

# Lessons Learned from Building a Recommender System

Jeremy Salfen



September 14, 2017



## PROBLEM #1

Building at production scale

## PROBLEM #2

Escaping a local optimum

## PROBLEM #3

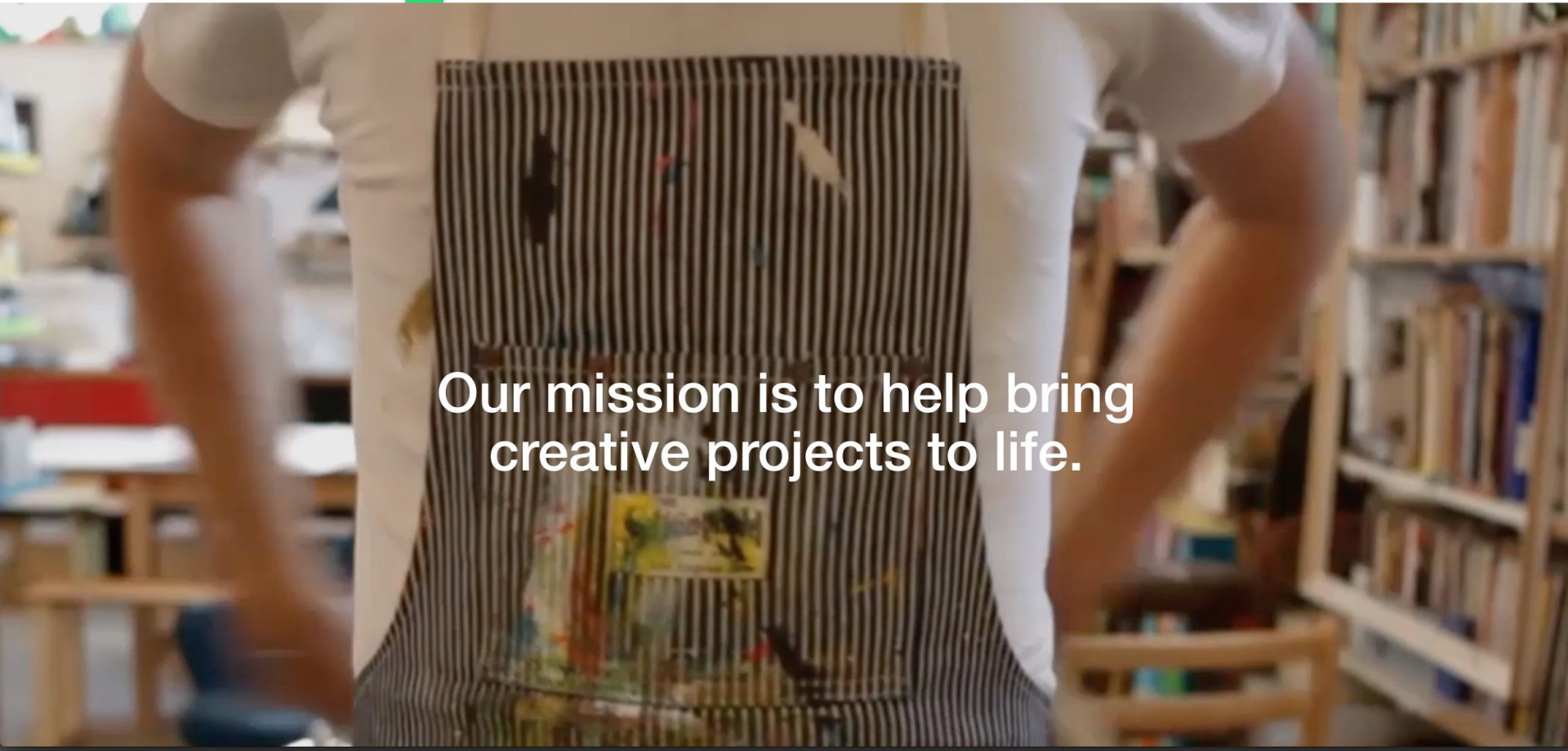
Knowing when worse is good enough

## PROBLEM #4

Fixing a cold start

## PROBLEM #5

Blending it all together

[Explore](#)[Start a project](#)[About us](#)**KICKSTARTER**[Log in](#)[Sign up](#)[Hello](#)[Charter](#)[Impact](#)[Team](#)[Jobs](#)[Press](#)[Stats](#)[Happening](#)A background image showing the back of a person wearing a white t-shirt and a black and white striped apron. The apron is covered in various paint splatters and has a pocket with a colorful abstract painting. The person is standing in a room with wooden bookshelves filled with books, suggesting a library or a creative workspace.

