

# Open Review Toolkit:

Better books, higher sales, increased access to knowledge

Matthew J. Salganik

Department of Sociology  
Princeton University

June 26, 2018  
Summer Institute in Computational Social Science

The Open Review Toolkit is supported by a grant from the Alfred P. Sloan Foundation. I would like to thank the Agathon Group—especially Luke Baker and Paul Yuen—for web development. I would also like to thank Princeton University Press—especially Meagan Levinson—for many valuable conversations.



Imagine a world in which every single person on the planet is given free access to the sum of all human knowledge. – Jimmy Wales



[http://commons.wikimedia.org/wiki/File:The\\_Earth\\_seen\\_from\\_Apollo\\_17.jpg](http://commons.wikimedia.org/wiki/File:The_Earth_seen_from_Apollo_17.jpg)

Imagine a world where you have the job of your dreams.



[http://commons.wikimedia.org/wiki/File:Nassau\\_Hall\\_Princeton.JPG](http://commons.wikimedia.org/wiki/File:Nassau_Hall_Princeton.JPG)

If you are a university press, imagine a world where you have a sustainable business model that is consistent with your values.



[https://commons.wikimedia.org/wiki/File:Money\\_Cash.jpg](https://commons.wikimedia.org/wiki/File:Money_Cash.jpg)

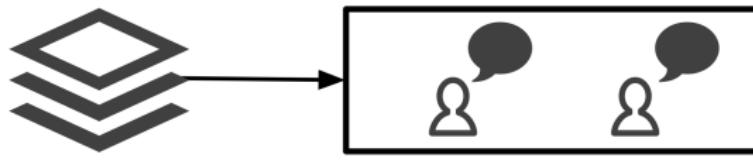


# OPEN REVIEW TOOLKIT

[www.openreviewtoolkit.org](http://www.openreviewtoolkit.org)



