



# Relevance Debugging and Explaining at LinkedIn

May 20<sup>th</sup>, 2019

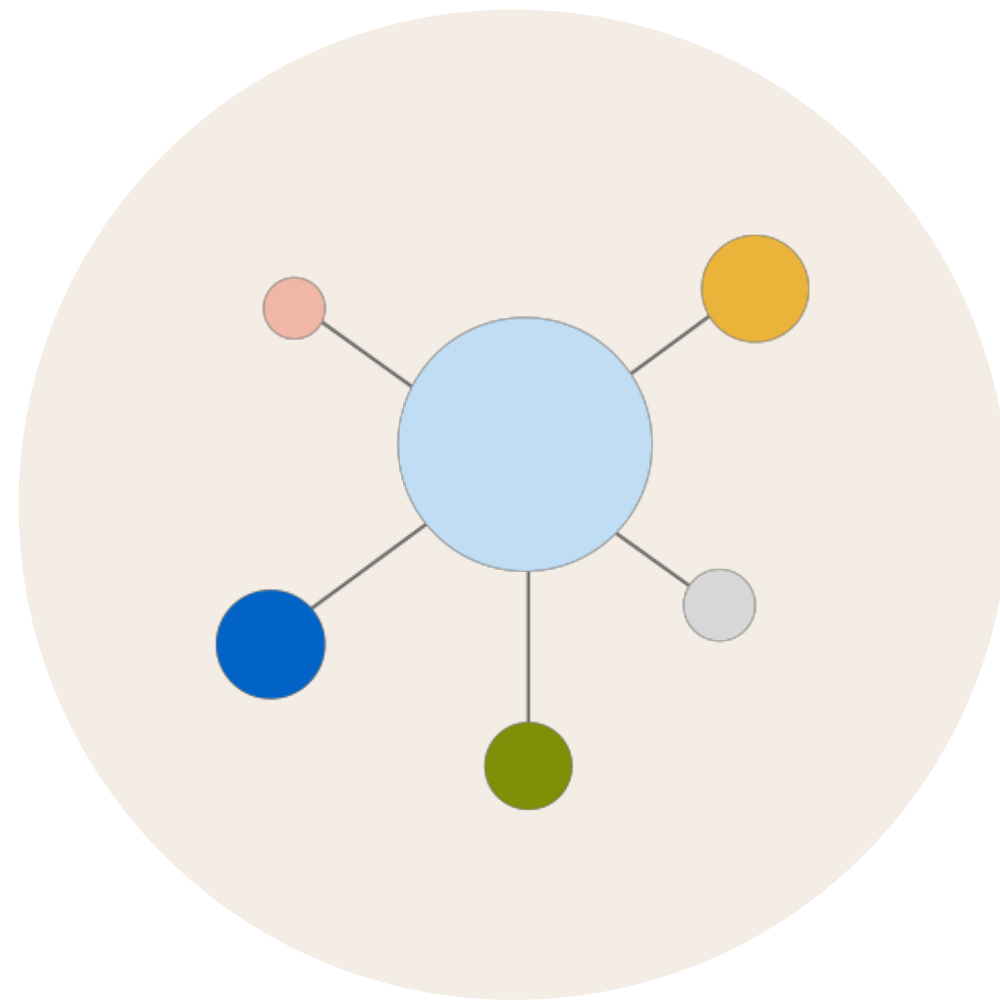


Daniel Qiu  
Software Engineer



Yucheng Qian  
Sr. Software Engineer

