

Crowd Surfing Tweets

08/05/2017 - Big Data Day LA

Kivanc Yazan - @kyzn



Bogazici University Computer Engineering Department

<https://kyzn.org/slides/BDD/2017/twitter.pdf>

Bio

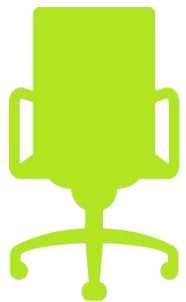
- Undergrad at Bogazici University, 2015
 - “SOS-LAB” Complex Systems Lab
 - Social Networks & Semantic Web

Bio

- Undergrad at Bogazici University, 2015
 - “SOS-LAB” Complex Systems Lab
 - Social Networks & Semantic Web
- Grad at UC Santa Cruz, 2016

Bio

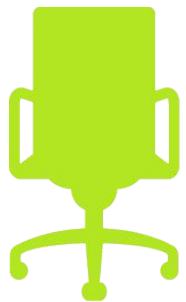
- Undergrad at Bogazici University, 2015
 - “SOS-LAB” Complex Systems Lab
 - Social Networks & Semantic Web
- Grad at UC Santa Cruz, 2016
- Software Engineer @ ZipRecruiter



ZipRecruiter[®]

Bio

- Undergrad at Bogazici University, 2015
 - “SOS-LAB” Complex Systems Lab
 - Social Networks & Semantic Web
- Grad at UC Santa Cruz, 2016
- Software Engineer @ ZipRecruiter



ZipRecruiter[®]



Perl
mongers

NoSQL?

- Not only NoSQL
- MySQL etc.
- Use case

Microblogs



Tweets	Following	Followers	Likes	Lists
677	302	485	159	2

Big Data Day LA 2017
@BigDataDayLA

The only free Big Data Conference in Los Angeles #BigDataDayLA! Our 5th annual event is on Aug 5, 2017 at University of Southern California. Stay tuned!

Los Angeles, CA
bigdatadayla.com
Joined April 2015

[Tweet to Big Data Day LA 2...](#)

114 Photos and videos

Tweets **Tweets & replies** **Media**

Big Data Day LA 2017 Retweeted

 **amyluca** @amyluca · Jul 26
Be sure to check out @theAmplify 'a own VP of Tech @BigDataDayLA on August 5th @USC He will be speaking about all things #emoji #socialmedia

Jeff Weintraub @jeff_weintraub
@BigDataDayLA can't wait! A solid #DataScience line-up it is 🔥🔥🔥
twitter.com/bigdatadayla/s...

3 2

 **Big Data Day LA 2017** @BigDataDayLA · Jul 25
Hadoop/Spark/Kafka Track is packed with sessions on operationalizing data science to accelerating real time decisions systems. Take a look!

Following Events



Wimbledon ✓
@Wimbledon



Follow

Will it result in a history making 8th
[#Wimbledon](#) title?



More keywords..

