

Open science: How libraries can support  
modern research practice

+

Supporting the modern research skillset:  
The summer of open science

---

Bret Davidson | Eka Grguric | Alison Blaine  
Jennifer Garrett | Lauren Di Monte

NCSU Libraries

[bretdavidson.github.io/dlf-2016](https://bretdavidson.github.io/dlf-2016)

# Agenda

- Open science as problem space
- Open science as modern research practice
- Open science at NC State
- Vagrant, Ansible, & Python

PHILOSOPHICAL  
TRANSACTIONS:  
GIVING SOME  
ACCOMPT  
OF THE PRESENT  
Undertakings, Studies, and Labours  
OF THE  
INGENIOUS  
IN MANY  
CONSIDERABLE PARTS  
OF THE  
WORLD.

---

*Vol I.*

*For Anno 1665, and 1666.*

---

*In the SAVOY,*

Printed by T. M. for John Maury at the Bell, a little with-  
out Temple-Bar, and *Jacob Elphry* in Duck-Lane,  
Thames at the Royal Exchange.

# Nullius in Verba

"Take nobody's word for it."

Only **6 out of 53** “landmark”  
cancer studies could be  
reproduced.

Nature, [www.nature.com/nature/journal/v483/n7391/full/483531a.html](http://www.nature.com/nature/journal/v483/n7391/full/483531a.html)

# How Science Goes Wrong

"Too much trusting, not enough verifying."

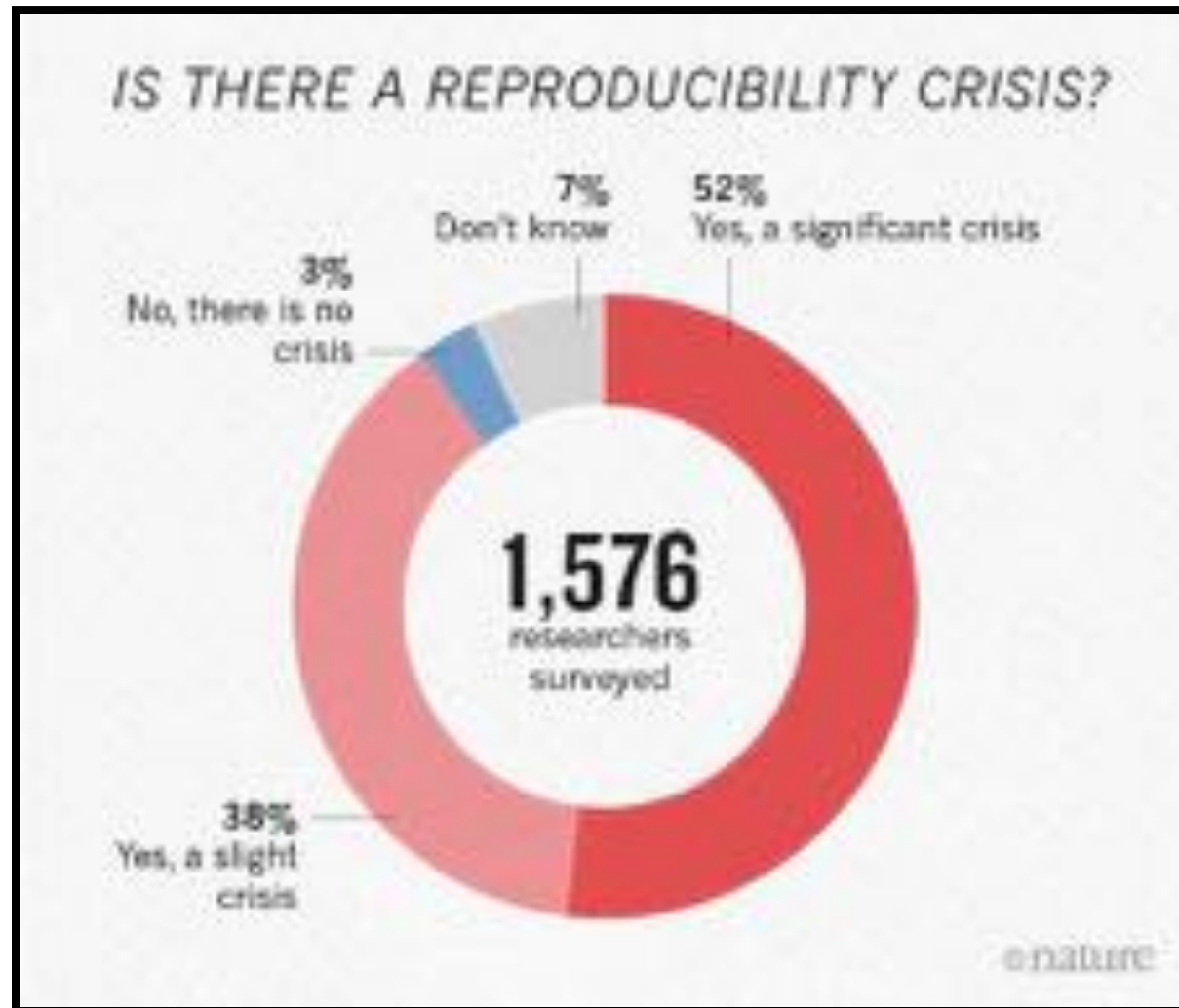
Economist, [www.economist.com/news/leaders/21588069-scientific-research-has-changed-world-now-it-needs-change-itself-how-science-goes-wrong](http://www.economist.com/news/leaders/21588069-scientific-research-has-changed-world-now-it-needs-change-itself-how-science-goes-wrong)

