



data is the



data is the **new bacon**

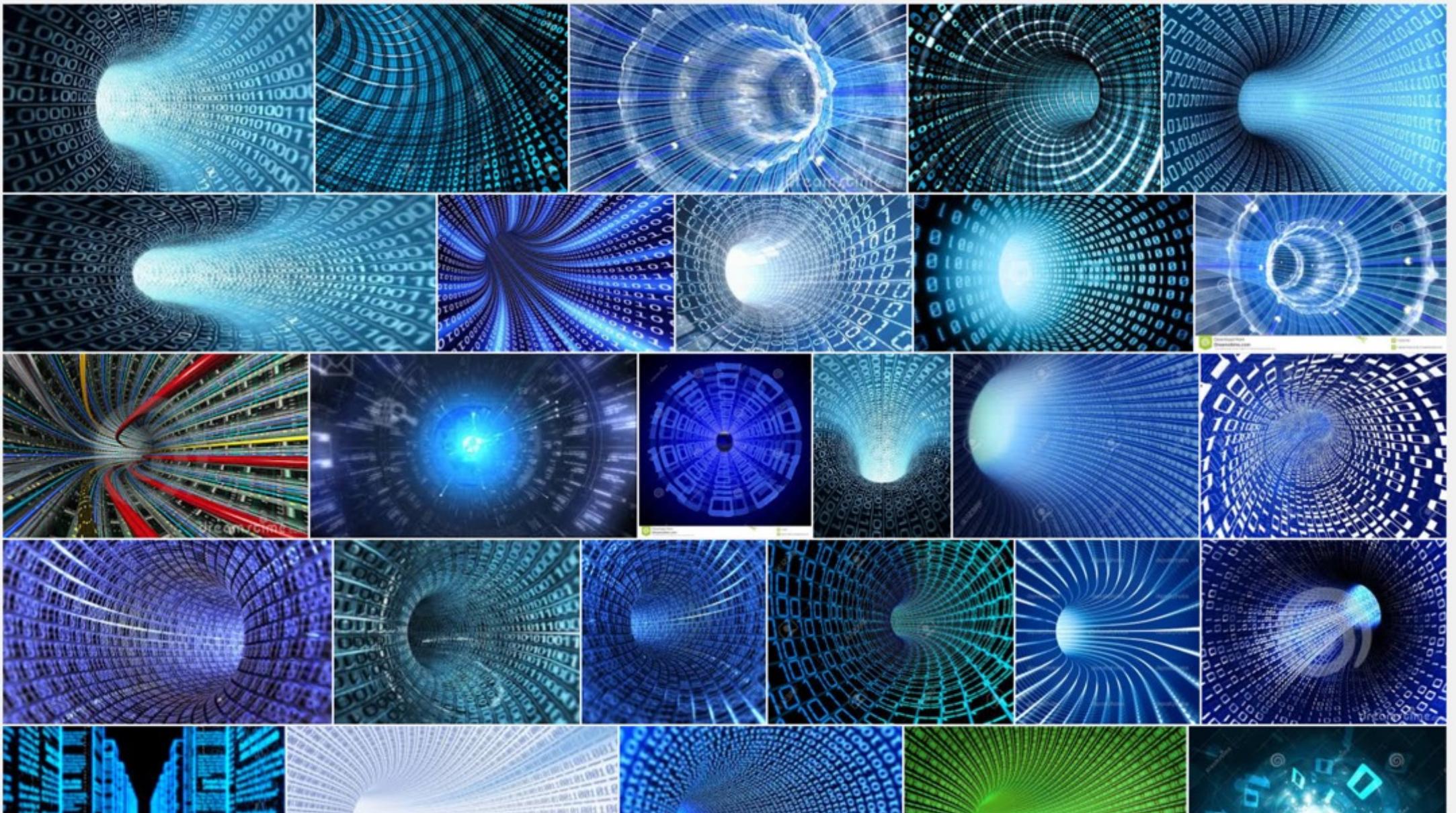
data is the **new oil**

data is the **new currency**

Unknown - [Use precise location](#)

Nothing makes sense
until it does.

data tunnel

[ALL](#) [IMAGES](#) [SHOPPING](#) [NEWS](#) [VIDEOS](#) [MAPS](#) [BOOKS](#) [FLIGHTS](#) [APPS](#) [SEARCH](#)

Big Data:

"More information is created faster than organisations can make sense of it."

Jeff Jonas

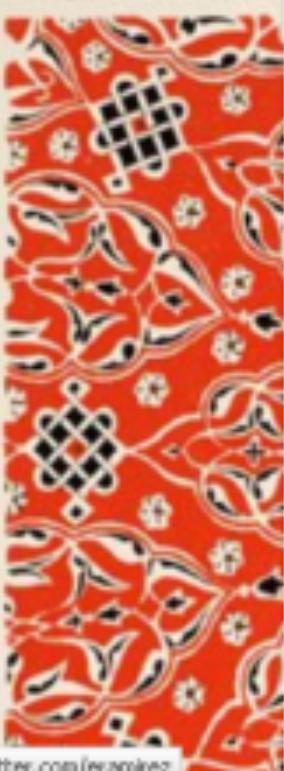
"When the entire industrial process is made explicit, when the metrics count for more than the object they measure

"
..."

Bruce Sterling

TURKISCH

TAFEL XXXVIII



[Ernesto Ramirez](#)

@eramirez

Follow

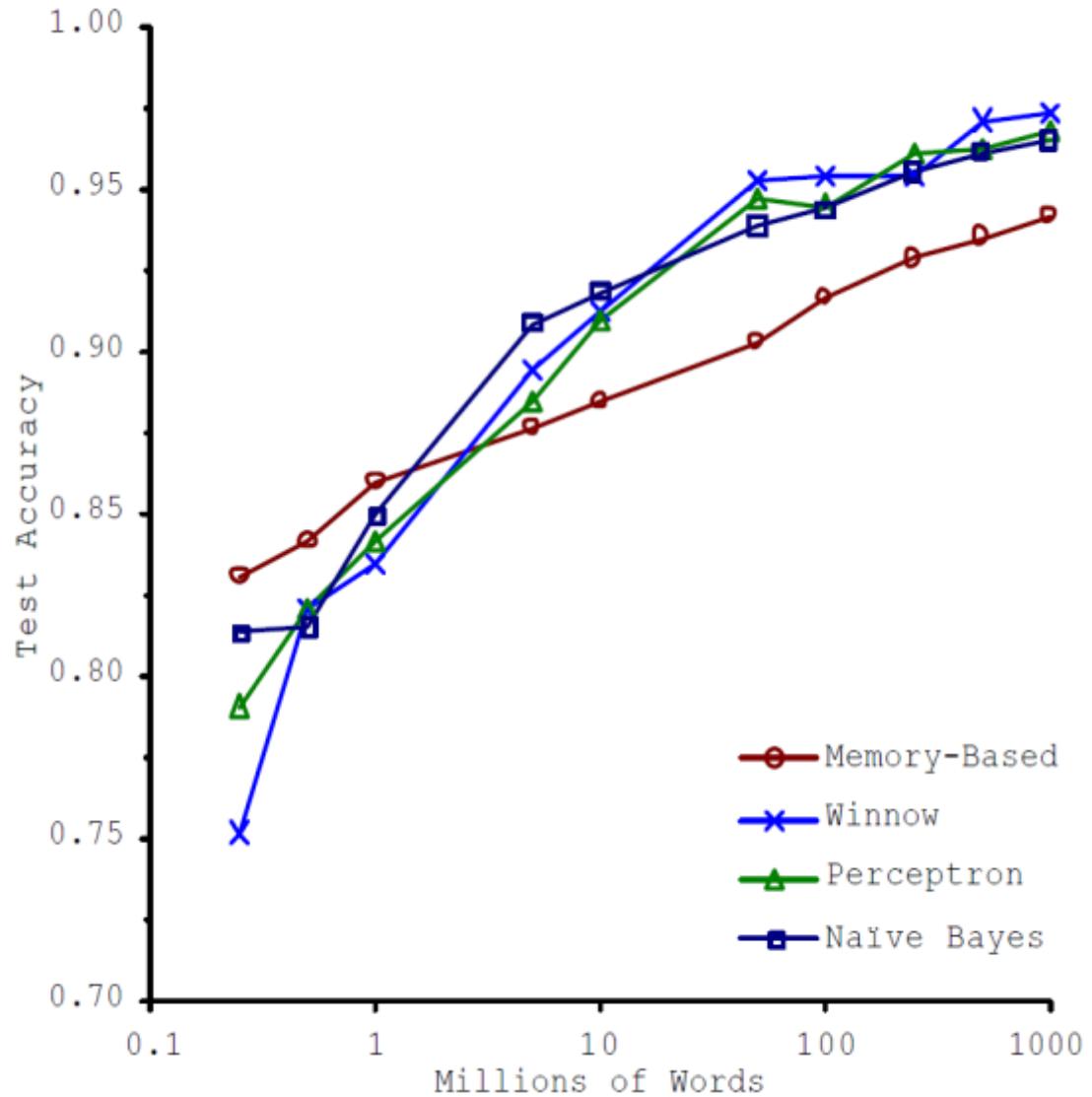
Data is made of people. Read this great post by [@wilbanks](#) on Larry Page and Pooling Medical Data: del-fi.org/post/801749671... #quantifiedself

