



Weekendpedia

A Chrome extension for event recommendation

Arthur Li
Insight Data Science

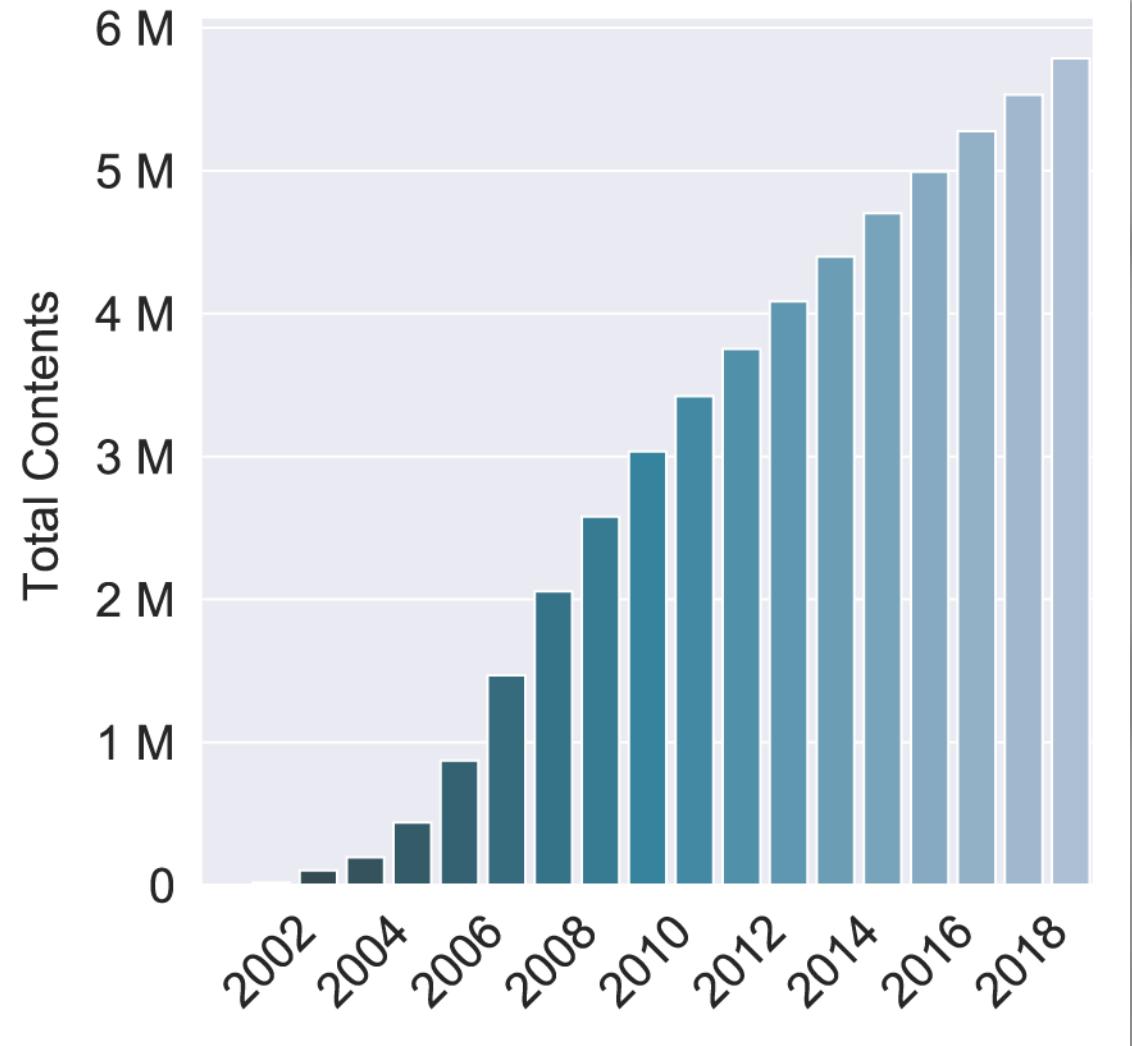
Galleries and museums in NYC



- NYC are has 8 million residents and 60 million visitors/year
- > 200 museums and galleries in NYC, hundreds of cultural events each week

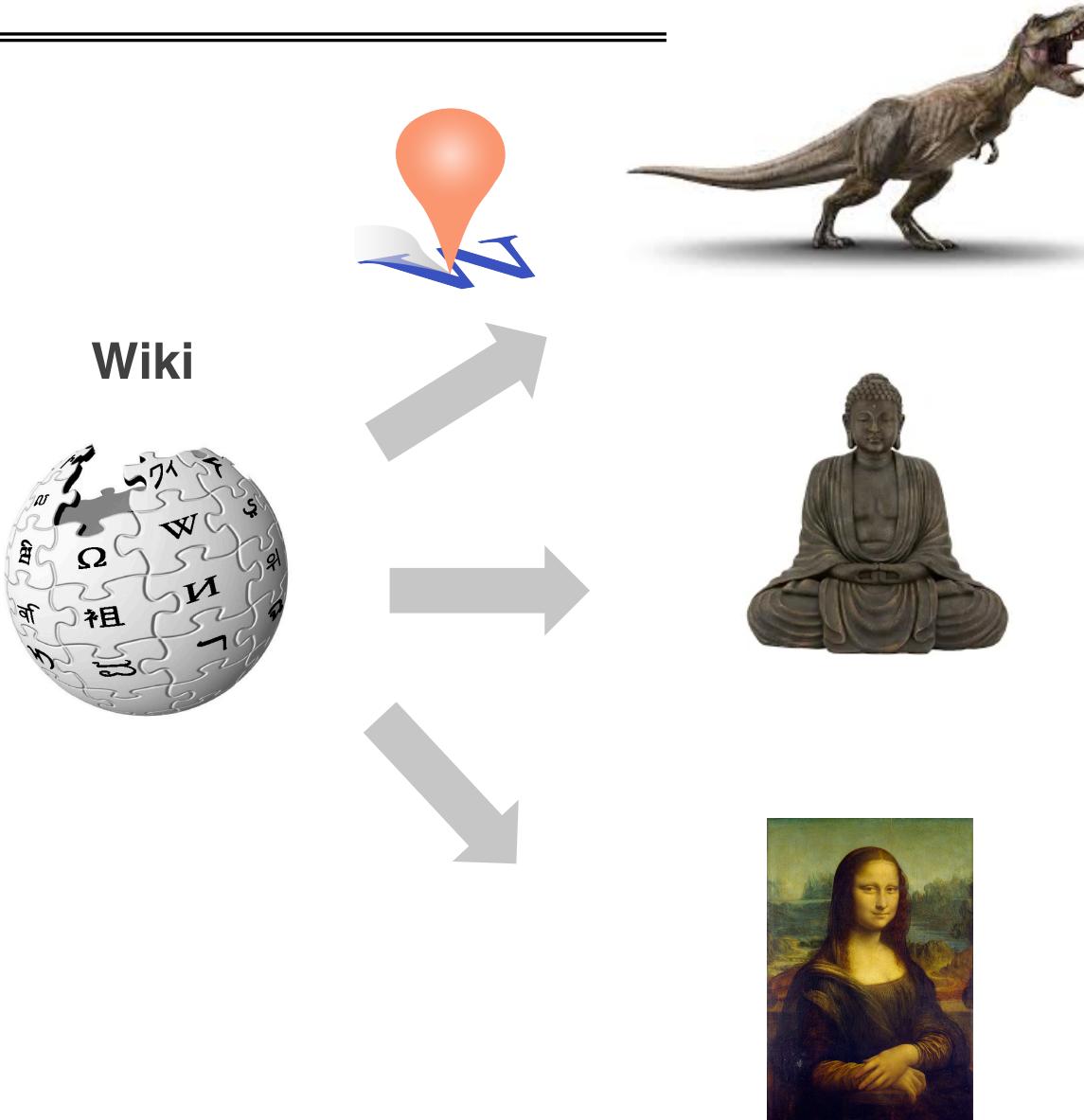
Wikipedia.org

- Wikipedia ranks 5th in global Internet
- Visitors spend on average
- Non-commercial