# Why relevance debugging and explaining



## Modeling

Improve the machine learning model



## Value

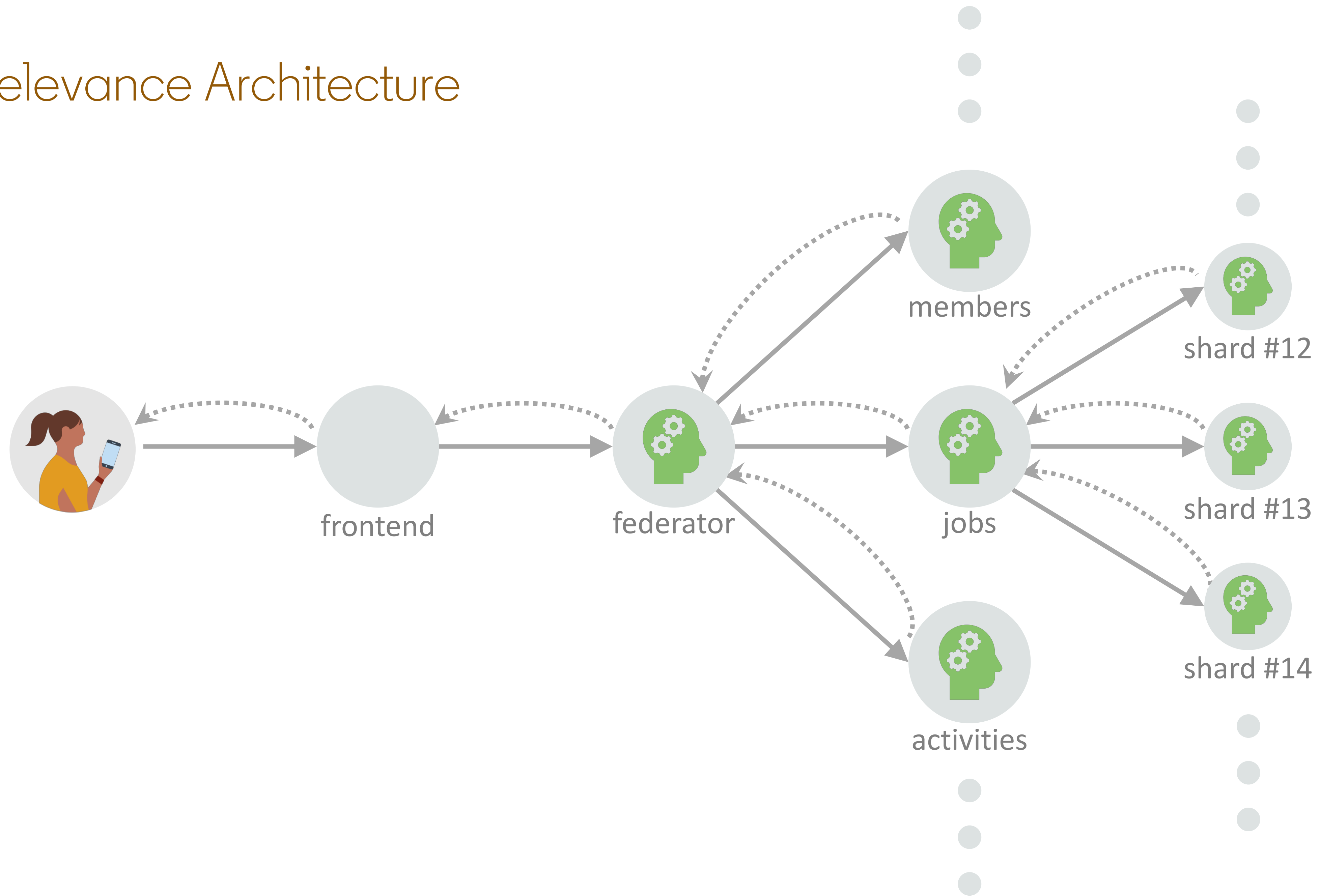
Bring value to our members by providing relevant experience



