ECE1724H S2: Empirical Software Engineering

Research Methods, Literature Review and Theory

Shurui Zhou

The Edward S. Rogers Sr. Department
of Electrical & Computer Engineering
UNIVERSITY OF TORONTO

- If you are in a different time zone, no need to attend the live lecture.
- We could arrange an office hour for you to participate and discuss.

How to Read a Paper

S. Keshav
David R. Cheriton School of Computer Science, University of Waterloo
Waterloo, ON, Canada
keshav@uwaterloo.ca

ABSTRACT

Researchers spend a great deal of time reading research papers. However, this skill is rarely taught, leading to much wasted effort. This article outlines a practical and efficient *three-pass method* for reading research papers. I also describe how to use this method to do a literature survey.

Categories and Subject Descriptors: A.1 [Introductory and Survey]

General Terms: Documentation.

Keywords: Paper, Reading, Hints.

1. INTRODUCTION

Researchers must read papers for several reasons: to review them for a conference or a class, to keep current in their field, or for a literature survey of a new field. A typical researcher will likely spend hundreds of hours every year reading papers.

Learning to efficiently read a paper is a critical but rarely taught skill. Beginning graduate students, therefore, must learn on their own using trial and error. Students waste much effort in the process and are frequently driven to frustration.

For many years I have used a simple approach to efficiently read papers. This paper describes the 'three-pass' approach and its use in doing a literature survey.

4. Glance over the references, mentally ticking off the ones you've already read

At the end of the first pass, you should be able to answer the *five Cs*:

1. *Category*: What type of paper is this? A measurement paper? An analysis of an existing system? A description of a research prototype?

2. *Context*: Which other papers is it related to? Which theoretical bases were used to analyze the problem?

3. *Correctness*: Do the assumptions appear to be valid?

4. *Contributions*: What are the paper's main contributions?

5. *Clarity*: Is the paper well written?

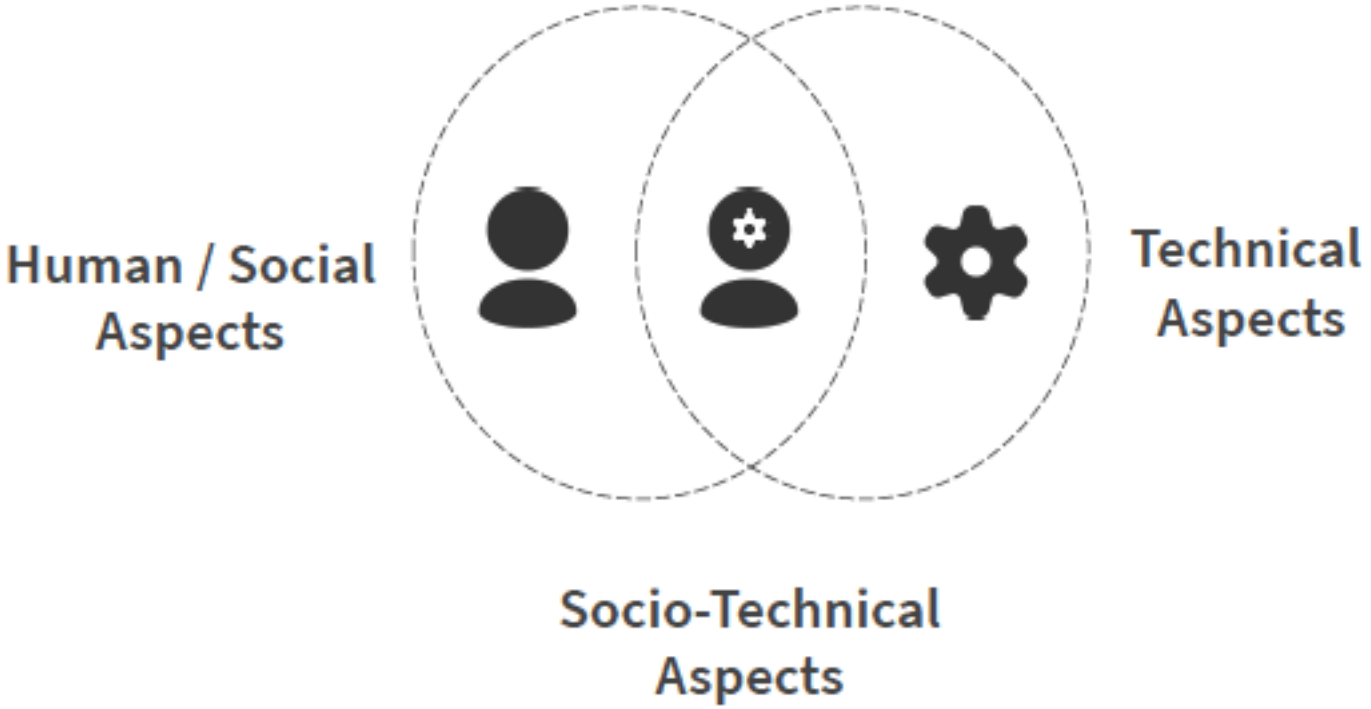
Using this information, you may choose not to read further. This could be because the paper doesn't interest you, or you don't know enough about the area to understand the paper, or that the authors make invalid assumptions. The first pass is adequate for papers that aren't in your research area, but may someday prove relevant.

Incidentally, when you write a paper, you can expect most reviewers (and readers) to make only one pass over it. Take care to choose coherent section and sub-section titles and

https://web.stanford.edu/class/cs244/papers/HowtoReadPaper.pdf

Any Questions?

Software Engineering **Design** Space



- **Social aspects of software engineering**
  - Human aspects of software engineering
  - Human-computer interaction
  - Distributed and collaborative software engineering
  - Agile methods and software processes
  - Software economics
  - C
  - E

☆ An Empirical Analysis of UI-based Flaky Tests
Alan Romano, Zihe Song, Sampath Grandhi, Wei Yang, Weihang Wang

☆ An Empirical Assessment of Global COVID-19 Contact Tracing Applications
Ruoxi Sun, Zach Wei Wang, Minhui (Jason) Xue, Gareth Tyson, Seyit Camtepe, Damith C. Ranasinghe

☆ An Empirical Study of Refactorings and Technical Debt in Machine Learning Systems
Yiming Tang, Raffi Khatchadourian, Mehdi Bagherzadeh, Rhia Singh, Ajani Stewart, Anita Raja
Media Attached

☆ An Empirical Study on Deployment Faults of Deep Learning Based Mobile Applications
Zhenpeng Chen, Huihan Yao, Yiling Lou, Yanbin Cao, Yuanqiang Liu, Haoyu Wang, Xuanzhe Liu

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 [*Stu-Merge*]
  - Needs an evaluation plan

Many available methods…

Common "in the lab" Methods	Common "in the wild" Methods
Controlled Experiments • Rational Reconstructions • Exemplars • Benchmarks • Simulations	Quasi-Experiments Case Studies Survey Research Ethnographies Action Research

- Artifact/Archive Analysis ("mining"!)

Qualitative vs Quantitative

- Often:
  - Words (qual) vs Numbers (quant)
  - Open-ended questions (qual interview questions)

    vs

    Closed-ended questions (quant hypotheses)

Mixed-Methods

- Qual. + Quant.
- Collecting both quantitative and qualitative data, integrating the two forms of data, and using distinct designs that may involve philosophical assumptions and theoretical frameworks.

Chapter 1 - Creswell, J. W., & Creswell, J. D. (2017). Research design: Qualitative, quantitative, and mixed methods approaches. Sage publications.

Agenda for Today

- Study Design Planning Checklist
- Homework debrief

Planning Checklist

- [x] Pick a topic
- [ ] Identify the research question(s)
- [ ] Check the literature
- [ ] Identify your philosophical stance
- [ ] Identify appropriate theories
- [ ] Choose the method(s)
- [ ] Design the study
  - Unit of analysis?
  - Target population?
  - Sampling technique?
  - Data collection techniques?
  - Metrics for key variables?
  - Handle confounding factors
- [ ] Critically appraise the design for threats to validity
- [ ] Get IRB approval
  - Informed consent?
  - Benefits outweigh risks?
- [ ] Recruit subjects / field sites
- [ ] Conduct the study
- [ ] Analyze the data
- [ ] Write up the results and publish them
- [ ] Iterate

Empirical Software Engineering at Microsoft Research

http://research.microsoft.com/en-us/projects/esm/

Christian Bird cbird@microsoft.com
Brendan Murphy bmurphy@microsoft.com
Nachiappan Nagappan nachin@microsoft.com
Thomas Zimmermann tzimmer@microsoft.com

Microsoft Research, Redmond, USA and Cambridge, UK

ABSTRA
We descr
neering (
our resea
research
triaging,
our relati
unique ab
the abilit
results of
our group

How Data Scientists Use Computational Notebooks for Real-Time Collaboration

APRIL YI WANG*, The University of Michig
ANANT MITTAL, The University of Michig
CHRISTOPHER BROOKS, The University
STEVE ONEY, The University of Michigan,

Effective collaboration in data science can lev
improve the quality and efficiency of the work
interactive solution for sharing and keeping tra
code, narrative text, visualizations, and other ric

A Multimodal Experimental Approach to Study CAD Collaboration

Vrushank S. Phadnis[1], David R. Wallace[2] and Alison Olechowski[3]

[1]Massachusetts Institute of Technology, vphadnis@mit.edu
[2]Massachusetts Institute of Technology, drwallac@mit.edu
[3]University of Toronto, olechowski@mie.utoronto.ca

Corresponding author: David Wallace, drwallac@mit.edu

Abstract. Computer-Aided Design (CAD) collaboration has been studied since the early days of CAD research, but standardized metrics and experimental procedures for synchronous, cloud-based CAD research are lacking in the literature. In this work, we lay out an empirical approach to investigate collaboration in CAD. Our work is unique in its relevance to the nuances of synchronous CAD collaboration. We first define metrics of interest: speed, quality, communication, satisfaction, and UI (user interaction). We then introduce an experimental toolkit that leverages automated and manual data capture methods. Lastly, we deploy our toolkit in a pilot study setting to reveal preliminary insights and validate the workings of our method. Although preliminary, our findings suggest that pairs were slower than single CAD users because of coordination overheads involving communication and model-tree-scanning.

What type of question are you asking?

- **Existence:**
  - Does X exist?
- **Description & Classification**
  - What is X like?
  - What are its properties?
  - How can it be categorized?
  - How can we measure it?
  - What are its components?
- **Descriptive-Comparative**
  - How does X differ from Y?
- **Frequency and Distribution**
  - How often does X occur?
  - What is an average amount of X?
- **Descriptive-Process**
  - How does X normally work?
  - By what process does X happen?
  - What are the steps as X evolves?
- **Relationship**
  - Are X and Y related?
  - Do occurrences of X correlate with occurrences of Y?
- **Causality**
  - Does X cause Y?
  - Does X prevent Y?
  - What causes X?
  - What effect does X have on Y?
- **Causality-Comparative**
  - Does X cause more Y than does Z?
  - Is X better at preventing Y than is Z?
  - Does X cause more Y than does Z under one condition but not others?
- **Design**
  - What is an effective way to achieve X?
  - How can we improve X?

What type of question are you asking?

Exploratory

→Existence:
- Does X exist?

→Description & Classification
- What is X like?
- What are its properties?
- How can it be categorized?
- How can we measure it?
- What are its components?

→Descriptive-Comparative
- How does X differ from Y?

Correlation

→Relationship
- Are X and Y related?
- Do occurrences of X correlate with occurrences of Y?

Causal Relationship

→Causality
- Does X cause Y?
- Does X prevent Y?
- What causes Y?
- What effect does X have on Y?

→Causality-Comparative
- Does X cause more Y than does Z?
- Is X better at preventing Y than is Z?
- Does X cause more Y than does Z under one condition but not others?

Descriptive-Process

→Frequency and Distribution
- How often does X occur?
- What is an average amount of X?

→Descriptive-Process
- How does X normally work?
- By what process does X happen?
- What are the steps as it evolves?

Design

→Design
- What is an effective way to achieve X?
- How can we improve X?

Putting the Question in Context

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?

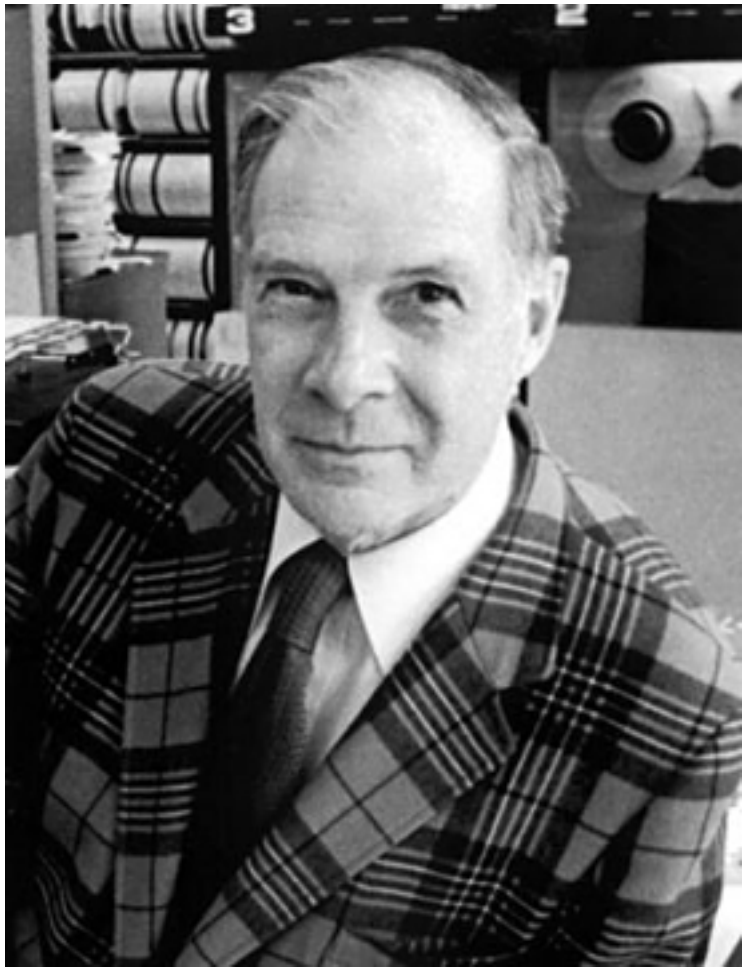
Systematical Literature Review (SLR)

- 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

"Indeed, one of my major complaints about the computer field is that whereas Newton could say, "If I have seen a little farther than others, it is because I have stood on the shoulders of giants,"

I am forced to say, "Today we stand on each other's feet." Perhaps the central problem we face in all of computer science is how we are to get to the situation where we build on top of the work of others rather than redoing so much of it in a trivially different way. Science is supposed to be cumulative, not almost endless duplication of the same kind of things".

---- Richard Hamming 1968 Turning Award Lecture



History of SLRs in Software Engineering

- Rather new, only since the 90's
- Inspired by evidence based medicine
- *'the conscientious, explicit, judicious use of current best evidence in making decisions about the care of individual patients.'* (Sackett et al. 1996)
-

Kassa *BMC Infectious Diseases* (2018) 18:216
https://doi.org/10.1186/s12879-018-3126-5

BMC Infectious Diseases

RESEARCH ARTICLE Open Access

Mother-to-child transmission of HIV infection and its associated factors in Ethiopia: a systematic review and meta-analysis

Getachew Mullu Kassa

Abstract

Background: Mother-to-child transmission (MTCT) is the main mode of HIV transmission in children under 15 years old. This problem is significant in the Sub-Saharan African countries, where more than 80% of children living with HIV are found. Previous studies in Ethiopia present inconsistent and inconclusive findings on the prevalence and associated factors of MTCT of HIV. Therefore, this study was conducted to determine the pooled prevalence of MTCT of HIV and its associated factors in Ethiopia.

Methods: Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guideline was followed. All published studies were retrieved using relevant search terms in MEDLINE, PUBMED, Cochrane Library, EMBASE, Google Scholar, CINAHL, and African Journals Online databases. Joanna Briggs Institute Meta-Analysis of Statistics Assessment and Review Instrument (JBI-MAStARI) was used to critically appraise articles. STATA version 14 software was used to perform the Meta-analysis. The I^2 statistics was used to test heterogeneity and publication bias was assessed using Begg's and Egger's tests. Odds ratio (OR) with 95% confidence interval (CI) was presented using forest plots.

Results: A total of nine studies, 3688 mother-baby pairs, were included in this meta-analysis. The pooled prevalence of MTCT of HIV in Ethiopia was 9.93% (95% CI: 7.29, 12.56). The subgroup analysis showed a higher prevalence of MTCT of HIV in Dire Dawa City Administration (15.7%) and lowest in Southern Nations, Nationality and Peoples Region (SNNPR) (4.16%). Associated factors with MTCT of HIV include: mixed feeding, OR = 7.46 (95%CI: 4.71, 11.81), absence of infant ARV prophylaxis, OR = 7.89 (95%CI: 4.32, 14.42), home delivery, OR = 5.08 (95%CI: 2.32, 11.15), and absence of maternal PMTCT intervention, OR = 7.13 (95% CI: 3.31, 15.35).

Conclusions: Almost one in ten HIV exposed infants become HIV positive in Ethiopia. Factors like: mixed feeding, the absence of infant ARV prophylaxis, home delivery and absence of mother's PMTCT intervention were significantly associated with MTCT of HIV. Therefore, the governmental and non-governmental organizations need to focus on the identified factors and work towards improving the prevention of mother to child transmission of HIV (PMTCT) program.