Matching Wiki articles to events

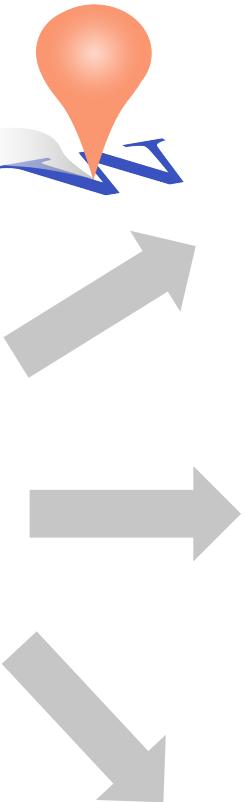
Events



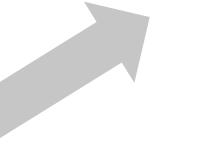
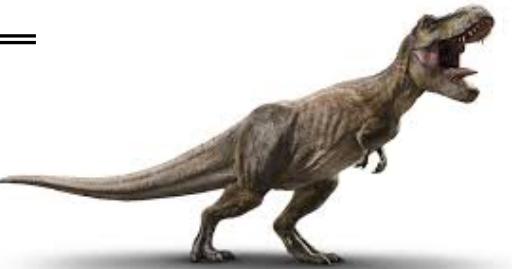
Matching Wiki articles to events



Wiki



Events



Feature engineering with NLP

The collection of Greek and Roman art and ancient sculptures at the Metropolitan Museum is among the most comprehensive in North America.

Feature engineering with NLP

Keywords

The collection of Greek and Roman art and ancient sculptures at the Metropolitan Museum is among the most comprehensive in North America.

Feature engineering with NLP

Feature engineering with NLP

Named entity recognition

The collection of Greek and Roman art and ancient sculptures at the Metropolitan Museum is among the most comprehensive in North America.

Feature engineering with NLP

Named entity recognition

The collection of Greek and Roman art and ancient sculptures at the Metropolitan Museum is among the most comprehensive in North America.

Feature engineering with NLP

Named entity recognition

The collection of Greek and Roman art and ancient sculptures at the Metropolitan Museum is among the most comprehensive in North America.

Part of speech tagging

The collection of Greek and Roman art and ancient sculptures at the Metropolitan Museum is among the most comprehensive in North America.

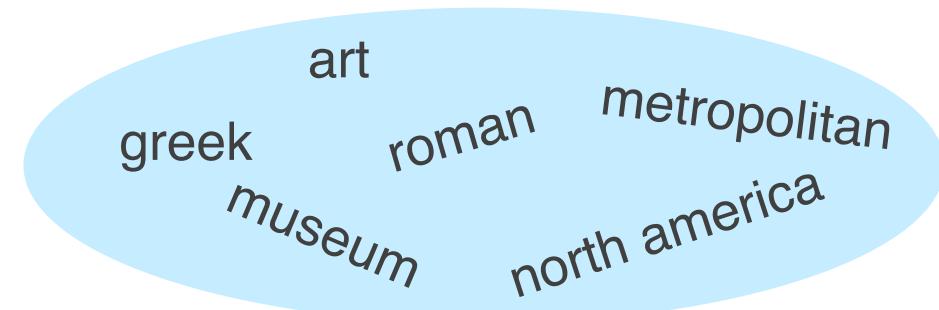
Feature engineering with NLP

Named entity recognition

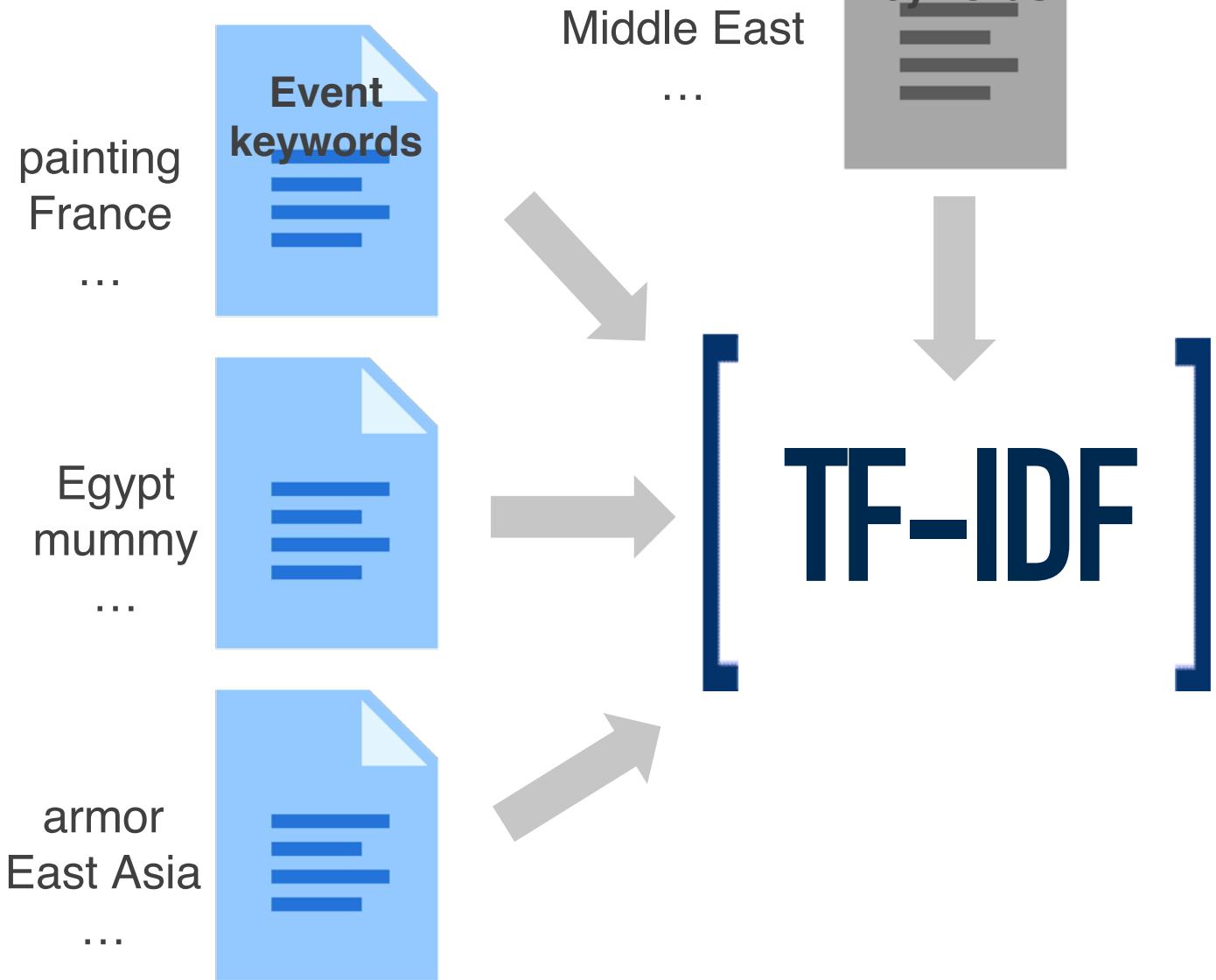
The collection of Greek and Roman art and ancient sculptures at the Metropolitan Museum is among the most comprehensive in North America.