# Reproducibility Crisis

Empirical estimates suggest  
most published medical  
research is true.

arXiv, [arxiv.org/abs/1301.3718](https://arxiv.org/abs/1301.3718)





Nature, [www.nature.com/news/1-500-scientists-lift-the-lid-on-reproducibility-1.19970](http://www.nature.com/news/1-500-scientists-lift-the-lid-on-reproducibility-1.19970)

Open Science can increase  
reproducibility.

# What is Open Science?

- Open Access
- Open Data
- Open Notebooks
- Open Source

Open Science is a return to first principles of scientific practice.

Open practices require  
new skillsets.

The way that research is  
carried out.

The way that research is  
disseminated.

How digital technologies are affecting the practice of science.







open science



# Paul David

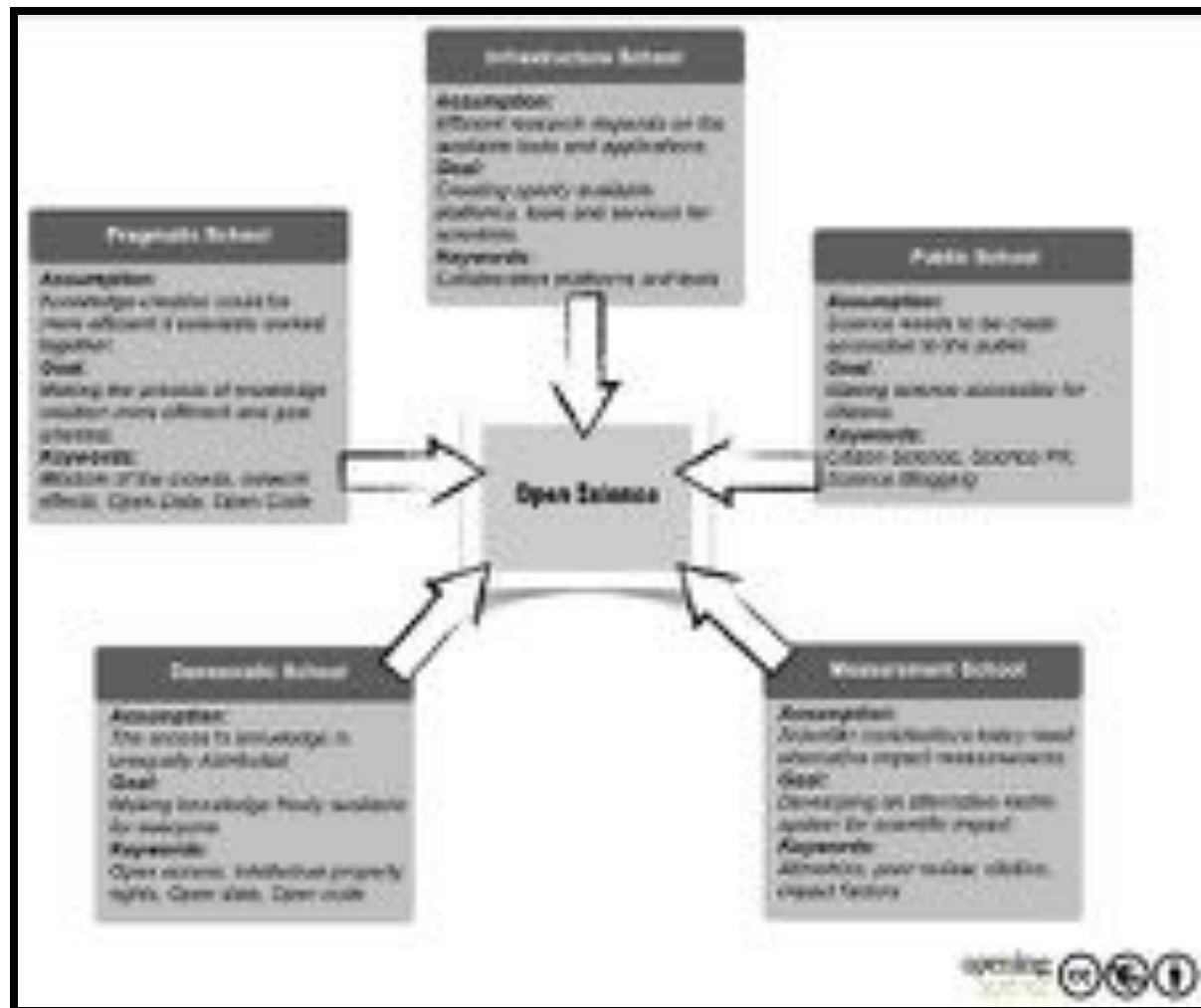
Economist / Historian

The Knowledge Economy

# Five schools of thought

by Sönke Bartling & Sascha Friesike

Editors, <http://book.openingscience.org/>



# The Five Schools of Thought

# Infrastructure

tech. architecture

# Public

accessibility of knowledge  
creation + citizen science

# Measurement

alt. impact measurement



# Democratic

access to knowledge

# Pragmatic

collaborative research

# Recent Events

Life Science | 18 JULY 2014 10:41 AM

Health | Cancer | Health

# In cancer science, many "discoveries" don't hold up

NEW YORK (Reuters Health) -



© 2014 Reuters Health. All rights reserved.



THE U.S. GOVERNMENT'S SUPPORT FOR THE RESEARCH

CONDUCTED WITH THE

RESEARCHER'S INTEREST

WAS TO BE

A study conducted at Princeton University found that many basic studies of cancer — a field  
progressed that had already been — are conducted with just enough science to

SPAIN

## Is the evidence for austerity based on an Excel spreadsheet error?

[illegible]By Steve Paulson [spaulson@bloomberg.com](mailto:spaulson@bloomberg.com) and [spaulson@bloomberg.com](mailto:spaulson@bloomberg.com)

One of the most influential studies that others used as a guide for assessing the results of an intervention was *Chaffin*. The authors argued that

The paper in question is a classic statistical and somewhat ingenuitous exercise that states "The welfare state flows of help," which found that economic growth severely suffers when a country's welfare state and welfare are around 40% of GDP. This is



strongly