from Los Angeles, CA

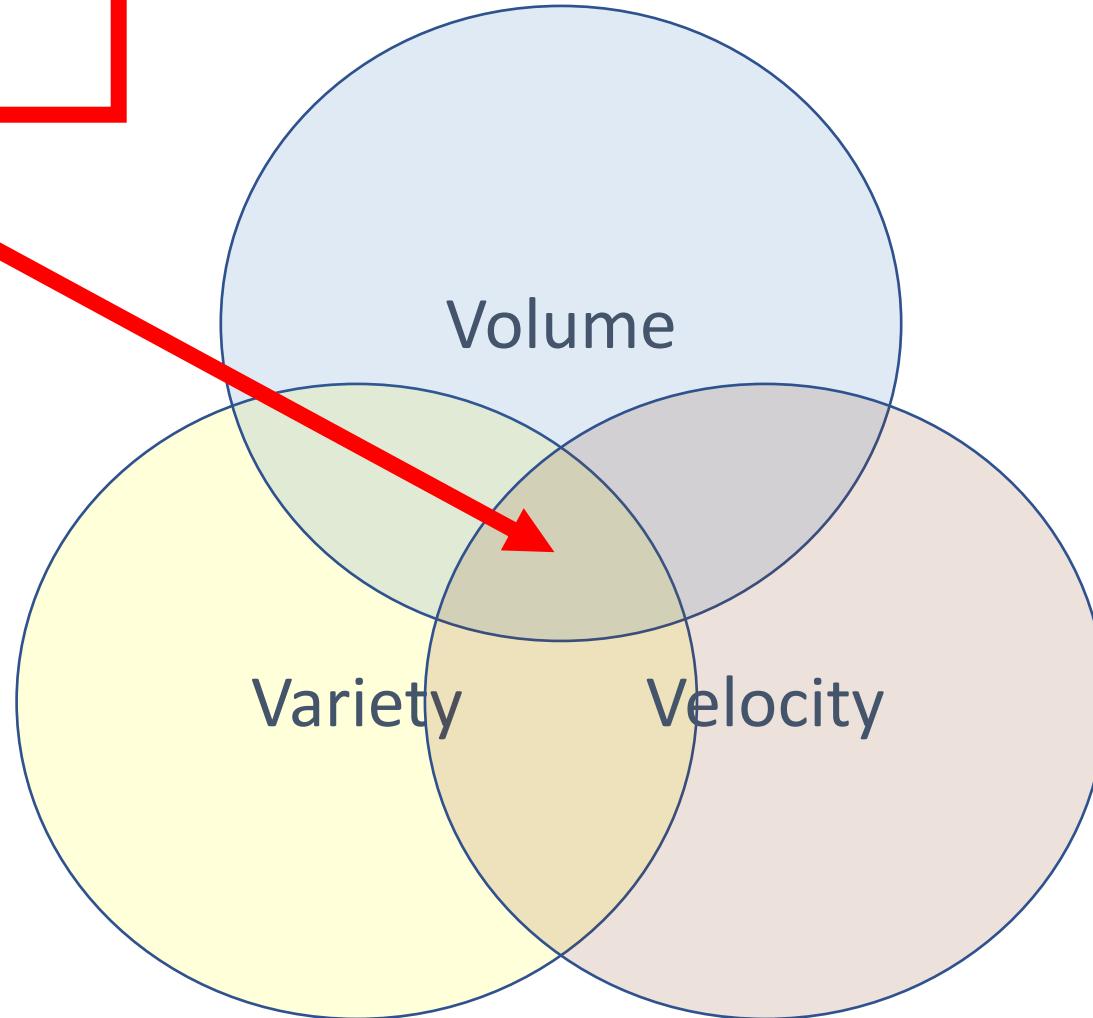
Reply Retweet Favorite More

twitter.com/eramirez

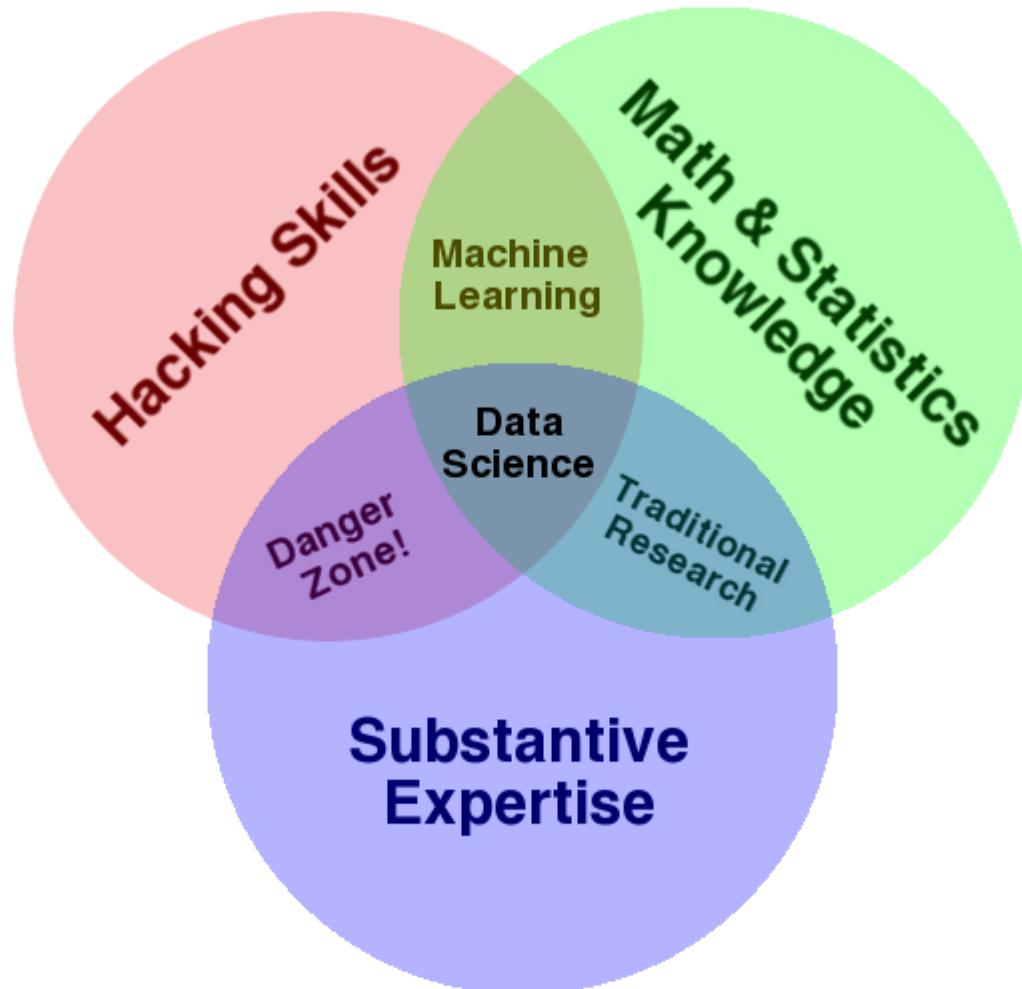
- 1) Data Science
- 2) Predictions
- 3) Content Generation



Big Data

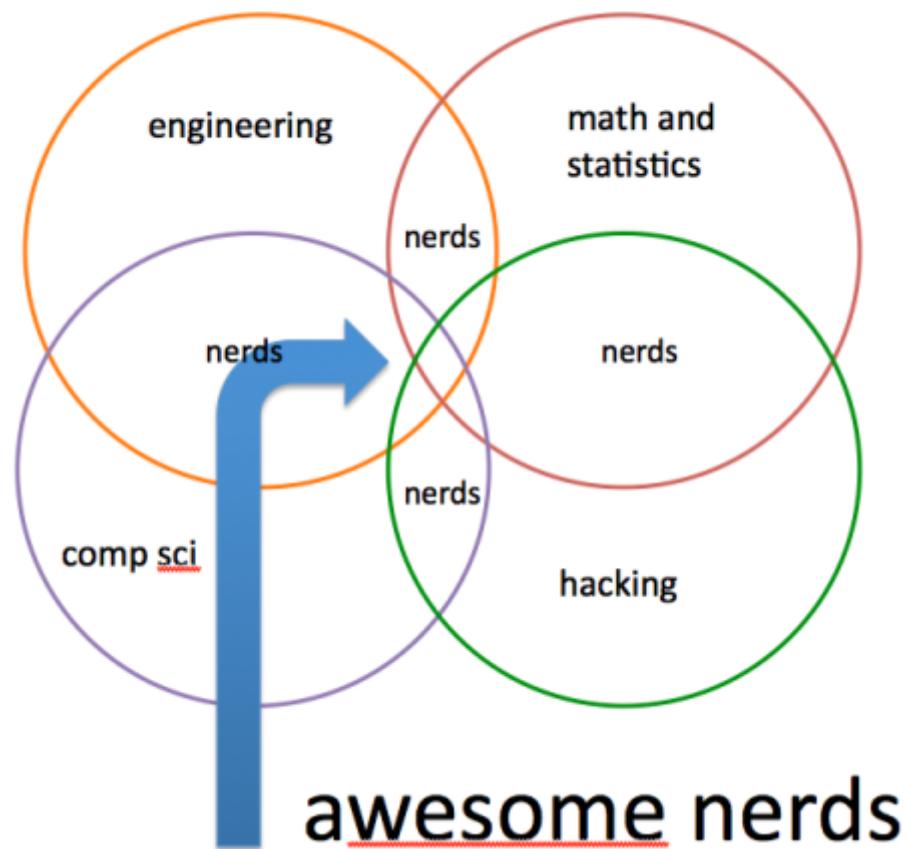


Data Science

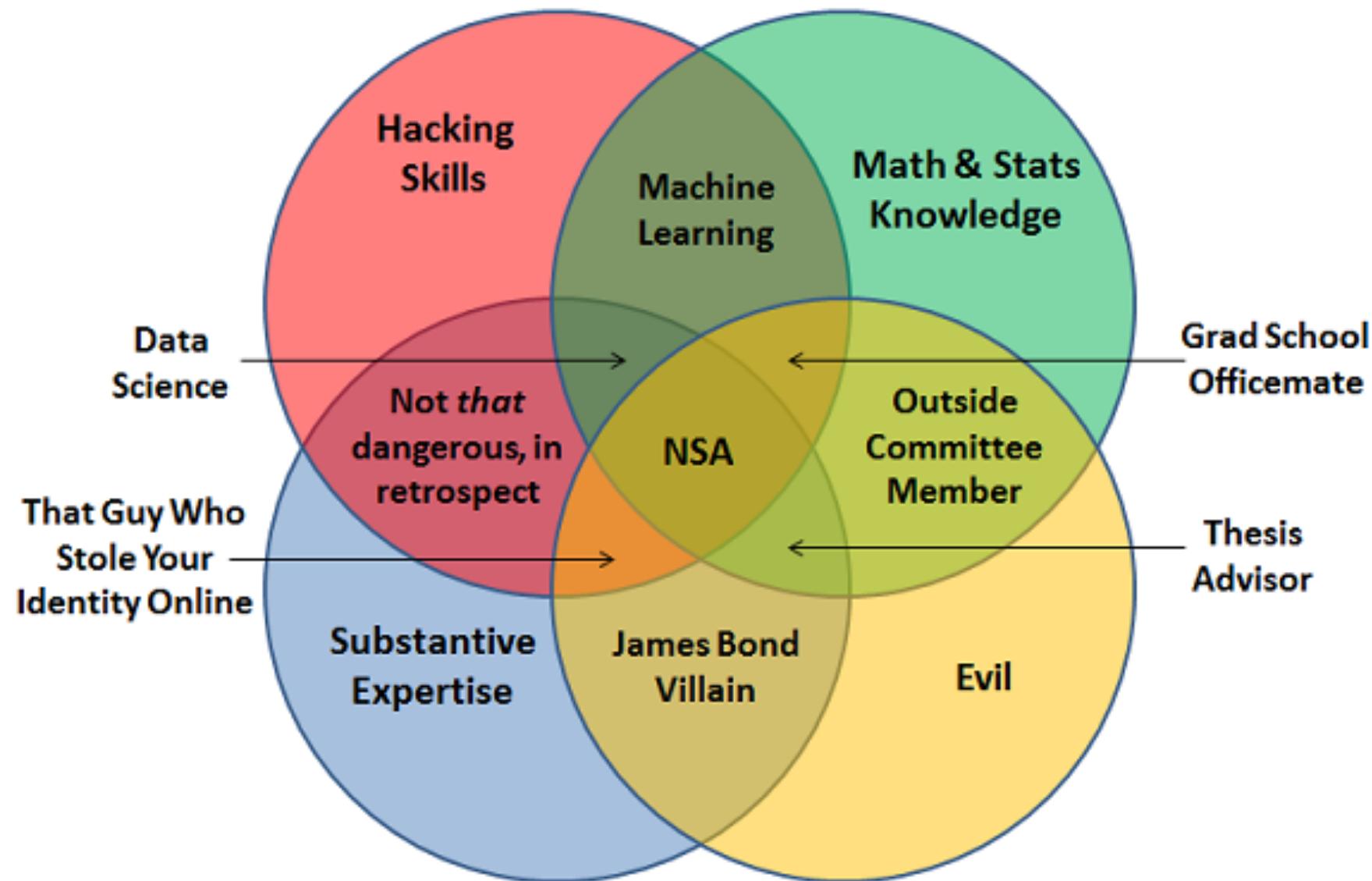


<http://drewconway.com/zia/2013/3/26/the-data-science-venn-diagram>

Data Scientists?