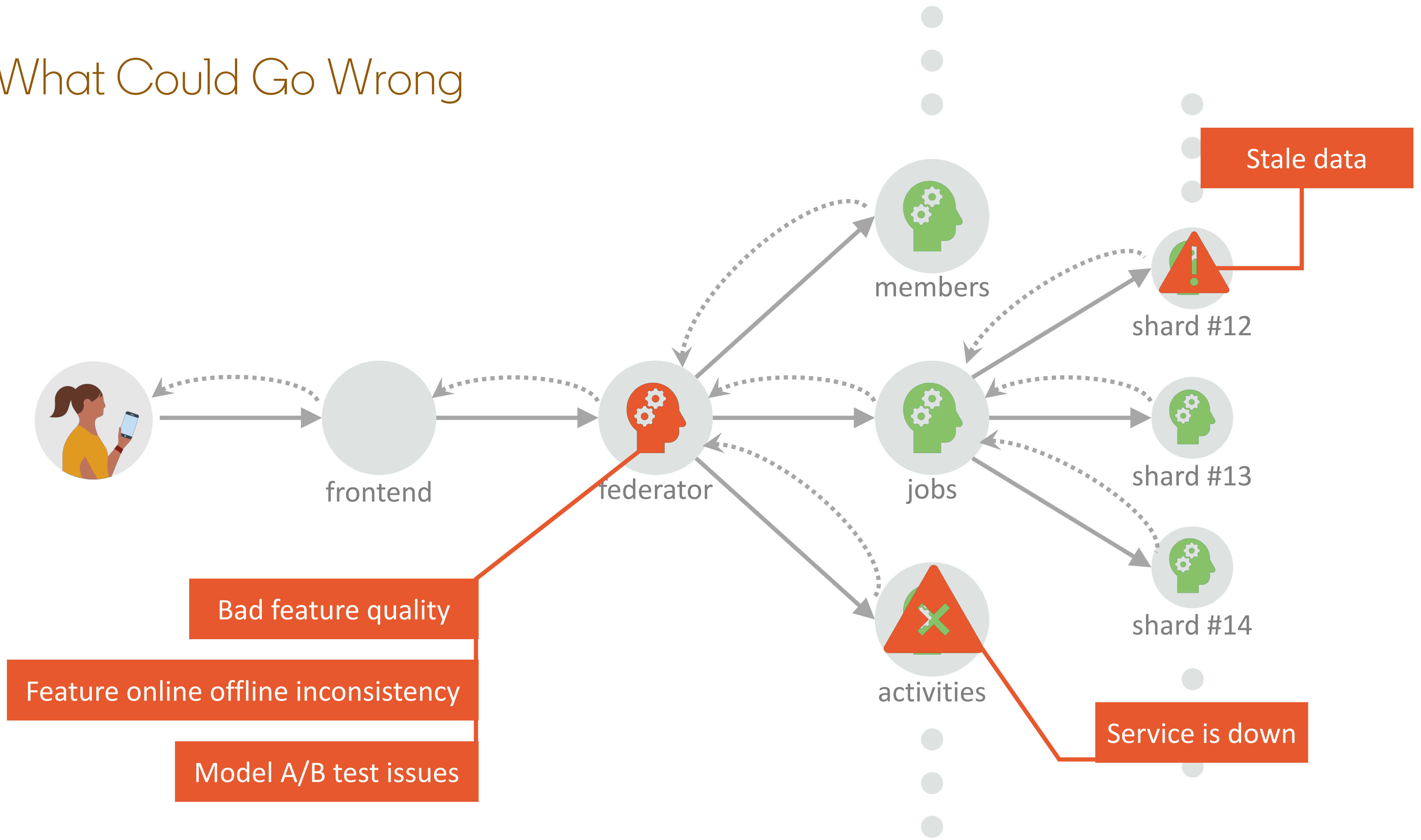
## Trust

Build trust with our members

# Relevance Architecture



# What Could Go Wrong



Bad feature quality

Feature online/offline inconsistency

Model A/B test issues

Stale data

Service is down



# Challenges



Complex Infrastructure