Part of speech tagging

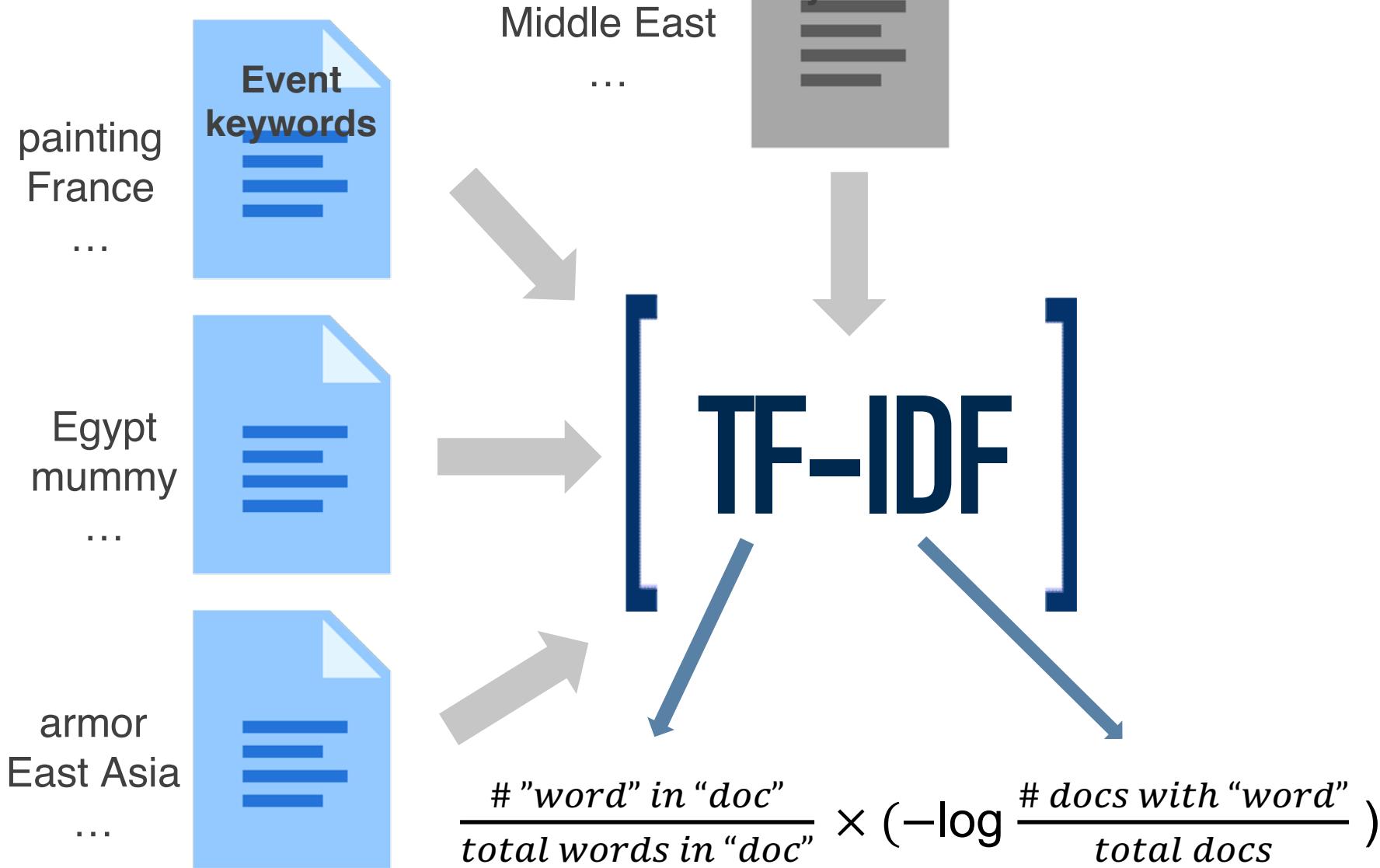
The collection of Greek and Roman art and ancient sculptures at the Metropolitan Museum is among the most comprehensive in North America.



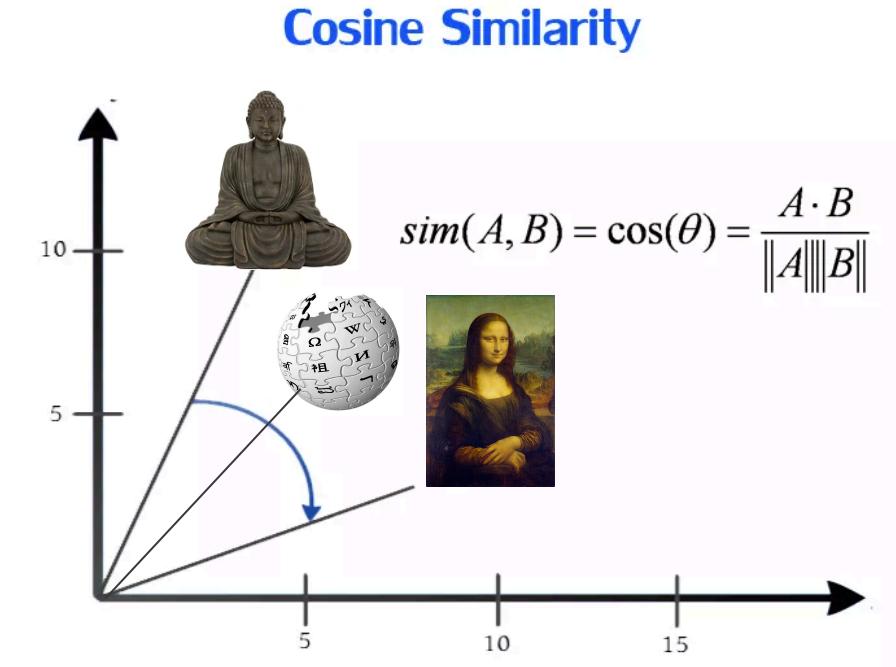
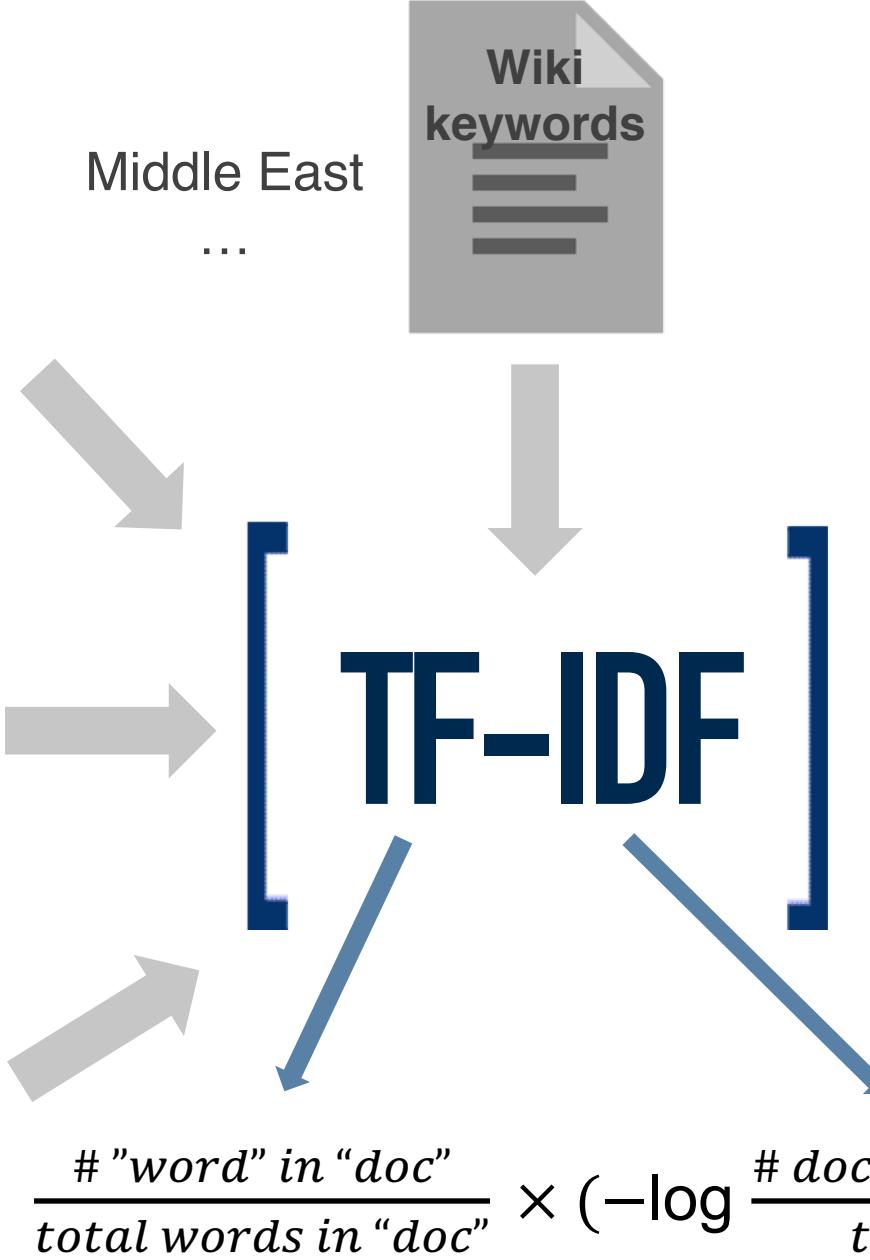
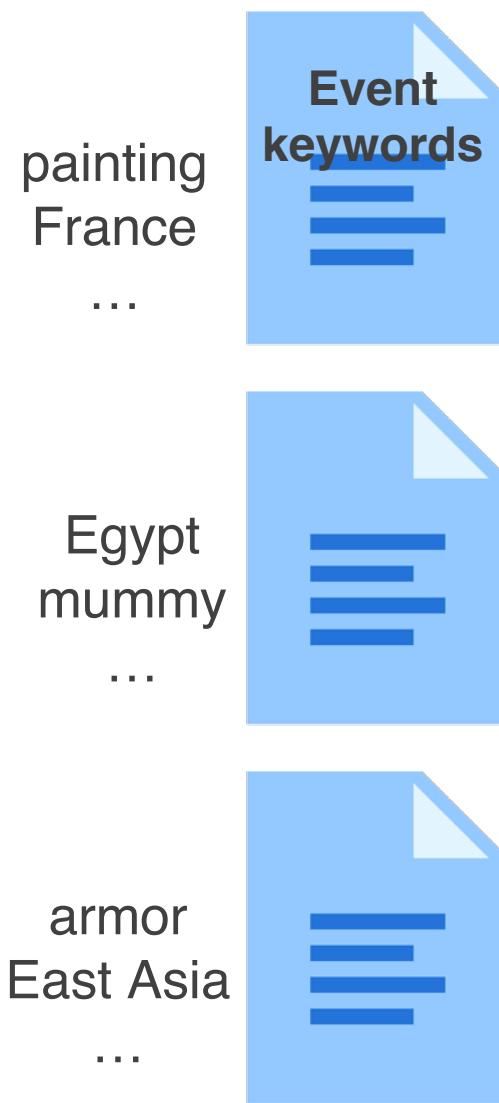
TF-IDF

