universität wien

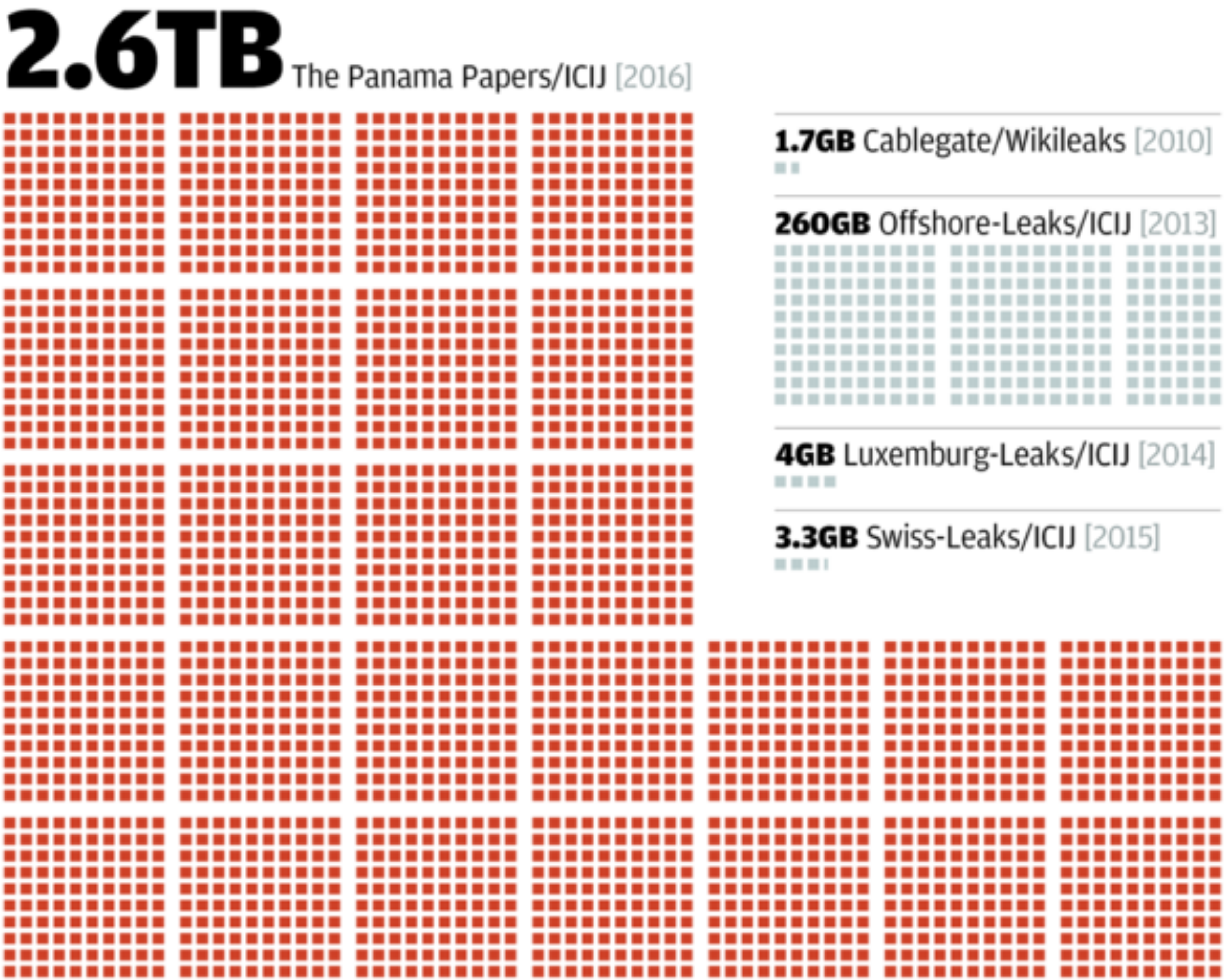
Text Visualization

Lecture in: “Visualization
and Visual Data Analysis”

thanks @ Elena Sofie Rudkowsky

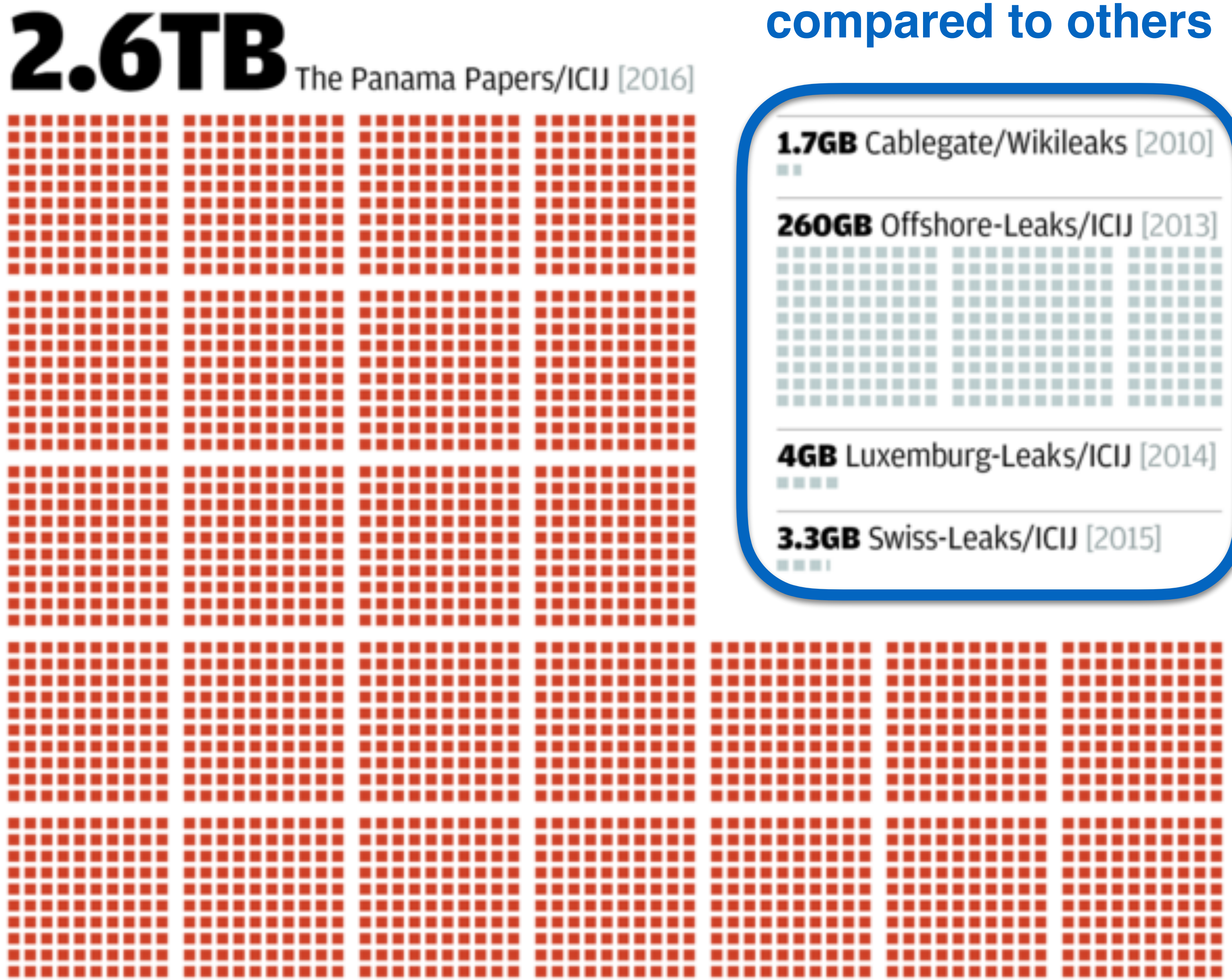
Introductory Example:

Panama Papers



http://www.computerworld.com/article/3053601/security/consider-the-panama-papers-breach-a-warning.html