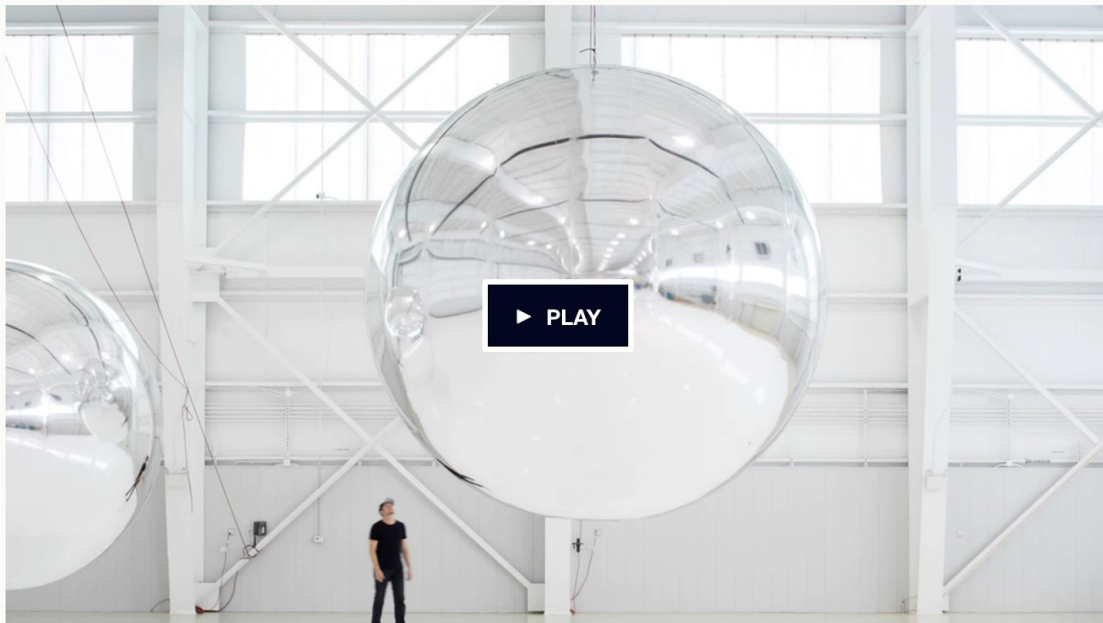
Our mission is to help bring  
creative projects to life.

[Explore](#)[Start a project](#)[About us](#)**KICKSTARTER**[Log in](#)[Sign up](#)

By **Nevada Museum of Art**  
First created

# Trevor Paglen: Orbital Reflector

In collaboration with Nevada Museum of Art, artist Trevor Paglen will launch the first satellite to exist purely as an artistic gesture

**\$12,115**

pledged of \$70,000 goal

**121**

backers

**35**

days to go

[Back this project](#)[♥ Remind me](#)[↗ Share](#)

All or nothing. This project will only be funded if it reaches its goal by Sun, October 8 2017 7:00 PM EDT.

