



Empirical Methods

Literature Reviews and Theory

Bogdan Vasilescu

09-06-2018

(many slides by Herbsleb and Easterbrook)

Agenda for Today

- Homework debrief
- Literature reviews
- Formulating research questions
- Theory in empirical research

Easterbrook et al Chapter

- Selecting Methods -

- Controlled Experiments (Quasi-experiments):
 - determine precisely how variables are related
 - or whether a cause–effect relationship exists
- Case Studies (Exploratory/Confirmatory)
 - offer in-depth understanding of how and why certain phenomena occur
- Survey Research
 - identify the characteristics of a broad population

Easterbrook et al Chapter

- Selecting Methods -

- Ethnographies
 - study a community of people to understand how the members of that community make sense of their social interactions
- Action Research
 - attempt to solve a real-world problem while simultaneously studying the experience of solving the problem

Bogart, et al “How to Break an API”

- What is the point of this paper?
- What is the methodology?
 - Why this choice of method?
- Do you trust the results?
 - Why or why not?
 - What are the risks of being misled?
 - How do you evaluate a study with this type of methodology?
 - What does it tell about other ecosystems?

Raemaekers, et al “Semantic Versioning versus Breaking Changes”

- What is the point of this paper?
- What is the methodology?
 - Why this choice of method?
- Do you trust the results?
 - Why or why not?
 - What are the risks of being misled?
 - How do you evaluate a study with this type of methodology?
 - What does it tell about other ecosystems?

Discussion

- How is the data collected?
- How is the data analyzed?
- Where do hypotheses come from?
- Reproducibility?
- Complementarity of methods?

Agenda for Today

- ~~Homework debrief~~
 - Literature reviews
 - Formulating research questions
 - Theory in empirical research

Choosing a Research Topic

- *Can* it be studied?

vs

- *Should* it be studied?
 - Does it add to the body of knowledge?
 - Who else besides you would care about results?

Literature Review

- Helps you choose a research topic:
 - Determine if the topic is worth studying
 - Limit the scope
- Ask yourself: How does this project contribute to the literature?
 - Addresses a new topic
 - Uses new data collection method
 - Extends the discussion
 - Replicates a study in a new situation
 - Refines / extends a theory

Purpose of Lit Review

- Share with reader results from related studies
- Relate your study to literature, filling in gaps
- Framework for establishing the importance of your study
- Benchmark for comparing results
- **Provide direction for your research questions and hypotheses**
 - E.g., describe the theory that will be used

Forms of Lit Review

- Integrate what others have done and said
- Criticize prior work
- Build bridges between related topics
- Identify the central issues in a field

How to Structure Lit Review

- Hook
- Coherent presentation of what has been done
- The gap, the need for exactly what you did

Chen et al, “Short and Tweet”

- What is the hook?
- How is the literature organized?
- What is the gap being addressed?

Chen et al, “Short and Tweet”

Introduction

- Information streams are increasingly popular.
- With an abundance of information comes the scarcity of attention.
- Need to filter the stream down.
- One approach is to recommend interesting content to users to better direct their attention.

Chen et al, “Short and Tweet”

- Recommenders as a solution to attention scarcity have been studied for years.
- Perhaps the most well-known approach is collaborative filtering (CF) - infers preference similarity from the overlap of rated items across users
- CF recommenders commonly suffer from little user rating overlap early on (“cold-start”); a common solution is to use other information
- There is a wealth of research on recommenders that utilize the textual content of items.
- Recommendations can be generated from explicit social information and social processes as well.

The Hook

- Because Twitter has both textual and social information available, key parts of the past work described above may be applicable for a Twitter recommender.
- However, most of them have not yet been implemented and evaluated on Twitter or information streams in general.
- As a result, it is unclear whether these techniques function well given the differences between their original domains and Twitter, or if some techniques need to be changed to fit the needs of Twitter users.
- Our work not only depict the design space for a Twitter recommender, but also better inform designers of recommenders for other information streams.

Mockus et al, "Two Case Studies"

- Open source is often characterized as a fundamentally new way to develop software
- The open source development process is radically different from the usual industrial style of development:
 - extreme case of geographically distributed development, where developers work in arbitrary locations, rarely or never meet face to face, and coordinate their activity almost exclusively by means of email and bulletin boards

The Hook

- What is perhaps most surprising about the process is that it lacks many of the traditional mechanisms used to coordinate software development, such as plans, system-level design, schedules, and defined processes.
- These “coordination mechanisms” are generally considered to be even more important for geographically distributed development than for colocated development [Herbsleb and Grinter 1999], yet here is an extreme case of distributed development that appears to eschew them all.
- Despite the very substantial weakening of traditional ways of coordinating work, the results from OSS development are often claimed to be equivalent, or even superior to software developed more traditionally.

The Hook (2)

- Although this existence proof (Apache, Linux) means that OSS processes can, beyond a doubt, produce high-quality and widely deployed software, the exact means by which this has happened, and the prospects for repeating OSS successes, are frequently debated
- If OSS really does pose a major challenge to the economics and the methods of commercial development, it is vital to understand it and to evaluate it.