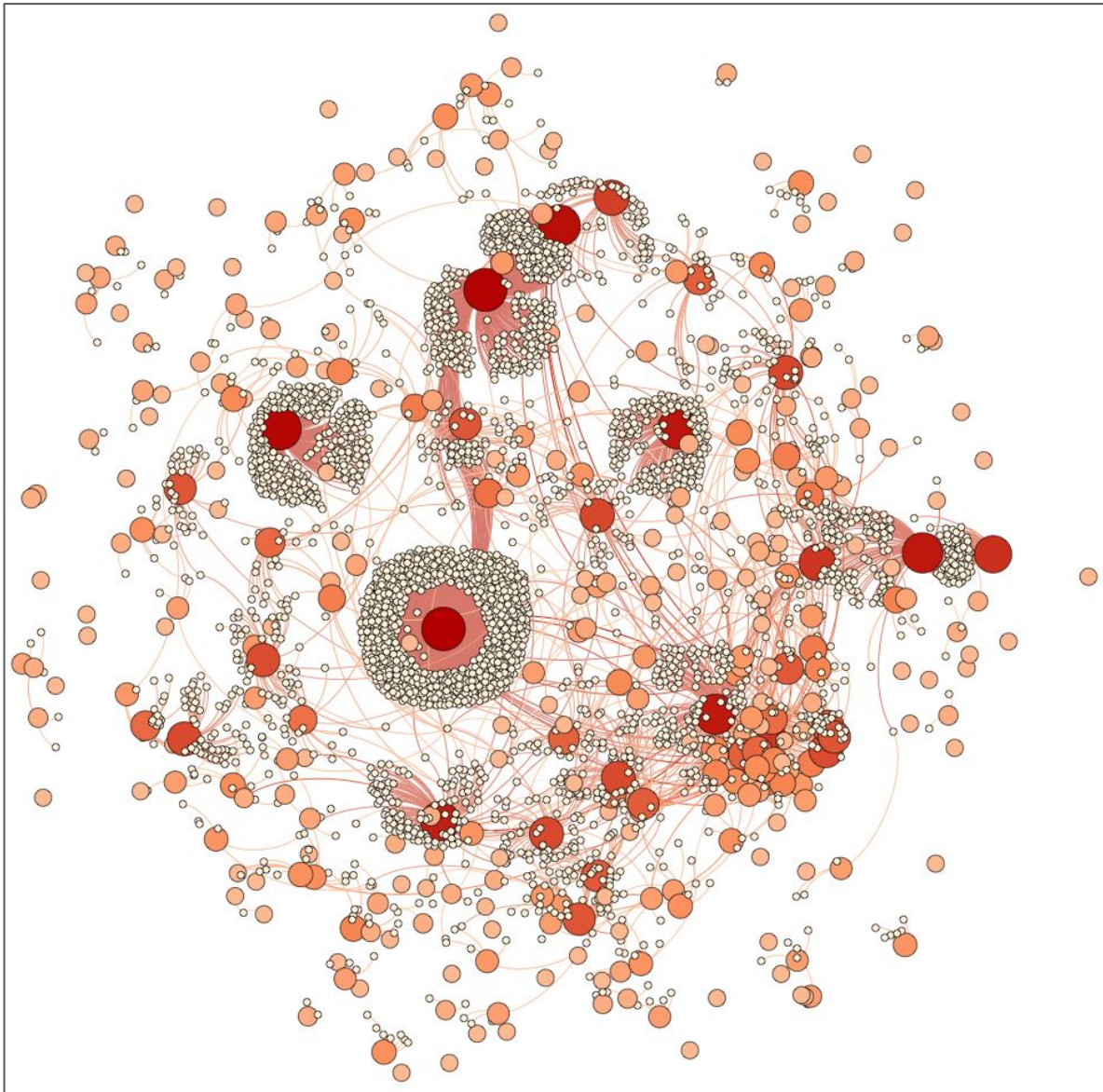
Nadal: "I lost & I'm sad for that. But that's sport. There are good moments & bad moments, today is a bad moment"



Traditional Search/Stream

- Give keywords
- Get results
- You will miss:
 - Unexpected events
 - Spontaneous hashtags
- Unless you update your query

