

# Workshop on Software Infrastructure for Reproducibility in Science

May 30th-31st, 2013

---

**May 30th:** *We will focus on understanding existing tools, how they fit in the reproducibility pipeline, their benefits and limitations. We will have demos/presentations of the tools interspersed with discussions.*

---

9:00 – 9:30 Welcome and introductions

9:30 – 10:15 Tools that capture the specification of experiments (Moderator: J. Freire)  
Brian Granger - iPython  
Carole Goble – Taverna  
Andrew Davison – Sumatra

10:15 – 10:30 Discussion and break

10:30 – 11:30 Tools that capture the specification of experiments (Moderator: C. Silva)  
James Taylor - Galaxy  
Sergey Fomel - Madagascar  
David Koop – VisTrails  
Carlos Scheidegger – Reproducibility for scripting

11:30 – 12:00 Discussion

12:00 – 1:00 Lunch

1:00 – 2:00 Environment capture, virtualization and publishing  
Mahadev Satyanarayanan (talk+demo) – VMs and Olive (Moderator: D. Shasha)  
Fernando Chirigati – ReProZip  
Victoria Stodden -- RunMyCode

2:00 – 2:15 Discussion and break

2:15 – 3:15 Talks on publishing code and data (Moderator: V. Stodden)  
Ana Nelson – Dexy  
Johanthan Marcow – DuraSpace  
Merce Crosas – Data Verse  
Geoffrey Brown – Virtual CD-ROM Collections

3:15 – 3:45 Discussion and break

3:45 – 4:30 Talks on test/validation, requirements and practices (Moderator: D. Koop)  
Brian Nosek – Open Science Framework  
Matthias Troyer – Reproducibility in Quantum Physics  
Kyle Cranmer – Reproducibility in High-Energy Physics

4:30 – 5:00 Discussion

5:00 – 6:00 Demos  
Jeff Spies – Open Science Framework  
Merce Crosas – Data Verse  
Fernando Chirigati – ReproZip  
David Koop – VisTrails  
Victoria Stodden -- RunMyCode

6:30 – 9:15 Dinner at Spice Market (<http://www.spicemarketnewyork.com>)  
*Take Subway A or C from Jay St/Metrotech to 14<sup>th</sup> Street and Walk to 403 W  
13th St, New York, NY 10014*

---

***May 31st:*** *We will focus on the design of a blueprint for a unified reproducibility framework and discuss models for creating and sustaining a community of developers for reproducibility tools.*

---

9:00 - 10:30 Discussion: What global architecture of tools can we use to enable working researchers capture computations easily, reproduce and vary them on different systems?

10:30 - 11:00 Coffee break

11:00 – 12:30 Breakout groups to define the ideal architecture of a reproducibility framework

*Three groups: capture/representation, archival, test/validation*

12:30 – 1:30 Lunch

1:30 – 3:00 Report from breakouts and discussion on proposed architecture

3:00 – 3:30 Coffee break

3:30 – 5:00 Discussion: How to create and sustain a community of developers of reproducibility tools? Work on white paper