# Entity Linking for Fun and Profit

Vaughan Kitchen

May 29, 2019

# Structure

- Background
- Techniques
- Maybe other stuff
- Maybe related Topics

# Background

### Why

Bag of words models are only so good. Wouldn't it be nice if we could search semantically?

### What

User searches for "michael motown first band"
Entities "Michael Jackson", "Motown Records", "The Jackson 5ive"

### How

*entity-by-entity linking*, *collective entity linking*, *knowledge base*, *graph theory*, *neural networks*

# Entity-by-Entity Linking

- One word at a time
- Simple
- Fast
- Poor results
- Comes from DB theory

# Collective Entity Linking

- ▶ Uses context
- ▶ Link all words?
- ▶ How to detect entities?
- ▶ How to make fast?
- ▶ Neural network it?

# How to Knowledge Base

- Use Wikipedia

# Queries and Documents

- Queries don't use the same language as documents
- "michael motown first band"
- Can they be linked the same way?
- Queries often short. Not enough context
- Results for partial disambiguation
- "michael singer"
- "George Michael", "Michael Jackson", "Michael Bublé", "Michael Alan Singer (Author)"
- Best candidate?
- "clinton falls asleep" $\rightarrow$ "clinton falls | asleep", "clinton | falls asleep"

# Questions?