# Distilling the Outcomes of Personal Experiences:

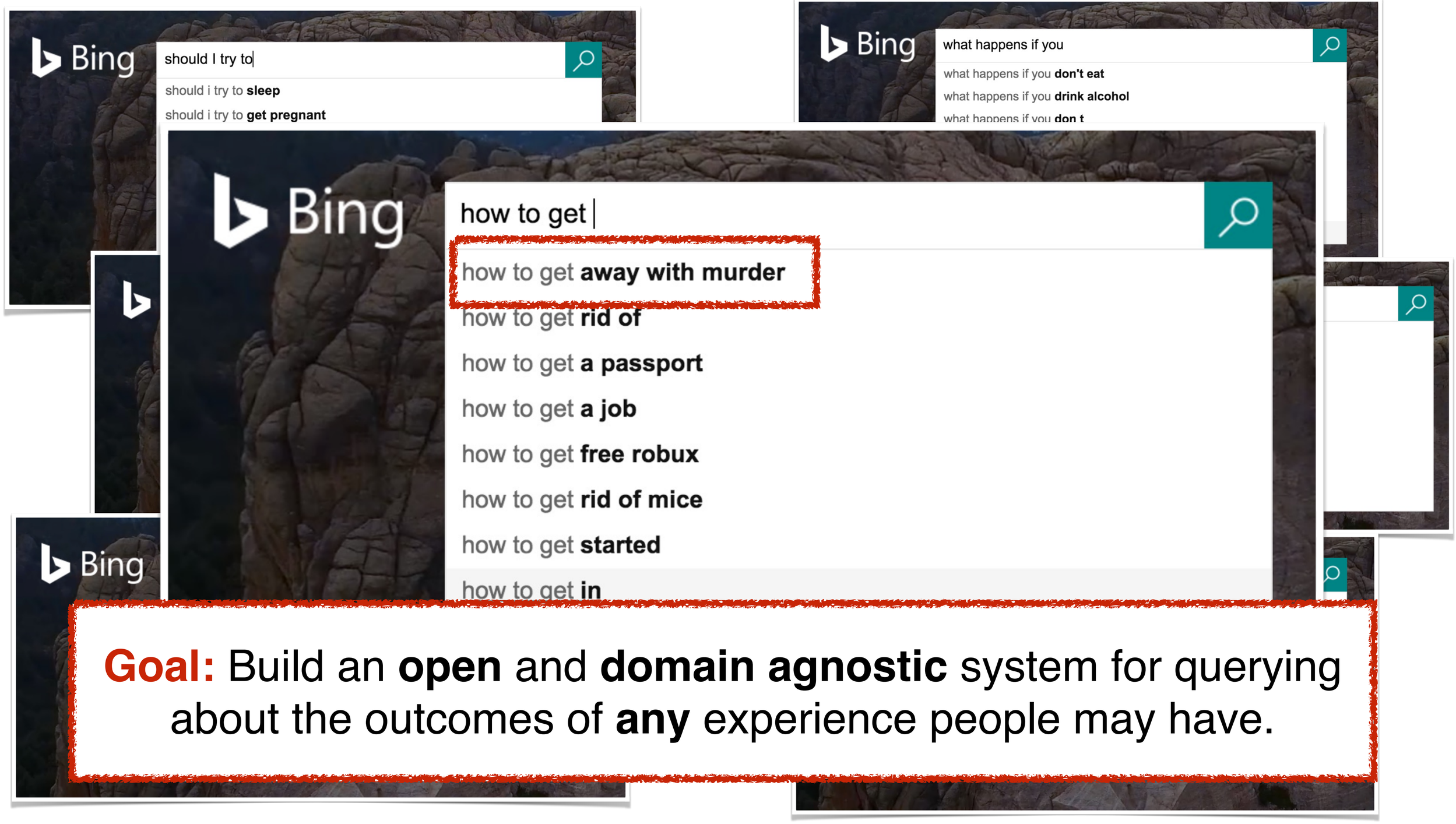
## A Propensity-scored Analysis of Social Media

***Alexandra Olteanu**
(IBM Research/EPFL)

***Onur Varol**
(Indiana University)

**Emre Kıcıman**
(Microsoft Research)

*** work done at Microsoft Research**




**Goal:** Build an **open** and **domain agnostic** system for querying about the outcomes of **any** experience people may have.

**Goal:** Build an **open** and **domain agnostic** system for querying about the outcomes of **any** experience people may have.

"Long-tail" of situations and experiences

- **Explore a situation: What happens …?**
  - when depressed, after disease diagnosis, after being fired, …
- **Understand the effects of a potential action: Should I …?**
  - get pregnant, ask for divorce, lose belly fat, change last name, …
- **Plan for outcomes/goals: How to …?**
  - lose weight, get admitted to MIT, increase income, find true love

Applications for individuals, policy makers, and others

# Social Media Posts: A Proxy to User Experiences

## Experiences & situations

**I ate** lots of fried things **today** and thoroughly enjoyed it. 🍔🍟🍗

**I**'m glad **I went** to the **show**. It was an experience **I** had to have, and **I**'m thankful.

**I had** my first **car accident** this morning

# Social Media Posts: A Proxy to User Experiences

**Experiences & situations**

**Post-hoc events (potential outcomes)**

**I ate** lots of fried things **today** and thoroughly enjoyed it. 🍔🍟🍗

Everyone got problems losing weight and **I** got problems **gaining weight** 😒

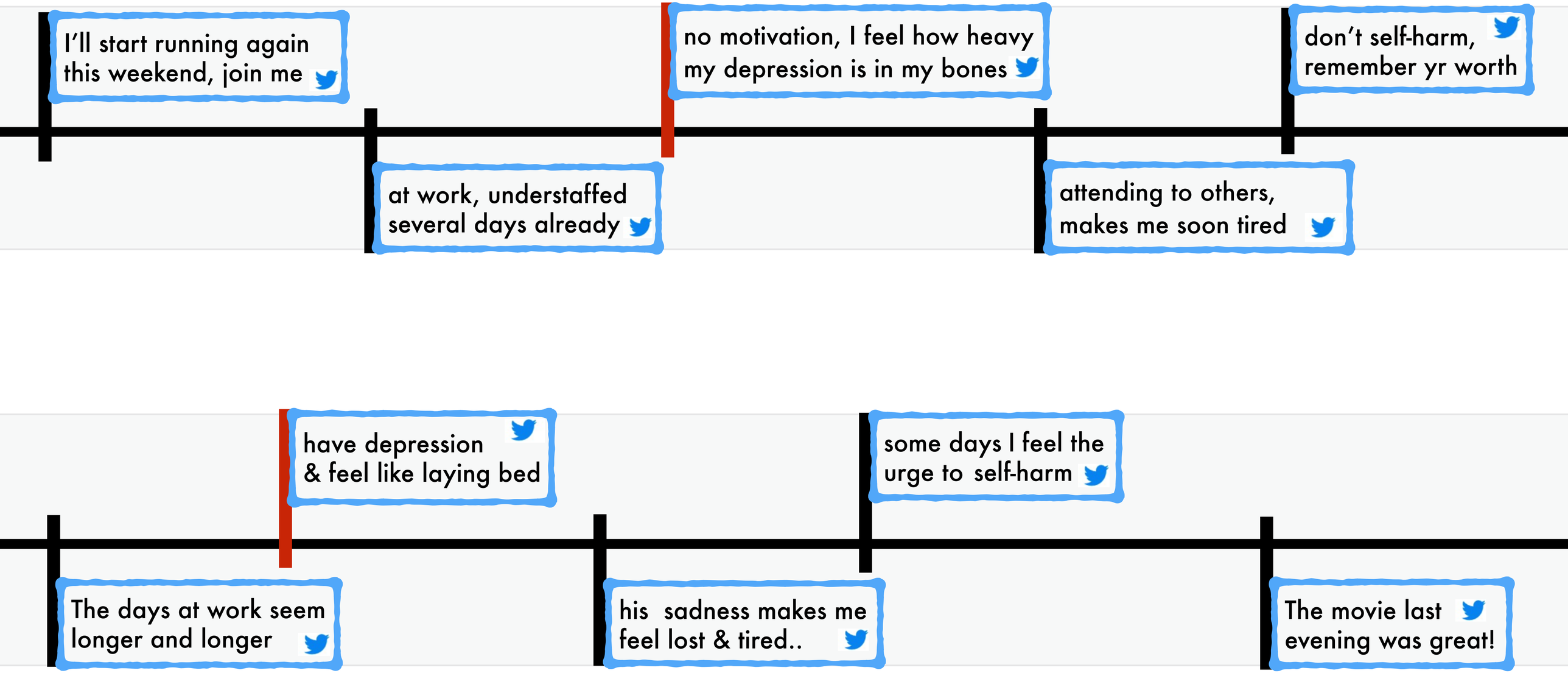
**I**'m glad **I went** to the **show**. It was an experience **I** had to have, and **I**'m thankful.

