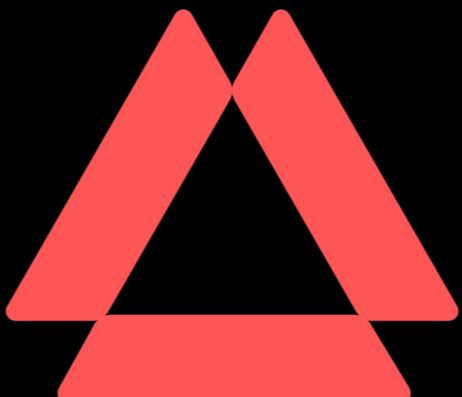


Machine Learning for Relevance and Serendipity

Jenny Rose Finkel
Prismatic



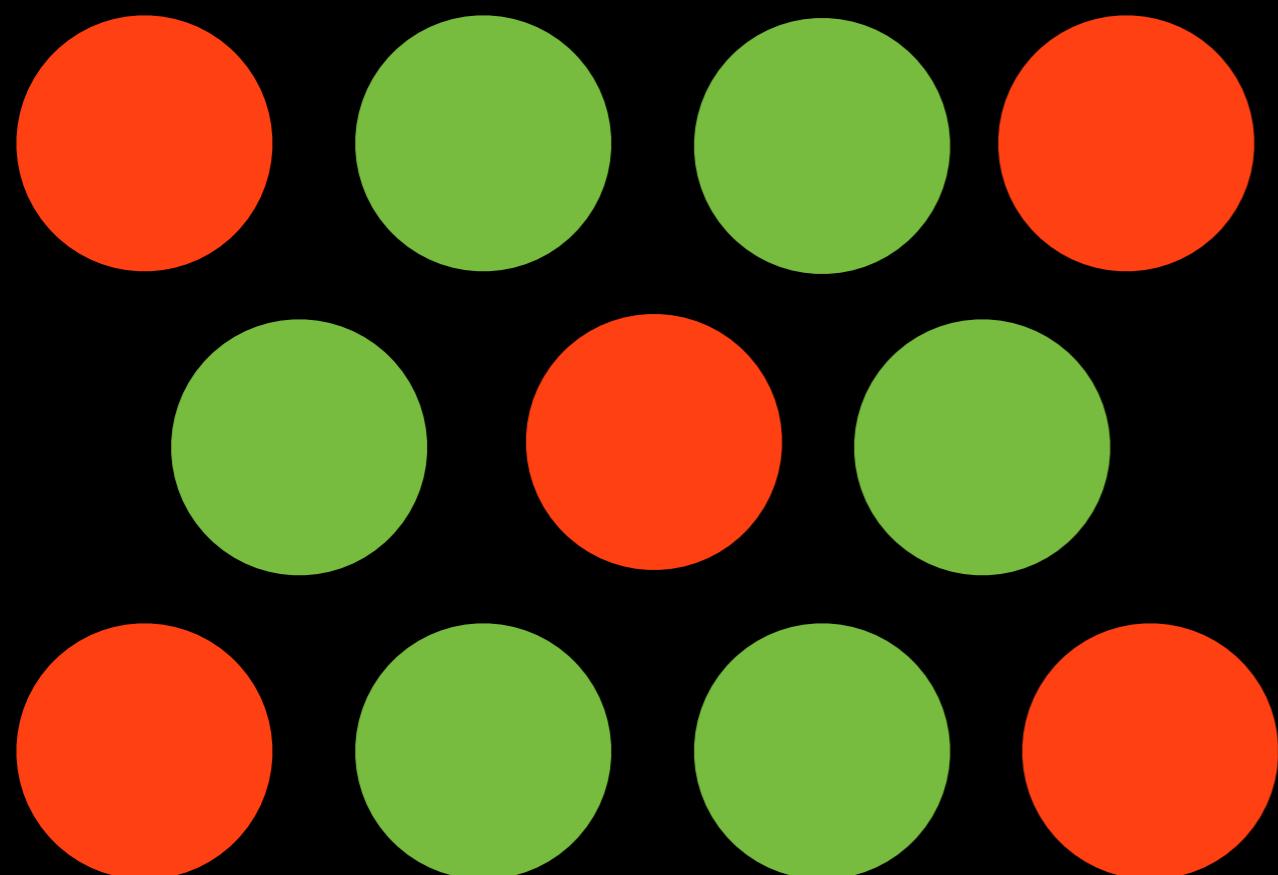
Prismatic

- We're very functional
 - clojure on the backend
 - clojurescript on the (web) front end
- We're very open-source
 - schema, plumbing, graph, hiphip, dommy
- This talk isn't about any of that