Scale of the Leak compared to others



http://www.computerworld.com/article/3053601/security/consider-the-panama-papers-breach-a-warning.html

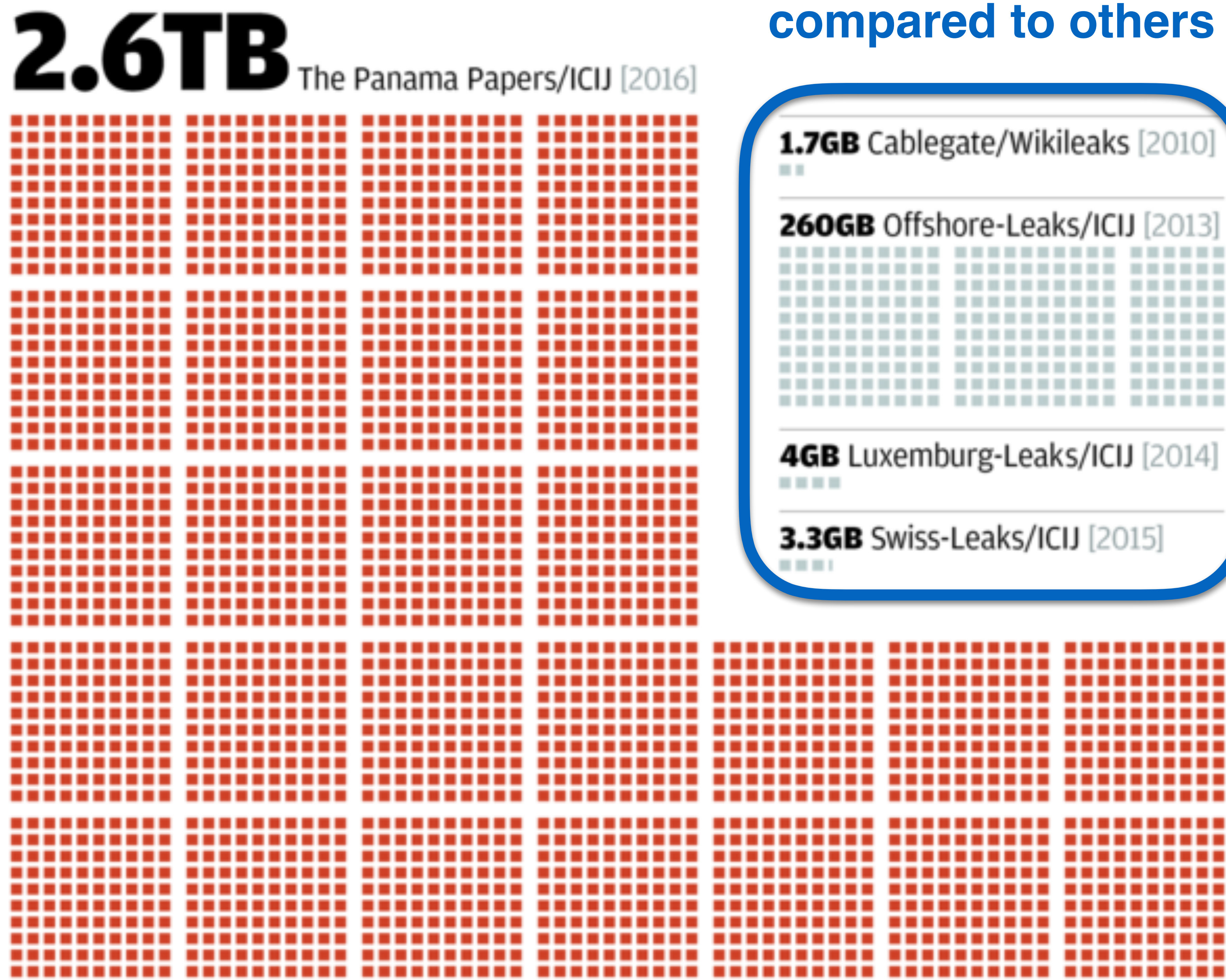
Panama Papers

Mossack Fonseca
- law office in Panama
- 214,000 letterbox companies in 21 tax havens

Letterbox Company
- address for tax purposes, money laundering…
- business carried on elsewhere/nowhere

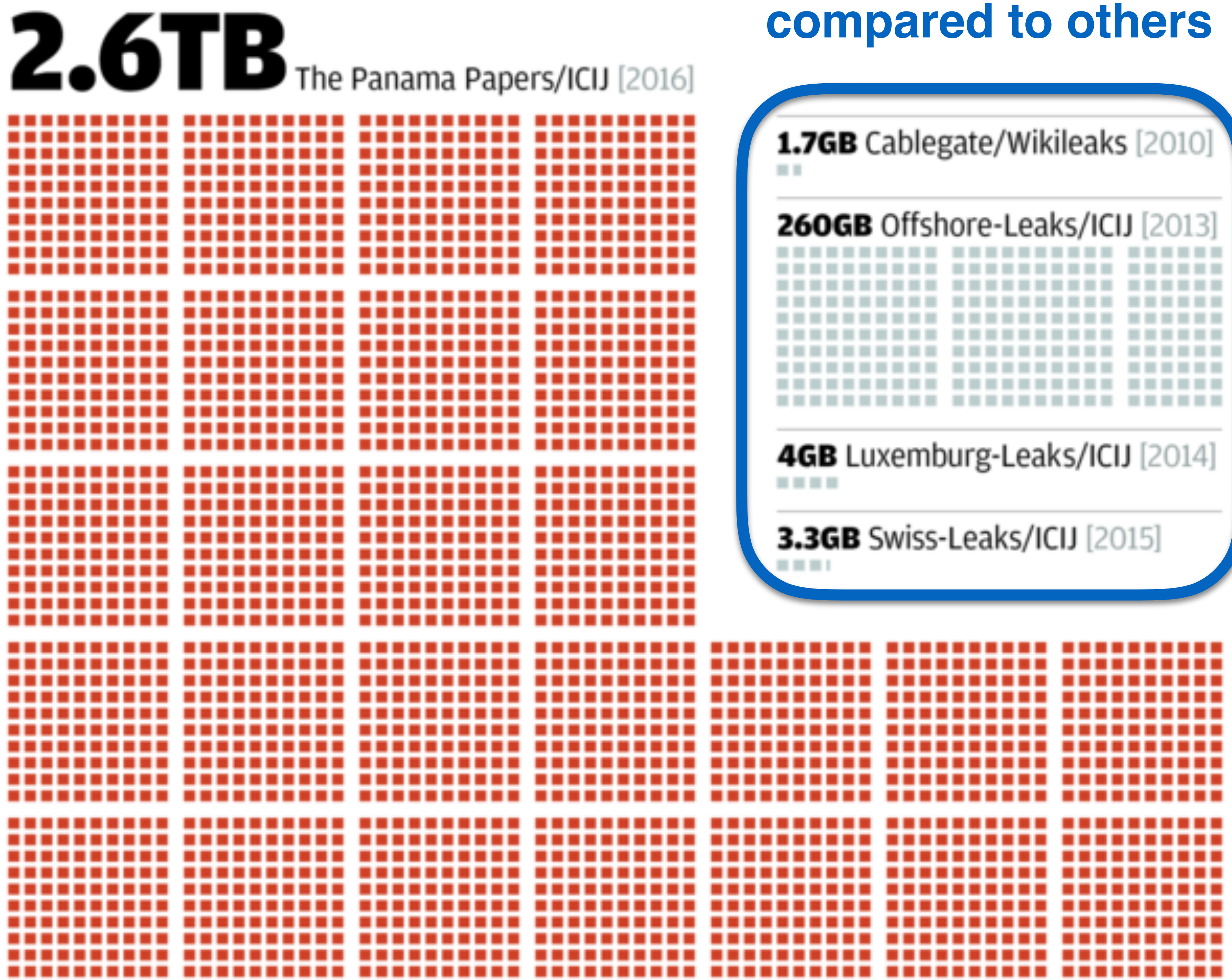
Prominent Owners
- Iceland: Prime Minister Gunnlaugsson
- Russia: Circle around president Putin
- Syria: Network around president al-Assad
- ….