Discussion

- Contrast these two literature reviews
 - How much prior work was published?
 - What kinds of gaps and questions were the papers addressing?
 - How did the authors choose the papers they discussed, and the specific points they focused on?

Agenda for Today

- ~~Homework debrief~~
- ~~Literature reviews~~
- Formulating research questions
- Theory in empirical research

Scenarios - Jane

- Topic: Effectiveness of a novel fisheye-view file navigator (compact format, less scrolling)
- Intuition: fisheye-view file navigator is more efficient for file navigation
- Critics: the more compact information is difficult to read; developers will not adopt it over the traditional file navigator.
- How can Jane find evidence that supports or refutes her intuition?

Scenarios - Joe

- Topic: understanding how developers in industry use (or not) UML diagrams during software design
- Research goals:
 - Explore how widely UML diagrams are used
 - Explore how UML diagrams are used as collaborative shared artefacts during design
- What strategy can Joe follow?

RQs – First Attempt

- Jane: “Is a fisheye-view file navigator more efficient than the traditional view for file navigation?”
- Joe: “How widely are UML diagrams used as collaborative shared artifacts during design?”
- Thoughts?

RQs – First Attempt

- Both questions are vague, because they make assumptions about:
 - the phenomena to be studied
 - the kinds of situation in which they occur

Better RQs

If Little Prior Knowledge

- Existence questions:
 - “Is file navigation something that (certain types of programmers) actually do?”
 - “Is efficiency actually a problem in file navigation?”
 - “Do collaborative shared artifacts actually exist?”
- Description and Classification questions
 - “How can we measure efficiency for file navigation?”
 - “What are all the types of collaborative shared artifacts?”
- Descriptive-Comparative:
 - “How do fisheye views differ from conventional views?”
 - “How do UML diagrams differ from other representations of design information?”

Better RQs

If Some Understanding of the Phenomena

- Frequency and distribution questions
 - “How many distinct UML diagrams are created in large software companies?”
- Descriptive-Process questions
 - “How do programmers navigate files with existing tools?”
- Relationship questions
 - “Does efficiency in file navigation correlate with the programmer’s familiarity with the environment?”
 - “Do managers’ claims about how often they use UML correlate with the actual use of UML?”

- Causality questions:
 - “Do fisheye-views cause an improvement in efficiency for file navigation?”
- Causality-Comparative questions:
 - “Do fisheye-views cause programmers to be more efficient at file navigation than conventional views?”
- Causality-Comparative Interaction questions:
 - “Do fisheye-views cause programmers to be more efficient at file navigation than conventional views when programmers are distracted?”
- Design questions:
 - “What is an effective way for teams to represent design knowledge to improve coordination?”

Agenda for Today

- ~~Homework debrief~~
- ~~Literature reviews~~
- ~~Formulating research questions~~
- Theory in empirical research

The Role of Theory

- A scientific theory identifies and defines a set of phenomena; makes assertions about their nature and the **causal** relationships between them.
- *Positivism*: science - verifying theories by testing hypotheses derived from them.
 - Strong predictive power
 - Generalized models of cause-and-effect as basis
- *Constructivism*: science - seeking local theories that emerge from (and explain) the data.
 - Strengthens an understanding of complex situations
 - Categorizations and analogies

Stol and Fitzgerald Paper

- Theory as both a driver for, and a result of empirical research
- Different research pathways involving theory are possible



You Gotta Have A Theory

Steve Easterbrook

sme@cs.toronto.edu

www.cs.toronto.edu/~sme



Meet Stuart Dent

→ Name:

- ↳ Stuart Dent (a.k.a. "Stu")

→ Advisor:

- ↳ Prof. Helen Back

→ Topic:

- ↳ Merging Stakeholder views in Model Driven Development

→ Status:

- ↳ 2 years into his PhD
- ↳ Has built a tool
- ↳ Needs an evaluation plan





Stu's Evaluation Plan

→ Formal Experiment

- ↳ Independent Variable: Stu-Merge vs. Rational Architect
- ↳ Dependent Variables: Correctness, Speed, Subjective Assessment
- ↳ Task: Merging Class Diagrams from two different stakeholders' models
- ↳ Subjects: Grad Students in SE
- ↳ H_1 : "Stu-Merge produces correct merges more often than RA"
- ↳ H_2 : "Subjects produce merges faster with Stu-Merge than with RA"
- ↳ H_3 : "Subjects prefer using Stu-Merge to RA"

→ Results

- ↳ H_1 accepted (strong evidence)
- ↳ H_2 & H_3 rejected
- ↳ Subjects found the tool unintuitive



Stu's Evaluation Plan

→ Formal Experiment

- ↳ Independent Variable: Stu-Merge vs. Rational Architect
- ↳ Dependent Variables: Correctness, Speed, Subjective Assessment
- ↳ Task: Merging Class Diagrams from two different stakeholders' models
- ↳ Subjects: Grad Students in SE
- ↳ H_1 : "Stu-Merge produces correct merges more often than RA"
- ↳ H_2 : "Subjects produce merges faster with Stu-Merge than with RA"
- ↳ H_3 : "Subjects prefer using Stu-Merge to RA"

→ Results

- ↳ H_1 accepted (strong evidence)
- ↳ H_2 & H_3 rejected
- ↳ Subjects found the tool unintuitive

What could have gone wrong?