The Discovery Problem

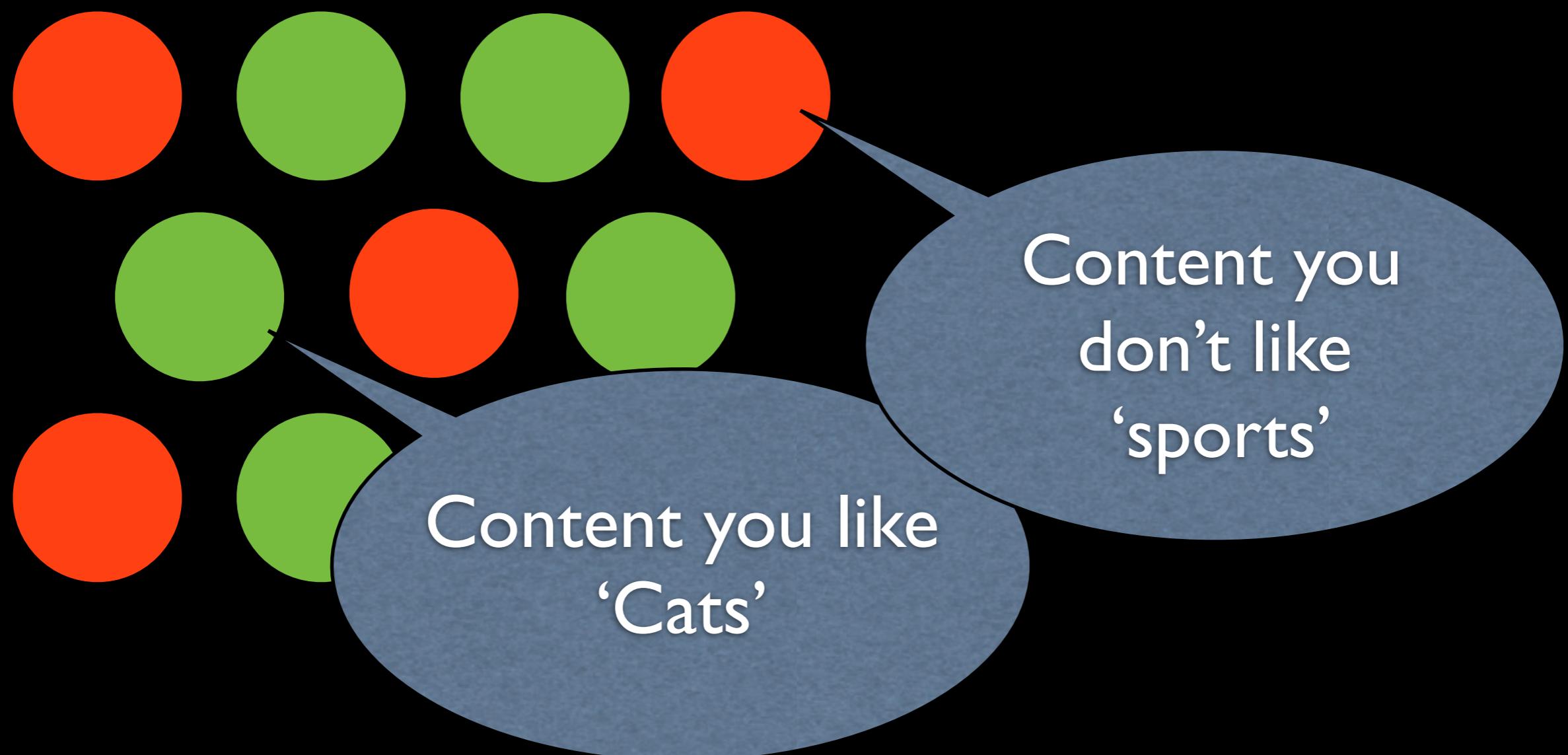
The Discovery Problem

Content You See



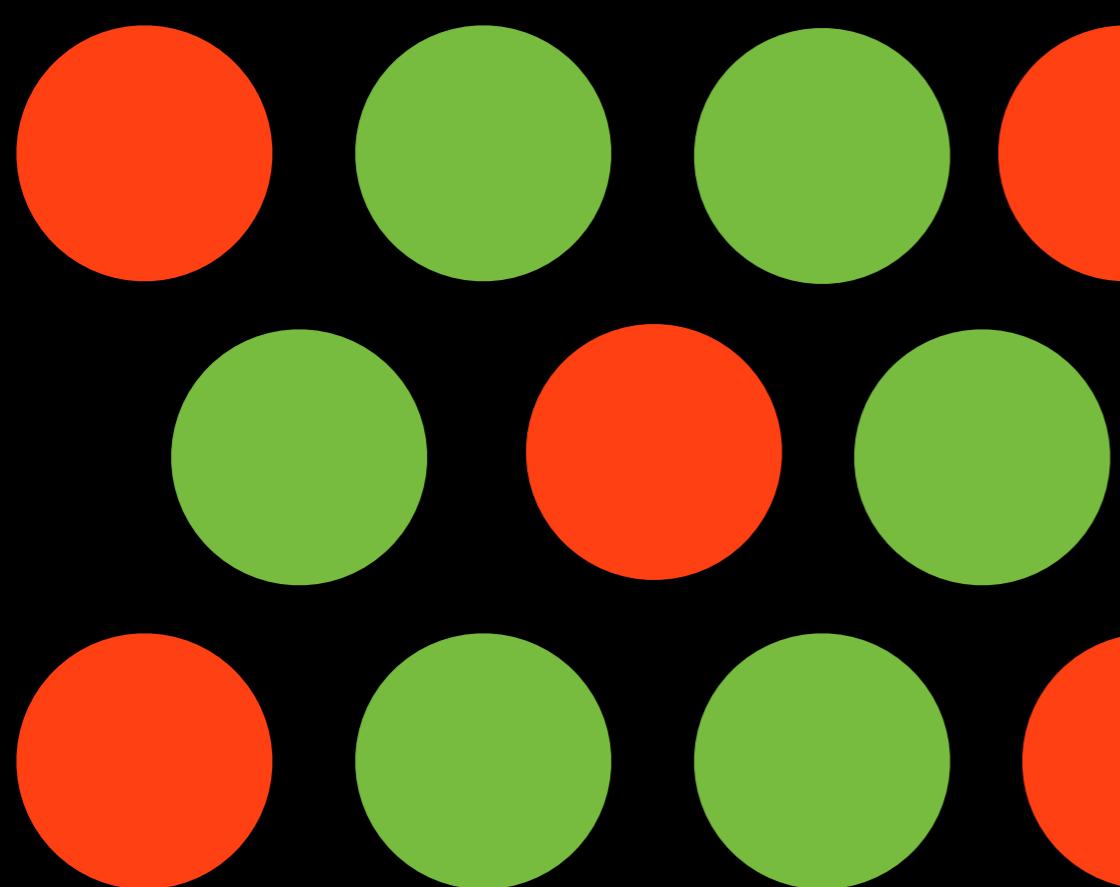
The Discovery Problem

Content You See

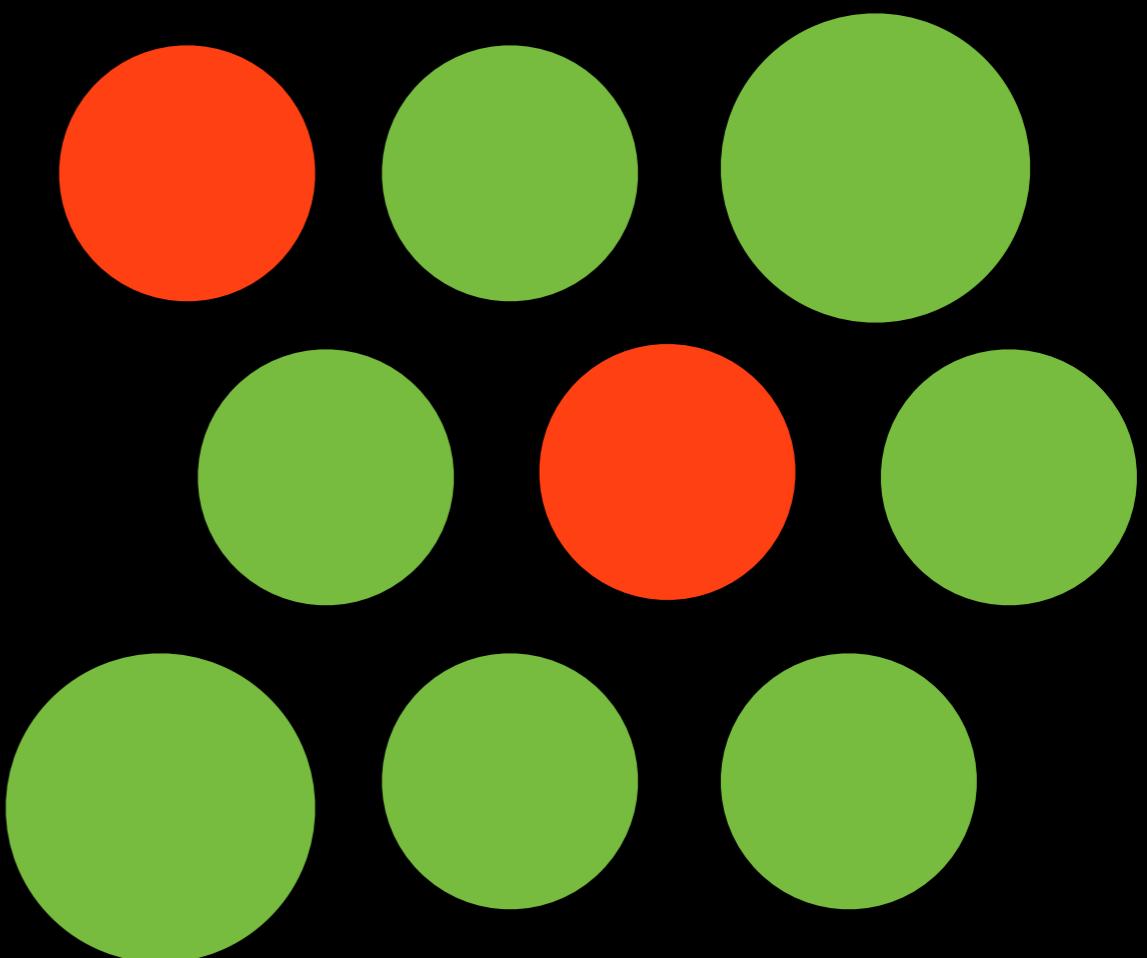


The Discovery Problem

Content You See

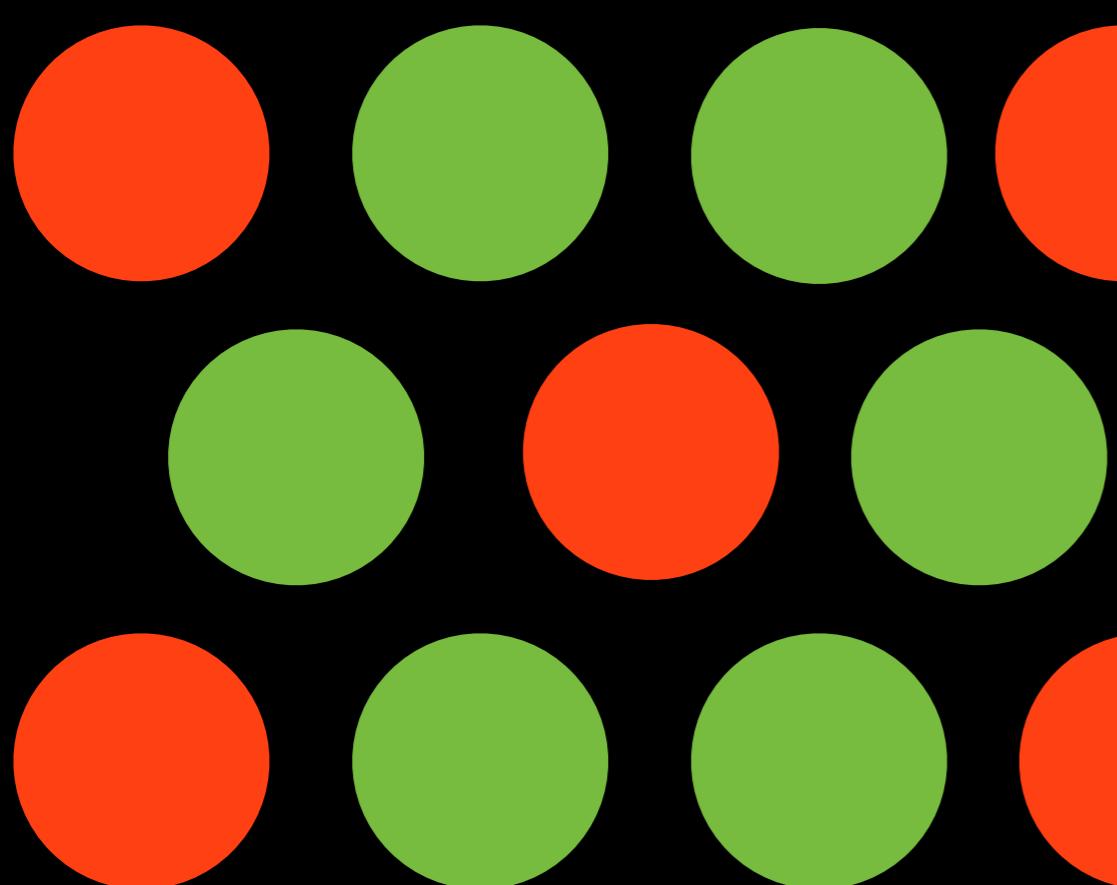


Content You Miss

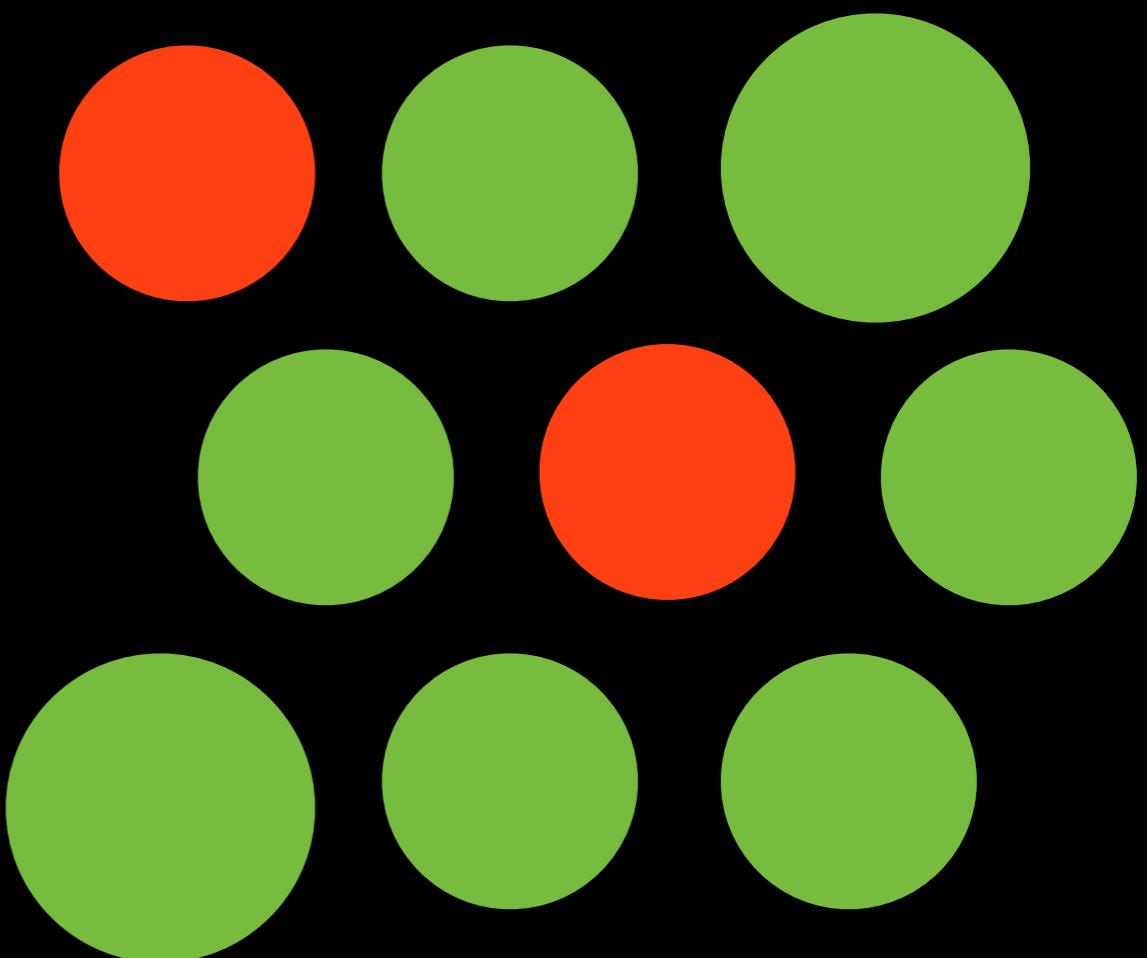


The Discovery Problem

Content You See



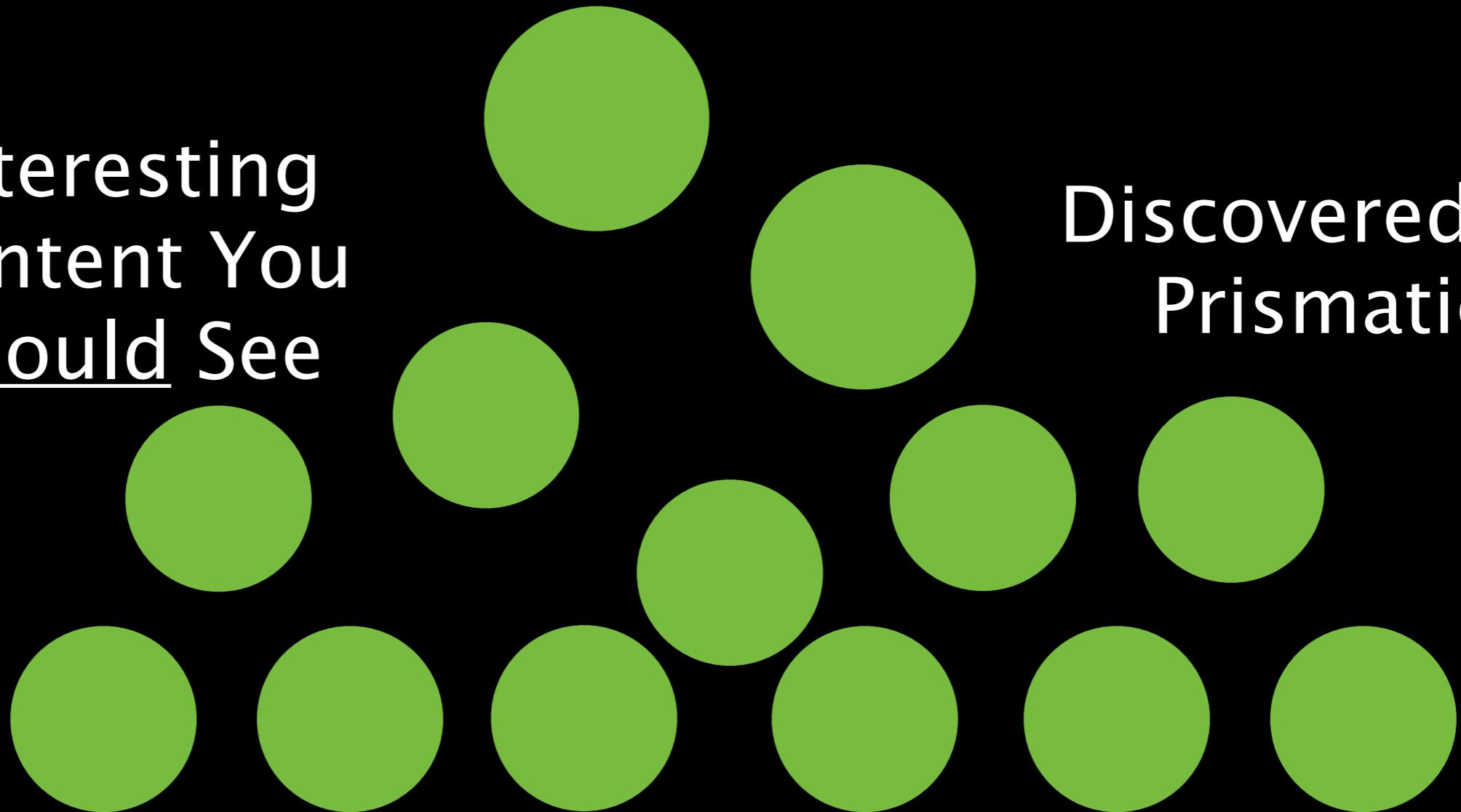
Content You Miss



The Discovery Problem

Interesting
Content You
Should See

Discovered by
Prismatic



Obese Cat Dives Into Workout Plan

 KTLA 5 by KTLA 5 Web Staff yesterday Cats Pets

While most cats hate water, a fat feline in Tennessee has taken the plunge to lose weight.

