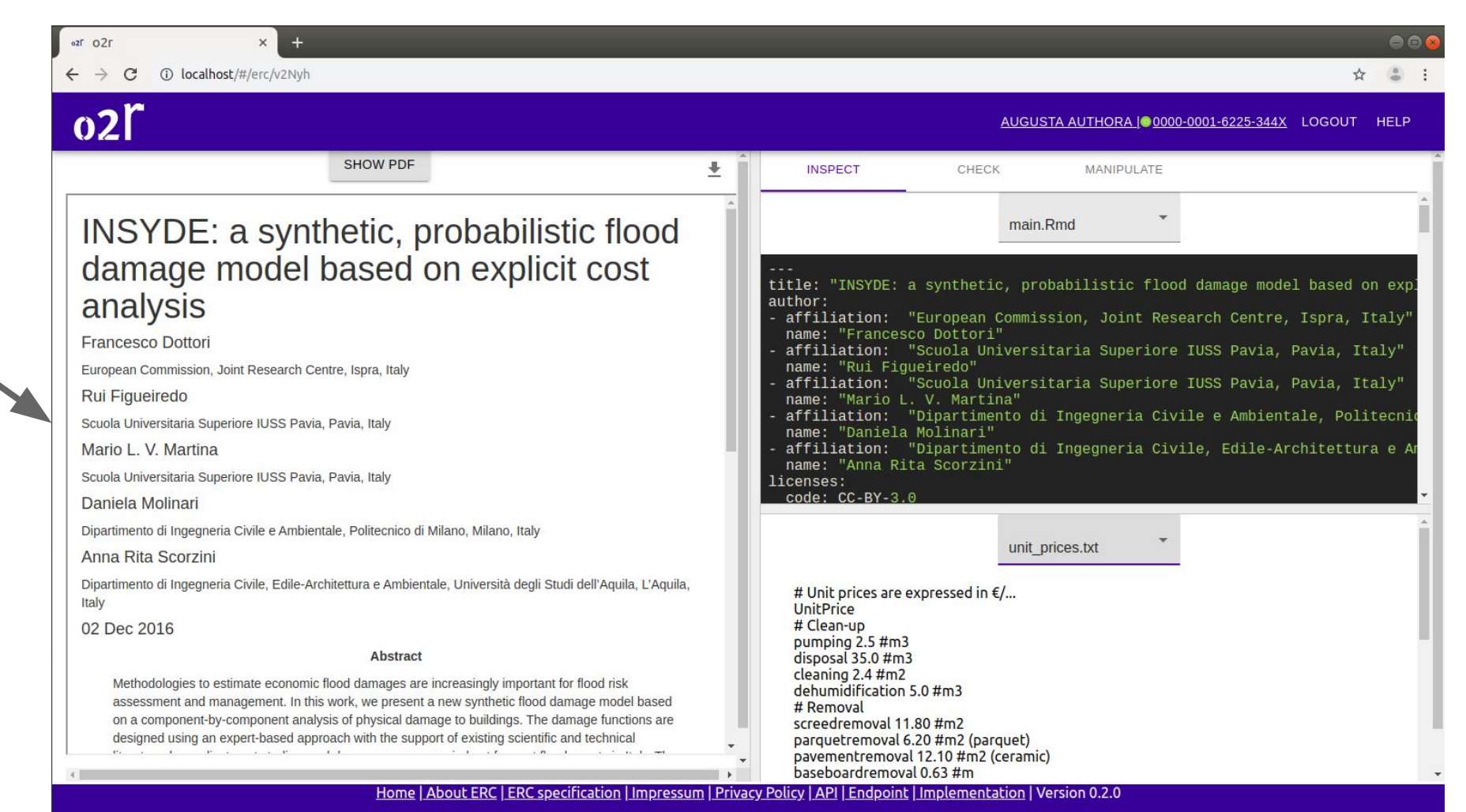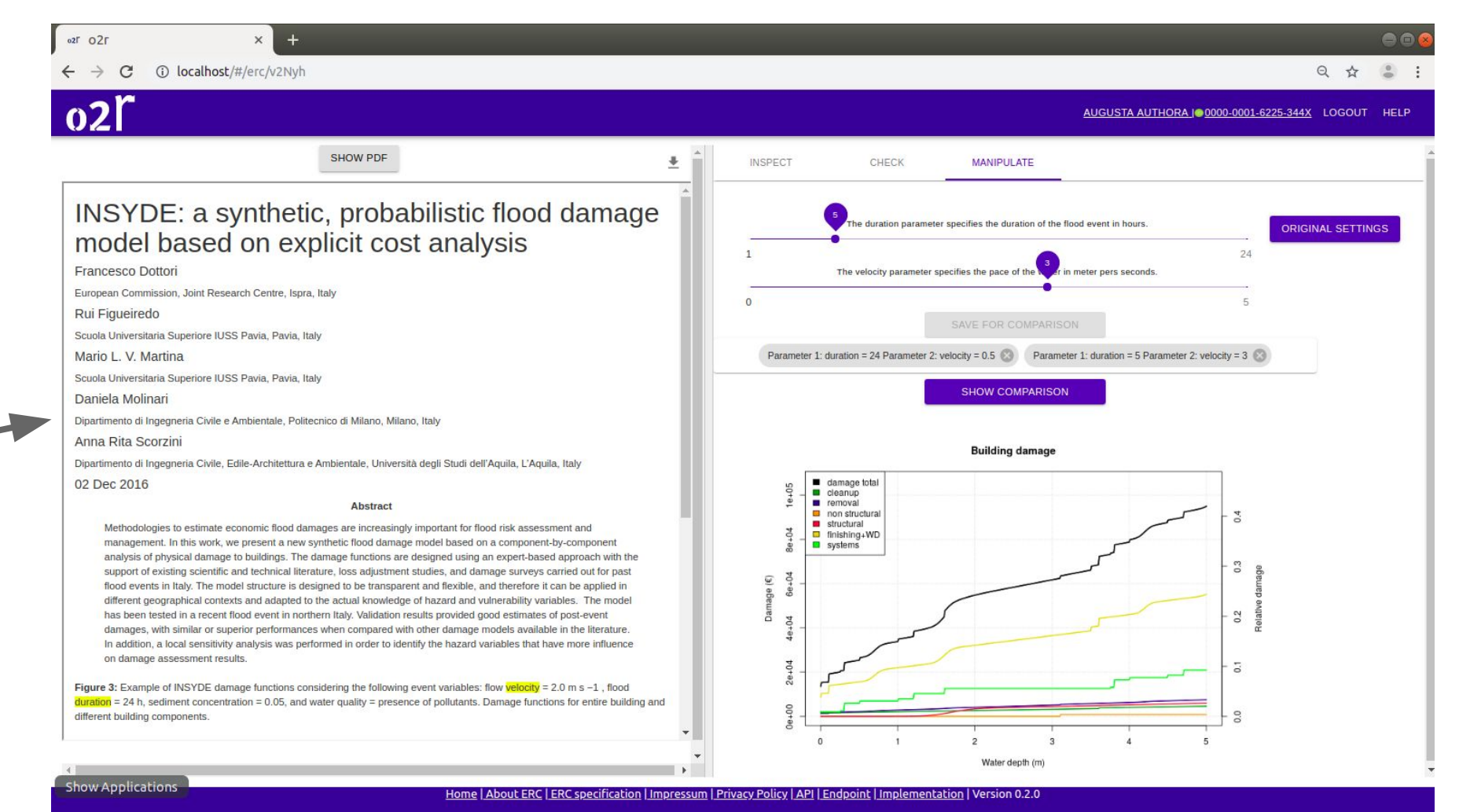