1. Investments are higher when a substantial flow of funds enters the economy.
2. The business sector is a demand-driven sector.
3. It's a high-tech, high-innovation industry. It has more workers than gold.
4. The technology base is growing rapidly.
5. When it's too slow, demand is depressed by the government.



† The authors declare no potential conflict of interest.



• The distribution of all scores is normal.



<sup>1</sup> [http://www.who.int/csr/don/2009\\_04\\_02/en/](http://www.who.int/csr/don/2009_04_02/en/)



THE CONSPIRACY

AND SUPPLIES

# Students Reveal How They Broke the Lead Contamination Case in Flint, Mich.

Officials in the city tried to keep 'secrets' from residents, but students in the city's schools found out the truth.

By Michael S. Schmidt and Michael C. Smith



RELATED ARTICLES



Flint, Mich., Is the Worst Case of Lead Contamination in the Nation



How Lead Contamination in Flint, Mich., Affects the Nation



How Lead Contamination in Flint, Mich., Affects the Nation

The Flint, Mich., Water Crisis: A Timeline

#### Open Peer Review

### Open drug discovery for the Zika virus (version 1) reference awaiting peer review

10/10/2019 11:11:11 AM  
10/10/2019 11:11:11 AM  
10/10/2019 11:11:11 AM  
10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

#### Abstract

The Zika virus (ZIKV) is a member of the Flaviviridae family, which is responsible for causing Zika disease. ZIKV is a single-stranded RNA virus that is transmitted by mosquitoes. It is a major public health concern because of its association with congenital defects and neurological complications. The purpose of this study was to develop a novel drug discovery approach for ZIKV. We used a combination of high-throughput screening and computational modeling to identify potential drug targets and compounds. We identified several novel targets and compounds that show promise for further development. This study represents a significant step forward in the discovery of new drugs for ZIKV.

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM



10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

#### Open Peer Review

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

10/10/2019 11:11:11 AM

# Why Libraries?





# Aligns with core library values

- information access
- peer review
- community-based knowledge creation
- the preservation and dissemination of research