Scale of the Leak compared to others



http://www.computerworld.com/article/3053601/security/consider-the-panama-papers-breach-a-warning.html

Scale of the Leak compared to others

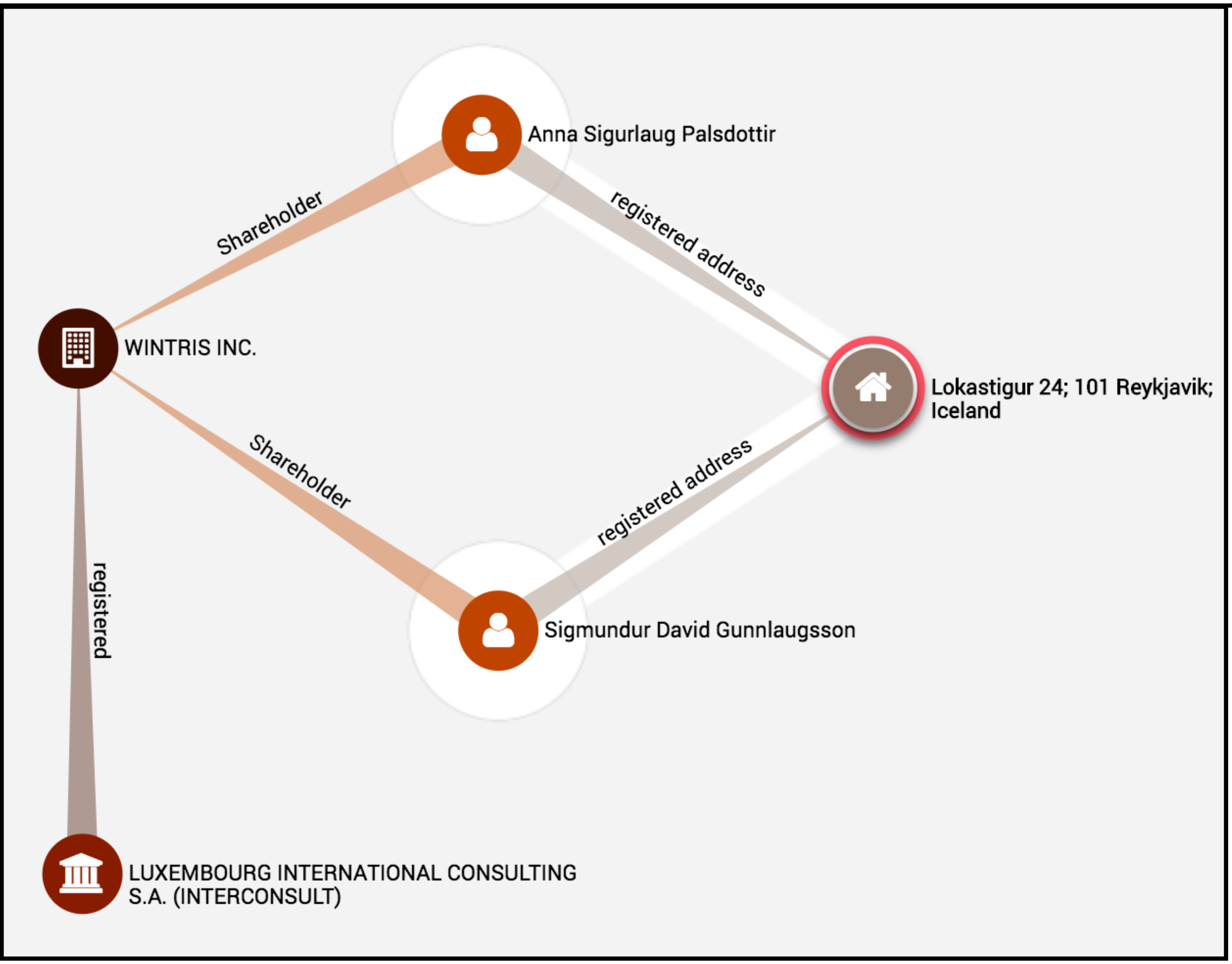


Analysis

- **400 journalists** from 80 countries
- **11.5 million documents** (PDFs, emails, text files, database format files, images…)
- **1 year research**
- **text analysis software** like Nuix

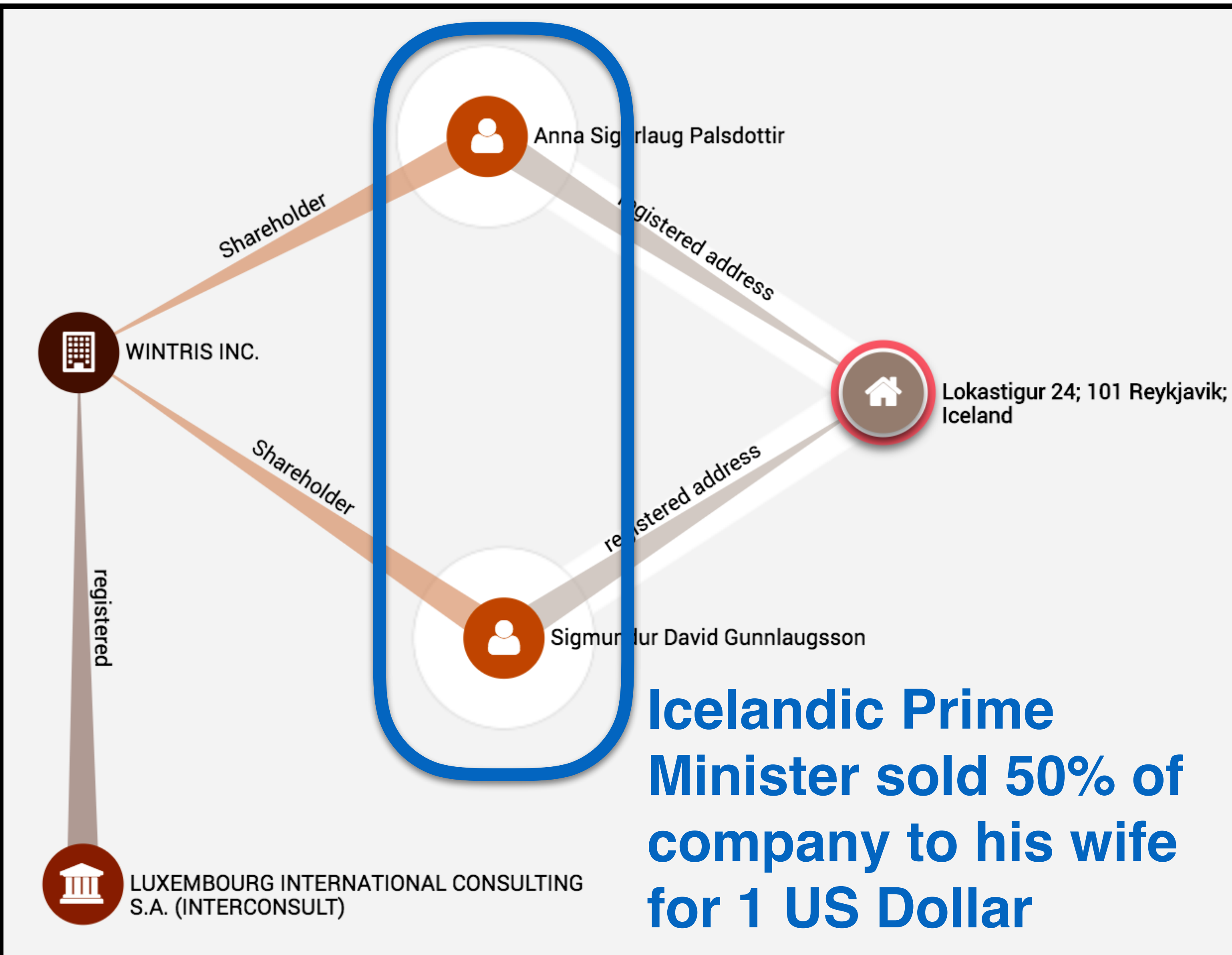
http://www.computerworld.com/article/3053601/security/consider-the-panama-papers-breach-a-warning.html