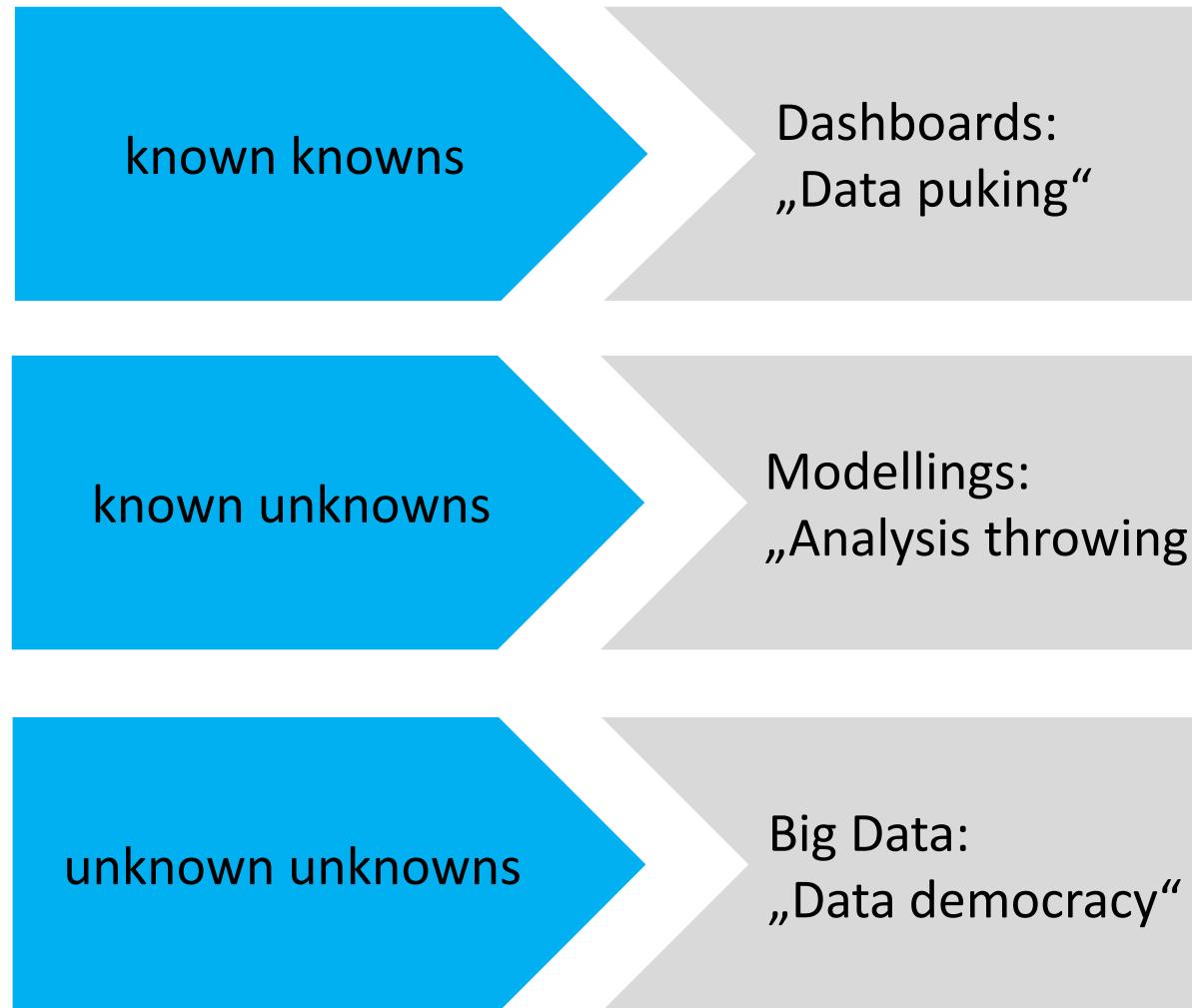


<http://www.forbes.com/sites/danwoods/2012/03/08/hilary-mason-what-is-a-data-scientist/>



As we know,
There are known knowns.
There are things we know we know.
We also know
There are known unknowns.
That is to say We know there are some things
We do not know.
But there are also unknown unknowns,
The ones we don't know
We don't know.

Donald Rumsfeld



Avinash Kaushik

As we know,
There are known knowns.
There are things we know we know.
We also know
There are known unknowns.
That is to say We know there are some things
We do not know.
But there are also unknown unknowns,
The ones we don't know
We don't know.

Donald Rumsfeld

But most dangerous are the things we
don't know that we know,
the unknown knowns.

Slavoj Žižek

Data Science 101

1. Generate / collect data
2. Store / retrieve data
3. Data munging
4. Data exploration
5. Prototype / visualization
6. Productive application

1. Generate / collect data

	Market Research / Statistics	Data Science
Generate and collect data	Field, survey, experiment, method	Parsing, crawling, scraping
Datenquelle	Tables, reports	API, corpus, "heap", "stream"
Datenbeschreibung	Code book	Specs, N.N.
Dienstleister	Field institute	Data Broker

1. Generate / collect data

- Data sources
- Parsing: APIs
 - Authentification (Token)
 - JSON
- Crawling
 - Spiders, crawlers, robots
- Scraping

Data Science 101

1. Generate / collect data
2. Store / retrieve data
3. Data munging
4. Data exploration
5. Prototype / visualization
6. Productive application

2. Store / retrieve data

	Market research / Statistics	Data Science
Format	Tables, RDBMS	Documents, files
Consistency	By definition (formating)	Functional
Structure	Cases, records	Networks, graphs
Data storage	local	cloud





Cloudservices

- Amazon
 - S3, EC2

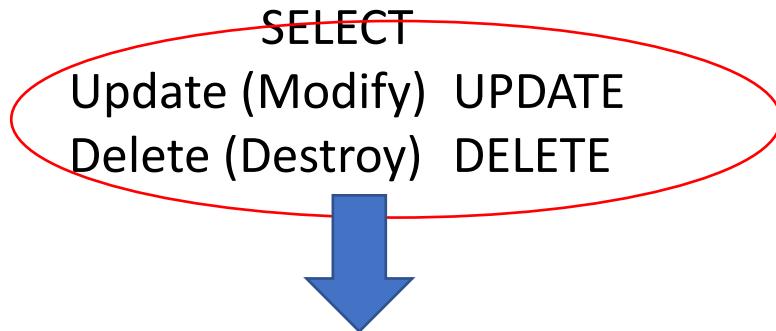


- Google
 - Storage, BigQuery



Immutability

<u>Operation</u>	<u>SQL</u>
Create	
	INSERT
Read (Retrieve)	
	SELECT
Update (Modify)	UPDATE
Delete (Destroy)	DELETE



"mutable"

"Each event happens at a particular time and is always true"

- Just C+R; nothing gets ever "updated"
 - Records are stored as files. Each record is a new file.
- 
- "immutable"

Document/Graph/Key-Value Store

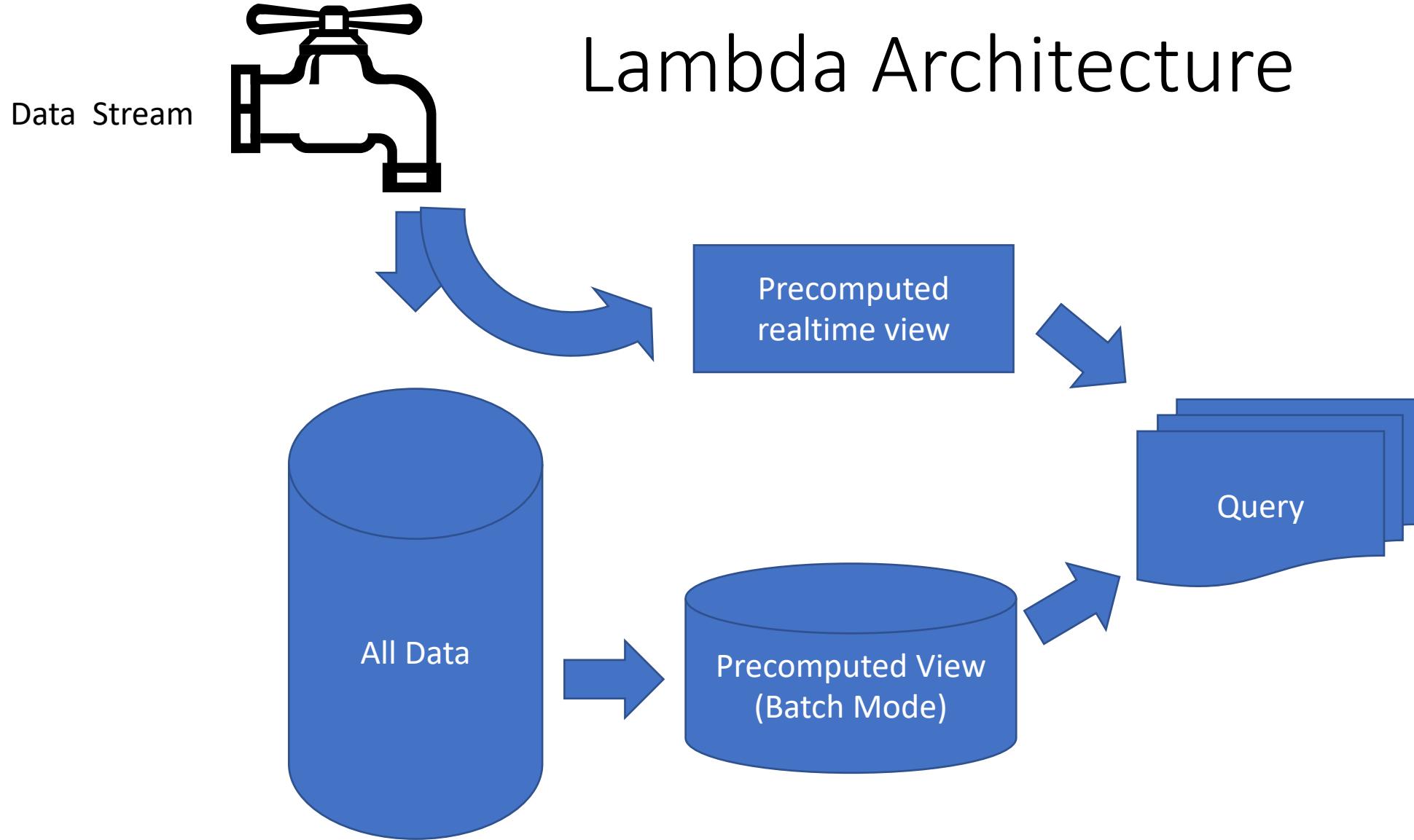
- Documents are indexed in a meaningful way
- Meta data are delivered in special formats (e.g. MPEG7)
- In the case of key-value structure the JSON or XML-files are transformed into a table-like structure



Hash-Table, Bitmap-Index

- Hash-Table: the index points to the data via a hash-function
- Bitmap-Index: all variables are transformed into a 0/1 structure („Dummy variables“)





http://radar.oreilly.com/2015/02/improving-on-the-lambda-architecture-for-streaming-analysis.html?imm_mid=0cc8e2&cmp=em-data-na-na-newsltr_20150211

Data Science 101

1. Generate / collect data
2. Store / retrieve data
3. Data munging
4. Data exploration
5. Prototype / visualization
6. Productive application

3. Data munging

- "Pervasive Data Gremlins":
 - Technical format: Endianness, Unicode-table, time encoding, OS-version, special characters, stopwords, etc.
 - Content: Missing values, time zones, day-light saving time, holidays, labels, addresses, area codes, altitude above sea level, etc.
 - Historic: Change in format (e.g. year 2000), changing granularity, etc.

Data Science 101

1. Generate / collect data
2. Store / retrieve data
3. Data munging
4. Data exploration
5. Prototype / visualization
6. Productive application

4. Data Analytics

	Market research / statistics	Data Science
Tools	SPSS, SAS, Excel	R, Python, D3, other frameworks
Method	Sampling, representativeness	Total universe, non-representative
Mathematik	Stochastics	Algorithms
Results	Finalized, permanent	Iterative

Agile analytics: a/b testing

Twitter's Thin Blue Line Is Already Wavering: A Sign Of More Changes To Conversation Threads To Come?

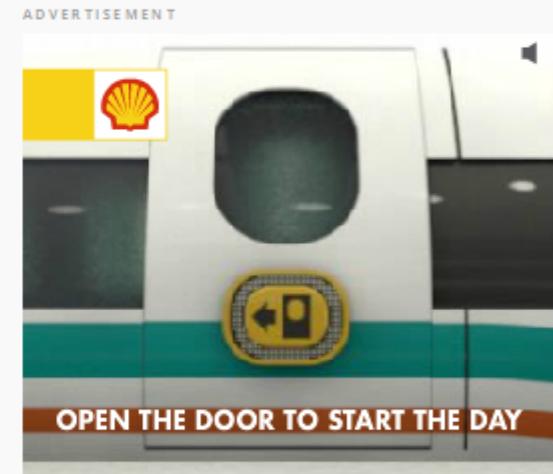
Posted Sep 4, 2013 by Ingrid Lunden (@ingridlunden)

3 Share 71 Share 0 Tweet 766



It was [only a week ago](#) that Twitter updated its web, Android and iOS versions with a new feature — a blue line — that changed how Twitter presented you with conversations around tweets from people you follow. Designed to “[make it easier to discover and follow conversations](#)” on the chatty service, it got a lot of [early negative feedback](#) that hasn’t abated for making the experience cluttered, noisy and confusing.