Threats to Validity

→ Construct Validity

- ↳ What do we mean by a merge? What is correctness?
- ↳ 5-point scale for subjective assessment - insufficient discriminatory power
 - (both tools scored very low)

→ Internal Validity

- ↳ Confounding variables: Time taken to learn the tool; familiarity
- ↳ Subjects were all familiar with RA, not with Stu-merge

→ External Validity

- ↳ Task representativeness: class models were of a toy problem
- ↳ Subject representativeness: Grad students as sample of what population?

→ Theoretical Reliability

- ↳ Researcher bias: subjects knew Stu-merge was Stu's own tool

More on validity in the backup slides at the end of the talk



What went wrong?

→ What was the research question?

↳ "Is tool A better than tool B?"

→ What would count as an answer?

→ What use would the answer be?

↳ How is it a "contribution to knowledge"?

→ How does this evaluation relate to the existing literature?

Document (8).pdf



Experiments as Clinical Trials

Why would we expect it to be better?

Why do we need to know?

What will we do with the answer?

Is drug A better than drug B?

Better at doing what?

Better in what situations?

Better in what way?



Why would
we expect it
to be better?

You gotta have a theory!



Science and Theory

→ A (scientific) theory is:

- ↳ more than just a description - it explains and predicts
- ↳ Logically complete, internally consistent, falsifiable
- ↳ Simple and elegant.

→ Components of a theory:

- ↳ concepts, relationships, causal inferences
 - E.g. Conway's Law- structure of software reflects the structure of the team that builds it. A theory should explain why.

→ Theories lie at the heart of what it means to do science.

- ↳ Production of generalizable knowledge
- ↳ Scientific method ⇔ Research Methodology ⇔ Proper Contributions for a Discipline

Document (8).pdf

→ Theory provides orientation for data collection

- ↳ Cannot observe the world without a theoretical perspective



The Role of Theory Building

→ Theories allow us to compare similar work

- ↳ Theories include precise definition for the key terms
- ↳ Theories provide a rationale for which phenomena to measure

→ Theories support analytical generalization

- ↳ Provide a deeper understanding of our empirical results
- ↳ ...and hence how they apply more generally
- ↳ Much more powerful than statistical generalization

→ ...but in SE we are very bad at stating our theories

- ↳ Our vague principles, guidelines, best practices, etc. could be strengthened into theories
- ↳ Every tool we build represents a theory



Theories are good for generalization...

Statistical Generalization

- First level generalization:
 - ↳ From sample to population
- Well understood and widely used in empirical studies
- Can only be used for quantifiable variables
- Based on random sampling:
 - ↳ Standard statistical tests tell you if results on a sample apply to the whole population
- Not useful when:
 - ↳ You can't characterize the population
 - ↳ You can't do random sampling
 - ↳ You can't get enough data points

Analytical Generalization

- Second level generalization:
 - ↳ From findings to theory
- Applicable to quantitative and qualitative studies
- Compares findings with theory
 - ↳ Do the data support or refute the theory?
 - ↳ Or: do they support this theory better than rival theories?
- Supports empirical induction:
 - ↳ Evidence builds if subsequent studies also support the theory (& fail to support rival theories)
- More powerful than stats
 - ↳ Doesn't rely on correlations
 - ↳ Examines underlying mechanisms



Stu's Theory

→ Background Assumptions

- ↳ Large team projects, models contributed by many actors
- ↳ Models are fragmentary, capture partial views
- ↳ Partial views are inconsistent and incomplete most of the time

→ Basic Theory

- ↳ (Brief summary:)
- ↳ Model merging is an exploratory process, in which the aim is to discover intended relationships between views. 'Goodness' of a merge is a subjective judgment. If an attempted merge doesn't seem 'good', many need to change either of the models, or the way in which they were mapped together.

→ Derived Hypotheses

- ↳ Useful merge tools need to represent relationships explicitly
- ↳ Useful merge tools need to be complete (work for any models, even if inconsistent)



Why Build a Tool?

→ Build a Tool to Test a Theory

↳ Tool is part of the experimental materials needed to conduct your study

→ Build a Tool to Develop a Theory

↳ Theory emerges as you explore the tool

→ Build a Tool to Explain your Theory

↳ Theory as a concrete instantiation of (some aspect of) the theory

Why did we
build tools?

Why did Stu build a
c/ool?

A circular portrait of Steve Easterbrook, a man with glasses and short hair, looking directly at the camera.

Why do you
build tools?