# Libraries are champions of open source



Libraries  
are about  
supporting their users

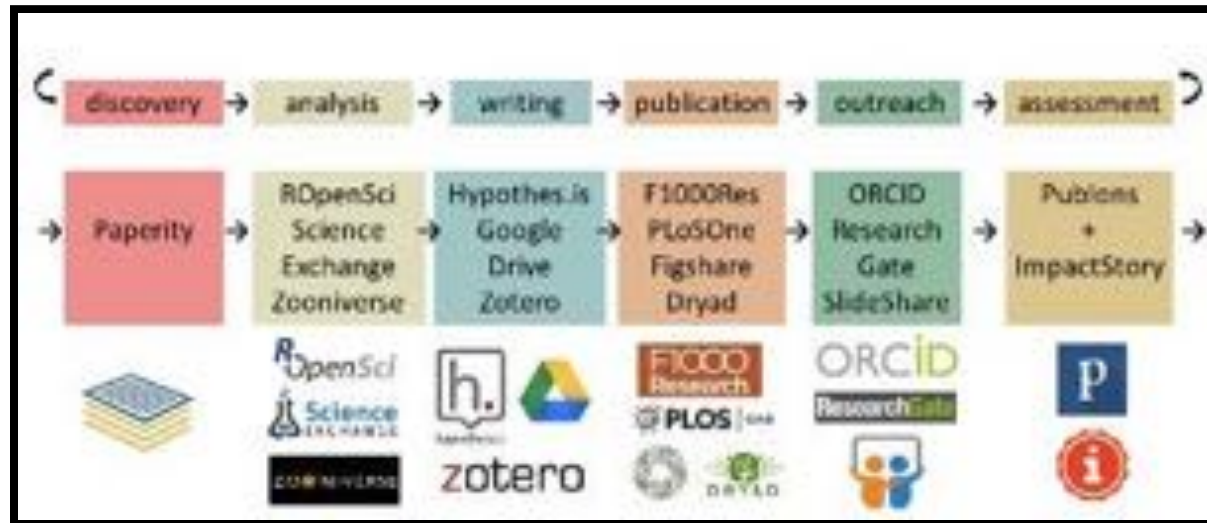
Academic Libraries

are about

supporting research practice

Ongoing disruption by digital  
technologies in modern research  
practice

# Hypothetical Open Science Workflow



101 Innovations in Scholarly Communication,  
<https://innoscholcomm.silk.co/>

# Policy Shifts in support of open



# Ecosystem of Support for Modern Research Practice at NCSU Libraries



## Research Support



### **COLLECTIONS**

Electronic Journals  
E-books, Digital  
Repositories



### **CITATION MANAGEMENT**

EndNote, Zotero,  
Mendeley, RefWorks



### **DATA ANALYSIS**

Probing, visualization of  
numeric data, GIS  
software assistance



### **DATA MANAGEMENT**

DMT backup, storage,  
file structure, data  
migration



Have a question?



### **GRANTS AND FUNDING**

Questionnaires,  
proposal writing,  
grant administration,  
compliance



### **MEASURING RESEARCH IMPACT**

Content analysis,  
citation indicators, web  
metrics, social  
media marketing



### **PUBLISHING AND COPYRIGHT**

Copyright  
agreements, Open  
Access, self-archiving



### **VISUALIZATION**

Big Data, network  
analysis, data  
visualization

### **CHANCELLOR'S FACULTY EXCELLENCE PROGRAM**

Consulting, Assessment, Data Support,  
Research Support

### **BUDGET SPECIALISTS**



Jeff Egan



Karen Gordin



Catherine Oswald



Mary Ann Wu

For a complete directory, go to



### **TEXT AND DATA MINING**

Classification, text and  
sentiment analysis







## WORKSHOPS

On Workshops

Expertise Programme

Workshop Introduction

Workshop Overview

Workshop Introduction

Workshop Introduction

View other events



2023

## Visualization Workshops

NOV

18

### Infographics: Visualizing Information

November 18  
11:00 AM to 12:30 PM

Dr. H. J.

ITC Club B, 2nd Floor, 1st Floor

DEC

2

### Getting Started with Data Visualization: Tools for Research

December 2  
11:00 AM to 12:30 PM

Dr. H. J.

ITC Club B, 2nd Floor, 1st Floor

DEC

2

### Getting Started with Data Visualization II: Elements of Design

December 2  
1:00 PM to 2:30 PM

Dr. H. J.

ITC Club B, 2nd Floor, 1st Floor

DEC

5

### R for Absolute Beginners

December 5  
1:00 PM to 2:30 PM

Dr. H. J.

## Makerspace

### D. H. HILL LIBRARY

A D-HH creation and collaboration space



### JAMES B. HUNT JR. LIBRARY

3D Printing services and more





**WOLFPACK  
CITIZEN  
SCIENCE  
CHALLENGE**

The NCSU Libraries'  
**Open Science Initiative**



# Goals

- explore open science practice at NCSU
- better understand researcher needs in context

We took a non-prescriptive  
**user-centered** approach.

Creating opportunities for  
communication.

# Open Science Unconference

[illegible]

# Follow-up Informal Interviews

Modern Research Skills Gap  
Insufficient Incentives



## EVENTS

Workshops & Seminars

Workshops

Workshops

Online

## EVENT SERIES

Default Sort

Summer of Open Science

Research Projects

Writing Studies

Water Studies

All Open Science Projects

Open Science

Writing Studies

## ABOUT

Workshops

Workshops

Workshops

Workshops

Workshops

Workshops

Workshops

Workshops

Workshops

Workshops

Workshops

Workshops

Workshops

Workshops

Workshops

Workshops

Workshops

Workshops

Workshops

Workshops

Workshops

## Summer of Open Science Event Series

The Summer of Open Science Event Series is a series of workshops and seminars that support student research projects through hands-on learning.

Workshops are designed to help students learn the basics of research and data analysis. They are designed to be hands-on and interactive, allowing students to learn by doing. The Summer of Open Science is designed to address these needs.