But now, many users are [reporting](#) that the blue line is no longer appearing, once again replaced by the old “View conversation” link.



CrunchBase

Twitter

Mobile A/B testing at LinkedIn: how members shape our apps

 Akhilesh Gupta January 27, 2014 Share 160  Tweet 2  Like 2  +1 7

At LinkedIn, we strive to build products that our members love. To help make informed product decisions, we do a lot of [A/B testing](#). Although it is a well understood process on the web, it presents a significant challenge on native iOS and Android apps because of their release cycles and singular shipped binary. In this post, we will give an overview of how we are able to modify the client side UI based on well-defined view based JSON responses without client iteration. We will also show how we use this ability along with the LinkedIn Experimentation infrastructure to do highly targeted context-aware A/B testing at scale through our phone server. Finally, we will provide a sneak-peek into how we use all the collected data to automatically generate sophisticated comparison graphs between different buckets to make decisions.

View based JSON

All of our mobile clients ([Android](#), [iPhone](#) and [Touch Web](#)) render content based on structured view based JSON. This means that these clients have a set of predefined views for which data can be populated with JSON associated with that view type. The top slot in the user's stream is called the *hero slot* view. One of the possible JSON responses for the *hero slot* looks like the following:



The vType (view type) tells the client which view to render and all the other fields are directly translatable to the requirements of that view such as the background picture, heading text and detail text. Notice that the fields are based on the visual design of the *hero slot* and you can directly map them to the visual presentation of the *hero slot*. This is different from a resource based JSON in which the entity being sent down from the server is a news resource like the following. In this case, the resource is then translated by the client into the required JSON for the specific view type.

Crowd sourcing



kaggle.com



Machine Learning

- Supervised Learning (Decision making)
 - Regression
 - Diskriminanzanalyse
 - k-Nearest Neighbours
 - Backpropagation (MLP)
- Unsupervised (Dimensionality reduction)
 - Clusteranalysis
 - Principal component analysis (PCA), Multi-dimensional scaling (MDS)
 - Self Organizing Maps



python™



Textanalysis

Word	Limburg	Stockdorf	Taunusstein
Aar	1		2
Aartalbahn			1
Aartalhalle			1
Abbild	1		
Abbildung	2		
aber	7	1	1
Abgaben	1		
abgegrenzten	1		
Abgeordnete	1		
Abgeordneter	1		
abgerissen		3	
Abgerufen	3		
abgeschlossen		1	
abgetrennt	1		
Abitur	2		
Abschluss	1	1	
Abschnitt	2		
Abschnitten	1		

$\text{Cos}(\text{Taunusstein}, \text{Stockdorf})$
0,75

$\text{Cos}(\text{Stockdorf}, \text{Limburg})$ 0,81

$\text{Cos}(\text{Taunusstein}, \text{Limburg})$
0,76

Example: Text mining

The screenshot shows a web browser window for 'Research - Google AI' at ai.google/research/. The page features a navigation bar with links for Google AI, About, Stories, Research, Education, Tools, Responsibilities, and Blog. Below the navigation is a secondary menu with links for Philosophy, Publications, People, Teams & Focus Areas, Join Us, and Outreach. The main content area has a decorative background image of colored dots and lines.

RESEARCH

Advancing the state of the art

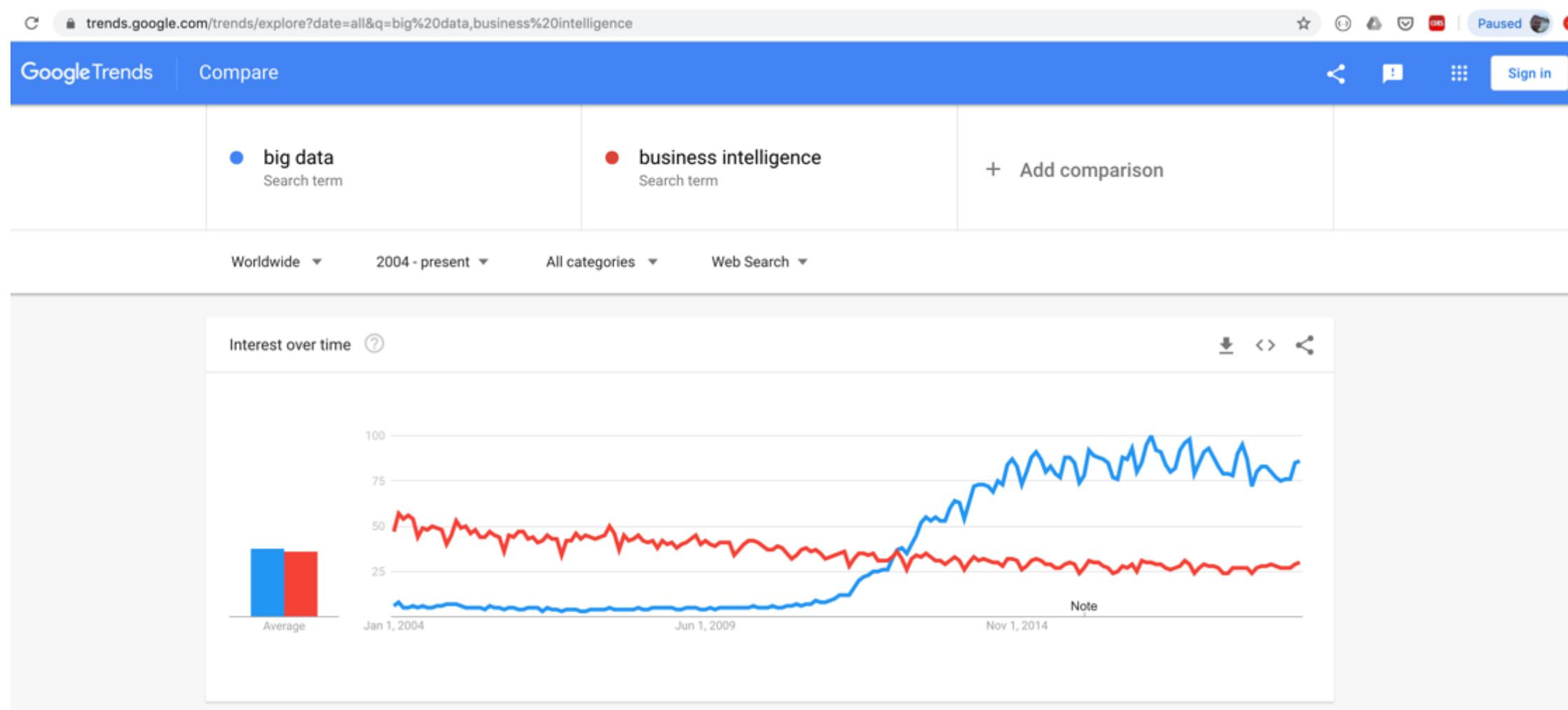
We work on computer science problems that define the technology of today and tomorrow.

Recent publications

[Publication database](#)

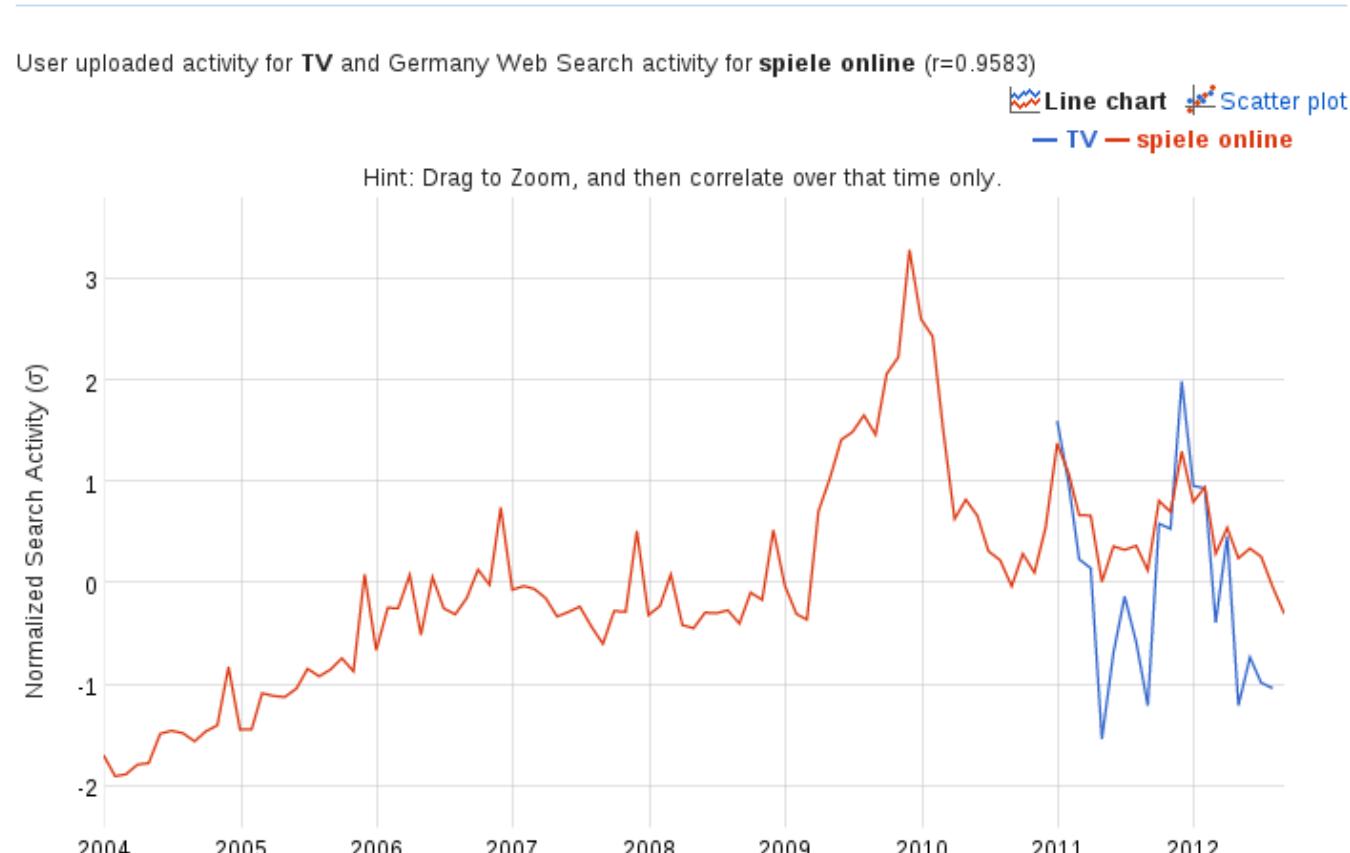
Non-delusional Q-learning and value-iteration Tyler Lu, Dale Schuurmans, Craig Boutilier • <i>Proceedings of the Thirty-second Conference on Neural Information Processing Systems (NeurIPS-18)</i> , Montreal, QC (2018), pp. 9971-9981			
Recurrent World Models Facilitate Policy Evolution David Ha, Jürgen Schmidhuber • <i>Neural Information Processing Systems</i> (2018)			
Grasp2Vec: Learning Object Representations from Self-Supervised Grasping Coline Manon Devin, Eric Jang, Sergey Levine, Vincent Vanhoucke • <i>CoRL</i> (2018)			
Linguistically-Informed Self-Attention for Semantic Role Labeling Emma Strubell, Pat Verga, Daniel Andor, David Weiss, Andrew McCallum • <i>Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing</i> (2018)			
Fluid Annotation: a human-machine collaboration interface for full image annotation Misha Andriukina, Jasper Uijlings, Vittorio Ferrari • <i>ACM Multimedia</i> (2018) (to appear)			