Take home messages

Articulate the theory(s) underlying your work

Be precise about your research question

Be explicit about your philosophical stance

Use the theory to guide the study design

Test the Theory not the Tool

Borrowed Theories: Transparency and Signaling

- Transparency
 - “accurate observability, of an organization’s low-level activities, routines, behaviors, output, and performance”*
- + Politics, finance, government, workplace
- + Performing in front of an audience
- + Accountability, coordination
- May hurt creativity and streamlining
- Information overload

GitHub: Transparency

- Open source hosting environment
- 28 million users, 85 million repositories
- Social media functionality
- Transparency

Example: Information Overload

The Ruby Programming Language <http://www.ruby-lang.org/>

37,555 commits 34 branches 290 releases 36 contributors

branch: trunk

* common.mk: Specify dependencies for generated C sources. ...
akr authored 8 hours ago latest commit 2a87fdedb7

File	Description	Time Ago
benchmark	benchmark/bm_hash_aref_sym*.rb: force static symbols	a month ago
bin	erb: multiline anchors	a month ago
bootstraptest	normalize reference to Timeout::Error	a month ago
ccan	ccan/container_of: add container_of_or_null	2 months ago
coverage	fix a typo [ci skip]	2 months ago
cygwin	* cygwin/GNUmakefile.in (uncommon.mk): link *.res.o.	2 years ago
defs	id.def: attrset	2 days ago
doc	* ext/dl/*: remove DL as it is replaced by Fiddle.	15 days ago
enc	Mark auogenerated part.	13 hours ago
ext	* internal.h: Include ruby.h and ruby/encoding.h to be	12 hours ago

Code Pull Requests (73) Pulse Graphs
HTTPS clone URL: <https://github.com/ruby/ruby> Clone in Desktop Download ZIP

You can clone with **HTTPS** or **Subversion**.