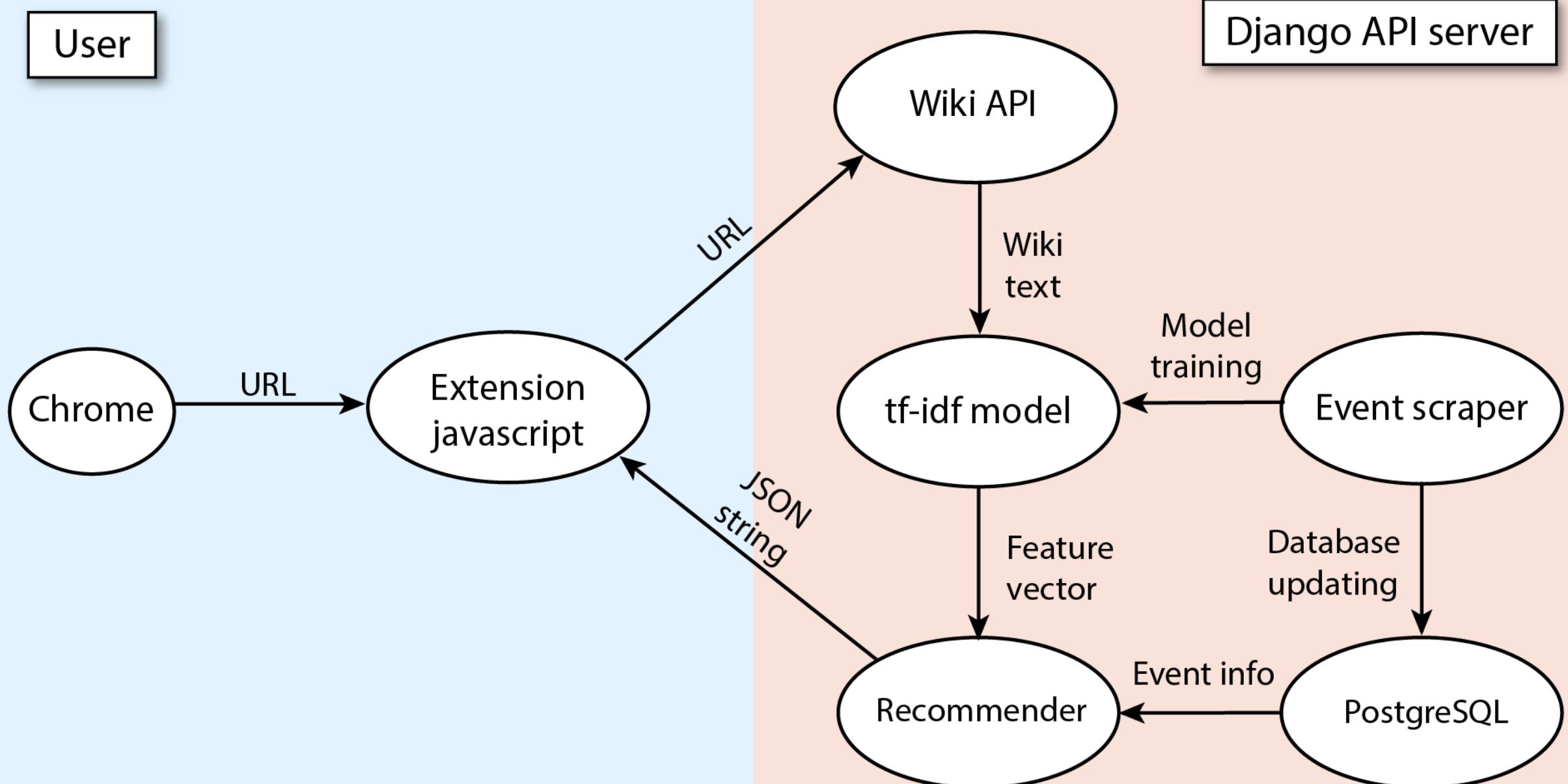


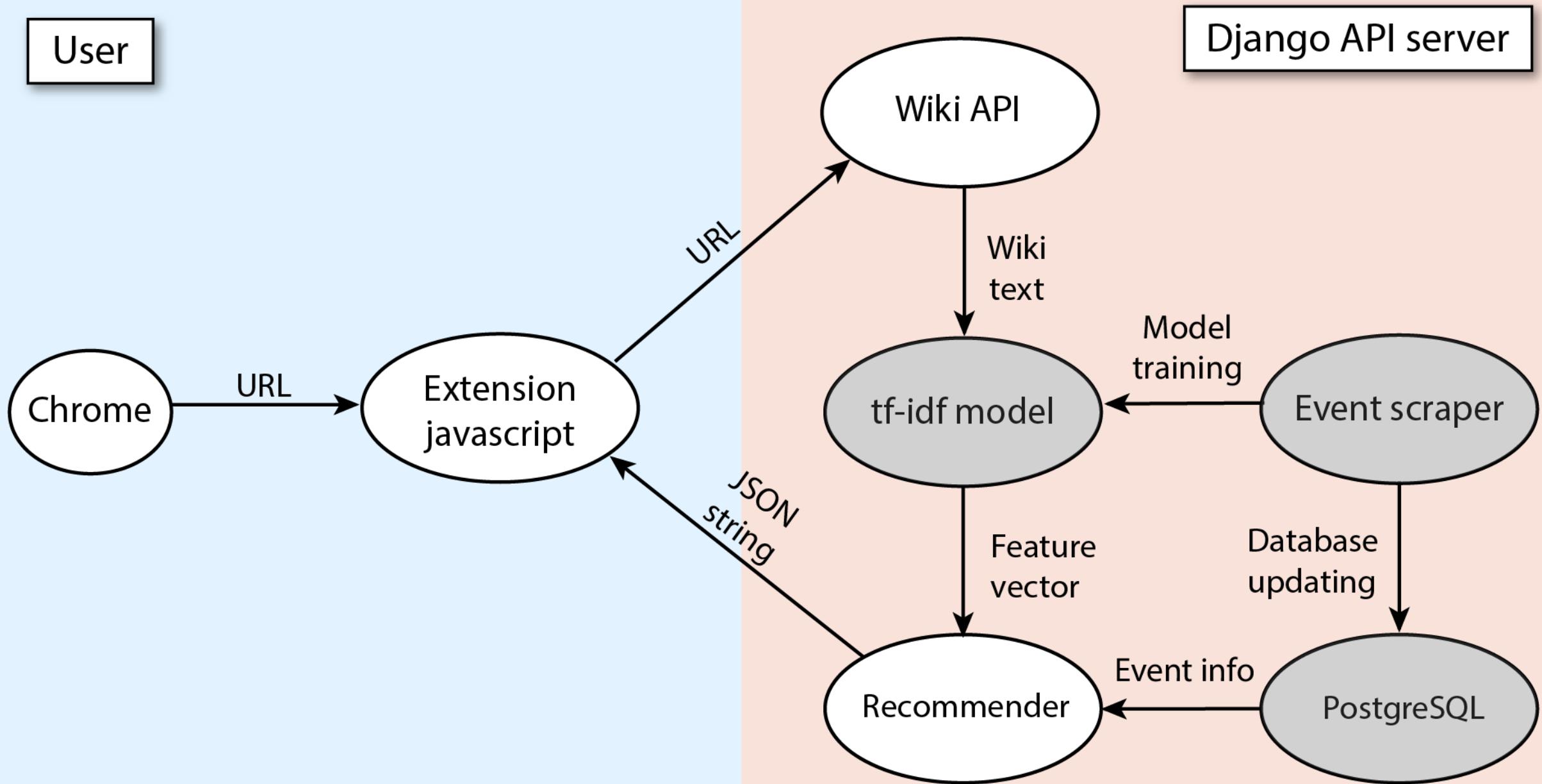
TF-IDF