Also.. Networks



or maps



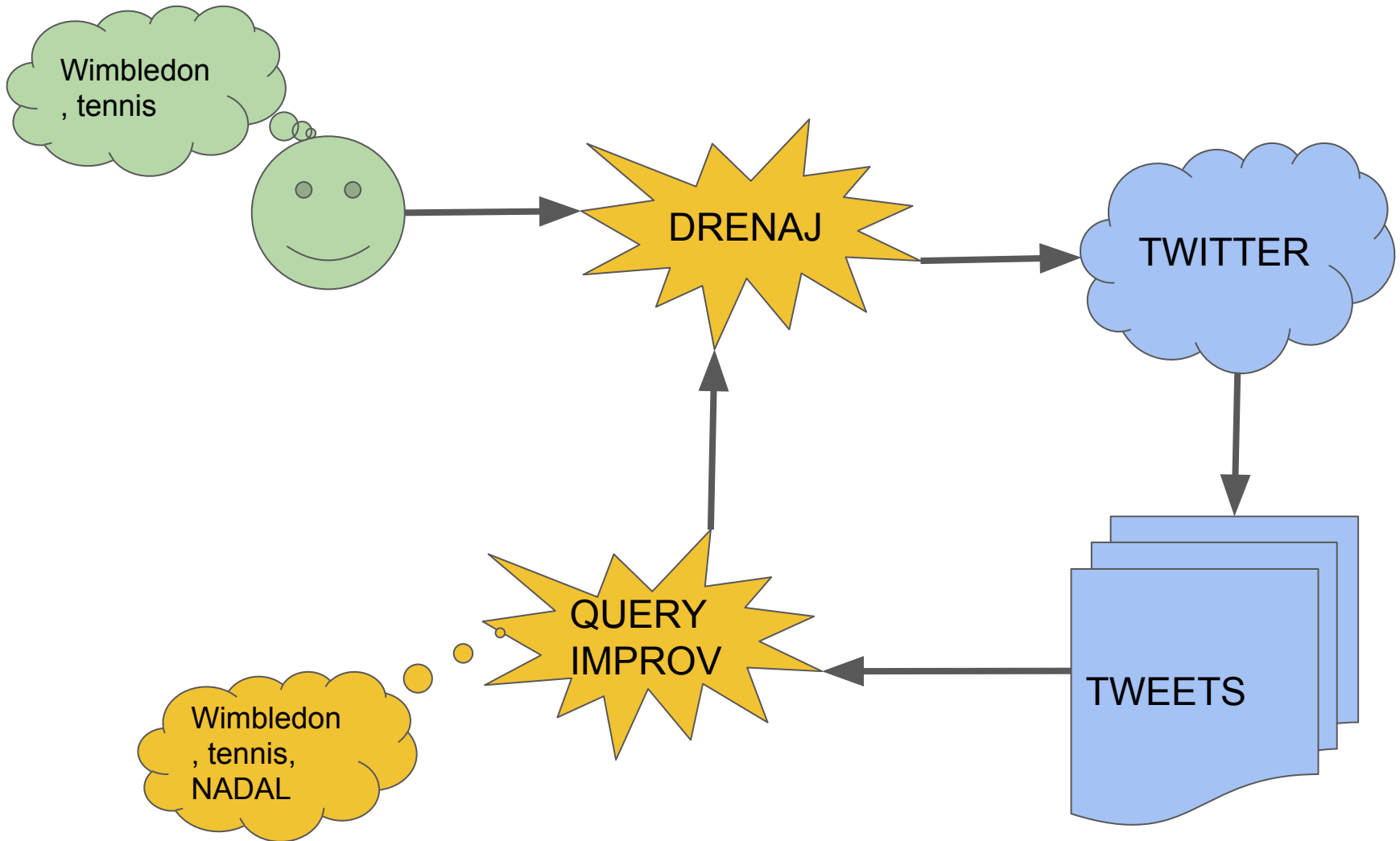
Drenaj: Distributed Streamers

- Decentralized clients streaming multiple searches
- Central storage (NoSQL)
 - RDB might break on new features
- Easy to work together / reusable data

Drenaj: Distributed Streamers

- Twitter gives you a fraction of data
 - More clients *might* mean more data
- Breaking changes can be done on APIs (2013)
 - Distributing update to all clients solves issue

QueryImprov



Updating query

- User starts a query with keywords
- Frequent words become “keyword candidate”
- Popular for a while, gets added to query
- Unpopular keywords gets dropped
- Report every 15-30-60min, rebuild the query

Eg: Debates

hillary clinton, hillary, donald trump, debate, trump, clinton

hillary clinton, hillary, donald trump, debate, trump, clinton

trump, clinton

clinton, trump

Eg: Grexit

varoufakis, referendum, minister, finance, vote,
greek, greece

minister, finance, vote, greek, greece

finance minister, vote, greek, greece

vote, greek, greece

vote, greek, greece

vote, greek, greece

greek, greece

greek, greece

Eg: Grexit

varoufakis, referendum, minister, finance, vote,
greek, greece



Yanis Varoufakis ✓

@yanisvaroufakis

Follow



Minister No More!

Minister No More!

The referendum of 5th July will stay in history as a unique moment when a small European nation rose up against debt-bondage.

yanisvaroufakis.eu

10:31 PM - 5 Jul 2015

16,864 Retweets **9,002** Likes



5.2K

17K

9.0K



Numbers

- Near 60M tweets in 2 languages
- 17.2M original content + 42M retweets
- ~20 searches about
 - Politics, election
 - Sporting events
 - Products etc.
- Paper WIP

Thanks

- Dr. Suzan Uskudarli (Adviser)
- Onur Gungor (Drenaj)
- SOSLAB

Questions?