

# Introduction to the Whole Tale Project (Part 2)

Victoria Stodden  
CIRSS Seminar  
September 2, 2016



I L L I N O I S

School of  
**Information Sciences**

The iSchool at Illinois

# Whole Tale “Philosophy”



- Why a platform?
- Computation is near-ubiquitous in research, yet we have few best practices or dissemination standards
- And it’s complex! Small changes in a computational implementation or in the data can have a dramatic impact on the result

# Some Bold Assertions



- Software is used as a tool of discovery in nearly all research today.
- When software is a key part of the discovery process, it should be subject to the same philosophy of transparency as any method.
- Software is an integral and inseparable component of the computational infrastructure in which most research takes place.
- Computational research is embedded in a social structure which includes many stakeholders.

# Example: Facebook Study



## Experimental evidence of massive-scale emotional contagion through social networks

Adam D. I. Kramer<sup>a,1</sup>, Jamie E. Guillory<sup>b,2</sup>, and Jeffrey T. Hancock<sup>b,c</sup>

<sup>a</sup>Core Data Science Team, Facebook, Inc., Menlo Park, CA 94025; and Departments of <sup>b</sup>Communication and <sup>c</sup>Information Science, Cornell University, Ithaca, NY 14853

Edited by Susan T. Fiske, Princeton University, Princeton, NJ, and approved March 25, 2014 (received for review October 23, 2013)

whereas people in the negativity-reduced condition should express increased positivity. Emotional expression was modeled, on a per-person basis, as the percentage of words produced by that person during the experimental period that were either positive or negative. Positivity and negativity were evaluated separately given evidence that they are not simply opposite ends of the same spectrum (8, 10). Indeed, negative and positive word use scarcely correlated [ $r = -0.04$ ,  $t(620,587) = -38.01$ ,  $P < 0.001$ ].

We examined these data by comparing each emotion condition to its control. After establishing that our experimental groups did not differ in emotional expression during the week before the experiment (all  $t < 1.5$ ; all  $P > 0.13$ ), we examined overall posting rate via a Poisson regression, using the percent of posts omitted as a regression weight. Omitting emotional content reduced the amount of words the person subsequently produced, both when positivity was reduced ( $z = -4.78$ ,  $P < 0.001$ ) and when negativity was reduced ( $z = -7.219$ ,  $P < 0.001$ ). This effect occurred both when negative words were omitted (99.7% as many words were produced) and when positive words were omitted (96.7%). An

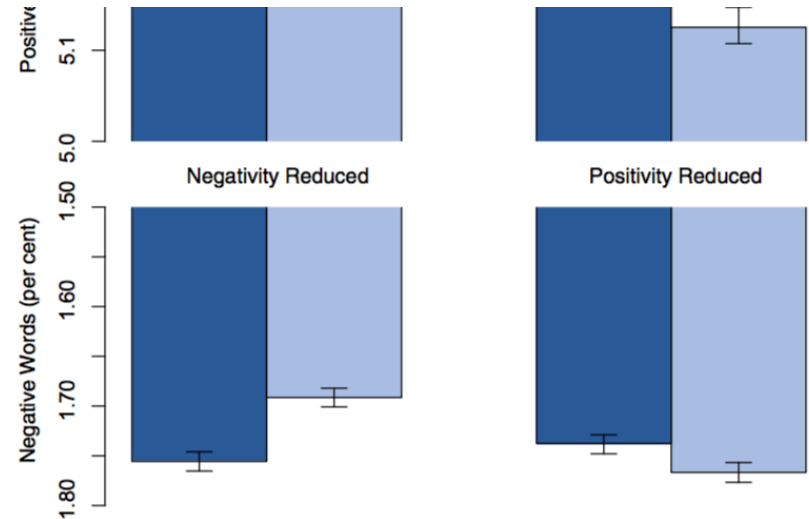


Fig. 1. Mean number of positive (*Upper*) and negative (*Lower*) emotion words (percent) generated per person, by condition. Bars represent standard errors.

# Dissemination is Incomplete



- Publications are missing details that are necessary to understand and verify published findings...
  - => credibility crisis in computational science
- How to reproduce the result? What's needed?

# Computational Reproducibility



Traditionally two branches to the scientific method:

Branch 1 (deductive): mathematics, formal logic.

Branch 2 (empirical): statistical analysis of controlled experiments.

Now, new branches due to technological changes?

Branch 3,4? (computational): large scale simulations / data driven computational science.

# The Ubiquity of Error



The central motivation for the scientific method is to root out error:

- Deductive branch: the well-defined concept of the proof,
- Empirical branch: the machinery of hypothesis testing, appropriate statistical methods, structured communication of methods and protocols.

**Claim: Computation presents only a *potential* third/ fourth branch of the scientific method, until the development of comparable standards.**

# Whole Tale?



## Proposed Solution:

- Capture computational steps / provide compute environment
- Provide unique identifiers to data/code/workflows associated with results
- Provide links to embed in the publication for discoverability
- Preserve digital scholarly objects



# So it all sounds pretty simple..



- What about big data?
- Complex codes?
- Reuse and bug fixes?
- Meta-analysis?
- Working with external groups, such as publishers?
- Incentives? What if they don't come?
- Allocating resources? Sustainability models?
- What does citation mean and how are contributions to be rewarded?

# Incompleteness



- “I ran all the stuff and it’s still the wrong answer!”
- “I got a different result, but I used your code and data!”
- “Your code doesn’t work!”
- “Where’s all the documentation? And the metadata? I can’t figure this thing out.”