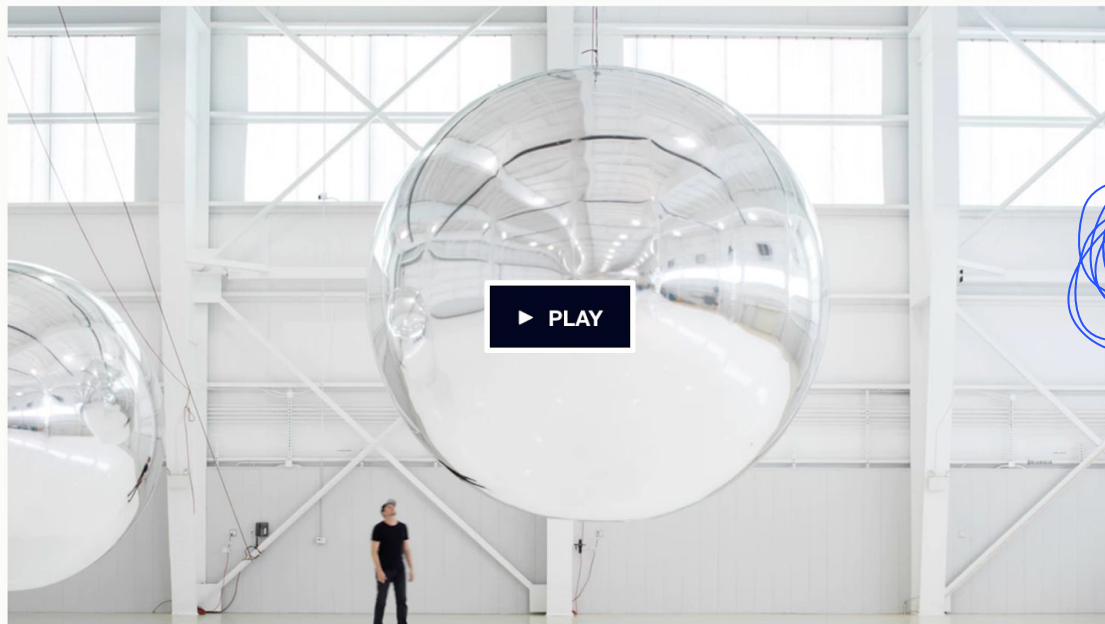
[Project We Love](#)[Projects of Earth](#)[Art](#)[Reno, NV](#)



By **Nevada Museum of Art**  
First created

# Trevor Paglen: Orbital Reflector

In collaboration with Nevada Museum of Art, artist Trevor Paglen will launch the first satellite to exist purely as an artistic gesture



**\$12,115**

pledged of \$70,000 goal

**121**

backers

**35**

days to go

Back this project

♥ Remind me

➦ Share

All or nothing. This project will only be funded if it reaches its goal by Sun, October 8 2017 7:00 PM EDT.



Project We Love



Projects of Earth

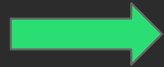


Art

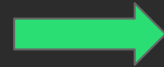


Reno, NV

Right  
projects



Right  
users



Right  
time



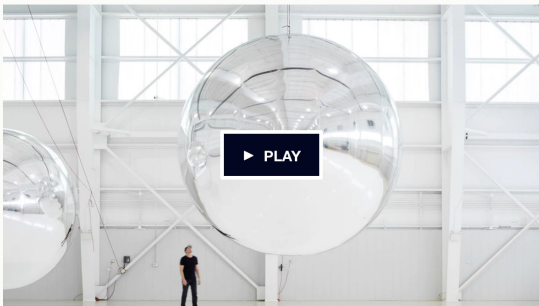




By Nevada  
Museum of Art  
First created

## Trevor Paglen: Orbital Reflector

In collaboration with Nevada Museum of Art, artist Trevor Paglen will launch the first satellite to exist purely as an artistic gesture



**K** Project We Love **E** Projects of Earth **A** Art **R**eno, NV

**\$14,789**

pledged of \$70,000 goal

**140**

backers

**33**

days to go

Back this project

Remind me

Share

All or nothing. This project will only be funded if it reaches its goal by Sun, October 8 2017 7:00 PM EDT.



By Spencer  
Merolla  
First created

## Coal Comforts: A Climate Change Art "Bakery"

Coal Comforts is a mobile art "bakery" where the sweet treats are not what they seem. These (inedible) desserts are made from coal ash.