Contributors

Commits

Code frequency

Punch card

Network

Members

Nov 21, 2004 – Nov 15, 2014

Contributions: Commits ▾

Contributions to master, excluding merge commits



dhh

3,414 commits / 202,650 ++ / 102,798 --

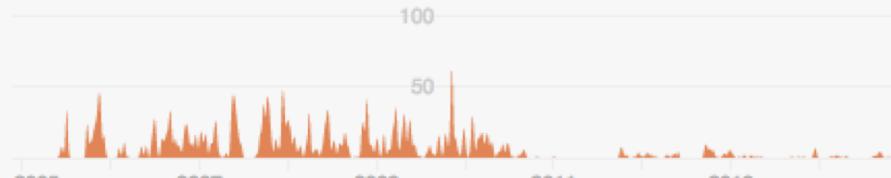
#1



jeremy

3,163 commits / 132,694 ++ / 110,179 --

#2



Contributors

Commits

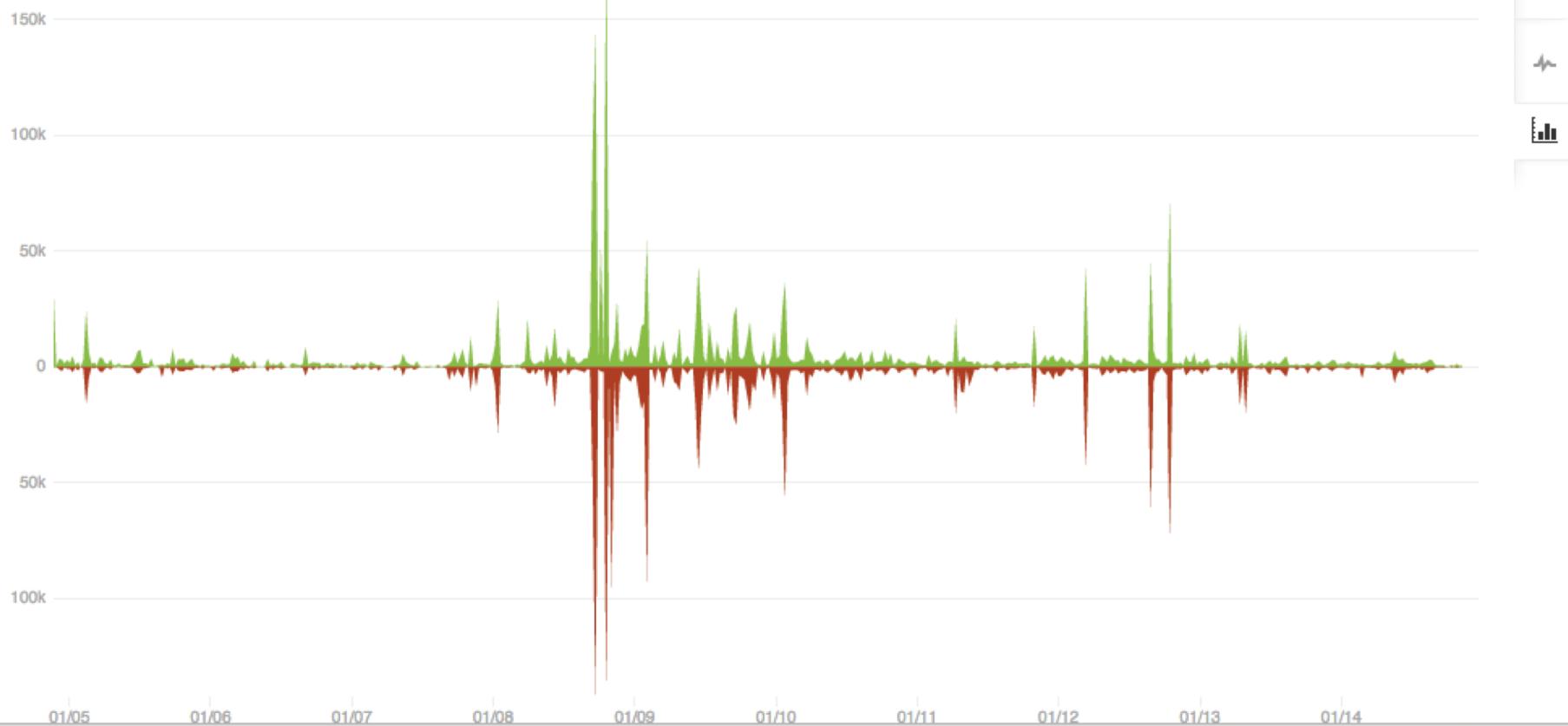
Code frequency

Punch card

Network

Members

Additions and Deletions per week



JonAbrams / synth

[Star](#) 952[Fork](#) 70[Contributors](#)[Commits](#)[Code frequency](#)[Punch card](#)[Network](#)[Members](#)