---



Hard to Reproduce

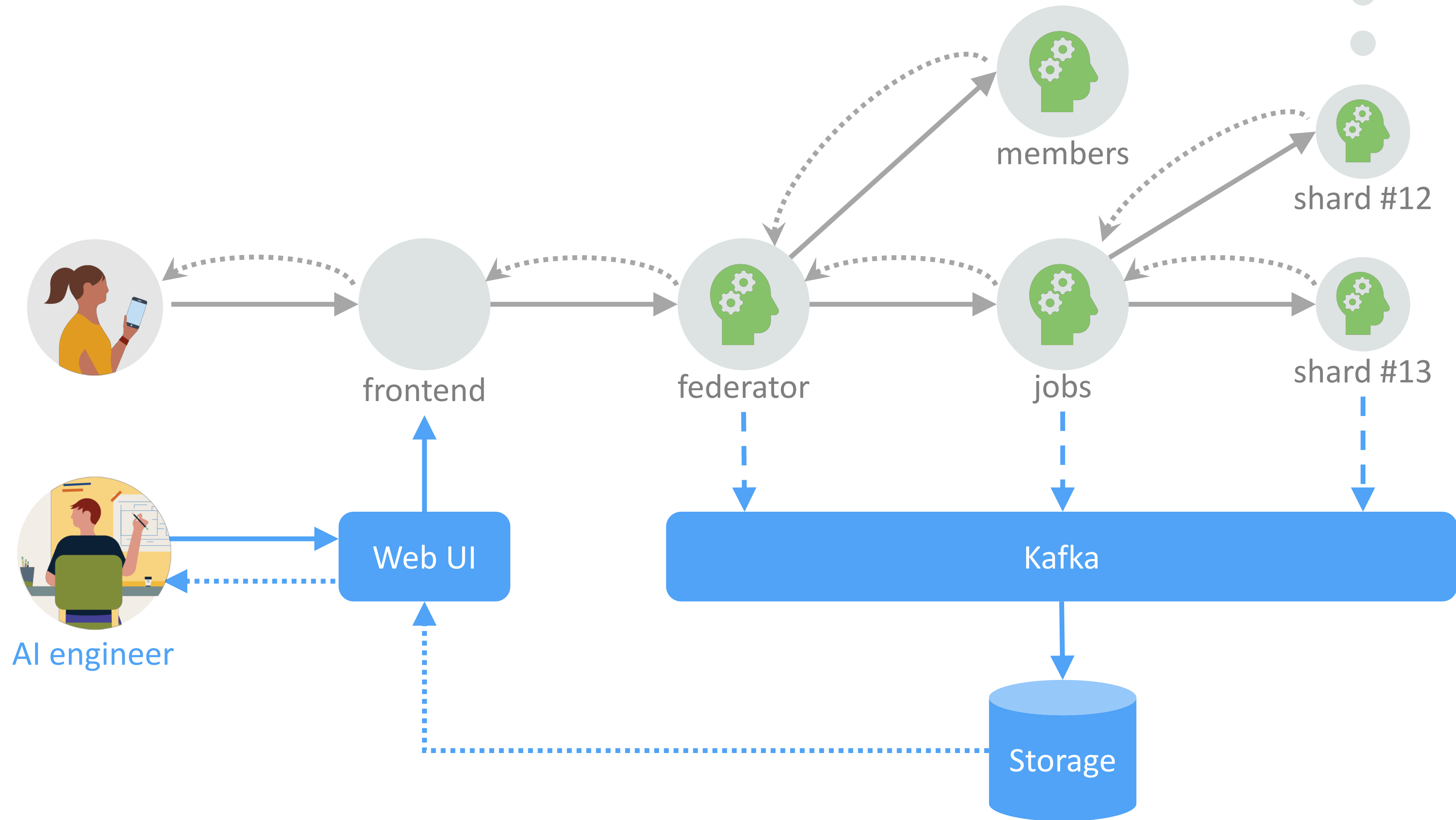
---



Time Consuming

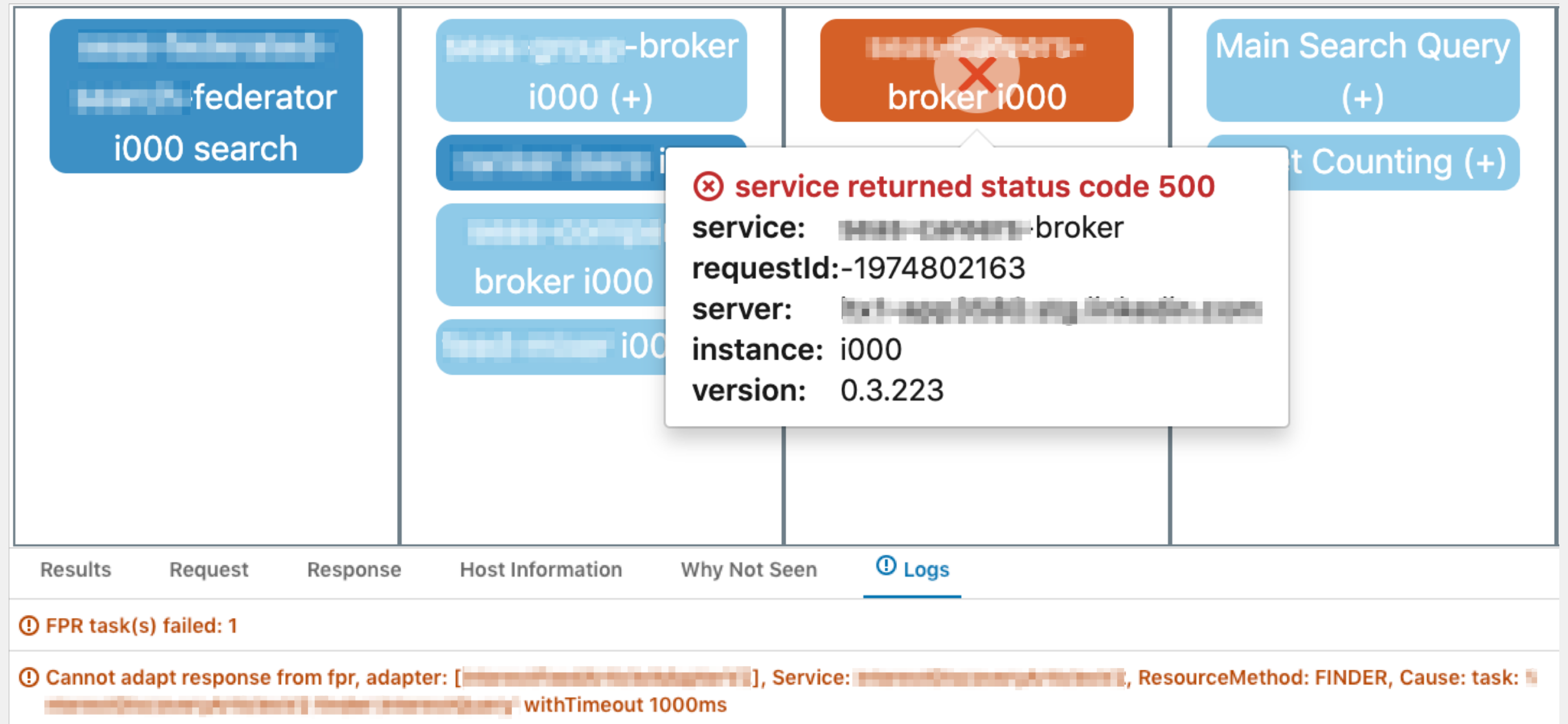
---

# Relevance Debugging and Explaining




# Call Graph

Understand the call graph and see what goes wrong in the service