**K** Project We Love **I**nstallations **R** Brooklyn, NY

**\$4,700**

pledged of \$3,750 goal

**41**

backers

**21**

days to go

Back this project

Remind me

Share

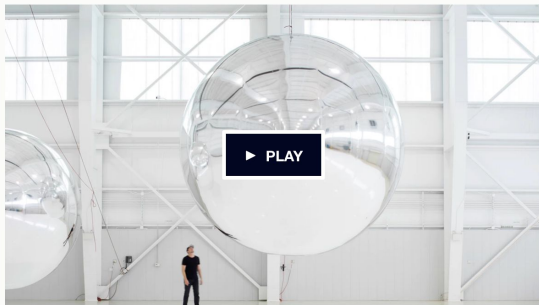
All or nothing. This project will only be funded if it reaches its goal by Tue, September 26 2017 12:31 PM EDT.



By Nevada  
Museum of Art  
First created

## Trevor Paglen: Orbital Reflector

In collaboration with Nevada Museum of Art, artist Trevor Paglen will launch the first satellite to exist purely as an artistic gesture



**K** Project We Love Projects of Earth Art Reno, NV

**\$14,789**

pledged of \$70,000 goal

**140**

backers

**33**

days to go

Back this project

Remind me

Share

All or nothing. This project will only be funded if it reaches its goal by Sun, October 8 2017 7:00 PM EDT.



By Spencer  
Merolla  
First created

## Coal Comforts: A Climate Change Art "Bakery"

Coal Comforts is a mobile art "bakery" where the sweet treats are not what they seem. These (inedible) desserts are made from coal ash.



**K** Project We Love Installations Brooklyn, NY

**\$4,700**

pledged of \$3,750 goal

**41**

backers

**11**

days to go

Back this project

Remind me

Share

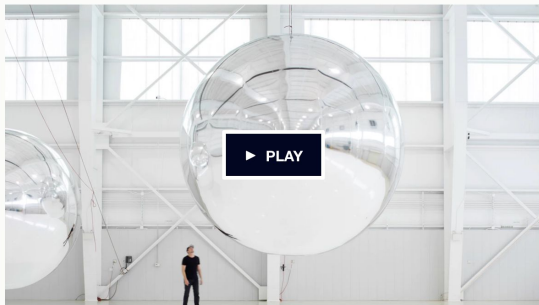
All or nothing. This project will only be funded if it reaches its goal by Tue, September 26 2017 12:31 PM EDT.



By Nevada  
Museum of Art  
First created

## Trevor Paglen: Orbital Reflector

In collaboration with Nevada Museum of Art, artist Trevor Paglen will launch the first satellite to exist purely as an artistic gesture



**K** Project We Love Projects of Earth Art Reno, NV

**\$14,789**  
pledged of \$70,000 goal

**140**  
backers

**33**  
days to go

Back this project

Remind me

Share

All or nothing. This project will only be funded if it reaches its goal by Sun, October 8 2017 7:00 PM EDT.



By Spencer  
Merolla  
First created

## Coal Comforts: A Climate Change Art "Bakery"

Coal Comforts is a mobile art "bakery" where the sweet treats are not what they seem. These (inedible) desserts are made from coal ash.



**K** Project We Love Installations Brooklyn, NY

**\$4,700**  
pledged of \$3,750 goal

**41**  
backers

**11**  
days to go

Back this project

Remind me

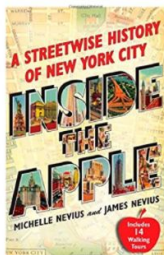
Share

All or nothing. This project will only be funded if it reaches its goal by Tue, September 26 2017 12:31 PM EDT.

# Collaborative Filter

## Customers who bought this item also bought

Page 1 of 20



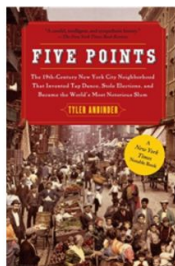
Inside the Apple: A Streetwise History of New York City

› Michelle Nevius

★★★★☆ 66

Paperback

\$11.80 ✓prime



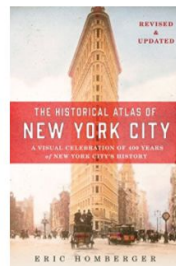
Five Points: The 19th Century New York City Neighborhood that...