### Python & R for Data Science

May 18, 2016  
10:00 AM - 12:00 PM  
Online (via Zoom) \$100 (includes materials)



### Python & R for Data Science

May 18, 2016  
10:00 AM - 12:00 PM  
Online (via Zoom) \$100 (includes materials)



### GitHub and Twitter for Data Science

May 18, 2016  
10:00 AM - 12:00 PM  
Online (via Zoom) \$100 (includes materials)

### Python & R for Data Science

May 18, 2016  
10:00 AM - 12:00 PM  
Online (via Zoom) \$100 (includes materials)

### Python & R for Data Science

May 18, 2016  
10:00 AM - 12:00 PM  
Online (via Zoom) \$100 (includes materials)



# Goals

- Support modern research practice through hands on skill building
- Provide networking opportunities
- Increase visibility of library spaces & services

# Skills

- Scholarly identity creation
- Scientific computing
- Building a website
- Data harvesting
- Code collaboration

# Tools

- Open source software
- Virtual machines

# The Planning Team

(Representation from Digital Library Initiatives, Makerspace, Research & Information Services, Libraries Fellows)

- **Ekatarina [Eka] Grguric (Project Lead)**
- **Lauren Di Monte (Project Manager)**
- **Alison Blaine (Content Development)**
- **Bret Davidson (Technical Lead)**
- **Jennifer Garrett (Community Development)**

# Summer of Open Science

- Workshops
  - Intro to the Command Line Interface
  - Web Scraping with Python
  - Understand and Build Your Scholarly Identity
  - Scientific Computing with Python & Raspberry Pi
  - Build Your Scholarly Website the Easy Way
- Events
  - Meetups
  - End-of-Summer Showcase

# Workshop Goals

- Skills for introductory-level users
- Hands-on practice
- Provide resources and opportunities for going deeper

# Workshop Structure



- Brief presentation with contextual overview
- Activities
- Roving instruction support and live demo from instructor
- Instruction handouts to allow for work at own pace





**OPENING TO THE CONNECTED LMS INTERFACE**  
 v. 3.01  
 10/20/2016  
 Copyright © 2016 by NCSU

10/20/2016

Version 3.01

Copyright © 2016 by NCSU. All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or by any information storage and retrieval system, without permission in writing from NCSU.

Copyright © 2016 by NCSU. All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or by any information storage and retrieval system, without permission in writing from NCSU.

Copyright © 2016 by NCSU. All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or by any information storage and retrieval system, without permission in writing from NCSU.

Copyright © 2016 by NCSU. All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or by any information storage and retrieval system, without permission in writing from NCSU.

## Introduction to the Connected LMS Interface

Version 3.01 (10/20/2016) by NCSU. All rights reserved.

Copyright © 2016 by NCSU. All rights reserved.

### Definition

Connected LMS Interface: A system that connects the LMS to the LMS.

Connected LMS Interface: A system that connects the LMS to the LMS.

Connected LMS Interface: A system that connects the LMS to the LMS.

Connected LMS Interface: A system that connects the LMS to the LMS.

Connected LMS Interface: A system that connects the LMS to the LMS.

Connected LMS Interface: A system that connects the LMS to the LMS.

Connected LMS Interface: A system that connects the LMS to the LMS.

Connected LMS Interface: A system that connects the LMS to the LMS.

Connected LMS Interface: A system that connects the LMS to the LMS.

Connected LMS Interface: A system that connects the LMS to the LMS.

Connected LMS Interface: A system that connects the LMS to the LMS.

Connected LMS Interface: A system that connects the LMS to the LMS.

Connected LMS Interface: A system that connects the LMS to the LMS.

Connected LMS Interface: A system that connects the LMS to the LMS.

Connected LMS Interface: A system that connects the LMS to the LMS.

Connected LMS Interface: A system that connects the LMS to the LMS.

Connected LMS Interface: A system that connects the LMS to the LMS.

go.ncsu.edu/...

# In-Workshop Logistics

- Shared materials folder on Google Drive
- Laptops provided via a laptop cart
- Evaluation forms for feedback
- At least 2 instructors

## Instructors



Brittany Johnson



Eka Grguric



Lauren Dillmore



Alison Blaine



Madison Sullivan



Will Cross



Todd Stauffer

# Marketing

- Slide on library front page linking to event website with logo
- Emails to various listservs
- E-board advertisements on campus
- Press release