# Targeted Query

Debug a specific service, data center, cluster, or host

**Endpoint Selection** 

**Endpoint**

**Fabric**

**Hostname**

**Partition**

**Tag**

# Timing Data

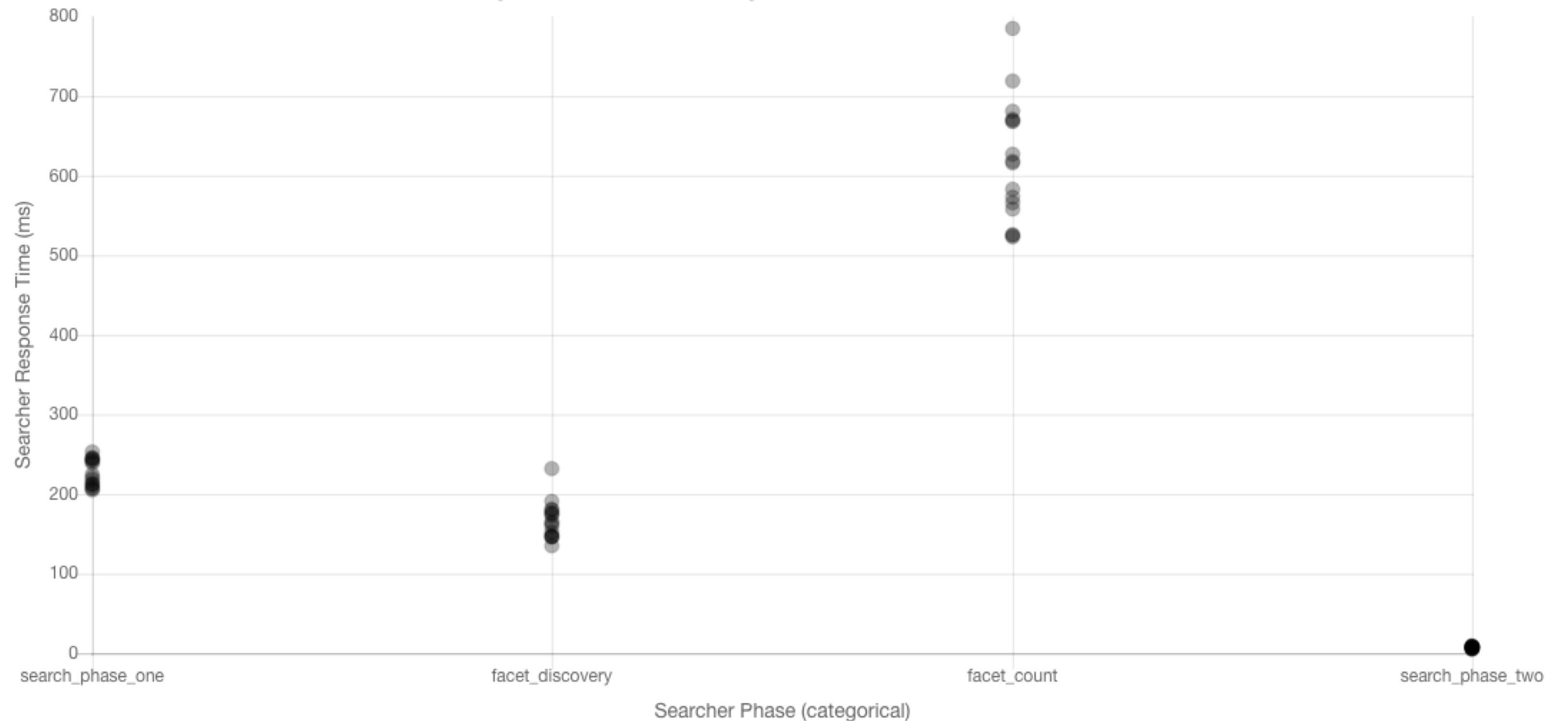
The statistics about the time spent on each phase of a search.