› Tyler Anbinder

★★★★☆ 94

Paperback

\$10.08 ✓prime



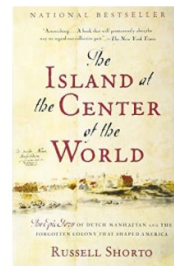
The Historical Atlas of New York City, Third Edition: A Visual Celebration of...

Eric Homberger

★★★★★ 4

Paperback

\$20.39 ✓prime



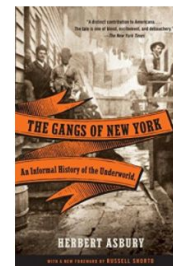
The Island at the Center of the World: The Epic Story of Dutch Manhattan and...

› Russell Shorto

★★★★☆ 461

Paperback

\$12.97 ✓prime



The Gangs of New York: An Informal History of the Underworld

› Herbert Asbury

★★★★☆ 99

Paperback

\$10.78 ✓prime

```

1 class RecommendationsController < ApplicationController
2
3   def index
4     @root_project = Project.find(params[:id])
5
6     @start = Time.now.to_f
7     project_ids = ActiveRecord::Base.connection.select_all <<-SQL
8       select sum(b1.amount * b2.amount) / s1.pledged_magnitude / s2.pledged_magnitude as factor, p2.id
9       from p as p1 join p as p2
10        join b as b1 on p1.id = b1.pid
11        join b as b2 on p2.id = b2.pid
12        join project_stats as s1 on p1.id = s1.pid
13        join project_stats as s2 on p2.id = s2.pid
14        where p1.id = #{@root_project.id}
15        and b1.bid = b2.bid
16        group by p2.id
17        order by factor desc
18        limit 20
19     SQL
20     @ending = Time.now.to_f
21
22     factors_by_id = {}
23     project_ids.each{|p| factors_by_id[p['id']] = p['factor'] }
24
25     @projects = Project.where(:id => factors_by_id.keys)
26     @projects = @projects.sort do |a,b|
27       factors_by_id[a.id] <=> factors_by_id[b.id]
28     end.reverse
29
30     render :layout => 'frame'
31   end
32 end

```

```
1 class RecommendationsController < ApplicationController
```

```
2  
3 def index
```

```
4   @root_project = Project.find(params[:id])
```

```
5  
6   @start = Time.now.to_f
```

```
7   project_ids = ActiveRecord::Base.connection.select_all <<-SQL
```

```
8     select sum(b1.amount * b2.amount) / s1.pledged_magnitude / s2.pledged_magnitude as factor, p2.id  
9     from p as p1 join p as p2  
10    join b as b1 on p1.id = b1.pid  
11    join b as b2 on p2.id = b2.pid  
12    join project_stats as s1 on p1.id = s1.pid  
13    join project_stats as s2 on p2.id = s2.pid  
14    where p1.id = #{@root_project.id}  
15    and b1.bid = b2.bid  
16    group by p2.id  
17    order by factor desc  
18    limit 20
```

```
19   SQL
```

```
20   @ending = Time.now.to_f
```

```
21  
22   factors_by_id = {}
```

```
23   project_ids.each{|p| factors_by_id[p['id']] = p['factor'] }
```

```
24  
25   @projects = Project.where(:id => factors_by_id.keys)
```

```
26   @projects = @projects.sort do |a,b|
```

```
27     factors_by_id[a.id] <=> factors_by_id[b.id]
```

```
28   end.reverse
```

```
29  
30   render :layout => 'frame'
```

PROBLEM #1

Building at production scale



Source projects	Target projects	Similarity scores
123456	322245	0.67
123456	345334	0.45
123456	912345	0.23
123456	932837	0.43
123456	339992	0.12
339992	273849	0.83
144453	675950	0.04