The obese feline named Buddha was caught on tape walking on a partially submerged treadmill to slim down.

Buddha was recorded walking on a treadmill. You may remember we first brought you Buddha's story in August. Since then he's become an international sensation.



 1  0 



Colorado: Hundreds Line Up For Free Joints In Marijuana Tax Protest

 Hemp News by Steve Elliott 16h ago Marijuana and Medical Marijuana Sales Tax

Hundreds of excited people lined up in Denver's Civic Center Park on Monday to get a free joint, as part of a protest against Colorado's plan to heavily tax recreational marijuana.

The protest was courtesy of the No On Proposition AA campaign, which opposes a plan calling for a 10 percent sales tax on recreational cannabis with the option of...



 1  0 



Functional Programming

 Clojure for the Brave and True 5h ago Functional Programming Lisp

If I were Mr. Miyagi and you were Daniel-san, the last chapter would be the equivalent of "wax on, wax off." Also, why haven't you cleaned my car?!? In this chapter, you'll begin to take your concrete experience with functions and data structures and integrate it in a new mindset, the

functional programming mindset. By the end of this chapter, you'll have learned: What pure functions are and why they're important; Why immutability...

 7  0 



Smitten Ice Cream to open new location within Los Altos Whole Foods

 Inside Scoop SF by Paolo Lucchesi 21h ago San Francisco Food Ice Cream Zoning



Give Smitten Ice Cream credit for opening up in unexpected location. The San Francisco ice cream shop known for its liquid nitrogen preparations opened in Hayes Valley's experimental Proxy project, an urban development that has been a huge success, to say the least.

Earlier this summer, Scoop reported on Smitten's next two projects, new sequels in Oakland and Los Altos, both due this fall. At the time, Smitten kept the exact locations under wraps. Today comes word of the Los Altos address. And in a fun little twist,...

 1  0 



Diddy Had A Life-changing Experience At Burning Man

 All | Complex.com by Kyle Kramer 6d ago Burning Man Festivals Camps and Camping



Burning Man, the renowned anti-capitalist art festival/utopian music and nudity experiment that takes place in the Nevada desert each summer, is a popular destination for people from all walks of life looking to come together and do stuff like scrape the dirt off of each others' naked bodies.

And, as anyone who's ever been cornered at a party by some New Age zealot or faux-bohemian tech entrepreneur knows, no matter who you are, you will invariably come back from your Burning Man...

 3  1 



feedback logout

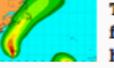
Search...



World News

 Bitcoin ATMs Arriving in Canada
 Mashable 5h ago

 Tropical Storm HUMBERTO Forecast Discussion
 National Hurricane Center 15h ago

 Tropical Storm Gabrielle reforms; Humberto to become hurricane
 CBS News 5h ago

 Syria accepts Russian chemical weapons proposal - Interfax
 Yahoo! News 6h ago

 Pakistan 'to free Taliban's Mullah Abdul Ghani Baradar'
 BBC News - Asia 6h ago

 Spurred By 'Puppy Mill' Report, Feds Crack Down On Dog Breeders Who Sell Animals Online
 The Blog 3h ago

Suggestions

 TechCrunch

 Shaun McInnis

 Social Media

 4EverFascinated

 Startups

 James Yu

 The New York Times

 Chris Curtis

 Twitter

 Anne Bell Jacks

 Gawker

 belvidge

 Technology

How we do it

How we do it

from unguarded warheads in Hawaii
bit.ly/10XbCUI via @instapundit

Expand

Farhad Manjoo @fmanjoo 5m
@cdixon Did you read @felixsalmon's post?
View conversation

chris dixon @cdixon 6m
I'd love to read a blog post where someone without an econ background explains why a fixed currency base will lead to deflation.
Expand

Peter Pham @peterpham 1h
The one and only @BenjLerer visiting @ScienceInc welcoming @jasonrapp to his first week! pic.twitter.com/GAEdSJm5Bi
Retweeted by Ben Lerer
View photo

Chris Ziegler @zpower 8m
I mean the second side of 'wider angle' is one of the sickest live sets I've ever heard. It is just mental.
Expand

Galen Ward @galenward 8m
Absolutely astounding: 3 years ago there was no tablet market - the iPad was 2 weeks old - and this year 200 million will be sold.
Expand

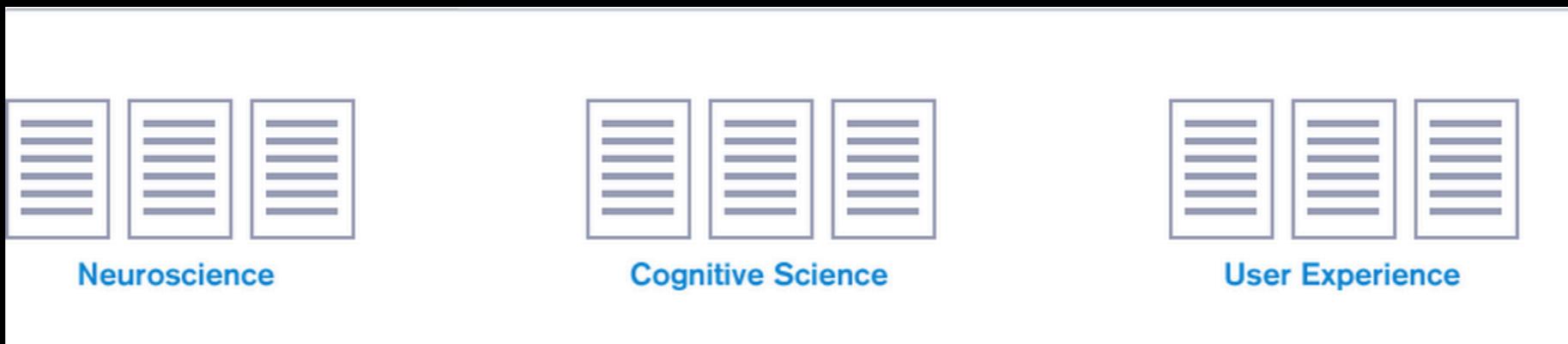
Lena Shaw @LenaShaw 52m
Help @Techcrunch Kick @Path & @davemorin Butt (And Support Teach For America) [@TFAbayarea](http://techcrunch.com/2013/04/10/help-tfa-bayarea) @davemorin #tfachallenge
Retweeted by Dave Morin
View photo

joshua schachter @joshu 8m
@fromedome especially the xiao long bao at DTF. Email me if you need further details



Crawl the social and
open web for
engaging content

How we do it



Analyze and categorize content
with hundreds of thousands of
fine-grained topics

How we do it

The screenshot shows a news feed interface with a header bar. The top bar has a house icon, the word "Home", and statistics: 242 interests, 112 following, and 72 viewed. Below this, the first article is titled "LEAKED MEMO: The New York Times Replaces Its Tech Editor Two Months After He Got The Job" by "SAI" (Nicholas Carlson) 8h ago. It includes tags for Newspapers, New York Times, and Time Inc. The text states that only two months after the new tech editor got the job, he has already replaced the old one. The second article is titled "'Secretbook' Lets You Encode Hidden Messages in Your Facebook Pics" by "Danger Room" (Robert Beckhusen) 17h ago. It includes tags for Social Media and Computer Vision. It discusses a browser extension that encodes secret messages into Facebook pictures. The third article is titled "With Thuzio, You Can Pay \$2,500 to Play Pick-Up With Antoine Walker".

LEAKED MEMO: The New York Times Replaces Its Tech Editor Two Months After He Got The Job
SAI by Nicholas Carlson 8h ago Newspapers New York Times Time Inc.
Only two months after he got the gig, the new New York Times tech editor is already the old New York Times editor. On January 16, the Times replaced tech editor Damon Darlin with assistant managing editor Glenn Kramon, who would move from New York to San Francisco...

'Secretbook' Lets You Encode Hidden Messages in Your Facebook Pics
Danger Room by Robert Beckhusen 17h ago Social Media Computer Vision
 Facebook is a place where you can share pictures of cute animals and fun activities. Now there's a browser extension that lets you encode those images with secret, hard-to-detect messages.

With Thuzio, You Can Pay \$2,500 to Play Pick-Up With Antoine Walker

Real-time personalization
based on user interests

Goal: Build a ranker to find
a small number of great
articles for a user from a
large document index.

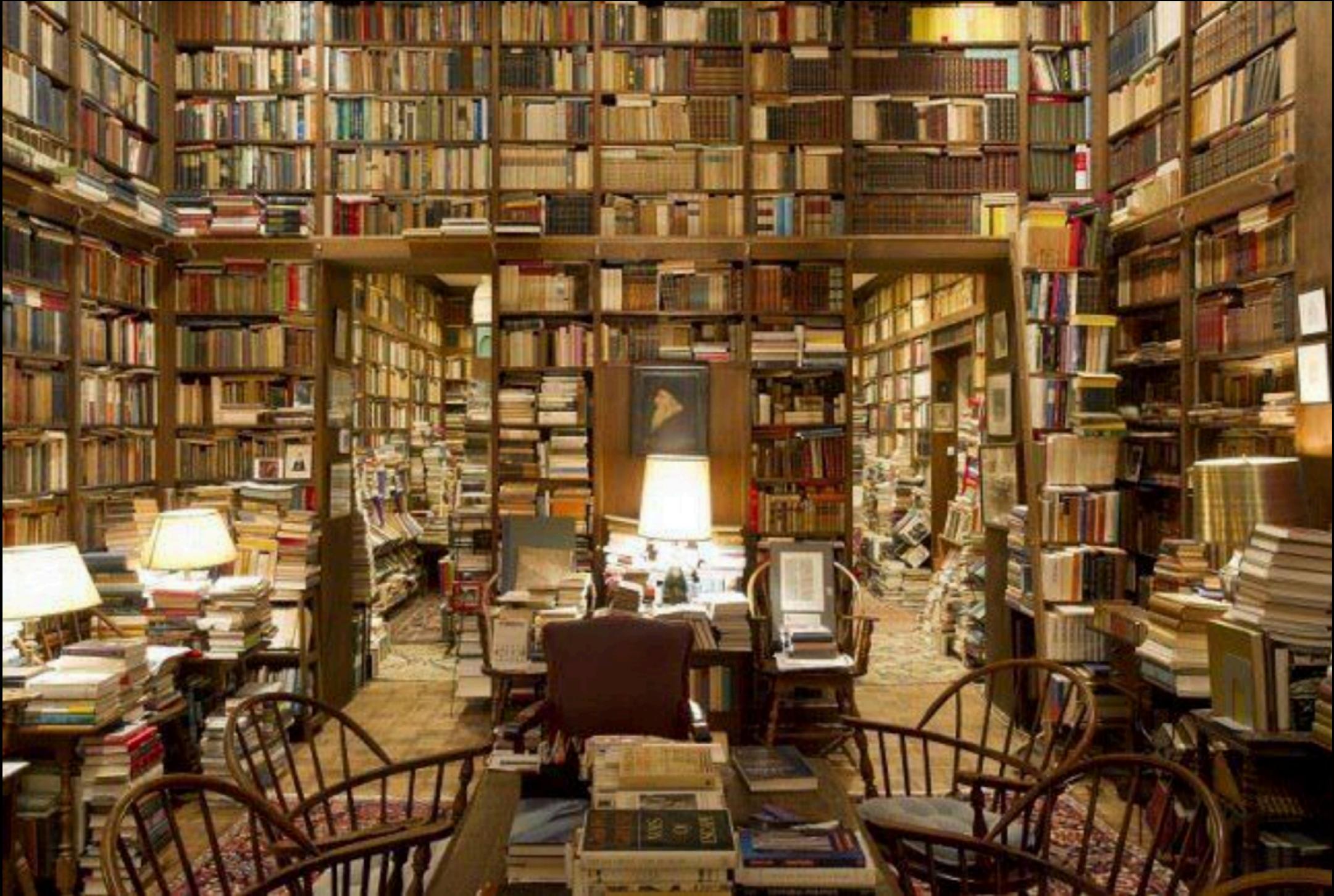
“People who use the term ‘secret sauce’ know about neither food nor technology”

- *Aria Haghghi*

Talk Outline

- What do we have to work with?
- What are some of our constraints?
- How do we know if we did a good job?
- How is our relevance system architected?
- What stumbling blocks did we encounter?
- How did we do?