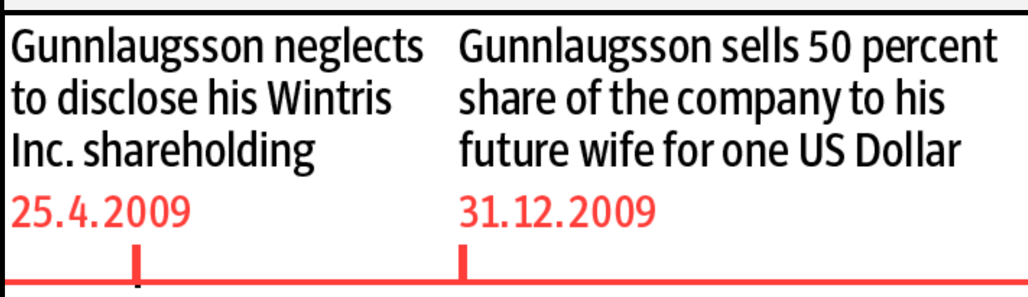
Panama Papers



https://panamapapers.icij.org/the_power_players/
http://panamapapers.sueddeutsche.de/articles/56fec0cda1bb8d3c3495adfc/

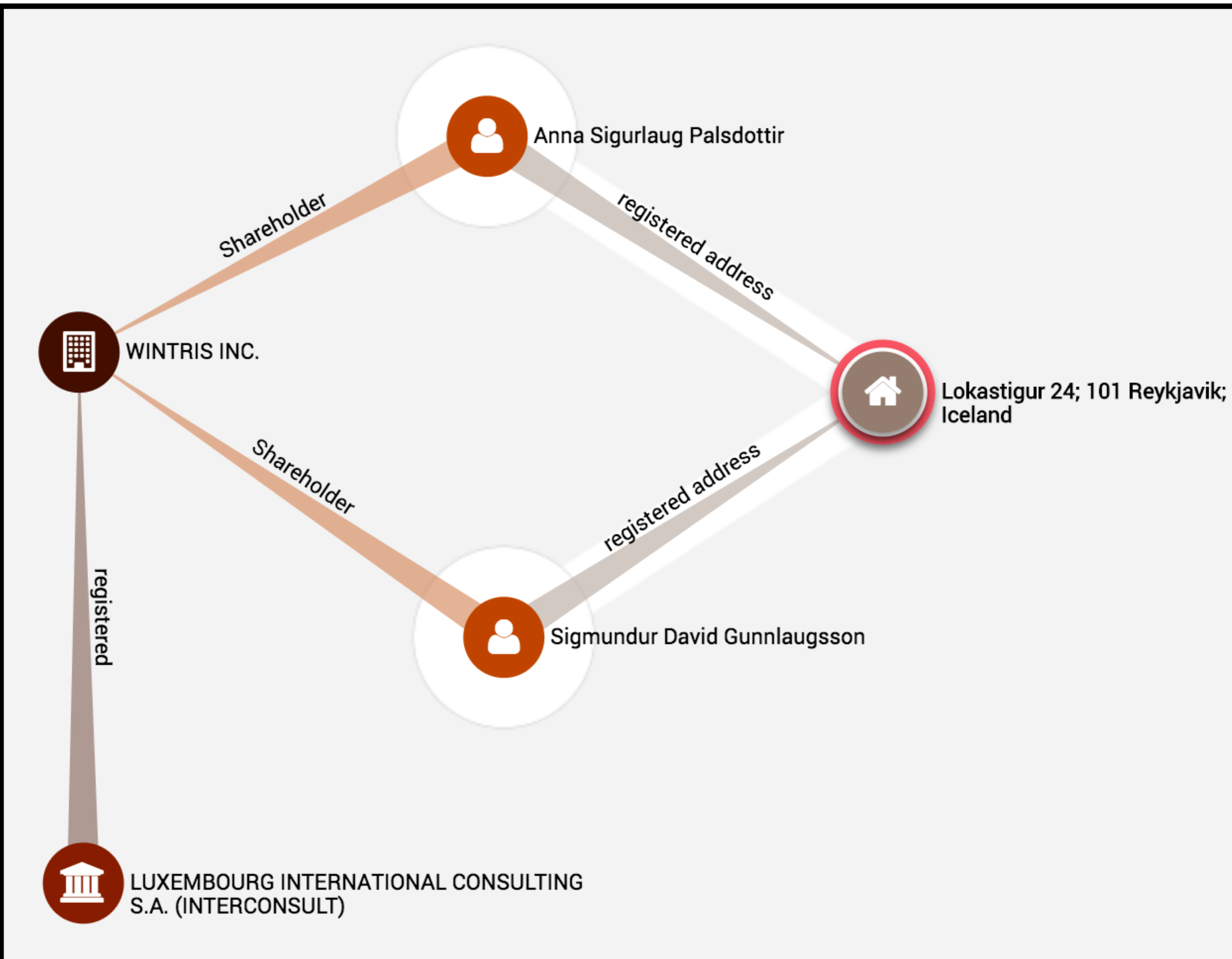
Panama Papers

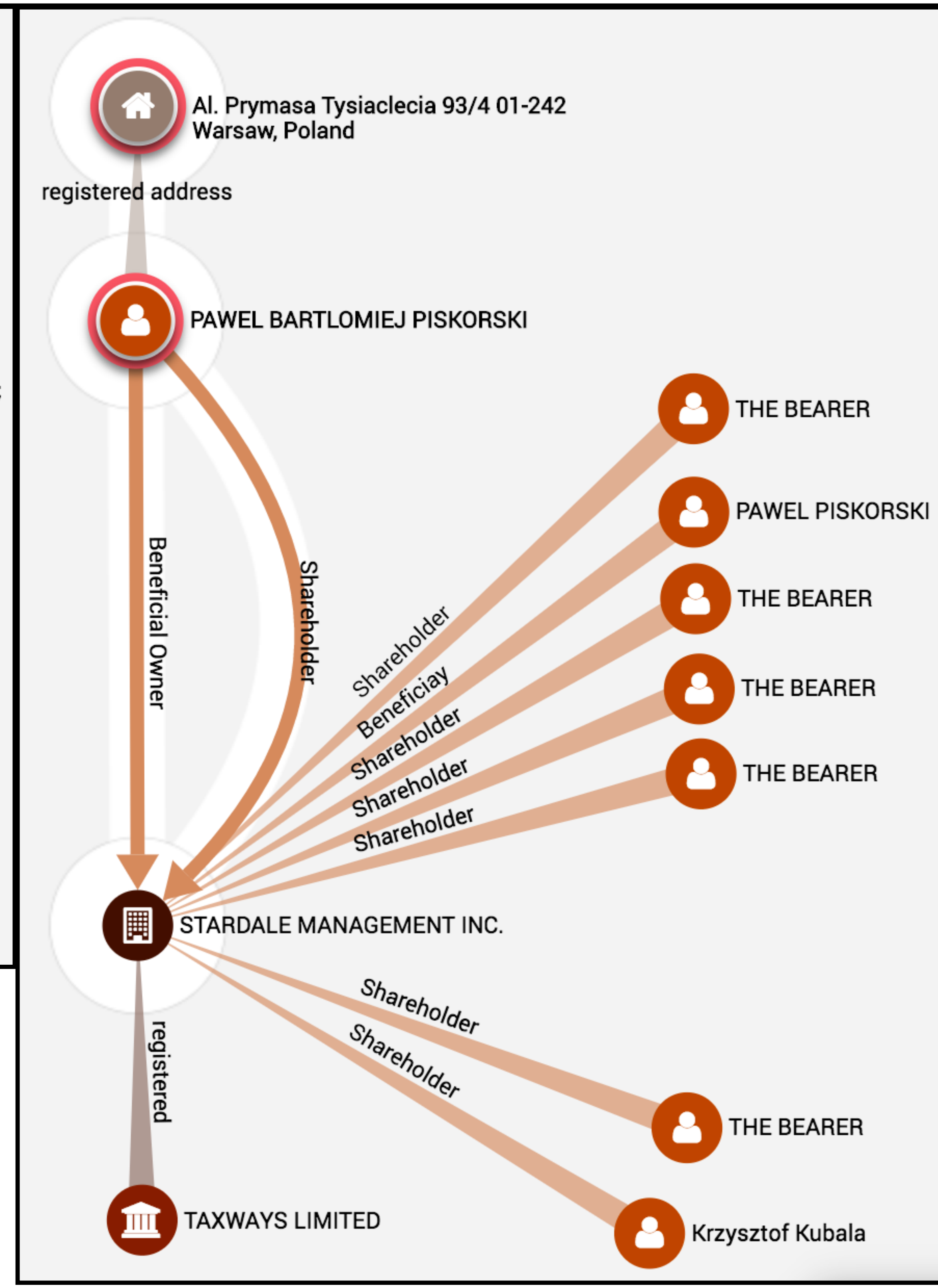




https://panamapapers.icij.org/the_power_players/
http://panamapapers.sueddeutsche.de/articles/56fec0cda1bb8d3c3495adfc/

Panama Papers





Gunnlaugsson neglects to disclose his Wintris Inc. shareholding

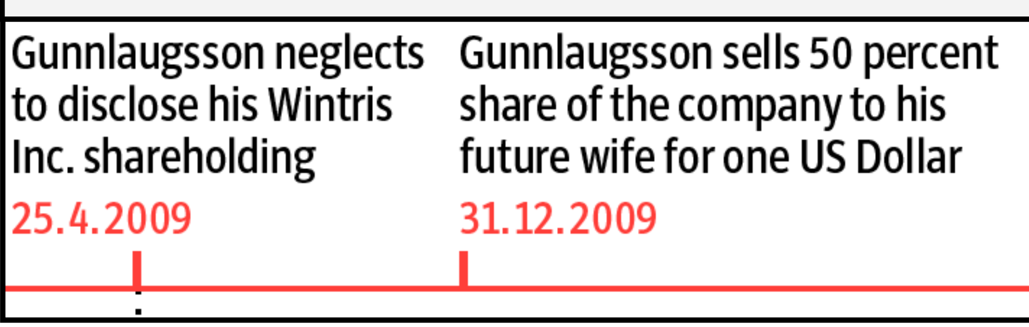