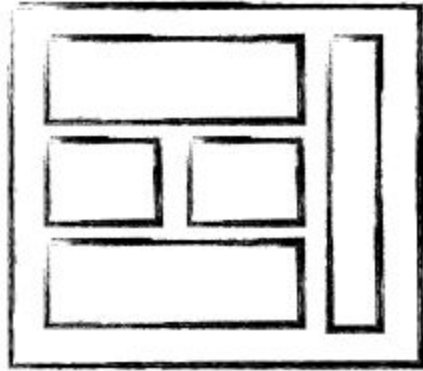




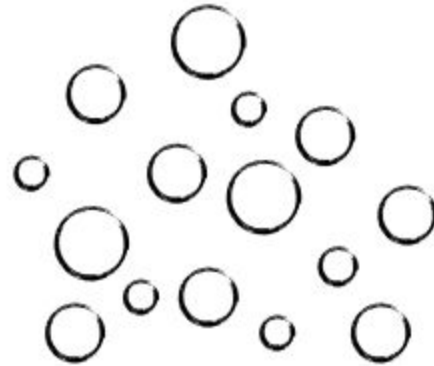


## PROBLEM #2

Escaping a local optimum



MONOLITHIC/LAYERED



MICRO SERVICES

## PROBLEM #3

Knowing when worse is good enough





## PROBLEM #4

Fixing a cold start



# Latent Semantic Indexing (LSI)

	apple	brass	car	dust	either	fox	groovy
document 1	1	1	0	0	1	0	1
document 2	1	0	0	1	0	1	0
document 3	0	0	0	0	0	0	0
document 4	1	1	1	0	0	0	0
document 5	0	0	1	1	1	0	0
document 6	0	1	0	1	0	0	1
document 7	0	1	0	0	0	0	1
document 8	0	0	0	0	1	0	0