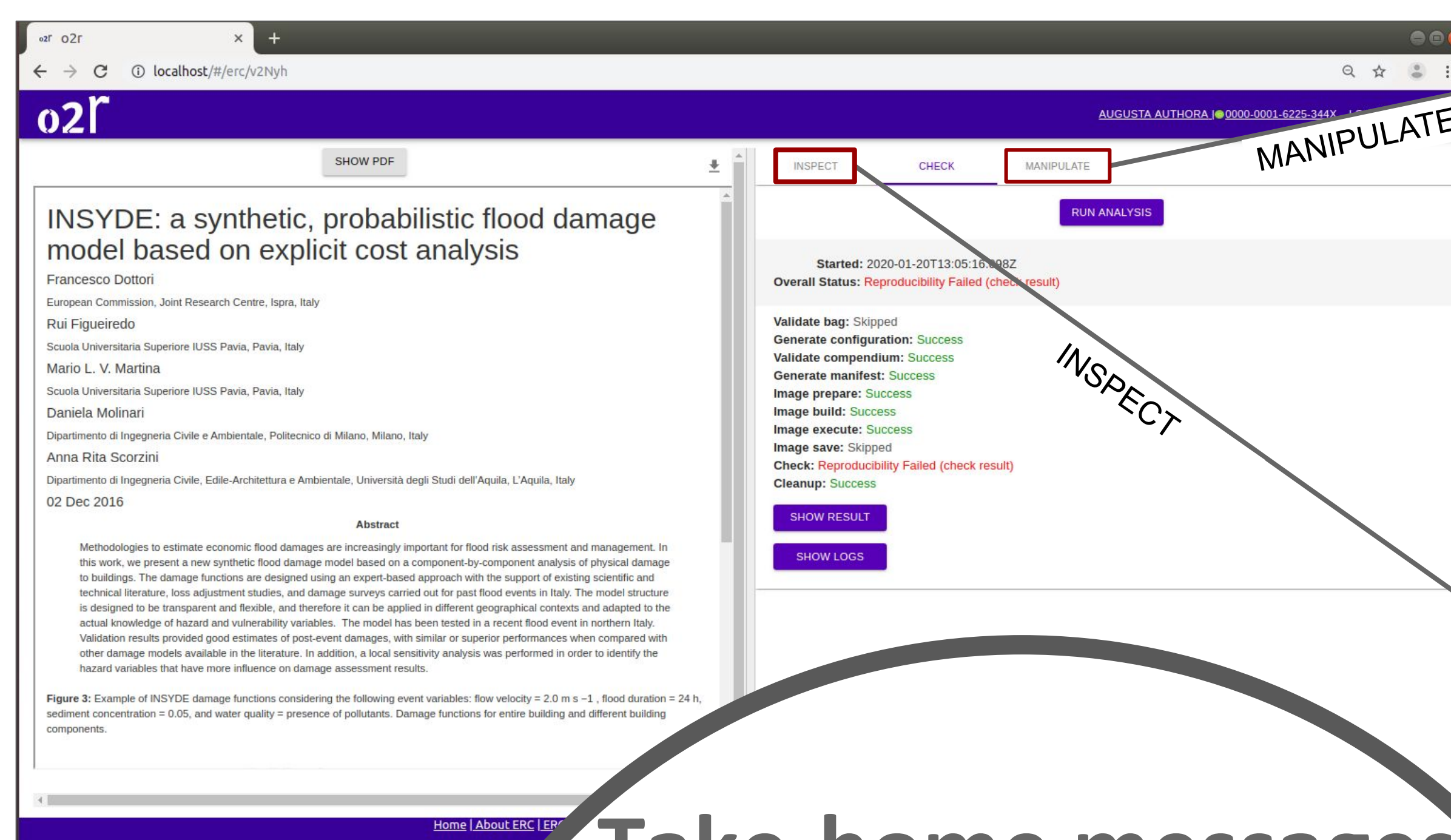
Executable Research Compendium