25.4.2009

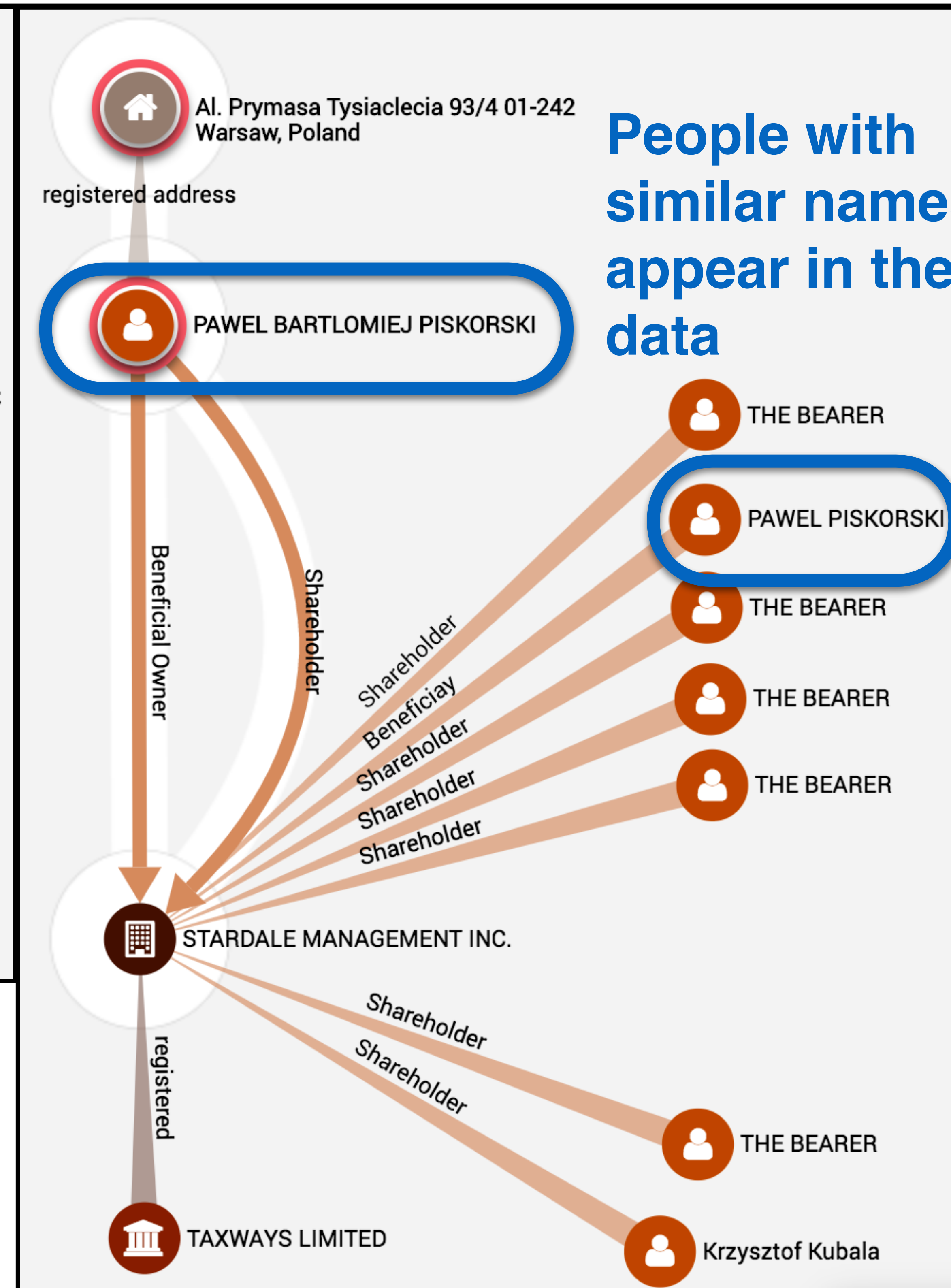
Gunnlaugsson sells 50 percent share of the company to his future wife for one US Dollar

31.12.2009

https://panamapapers.icij.org/the_power_players/
http://panamapapers.sueddeutsche.de/articles/56fec0cda1bb8d3c3495adfc/

Panama Papers





People with similar names appear in the data

https://panamapapers.icij.org/the_power_players/
http://panamapapers.sueddeutsche.de/articles/56fec0cda1bb8d3c3495adfc/

Question:

What can be extracted out of textual data?