Book  
Manuscript



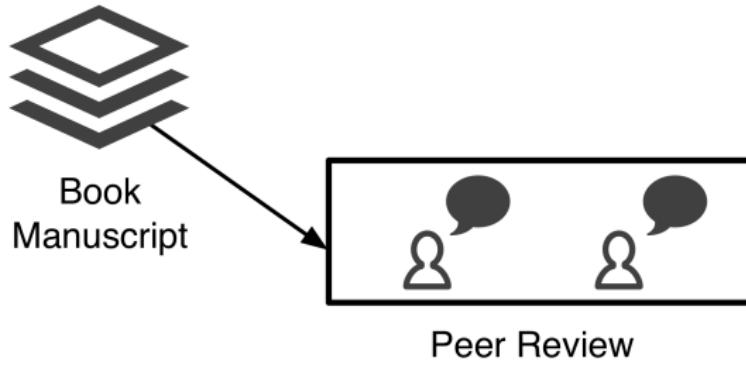
Book  
Manuscript

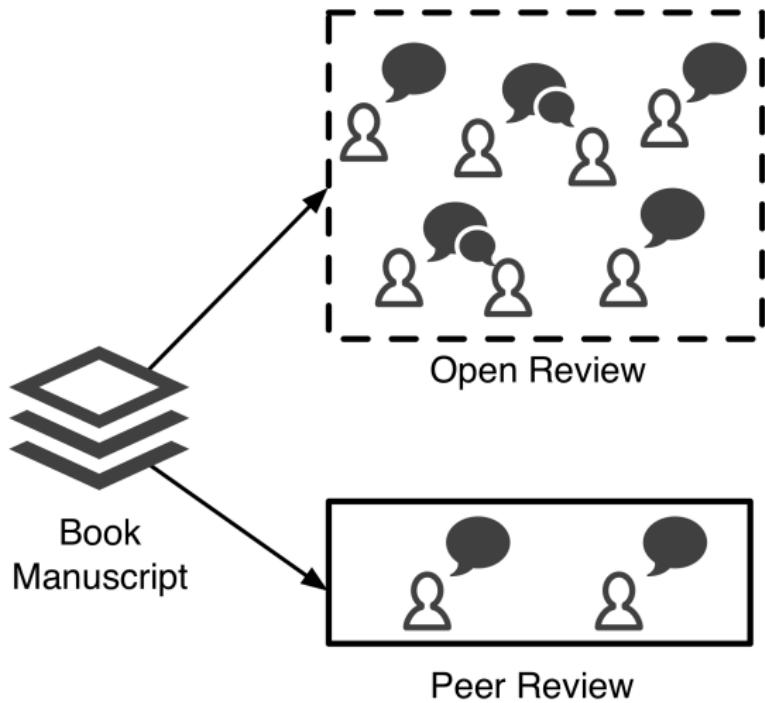
Peer Review

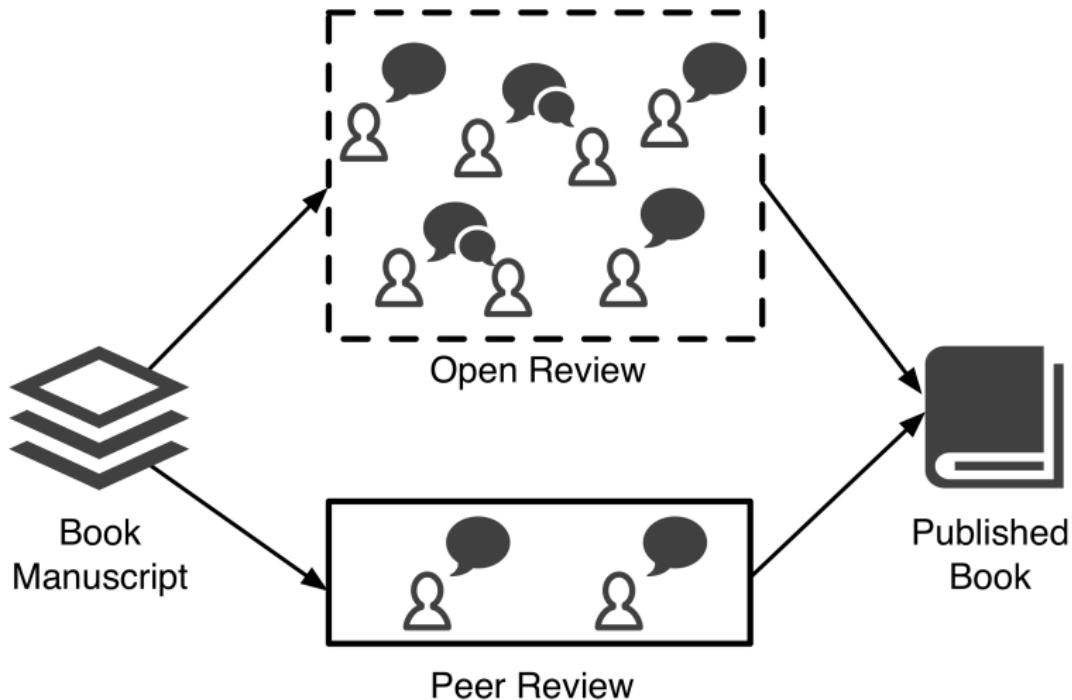




Book  
Manuscript







The Open Review process has three goals:

- ▶ Better books

The Open Review process has three goals:

- ▶ Better books
- ▶ Higher sales

The Open Review process has three goals:

- ▶ Better books
- ▶ Higher sales
- ▶ Increased access to knowledge

The screenshot shows a web browser window with the URL [www.bitbybitbook.com/en/introduction/ink-blot/](http://www.bitbybitbook.com/en/introduction/ink-blot/). The page title is "Bit By Bit - Social Research in the Digital Age". The main content area displays the first chapter, "1.1 An ink blot". The sidebar on the left lists chapters 1 through 7. The right side features a sidebar for annotations and a main content area with several annotations from users like "efosse" and "fasha".