Total time (ms): 1041

Number of garbage collection events: 0


	Start Time	End Time	Total Time	Resent?	Partitions	Min	Max	p50	p90
search_phase_one	7	266	259	false	16	205	253	223.0	245.5
facet_discovery	13	240	227	true	16	135	232	164.0	186.0
facet_count	262	1041	779	true	16	523	785	617.0	700.0
search_phase_two	266	274	8	false	15	5	9	8.0	9.0

Scatterplot of Searcher Response Time and Searcher Phase



# Features

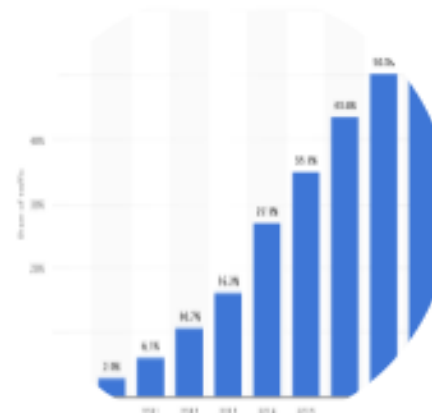
Inspect the features for the model

Group	Feature 	Value
SPR	activity_recent_click /	968
SPR	[REDACTED]	1
SPR	[REDACTED]	6.8762646
SPR	[REDACTED]	null
SPR	[REDACTED]	null
SPR	binary_activity_recent_click /	1
SPR	[REDACTED]	null
SPR	log_activity_recent_click /	6.8762646
SPR	[REDACTED]	0
SPR	[REDACTED]	0



# Model and Score

Inspect the models for ranking and scores of the items













## How to Conduct Mobile UX Research (and What to Do with It)

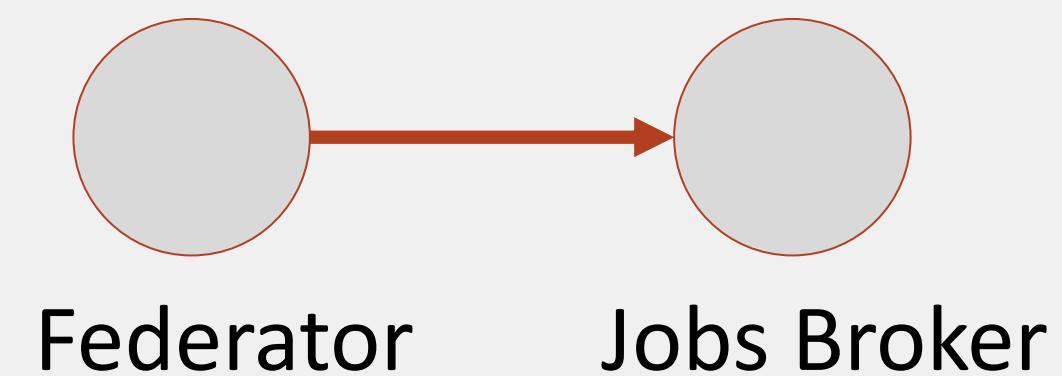
The number of people browsing and shopping on mobile devices continues to grow. During the past 10 years, smartphone and.....







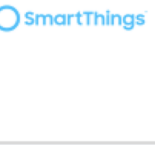



<b>Position</b>	#1
<b>Reference</b>	urn:li:article: [REDACTED]
<b>SPR Score</b>	0.0035542063
<b>Relevance Model</b>	[REDACTED]
<b>Source Type</b>	ORGANIC
<b>FPR Model</b>	[REDACTED]

# Entities Returned

See how entities are reranked and filtered between services.

	<b>Senior Software Engineer</b> Mountain View, California GroundTruth
	<b>Senior Platform Software Engineer</b> San Francisco, CA, US Juul Labs, Inc.
	<b>Senior Software Engineer iOS</b> San Jose, California Diligente Technologies
	<b>Senior / Software Engineer (Java)</b> Redwood City, CA, US PubMatic
	<b>Senior Software Engineer</b> Dallas/Fort Worth Area Riccione Resources
	<b>Senior Software Engineer</b> Sunnyvale, CA, US Walmart Labs
	<b>Senior Software Engineer - React...</b> 743 Clementina Street , San Francisc... Peerspace
	<b>Senior Software Engineer - Devops</b> Herndon, Virginia Cynet Systems Inc
	<b>Senior Software Engineer</b> Mountain View, CA, US Smartthings
	<b>Senior Software Engineer, Digital</b> New York City, NY, US Teachers Pay Teachers