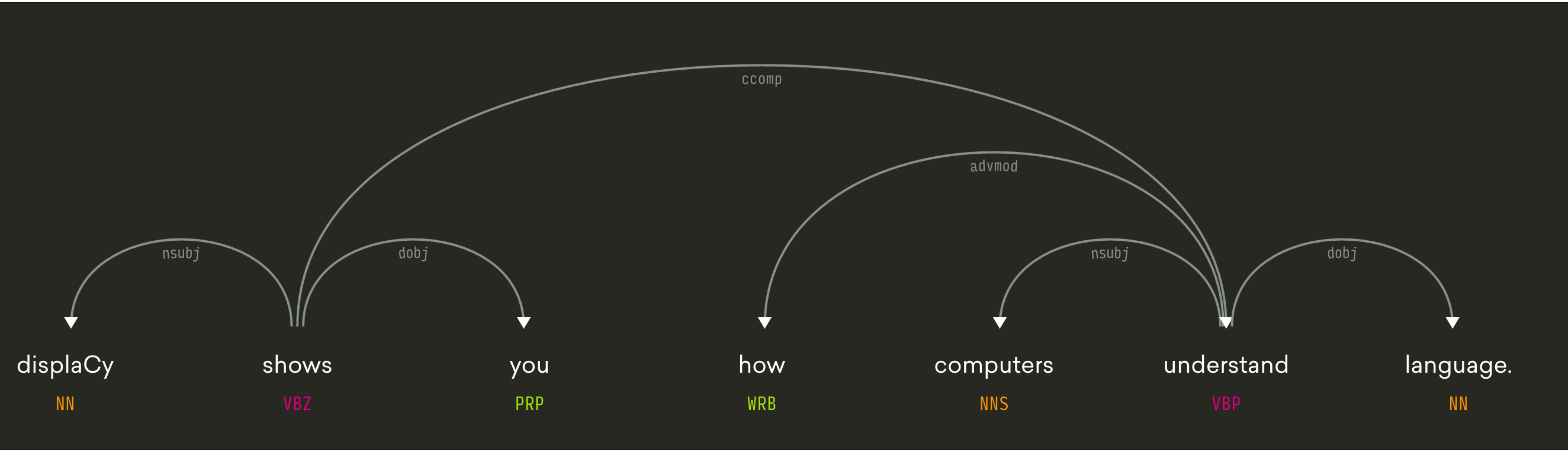
Agenda

- From Text to Structure
  - POS Tags and Parsing Trees
  - Bag-of-Words
  - Distributed Word Embeddings
  - Topic Models
  - Named Entities
  - Sentiment Analysis
  - Temporal Events
  - Deep Learning
- Projects

From Text to Structure

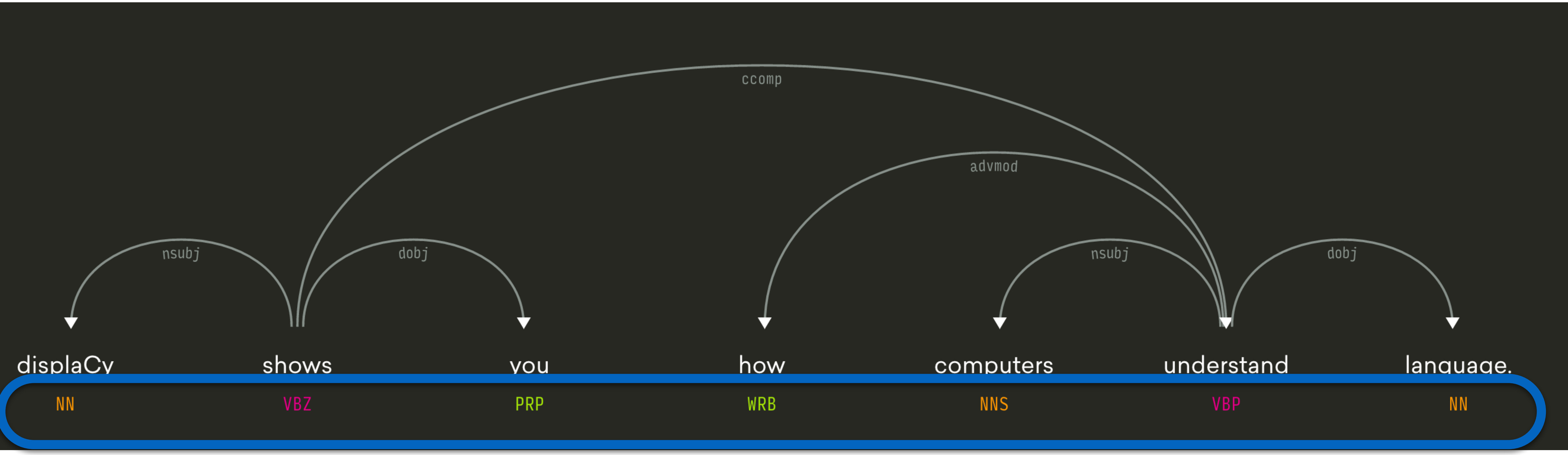
POS Tags and Parsing Trees

POS Tags and Parsing Trees



https://demos.explosion.ai/displacy/

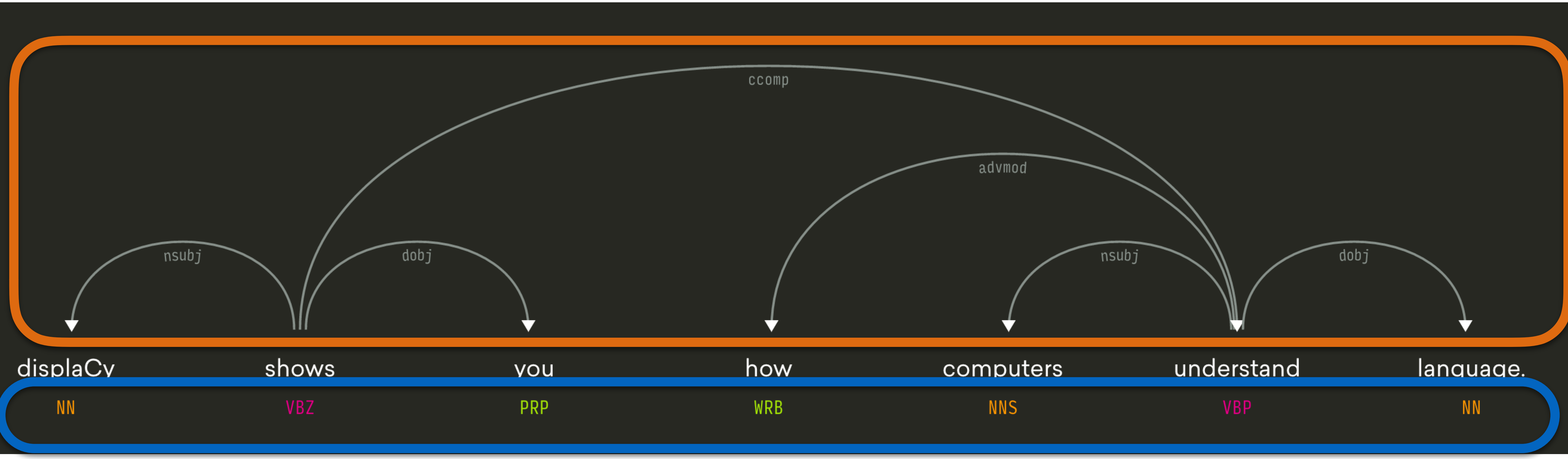
POS Tags and Parsing Trees



Part-of-speech Tags (POS Tags)
mark verbs, nouns, adjectives etc.

https://demos.explosion.ai/displacy/

Parsing (or Syntax) Trees show the syntactical structure of language like object, subject…



Part-of-speech Tags (POS Tags) mark verbs, nouns, adjectives etc.

https://demos.explosion.ai/displacy/

Question:

Are POS Tags and Syntax Trees interesting for Pattern Detection?

POS Tags and Parsing Trees

Mostly not interesting for text mining (pattern detection) in document collections!

Mostly not interesting for text mining (pattern detection) in document collections!

POS Tags and Parsing Trees

Mostly not interesting for text mining (pattern detection) in document collections!

https://www.facebook.com/NewYorkMag/videos/10154081648719826/

POS Tags and Parsing Trees

Mostly not interesting for text mining (pattern detection) in document collections!

TABLE 1

SENTENCE SIZE

Number of sentences spoken by each speaker and sentence word count statistics. Number of words in a sentence is shown by average and 50%/90% cumulative values for all, stop and non-stop words.

speaker	number of sentences	sentence size: all			stop			non-stop		
Hillary Clinton	1,206	15.7•	21	43	9.0•	12	26	7.0•	9	19
Donald Trump	1,970	10.9•	15	36	6.6•	9	22	4.6•	6	16
total	3,176	14.8	19	40	9.5	11	25	7.5	9	19

Fields with • (e.g. 155•) link to data files and Wordles. Hover over the field to show these links. See analysis.