- UI bindings
- documentation
- software
- data

o2r

INSPECT · MANIPULATE

Take-home messages

1. Many platforms provide one-click reproduction, e.g. **o2r**.

2. All platforms cover a broad range of stakeholder needs

3. You can use **o2r** to modify parameters and compare interactive Figures.

---

o2r
Provides one-click reproduction to manipulate parameters and compare results.

Binder
Launch analyses from Binder-ready repository and inspect the workflow in the browser

Code Ocean
Commercial platform to create "capsules" for several programming languages. Inspect, execute and manipulate the analysis in the browser

eLife RDS
Platform to publish scientific articles as executable documents with all code snippets on "Stencila"

Galaxy
Develop computational analyses without programming expertise using Jupyter Notebook

Gigantum
Packages analyses in Git repositories. Offers a commercial client application to create and execute analyses locally and a cloud to collaborate with peers

Manuscripts
Online tool to write executable collaborative documents based on literate programming within a UI

REANA
Self-hosting platform that provides a set of CLI commands to run large analyses on a remote cloud

ReproZip
Provides a set of CLI commands for encapsulating the analysis. It can be executed on a provided server or locally

Whole Tale
Create a "Tale" for the analysis requirements so users can inspect and execute it in the original environment in browser

Limitations

➔ Handling of sensitive (e.g. private) data
   solution: cloud based data enclaves or involvement of trustworthy authority
➔ Very large data sets and long computational times
➔ No guaranteed anonymity of authors within review process
   solution: anonymous version of the materials
➔ Need for specialised hardware
➔ Often no support for handling copyright
➔ Possibility of deleting materials after publication prevents permanent storage of code and data

Trends

➔ Self-hosting of the infrastructure
➔ Support for notebooks (e.g. RMarkdown, Jupyter)
➔ Provision of UI or integrated development environment instead of CLI commands

---

Konkol *et al.* (2020): Publishing computational Research - A review of infrastructures for reproducible and transparent scholarly communication. arXiv: 2001.00484

DFG Deutsche Forschungsgemeinschaft