Keywords: HIV, PMTCT, MTCT, Prevalence of MTCT of HIV, Vertical HIV transmission, HIV-exposed infant, Systematic review, Meta-analysis, Ethiopia

Correspondence: gechm2005@gmail.com
College of Health Sciences, Debre Markos University, Debre Markos, Ethiopia

5 reasons SLRs are a good thing

Researchers get an **overview** of their area of interest

Highlight areas for **further work**

Knowledge **organization**

Service to the research community

They are (usually) **well-cited**

https://research-seminar.github.io/slides/EiriniKalliamvakou_SystematicLiteratureReviews.pdf

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

Glossary (1 of 2)

- ***Primary study.*** (In the context of evidence) An empirical study investigating a specific research question.
- ***Secondary study.*** A study that reviews all the primary studies relating to a specific research question with the aim of integrating/synthesising evidence related to a specific research question.
- ***Tertiary study*** (also called a tertiary review) is a review of secondary studies related to the same research question. (also called a tertiary review). A review of secondary studies related to the same research question.

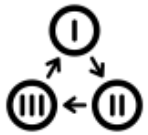
Glossary (2 of 2)

- "***Systematic mapping study*** (also referred to as a scoping study): A broad review of primary studies in a specific topic area that aims to identify what evidence is available on the topic." (Kitchenham and Charters, 2007)
- "***Systematic literature* review:** (also referred to as a systematic review). A form of secondary study that uses a well-defined methodology to identify, analyse and interpret all available evidence related to a specific research question in a way that is unbiased and (to a degree) repeatable." (Kitchenham and Charters, 2007)
- "***Reliability*:** Demonstrating that the operations of a study – such as the data collection procedures – can be repeated, with the same results." (Yin, 2009)
-

Typology of literature reviews

Narrative literature review

Situates a study within the relevant literature, non-systematic



Systematic literature review

Provides a comprehensive summary of literature

Systematic **mapping** review

Characterizes quantity and themes of research in an area

Systematic **scoping** review

Similar to mapping, but considered preliminary

For even more types: http://bit.ly/2h2IVqE

A common point of confusion!

- Systematic literature mapping studies → structure a research area
- Systematic literature reviews gather and synthesize evidence

	Mapping studies:	Systematic literature reviews:
Research questions	General questions about the topic, what has been done	More specific, aim to aggregate evidence
Search process	Considers the landscape of research/topics/area	Driven by a research question
Quality assessment	Less important to do (but may be discussed)	Rigor and relevance of primary studies is very important
Results	Descriptive	Theoretical insights, framework, synthesizes evidence, may lead to new hypotheses

General advice

- Do a mapping study before a systematic literature review…

Literature Map of Research

- A literature map is a visual summary of existing research on a topic
- The structure of the literature map may be:
  - a hierarchical pattern
  - a flowchart layout
  - a series of circles

Mapping study steps

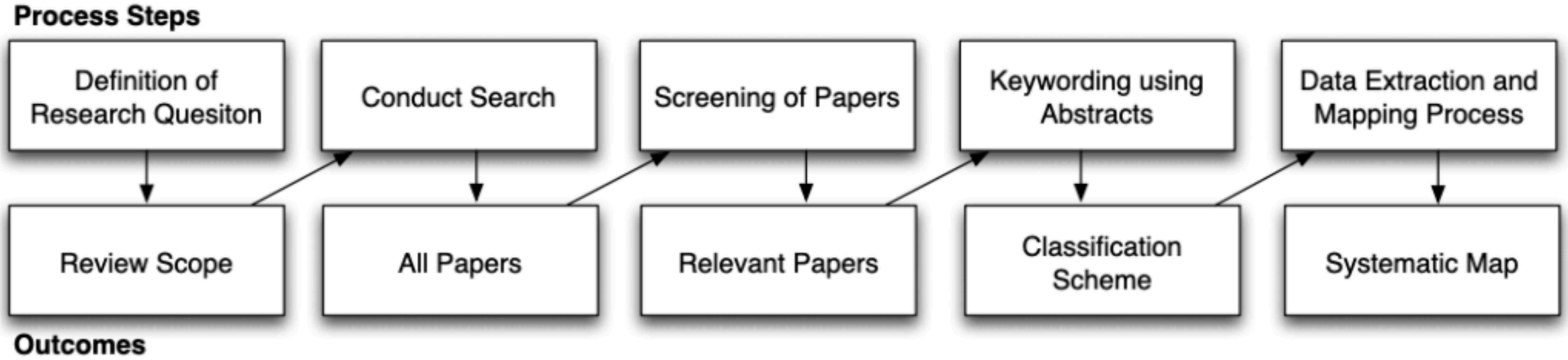


FIGURE 1: The Systematic Mapping Process

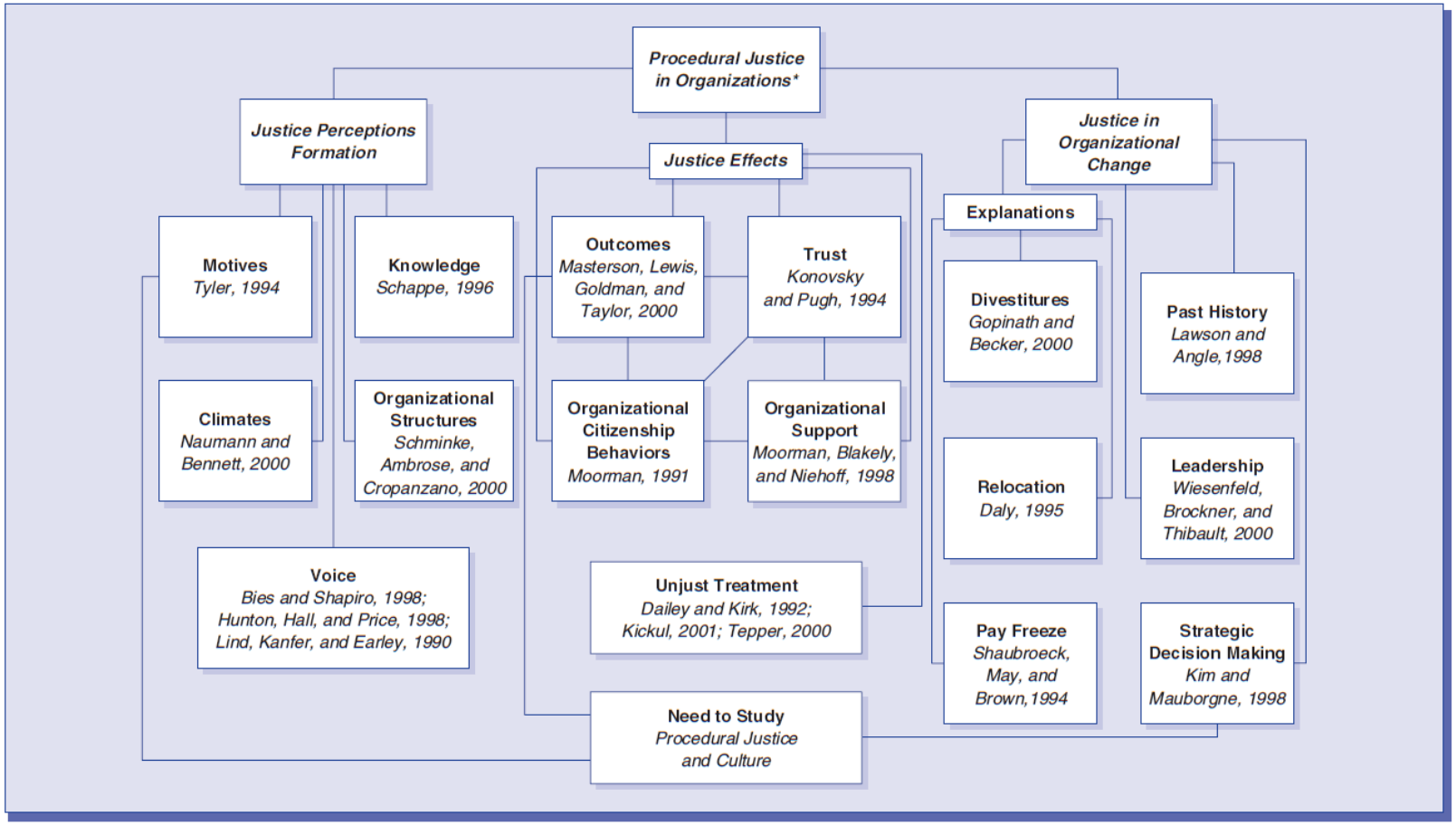
http://robertfeldt.net/publications/petersen_ease08_sysmap_studies_in_se.pdf

Selecting Literature Material

- Start with broad syntheses (such as encyclopedias) if you are new to the topic
- Turn to journal articles in national journals
  - Best source for research reports
- Next consider books
- Then examine conference papers
- Scan for dissertations
- Last consider reports on the web

An Example of Literature Map

Figure 2.1 An Example of a Literature Map



*Employees' concerns about the fairness of and the making of managerial decisions

SOURCE: Janovec (2001). Reprinted by permission.

What to watch out for

Say hello to a flood of papers

But don't despair yet

What to watch out for

It is easy to get overwhelmed

Stick with your questions

What to watch out for

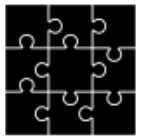
It is easy to get lost in details

Stick with your questions

WRITING MY FIRST LITERATURE REVIEW

https://sites.lafayette.edu/econ408-fa19-jarretlm/2019/12/05/writing-my-first-literature-review/

Optional step: snowball search



The aim is to support the **completeness** of the search

This is an additional search based on some of the **reviewed** papers in 

Backward snowball: papers that paper X cites

Apply the existing selection criteria

Forward snowball: papers that cite paper X

Apply the existing selection criteria

The real work

Probably the most **labor-intensive** step

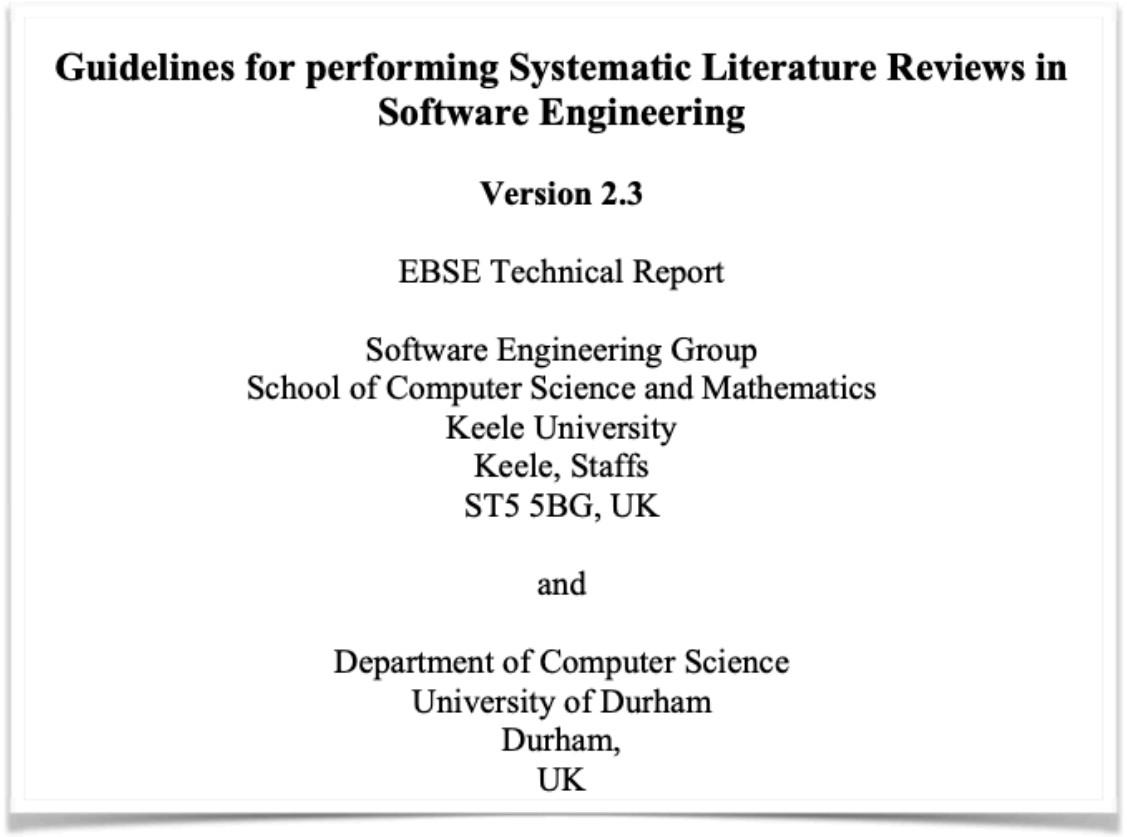
Extract the data and create an **annotated bibliography**

http://guides.library.cornell.edu/annotatedbibliography

Data **processing** can take many forms

qualitative coding, quantitative analysis, etc

Guidelines for SLRs in Software Engineering



Guidelines for performing Systematic Literature Reviews in Software Engineering

Version 2.3

EBSE Technical Report

Software Engineering Group
School of Computer Science and Mathematics
Keele University
Keele, Staffs
ST5 5BG, UK

and

Department of Computer Science
University of Durham
Durham,
UK

Things to remember

The methodology behind SLRs is meant to lessen **bias**

Bias in the primary studies can still exist though

The SLR reporting should support **replication**

Present your review protocol clearly

You may forget **tacit** details

Document everything

SLRs may detect **effects** that individual studies cannot

This applies more when assessing quality

Yes, there is such a thing as a **SLR of SLRs**

It's called a "tertiary review"

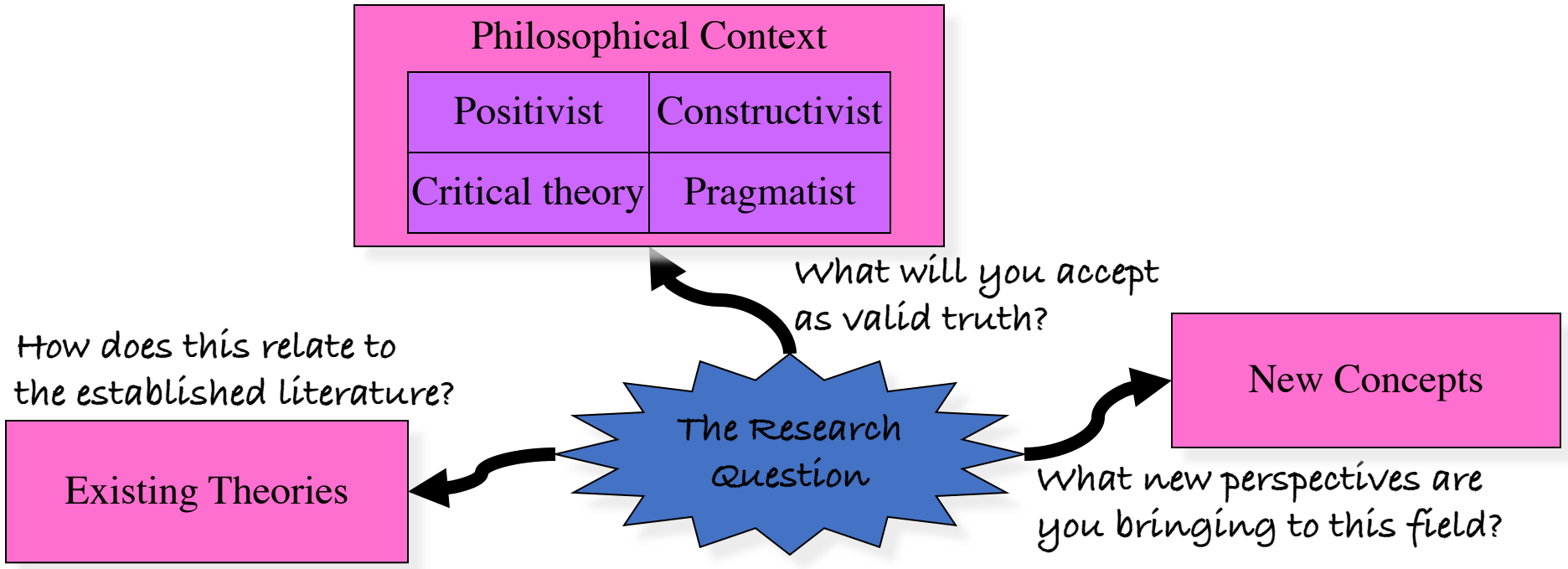
Good examples

- Kitchenham, B., Brereton, O. P., Budgen, D., Turner, M., Bailey, J., & Linkman, S. (2009**). Systematic literature reviews in software engineering–a systematic literature review**. Information and software technology, 51(1), 7-15.
- Beecham, S., Baddoo, N., Hall, T., Robinson, H., & Sharp, H. (2008). **Motivation in Software Engineering: A systematic literature review. Information and software technology**, 50(9), 860-878.
- Hossain, E., Babar, M. A., & Paik, H. Y. (2009, July). **Using scrum in global software development: a systematic literature review**. In Global Software Engineering, 2009. ICGSE 2009. Fourth IEEE International Conference on (pp. 175-184). Ieee.
- Dybå, T., & Dingsøyr, T. (2008). **Empirical studies of agile software development: A systematic review.** Information and software technology, 50(9), 833-859. Chicago
- Šmite, D., Wohlin, C., Gorschek, T., & Feldt, R. (2010**). Empirical evidence in global software engineering: a systematic review.** Empirical software engineering, 15(1), 91-118.