**i was** just woken up to a strawberry milkshake and a **relaxed** household. this is nice.

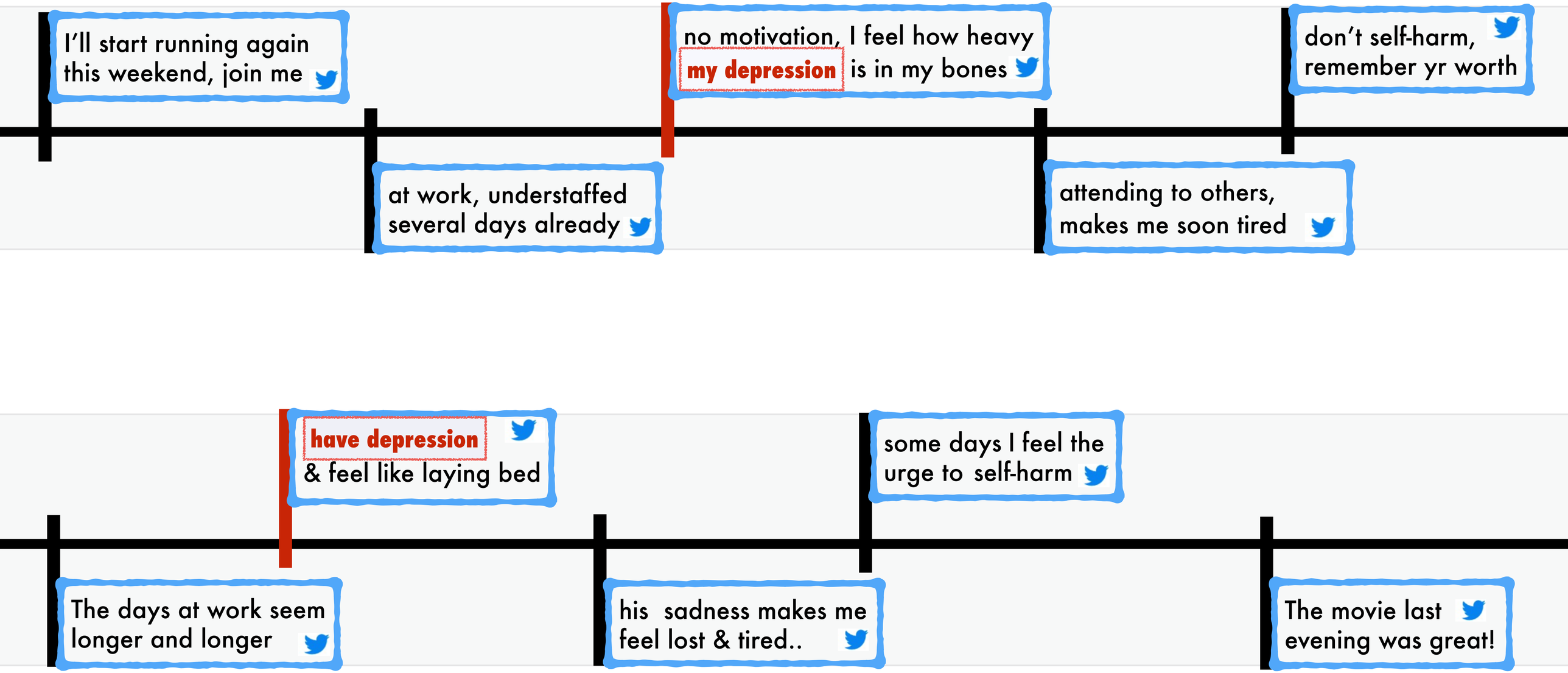
**I had** my first **car accident** this morning

Not having a car this week and maybe next week will be the longest, most hardest thing ever 😒 What a great Valentines Day 😢

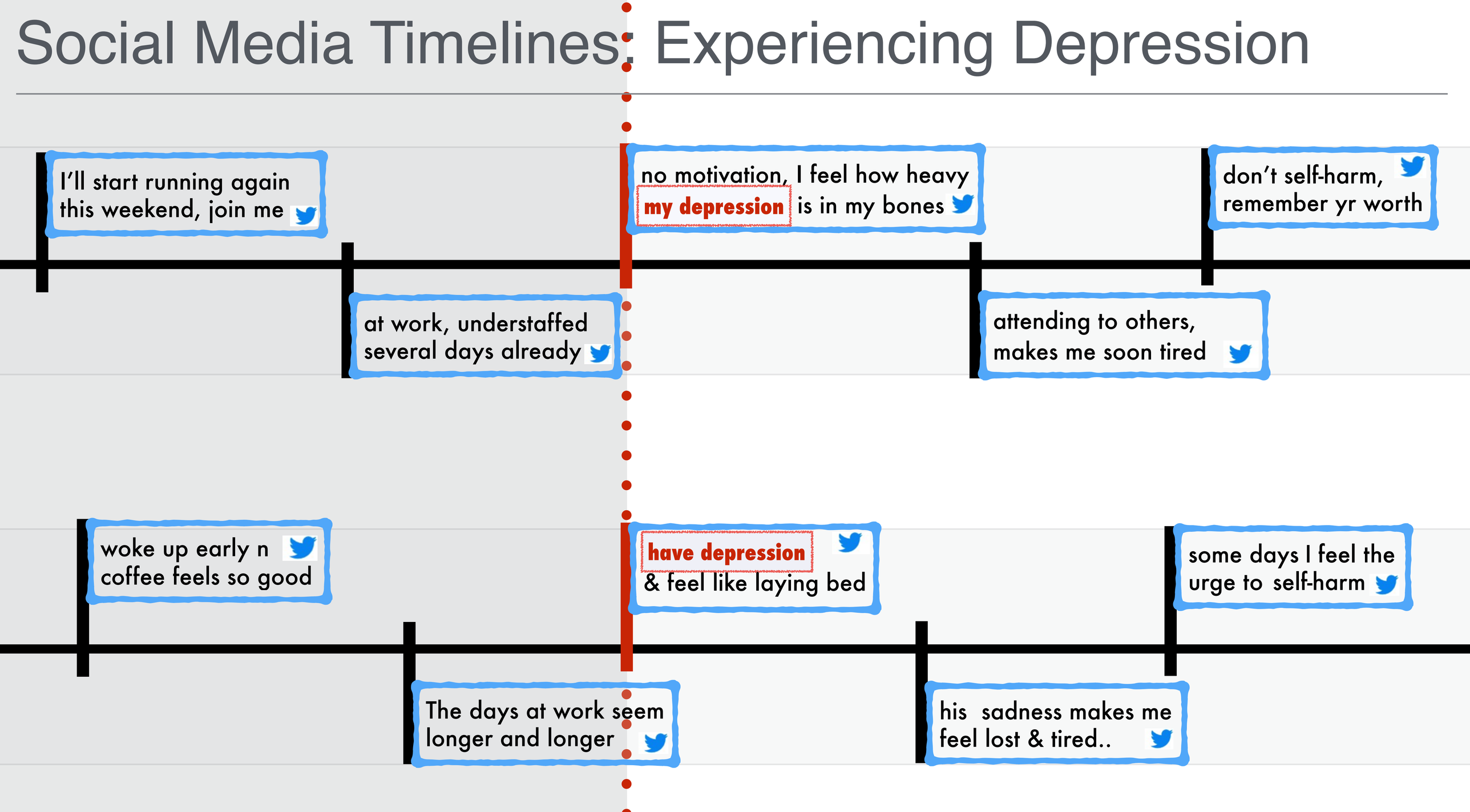
# Social Media Timelines: Experiencing Depression