Our Data



~1 million timely articles in our index

5 Ancient Cars We Can't Believe Are Still Sold Today

by Car Throttle by Matt Robinson 13h ago Cars Automobile Design



When you think about it, the average shelf life for most cars isn't all that long. A new car will have barely been in the showrooms before the manufacturer begins feverishly cooking up its eventual replacement.

That's not the case for every vehicle, however. For one reason or another, some car designs just refuse to die, and despite being pretty damned long in the tooth, you can still buy them brand new today.

2 Q0



5 Ancient Cars We Can't Believe Are Still Sold Today

Car Throttle

by Matt Robinson 13h ago

Cars

Automobile Design



When you think about it, the average shelf life for most cars isn't all that long. A new car will have barely been in the showrooms before the manufacturer begins feverishly cooking up its eventual replacement.

That's not the case for every vehicle, however. For one reason or another, some car designs just refuse to die, and despite being pretty damned long in the tooth, you can still buy them brand new today.

2 Q0



the publisher

5 Ancient Cars We Can't Believe Are Still Sold Today

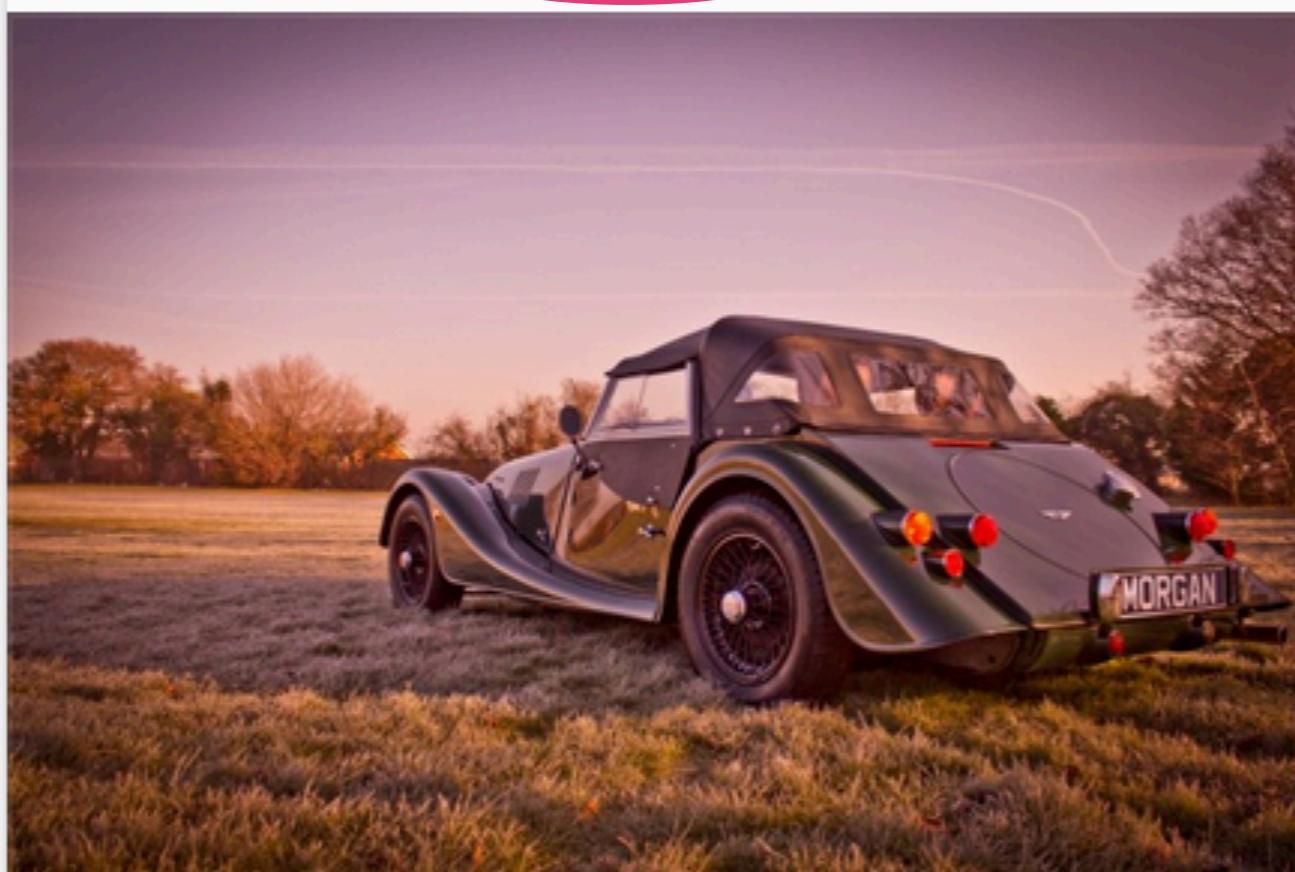
Car Throttle

by Matt Robinson

13h ago

Cars

Automobile Design



When you think about it, the average shelf life for most cars isn't all that long. A new car will have barely been in the showrooms before the manufacturer begins feverishly cooking up its eventual replacement.

That's not the case for every vehicle, however. For one reason or another, some car designs just refuse to die, and despite being pretty damned long in the tooth, you can still buy them brand new today.

2 Q0



the author

5 Ancient Cars We Can't Believe Are Still Sold Today

Car Throttle by Matt Robinson 13h ago Cars Automobile Design



When you think about it, the average shelf life for most cars isn't all that long. A new car will have barely been in the showrooms before the manufacturer begins feverishly cooking up its eventual replacement.

That's not the case for every vehicle, however. For one reason or another, some car designs just refuse to die, and despite being pretty damned long in the tooth, you can still buy them brand new today.

2 Q0



the date

5 Ancient Cars We Can't Believe Are Still Sold Today

Car Throttle by Matt Robinson 13h ago **Cars** Automobile Design



When you think about it, the average shelf life for most cars isn't all that long. A new car will have barely been in the showrooms before the manufacturer begins feverishly cooking up its eventual replacement.

That's not the case for every vehicle, however. For one reason or another, some car designs just refuse to die, and despite being pretty damned long in the tooth, you can still buy them brand new today.

2 Q0



topics

Our Topic Classifier

- Can very accurately classify documents into over 5000 topics (soon will be 50x that), and give a degree of confidence
- Built entirely in-house
- No time to explain how it works in this talk
- Critical for exploring new interests

5 Ancient Cars We Can't Believe Are Still Sold Today

by Car Throttle by Matt Robinson 13h ago Cars Automobile Design



When you think about it, the average shelf life for most cars isn't all that long. A new car will have barely been in the showrooms before the manufacturer begins feverishly cooking up its eventual replacement.

That's not the case for every vehicle, however. For one reason or another, some car designs just refuse to die, and despite being pretty damned long in the tooth, you can still buy them brand new today.

2 Q 0



↓ ×

social signals from both internal and external networks

5 Ancient Cars We Can't Believe Are Still Sold Today

Car Throttle by Matt Robinson 13h ago Cars Automobile Design



When you think about it, the average shelf life for most cars isn't all that long. A new car will have barely been in the showrooms before the manufacturer begins feverishly cooking up its eventual replacement.

That's not the case for every vehicle, however. For one reason or another, some car designs just refuse to die, and despite being pretty damned long in the tooth, you can still buy them brand new today.

2 Q0



actions on prismatic: click, like, save, share, remove

5 Ancient Cars We Can't Believe Are Still Sold Today

Car Throttle by Matt Robinson 13h ago Cars Automobile Design



When you think about it, the average shelf life for most cars isn't all that long. A new car will have barely been in the showrooms before the manufacturer begins feverishly cooking up its eventual replacement.

That's not the case for every vehicle, however. For one reason or another, some car designs just refuse to die, and despite being pretty damned long in the tooth, you can still buy them brand new today.

2 Q0



the title

5 Ancient Cars We Can't Believe Are Still Sold Today

Car Throttle by Matt Robinson 13h ago Cars Automobile Design



When you think about it, the average shelf life for most cars isn't all that long. A new car will have barely been in the showrooms before the manufacturer begins feverishly cooking up its eventual replacement.

That's not the case for every vehicle, however. For one reason or another, some car designs just refuse to die, and despite being pretty damned long in the tooth, you can still buy them brand new today.

2 Q 0



the text

	jrfinkel 		61 interests	230 following	101 followers
	invite friends	settings			
	Ryan Mason				
	adam				
	Matt Johnson				
	Ben Seck				
	Shane John David				
	Greg Rahn				
	Josh Persky				
	Petra Cross				
	Dan Kaminsky				
	mike ihbe				
	Jackie Bavaro				
	Dave Bogaty				
	Evan Hindman				
	Stanford NLP Group				
	Sara Elaine Stephain				
	Brendan Wypich				
	Philip yell				
	Diana Dinh				
	Anne Legg				
	Abel Allison				
	Tim Bavaro				
	Josh Crowder				
	John Stockdale				
	Nundu Janakiram				
	Joshua Gourneau				
	Brian House				
	Timothy Finne				
	evan rosen				
	Tristan Harris				
	Kiira Mancasola				
	Ian Ward Comfort				
	Bob Smith				
	lizneiman				
	Erica Ford				
	tanya				
	chris wiggins				