Planning Checklist

- [x] Pick a topic
- [x] Identify the research question(s)
- [x] Check the literature
- [ ] Identify your philosophical stance
- [ ] Identify appropriate theories
- [ ] Choose the method(s)
- [ ] Design the study
  - Unit of analysis?
  - Target population?
  - Sampling technique?
  - Data collection techniques?
  - Metrics for key variables?
  - Handle confounding factors
- [ ] Critically appraise the design for threats to validity
- [ ] Get IRB approval
  - Informed consent?
  - Benefits outweigh risks?
- [ ] Recruit subjects / field sites
- [ ] Conduct the study
- [ ] Analyze the data
- [ ] Write up the results and publish them
- [ ] Iterate

Putting the Question in Context



Easterbrook et al Chapter
- Selecting Methods -

Choice of method depends on the **research question** being asked (exploratory, confirmatory, relationship) as well as the **researcher's philosophical perspective**

Research Designs

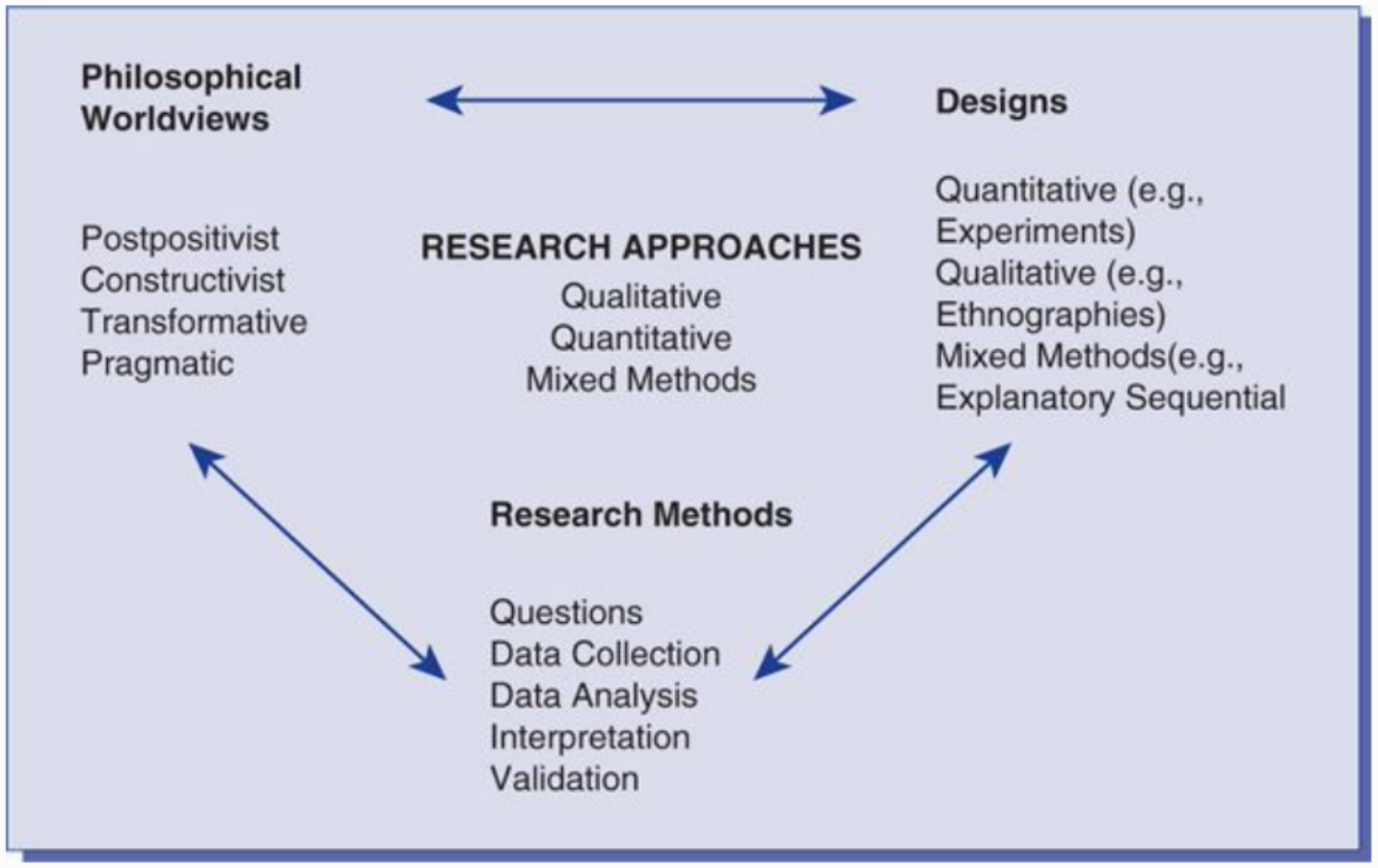
- Step 1 to **design** empirical research:

  adopt a general (and guiding) approach

- Three main approaches:
  - **Qualitative**
  - **Quantitative**
  - **Mixed Methods**
- Qual. vs Quant.:
  - Not dichotomies
  - Rather, different ends on a continuum

Chapter 1 - Creswell, J. W., & Creswell, J. D. (2017). Research design: Qualitative, quantitative, and mixed methods approaches. Sage publications.

A Framework for Research—
The Interconnection of Worldviews, Design, and Research Methods



"A paradigm is a shared world view that represents the beliefs and values in a discipline and that guides how problems are solved."

– Schwandt, 2001

Scientific method

Evidence-based reality

Theory verification and falsification

Quantitative over qualitative

Paradigms – **Postpositivism**

Positivist View

- Traditional form of research, also referred to as the ***scientific method***; or empirical science; or postpositivism

Positivist View

- Traditional form of research, also referred to as the ***scientific method***; or empirical science; or postpositivism
  - More often quant. research than qual.
- Never absolute truth of knowledge
  - Can't be "positive" about our claims of knowledge when studying human behavior and actions
  - Do not prove a hypothesis; instead, fail to reject the hypothesis

Positivist View (2)

- Deterministic philosophy: causes probably determine effects or outcomes
  - Research seeks to identify the causes that influence outcomes (e.g., experiments)
- Reductionistic in nature:
  - small, discrete set of ideas to test, e.g., variables that comprise hypotheses and research questions

Positivist View (3)

- Knowledge develops through careful observation and measurement of the objective reality “out there” in the world.
- Laws or theories govern the world
    - need to be tested, verified, refined so that we can understand the world
- Accepted approach:

begin with a theory → collect data that either supports or refutes the theory → make necessary revisions → perform additional tests

Reality is subjective and experiential

Theory generation

Biases are expected and made explicit

Qualitative over quantitative

Paradigms – **Constructivism**

(Social) Constructivist View

- Typically qual research
- Individuals develop subjective meanings of their experiences, directed toward objects or things
  - Meanings are varied and multiple, look for complexity of views
  - Meanings are not imprinted but rather formed through interaction with others (hence social constructivism)
- The goal is to rely as much as possible on the participants' views of the situation being studied.
- Inductive

(Social) Constructivist View 2

- Researchers want to make sense of (or interpret)
- the meanings others have about the world.
  - generate or inductively develop a theory rather than start with one

Change oriented

Collaborative

Shaped by political and social lenses

Qualitative and quantitative

Paradigms – **Advocacy / Participatory/ Transformative**

Transformative View

- Research inquiry needs to be intertwined with politics
  - the research contains an action agenda for reform that may change the lives of the participants
  - focuses on the needs of groups and individuals that may be marginalized or disenfranchised
  - addresses important social issues of the day, such as empowerment, inequality, oppression

Problem centered

Real-world practice oriented

Chooses methods as needed

Paradigms – **Pragmatism**

Pragmatic View

- Arises out of actions, situations, and consequences rather than antecedent conditions (as in postpositivism)
- Instead of focusing on methods, researchers emphasize the research problem and use all approaches available to understand it
  - Typical for mixed methods studies
- “Researchers would simply like to change the subject”

Philosophical Worldviews

Chapter 1 - Creswell, J. W., & Creswell, J. D. (2017). Research design: Qualitative, quantitative, and mixed methods approaches. Sage publications.

Postpositivism	Constructivism
• Determination • Reductionism • Empirical observation and measurement • Theory verification	• Understanding • Multiple participant meanings • Social and historical construction • Theory generation
Transformative	**Pragmatism**
• Political • Power and justice oriented • Collaborative • Change-oriented	• Consequences of actions • Problem-centered • Pluralistic • Real-world practice oriented

Which do you subscribe to?

Positivist (or "Post-positivist") • Knowledge is objective • "Causes determine effects/outcomes" • Reductionist: study complex things by breaking down to simpler ones • Prefer quantitative approaches • Verifying (or Falsifying) theories	**Constructivist/Interpretivist** • Knowledge is socially constructed • Truth is relative to context • Theoretical terms are open to • interpretation • Prefer qualitative approaches • Generating "local" theories
Advocate / Critical Theorist • Research is a political act • Knowledge is created to empower groups/individuals • Choose what to research based on who it will help • Prefer participatory approaches • Seeking change in society	**Pragmatist** • Research is problem-centered • "All forms of inquiry are biased" • Truth is what works at the time • Prefer multiple methods / multiple perspectives • Seeking practical solutions to problems

Planning Checklist

- [x] Pick a topic
- [x] Identify the research question(s)
- [x] Check the literature
- [x] Identify your philosophical stance
- [ ] Identify appropriate theories
- [ ] Choose the method(s)
- [ ] Design the study
  - Unit of analysis?
  - Target population?
  - Sampling technique?
  - Data collection techniques?
  - Metrics for key variables?
  - Handle confounding factors
- [ ] Critically appraise the design for threats to validity
- [ ] Get IRB approval
  - Informed consent?
  - Benefits outweigh risks?
- [ ] Recruit subjects / field sites
- [ ] Conduct the study
- [ ] Analyze the data
- [ ] Write up the results and publish them
- [ ] Iterate

Identify Appropriate Theories

Where do theories come from?

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

A simpler definition

A (good) Theory is the best explanation of all the available evidence

The Use of Theory

- Quantitative tests theories as explanations
- Qualitative studies may generate the theory
- Mixed methods studies may have no theories at all or a theoretical framework in which both quantitative and qualitative data are collected

Variables in Quantitative Research

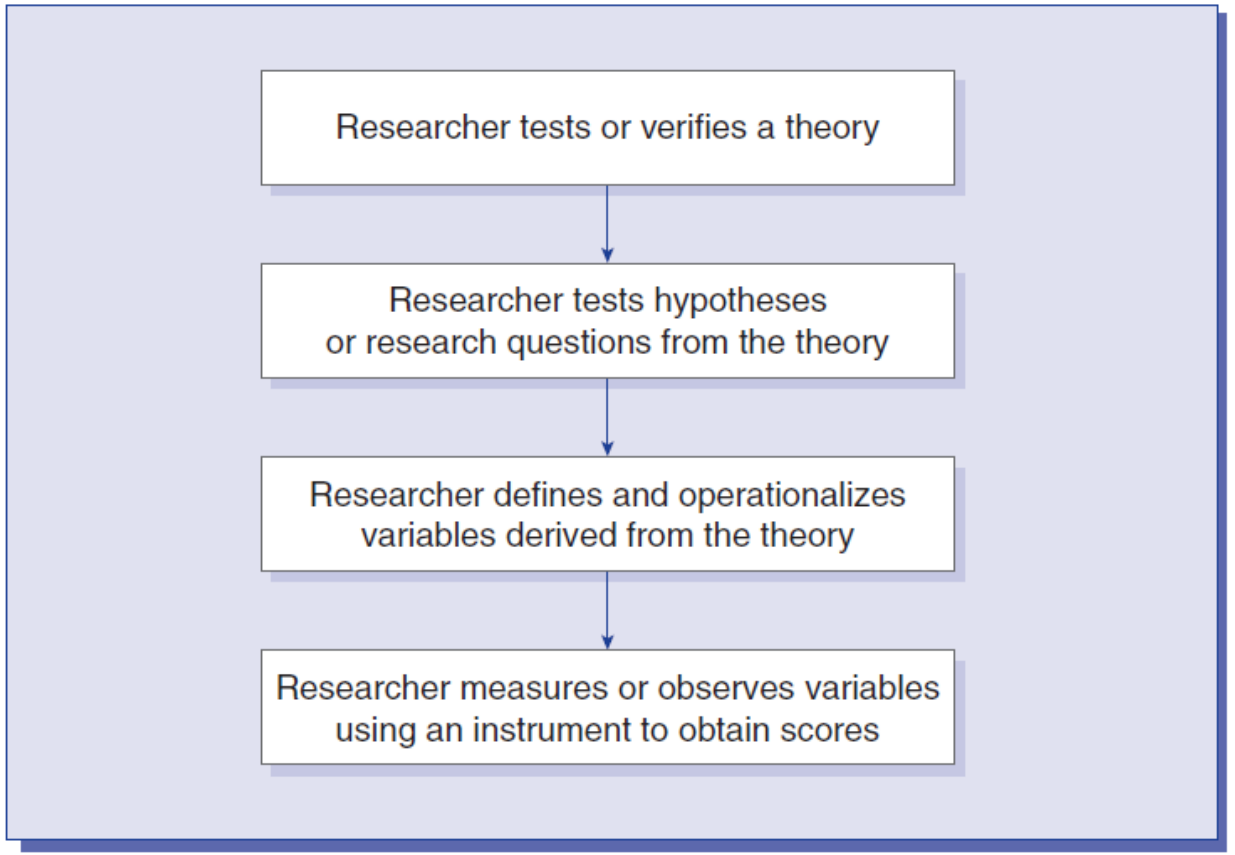
- A variable is a characteristic of an individual or group that is measurable.
- Examples of variables in a study may include age, gender and socioeconomic status.
- Variables may have temporal order, or be measurable or observable.
- Another term for a variable is a construct.

Theories in Quantitative Research

- Here a theory is a scientific predication or explanation of what the Researcher expects to find out about a set of interrelated **variables**.
- When a Researcher uses a theory to **predict** the anticipated outcome of a study he/she has created the **hypothesis.**
- As the Researcher conducts the study and puts forward an **explanation** for his/her predictions and study findings, he/she is presenting his/her **theory.**

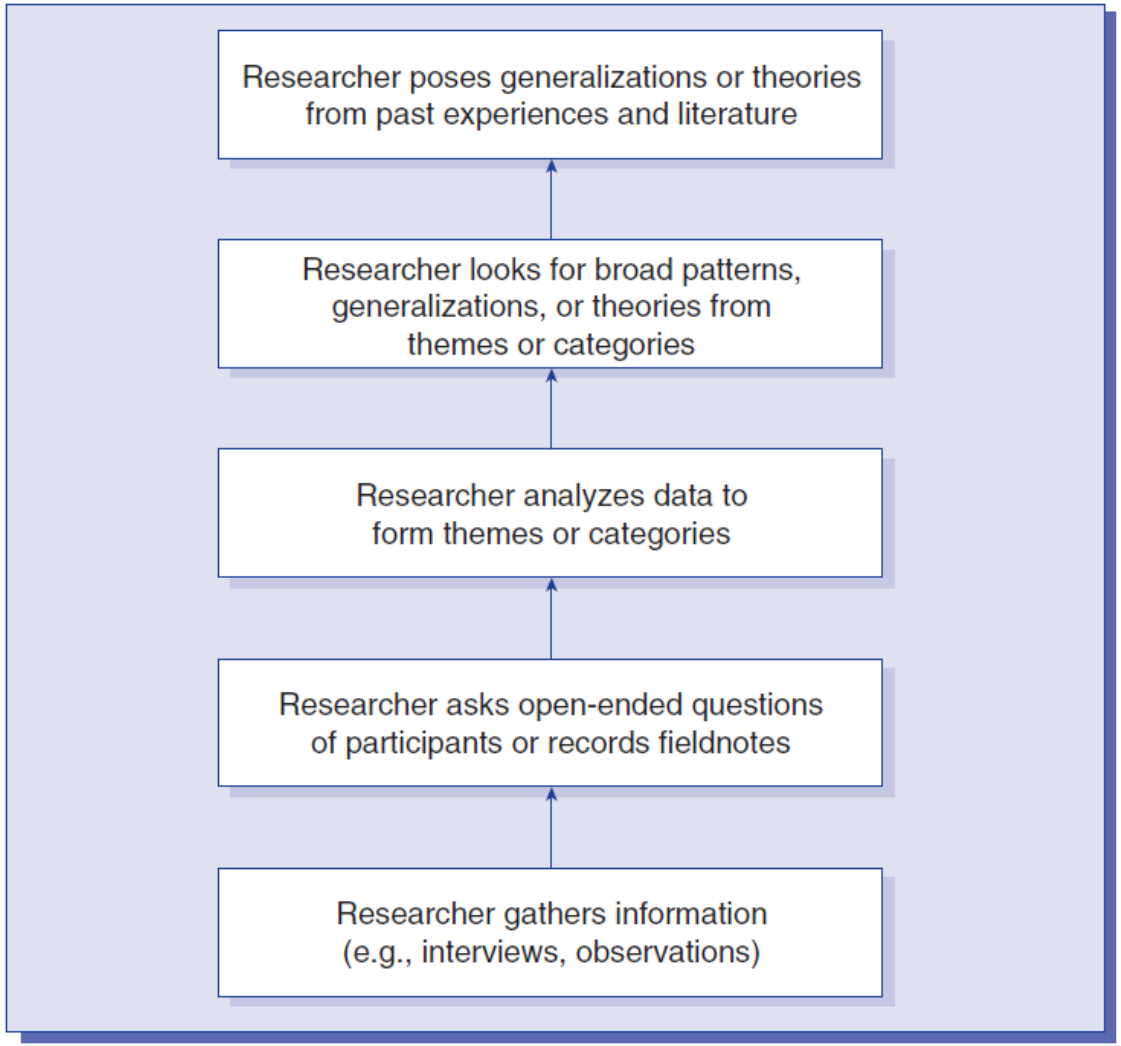
The Deductive Approach Used in Quantitative Research

Figure 3.4 The Deductive Approach Typically Used in Quantitative Research



Qualitative Theory Use

Figure 3.5 The Inductive Logic of Research in a Qualitative Study




EXAMPLE

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

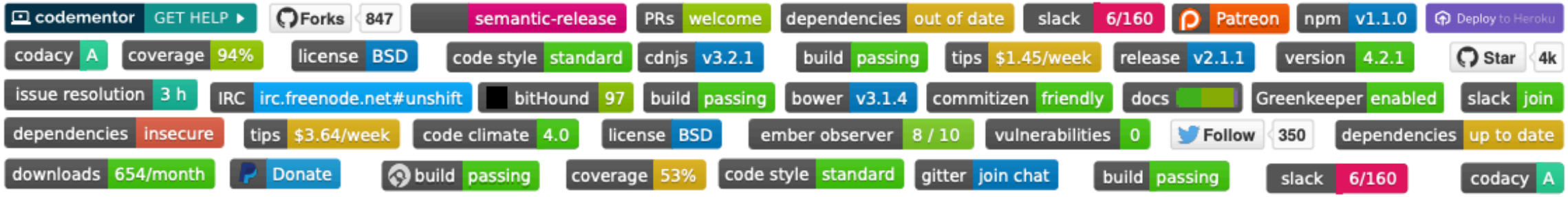
- May hurt creativity and streamlining

- Information overload

*Bernstein, E. S. (2012). The transparency paradox a role for privacy in organizational learning and operational control. Administrative Science Quarterly, 57(2), 181-216.