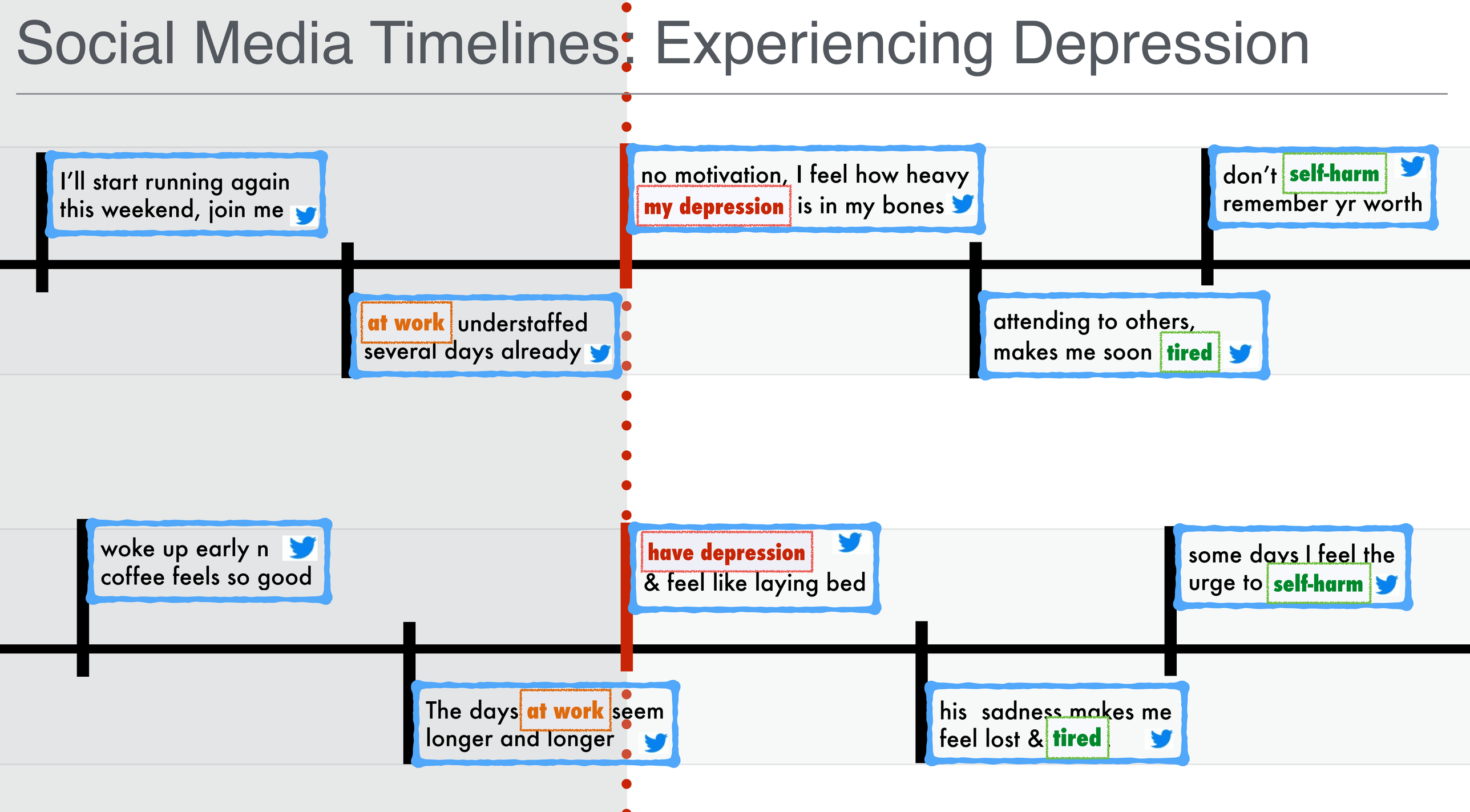
# Social Media Timelines: Experiencing Depression



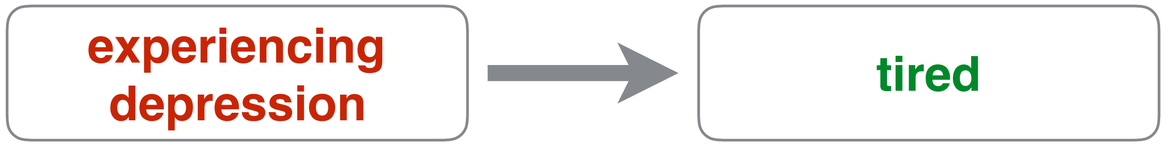
# Social Media Timelines: Experiencing Depression

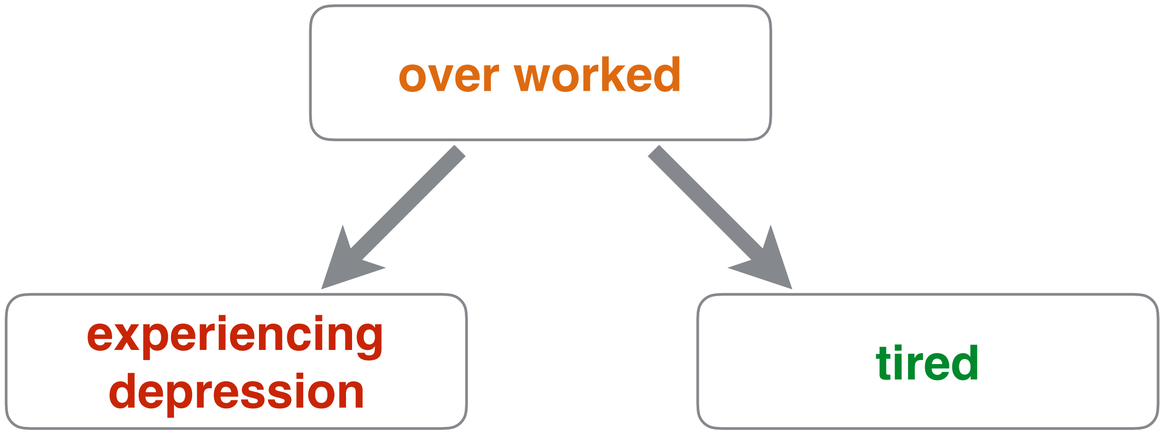


# Social Media Timelines: Experiencing Depression

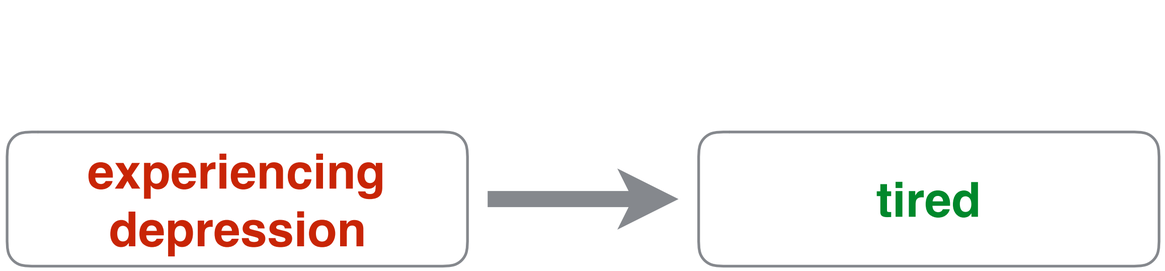


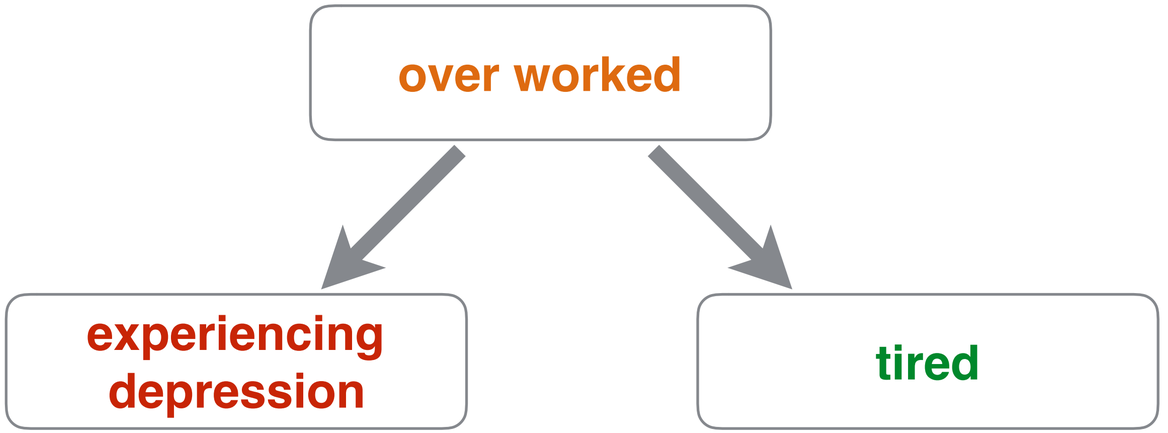
# Confounding Bias & Matching



# Confounding Bias & Matching




**Matching**: For every user with a particular experience, find another user with identical characteristics (prior to the experience) who didn't have the experience.