TABLE 2

PART OF SPEECH COUNT

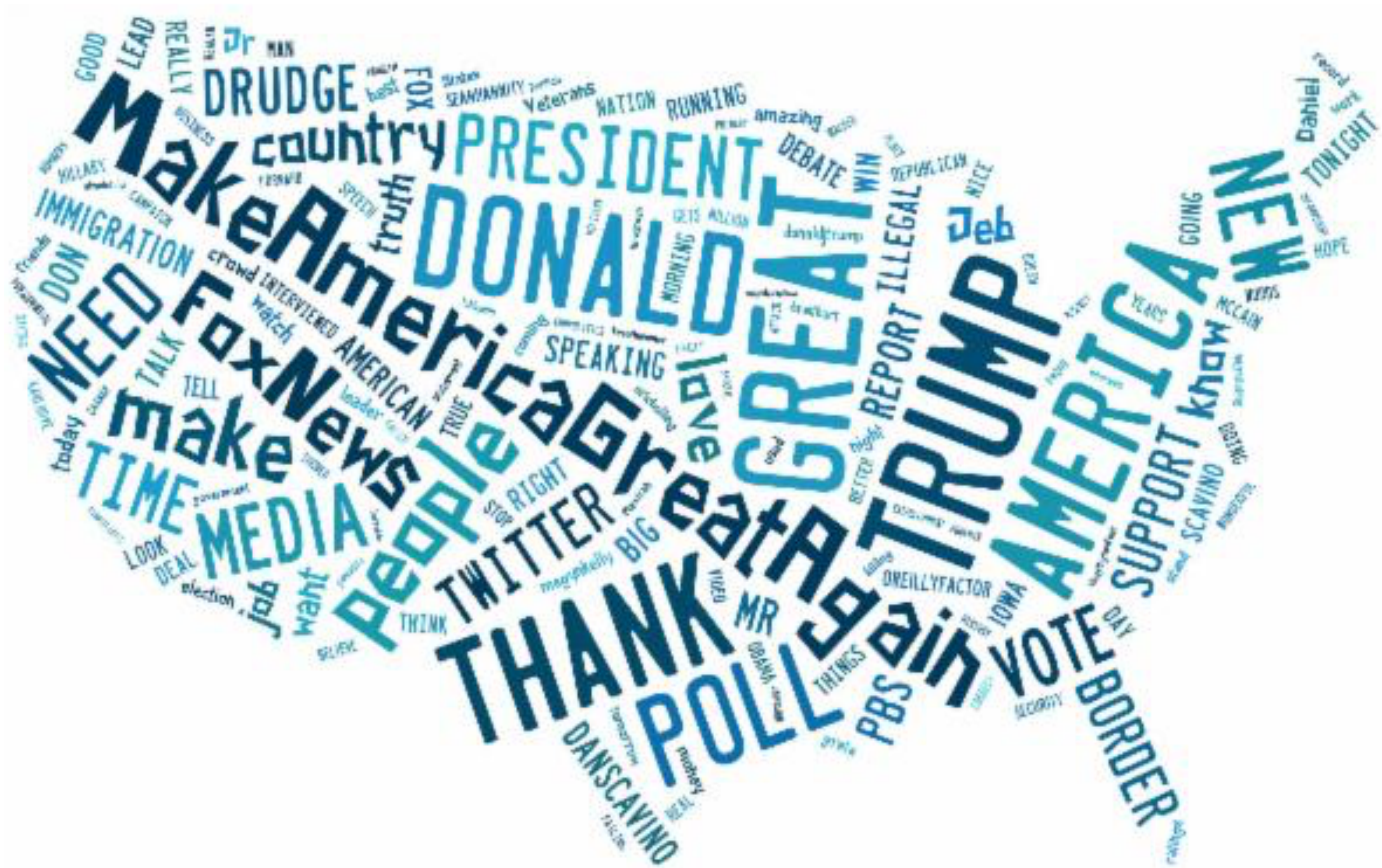
Count of words categorized by part of speech (POS).

	part of speech: n+v+adj+adv		nouns (n)		verbs (v)		adjectives (adj)		adverbs (adv)	
Hillary Clinton	7,636• 40.5%	2,173 28.5%	3,581• 46.9%	1,117 31.2%	2,322• 30.4%	799 34.4%	1,306• 17.1%	543 41.6%	427• 5.6%	97 22.7%
Donald Trump	8,158• 37.9%	1,752 21.5%	3,639• 44.6%	953 26.2%	2,375• 29.1%	588 24.8%	1,621• 19.9%	550 33.9%	523• 6.4%	83 15.9%
total	15,794• 39.1%	3,008 19.0%	7,220• 45.7%	1,635 22.6%	4,697• 29.7%	1,102 23.5%	2,927• 18.5%	886 30.3%	950• 6.0%	127 13.4%

Fields with • (e.g. 155•) link to data files and Wordles. Hover over the field to show these links. See analysis.

http://mkweb.bcgsc.ca/debates2016/

Bag-Of-Words (BOW)


MakeAmericaGreatAgain
DONALD
TRUMP
GREAT
AMERICA
THANK
PRESIDENT
FoxNews
DRUDGE
country
NEED
make
people
MEDIA
TIME
POLL
TWITTER
NEW
IMMIGRATION
SPEAKING
VOTE
BORDER
SUPPORT
REPORT
ILLEGAL
love
Jeb
DANSCAVINO
PBS
MR
AMERICAN
RIGHT
BIG
TONIGHT
know
SCAVINO
IOWA
OREILLYFACTOR
DEBATE
RUNNING
NATION
Veterans
amazing
truth
crowd
watch
job
want
DEAL
LOOK
TALK
TELL
TRUE
THINK
DON
today
GOOD
LEAD
REALLY
FOX
WIN
GOING
HOPE
MCCAIN
Daniel
THINGS
OBAMA
election

Word Cloud

^ All nouns in debates, colored by contributing speaker (Clinton: blue, Trump: red, spoken by both: grey).

^ All verbs in debates, colored by contributing speaker (Clinton: blue, Trump: red, spoken by both: grey).

http://mkweb.bcgsc.ca/debates2016/

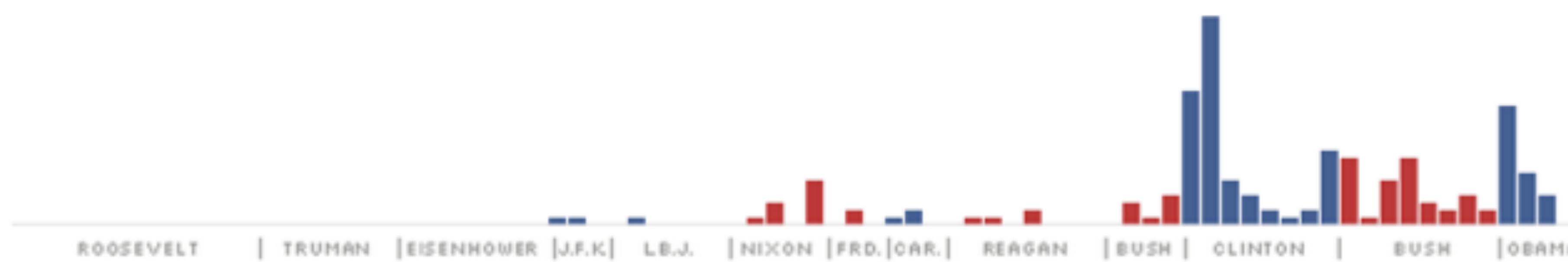
Word Frequencies over Time

Patterns of Speech: 75 Years of the State of the Union Addresses
(Data Story, Washington Post 2011)

'health care'

The expansion of health insurance coverage remains unpopular with nearly half the country, but Mr. Obama defended the health care law in his 2011 speech, though he added that he was willing with Republicans to improve it. In 1994, Mr. Clinton promoted his plan, which collapsed that year.