Owners

Sep

Oct

Nov

20 25 26 27 29 13 20 22 24 26 29 10 3

JonAbrams



etanlubeck

ky0615

DanDeMicco

kumagi

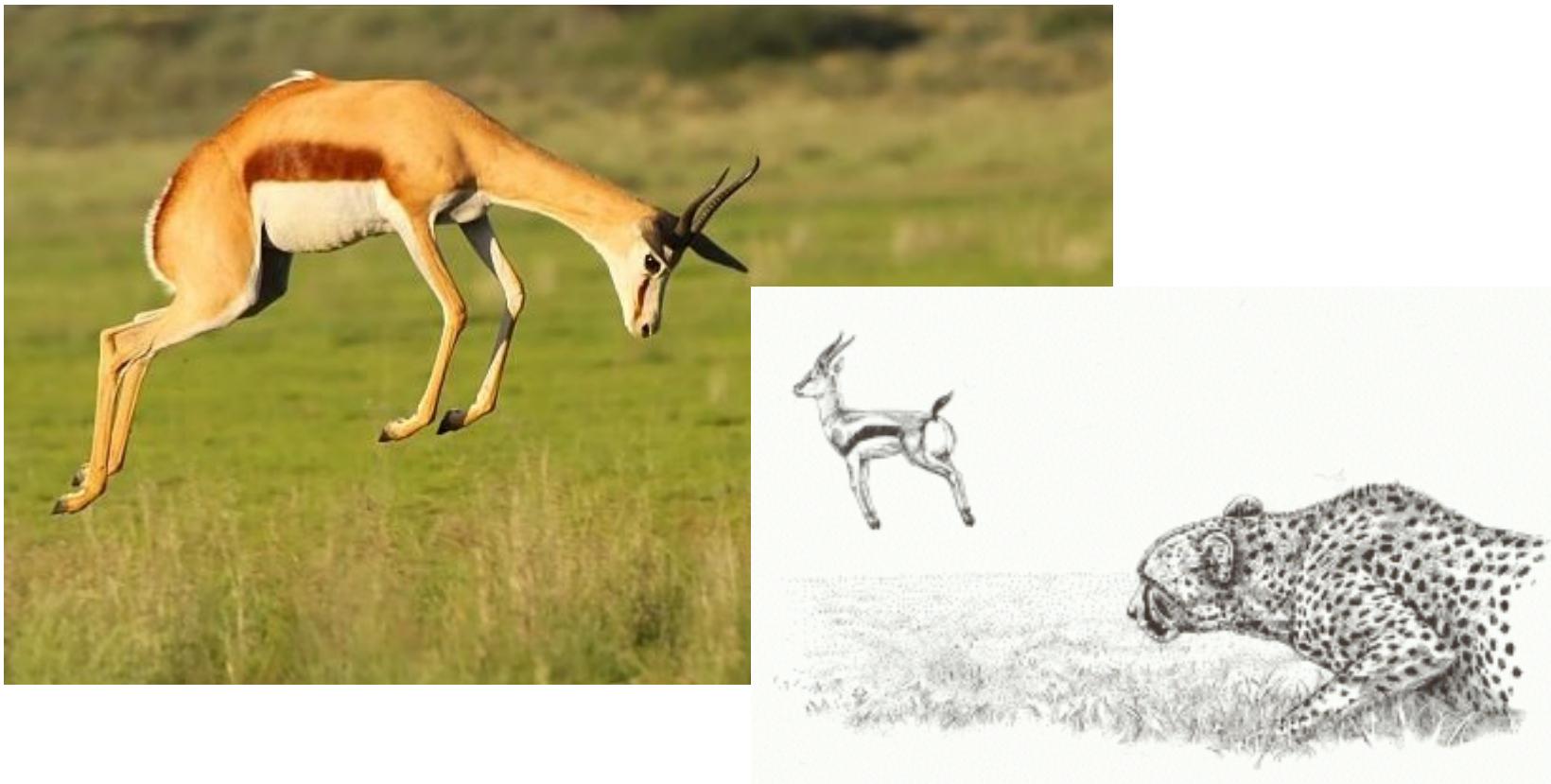
rzmz

stevemanuel

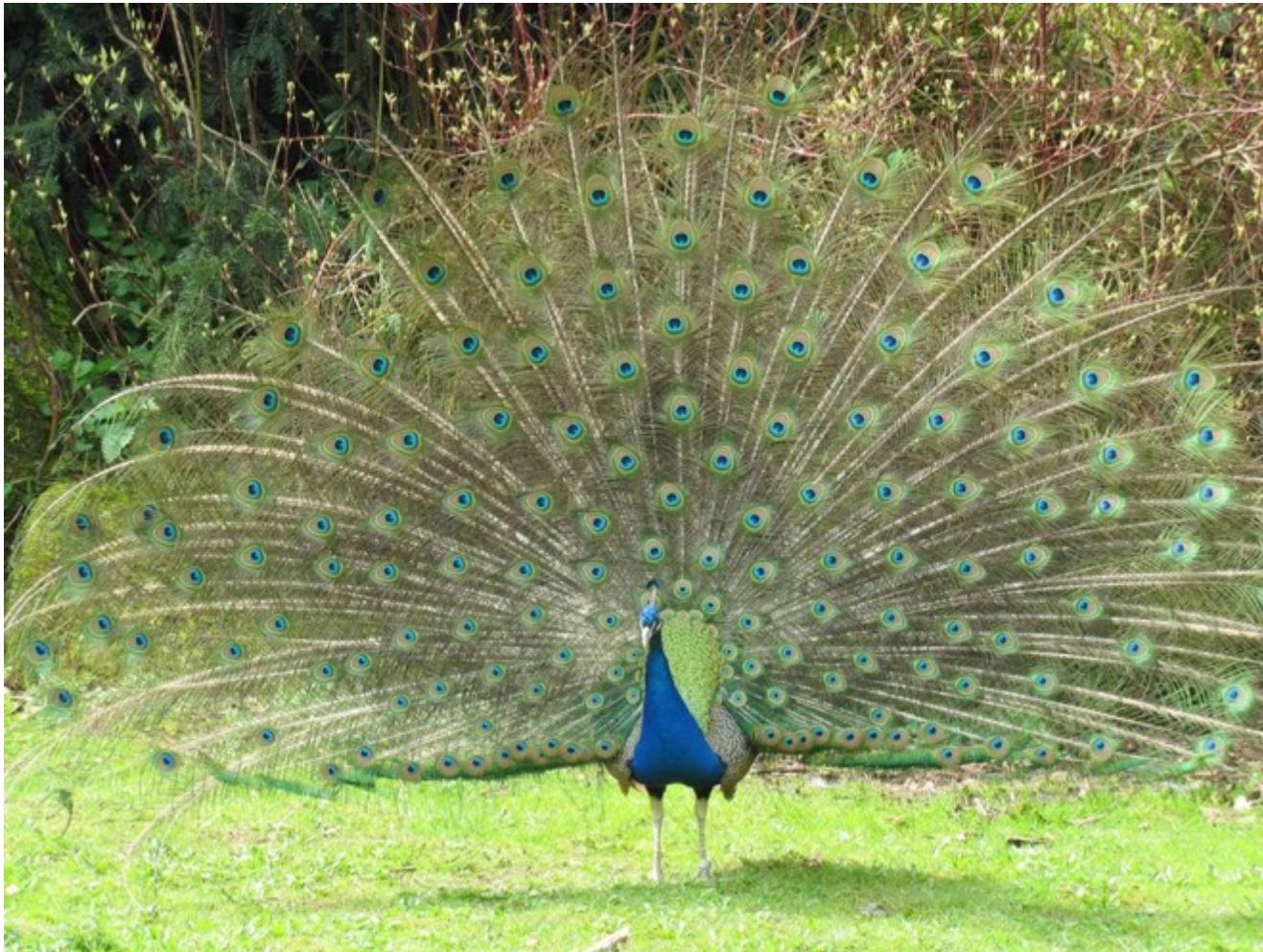
Transparency and Signaling

- Signals
 - Original idea from evolutionary biology
 - Visible clues that imply hidden quality
 - Types of signals
 - Assessment: visible clue cannot be produced without hidden quality
 - Conventional: meaning is agreed upon, will continue to exist only if enforced by norms

Stotting as Honest Signal



Peacock Tail Feathers as Handicap Signal



Avoid Colorful Snakes



a) Eastern coral snake (poisonous)

Copyright © 2007 Pearson Education Inc., publishing as Pearson Benjamin Cummings. All rights reserved.