In high-dimensional spaces (e.g., textual data from social media), identical matches cannot be found; instead project units to a single dimension.

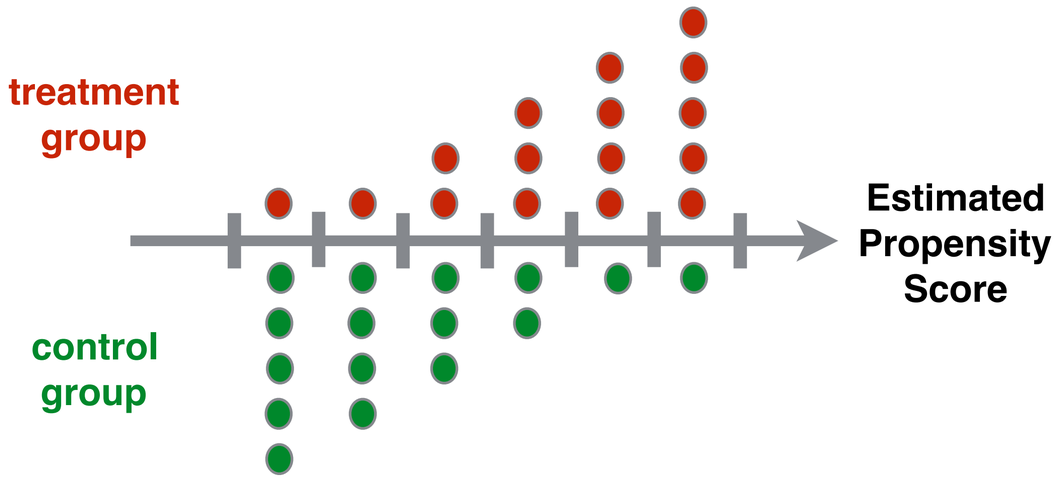
# Propensity Score Matching

**Propensity score (PS)**: the probability of a user to have a given experience.

**PS estimator:** learn an estimator that uses the terms used by a user in the past to determine her propensity to have the experience.

**Match and compare (treatment group)** users with a given experience **with (control group)** users who did not have the experience, but with the same or similar propensity score.

**Determine average outcomes** across user groups with different propensities to have the experience.



# Does It Work?

**Precision:** Are the outcomes we distill likely to be correct across semantic domains?

**Coverage:** How well do we cover the breadth of outcomes likely to occur after an experience?

**Types:** What kinds of outcomes (words) do we recover? Do they capture actual causal relationships?

**Exploration:** Are there other insights we can gain about the outcomes?

# Evaluation Setup

## Selected 39 experiences across 9 semantic domains

- (**should, why, what, why, vs, or**)-search queries from **Bing.com** were semantically annotated
- Selected top decision-queries by popularity within domain

## Twitter user identification

- Firehose Twitter data, March to May 2014

```
1st person personal and possessive pronoun
AND
experience phrase
AND
verbs and verbal phrases
```

- Many report on mental health issues, just a few discuss constructions tasks

| | Category | Experience / Event | Treatment Users | Treatment Msgs |
|---|---|---|---|---|
| Business | Construct. and Maintenan. | *Building stairs | 24 | 8.3K |
| | | *Cleaning countertops | 8 | 3.7K |
| | | *Installing a garbage disposal | 29 | 8.2K |
| | | Painting the deck | 592 | 164K |
| | Financial Services | Owning a good credit card | 2291 | 920K |
| | | Paying credit card debts | 414 | 233K |
| | | Buying life insurance | 1881 | 561K |
| | | Having pension | 2344 | 796K |
| | | *Incorporating one's business | 28 | 9.1K |
| | Investing | Becoming a broker | 855 | 355K |
| | | Investing money | 23981 | 9.8M |
| | | **Total** | 32447 | 12.8M |
| Health | Diseases | Having high blood pressure | 5279 | 1.9M |
| | | Having gout | 364 | 118K |
| | | Having high cholesterol | 1384 | 522K |
| | | Having kidney stone | 727 | 259K |
| | | *Having high triglycerides lev. | 27 | 12K |
| | Mental | Suffering from depression | 25207 | 10.5M |
| | | Suffering from OCD | 11429 | 4.8M |
| | | Being a sociopath | 1491 | 676K |
| | | Being a psychopath | 2895 | 1.3M |
| | | Suffering from anxiety | 53983 | 22.6M |
| | | Suffering from bipolar disord. | 13723 | 6.3M |
| | Pharmacy | Taking Prozac | 617 | 222K |
| | | Taking Lorazepam | 47 | 19.4K |
| | | Taking Promethazine | 242 | 118K |
| | | Taking Tramadol | 397 | 161K |
| | | Taking Xanax | 3300 | 1.4M |
| | | **Total** | 121112 | 50.9M |
| Society | Issues | Losing belly fat | 93 | 24.8K |
| | | Increasing gross income | 135 | 93.2K |
| | Law | Getting divorced | 2717 | 1.2M |
| | | Becoming a notary | 65 | 22.7K |
| | | Applying for social security | 6172 | 2.3M |
| | | Filing for bankruptcy | 921 | 347K |
| | | *Having a living trust | 18 | 7.5K |
| | Relationsh. | Finding true love | 1885 | 654K |
| | | *Recovering after adultery | 9 | 4.4K |
| | | Filing divorce | 422 | 178K |
| | | Dealing with jealousy issues | 789 | 370K |
| | | Changing last name | 1019 | 403K |
| | | **Total** | 14245 | 5.7M |

# **Raw Results**: Examples

Society\Issues\Belly Fat

| Outcome | Count | Absolute Increase | Z-Score |
|---|---|---|---|
| Burn | 156 | 62.2% | 8.96 |
| Ab_workout | 13 | 8.5% | 5.82 |
| Workout_lose | 13 | 8.5% | 5.82 |
| Help_burn | 8 | 11.1% | 5.82 |
| add_video | 26 | 14.0% | 5.75 |
| url_playlist | 26 | 14.0% | 5.75 |
| Fitness | 39 | 18.6% | 5.51 |
| Ab | 43 | 19.1% | 5.51 |
| Playlist_metion | 30 | 15.3% | 5.39 |
| Biceps | 7 | 4.7% | 4.74 |

Health\Diseases\Gout

| Outcome | Count | Absolute Increase | Z-Score |
|---|---|---|---|
| Flare_up | 35 | 4.1% | 12.33 |
| Uric_acid | 27 | 2.9% | 10.36 |
| Uric | 28 | 2.9% | 10.11 |
| Flare | 81 | 4.9% | 9.92 |
| Big_toe | 38 | 2.9% | 9.86 |
| Joint | 301 | 7.2% | 7.22 |
| Aged | 32 | 1.7% | 6.51 |
| Correlation | 45 | 2.8% | 6.11 |
| Bollock | 53 | 2.5% | 5.96 |
| Shite | 108 | 3.4% | 5.93 |

It is easy to observe that the extracted outcomes are **topically related**.