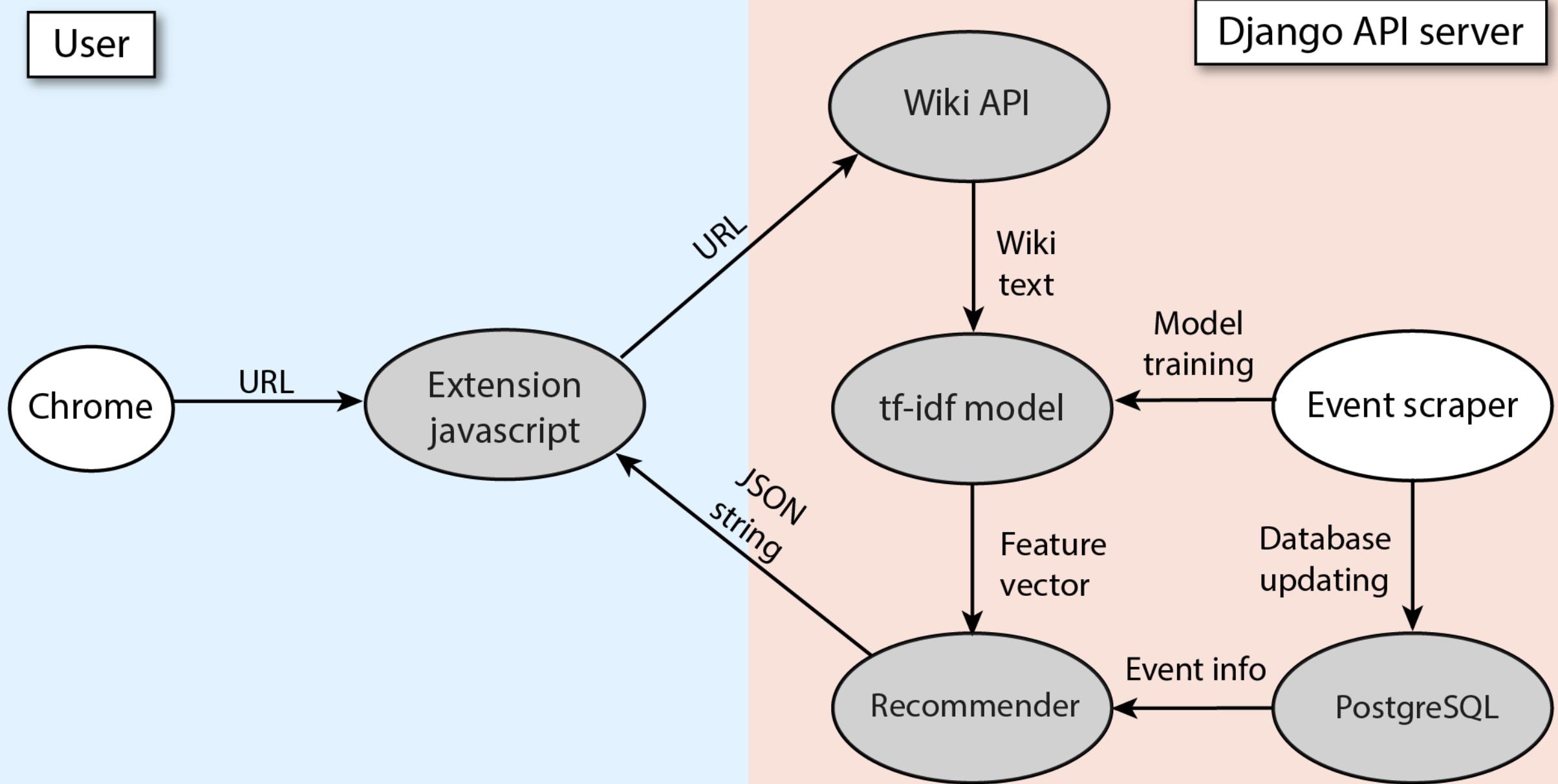


TF-IDF









Validation



Validation

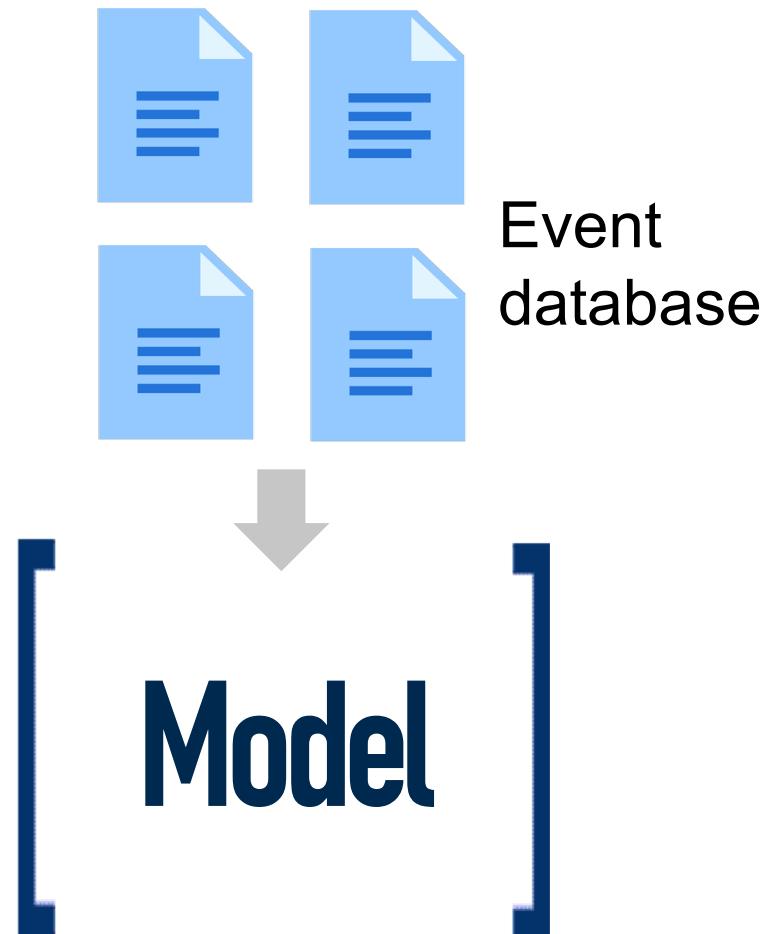