Big Data instead of Business Intelligence

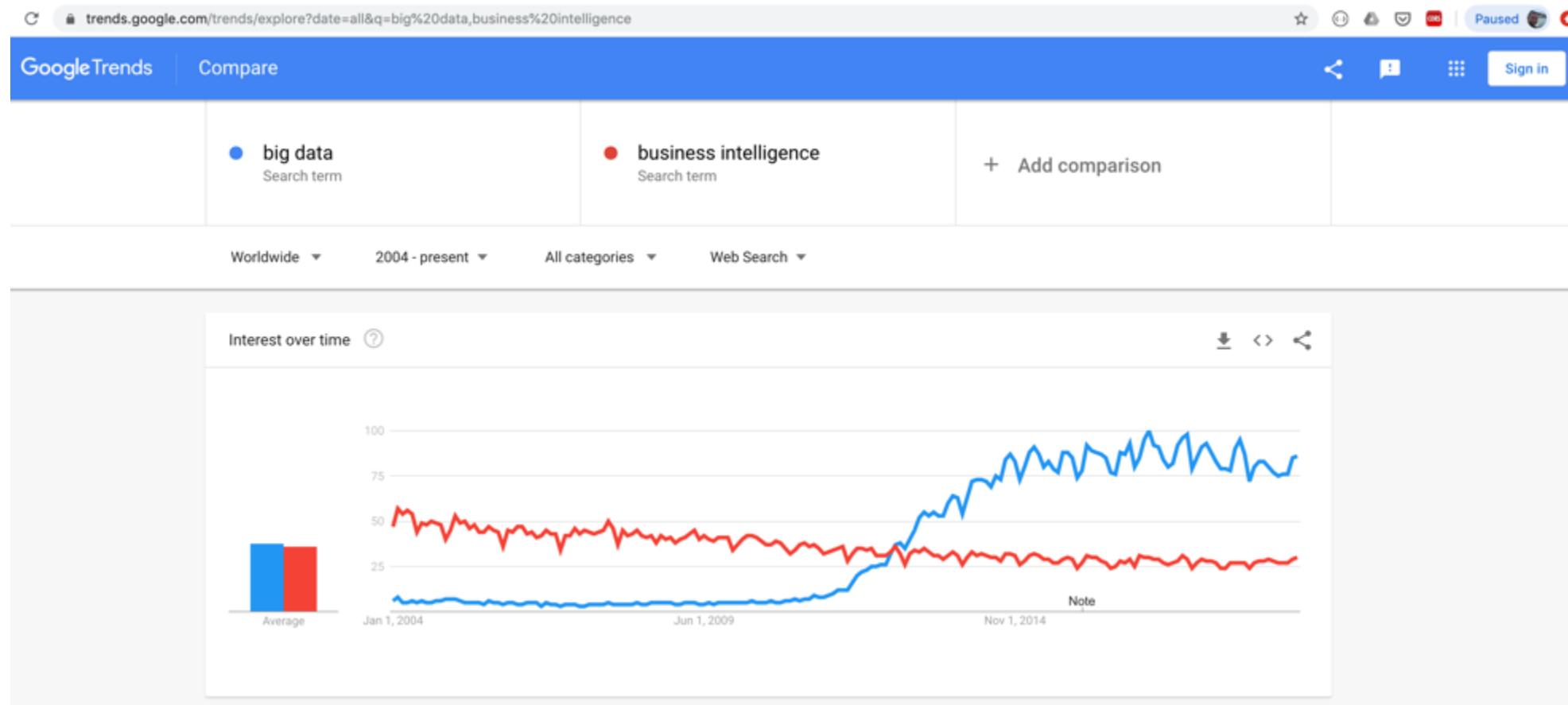


Google Correlate

0.9583 [spiele online](#)
0.9470 [garten eden dietzenbach](#)
0.9389 [sauna dülmen](#)
0.9358 [sauna kitzingen](#)
0.9316 [thermalbad waldbronn](#)
0.9312 [magnat monitor supreme](#)
0.9298 [galli](#)
0.9288 [7 64bit](#)
0.9285 [netzwerkschlüssel vergessen](#)
0.9257 [wf trac](#)
0.9254 [magnat monitor](#)
0.9248 [bowling reutlingen](#)
0.9231 [farbwürfel](#)
0.9223 [thermalbad bad bergzabern](#)
0.9222 [formatieren ntfs](#)
0.9221 [fagor schnellkochtopf](#)
0.9212 [thermalbad bad wildbad](#)
0.9211 [streckenplaner carrera](#)
0.9204 [saunagarten](#)
0.9197 [blocksberg](#)



N-Gram



Google Ngram Viewer <http://books.google.com/ngrams> + DB <http://books.google.com/ngrams/datasets>

WordNet

- WordNet:
semantische
und lexikalische
Bedeutung von
Wörtern

- Daraus z.B.
Wörter mit
Stimmungen
identifizierbar
(WN Affect)

WordNet Search - 3.1
- [WordNet home page](#) - [Glossary](#) - [Help](#)

Word to search for:

Display Options:

Key: "S:" = Show Synset (semantic) relations, "W:" = Show Word (lexical) relations
Display options for sense: (frequency) {offset} <lexical filename> [lexical file number] (gloss)
"an example sentence"
Display options for word: word#sense number (sense key)

Noun

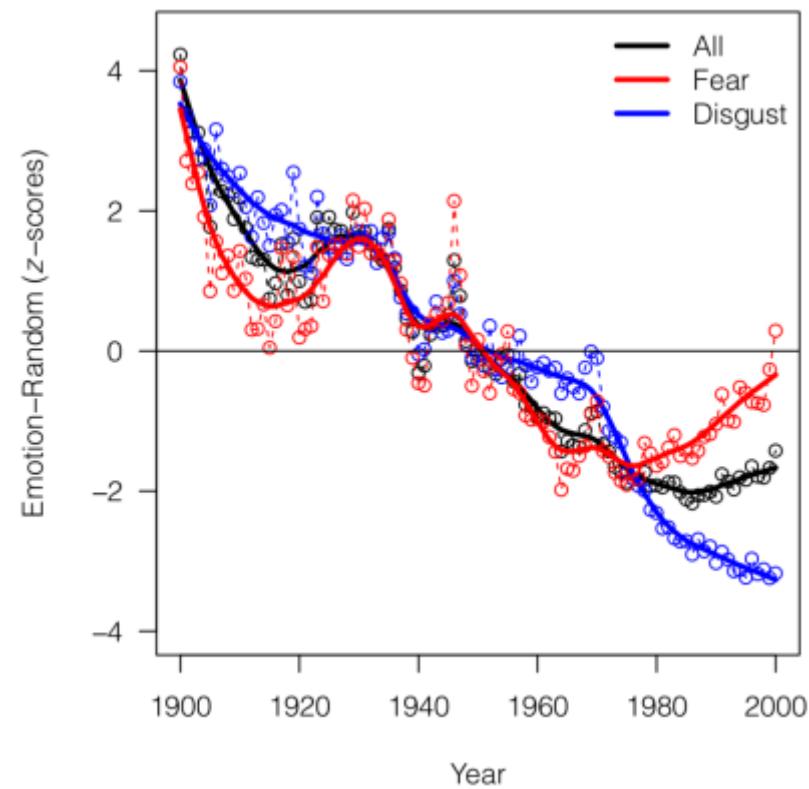
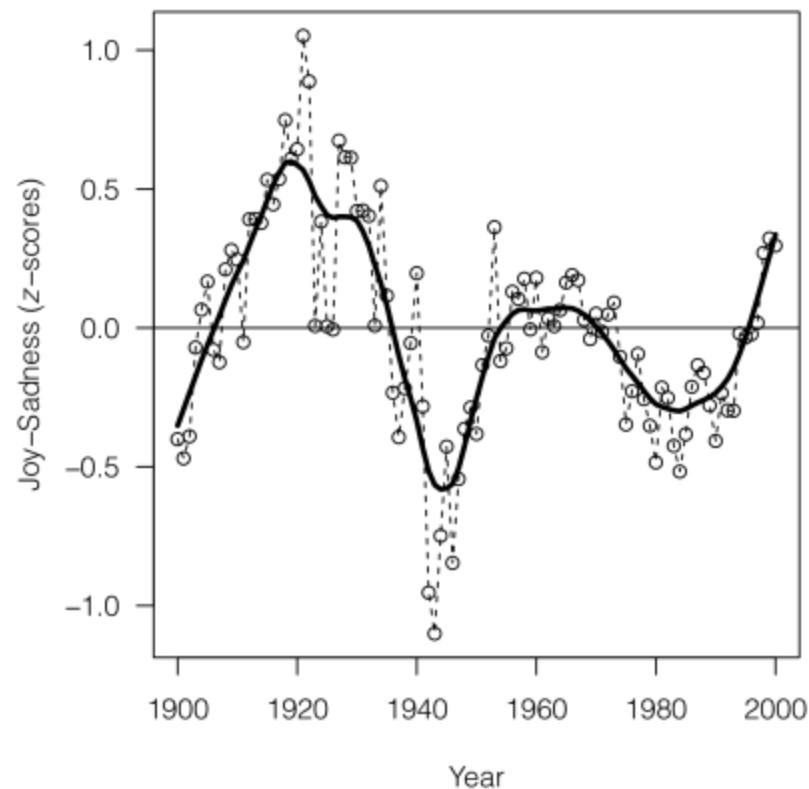
- (73){07534492} <noun.feeling>[12] S: (n) **fear#1 (fear%1:12:00::), fearfulness#1 (fearfulness%1:12:00::), fright#1 (fright%1:12:00::)** (an emotion experienced in anticipation of some specific pain or danger (usually accompanied by a desire to flee or fight))
- (7){07539768} <noun.feeling>[12] S: (n) **concern#2 (concern%1:12:01::), care#3 (care%1:12:01::), fear#2 (fear%1:12:01::)** (an anxious feeling) "care had aged him"; "they hushed it up out of fear of public reaction"
- {07536278} <noun.feeling>[12] S: (n) **fear#3 (fear%1:12:02::), reverence#1 (reverence%1:12:00::), awe#2 (awe%1:12:00::), veneration#1 (veneration%1:12:00::)** (a feeling of profound respect for someone or something) "the fear of God"; "the Chinese reverence for the dead"; "the French treat food with gentle reverence"; "his respect for the law bordered on veneration"

Verb

- (23){01784548} <verb.emotion>[37] S: (v) **fear#1 (fear%2:37:03::)** (be afraid or feel

N-Gramme + WordNet

- Emotionen im Zeitverlauf

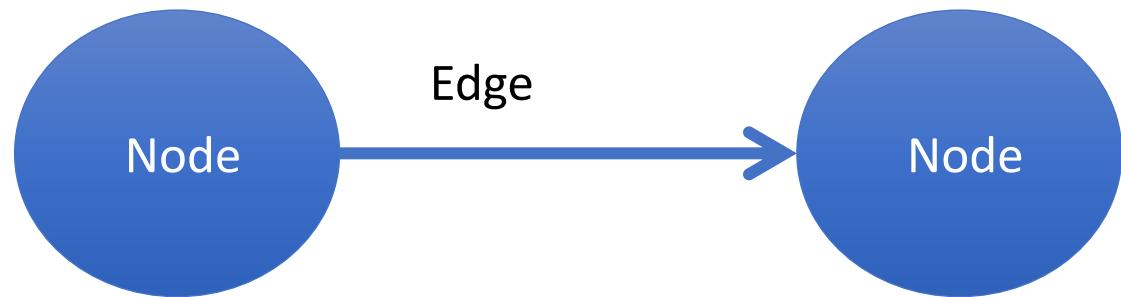


Machine Translation; Large Scale;
Neural Networks; Speech
Recognition; Language Modeling;
Deep Neural; Dependency Parsing;
Finite-State Transducers; Grammar
Induction; Latent Variable; Statistical
Machine; Information Retrieval;
Public Opinion; Automatic Speech;
search engine; Domain Adaptation;
Image Annotation; Revenue
Maximization; Voice Search; Social
Networks

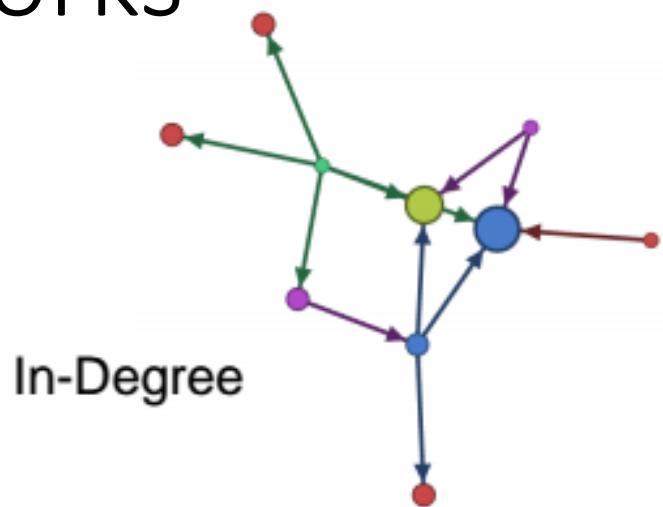
```
text.generate(100)
```

Abstract : We present a new multi-stroke gesture recognition algorithms and show how collected service usage (service access) information, as well as the 2-expansion lemma, menger 's theorem and gallai 's theorem. These two parties, it may be mediated by position-specific sequence features. We present an in-home observation and in-context research study for breast cancer prognosis.