GitHub: Transparency

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



Adding Sparkle to Social Coding:

An Empirical Study of Repository Badges in the npm Ecosystem

Asher Trockman, Shurui Zhou, Christian Kästner, Bogdan Vasilescu

ICSE '18, May 27-June 3, 2018, Gothenburg, Sweden

STRUDEL
SOCIO-TECHNICAL RESEARCH USING DATA EXCAVATION LAB

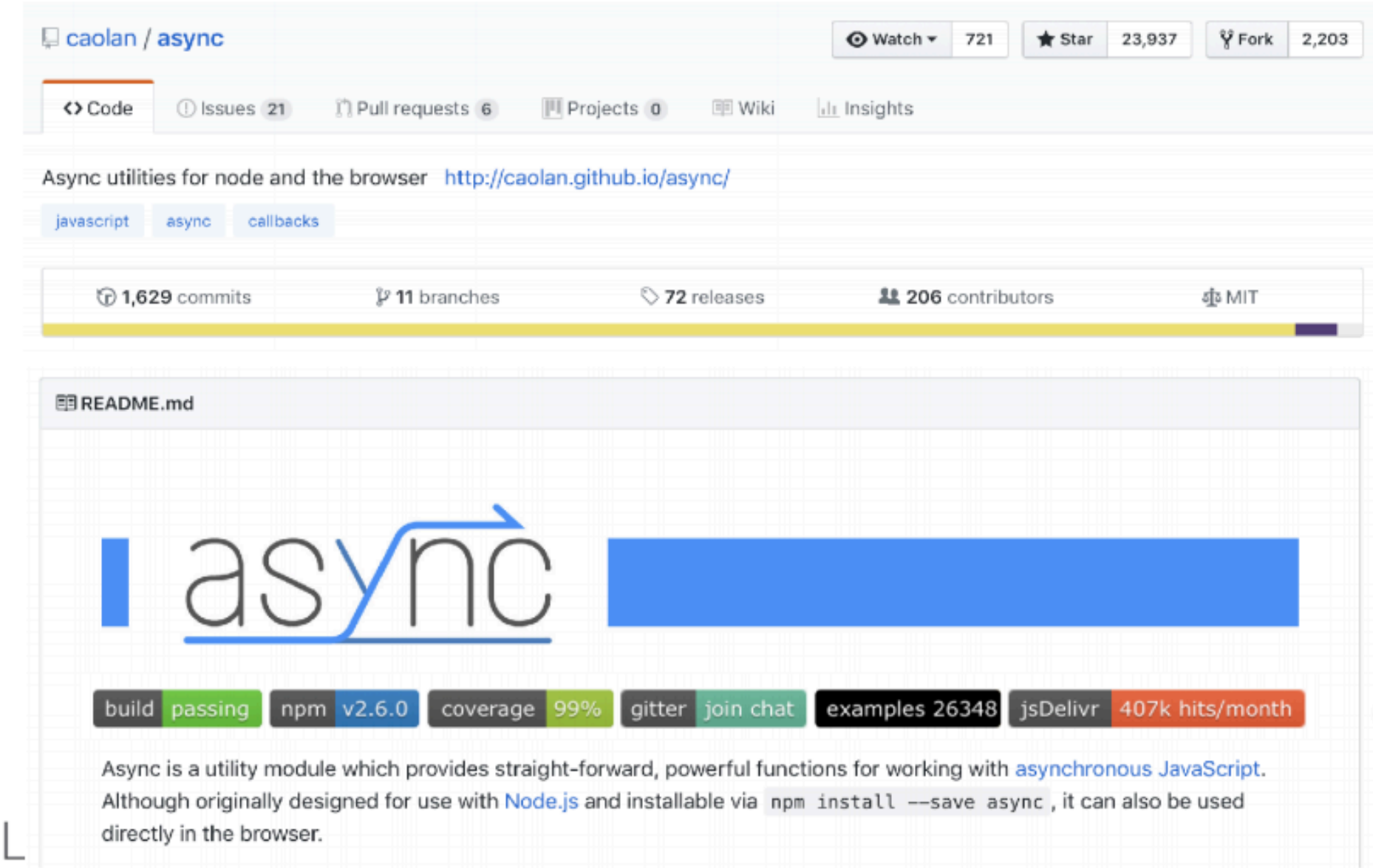
UNIVERSITY OF EVANSVILLE
Civic Mission...Sacred Trust

Carnegie Mellon University

isr institute for SOFTWARE RESEARCH

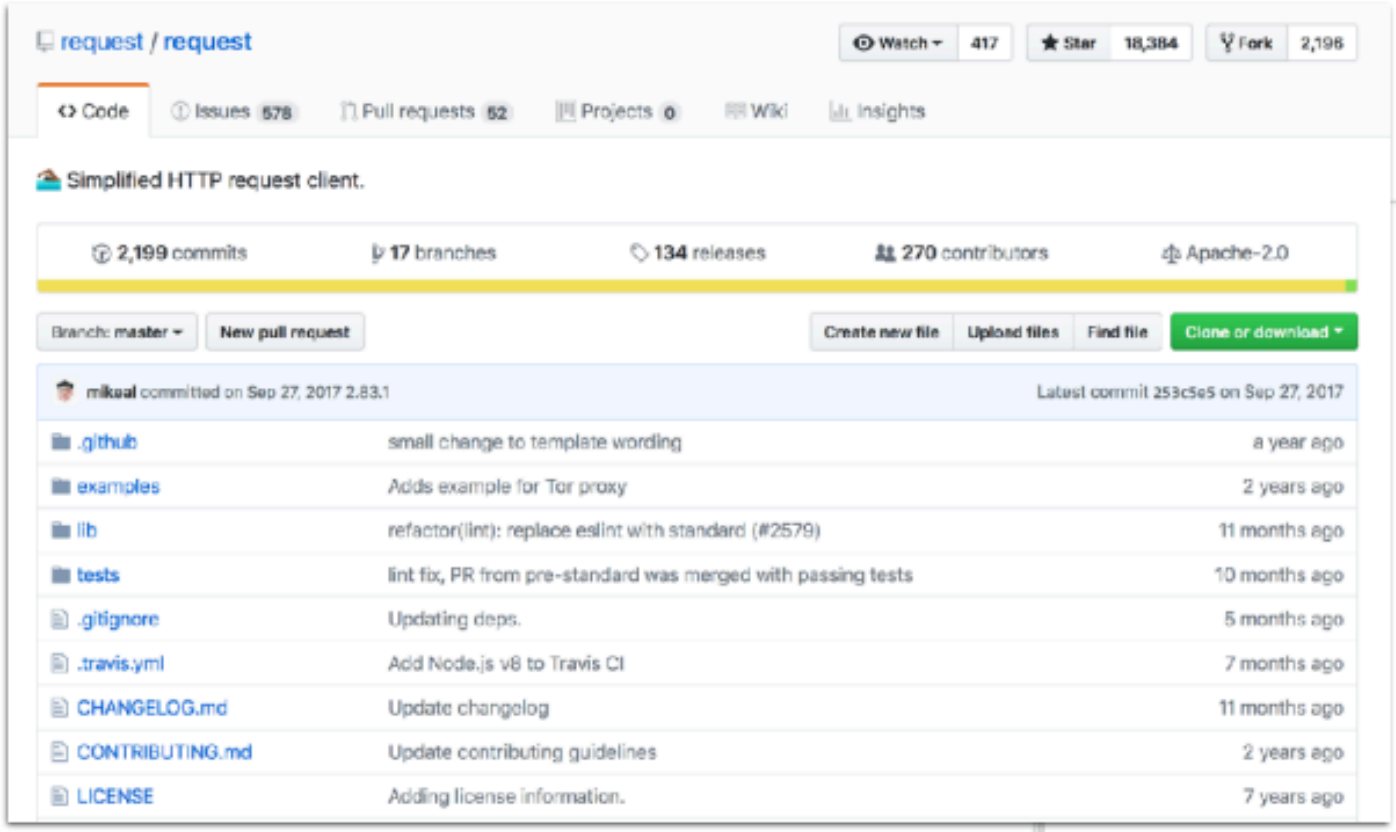
The Edward S. Rogers Sr. Department of Electrical & Computer Engineering
UNIVERSITY OF TORONTO

GitHub Repository Badges



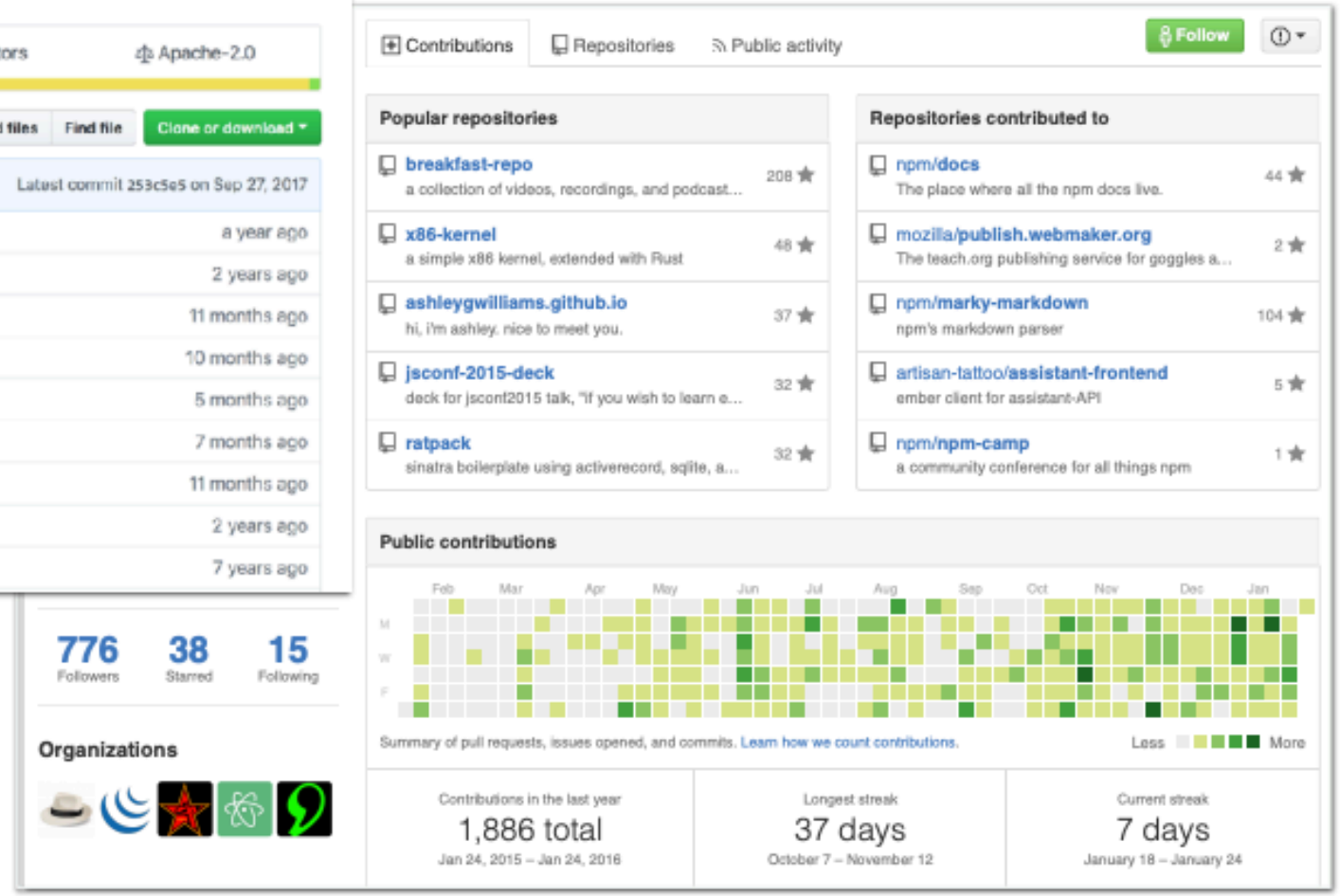
Enlarged to show detail.