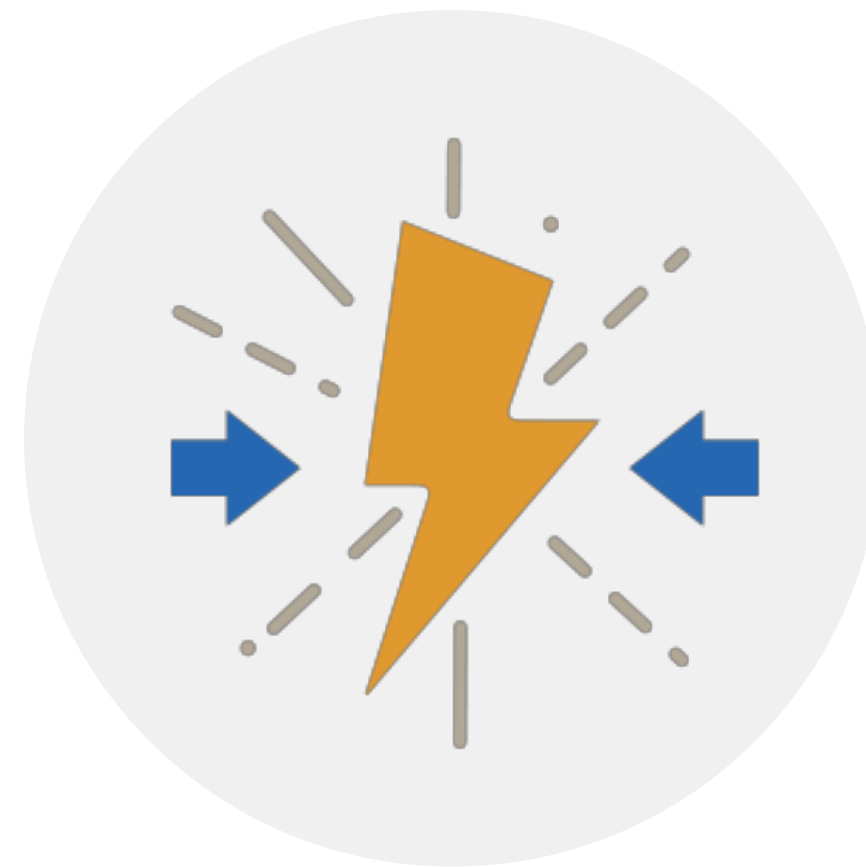
	<b>Senior Software Engineer iOS</b> San Jose, California Diligente Technologies
	<b>Senior / Software Engineer (Java)</b> Redwood City, CA, US PubMatic
	<b>Senior Software Engineer</b> Dallas/Fort Worth Area Riccione Resources
	<b>Senior Software Engineer</b> Sunnyvale, CA, US Walmart Labs
	<b>Senior Software Engineer - React...</b> 743 Clementina Street , San Francisc... Peerspace
	<b>Senior Software Engineer - Devops</b> Herndon, Virginia Cynet Systems Inc
	<b>Senior Software Engineer</b> Mountain View, CA, US Smartthings
	<b>Senior Software Engineer, Digital</b> New York City, NY, US Teachers Pay Teachers
	<b>Senior Software Engineer</b> Sunnyvale, CA, US Walmart
	<b>Senior Software Engineer (San J...</b> San Jose, California, United States TekVizor

# Advanced Use Cases



Perturbation

---



Comparison

---



Replay

---

# System Perturbation

Turning knobs and see what happens.

## 1. Inject

---

Injected as part of the request

## 2. Relay

---

Passed to downstream service

## 3. Overwrite

---

Overwrite the system behavior

# Perturbation Example

- Override A/B test settings
- Model selection
- Feature override

# Side-by-Side Comparison

## Compare Model

Compare results of 2 different queries/models

## Compare Items

Compare features and scores of 2 different items, from the same query or different queries



Position changes: 3 | New items: 11

Click to view details, or select to compare.