Networks



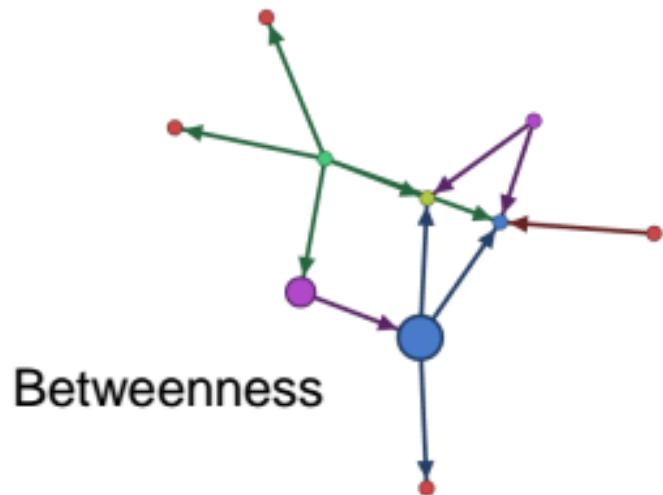
Networks



In-Degree



Out-Degree

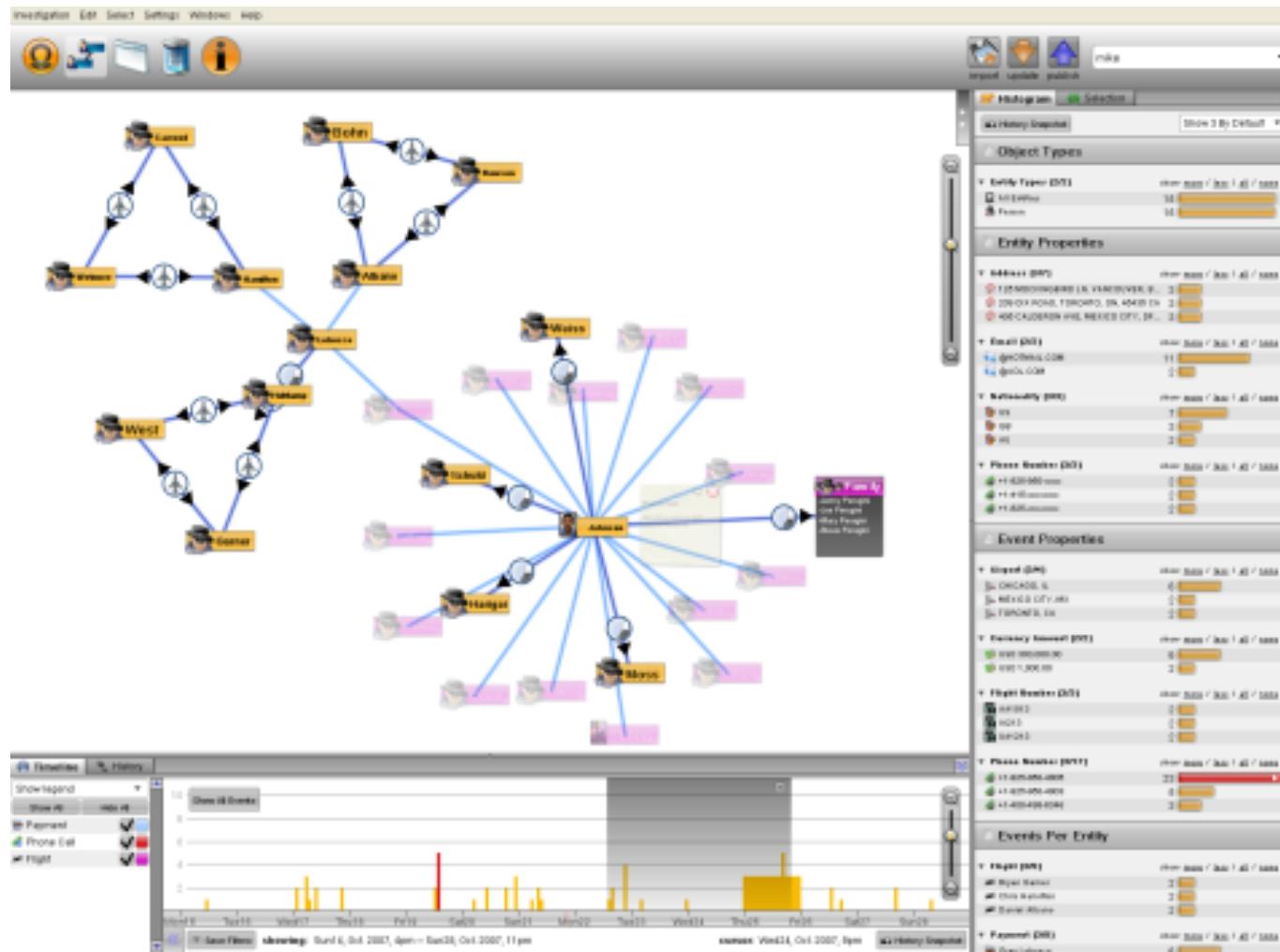


Betweenness



Page Rank

Example: Palantir

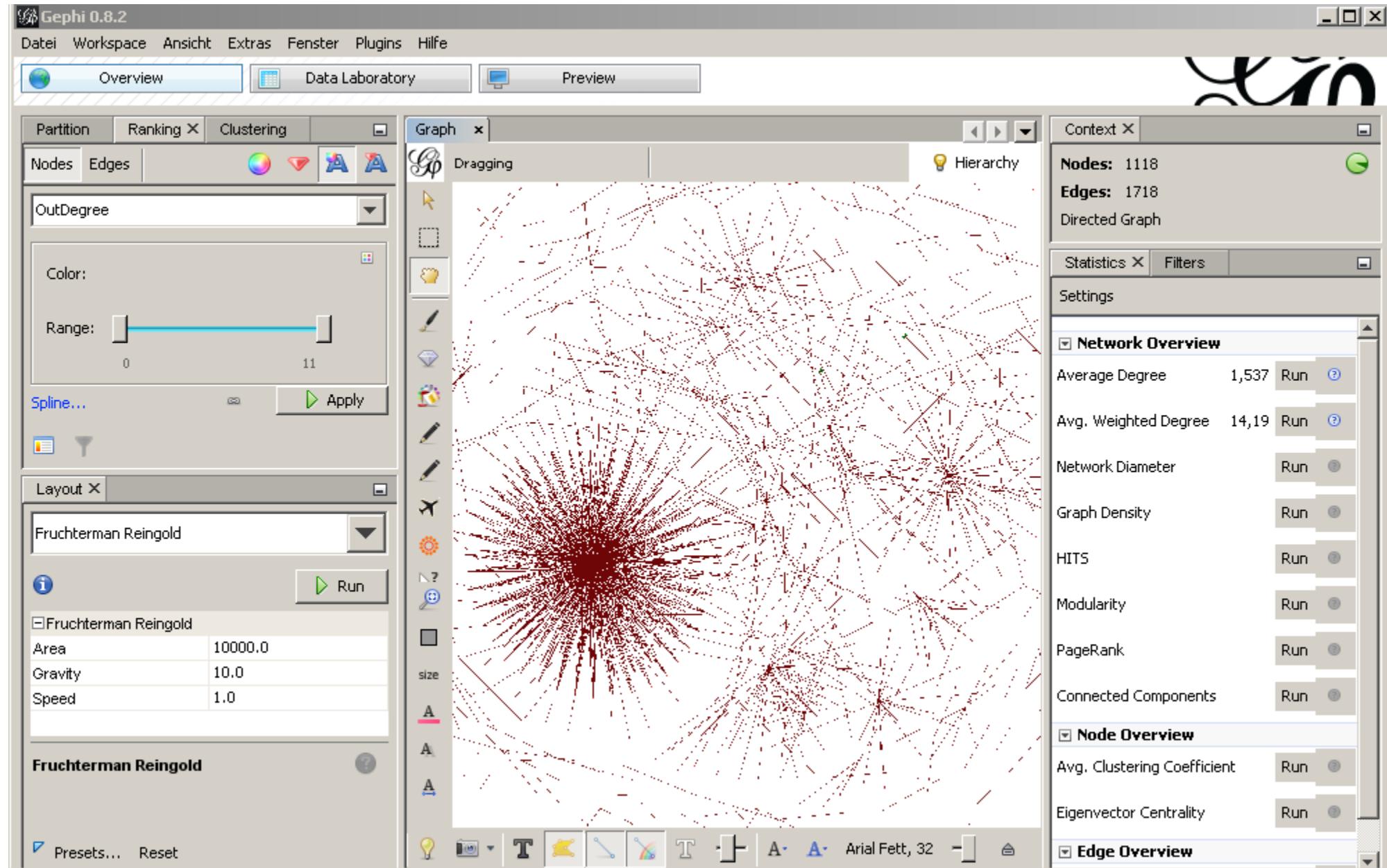


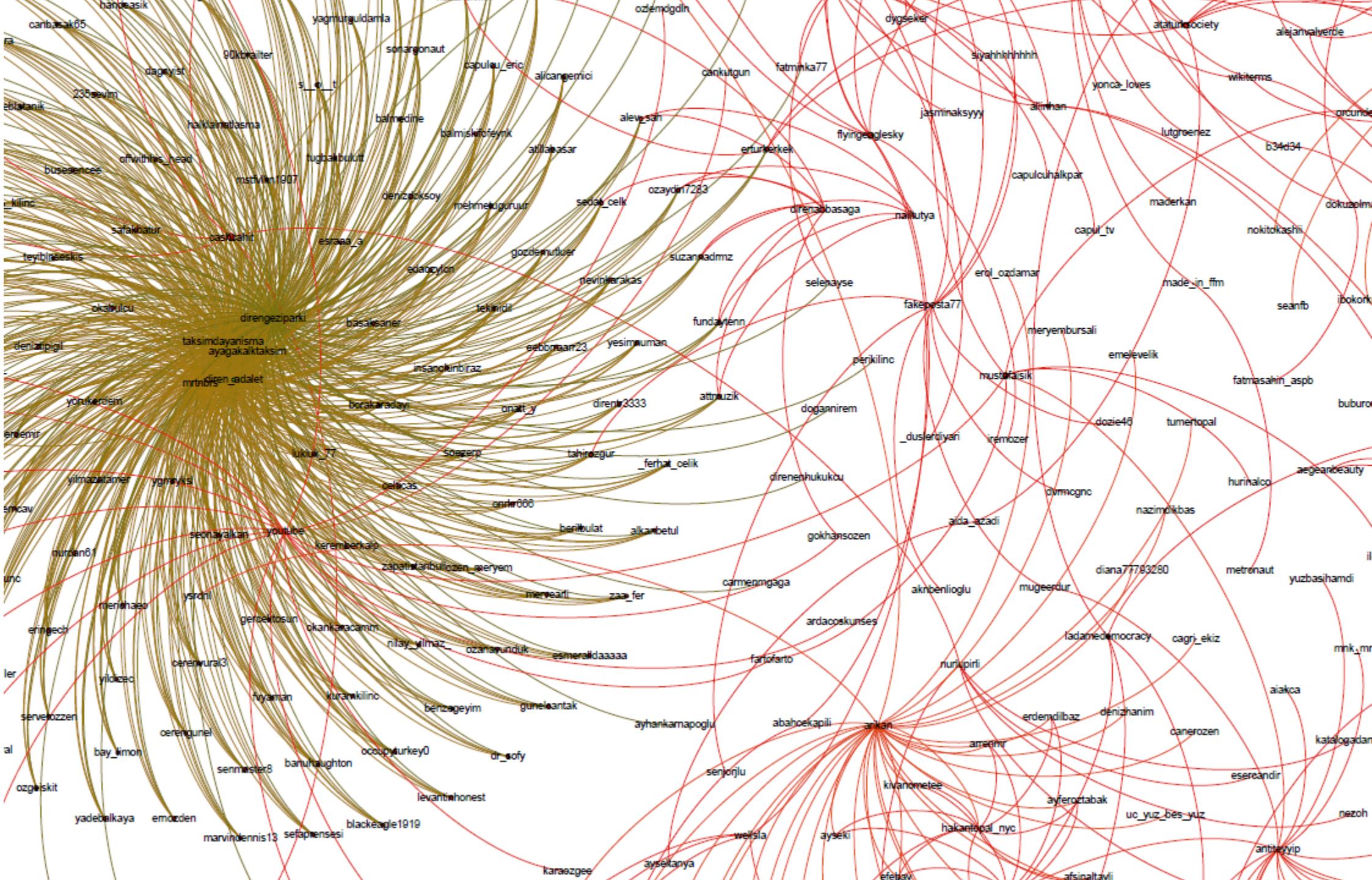
Data Table x Graph x

Nodes Edges Configuration Add node Add edge Search/Replace Import Spreadsheet Export table More actions

Source	Target	Type	Id
atmuzik	direnabbasaga	Directed	255
atmuzik	diren_adalet	Directed	256
atmuzik	ayagakalktaksim	Directed	257
atmuzik	taksimdayanisma	Directed	258
atmuzik	direngeziparki	Directed	259
atmuzik	nailtutya	Directed	260
diren_adalet	ayagakalktaksim	Directed	2057
diren_adalet	taksimdayanisma	Directed	2058
diren_adalet	direngeziparki	Directed	2059
diren_adalet	hakansivlim	Directed	1715
diren_adalet	lasciateimpensr	Directed	821
direngeziparki	parklarbizim	Directed	2081
ernozden	diren_adalet	Directed	261
ernozden	ayagakalktaksim	Directed	262
ernozden	taksimdayanisma	Directed	263
ernozden	direngeziparki	Directed	264