Manually tag events

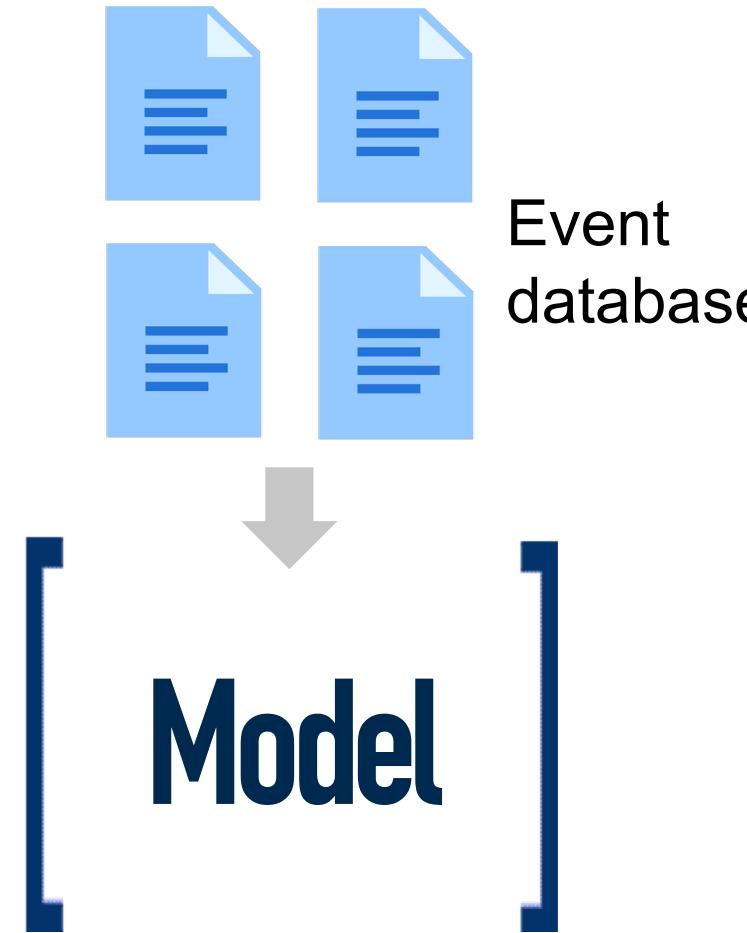
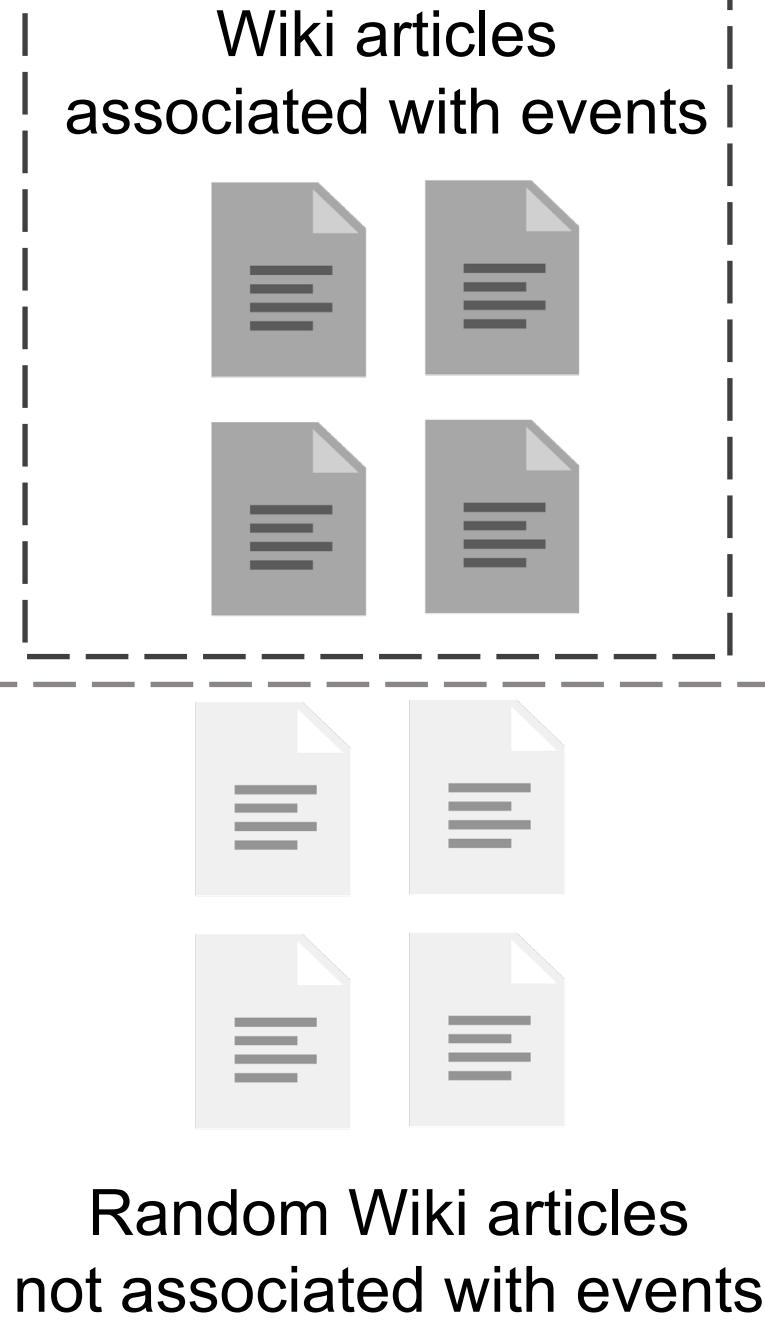
Wiki