Key features: Transparency & signaling



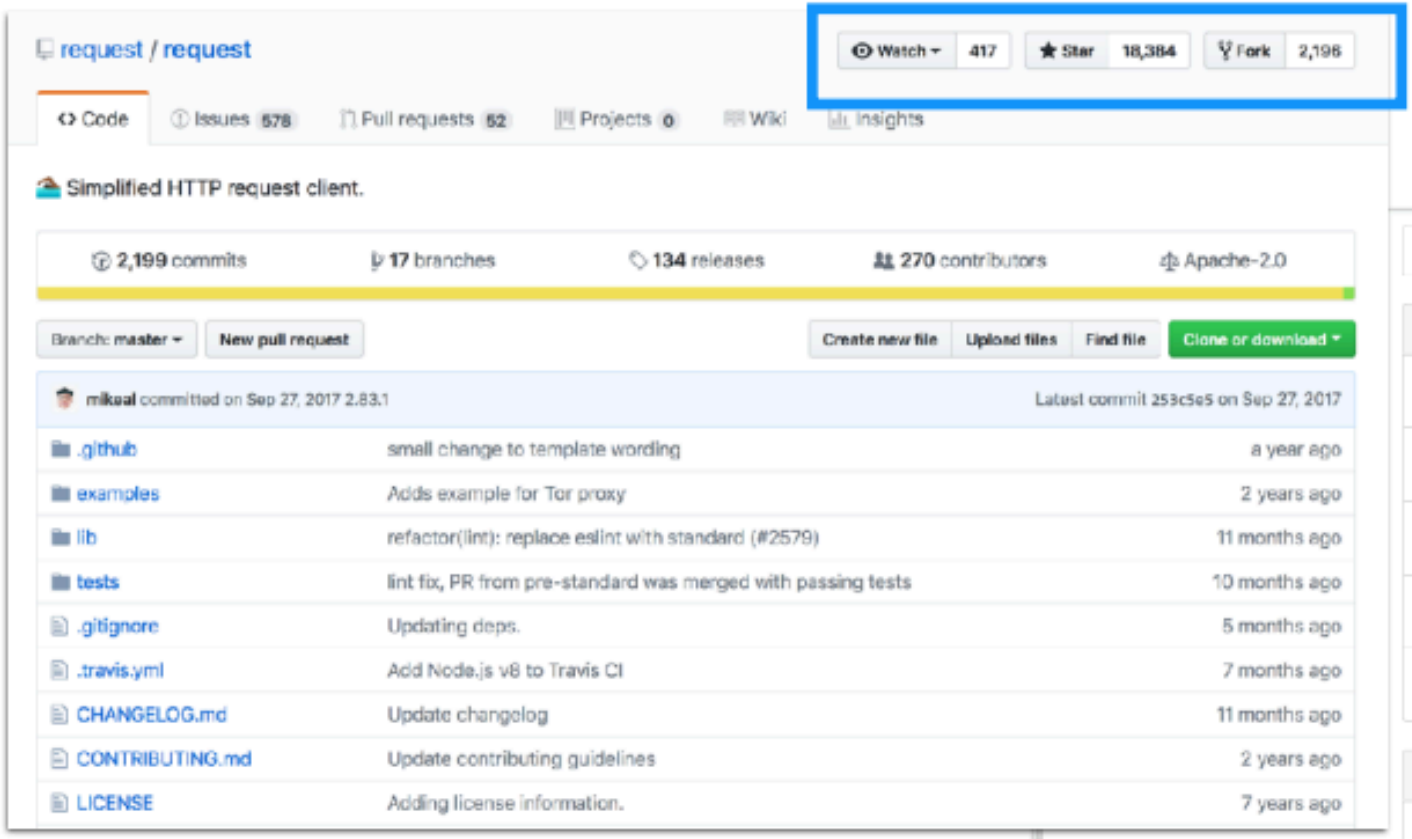
Projects

Users



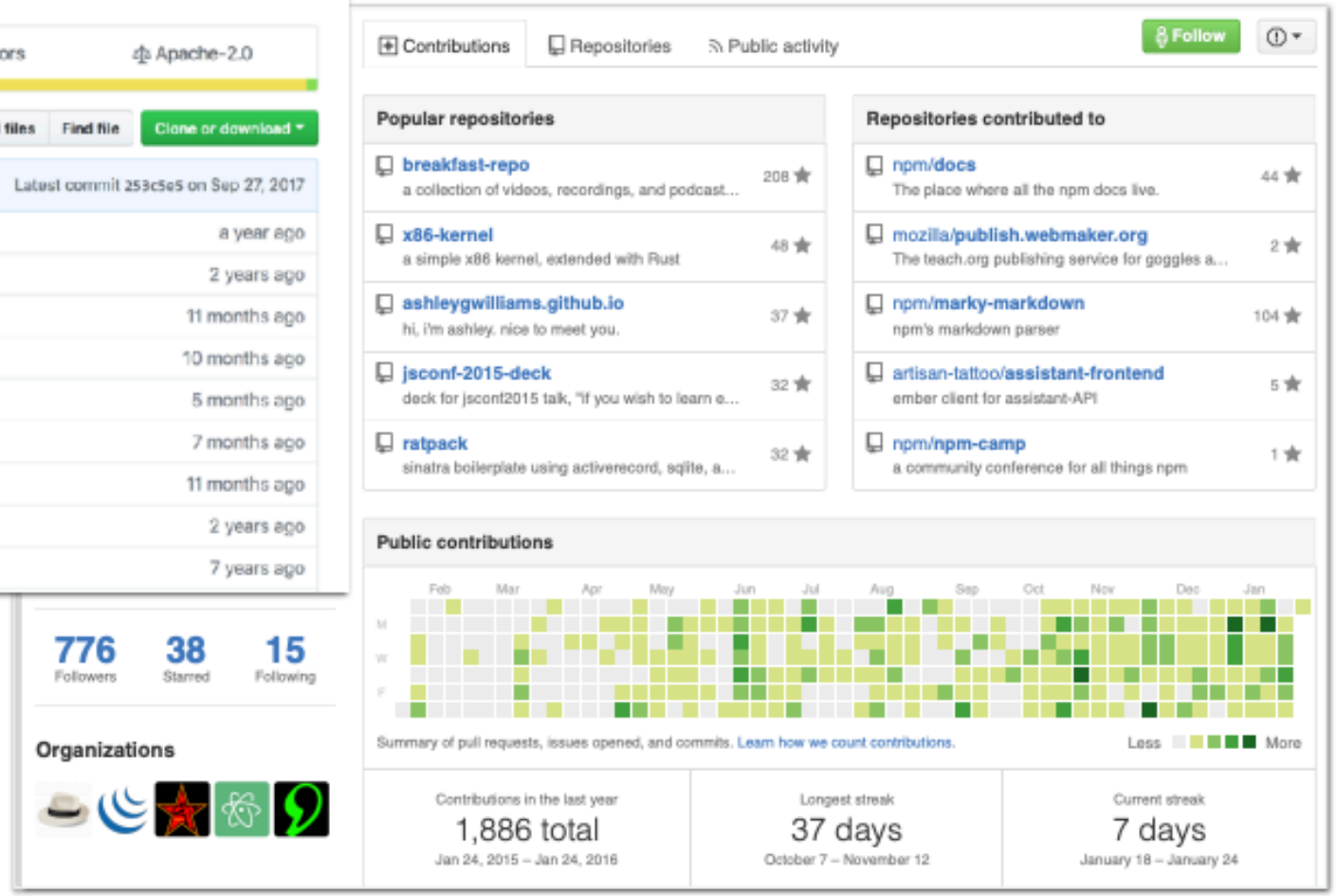
STRUDEL
Carnegie Mellon University

Key features: Transparency & signaling



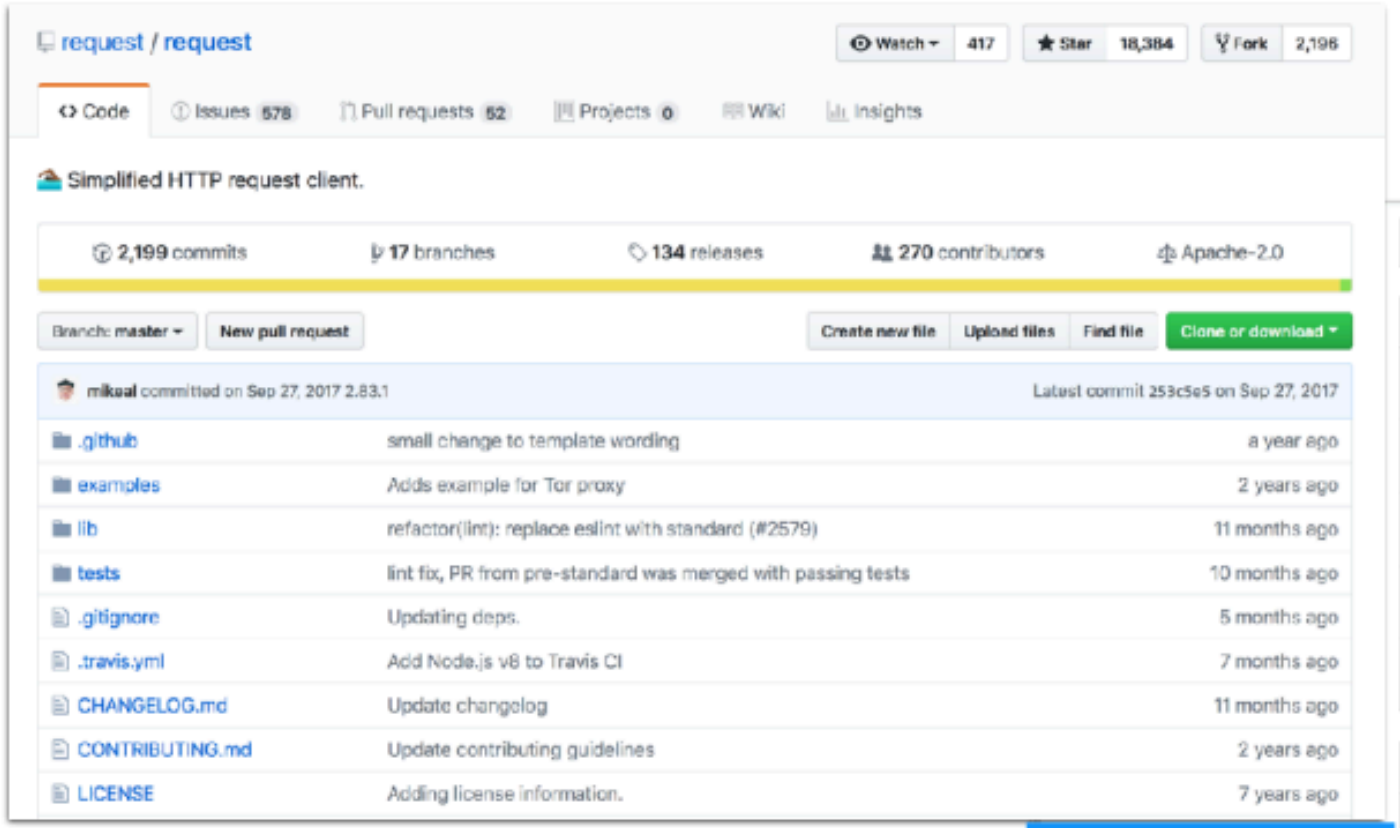
Projects

Users



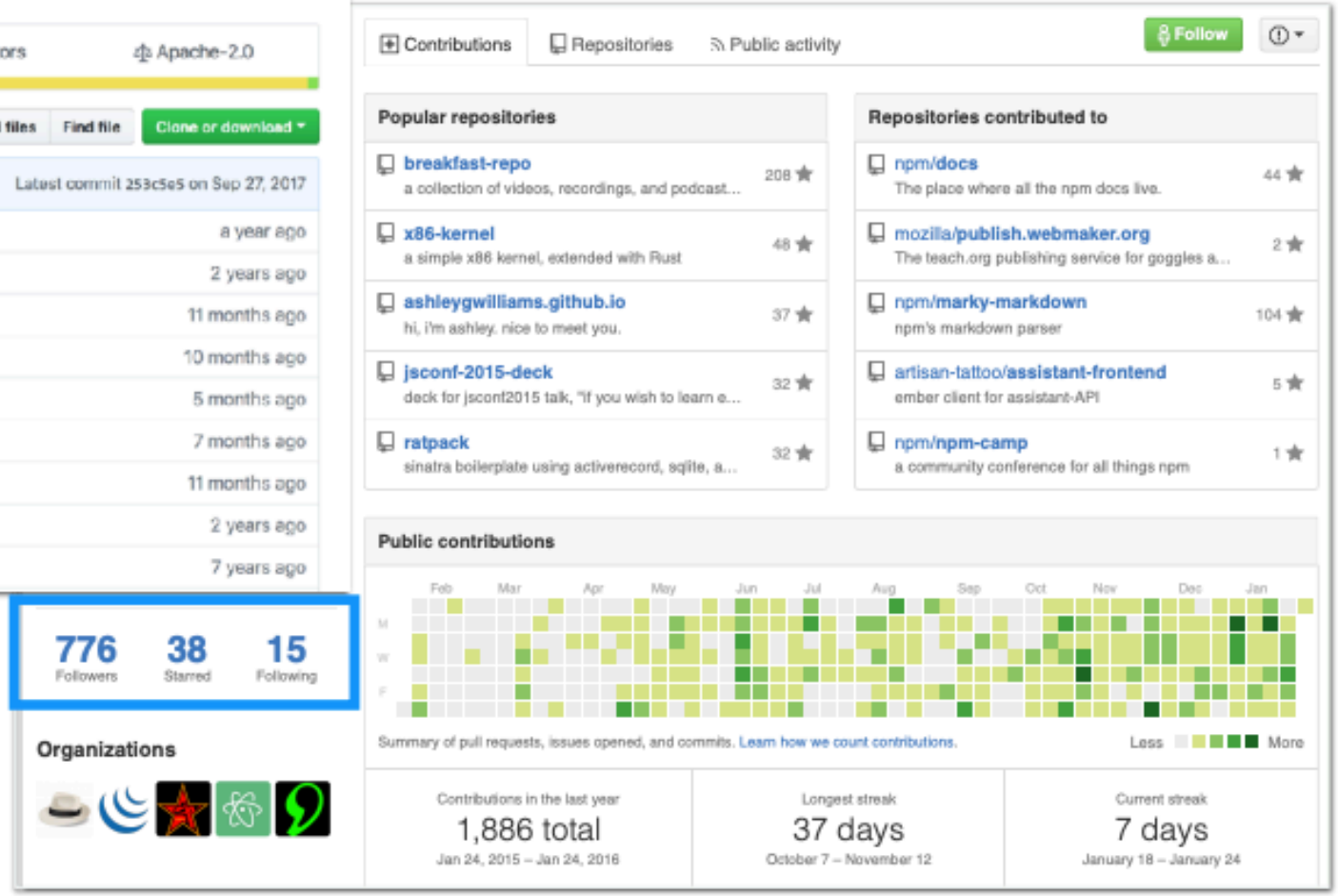
STRUDEL
Carnegie Mellon University

Key features: Transparency & signaling



Projects

Users



STRUDEL
Carnegie Mellon University

Badges are Reliable Signals

Mostly

- 

of the presence of **tests**
- 

of **up-to-date** and secure **dependencies**
- 

of the presence of **tests in pull requests**
- 

of **popularity**

Mixed methods study

+

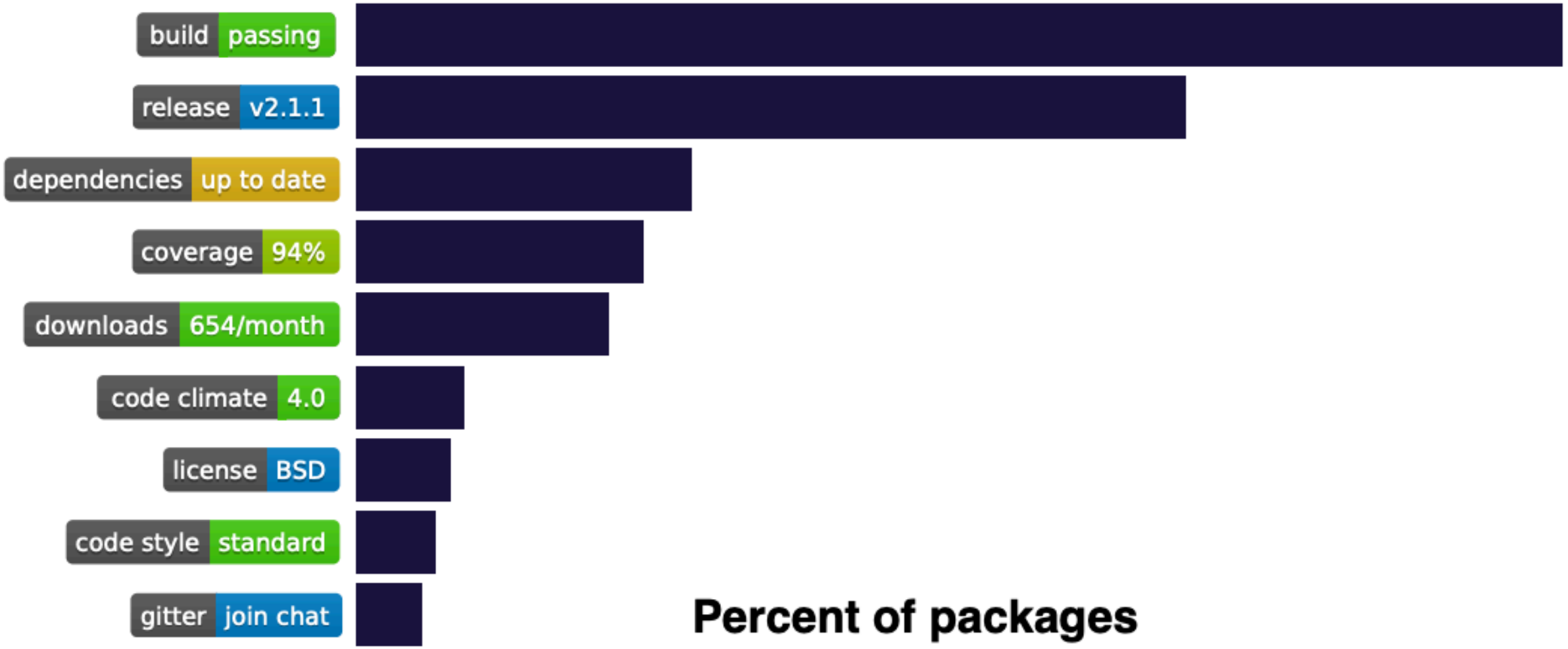
Survey

- 32 maintainers, 57 contributors
- Maintainers:
  - What do you intend to signal?
  - What effects do you expect?
- Contributors:
  - What do badges tell you?

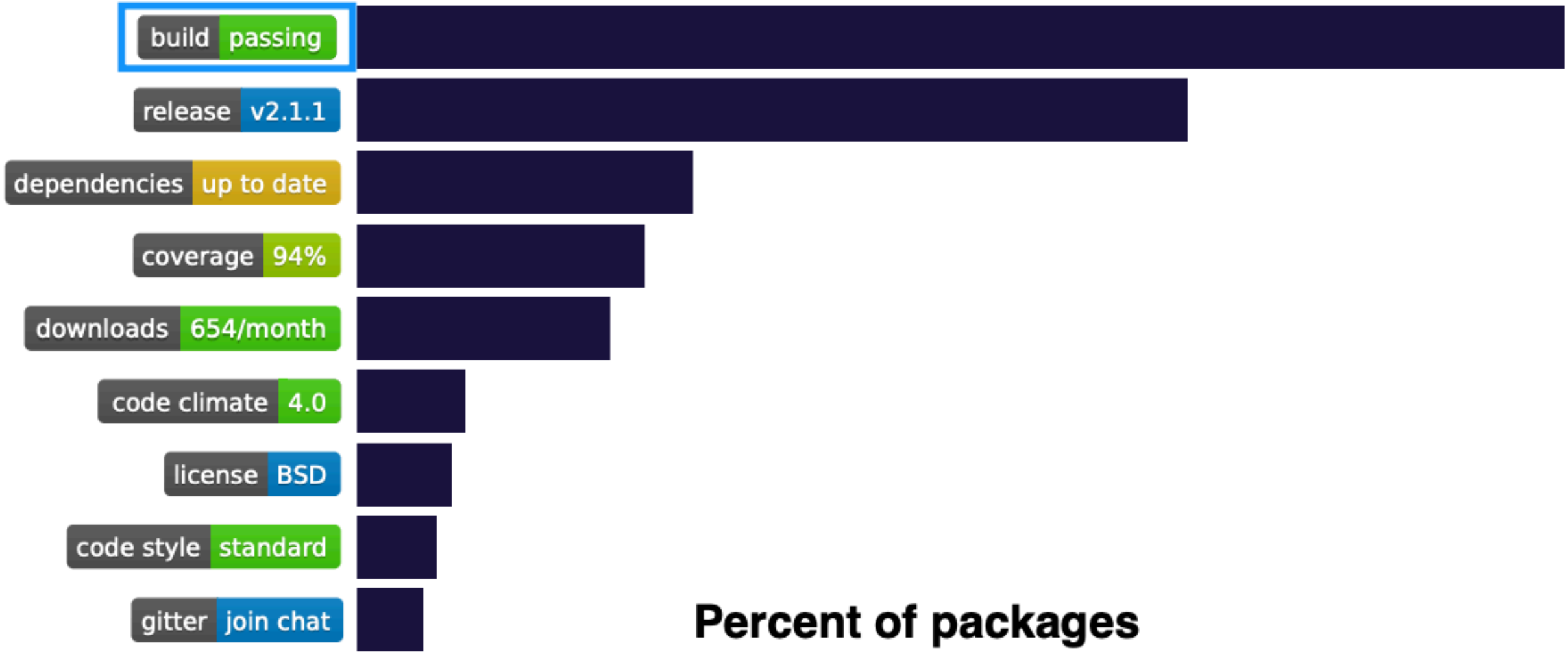
Repository Mining

- 294,941 *npm* packages
- Mined badge adoptions/removals from README files
- Measured proxies for code quality, test suite quality, popularity, dependency freshness, …