ernozden	mrtnbrs	Directed	265
mrtnbrs	direngeziparki	Directed	2065
mrtnbrs	mrtnbrs	Directed	2066
insanolunbiraz	diren_adalet	Directed	266
insanolunbiraz	ayaqakalktaksim	Directed	267

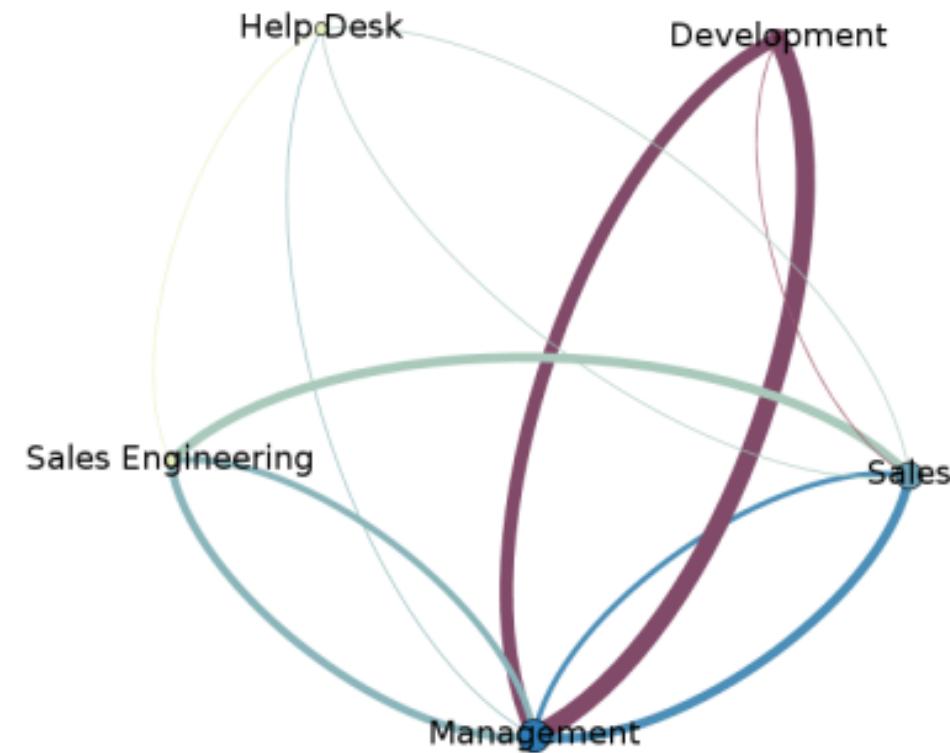
Add column Merge columns Delete column Clear column Copy data to other column Fill column with a value Duplicate column Create a boolean column from regex match Create column with list of regex matching groups



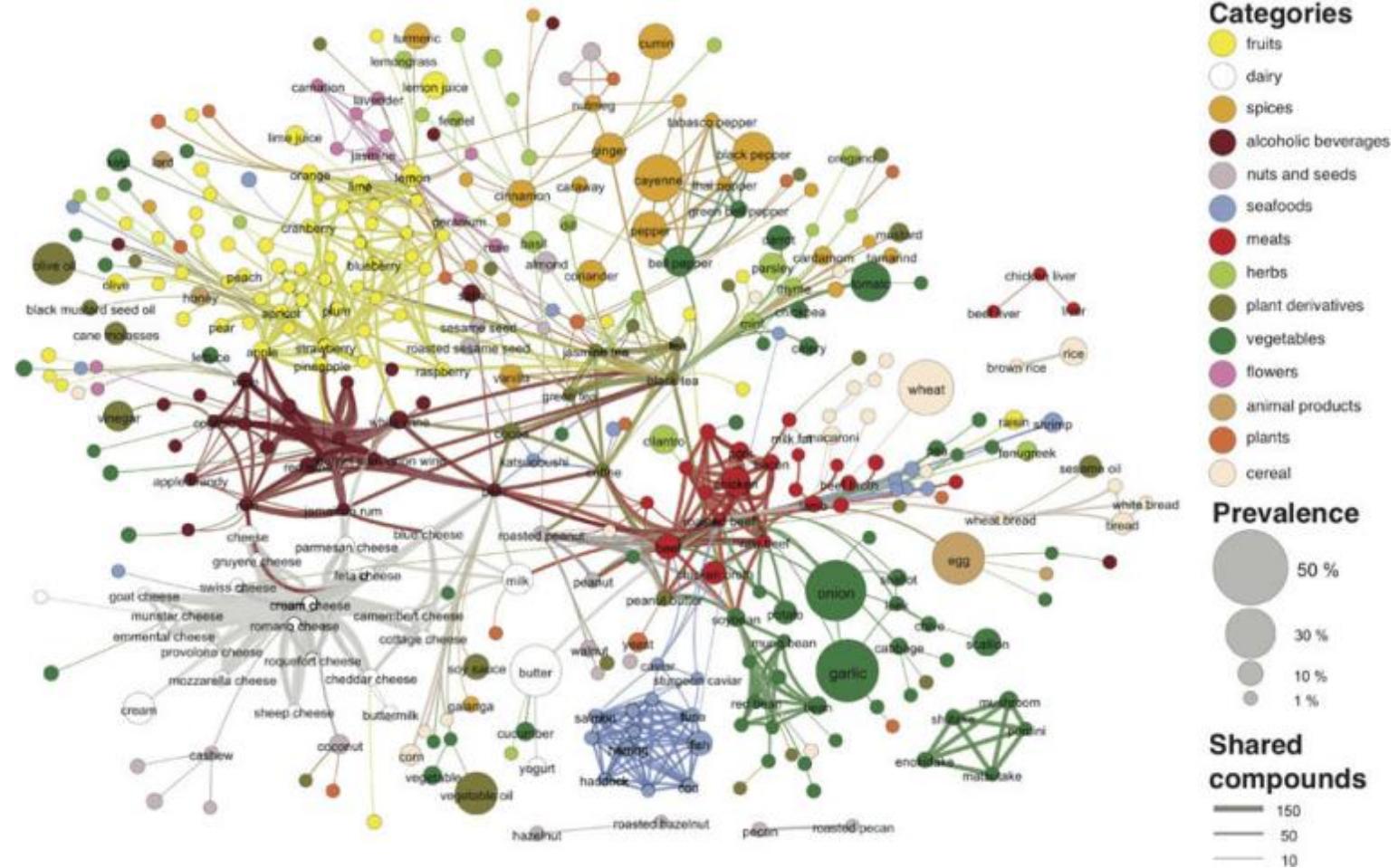


Example: Communication in a company

- Who sends mails to whom?



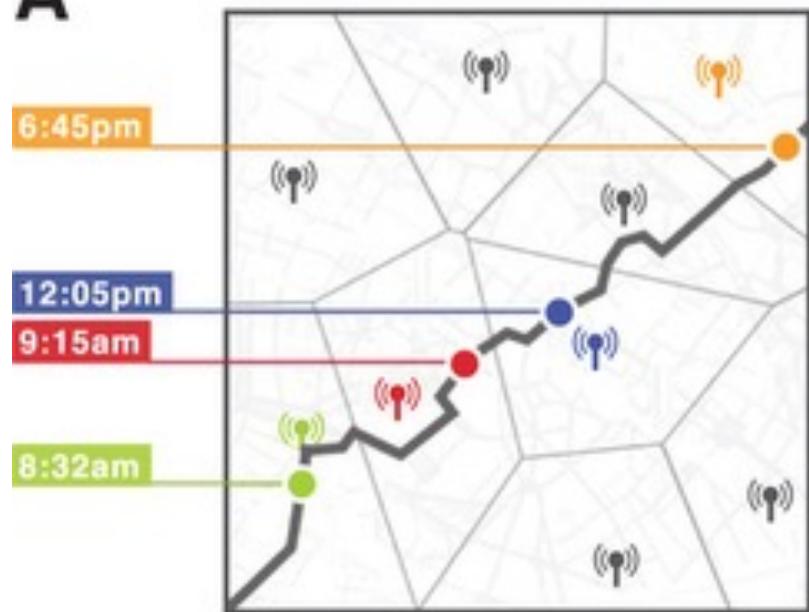
Food Pairings



Ahn et al 2011 <http://www.nature.com/srep/2011/111215/srep00196/full/srep00196.html>

Data Fingerprinting

A



B



The full process

1. Generate / collect data

- Crawler, APIs, etc.

2. Save / retrieve data

- Hadoop, NoSQL, ...

3. Data munging

- Python

4. Data exploration

- R, Python, Gephi, ...

5. Prototype / visualization

- Gephi, R, Python, ...

6. Productive application

- C#, Java, HTML5, ...