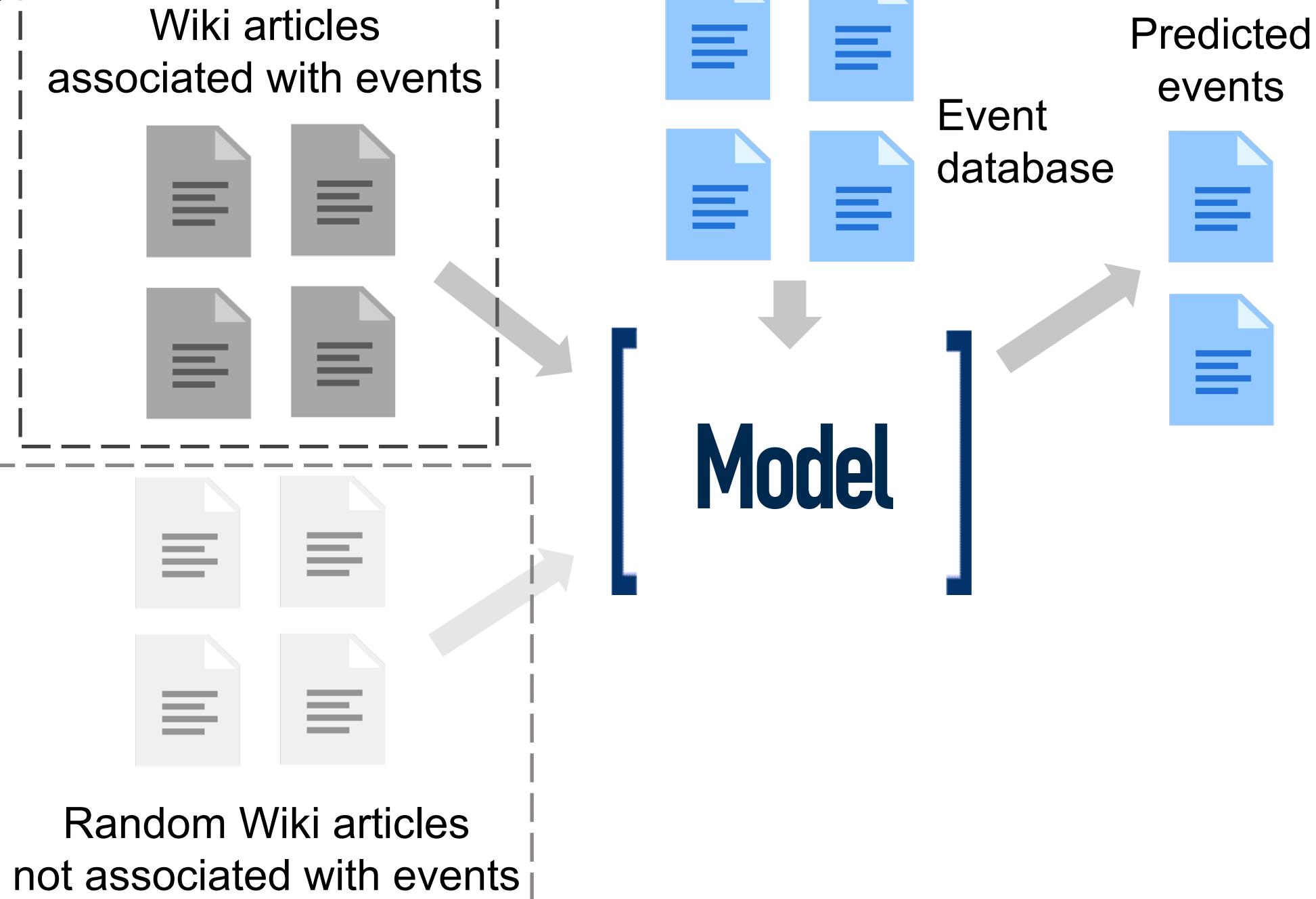


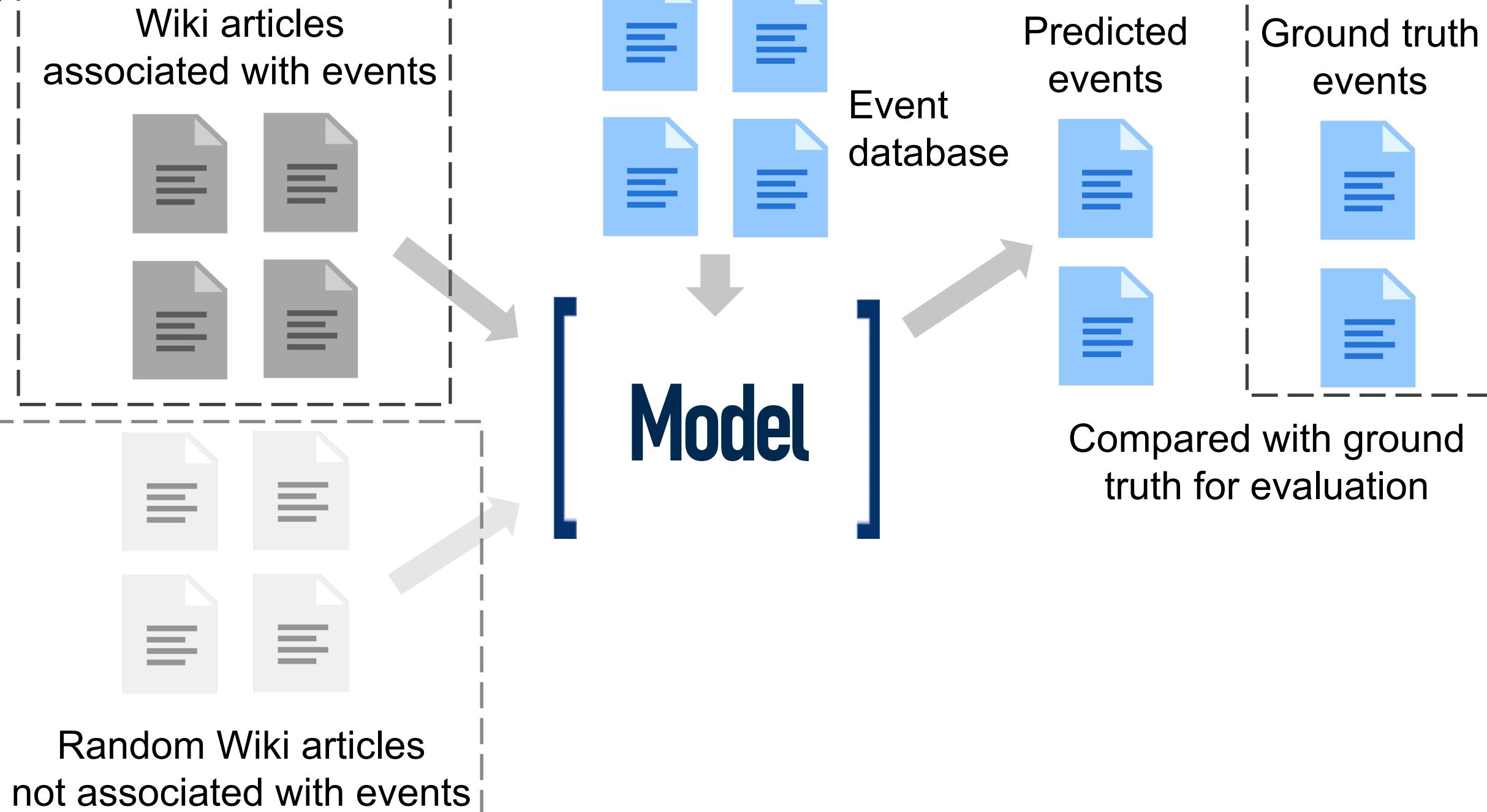
Events

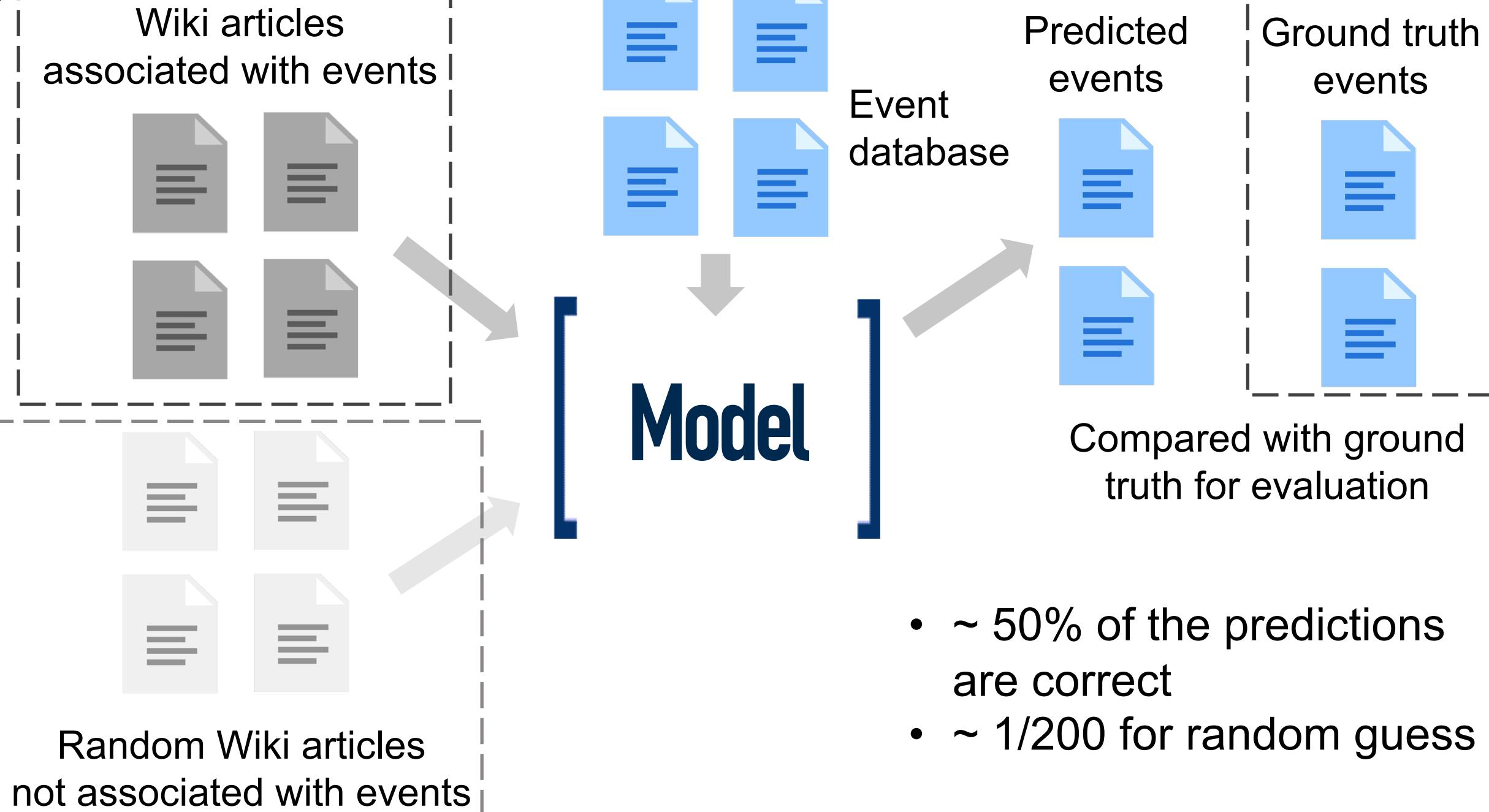








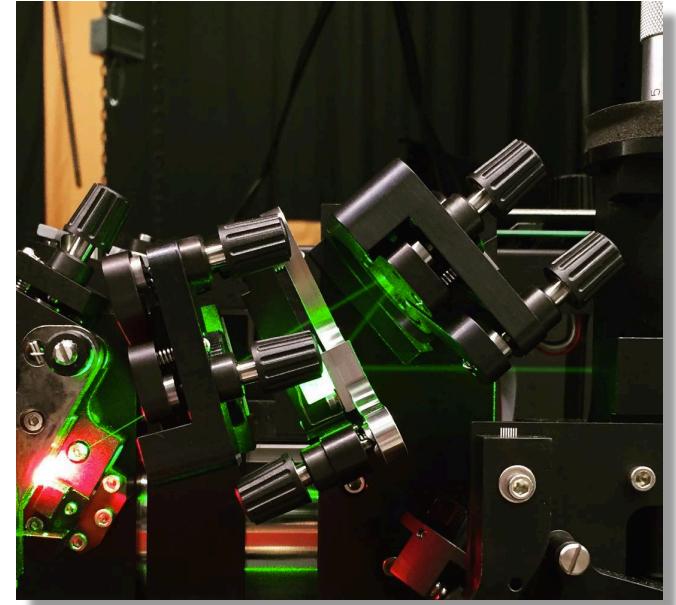




Demo

<https://youtu.be/TJuDtRmjgR0>

Arthur Li, PhD



My PhD research was on experimental material physics, and searched for signals of novel physical phenomena.

I worked with lasers.

Data sources

- nyc.com
- 200 cultural events scraped for one month

The screenshot displays a dark-themed website interface featuring two event cards.

Bindlestiff Family Cirkus: Brooklyn Abridged

JUN 06 THU

12:00 AM at Connelly Theater - 220 E 4th St

Bindlestiff Family Cirkus, New York City's acclaimed variety arts ensemble, returns to the Connelly Theater (220 East 4th Street, New York City) with an all-new production, "Brooklyn Abridged," which combines history, humor and heart with an eclectic... *read more*