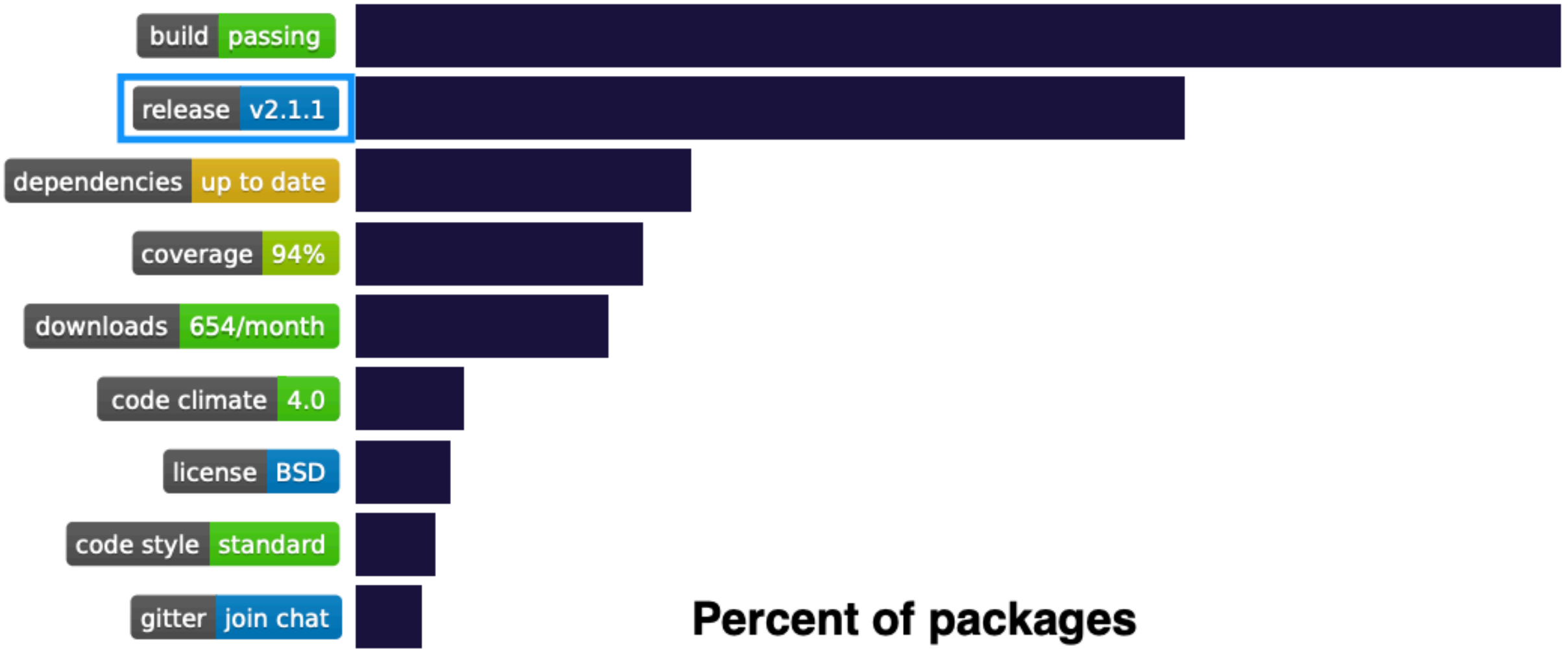
Popular Badges in npm



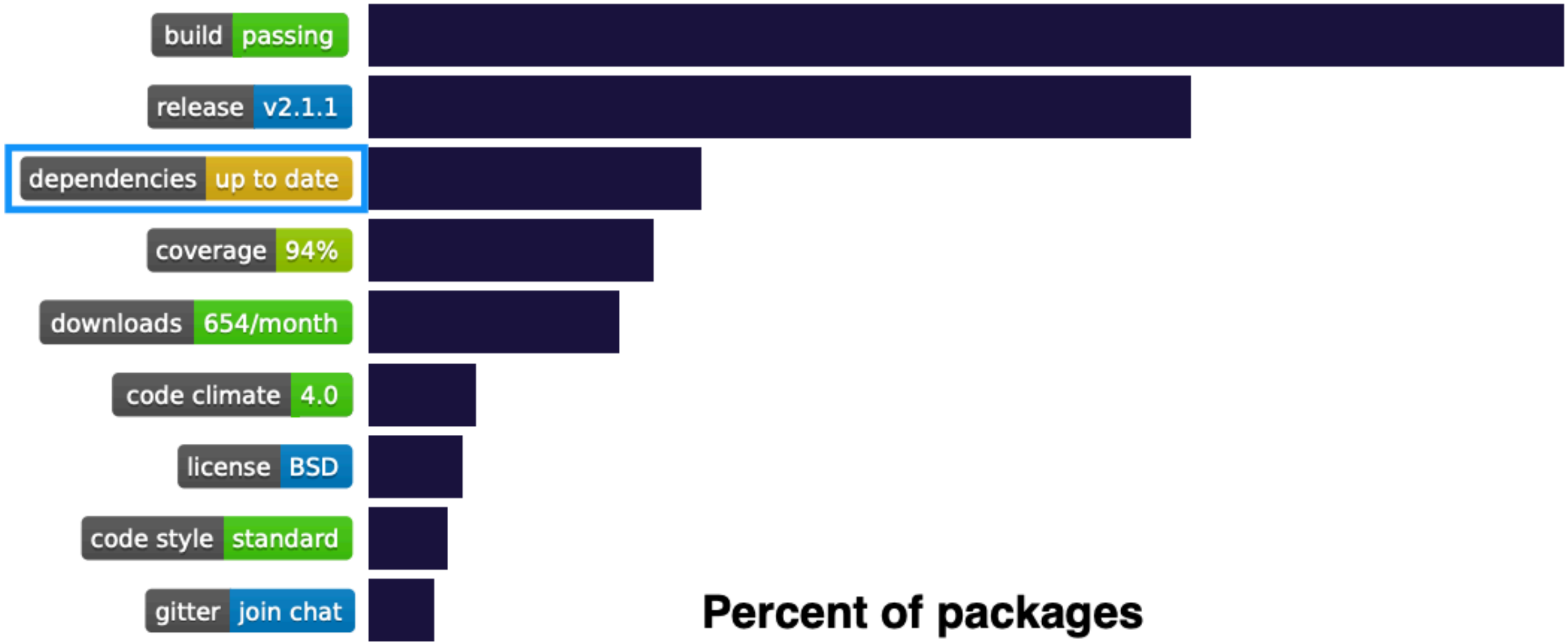
Popular Badges in npm



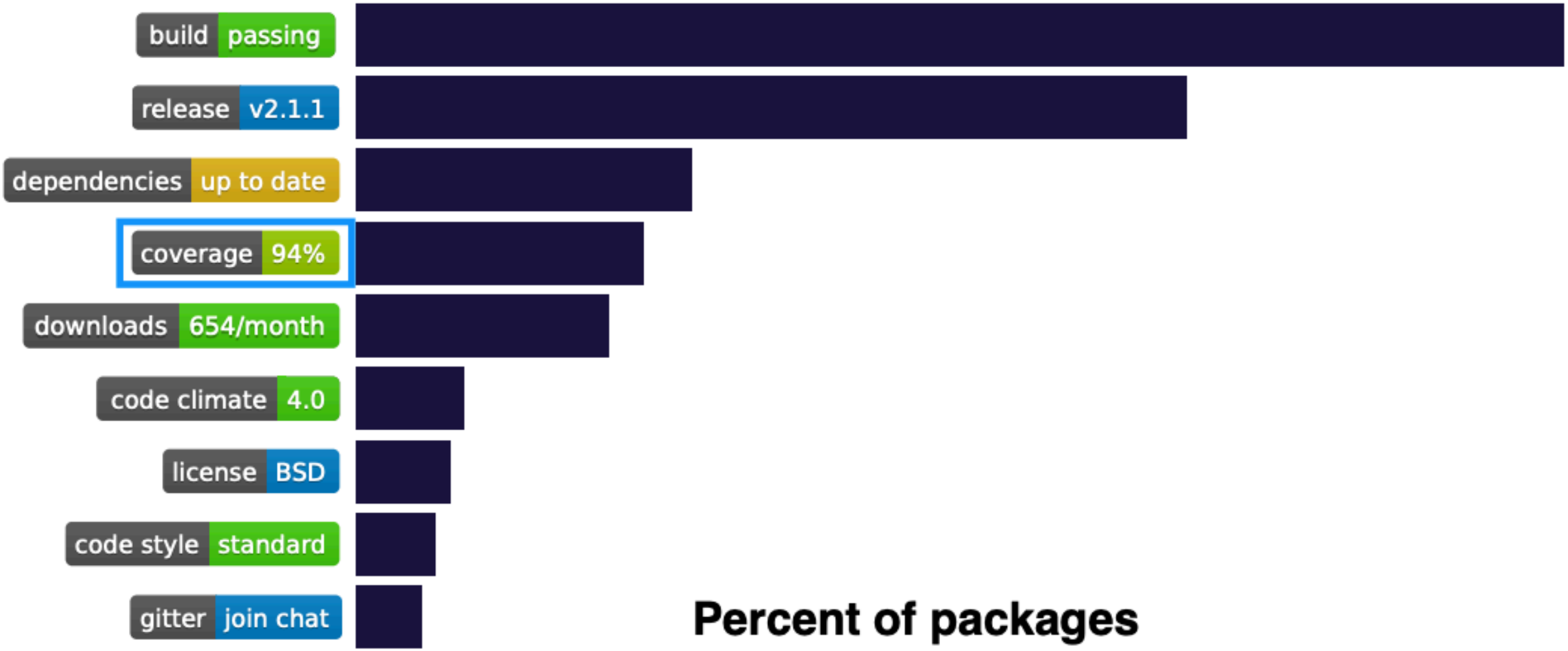
Popular Badges in npm



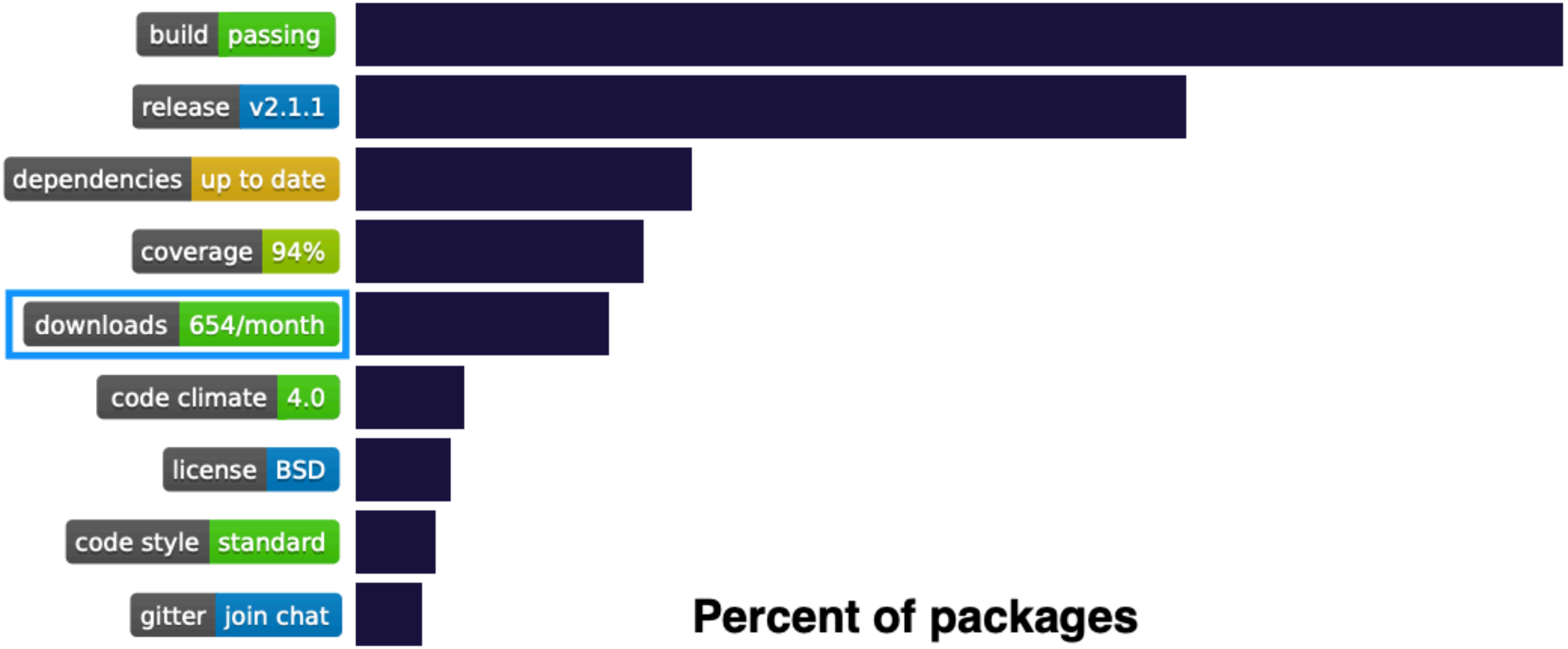
Popular Badges in npm



Popular Badges in npm



Popular Badges in npm



What do developers expect from badges?

- *32* Maintainers
  - What do you intend to signal?
  - What effects do you expect?
- *57* Contributors
  - What do badges tell you?

"welcoming contributions"

"expectations of contribution quality"

"dedicated to offering support"

"reduced chances of conflicting versions of dependencies"

"indicator of product quality"

STRUDEL
Carnegie Mellon University
UNIVERSITY OF TORONTO

Analysis

Correlation

If all you saw was the badge, how much would that tell you?

Regression Analysis

How much more does the badge tell you, relative to existing signals?

Time Series Analysis

How do things change after adding the badge?

Analysis

Correlation

If all you saw was the badge, how much would that tell you?

Regression Analysis

How much more does the badge tell you, relative to existing signals?

Time Series Analysis

How do things change after adding the badge?

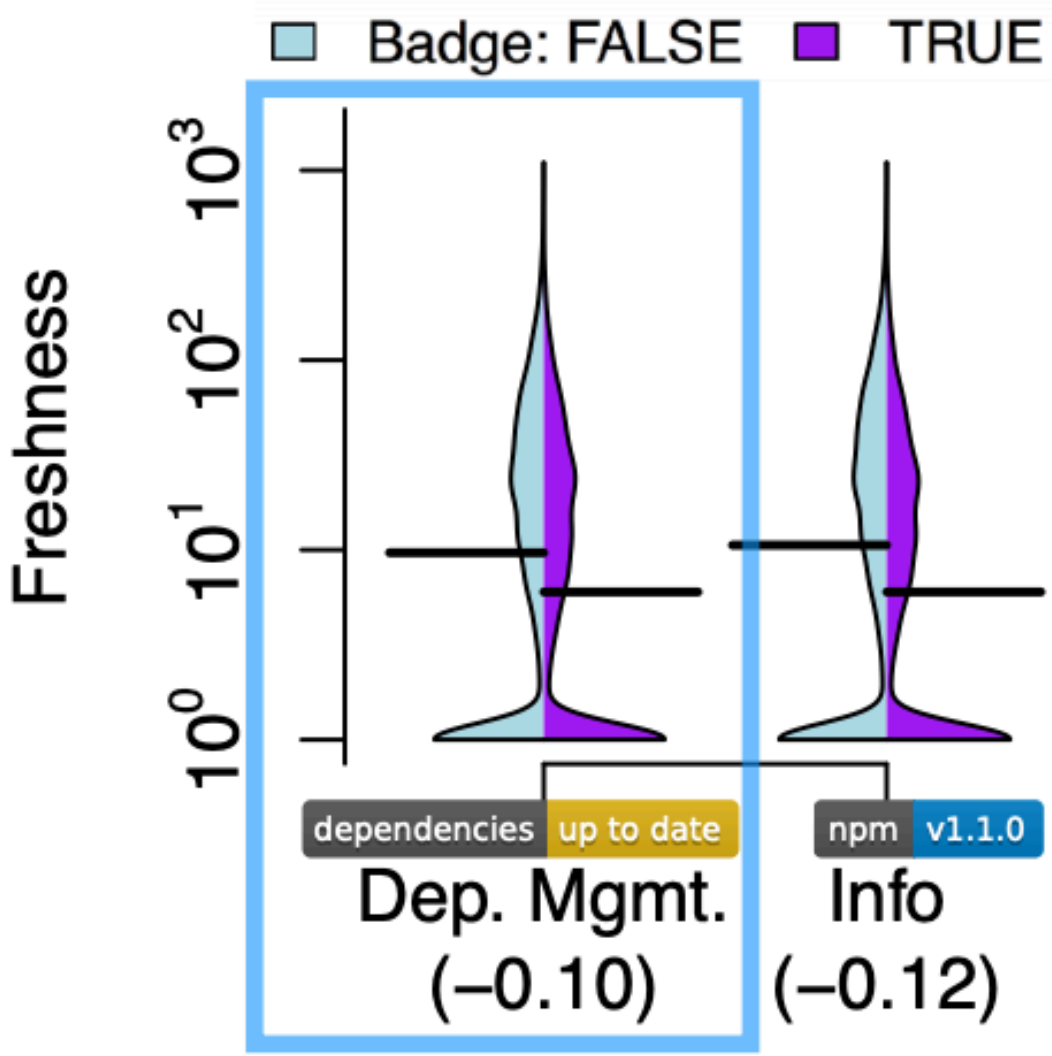
STRUDEL
Carnegie Mellon University

Step 1: Correlation

Signals of fresh dependencies

- **Based on survey:** The adoption of dependency management badges correlates with fresher dependencies
- **Freshness metric:** *lower* is better
  - (More up-to-date deps.)

Result: Dep. badges correlate with fresher dependencies



STRUDEL
Carnegie Mellon University

Step 2: Regression Analysis

Signals of fresh dependencies

- **Based on survey:** The adoption of dependency management badges correlates with fresher dependencies
- **Freshness metric:** *lower* is better
  - (More up-to-date deps.)

Result: Dep. badges are *the best* signals of fresh dependencies

	Basic Model response: *freshness = 0* 17.3% deviance explained		**Full Model** response: *freshness = 0* 17.4% deviance explained	
	Coeffs (Err.)	LR Chisq	Coeffs (Err.)	LR Chisq
(Interc.)	$3.54\ (0.03)^{***}$		$3.50\ (0.03)^{***}$	
Dep.	$-1.78\ (0.01)^{***}$	32077.8^{***}	$-1.79\ (0.01)^{***}$	32292.8^{***}
RDep.	$0.22\ (0.01)^{***}$	610.3^{***}	$0.21\ (0.01)^{***}$	560.6^{***}
Stars	$-0.08\ (0.00)^{***}$	301.4^{***}	$-0.09\ (0.00)^{***}$	311.2^{***}
Contr.	$-0.24\ (0.01)^{***}$	500.5^{***}	$-0.25\ (0.01)^{***}$	548.7^{***}
lastU	$-0.65\ (0.01)^{***}$	12080.9^{***}	$-0.64\ (0.01)^{***}$	11537.9^{***}
dependencies up to date			$0.24\ (0.03)^{***}$	116.1^{***}
npm v1.1.0			$0.11\ (0.02)^{***}$	48.3^{***}
dependencies up to date : npm v1.1.0			$-0.05\ (0.04)$	1.9
hasOther			$0.01\ (0.01)$	

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

STRUDEL
Carnegie Mellon University

Analysis

Correlation

If all you saw was the badge, how much would that tell you?