# But, How Well Do We Do?

| Experience | Outcome |
|---|---|
| taking Lorazepam | stage |
| taking Tramadol | pharmacy |
| taking Xanax | liquor store |
| taking Prozac | mental illness |

**Domain**: Health\Pharmacy

# But, How Well Do We Do?

| | Experience | Outcome |
|---|---|---|
| ✗ | taking Lorazepam | stage |
| ✓ | taking Tramadol | pharmacy |
| ✗ | taking Xanax | liquor store |
| ✓ | taking Prozac | mental illness |

**Domain**: Health\Pharmacy

# But, How Well Do We Do?

| | Experience | Outcome | Experience tweet | Outcome Tweet |
|---|---|---|---|---|
| ✘ | taking Lorazepam | stage | I was freaking out earlier.. **took 2 lorazepam** to calm down | lately I'm better at managing my **stage** fright, last talks were much better |
| ✔ | taking Tramadol | pharmacy | the headache from **taking tramadol** is worse than the pain for which I took it | got my **pharmacy** tech license today! to start new job beginning of may |
| ✘ | taking Xanax | liquor store | forensics swap team came & examined it.. I was like k going to jail but **I take my xanax** w me! | the **liquor store** is right next to my building and I'm sooooo lazy |
| ✔ | taking Prozac | mental illness | the fact that she is no longer here still devastates me. the week she passed **I was on a prozac** diet | I'm no doctor but **mental illness** is real & getting through loss is difficult |

**Domain**: Health\Pharmacy

# But, How Well Do We Do?

| | Experience | Outcome | Experience tweet | Outcome Tweet |
|---|---|---|---|---|
| ✔ | taking Lorazepam | stage | I was freaking out earlier.. **took 2 lorazepam** to calm down | lately I'm better at managing my **stage** fright, last talks were much better |
| ✔ | taking Tramadol | pharmacy | the headache from **taking tramadol** is worse than the pain for which I took it | got my **pharmacy** tech license today! to start new job beginning of may |
| ✘ | taking Xanax | liquor store | forensics swap team came & examined it.. I was like k going to jail but **I take my xanax** w me! | the **liquor store** is right next to my building and I'm sooooo lazy |
| ✔ | taking Prozac | mental illness | the fact that she is no longer here still devastates me. the week she passed **I was on a prozac** diet | I'm no doctor but **mental illness** is real & getting through loss is difficult |

**Domain**: Health\Pharmacy

# But, How Well Do We Do?

| | Experience | Outcome | Experience tweet | Outcome Tweet |
|---|---|---|---|---|
| ✔ | taking Lorazepam | stage | I was freaking out earlier.. **took 2 lorazepam** to calm down | lately I'm better at managing my **stage** fright, last talks were much better |
| ? | taking Tramadol | pharmacy | the headache from **taking tramadol** is worse than the pain for which I took it | got my **pharmacy** tech license today! to start new job beginning of may |
| ✘ | taking Xanax | liquor store | forensics swap team came & examined it.. I was like k going to jail but **I take my xanax** w me! | the **liquor store** is right next to my building and I'm sooooo lazy |
| ✔ | taking Prozac | mental illness | the fact that she is no longer here still devastates me. the week she passed **I was on a prozac** diet | I'm no doctor but **mental illness** is real & getting through loss is difficult |

**Domain**: Health\Pharmacy

# But, How Well Do We Do?

| | Experience | Outcome | Experience tweet | Outcome Tweet |
|---|---|---|---|---|
| ✔ | taking Lorazepam | stage | I was freaking out earlier.. **took 2 lorazepam** to calm down | lately I'm better at managing my **stage** fright, last talks were much better |
| ? | taking Tramadol | pharmacy | the headache from **taking tramadol** is worse than the pain for which I took it | got my **pharmacy** tech license today! to start new job beginning of may |
| ? | taking Xanax | liquor store | forensics swap team came & examined it.. I was like k going to jail but **I take my xanax** w me! | the **liquor store** is right next to my building and I'm sooooo lazy |
| ✔ | taking Prozac | mental illness | the fact that she is no longer here still devastates me. the week she passed **I was on a prozac** diet | I'm no doctor but **mental illness** is real & getting through loss is difficult |

**Evaluating the outcomes is challenging.**

**Domain**: Health\Pharmacy

# **Precision**: Are Outcomes Likely To Be Correct?

**Ask crowd-workers**: Someone **having gout** will later on be more likely to talk about **big toe**?

## Pairs of tweets

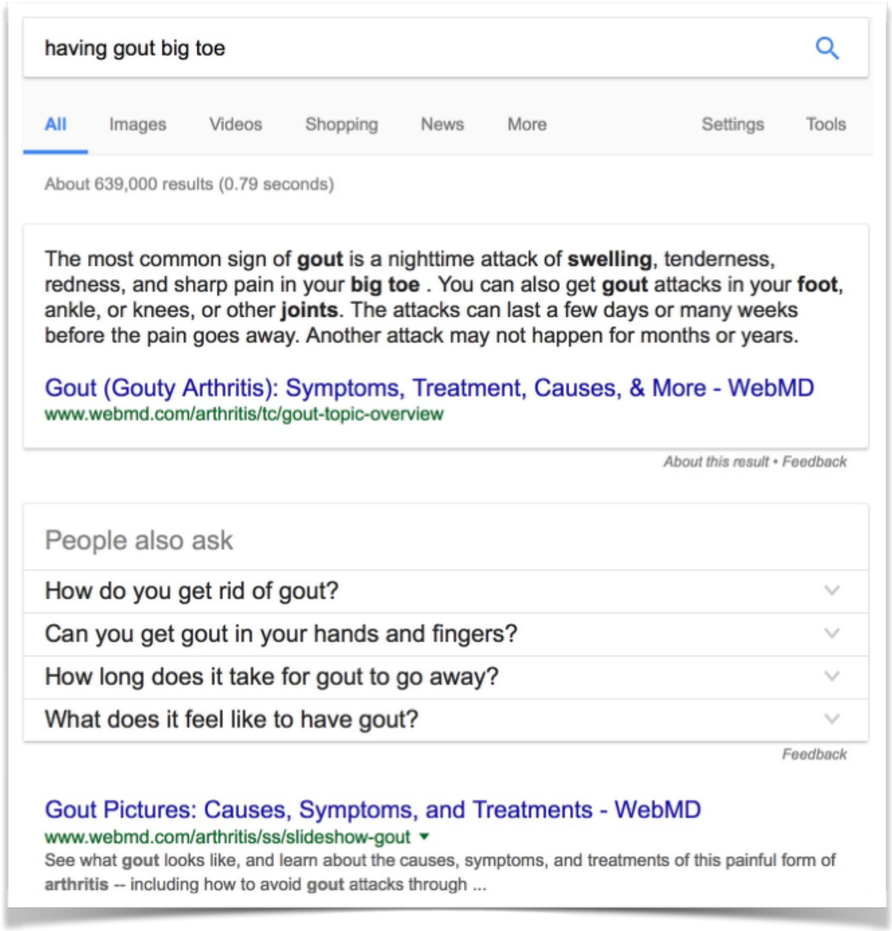
I thought uric acid was just a number. Tsk. I'm too young to **have gout**!