**Annotations:**

- efosse** (Sep 9): I like the Blumenstock et al. example, but I think the introduction would show the immense change going on with a parallel example from the analog age. E.g., compare Blau and Duncan's work on the American Occupational Structure, which required specifying hypotheses weeks in advance and entailed slow computation with punch cards.
- fasha** (Oct 8): They used the survey data to train a machine learning model to predict someone's wealth from their call data, and the ... [More](#)
- fasha** (Oct 8): I have some background in machine learning—which may be an aid or an obstacle in understanding this description of Blumenstock's work—but it's not obvious to me what “call logs” and “call data” mean here.
- fasha** (Oct 8): The researchers collected survey data on a small sample of the entire population—ok.
- fasha** (Oct 8): They trained a model that consumes “call data” and estimates wealth—but what is “call data” here? The survey data without the wealth data? Or is it cell phone usage data from the cell phone company?
- fasha** (Oct 8): This is where the model on data for all customers. This is what makes me think the model was trained on some kind of cell phone data, because that's the only data available for people outside the survey.

## Better books

- ▶ hundreds of annotations from dozens of people

The screenshot shows a web browser window with the URL [www.bitbybitbook.com/en/introduction/ink-blot/](http://www.bitbybitbook.com/en/introduction/ink-blot/). The page title is "Bit By Bit - Social Research in the Digital Age". The main content area displays the first chapter, "1.1 An ink blot". The sidebar on the left lists chapters 1 through 7. The right side of the screen shows an annotation interface where users can add comments and notes to specific sections of the text.

**Annotations:**

- efosse** (Sep 9)
  - 1.1 An ink blot

I like the Blumenstock et al. example, but I think the introduction would show the immense change going on with a parallel example from the analog age. E.g., compare Blau and Duncan's work on the American Occupational Structure, which required specifying hypotheses weeks in advance and entailed slow computation with punch cards.
- fashfa** (Oct 8)
  - They used the survey data to train a machine learning model to predict someone's wealth from their call data, and the ... [More](#)
  - I have some background in machine learning—which may be an aid or an obstacle in understanding this description of Blumenstock's work—but it's not obvious to me what “call logs” and “call data” mean here.
  - The researchers collected survey data on a small sample of the entire population—ok.
  - They trained a model that consumes “call data” and estimates wealth—but what is “call data” here? The survey data without the wealth data? Or is it cell phone usage data from the cell phone company?
  - This is what makes me think the model was trained on some kind of cell phone data, because that's the only data available for people outside the survey.

## Better books

- ▶ hundreds of annotations from dozens of people
- ▶ implicit feedback

The screenshot shows a web browser window with the URL [www.bitbybitbook.com/en/introduction/ink-blot/](http://www.bitbybitbook.com/en/introduction/ink-blot/). The page is titled "Bit By Bit: Social Research in the Digital Age". The main content area displays the first section of the book, "1.1 An ink blot". The sidebar on the left lists chapters from "1.1 An ink blot" to "7. The future". On the right, there is a sidebar for annotations and a main content area with two annotations by users "efosse" and "tasha".

**Annotations:**

- efosse** (Sep 9): I like the Blumenstock et al. example, but I think the introduction would show the immense change going on with a parallel example from the analog age. E.g., compare Blau and Duncan's work on the American Occupational Structure, which required specifying hypotheses weeks in advance and entailed slow computation with punch cards.
- tasha** (Oct 8): They used the survey data to train a machine learning model to predict someone's wealth from their call data... More
- tasha** (Oct 8): I have some background in machine learning—which may be an aid or an obstacle in understanding this description of Blumenstock's work—but it's not obvious to me what “call logs” and “call data” mean here.
- tasha** (Oct 8): The researchers collected survey data on a small sample of the entire population—ok.
- tasha** (Oct 8): They trained a model that consumes “call data” and estimates wealth—but what is “call data” here? The survey data without the wealth data? Or is it cell phone usage data from the cell phone company?
- tasha** (Oct 8): This is where the model on data for all customers. This is what makes me think the model was trained on some kind of cell phone data, because that's the only data available for people outside the survey.