Regression Analysis

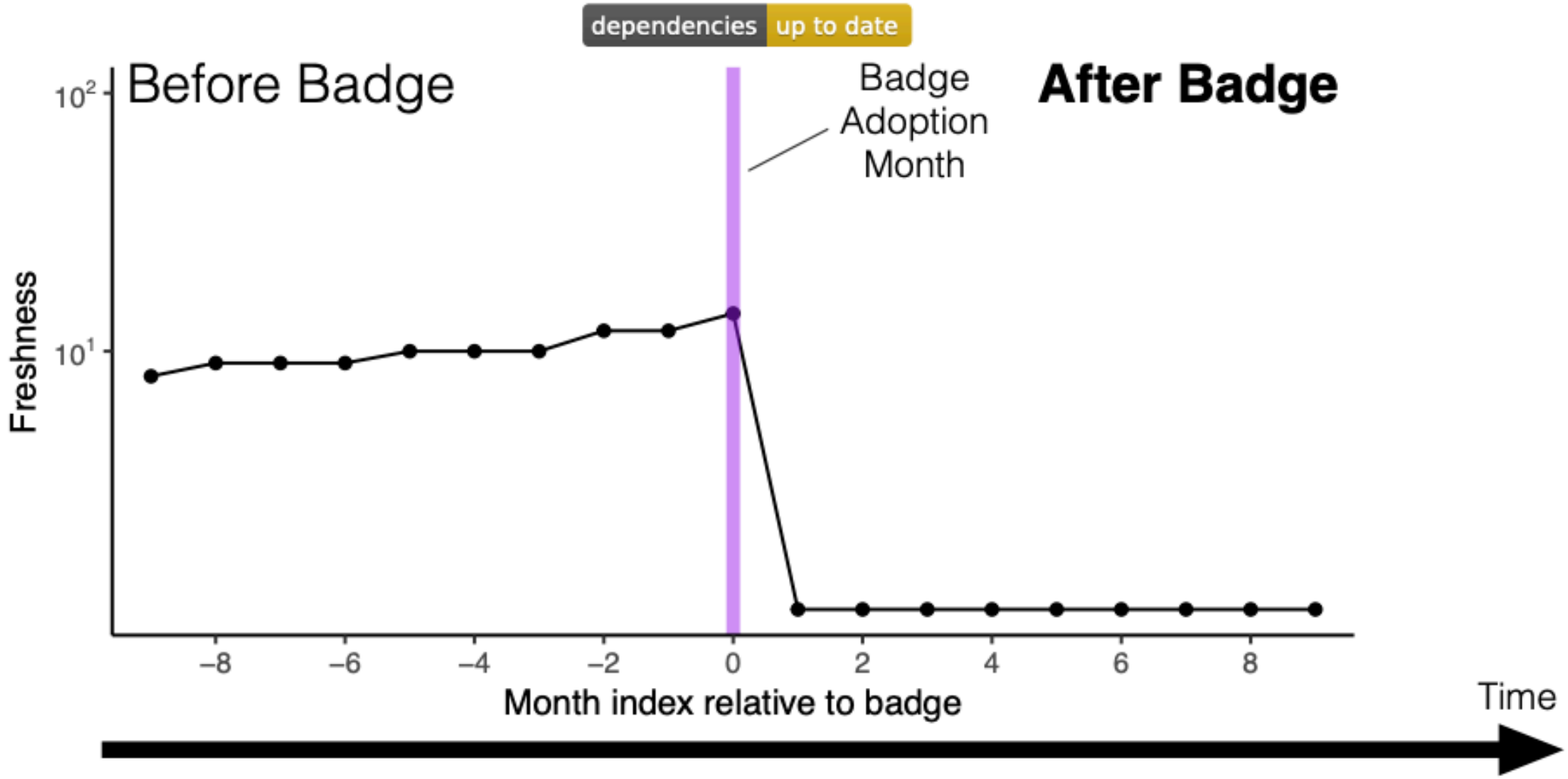
How much more does the badge tell you, relative to existing signals?

Time Series Analysis

How do things change after adding the badge?

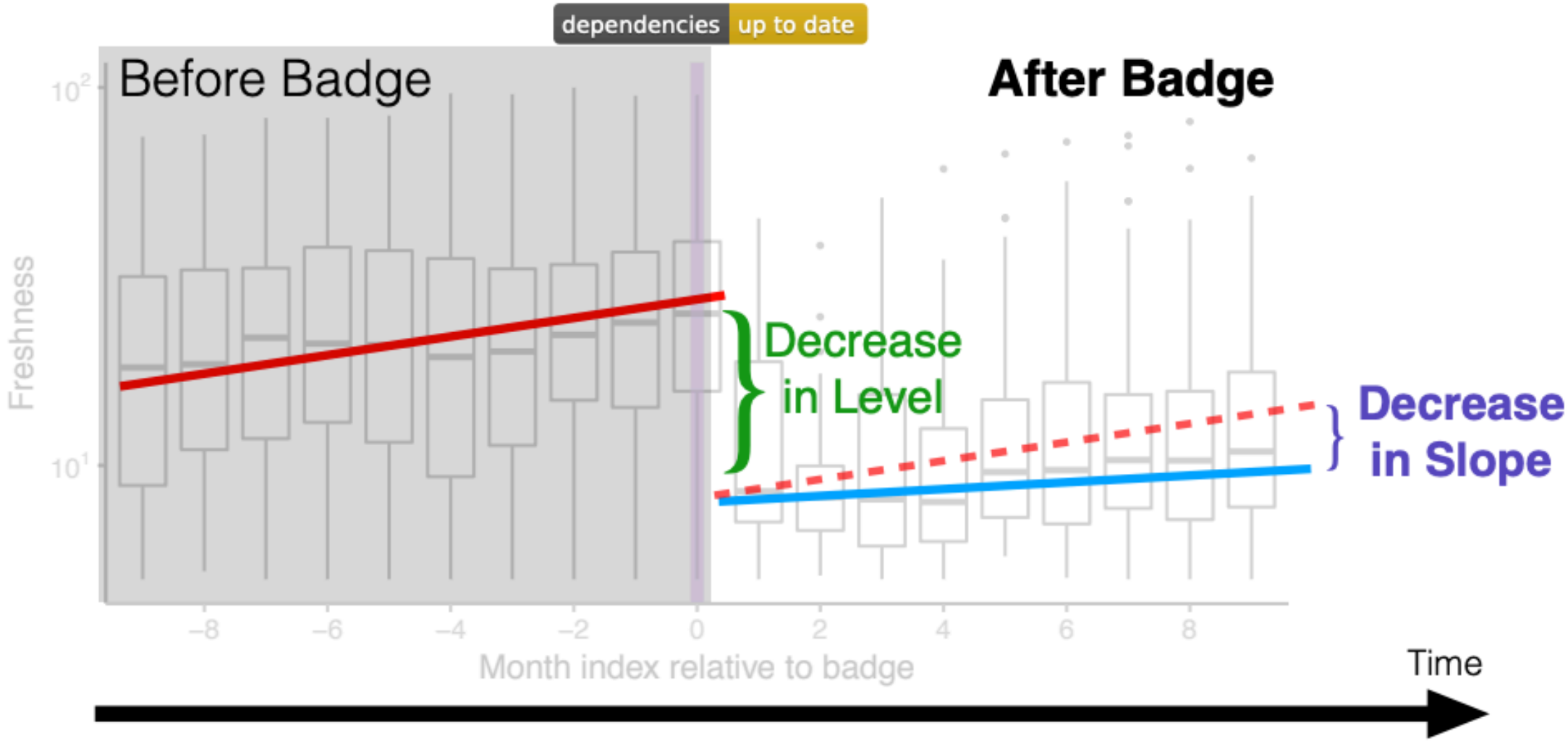
Step 2: Time Series Analysis

Signals of fresh dependencies



Step 2: Time Series Analysis

Signals of fresh dependencies



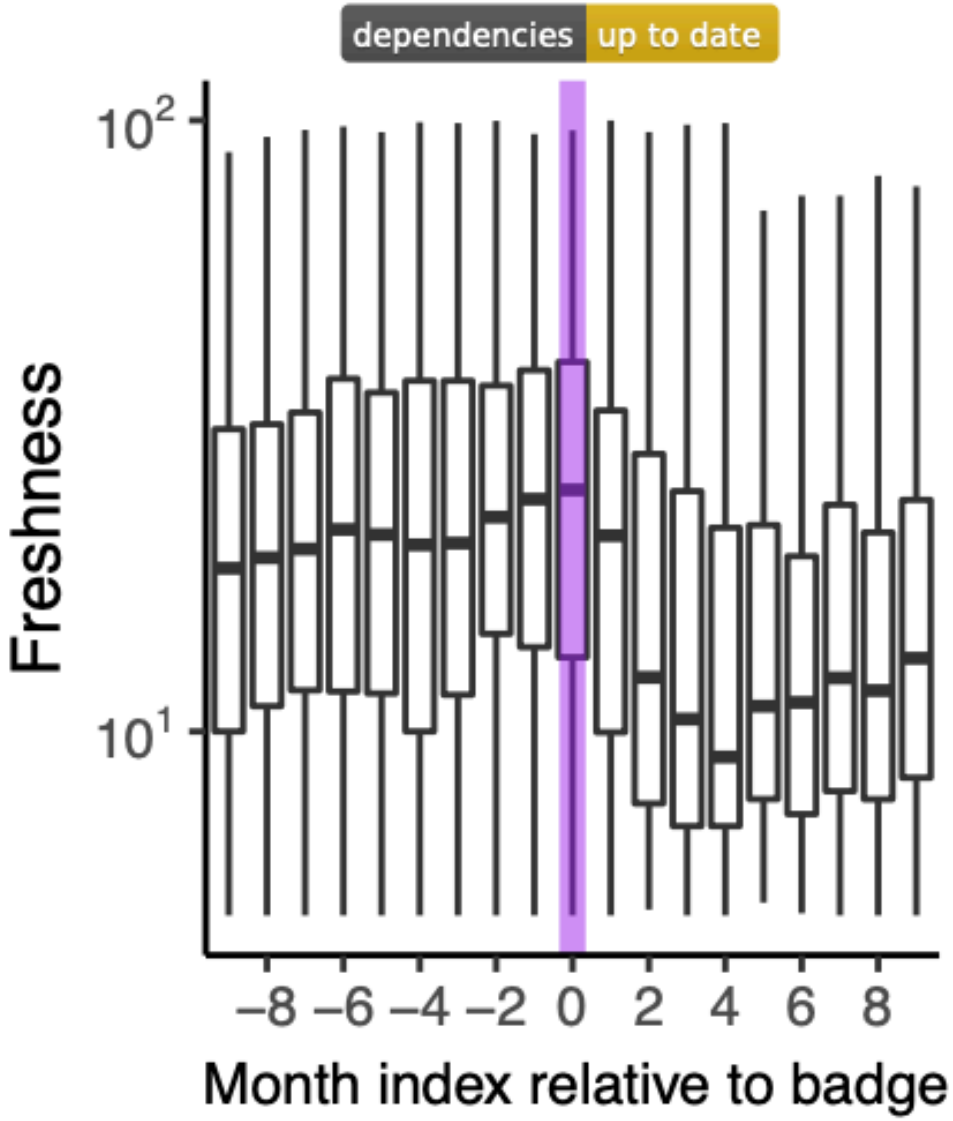
STRUDEL
Carnegie Mellon University

Step 2: Time Series Analysis

Signals of fresh dependencies

- **Based on survey:** The adoption of dependency management badges correlates with fresher dependencies
- **Freshness metric:** *lower* is better
  - (More up-to-date deps.)

Result: Dep. badges indicate *improved dep. management practices*



STRUDEL
Carnegie Mellon University

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

Take-aways

Badges with **underlying analyses:**

assessment signals

are **stronger predictors** than badges that merely state intentions or provide links:

conventional signals

STRUDEL
Carnegie Mellon University

Take-aways

When possible,
design or choose the badge that takes the **most work:**

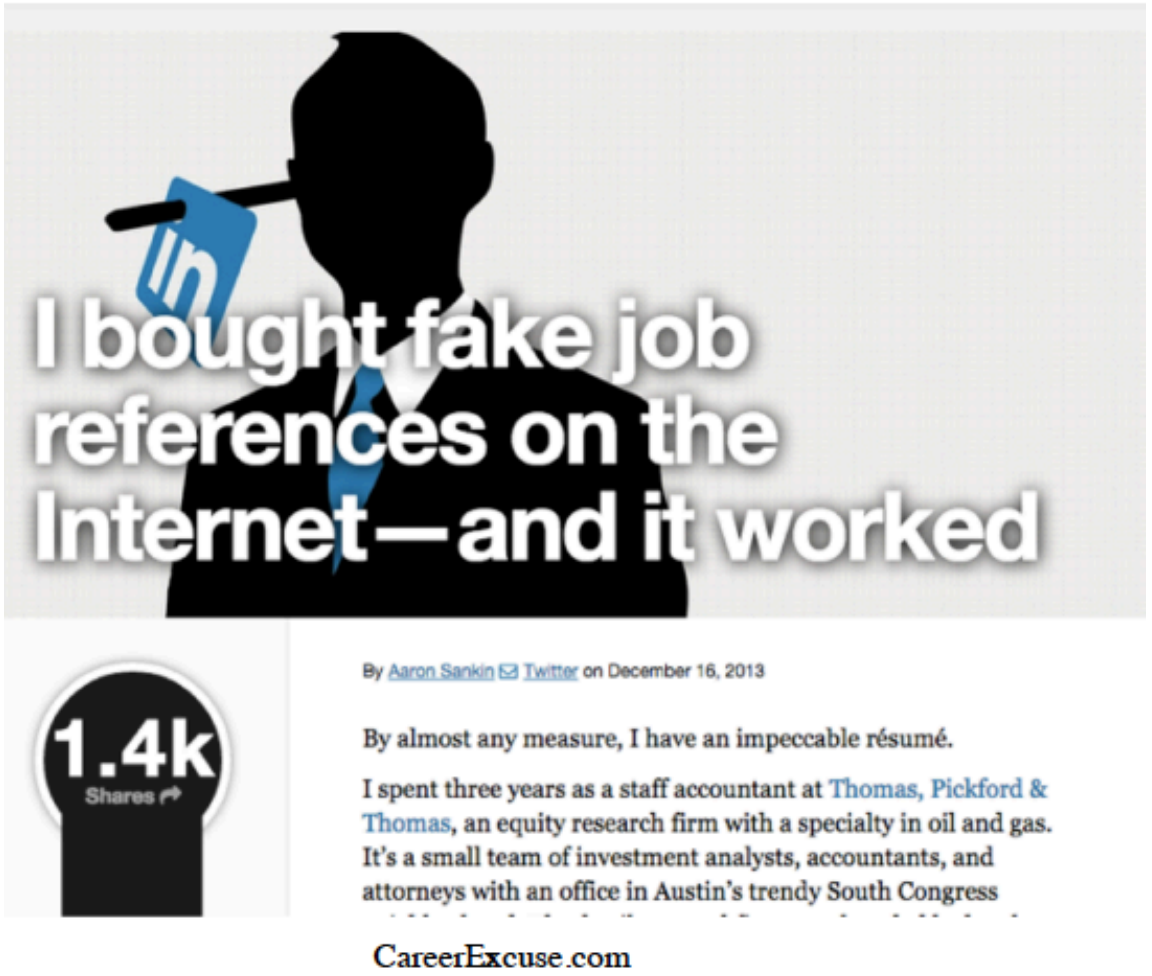
assessment signal

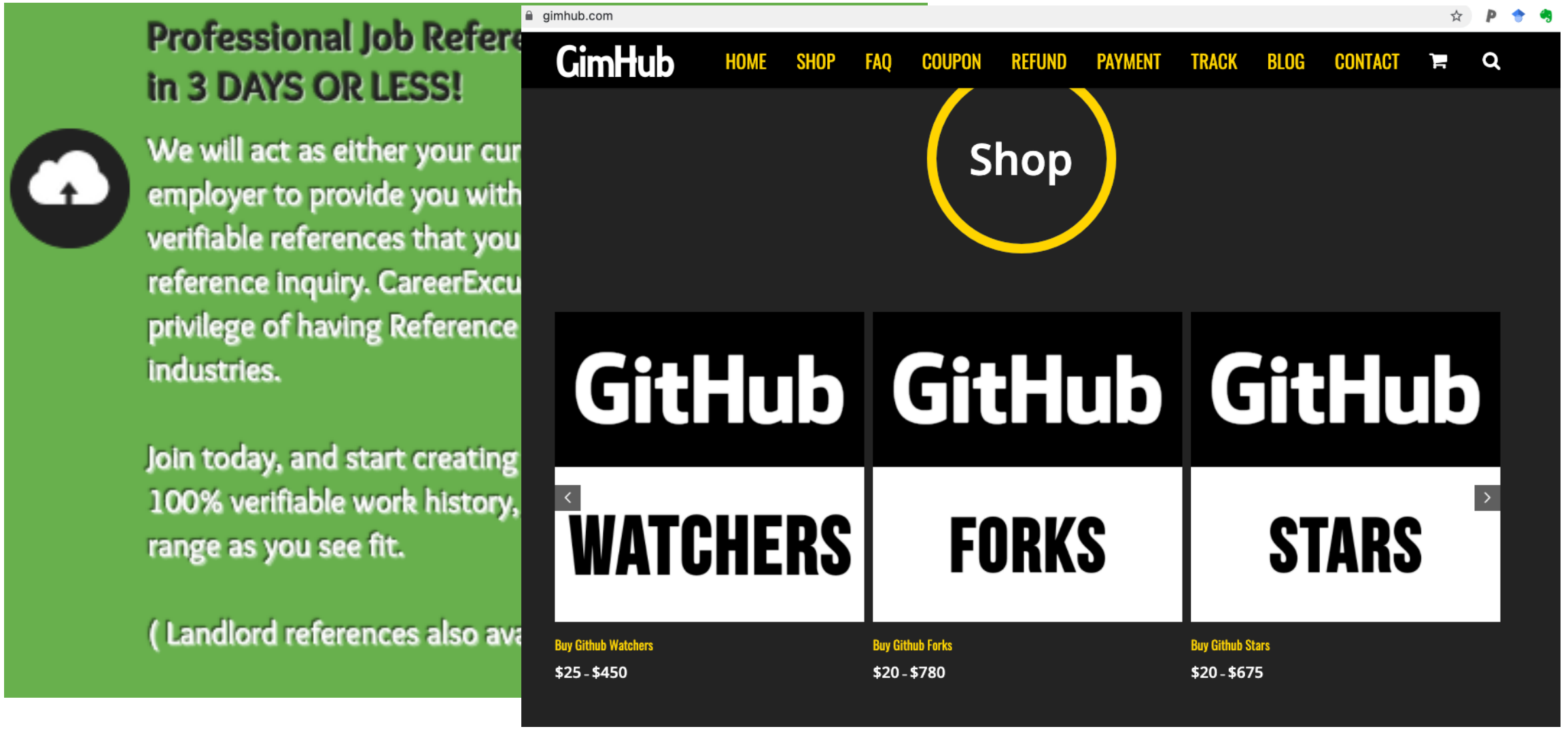
>

conventional signal

https://www.careerexcuse.com/

Conventional Signals – Trustworthy?