Occurs 3 more times through Jun 09

Smith Haven Mall Carnival

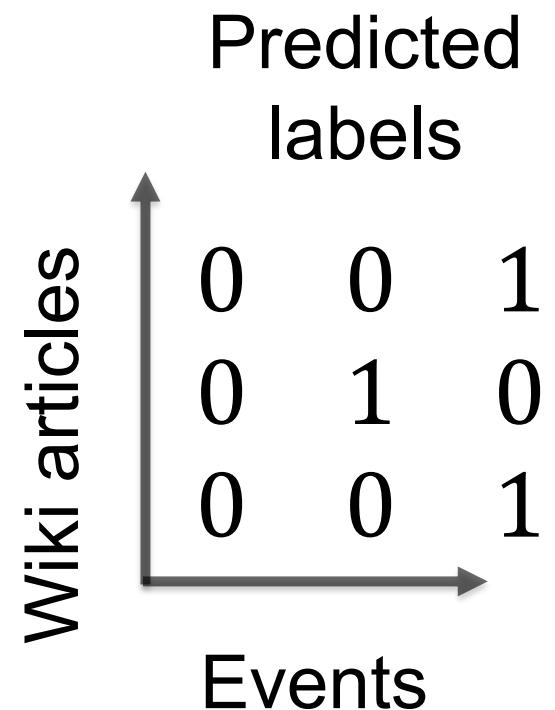
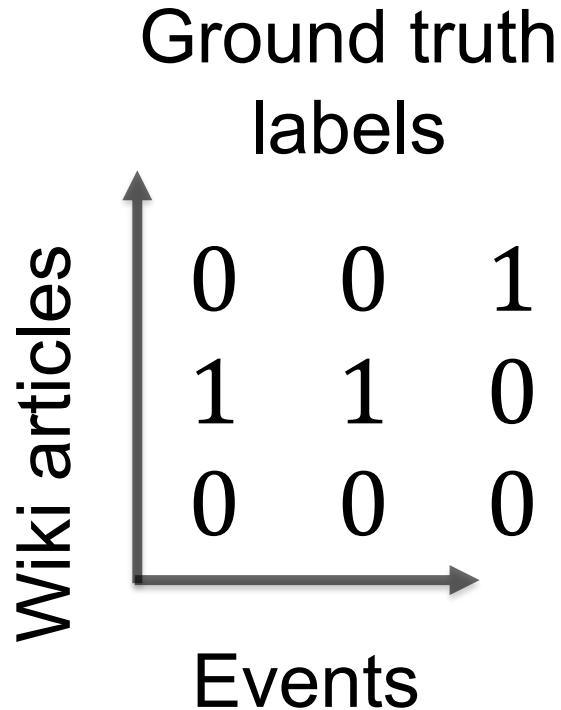
JUN 06 THU

12:00 AM at Smith Haven Mall

Enjoy family-friendly entertainment with daredevil circus shows, midway rides and games for all ages, and favorite fair food. There is no charge for carnival parking, admission and Victoria Circus shows featuring exciting high-wire and aerial acts... *read more*

Get Tickets

Get Tickets



Confusion matrix

Actual	Predicted	
	TP = 211	FP = 235
FN = 186		TN = 124547

Precision = 0.47; recall = 0.53, F1 = 0.50
150 events in total;
400 wiki articles associated with events;
600 random wiki articles.