jrfinkel

invite friends

settings

58
interests

230
following

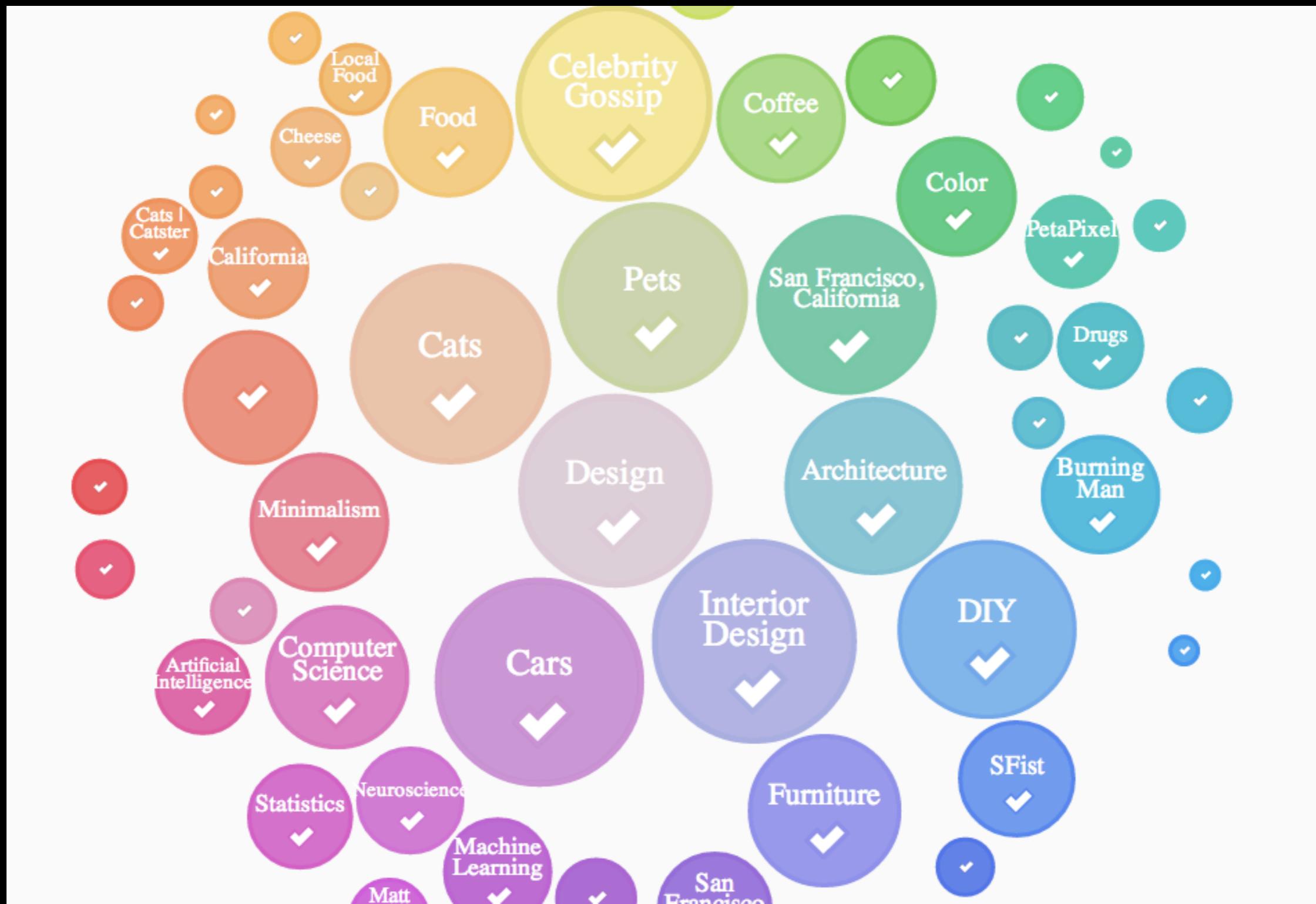
101
followers



- Cars
- Design
- Pets
- Architecture
- San Francisco, California
- Computer Science
- Coffee
- Color
- San Francisco
- Marijuana and Medical Mari...
- Neuroscience
- Artificial Intelligence
- Burning Man
- Drugs

- Barack Obama
- Interior Design
- Celebrity Gossip
- Cats
- DIY
- Furniture
- Minimalism
- Food
- SFist
- Happiness
- Machine Learning
- Statistics
- California
-  PetaPixel

Interest Scores



Saved Session Data

- We snapshot every session
 - the articles shown, and their order
 - what actions the user took
 - the user's state at the beginning of the session
 - favorites
 - interest scores

Our Constraints



so many documents



so little (computation) time

What I want to be reading

BBC News Sport Weather Ca

NEWS MIDDLE EAST

Home | US & Canada | Latin America | UK | Africa | Asia | Europe | Mid-East | Bu

16 September 2013 Last updated at 15:26 ET

Q&A: Syria chemical weapons disarmament deal



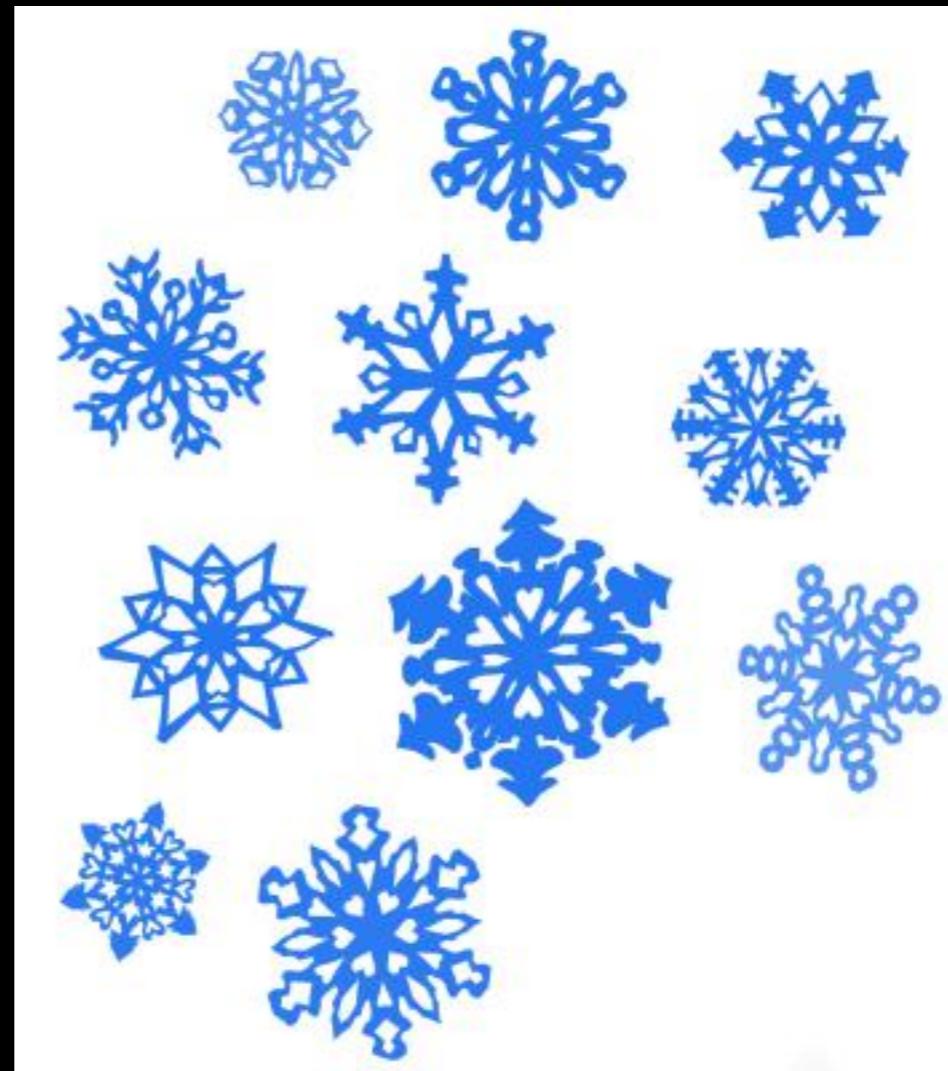
John Kerry and Sergei Lavrov reached an agreement after three days of talks

The United States and Russia agreed on 14 September a framework for the removal and destruction of Syria's arsenal of chemical weapons. The deal indefinitely stalled the prospect of US-led

What I am reading



You are all unique snowflakes



Evaluation

How do we know if we've done a good job?

- **Real Goal:** Improved user engagement
- It takes a long time to change user behavior
 - It takes even longer to confidently measure changes in user behavior
- Can't do machine learning on it

How do we know if we've done a good job?

- **Surrogate Goal:** More user interactions
- Easy to measure from saved sessions
- Divide data into *training data* and *test data*
 - Build the model using the training data
 - Use the model to re-order the articles in the sessions from the test data
 - Evaluate on the re-ordered articles

Architecture

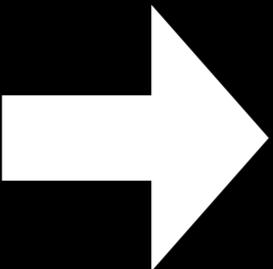
Formula for a Machine Learning Problem

- Step #1: Collect Data
- Step #2: “Featurize” the data
- Step #3: Do numeric optimization on objective function of featurized data to learn feature weights
- Step #4: Use feature weights to classify new data

Step I: Collect Data

- Save session information (described earlier)
 - docs viewed, in order, and user's actions
 - user's favorites & interaction scores
- Transform data into a set of article-user pairs
 - *Positive examples* are when a user clicks, shares, or saves the article
 - *Negative examples* are when a user Xs out the article (scrolling past an article is a weak negative example)

session
log



...

Step 2: Featurize Data

Represent user-article pairs as a vector of numbers

<i>feature</i>	<i>feature value</i>		
article has a topic that the user follows?	1		
# times user previously clicked on this publisher	0		
# times user previously removed this publisher	2		
article has a picture?	1		
# of prismatic users who have liked this article	2		
...	...		

Step 2: Featurize Data

Represent user-article pairs as a vector of numbers

<i>feature</i>	<i>feature value</i>
article has a topic that the user follows?	1
# times user previously clicked on this publisher	0
# times user previously removed this publisher	2
article has a picture?	1
# of prismatic users who have liked this article	2
...	...

Step 2: Featurize Data

How likely is *this user* to like *this article*?

<i>feature</i>	<i>feature value</i>	<i>feature weight</i>	
article has a topic that the user follows?	1	+1.2	
# times user previously clicked on this publisher	0	+1.5	
# times user previously removed this publisher	2	-3.4	
article has a picture?	1	+0.5	
# of prismatic users who have liked this article	2	+0.8	
...	

Step 2: Featurize Data

How likely is *this user* to like *this article*?

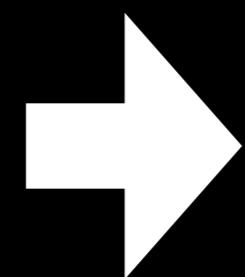
<i>feature</i>	<i>feature value</i>	<i>feature weight</i>	<i>product</i>
article has a topic that the user follows?	1	+1.2	1.2
# times user previously clicked on this publisher	0	+1.5	0
# times user previously removed this publisher	2	-3.4	-6.8
article has a picture?	1	+0.5	0.5
# of prismatic users who have liked this article	2	+0.8	1.6
...
<i>total score (dot-product of features and weights)</i>			-3.5

Step 2: Featurize Data

How likely is *this user* to like *this article*?

<i>feature</i>	<i>feature value</i>	<i>feature weight</i>	<i>product</i>
article has a topic that the user follows?	1	?	1.2
# times user previously clicked on this publisher	0	?	0
# times user previously removed this publisher	2	?	-6.8
article has a picture?	1	?	0.5
# of prismatic users who have liked this article	2	?	1.6
...
<i>total score (dot-product of features and weights)</i>			-3.5