Query 1

cURL Calltree

Query 2

cURL Calltree

#1.1 → #1.4 SPR: 0.017552437



Lead Software Engineer – Platform  
Confidential

#1.2 → #1.1 SPR: 0.017409125



Test Engineering Software Development Lead  
Flextronics

#2 → SPR: 0.0068453606

Sponsored

Decorator for URN family unavailable

#3 → SPR: 0.04593608



#4 → SPR: 0.02149221

#1.2 → #1.1 SPR: 7.2239555E-4



Test Engineering Software Development Lead  
Flextronics

→ #1.2 SPR: 6.688792E-4



Software Engineer - Application Backend  
Yelp

→ #1.3 SPR: 6.687663E-4



Software Engineer - Messaging Services  
Twilio

#1.1 → #1.4 SPR: 6.686083E-4



Lead Software Engineer – Platform  
Confidential

Compare 2 queries

Query 1

**flex** Test Engineering Software Development Lead  
Flextronics

Position #1.2

Reference urn:li:jobPosting: [redacted]

SPR Score 0.017409125

Relevance Model [redacted]

Source Type ORGANIC

FPR Model [redacted]

Query 2

**flex** Test Engineering Software Development Lead  
Flextronics

Position #1.1

Reference urn:li:jobPosting: [redacted]

SPR Score 7.2239555E-4

Relevance Model [redacted]

Source Type ORGANIC

FPR Model [redacted]

All Groups ▼ 🔍 Search feature 
 Shared features only  Different values only

Group	Feature <span style="font-size: 0.8em;">↕</span>	Item 1	Item 2	% Change <span style="font-size: 0.8em;">↕</span>
SPR	responsePenalty /	4.0601455e-7	0.009018197	2221051.19
SPR	response	5.2125584e-9	0.000011580406	222063.57
SPR	score_response_viral	5.2125584e-9	0.000011580406	222063.57
SPR	diffHoursSinceLvFiveAndAgeInHour /	-3.0348454	-50.475624	1563.2

Compare same item in 2 queries

# Replay

Scoring and ranking could be time sensitive

- Relevance services emit features and scores using Kafka.
- Consume Kafka event using Samza and store data in the database
- Query database for historical data when debugging

# Feed Replay UI

### Feed Replay

Viewer ID:

**Viewer ID must be a LinkedIn employee.**

Start Time (Pacific Time):




End Time (Pacific Time):

[Load Sessions](#)

- 2019-03-26 13:12:30 PDT**  
Finder: UseCase  
DESKTOP\_HOMEPAGE\_NEPTUNE
- 2019-03-26 17:12:48 PDT  
Finder: UseCase  
DESKTOP\_HOMEPAGE\_NEPTUNE
- 2019-03-27 17:49:32 PDT  
Finder: UseCase  
PHONE\_HOMEPAGE\_VOYAGER
- 2019-03-27 17:56:05 PDT  
Finder: UseCase  
DESKTOP\_HOMEPAGE\_NEPTUNE
- 2019-03-27 18:28:51 PDT  
Finder: UseCase  
PHONE\_HOMEPAGE\_VOYAGER
- 2019-03-27 18:28:51 PDT  
Finder: UseCase  
PHONE\_HOMEPAGE\_VOYAGER
- 2019-03-28 10:12:35 PDT  
Finder: UseCase  
PHONE\_HOMEPAGE\_VOYAGER
- 2019-03-29 16:32:18 PDT  
Finder: UseCase  
DESKTOP\_HOMEPAGE\_NEPTUNE

### cURL

Calltree not available

- 1** urn:li:activity: [redacted] [Details](#) [Features](#) [Traces](#)  
 linkedin:group-post  
urn:li:groupPost: [redacted]  
Relevance Model: nus:homepage\_federator\_relevance\_463\_ramp  
FPR Model: m124\_v2\_multi\_pass
- 2** sponsored urn:li:sponsoredContentV2:  
(urn:li:activity:[redacted],  
urn:li:sponsoredCreative:[redacted]) [Details](#) [Features](#) [Traces](#)  
Decorator for URN family unavailable  
Relevance Model: nus:homepage\_federator\_relevance\_463\_ramp  
FPR Model: su:2700601;pc:sc\_003!100000;
- 3** urn:li:activity: [redacted] [Details](#) [Features](#) [Traces](#)  
 linkedin:like  
urn:li:activity: [redacted]  
Relevance Model: nus:homepage\_federator\_relevance\_463\_ramp  
FPR Model: m124\_v2\_multi\_pass
- 4** urn:li:activity: [redacted] [Details](#) [Features](#) [Traces](#)  
 linkedin:react  
urn:li:groupPost: [redacted]  
Relevance Model: nus:homepage\_federator\_relevance\_463\_ramp





## Protecting members' privacy

- GDPR compliant
- Controlled and audited access
- Limited data retention time
- Replay for internal employees only



# Customer AI teams

- Search
- Feed
- Comments
- People you may know
- Job you may be interested in
- Notification



Q&A