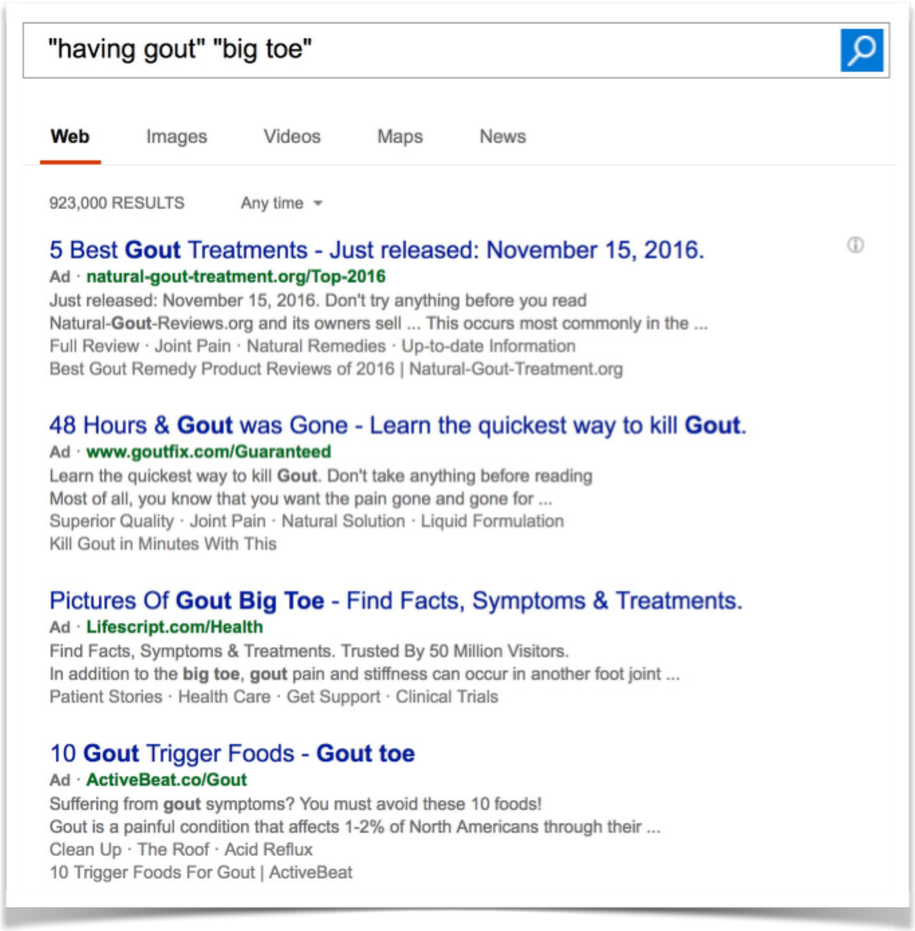
I'm feeling a little something in my right **big toe**, and I have a drinking session tomorrow.

doctor says **I have gout**.. and I wanted to go to a marathon next month

Wow, another attack of gout on my **big toe** feel sick with the pain #gettingold