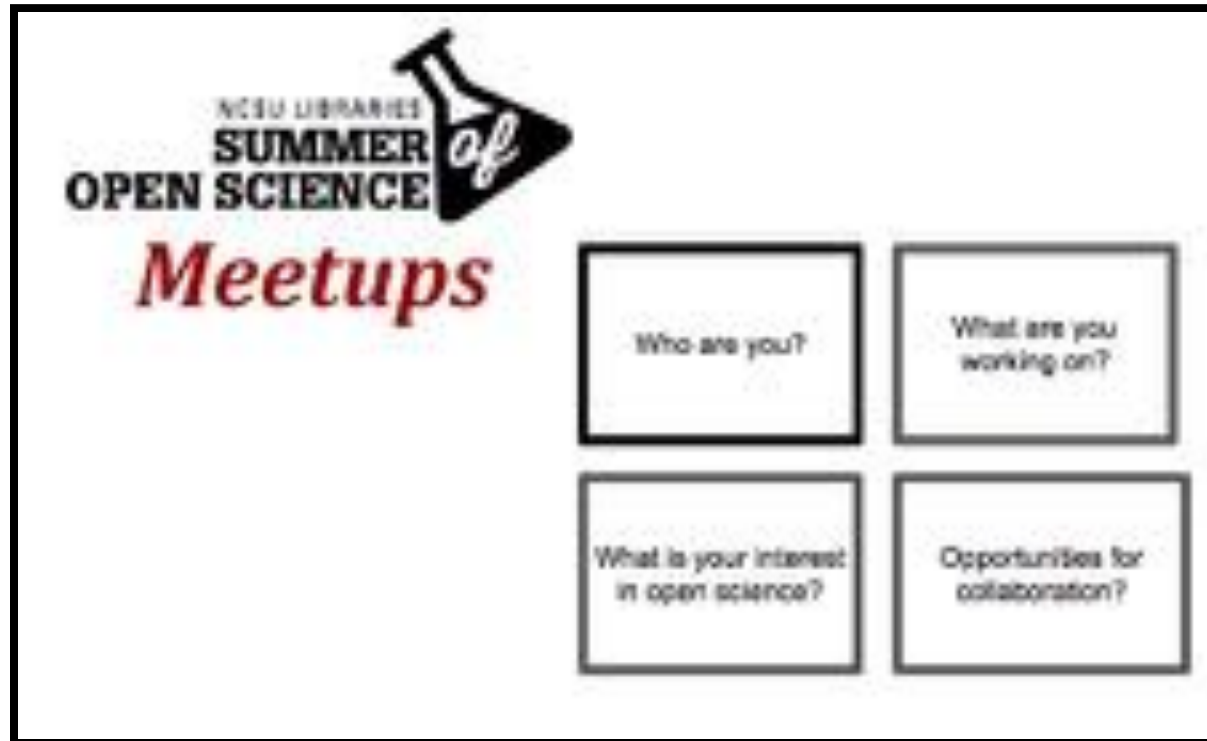




## **Scientific Computing with Python & Raspberry Pi:**

40 person waitlist

# Meetups



- Monthly informal networking and discussion
- Held at a coffee shop on campus
- Rich conversations, although small numbers

# Final Showcase



- Food & mingling
- Lightning talks
  - Agriculture and Accidents: Digital Learning in Summer 2016
  - Modeling Individual Developer Concept Knowledge Using Public Git Repositories
  - SciBridge: Bringing together African and U.S. scientists

# Takeaways





**Interdisciplinary Need:**  
over 40 departments across ~16 colleges

# Takeaways

- "Open Science" attracted non-scientists as well as scientists
- High demand for introductory coding skills (Python)
- Interest among graduate students for opportunities for interdisciplinary research sharing
- Summer presents interesting opportunities and challenges

# Web Scrapping with Python

Technical workshops are  
ripe for disaster.

# What could go wrong?

- OS images reset overnight
- Improper permissions
- Network connectivity issues
- Python 2 vs Python 3
- Missing packages

This is not theoretical!

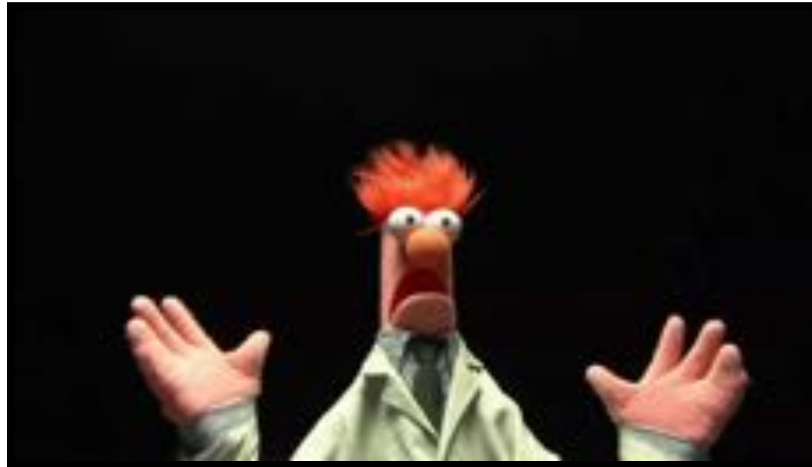


Image courtesy of Tumblr, [animisensa](#)

# Instructor Challenges

- Consistency across user environments
- Consistency of course materials
- Time to provision computing environments
- Ease of collaboration



# Student Challenges

- Basic Python data types and structures
- Python module system
- Retrieve a web page with Requests
- Parse content with BeautifulSoup
- Generate a word cloud with matplotlib
- Control Structures
- Exception Handling
- Working with file system