Avoid Colorful Snakes



a) Eastern coral snake (poisonous)

Copyright © 2007 Pearson Education Inc., publishing as Pearson Benjamin Cummings. All rights reserved.



(b) Scarlet king snake (nonpoisonous)

Conventional Signals



James D. Herbsleb

Professor, Institute for Software Research
School of Computer Science
Carnegie Mellon University

[Home](#)
[Bio](#)
[Research](#)
[Publications](#)
[Courses](#)
[People](#)
[Presentations](#)
[Links](#)
[Contact](#)

Jim Herbsleb's Home Page

Software is in everything. Software shapes the digital environment, which in turn shapes how we find information, conduct commerce, share and socialize, do our work, and amuse ourselves.

Our old ways of designing software for a specific known purpose are no longer adequate.

Rather than ask

How can I specify, design, and build the system that my stakeholders need?

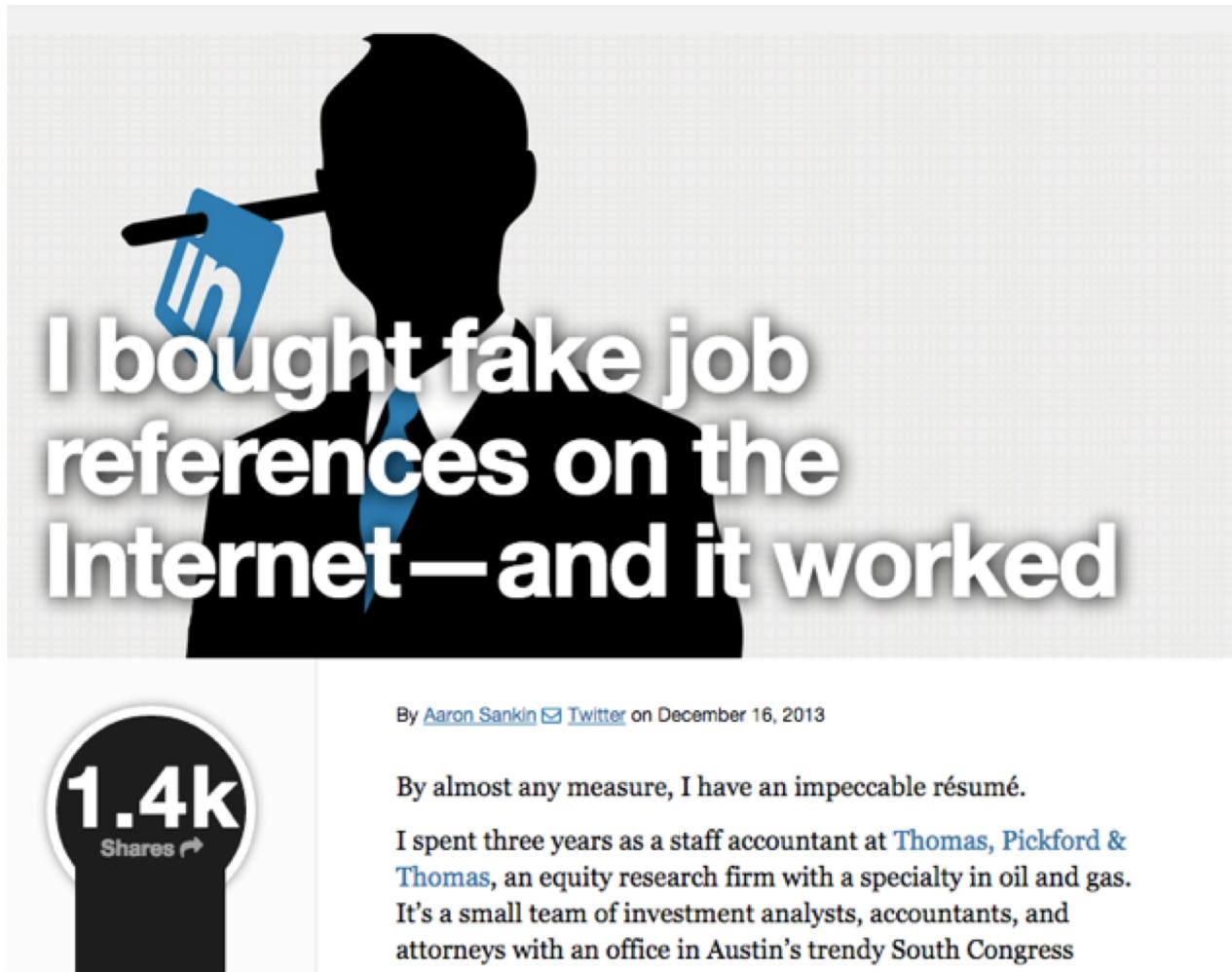
Maybe we should ask

How can I set up the socio-technical ecosystem that will allow users, developers, businesses, and everyone else to cooperate and compete to build what everyone needs?