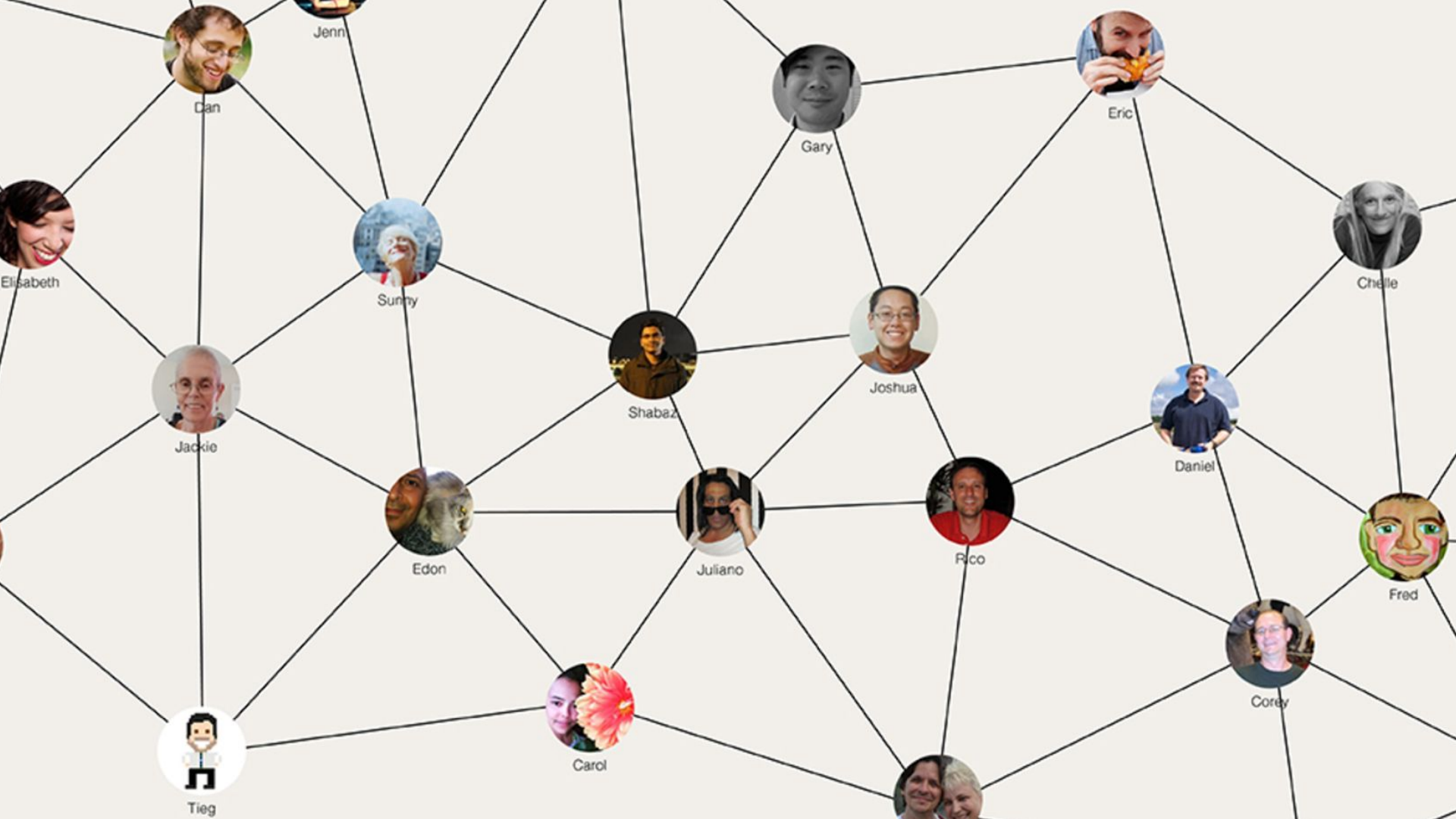
**LSI**

	apple	brass	car	dust	either	fox	groovy
document 1	1	1	0	0	1	0	1
document 2	1	0	0	1	0	1	0
document 3	0	0	0	0	0	0	0
document 4	1	1	1	0	0	0	0
document 5	0	0	1	1	1	0	0
document 6	0	1	0	1	0	0	1
document 7	0	1	0	0	0	0	1
document 8	0	0	0	0	1	0	0

**LSI**

**SVD**

document 1	0.53	-0.97
document 2	0.22	-1.21
document 3	1.54	0.65
document 4	-0.31	1.43
document 5	-1.22	-0.03
document 6	0.19	1.45
document 7	0.67	1.03
document 8	-0.34	-0.62



## PROBLEM #5

Blending it all together





Looking forward



Looking back

# Looking back

## 1 Clarify your goals

## Looking back

- 1 Clarify your goals
- 2 Keep it simple

## Looking back

- 1 Clarify your goals
- 2 Keep it simple
- 3 Be nimble

## Looking back

- 1 Clarify your goals
- 2 Keep it simple
- 3 Be nimble
- 4 Build partnerships

# Data Infrastructure Engineer

Brooklyn, NY

Kickstarter is looking for an experienced data engineer to help us scale up our data infrastructure.

## About the team

You'll be joining Kickstarter's Data team, a collaborative group of people who are passionate about using data to help bring creative projects to life. We value peer feedback, continual learning, and blameless post-mortems. We believe in democratizing data access across the organization and providing tools and training to support data-informed decisions. We prefer proven technology and have built most of our data infrastructure on AWS, but we also care about finding the right tools for the job, so we invest in cutting-edge technology when it's a good fit.

## About the role

In this role you will:

- Build data pipelines in AWS to support machine learning and analytics
- Optimize the performance of databases like Redshift and file systems like S3
- Design architecture for distributed batch and stream processing, creating systems that will empower your teammates to write their own custom ETL jobs
- Collaborate with DevOps engineers on AWS platform improvements
- Help product teams grow a suite of data microservices like spam classification and recommendations
- Work with others on the Data team to support analytics and to ship models to production

What's it like to work at  
Kickstarter?

[Come find out](#) →

## Current openings

[Brand Photographer-Videographer](#)

[Brand Producer-Editor](#)

[Front End Engineer](#)

[iOS Engineer](#)

[Engineering Manager, Native Apps](#)

[Full Stack Engineer](#)

[Marketing Associate](#)

[iOS Engineer](#)

[Product Copywriter](#)

**Data Infrastructure Engineer**

[Senior Operations Engineer](#)

[Games Content Strategist](#)

[Operations Engineer](#)

[Talent Acquisition Lead](#)

[Product Designer](#)

[Senior Back End Engineer](#)

[www.kickstarter.com/jobs/data-engineer](https://www.kickstarter.com/jobs/data-engineer)

# Thank you

[jeremy@kickstarter.com](mailto:jeremy@kickstarter.com)  
[twitter.com/jsalfen](https://twitter.com/jsalfen)