# Many Options

- Custom OS Images
- Custom Distributions, e.g. Anaconda
- Interactive Environments, e.g. Jupyter

# Our Approach

- Vagrant for managing OS
- Ansible for provisioning and configuration
- Course or lab specific packages and resources

# Easy!

1. Install Vagrant
2. Install VirtualBox
3. Clone project repo
4. ``vagrant up``
5. ``vagrant ssh``
6. Execute code!

This is reproducible computing!

github.com/NCSU-Libraries/python-vagrant



# Benefits

- Consistent environment user to user
- Single target for course materials
- Faster provisioning for new workshops
- Reproducible

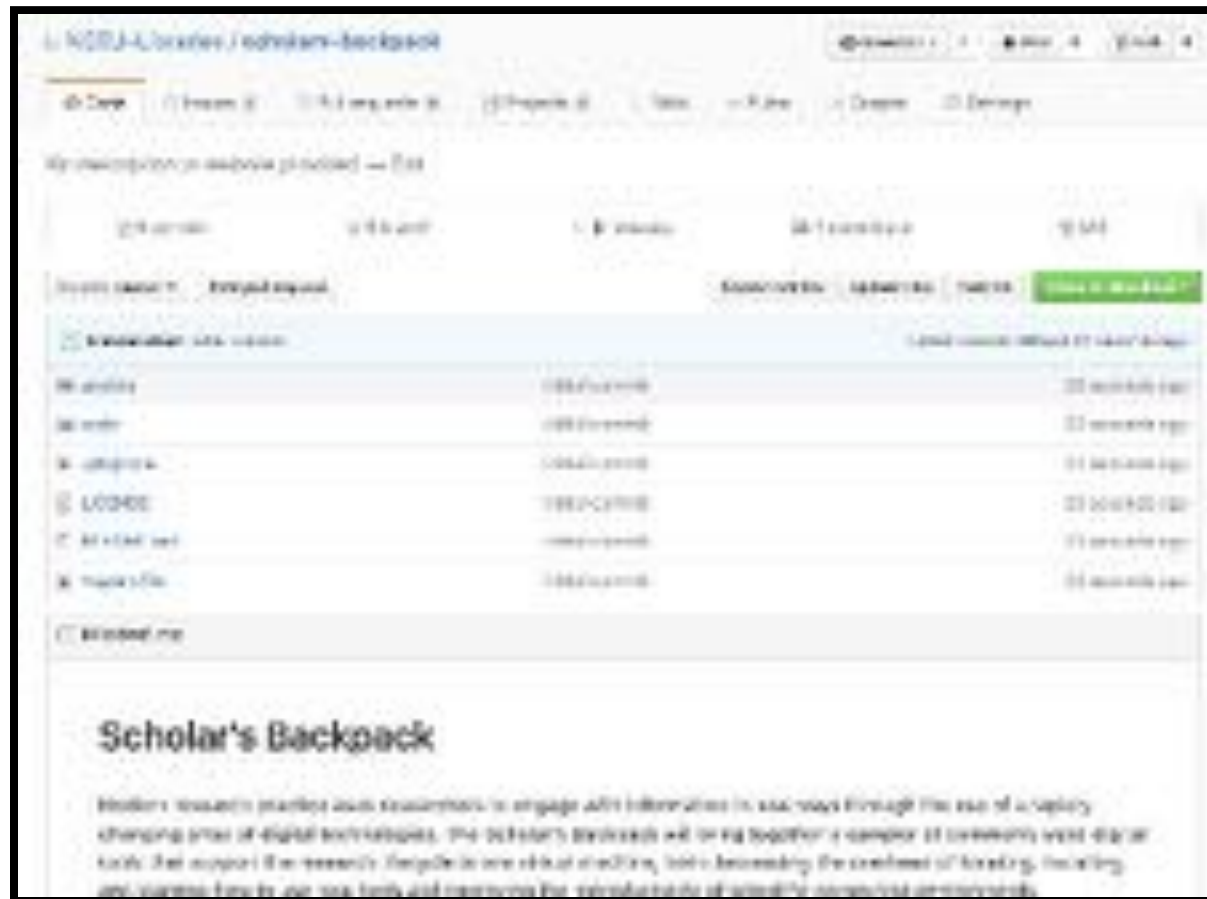
# Rise of Scholarly Code



# Researcher Challenges

- Consistency across lab environments
- Ability to see results of code
- Consistency across time
- Ease of collaboration

[github.com/NCSU-Libraries/scholars-backpack](https://github.com/NCSU-Libraries/scholars-backpack)



# Features

- Python 3
- Jupyter Notebook Server
- R and R Studio
- Example Notebooks

# Benefits

- Decrease overhead
- Improve reproducibility
- Encourage collaboration

# Vagrant



Create and configure lightweight,  
**reproducible**, and portable  
development environments.