Even though those needs are currently unknowable and evolving . . .

These are the kinds of questions that drive my research group. (See my [research page](#).) We are funded by the National Science

Conventional Signals – Trustworthy?



I bought fake job references on the Internet—and it worked

By [Aaron Sankin](#) [Twitter](#) on December 16, 2013

1.4k Shares

By almost any measure, I have an impeccable résumé.

I spent three years as a staff accountant at [Thomas, Pickford & Thomas](#), an equity research firm with a specialty in oil and gas. It's a small team of investment analysts, accountants, and attorneys with an office in Austin's trendy South Congress

[CareerExcuse.com](#)

[LOGIN](#)**PH: (800) 637-7709**[REQUEST A CALL](#)[Home](#)[Services](#)[About Us](#)[Contact](#)[Pricing](#)[Blog](#)

Professional Fake Job References in 3 Days or Less!

- We Answer Calls Per Your Resume
- Eliminate Gaps in Resume
- Helps you get more interviews faster
- Providing Fake Job References Since July, 2009

[Looking for References?](#)[Check Out Real Demo!](#)

Norms, Enforcement

Indiatimes | The Times of India | **The Economic Times** | More ▾ Log In / Join @ f t g+ in FOLLOW ET: f t s in

THE ECONOMIC TIMES

Jobs

04:04 PM | 14 Nov **EOB** SENSEX **28,046.66** ▲ 106.02 | NIFTY **8,389.90** ▲ 32.05 | GOLD (MCX) (Rs/10g.) **26,443.00** ▲ 621.00 | USD/INR **61.72** ▼ 0.15 | Login to Track your Investment

04:04 PM | 14 Nov **EOB** SENSEX **28,046.66** ▲ 106.02 | NIFTY **8,389.90** ▲ 32.05 | GOLD (MCX) (Rs/10g.) **26,443.00** ▲ 621.00 | USD/INR **61.72** ▼ 0.15 | Login to Track your Investment

You are here: Home > Collections

Fake CV? Chances are your company will catch you out

Mishita Mehra, ET Bureau Jan 8, 2012, 02.27AM IST

Tags: Sunday ET | Steve Jobs | resume | Fake CV | CV | company

Did you say deception? It is not even a lie, not in the lowly, how-can-you-do-this sort of way. Call it an exaggeration, if you must - the stretched truth that boosts your resume. So what if you forget the 'assistant' in your designation and just write manager or add a few thousands to your current pay. Or extend a job tenure to cover three months of vegetating before the TV - even Steve Jobs needed his psychedelic breaks, didn't he?

RELATED ARTICLES

Check on job applicants gets deeper as hiring activity goes...
November 28, 2009



For next time

- A mini lit review
 - The hook
 - Coherent presentation of literature
 - Identify the gap you are addressing

Summary

- Literature review
 - Identify gaps
 - Position your work
 - Present theory
 - Helps formulate research questions / hypotheses
- Theory in empirical research
 - Both a driver and a result
 - You gotta have one

Bonus: Theory versus Evidence

- Evidence-based medicine

To Bleed or not to Bleed . . .

- Late 18th century
- Francois Joseph Victor Broussais
 - Chief physician Paris military hospital
 - Promoted bleeding of “affected organ”
- Pierre-Charles-Alexandre Louis
 - Actual data collection about outcomes
 - Bleeding is not such a great idea

Statistics, Medicine, Science

- Pierre Louis promoted use of correlation of treatment and outcome to evaluate effectiveness
- Useful, but not science
 - Discovery of correlation not science
 - Science requires understanding the causal connection
- Joseph Lister – outcomes of antiseptic surgery in Edinburgh
 - Mortality rates decreased from 45.7% to 15%
 - Technique based on Louis Pasteur’s “germ theory”

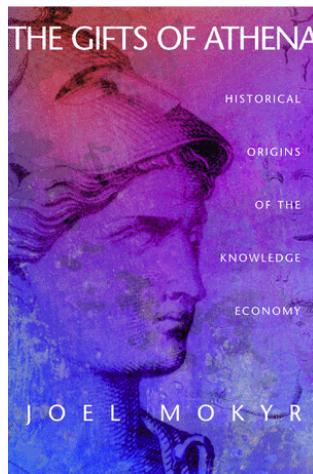
History of Useful Knowledge

Technique

- Make iron (from 2000 BC)
 - Mix ore, charcoal
 - Apply heat
 - Pour when ready

Epistemic Base

- Metallurgy
 - Eliminate phosphorus
 - Add carbon at right time
 - Reduce oxygen
 - Siemens Martin process (1865)



History of Useful Knowledge

Technique

- Analgesic (1763)
 - Ingest willow bark
 - Pain relief
 - Side effects

Epistemic Base

- Chemistry
 - Salicin
 - Explore related compounds
 - Salicylic acid (1835)