Professional Job Refere
in 3 DAYS OR LESS!
We will act as either your cur
employer to provide you with
verifiable references that you
reference inquiry. CareerExcu
privilege of having Reference
industries.
Join today, and start creating
100% verifiable work history,
range as you see fit.
(Landlord references also ava
gimhub.com
GimHub
HOME
SHOP
FAQ
COUPON
REFUND
PAYMENT
TRACK
BLOG
CONTACT
Shop
GitHub
WATCHERS
Buy Github Watchers
$25 - $450
GitHub
FORKS
Buy Github Forks
$20 - $780
GitHub
STARS
Buy Github Stars
$20 - $675

Indiatimes | The Times of India | The Economic Times | More ▾ Log In / Join FOLLOW ET:

THE ECONOMIC TIMES

Jobs

News | Industry | Politics | Markets | Wealth | MF | Tech | Jobs | Opinion | Blogs | NRI | Slideshows | Magazines | ET NOW | ET Speed | ET Portfolio

04:04 PM | 14 Nov EOD MARKET STATS

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

Submit

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...

The Theoretical Lens

- Our Theories impact how we see the world
  - Real-world phenomena too rich and complex
  - Need a way of filtering our observations
  - The theory guides us, whether it is explicitly stated or not
- In Quantitative Methods:
  - Theoretical lens tells you what variables to measure…
  - …and which to ignore or control
- In Qualitative Methods:
  - Theoretical lens usually applied after data is collected
  - …and used to help with labeling and categorizing the data

Theories are good for generalization…

Statistical Generalization

- Generalize from sample to population
- Can only be used for quantifiable variables
- Based on random sampling:
  - Test whether 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

- Generalize from findings to theory
- Applicable to quantitative and qualitative studies
- Compares findings with theory
  - Do the data support/refute the theory?
  - Do they support this theory better than rival theories?
- Supports empirical induction:
  - Evidence builds if subsequent studies also support the theory
- More powerful than stats
  - Doesn't rely on correlations
  - Examines underlying mechanisms

The Role of Theory Building

- Theories lie at the heart of what it means to do science.
  - Production of generalizable knowledge
- Theory provides orientation for data collection
  - Cannot observe the world without a theoretical perspective
- 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

Take home messages

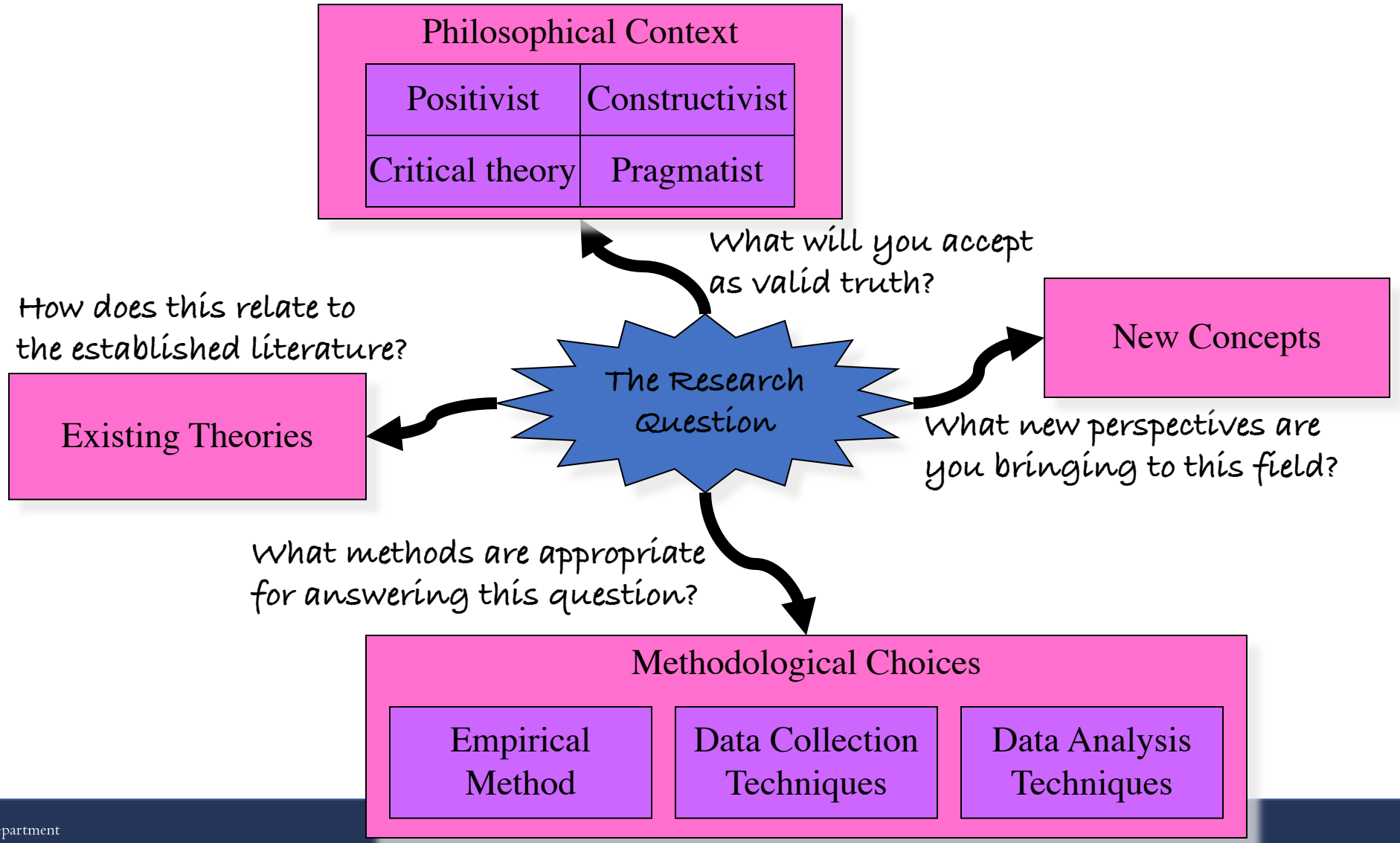
- Articulate the theory(s) underlying your work
- Be precise about your research questions
- Be explicit about your philosophical stance
- Use the theory to guide the study design

Test the Theory not the Tool

Planning Checklist

- [x] Pick a topic
- [x] Identify the research question(s)
- [x] Check the literature
- [x] Identify your philosophical stance
- [x] Identify appropriate theories
- [ ] Choose the method(s)
- [ ] Design the study
  - Unit of analysis?
  - Target population?
  - Sampling technique?
  - Data collection techniques?
  - Metrics for key variables?
  - Handle confounding factors
- [ ] Critically appraise the design for threats to validity
- [ ] Get IRB approval
  - Informed consent?
  - Benefits outweigh risks?
- [ ] Recruit subjects / field sites
- [ ] Conduct the study
- [ ] Analyze the data
- [ ] Write up the results and publish them
- [ ] Iterate

Putting the Question in Context



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

Choose a Method…

- Exploratory
  - Used to build new theories where we don't have any yet
  - E.g. What do CMM level 3 organizations have in common?
  - E.g. What are the experiences of developers who have adopted Ruby?
- Descriptive
  - Describes sequence of events and underlying mechanisms
  - E.g. How does pair programming actually work?
  - E.g. How do software immigrants naturalize?
- Causal
  - Determines whether there are causal relationship between phenomena
  - E.g. Does tool X lead to software with fewer defects?
  - E.g. Do requirements traceability tools help programmers find information more rapidly?
- Explanatory
  - Adjudicates between competing explanations (theories)
  - E.g. Why does software inspection work?
  - E.g. Why do people fail to document their requirements?

How will you substantiate your claims?

Common "in the lab" Methods	Common "in the wild" Methods
• Controlled Experiments • Rational Reconstructions • Exemplars • Benchmarks • Simulations	• Quasi-Experiments • Case Studies • Survey Research • Ethnographies • Action Research

- Artifact/Archive Analysis ("mining"!)

Planning Checklist

- [x] Pick a topic
- [x] Identify the research question(s)
- [x] Check the literature
- [x] Identify your philosophical stance
- [x] Identify appropriate theories
- [x] Choose the method(s)
- [ ] Design the study
  - Unit of analysis?
  - Target population?
  - Sampling technique?
  - Data collection techniques?
  - Metrics for key variables?
  - Handle confounding factors
- [ ] Critically appraise the design for threats to validity
- [ ] Get IRB approval
  - Informed consent?
  - Benefits outweigh risks?
- [ ] Recruit subjects / field sites
- [ ] Conduct the study
- [ ] Analyze the data
- [ ] Write up the results and publish them
- [ ] Iterate

Unit of Analysis

- Defines what phenomena you will analyze
  - Choice depends on the primary research questions
  - Choice affects decisions on data collection and analysis
  - Hard to change once the study has started (but can be done if there are compelling reasons)
  - If possible, use same unit of analysis as previous studies (why?)
- Often many choices:
  - E.g. for an exploratory study of agile programming:
    - Unit of analysis = individual developer (study focuses on a person's participation in the project)
    - Unit of analysis = a team (study focuses on team activities)
    - Unit of analysis = a decision (study focuses on activities around that decision)
    - Unit of analysis = a process (study examines how user stories are collected and prioritized)
    - …

Examples of Units of Analysis

- For a study of how software immigrants naturalize
  - Individuals?
  - … or the Development team?
  - … or the Organization?
- For a study of pair programming
  - Programming episodes?
  - … or Pairs of programmers?
  - … or the Development team?
  - … or the Organization?
- For a study of software evolution
  - A Modification report?
  - … or a File?
  - … or a System?
  - … or a Release?
  - … or a Stable release?

Planning Checklist

- [x] Pick a topic
- [x] Identify the research question(s)
- [x] Check the literature
- [x] Identify your philosophical stance
- [x] Identify appropriate theories
- [x] Choose the method(s)
- [ ] Design the study
  - [x] Unit of analysis?
  - [ ] Target population?
  - [ ] Sampling technique?
  - [ ] Data collection techniques?
  - [ ] Metrics for key variables?
  - [ ] Handle confounding factors
- [ ] Critically appraise the design for threats to validity
- [ ] Get IRB approval
  - [ ] Informed consent?
  - [ ] Benefits outweigh risks?
- [ ] Recruit subjects / field sites
- [ ] Conduct the study
- [ ] Analyze the data
- [ ] Write up the results and publish them
- [ ] Iterate

Target Population

- Determines scope of applicability of your results
  - If you don't define the target population…
  - …nobody will know whether your results apply to anything at all
- From what population are your units of analysis drawn?
  - UoA = "developer using agile programming"
  - Population =
    - All software developers in the world?
    - All developers who use agile methods?
    - All developers in Canadian Software Industry?
    - All developers in Small Companies in Ontario?
    - All students taking SE courses at U of T?
- Choice closely tied to choice of sampling method…

Planning Checklist

- [x] Pick a topic
- [x] Identify the research question(s)
- [x] Check the literature
- [x] Identify your philosophical stance
- [x] Identify appropriate theories
- [x] Choose the method(s)
- [ ] Design the study
  - [x] Unit of analysis?
  - [x] Target population?
  - [ ] Sampling technique?
  - [ ] Data collection techniques?
  - [ ] Metrics for key variables?
  - [ ] Handle confounding factors
- [ ] Critically appraise the design for threats to validity
- [ ] Get IRB approval
  - Informed consent?
  - Benefits outweigh risks?
- [ ] Recruit subjects / field sites
- [ ] Conduct the study
- [ ] Analyze the data
- [ ] Write up the results and publish them
- [ ] Iterate

Sampling Method

- Used to select representative set from a population
  - Simple Random Sampling - choose every kth element
  - Stratified Random Sampling - identify strata and sample each
  - Clustered Random Sampling - choose a representative subpopulation and sample it
  - Purposive Sampling - choose the parts you think are relevant without worrying about statistical issues (see next slide…)
- Sample Size is important
  - balance between cost of data collection/analysis and required significance
- Process:
  - Decide what data should be collected
  - Determine the population
  - Choose type of sample
  - Choose sample size

Purposive Sampling

- Typical Case
  - Identify typical, normal, average case
- Extreme or Deviant Case
  - E.g outstanding success/notable failures, exotic events, crises.
- Critical Case
  - if it's true of this one case it's likely to be true of all other cases.
- Intensity
  - Information-rich examples that clearly show the phenomenon (but not extreme)
- Maximum Variation
  - choose a wide range of variation on dimensions of interest
- Homogeneous
  - Instance has little internal variability - simplifies analysis
- Snowball or Chain
  - Select cases that should lead to identification of further good cases
- Criterion
  - All cases that meet some criterion
- Confirming or Disconfirming
  - Exceptions, variations on initial cases
- Opportunistic
  - Rare opportunity where access is normally hard/impossible
- Politically Important Cases
  - Attracts attention to the study
- Convenience sampling
  - Cases that are easy/cheap to study
  - (May reduce credibility)
- …Or any combination of the above

Planning Checklist

- [x] Pick a topic
- [x] Identify the research question(s)
- [x] Check the literature
- [x] Identify your philosophical stance
- [x] Identify appropriate theories
- [x] Choose the method(s)
- [ ] Design the study
  - [x] Unit of analysis?
  - [x] Target population?
  - [x] Sampling technique?
  - [ ] Data collection techniques?
  - [ ] Metrics for key variables?
  - [ ] Handle confounding factors
- [ ] Critically appraise the design for threats to validity
- [ ] Get IRB approval
  - Informed consent?
  - Benefits outweigh risks?
- [ ] Recruit subjects / field sites
- [ ] Conduct the study
- [ ] Analyze the data
- [ ] Write up the results and publish them
- [ ] Iterate

Data Collection Techniques

- Direct Techniques
  - Brainstorming / Focus Groups
  - Interviews
  - Questionnaires
  - Conceptual Modeling
  - Work Diaries
  - Think-aloud Sessions
  - Shadowing and Observation
  - Participant Observation
- Indirect Techniques
  - Instrumented Systems
  - Fly on the wall
- Independent Techniques
  - Analysis of work databases
  - Analysis of tool usage logs
  - Documentation Analysis
  - Static and Dynamic Analysis

Planning Checklist

- [x] Pick a topic
- [x] Identify the research question(s)
- [x] Check the literature
- [x] Identify your philosophical stance
- [x] Identify appropriate theories
- [x] Choose the method(s)
- [ ] Design the study
  - [x] Unit of analysis?
  - [x] Target population?
  - [x] Sampling technique?
  - [x] Data collection techniques?
  - [ ] Metrics for key variables?
  - [ ] Handle confounding factors
- [ ] Critically appraise the design for threats to validity
- [ ] Get IRB approval
  - Informed consent?
  - Benefits outweigh risks?
- [ ] Recruit subjects / field sites
- [ ] Conduct the study
- [ ] Analyze the data
- [ ] Write up the results and publish them
- [ ] Iterate

How will you measure things?

Type	Meaning	Admissible Operations
Nominal Scale	Unordered classification of objects	=
Ordinal Scale	Ranking of objects into ordered categories	=, <, >
Interval Scale	Differences between points on the scale are meaningful	=, <, >, difference, mean
Ratio Scale	Ratios between points on the scale are meaningful	=, <, >, difference, mean, ratio
Absolute Scale	No units necessary - scale cannot be transformed	=, <, >, difference, mean, ratio

Planning Checklist

- [x] Pick a topic
- [x] Identify the research question(s)
- [x] Check the literature
- [x] Identify your philosophical stance
- [x] Identify appropriate theories
- [x] Choose the method(s)
- [ ] Design the study
  - [x] Unit of analysis?
  - [x] Target population?
  - [x] Sampling technique?
  - [x] Data collection techniques?
  - [x] Metrics for key variables?
  - [ ] Handle confounding factors
- [ ] Critically appraise the design for threats to validity
- [ ] Get IRB approval
  - Informed consent?
  - Benefits outweigh risks?
- [ ] Recruit subjects / field sites
- [ ] Conduct the study
- [ ] Analyze the data
- [ ] Write up the results and publish them
- [ ] Iterate

What could go wrong?

- Many phenomena might affect your results
- Must be able to distinguish:
  - My results follow clearly from the phenomena I observed
  - My results were caused by phenomena that I failed to observe
- Identify all (likely) confounding variables
- For each, decide what to do:
  - Selection/Exclusion
  - Balancing
  - Manipulation
  - Ignore (with justification)

Agenda for Today

- Study Design Planning Checklist
- Homework debrief


8%
45%
17%
$
2%
28%

Discussion

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