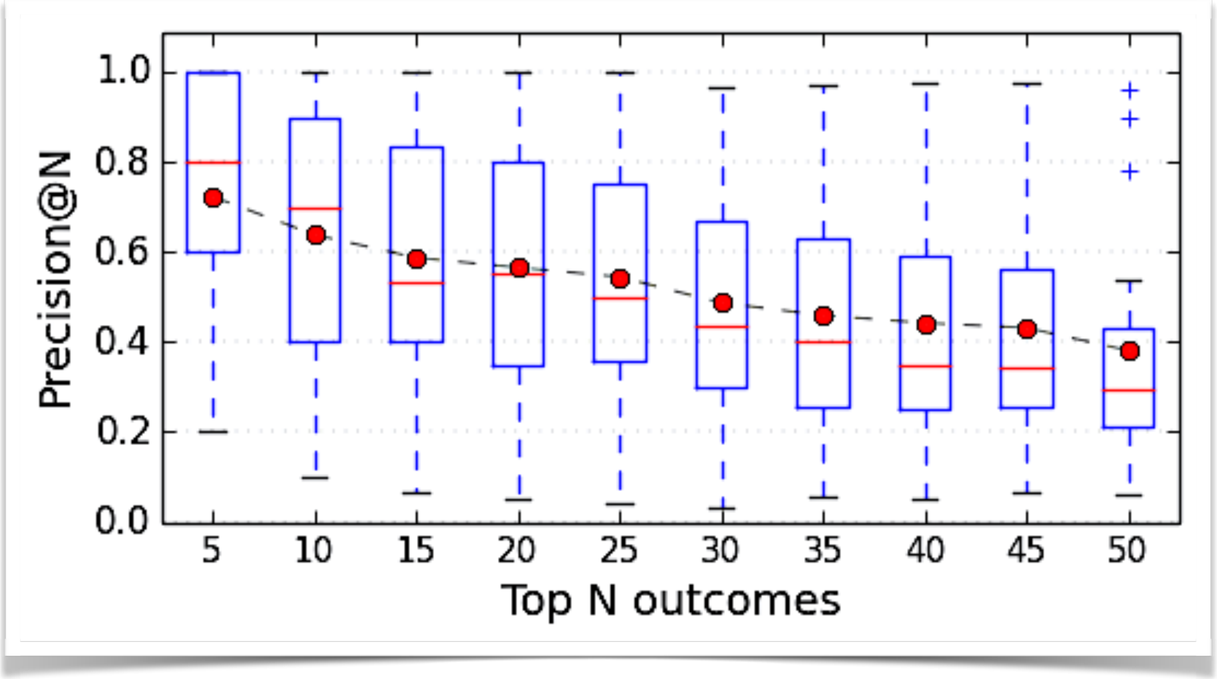
## Search results



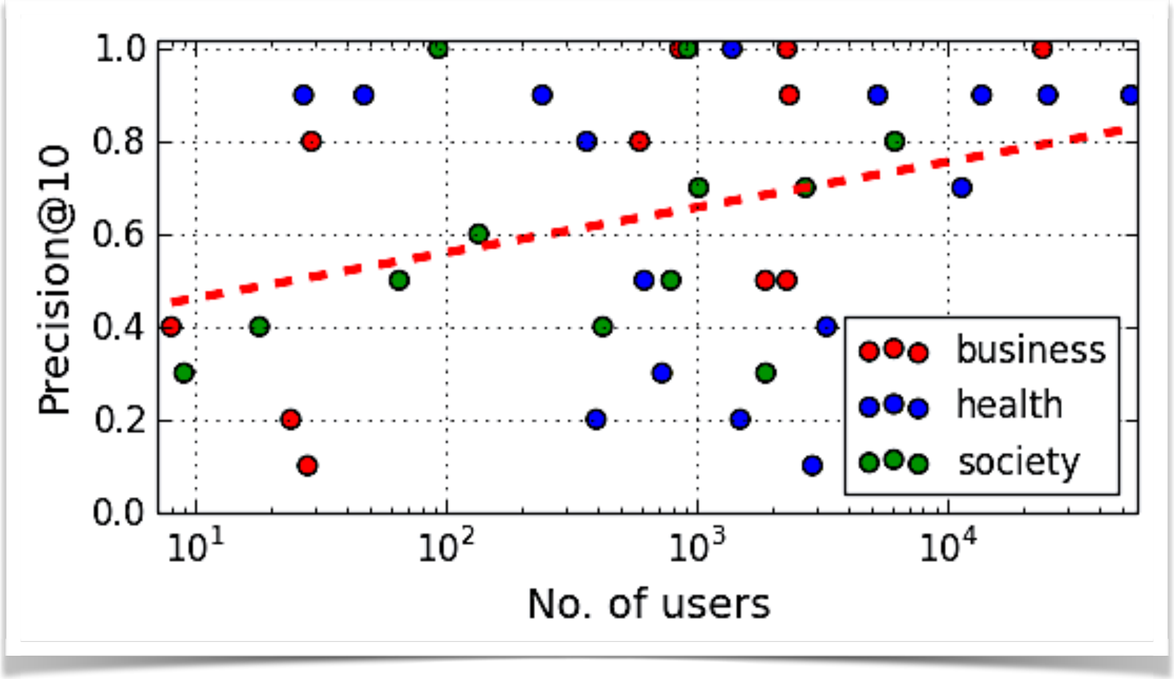



# **Precision**: Are the Outcomes Correct?

Precision @rank by treatment effect




Precision@10 vs. data volume

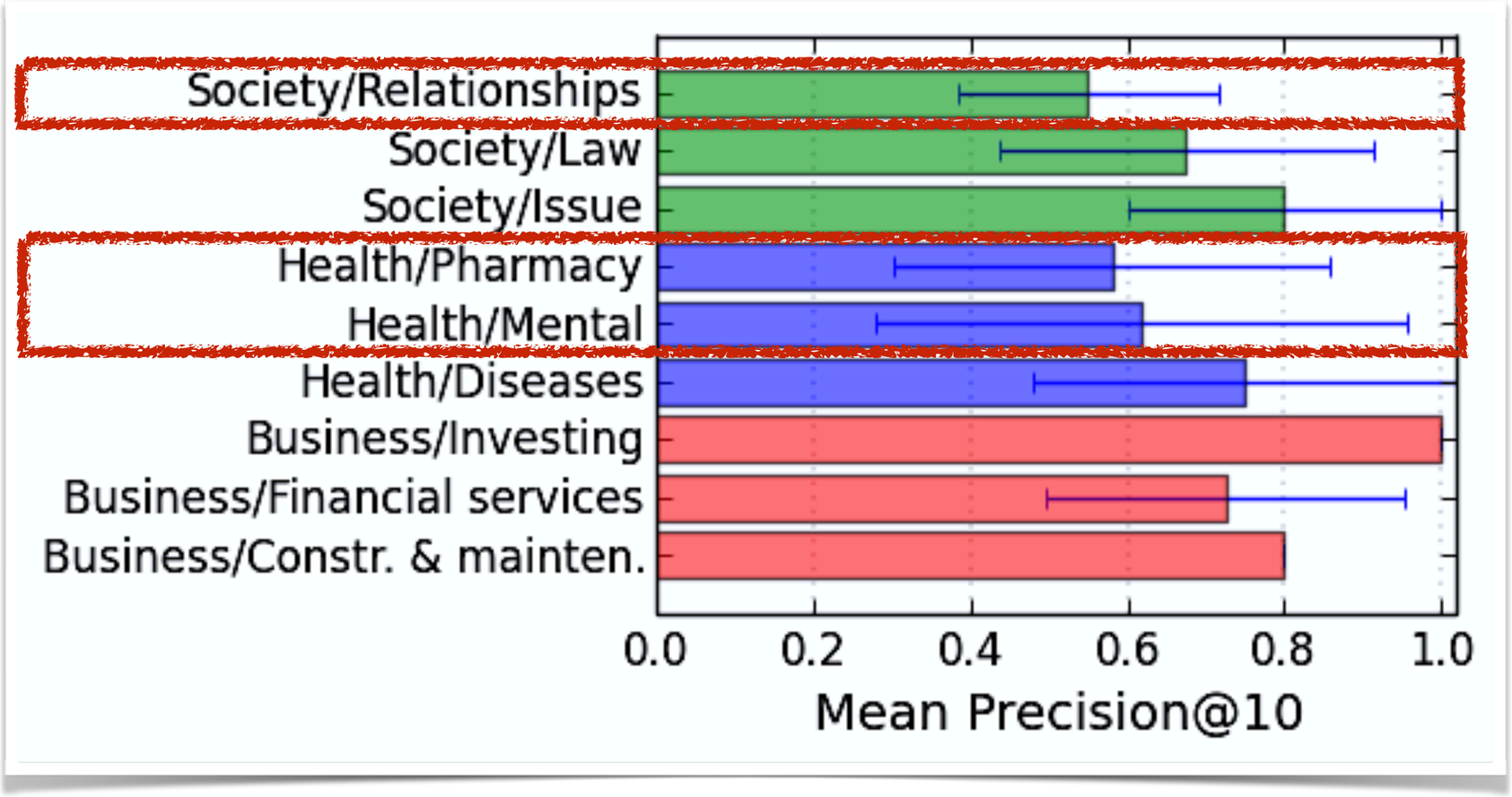



Outcomes with larger treatment effects
are more likely to be perceived as correct.

# **Precision**: Are the Outcomes Correct?

E.g., finding true love, dealing with jealousy

E.g., suffering from bipolar disorder, taking lorazepam

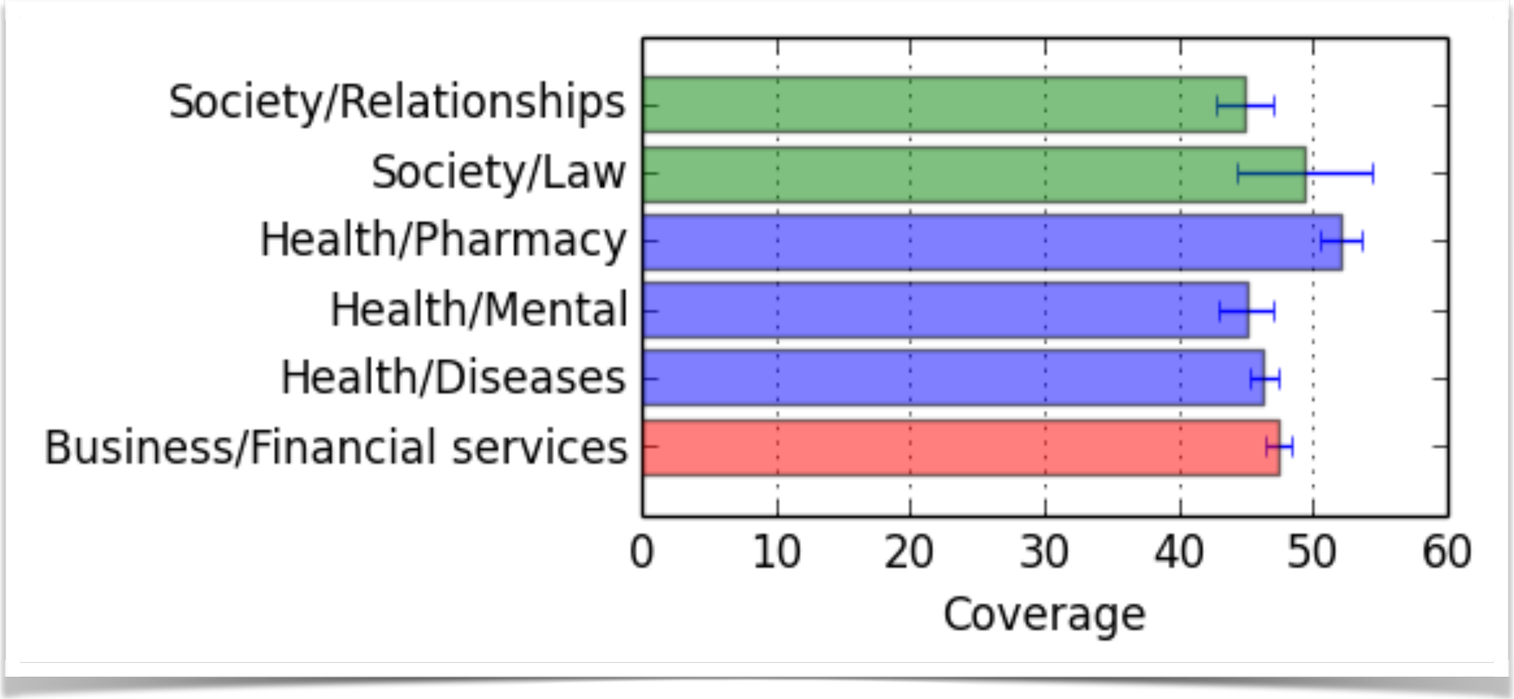



Mean Precision@10 across domains

# **Coverage**: Do We Recover Known Outcomes?

ConceptNet5 (encyclopedia-based knowledge graph)

- coverage of both related concepts and shared relationships e.g.,
  `[Xanax` **`UsedFor`** `anxiety], [divorce` **`Causes`** `drink], [Lorazepam` **`isA`** `drug]`
- higher discovery rates for outcomes with larger treatment effects



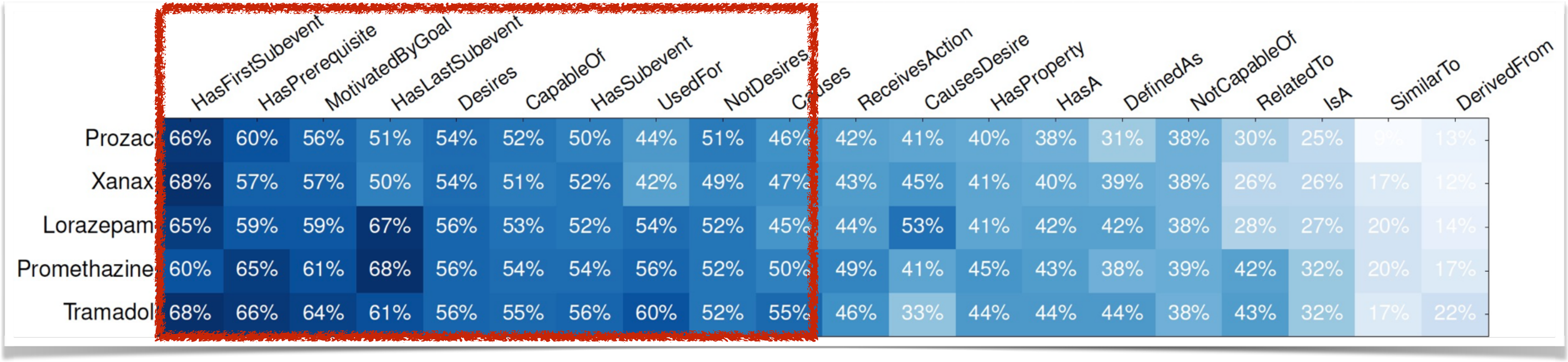
# **Types**: What Kind of Outcomes We Distill?