## Better books

- ▶ hundreds of annotations from dozens of people
- ▶ implicit feedback
- ▶ psychological fuel

Bit By Bit - Introduction - 1.1 An ink blot

www.bitbybitbook.com/en/introduction/ink-blot/

Bit By Bit: Social Research in the Digital Age | Open Review Draft | About | Languages | Buy the book | Public | Sign up / Log in | Annotations 8 | Page Notes 1

Preface

1.1 An ink blot

1.2 Welcome to the digital age

1.3 Research design

1.4 Themes of this book

1.5 Outline of the book

2 Observing behavior

3 Asking questions

4 Running experiments

5 Mass collaboration

6 Ethics

7 The future

## 1.1 An ink blot

In the summer of 2009, mobile phones were ringing all across Rwanda. In addition to the millions of calls between family, friends, and business associates, about 1,000 Rwandans received a call from Joshua Blumenstock and his colleagues. The researchers were studying wealth and poverty by conducting a survey of people who had been randomly sampled from a database of 1.5 million customers from Rwanda's largest mobile phone provider. Blumenstock and colleagues asked the participants if they wanted to participate in a survey, explained the nature of the research to them, and then asked a series of questions about their demographic, social, and economic characteristics.

Everything I have said up until now makes this sound like a traditional social science survey. But, what comes next is not traditional, at least not yet. They used the survey data to train a machine learning model to predict someone's wealth from their call data, and then they used this model to estimate the wealth of all 1.5 million customers. Next, they estimated the place of residence of all 1.5 million customers by using the geographic information

To annotate this document [create a free account](#) or [log in](#)

Annotations 8 | Page Notes 1

**efosse** | 1.1 An ink blot | Sep 9

I like the Blumenstock et al. example, but I think the introduction should show the immense change going on with a parallel example from the analog age. E.g., compare Blau and Duncan's work on the American Occupational Structure, which required specifying hypotheses weeks in advance and entailed slow computation with punch cards.

**tasha** | Oct 6

They used the survey data to train a machine learning model to predict someone's wealth from their call data, and the ... More

I have some background in machine learning—which may be an aid or an obstacle in understanding this description of Blumenstock's work—but it's not obvious to me what "call logs" and "call data" mean here.

The researchers collected survey data on a small sample of the entire population—ok.

They trained a model that consumes "call data" and estimates wealth—but what is "call data" here? The survey data without the wealth data? Or is it cell phone usage data from the cell phone company?

This is where the model on data for all customers. This is what makes me think the model was trained on some kind of cell phone data, because that's the only data available for people outside the survey.

## Better books

- ▶ hundreds of annotations from dozens of people
- ▶ implicit feedback
- ▶ psychological fuel

More information: <http://blog.press.princeton.edu/2018/04/18/matthew-salganik-the-open-review-of-bit-by-bit-part-1-better-books/>

# **Bit by Bit: Social Research in the Digital Age** Hardcover –

December 5, 2017

by [Matthew J. Salganik](#) (Author)

#1 New Release in [Social Sciences Methodology](#)

## **Higher sales**

- ▶ email addresses

# **Bit by Bit: Social Research in the Digital Age** Hardcover –

December 5, 2017

by [Matthew J. Salganik](#) (Author)

#1 New Release in [Social Sciences Methodology](#)

## **Higher sales**

- ▶ email addresses
- ▶ market intelligence

# **Bit by Bit: Social Research in the Digital Age** Hardcover –

December 5, 2017

by [Matthew J. Salganik](#) (Author)

#1 New Release in [Social Sciences Methodology](#)