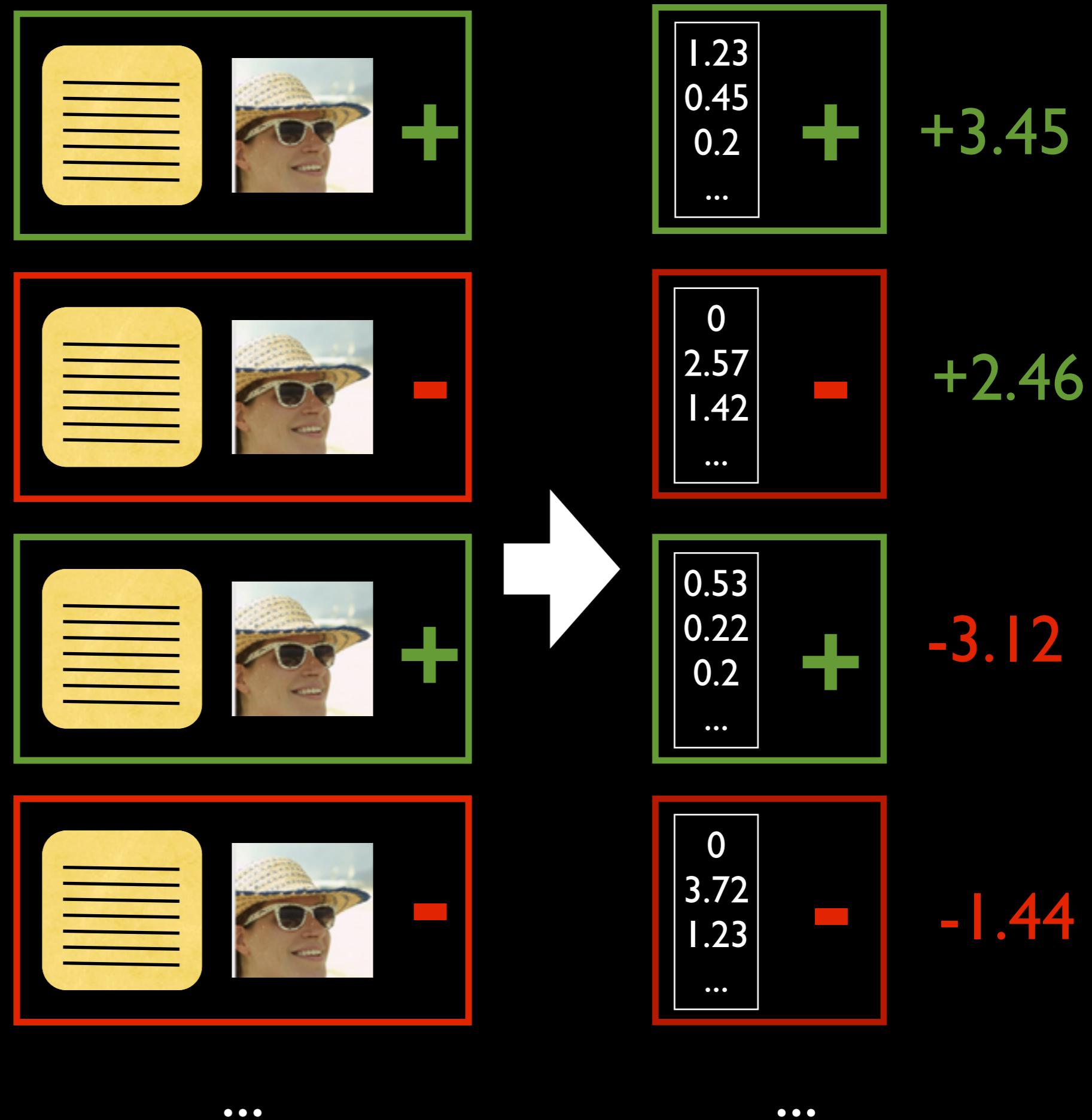
session log



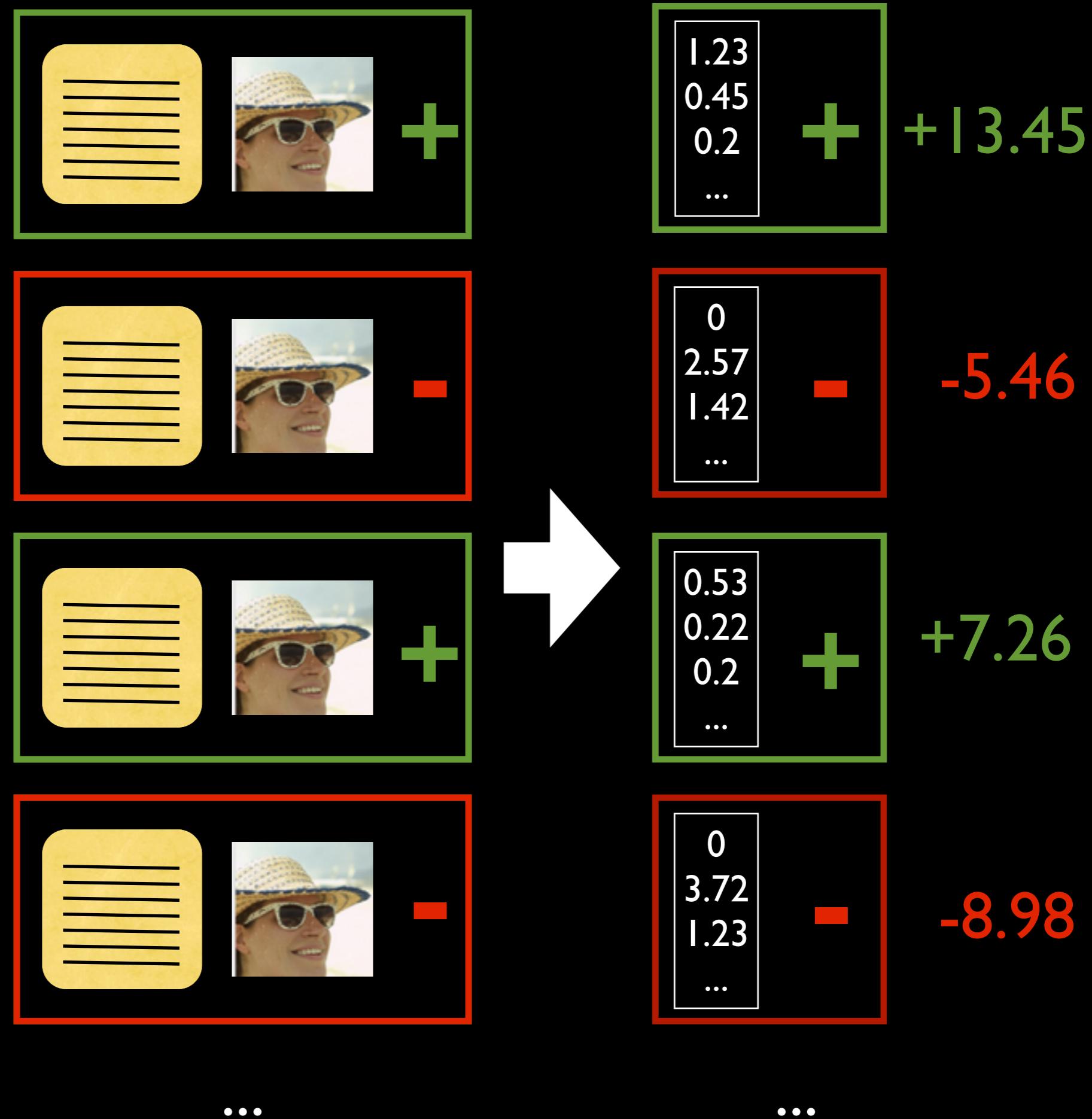
...

...

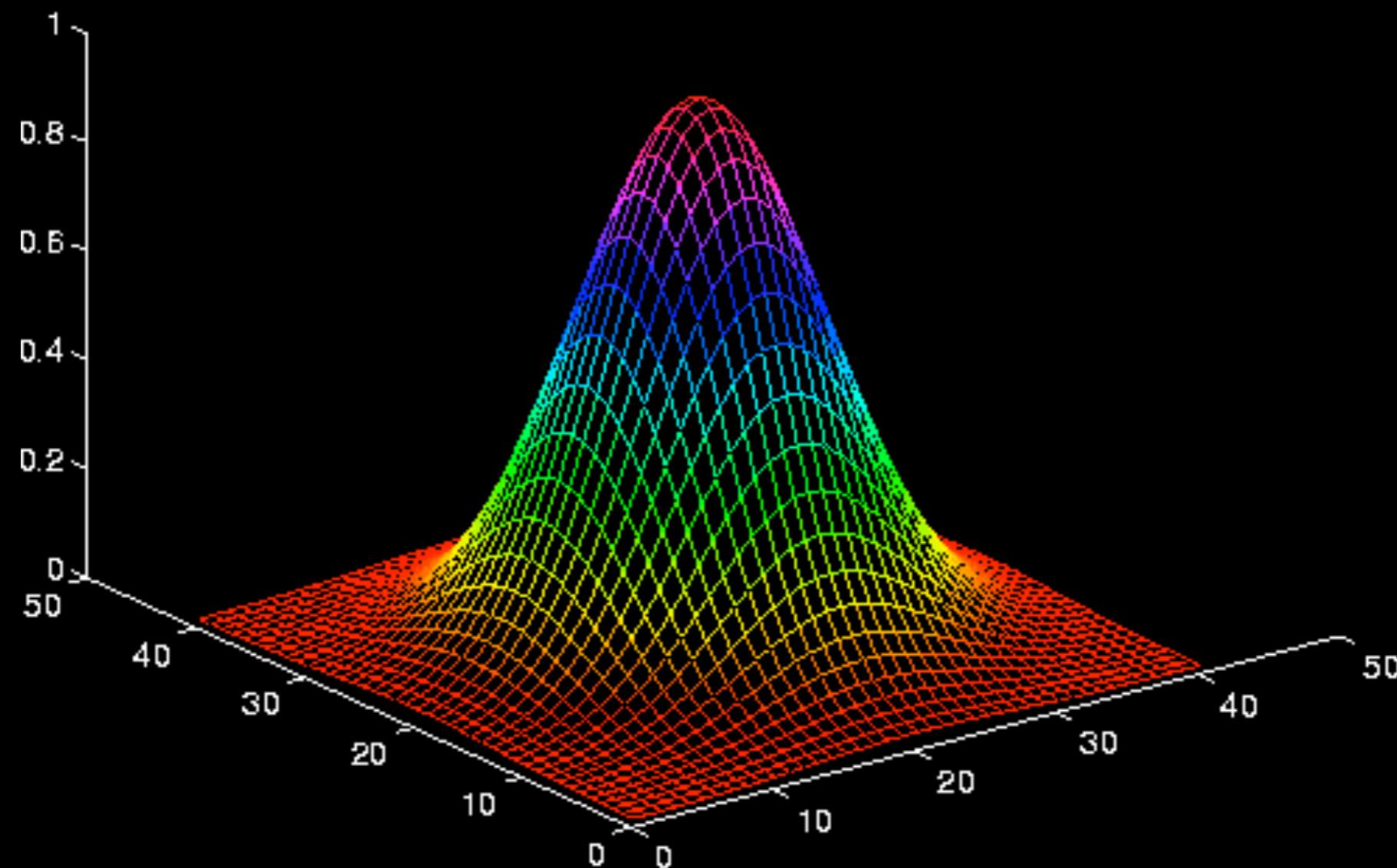
session log



session log



Step 3: Training the model (numeric optimization)



Step 4: Classify New Data

- Straightforward ranking process is simple
 - Use learned weights to score articles for the current user
 - Sort articles by score
- Simple classification process is too slow
 - Must constructing a new object (the feature vector) for hundreds of thousands of articles during ranking

original formulation

weight vector is universal, features are user-specific

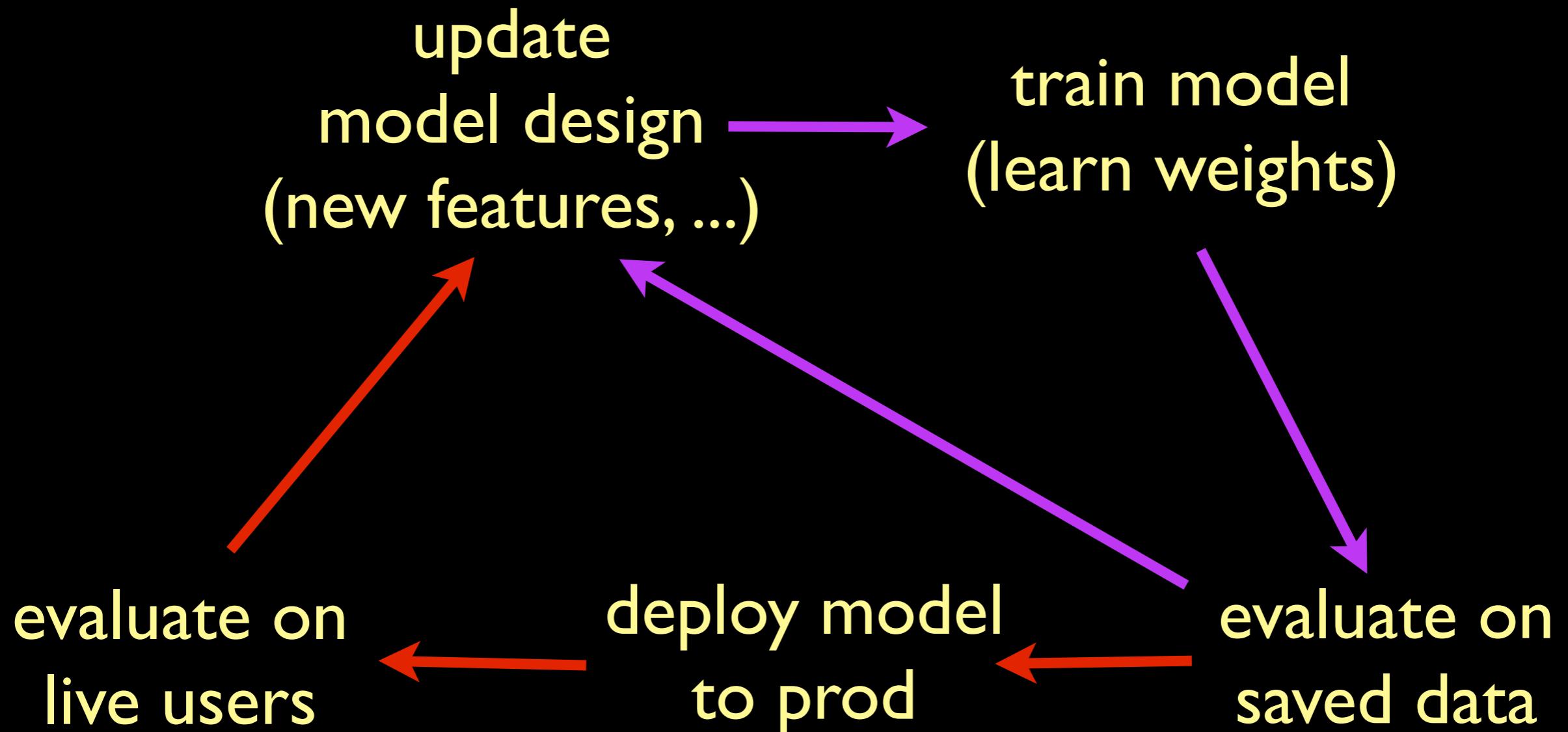
<i>feature</i>	<i>value</i>	<i>weight</i>
article has the topic ‘Cats’?	I	0.5
article has the topic ‘Photography’?	I	0.75
# of article topics user has favorited	I	1.25
<i>total score</i>		2.5

faster formulation

weight vector is user-specific, features are universal

<i>feature</i>	<i>value</i>	<i>weight</i>
article has the topic ‘Cats’?	I	0.5 + 1.25
article has the topic ‘Photography’?	I	0.75
<i>total score</i>		2.5

The Development Cycle



purple = fast, frequent cycle red = slow, infrequent cycle

Challenges We Faced

- Data bugs (are the worst bugs)
- Presentation bias
- Ranking vs “normal” machine learning
- Statistical bleeding
- Simpson’s paradox

Data bugs (are the worst)



At first we were like this ...