## Causal-like relations discovered at higher rates

- **implementation steps** e.g., HasSubEvent, HasFirstSubEvent
- **motivations and prerequisites** e.g., MotivatedByGoal, HasPrerequisite
- **implications** e.g., Desires, NotDesires, CapableOf, UsedFor, Causes

## We miss more conceptual or descriptive relationships

- **definitions, alternate names or similar actions** e.g., DefinedAs, RelatedTo, IsA, SimilarTo



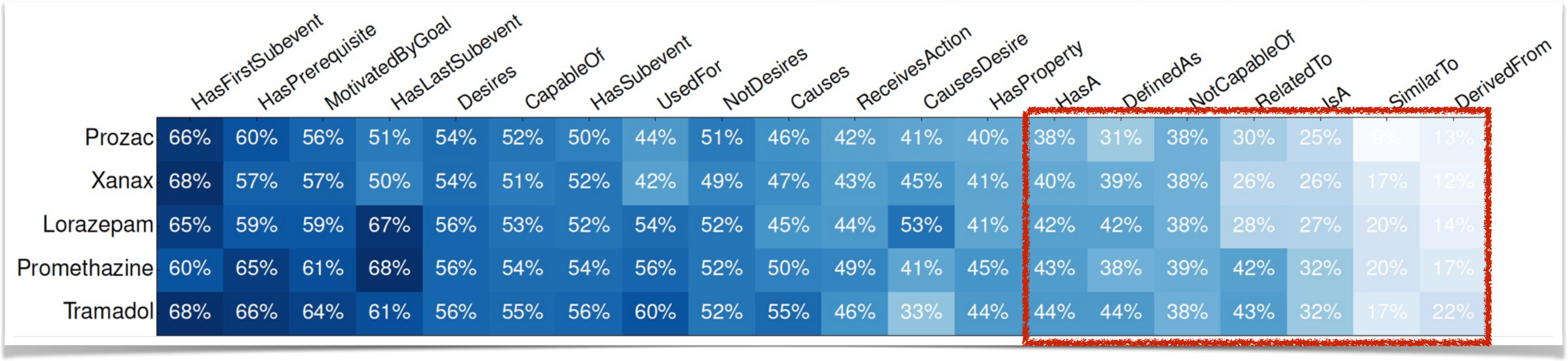
# **Types**: What Kind of Outcomes We Distill?

## Causal-like relations discovered at higher rates

- **implementation steps** e.g., HasSubEvent, HasFirstSubEvent
- **motivations and prerequisites** e.g., MotivatedByGoal, HasPrerequisite
- **implications** e.g., Desires, NotDesires, CapableOf, UsedFor, Causes

## We miss more conceptual or descriptive relationships

- **definitions, alternate names or similar actions** e.g., DefinedAs, RelatedTo, IsA, SimilarTo

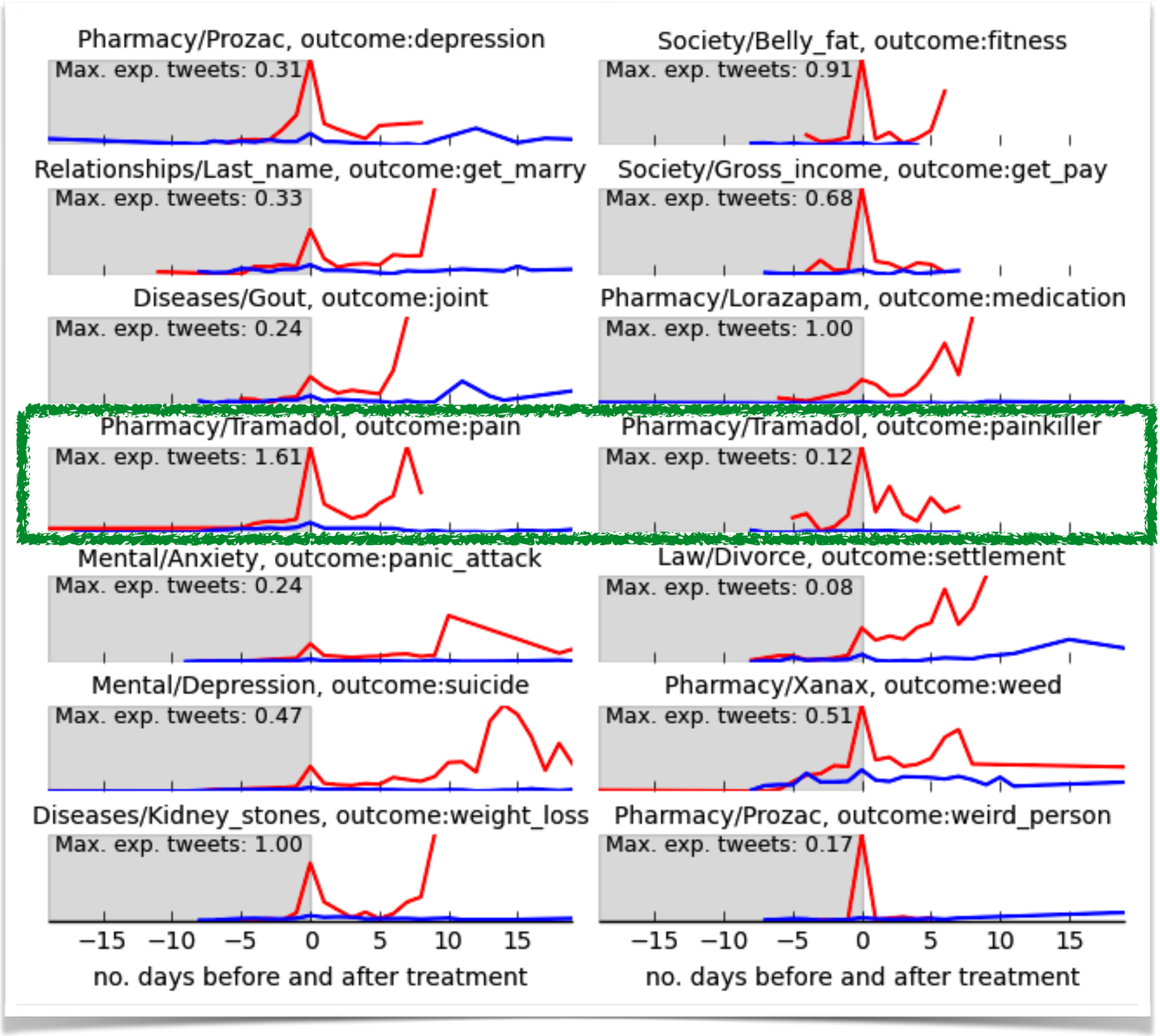


# **Exploration:** When Do They Occur?

## Temporal relationships with the target experience

- likely to occur with the experience
- more likely to occur several days later

## Temporal interplay among the outcomes of an experience




**treatment**

**control**

# Assumptions & Limitations

All confounding variables are included in the observed covariates (the terms used by a user)

➡ Difficult to argue that all relevant aspects of individuals' lives are captured in their Twitter streams

One person outcome is independent of whether another person has the target experience

➡ Conversations may contain e.g., retweets or use same hashtags

+ use of predefined time periods, ignore repeating events, etc.

# Summary

Propensity score analysis can be applied to social media to distill relevant outcomes.

Results are challenging to evaluate, discussion context and topical information helps human judges.

Outcomes with larger treatment effects and larger statistical significance are more likely to be perceived as correct.

More likely to discover causally related concepts & known outcomes.

Applications for individuals, policy makers, and others.

Emails:
alexandra.olteanu@ibm.com
ovarol@indiana.edu
emrek@microsoft.com

Twitter:
@o_saja
@onurvarol
@emrek

KEEP
CALM
AND
QUESTION
EVERYTHING