## **Higher sales**

- ▶ email addresses
- ▶ market intelligence
- ▶ course adoptions

# **Bit by Bit: Social Research in the Digital Age** Hardcover –

December 5, 2017

by [Matthew J. Salganik](#) (Author)

#1 New Release in [Social Sciences Methodology](#)

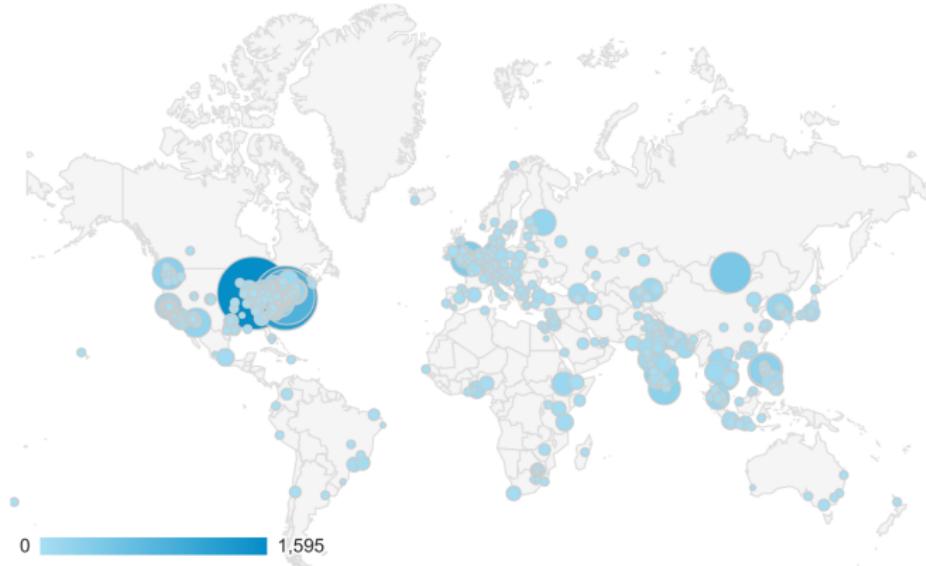
## **Higher sales**

- ▶ email addresses
- ▶ market intelligence
- ▶ course adoptions

## More information:

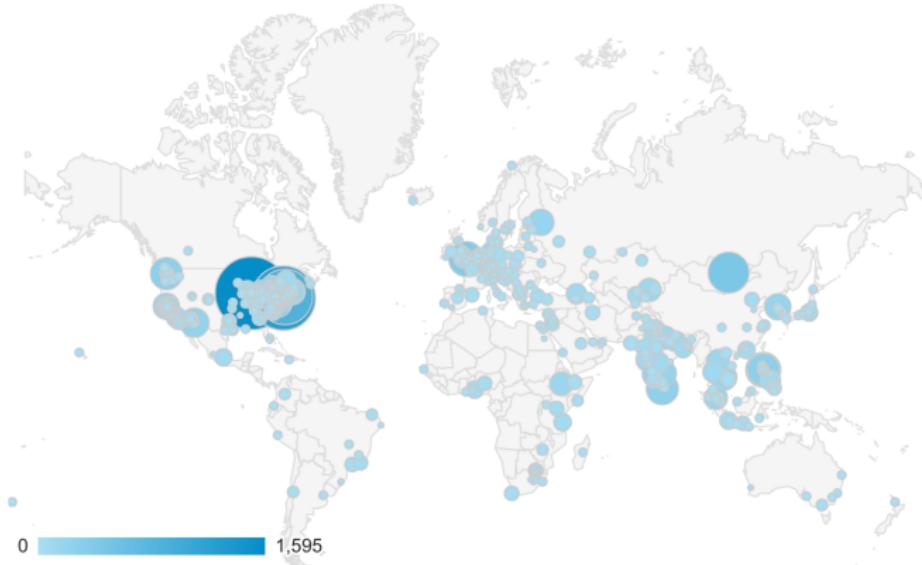
<http://blog.press.princeton.edu/2018/04/19/>