Data bugs (are the worst)



... but then we were like this.

Data bugs (are the worst)

during training, what the feature value should be for every clicked-on article with a previously unseen publisher

user's interest score for doc's publisher

0

during training, what the feature value actually (incorrectly) was for every clicked-on article with a previously unseen publisher

user's interest score for doc's publisher

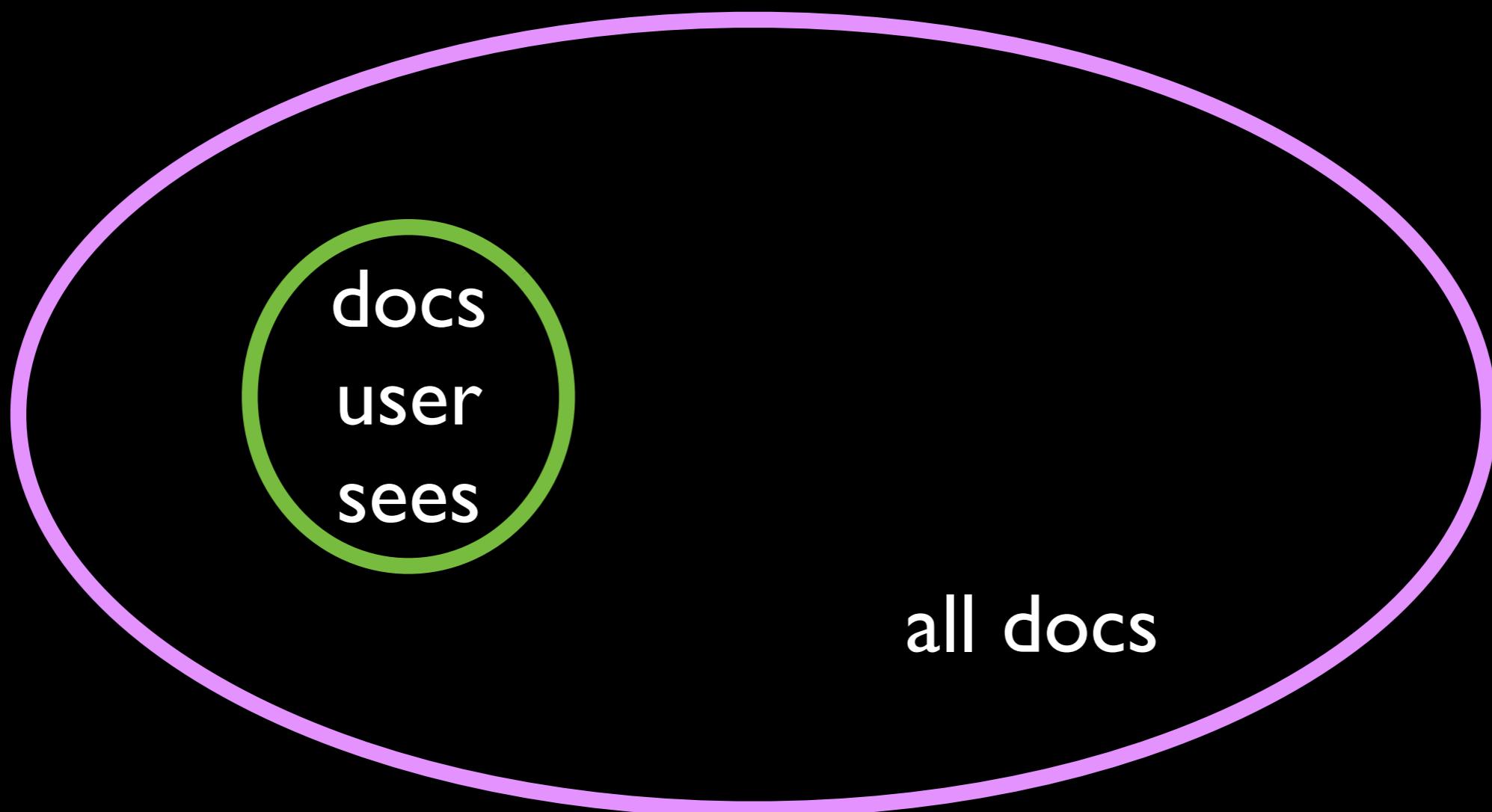
0.2

Result: the numeric optimization process exploits this bug, resulting in a large positive weight for this feature

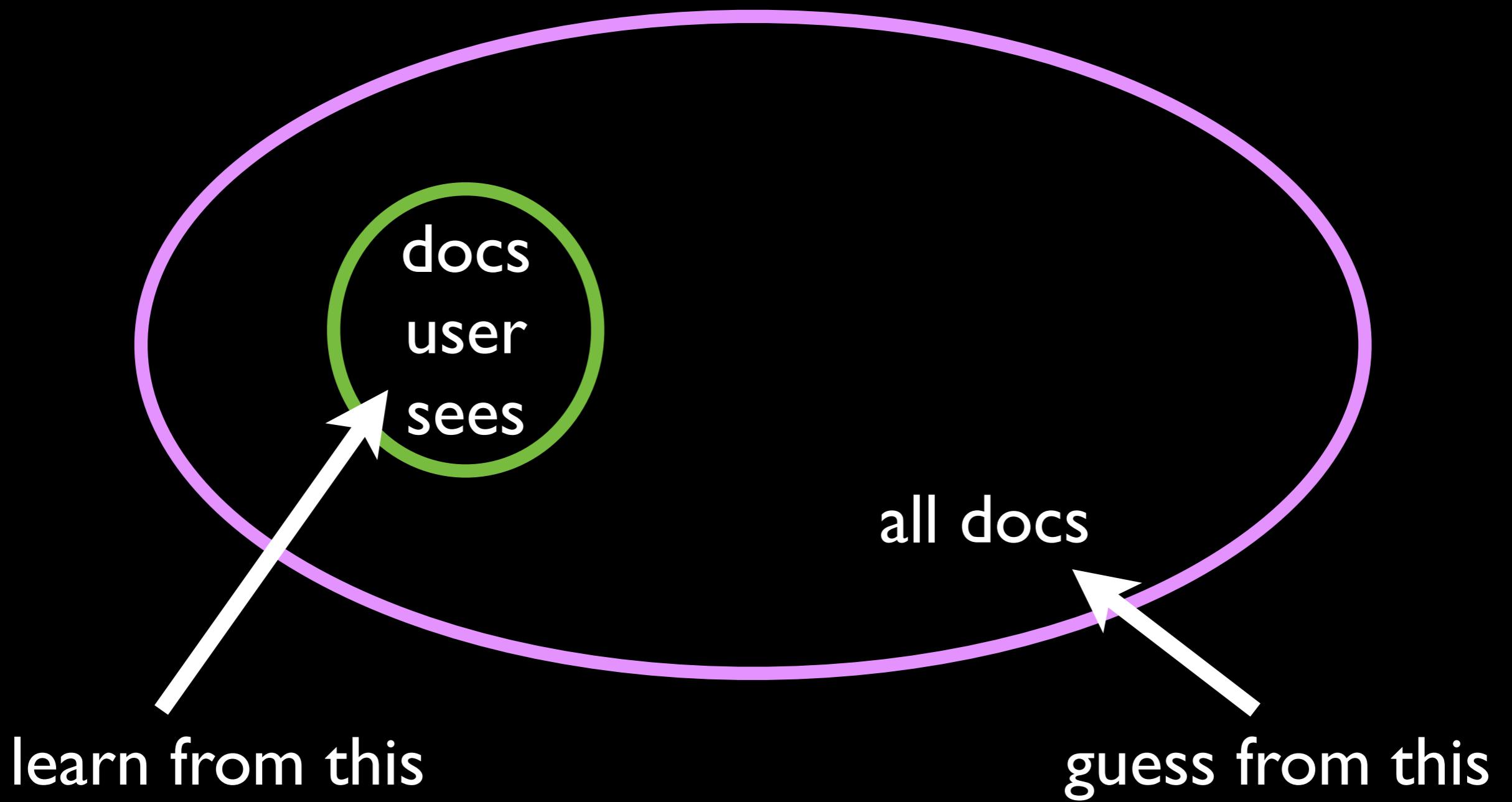
Presentation bias

- Users are more likely to click earlier articles in a feed
- Adds noise to the training data
 - model assumes user clicked because of interest in the article and not it's location
 - Makes live ranker results look better than they are