# Usage

- Easy installation through binary package.
- Flexible configuration via **text file**.
- Single command: ``vagrant up``

```

1  vagrant.configure(2) do |config|
2
3      config.vm.box = "bco-ubuntu/centos72-desktop"
4
5      config.vm.provider "virtualbox" do |v|
6          v.guest__ = true
7      end
8
9      config.vm.hostname = "scholars-backpack"
10
11      config.vm.network "private_network", ip: "192.168.56.10"
12
13      config.vm.network "forwarded_port", guest: 80, host: 1080, auto_correct: true
14      config.vm.network "forwarded_port", guest: 1338, host: 10487, auto_correct: true
15      config.vm.network "forwarded_port", guest: 5000, host: 10482, auto_correct: true
16      config.vm.network "forwarded_port", guest: 8888, host: 10484, auto_correct: true
17
18      config.vm.provider "ansible_local" do |ansible|
19          ansible.playbook = "ansible/playbooks.yml"
20          ansible.inventory_path = "ansible/inventories/development.ini"
21          ansible.limit = "all"
22      end
23  end

```



```

9  vagrant.configure[2] do |config|
10
11      config.vm.box = "bbs-ubuntu/canbus72-desktop"
12
13      config.vm.provider "virtualbox" do |v|
14          v.guest__ = true
15      end
16
17      config.vm.hostname = "scholarship-backpack"
18
19      config.vm.network "private_network", ip: "192.168.56.10"
20
21      config.vm.network "forwarded_port", guest: 80, host: 16182, auto_correct: true
22      config.vm.network "forwarded_port", guest: 1000, host: 16182, auto_correct: true
23      config.vm.network "forwarded_port", guest: 5000, host: 16182, auto_correct: true
24      config.vm.network "forwarded_port", guest: 8080, host: 16184, auto_correct: true
25
26      config.vm.provider "ansible_local" do |ansible|
27          ansible.playbook = 'ansible/playbooks.yml'
28          ansible.inventory_path = 'ansible/inventories/development.ini'
29          ansible.limit = 'all'
30      end
31  end

```

```

9  vagrant.configure[2] do |config|
10
11     config.vm.box = "bbs-ubuntu/canonical-22-desktop"
12
13     config.vm.provider "virtualbox" do |v|
14         v.guest__add__ = true
15     end
16
17     config.vm.hostname = "scholar-backpack"
18
19     config.vm.network "private_network", ip: "192.168.56.10"
20
21     config.vm.network "forwarded_port", guest: 80, host: 16182, auto_correct: true
22     config.vm.network "forwarded_port", guest: 1000, host: 16182, auto_correct: true
23     config.vm.network "forwarded_port", guest: 5000, host: 16482, auto_correct: true
24     config.vm.network "forwarded_port", guest: 8080, host: 16484, auto_correct: true
25
26     config.vm.provider "ansible_local" do |ansible|
27         ansible.playbook = "ansible/playbooks.yml"
28         ansible.inventory_path = "ansible/inventories/development.ini"
29         ansible.limit = "all"
30     end
31 end

```

# Ansible

"Automation engine" for provisioning  
and configuration management.

# Provisioning

"To make something available."

Installation!

# Configuration Management

"Establish and maintain **consistency** of an environment."

# playbook.yml

```
1  ---
2  - hosts: scholars-backpack
3    gather_facts: yes
4    become: yes
5    become_method: sudo
6    become_user: root
7    vars:
8      project_name: "scholars-backback"
9    roles:
10      - role: basic-setup
11      - role: python
12      - role: jupyter
13      - role: r
14      - role: desktop
15    handlers:
16      - include: handlers/handlers.yml
```

```
1  ***
2  - name: install epel-release
3    yum:
4      name: epel-release
5      state: present
6  - name: install required packages using yum
7    yum:
8      pkg: "{{item}}"
9      state: present
10     update_cache: yes
11   with_items:
12     - firefox
13     - git-core
14     - vim
15     - curl
```

```
50 - name: set default vagrant ssh directory
51   lineinfile: dest=/home/vagrant/.bashrc line="cd /vagrant/code"
52   tags:
53     - set_default_dir
54 - name: set $DISPLAY environment variable
55   lineinfile: dest=/home/vagrant/.bashrc line="export DISPLAY=:0.0"
```



Future Work

# Richer Environment

- Broader scientific computing
- Improved adherence to best practices
- Docker containers for portability

# Embedded Use

- Curricular use
- Laboratory use

# Summary

Emphasis on reproducibility has  
ignited a shift toward new  
practices.

With these new practices come  
new requirements for  
researchers.

Reproducible and portable  
computing environments are  
critical for future success.

Tools like Vagrant and Ansible  
can help researchers develop  
the scientific environments they  
need to be productive.



# Thanks!

eka\_grguric@ncsu.edu | @egrguric

alison\_blaine@ncsu.edu | @Bigggggggg\_Al

bret\_davidson@ncsu.edu | @brtdvdsn

[github.com/NCSSU-Libraries/scholars-backpack](https://github.com/NCSSU-Libraries/scholars-backpack)