**YOU
NO LONGER
NEED
YOUR
PRE-INTERNET
BRAIN**

London Evening Standard

[HOME](#) | [NEWS](#) | [SPORT](#) | [FOOTBALL](#) | [SHOWBIZ](#) | [GOING OUT](#) | [LIFESTYLE](#) | [FASHION & BEAUTY](#) | [VIDEOS](#) | [BUSINESS](#) | [COMME](#)[London](#) | [Crime](#) | [Politics](#) | [Mayor](#) | [Transport](#) | [Education](#) | [Health](#) | [TECHNOLOGY](#) | [UK](#) | [World](#) | [Campaigns](#) | [Diary](#)**BOMB DISPOSAL OFFICERS BLOW UP BAG OUTSIDE IMPERIAL WAR MUSEUM AMID REMEMBRANCE DAY SECURITY**

Average smartphone user checks device 221 times a day, according to research



MOST READ



system



afterparty





Twitter account settings page for @jbenno.

Personalization and data

Control how Twitter personalizes content and collects and shares certain data.

Personalization and data (Toggle switch)

This will enable or disable all of the settings on this page.

Personalization

Personalized ads (checkbox)

You will always see ads on Twitter based on your Twitter activity. When this setting is enabled, Twitter may further personalize ads from Twitter advertisers, on and off Twitter, by combining your Twitter activity with other online activity and information from our partners. [Learn more](#)

Personalize based on your inferred identity (checkbox)

Twitter will always personalize your experience based on information you've provided, as well as the devices you've used to log in. When this setting is enabled, Twitter may also personalize based on other inferences about your identity, like devices and browsers you haven't used to log in to Twitter or email addresses and phone numbers similar to those linked to your Twitter account. [Learn more](#)

Personalize based on the places you've been (checkbox)

Twitter always uses some information, like where you signed up and your current location, to help show you more relevant content. When this setting is enabled, Twitter may also personalize your experience based on other places you've been.

Settings

- @jbenno
- Account
- Privacy and safety
- Notifications
- Content preferences

General

- Display
- Data usage
- Accessibility
- About Twitter

17:59
techblog.netflix.com

the way we score items, how we determine what items to include in them, and in what order to place those items.

Take as a first example the Top 10 row: this is our best guess at the ten titles you are most likely to enjoy. Of course, when we say "you", we really mean everyone in your **household**. It is important to keep in mind that Netflix' personalization is intended to handle a household that is likely to have different people with different tastes. That is why when you see your Top10, you are likely to discover items for dad, mom, the kids, or the whole family. Even for a single person household we want to appeal to your range of interests and moods. To achieve this, in many parts of our system we are not only optimizing for accuracy, but also for **diversity**.

Personalization awareness

Diversity

Another important element in Netflix' personalization is **awareness**. We want members to be aware of how we are adapting to their tastes. This not only promotes trust in the system, but encourages members to give feedback that will result in better recommendations. A different way of promoting trust with the personalization component is to provide **explanations** as to why we decide to recommend a given movie or show. We are not recommending it because it suits our business needs, but because it matches the information we have from you: your explicit taste preferences and ratings, your viewing history, or even your friends' recommendations.



The New York Times

Tuesday, September 24, 2013 Last Update: 9:34 AM ET

[SUBSCRIBE TO THE NEW YORK TIMES.](#)[GET UNLIMITED ACCESS ▾](#)UNDER €1 FOR YOUR
FIRST 4 WEEKS.GLOBAL EDITION
The New York Times SearchFollow Us [f](#) [t](#) | Personalize Your WeatherINVESTMENT BANKING | JULY 30, 2013, 9:17 PM | [668 Comments](#)

Over a Million Are Denied Bank Accounts for Past Errors

BY JESSICA SILVER-GREENBERG

Mistakes like a bounced check or a small overdraft have effectively blacklisted more than a million low-income Americans from the mainstream financial system for as long as seven years as a result of little-known private databases that are used by the nation's major banks.

"All models are wrong
but some are useful."

AMAZON ECHO IST NUR PER EINLADUNG VERFÜGBAR Mehr dazu ▾



Amazon Echo, Schwarz

von Amazon

79 beantwortete Fragen

Teilen

Einladung anfordern

Preis: **EUR 179,99** GRATIS Lieferung für Prime-Mitglieder sobald verfügbar
Alle Preisangaben inkl. MwSt.

Dieser Artikel erscheint am 26. Oktober 2016.

Bestellen Sie jetzt.

Verkauf und Versand durch Amazon EU Sarl. Geschenkverpackung verfügbar.

Farbe: **Schwarz**



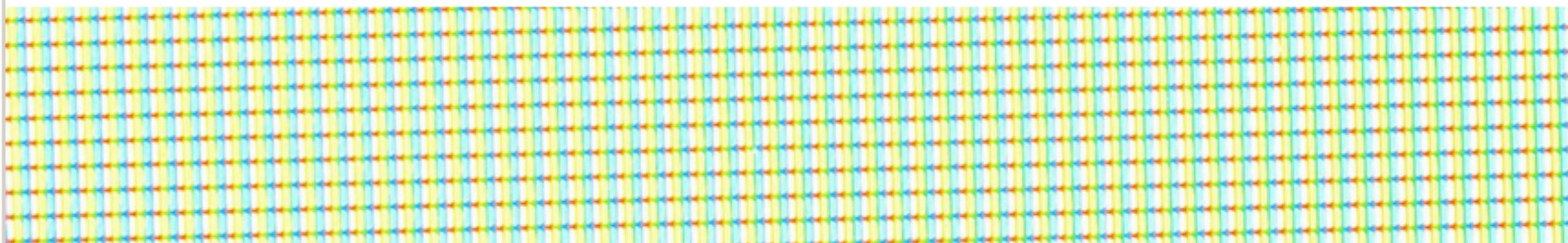
- Spielt all Ihre Musik von Prime Music, Spotify, Tuneln und vielen weiteren ab - allein mithilfe Ihrer Stimme
- Erfüllt den Raum mit einem 360° omnidirektionalen Klang
- Kann allein mit Ihrer Stimme gesteuert werden
- Hört Sie dank der Fernfeld-Spracherkennung auch von der anderen Seite des Raumes, sogar in lauter Umgebung oder wenn Musik läuft
- Beantwortet Fragen, liest Hörbücher, liefert Nachrichten, Verkehrs- und Wetterinformationen, informiert Sie über Sportergebnisse, Spielpläne und mehr mithilfe des Alexa Voice Service
- Steuert Lampen, Lichtschalter, Thermostate und mehr mit kompatiblen verbundenen Geräten von WeMo, Philips Hue, Netatmo, Innogy, tado° und weiteren
- Lernt immer mehr dazu - Alexa wird über die Cloud automatisch aktualisiert, lernt ständig dazu und erhält neue Funktionen und Skills

Gehe zu: [Highlights](#) | [Technische Details](#)

Möchten Sie Ihr Elektro- und Elektronik-Gerät kostenlos recyceln? ([Erfahren Sie mehr.](#))

Botnet defrauds advertisers of £4m per month, may be run by publishers

TECHNOLOGY / 19 MARCH 13 / by OLIVIA SOLON 





Lokales



Tegernsee



Tegernsee

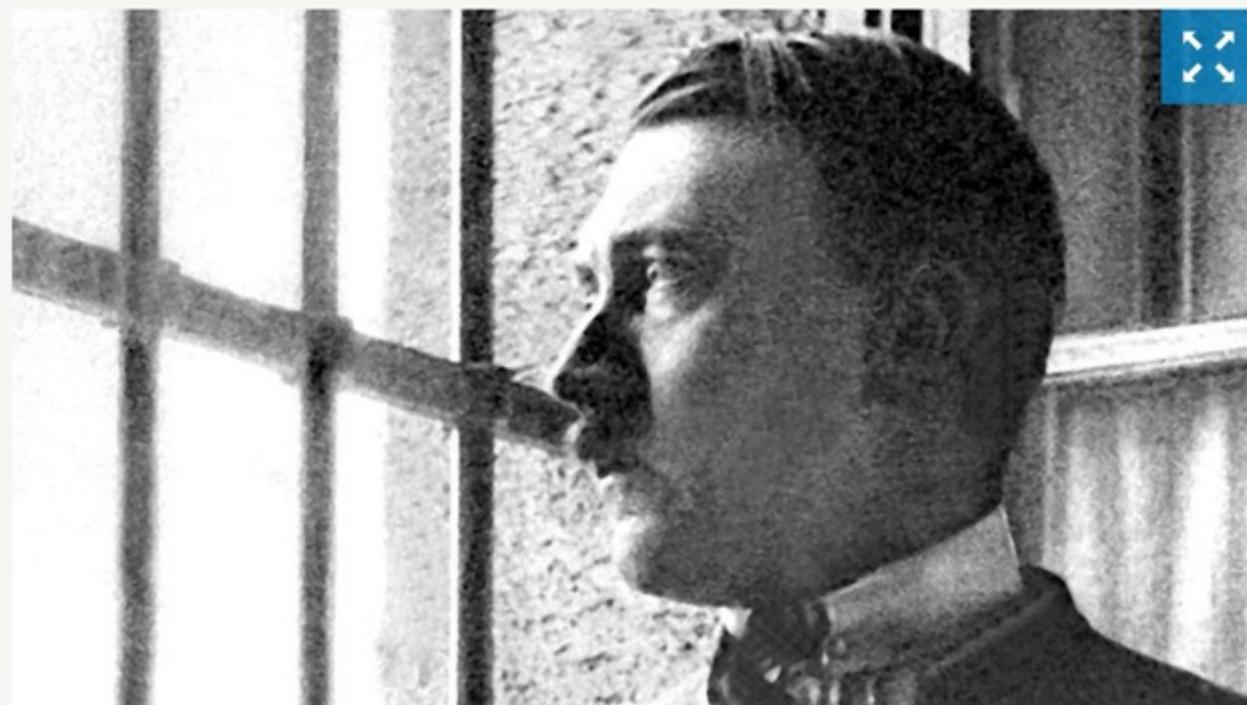


Verwöhntag Hotel Hessen »

hotel-ziegelruh.de



Spa Angebote mit Luxus und Wellness
inkl. Tageszimmer mit eigener Sauna



"So strange they didn't think
Facebook was real."

William Gibson

DARK DATA

MARC 2018

**CRIMINALS, TERRORISTS
AND HACKERS UNDERSTAND
THE POWER OF OUR
INTERCONNECTIVITY:
IF YOU CONTROL THE CODE
YOU CONTROL THE WORLD.**

lution
and
of
ers

tools

t

le

, women, and children, were perhaps the most attacks to date. The terrorists were armed, not just yes, but also with satellite phones, BlackBerrys, ery. They leveraged these technologies in real people with great





All dash



Shop by
Department ▾

Your Amazon.com

Today's Deals

Gift Cards

Sell

Help



All Electronics Deals Best Sellers TV & Video Audio & Home Theater Computers Camera & Photo Wearable Technology Car Electronics & GPS Portable Audio Cell Phones Office Elec

◀ Back to search results for "dash"



Tide Dash Button - Limited Release

by [Amazon](#)

[Be the first to review this item](#)

Price: **\$4.99**

In Stock.

Ships from and sold by Amazon.com exclusively for Prime members. [Details](#)

Want it Tuesday, Aug. 11? Order within **37 hrs 10 mins** and choose Two-Day Shipping at checkout. [Details](#)

- Limited release for Prime members - while supplies last
- Just press and never run out
- Buy from Amazon with confidence - get the same low prices we offer on Amazon.com
- Choose from favorite household products
- Compatible with your iOS or Android phone

Setup your Dash Button and Shop

Discover how your Dash Button works: [Learn more](#).

See which products work with your Dash Button: [Shop now](#).

URRY,
IT LOST
NA
TUBE
WARP

THE NEXT
MASSIVE NEW
TECHNOLOGY
IS ALREADY
ON THE WAY.
WE JUST DON'T KNOW
WHAT IT IS.

**THE FUTURE
LOVES
YOU
BUT IT DOESN'T
NEED
YOU**