[matthew-salganik-the-open-review-of-bit-by-bit-part-2-higher-sales/](http://matthew-salganik-the-open-review-of-bit-by-bit-part-2-higher-sales/)



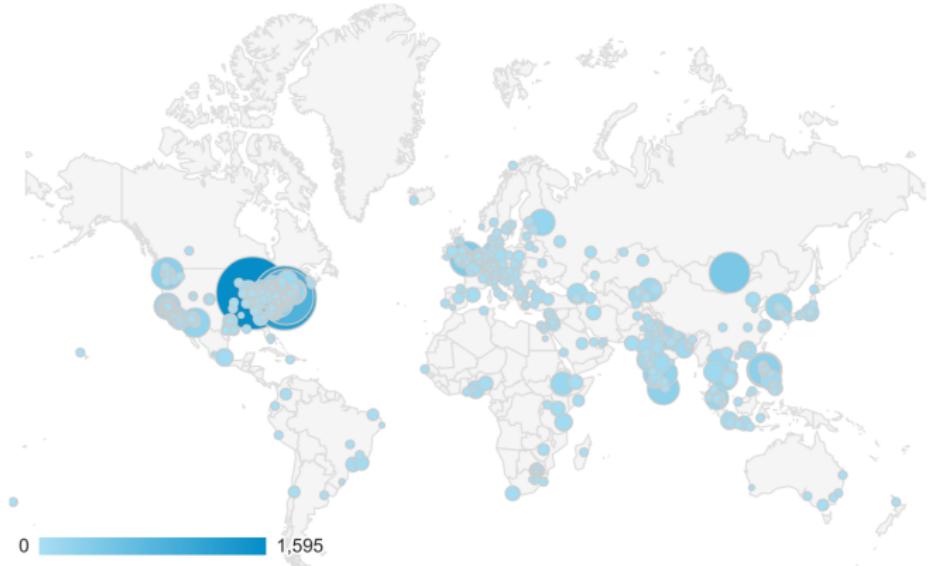
## Increased access to knowledge

- ▶ English version



## **Increased access to knowledge**

- ▶ English version
  - ▶ Machine translations



## Increased access to knowledge

- ▶ English version
- ▶ Machine translations

More information:

<http://blog.press.princeton.edu/2018/04/20/>

[matthew-salganik-the-open-review-of-bit-by-bit-part-3-increased-access-to-knowledge/](#)

Review stage is the ideal place to inject openness

- ▶ minimizes risk of cannibalizing sales
- ▶ maximizes benefits to authors

Aren't you giving your book away for free?

Open Review Toolkit =  
Hypothes.is + Pandoc + Google Analytics + Google Forms +  
Plumbing

Code available: [www.github.com/open-review-toolkit](https://www.github.com/open-review-toolkit)

Hire a developer: [www.openreviewtoolkit.org/hire-a-developer/](https://www.openreviewtoolkit.org/hire-a-developer/)



## BETTER BOOKS, HIGHER SALES, & INCREASED ACCESS TO KNOWLEDGE

The Open Review Toolkit is a platform designed to enhance the review process for academic books. It features a modular interface where reviews can be broken down into smaller, manageable components. The toolkit includes tools for managing annotations, tracking review progress, and generating reports. It also provides a space for open discussion and collaboration among reviewers.

**Annotations** feature allows users to add notes, questions, or comments directly onto the text of the book. These annotations can be shared with other reviewers or used for future reference. The platform supports multiple languages and offers a range of customization options for different types of reviews.

The toolkit is built with a focus on accessibility and user experience, making it easier for scholars to engage with their peers and contribute to the academic conversation. By providing a central hub for reviews, the Open Review Toolkit aims to improve the quality and efficiency of the peer review process.

[www.openreviewtoolkit.org](http://www.openreviewtoolkit.org)