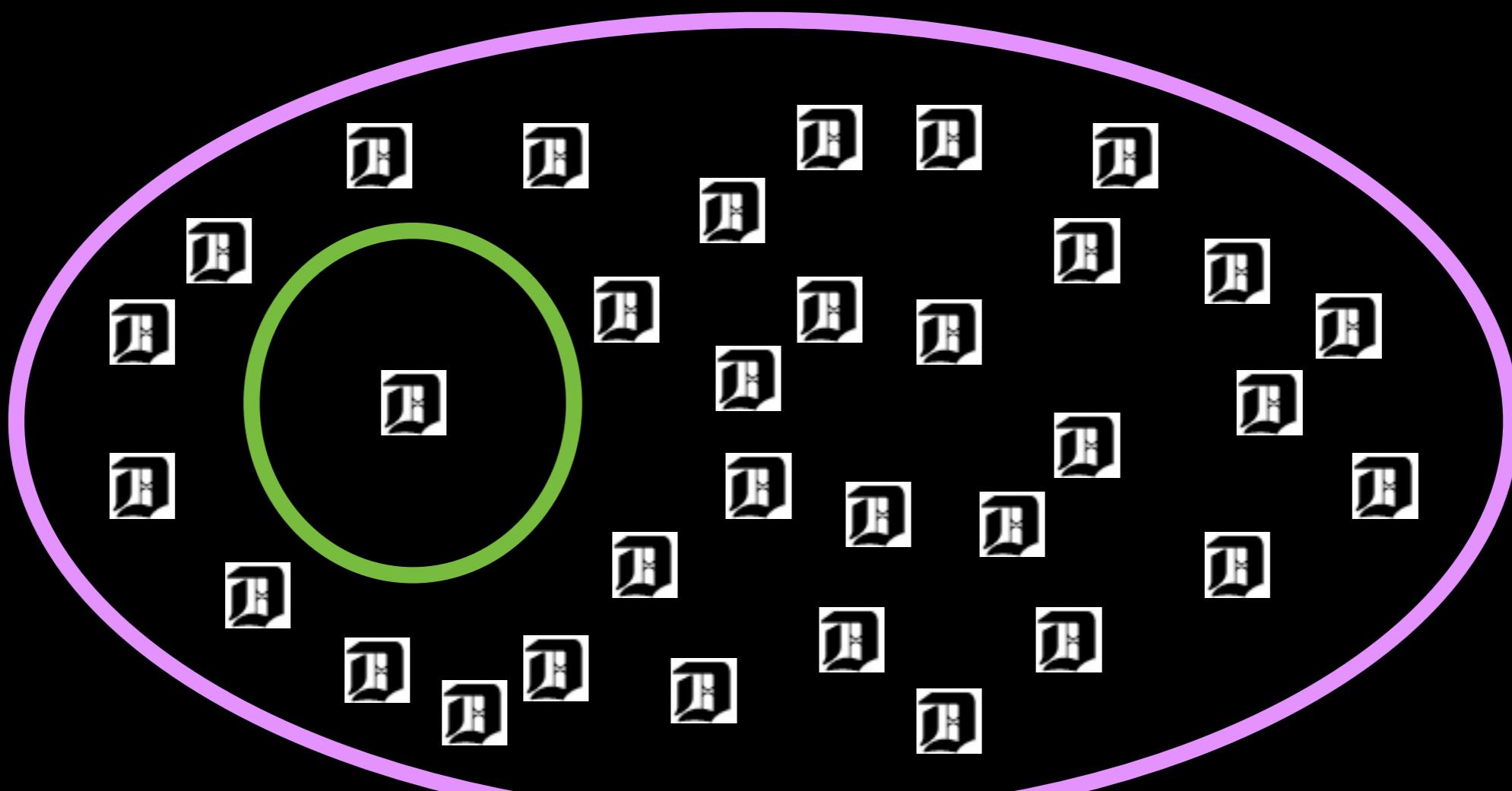
Ranking vs Normal ML



Ranking vs Normal ML

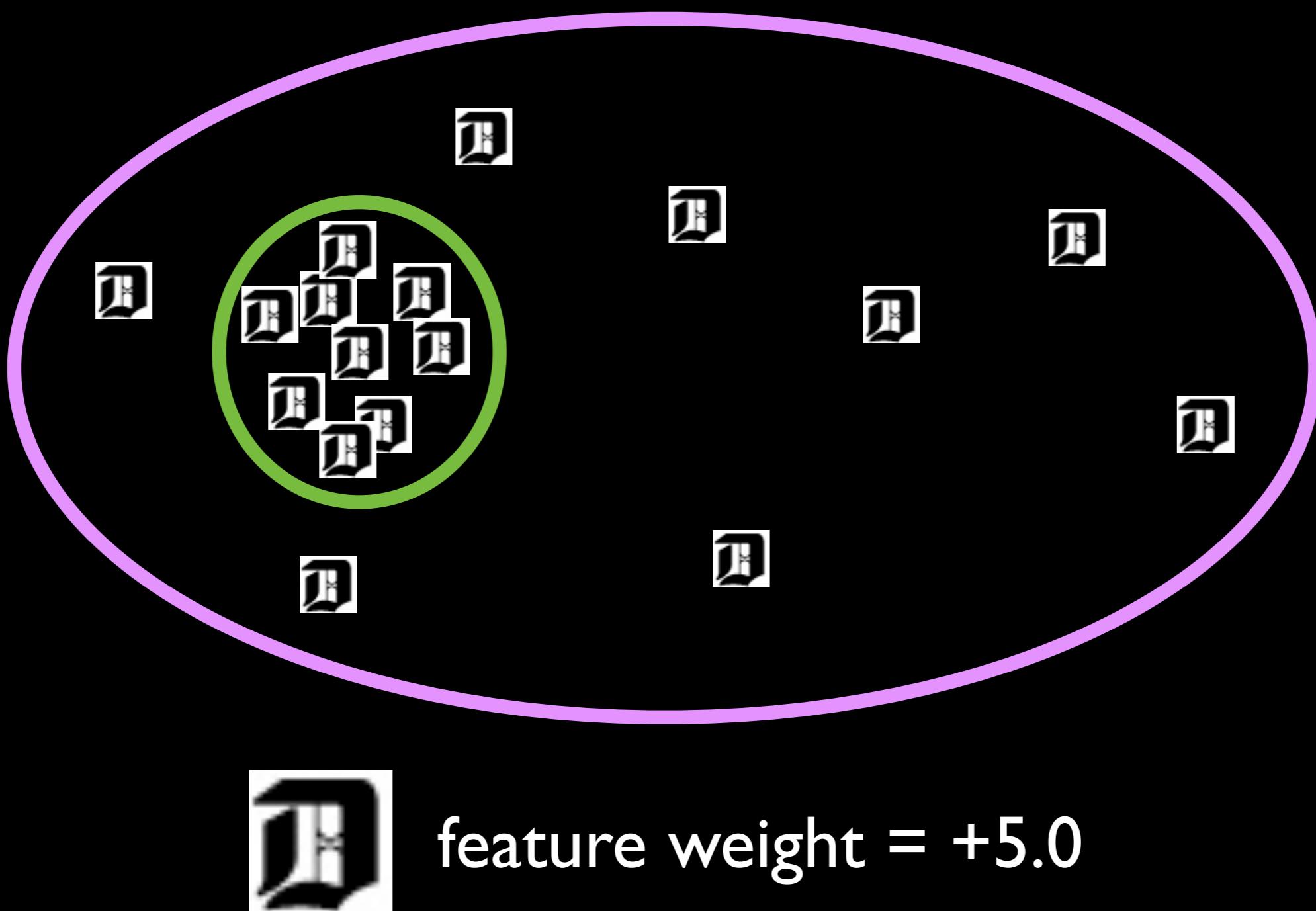


Ranking vs Normal ML

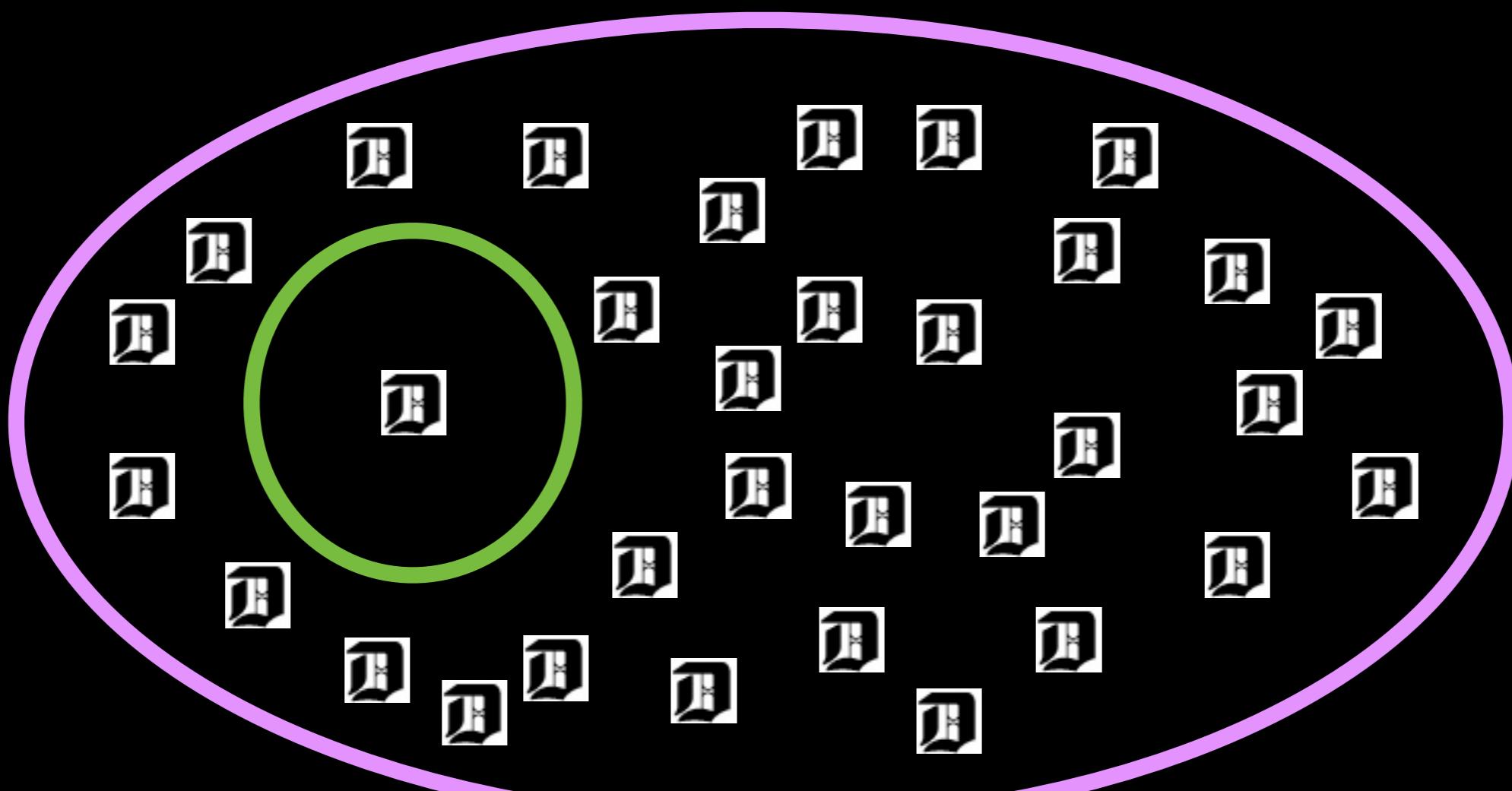


feature weight = -5.0

Ranking vs Normal ML



Ranking vs Normal ML



feature weight = -5.0

Statistical Bleeding

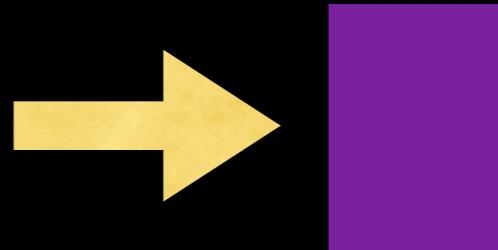


the explorer

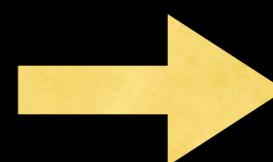


the exploiter

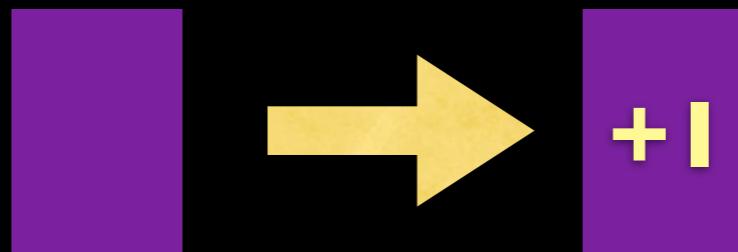
I. Brand new article
enters the index



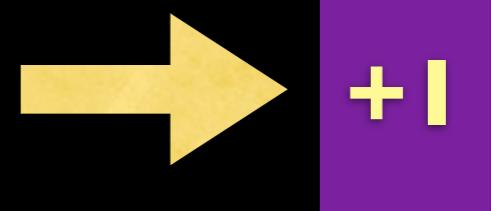
2. The explorer shows
the article to a user
who then clicks on it



3. The statistics on the
article get updated



4. The exploiter now
sees that this is a
good article, shows it
to other users



Simpsons Paradox

	Applicants	Admitted
Men	8442	44%
Women	4321	35%

[http://en.wikipedia.org/wiki/Simpson's_paradox]

Simpsons Paradox

	Applicants	Admitted
Men	8442	44%
Women	4321	35%

Departmant	Men		Women	
	Applicants	Admitted	Applicants	Admitted
A	825	62%	108	82%
B	560	63%	25	68%
C	325	37%	593	34%
D	417	33%	375	35%
E	191	28%	393	24%
F	272	6%	341	7%

[http://en.wikipedia.org/wiki/Simpson's_paradox]

Simpsons Paradox

	Overall Results	
	# sessions	score
ranker 1	100	0.85
ranker 2	100	0.83

	iPhone Results		Web Results	
	# sessions	score	# sessions	score
ranker 1	75	0.86	25	0.80
ranker 2	25	0.89	75	0.81

(*) not real numbers - meant to be illustrative

Results

- 50% more shares
- 25% more likely that first doc is good
- 14% longer sessions
- 15% fewer “total crap” sessions

Summary

- We collect information about our users' preferences and dispreferences
- We analyze all articles in our index to understand what they are about
- We snapshot user sessions to learn from
- We extract features from our data, which are hints which should help us decide if a user will like an article
- We learn weights for these features, which determine how the features trade off between each other
- We use these weights to rank articles for a user

Prismatic Backend Team



Jason
Wolfe

Dave
Golland

Aria
Haghghi

Ian
Wong

Jenny
Finkel (me)

Sound like fun? <http://getprismatic.com/jobs>