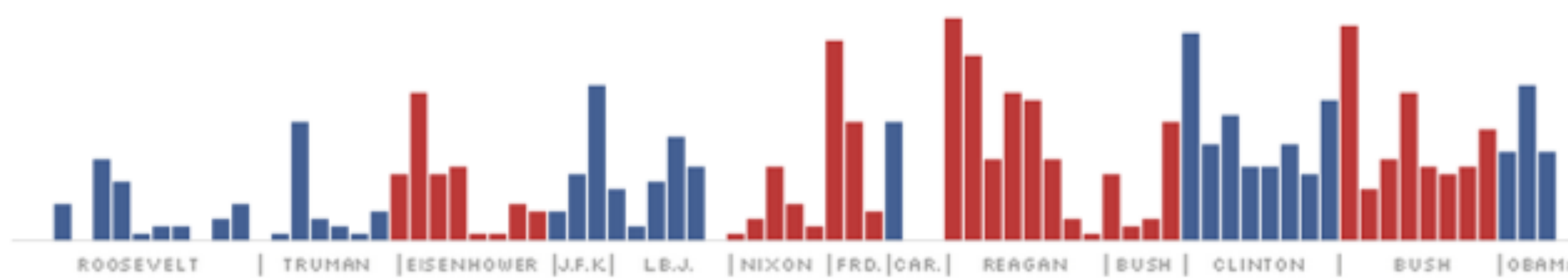
HEALTH CARE



'tax'

Presidents have used the word every year since 1981, when Mr. Reagan uttered it 30 times, detailing his plan to reduce taxes and government spending.

TAX, TAXED, TAXES, TAXING



Patterns of Speech: 75 Years of the State of the Union Addresses
(Data Story, Washington Post 2011)

‘health care’
became popular late

Document Term Matrix

	Document 1	Document 2	Document 3	Document 4	Document 5	Document 6	Document 7	Document 8
Term(s) 1	10	0	1	0	0	0	0	2
Term(s) 2	0	2	0	0	0	18	0	2
Term(s) 3	0	0	0	0	0	0	0	2
Term(s) 4	6	0	0	4	6	0	0	0
Term(s) 5	0	0	0	0	0	0	0	2
Term(s) 6	0	0	1	0	0	1	0	0
Term(s) 7	0	1	8	0	0	0	0	0
Term(s) 8	0	0	0	0	0	3	0	0

Word Vector
(Passage Vector)

Document Vector

Bag of Words (BOW)
counts words per document

Document Term Matrix

document corpus

vocabulary= terms within the corpus

	Document 1	Document 2	Document 3	Document 4	Document 5	Document 6	Document 7	Document 8
Term(s) 1	10	0	1	0	0	0	0	2
Term(s) 2	0	2	0	0	0	18	0	2
Term(s) 3	0	0	0	0	0	0	0	2
Term(s) 4	6	0	0	4	6	0	0	0
Term(s) 5	0	0	0	0	0	0	0	2
Term(s) 6	0	0	1	0	0	1	0	0
Term(s) 7	0	1	8	0	0	0	0	0
Term(s) 8	0	0	0	0	0	3	0	0

Word Vector (Passage Vector)

Document Vector

Document Term Matrix

	Document 1	Document 2	Document 3	Document 4	Document 5	Document 6	Document 7	Document 8
Term(s) 1	10	0	1	0	0	0	0	2
Term(s) 2	0	2	0	0	0	18	0	2
Term(s) 3	0	0	0	0	0	0	0	2
Term(s) 4	6	0	0	4	6	0	0	0
Term(s) 5	0	0	0	0	0	0	0	2
Term(s) 6	0	0	1	0	0	1	0	0
Term(s) 7	0	1	8	0	0	0	0	0
Term(s) 8	0	0	0	0	0	3	0	0

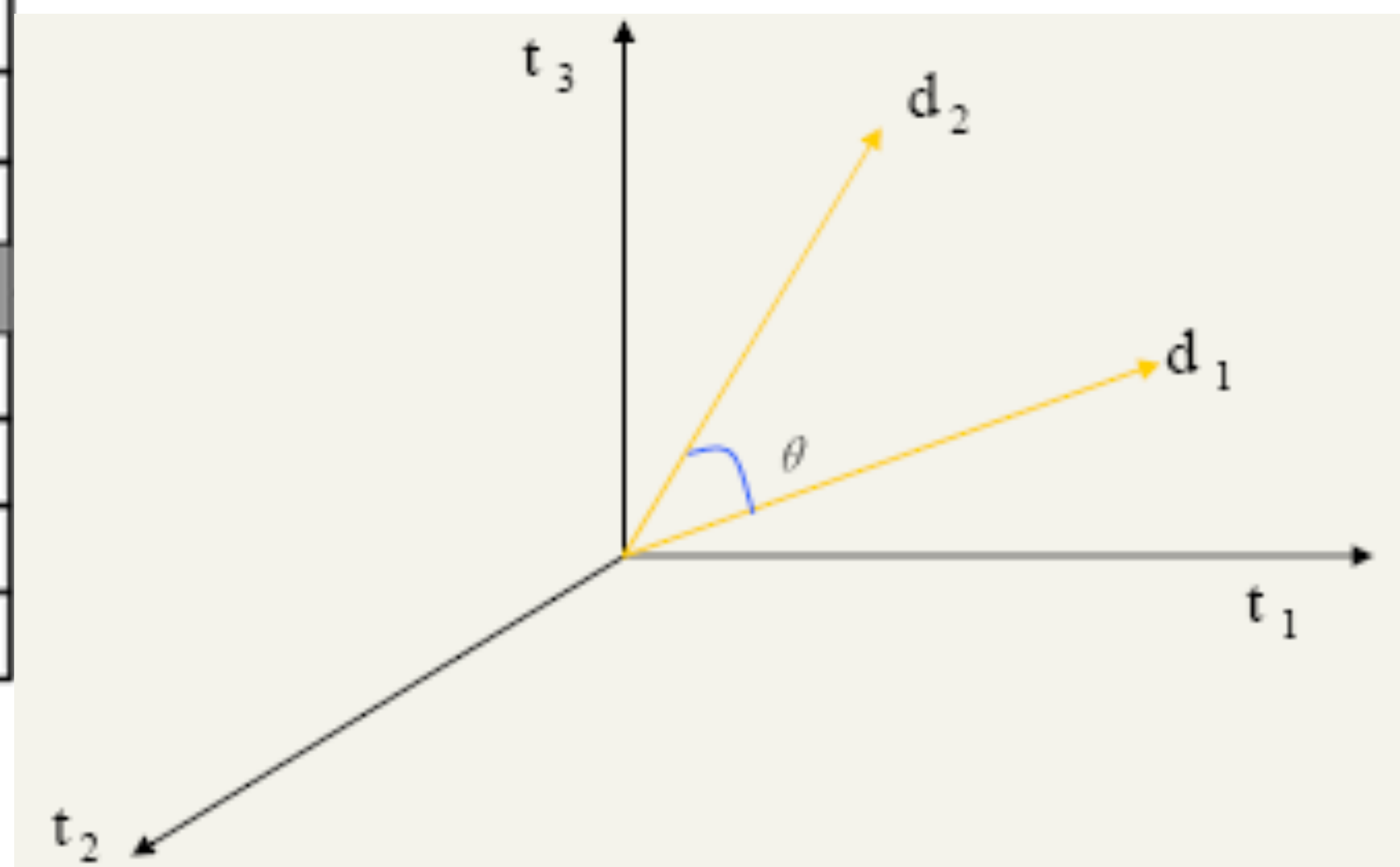
Word Vector
(Passage Vector)

Document Vector

Document Term Matrix

	Document 1	Document 2	Document 3	Document 4	Document 5	Document 6	Document 7	Document 8
Term(s) 1	10	0	1	0	0	0	0	2
Term(s) 2	0	2	0	0	0	18	0	2
Term(s) 3	0	0	0	0	0	0	0	2
Term(s) 4	6	0	0	4	6	0	0	0
Term(s) 5	0	0	0	0	0	0	0	2
Term(s) 6	0	0	1	0	0	1	0	0
Term(s) 7	0	1	8	0	0	0	0	0
Term(s) 8	0	0	0	0	0	3	0	0

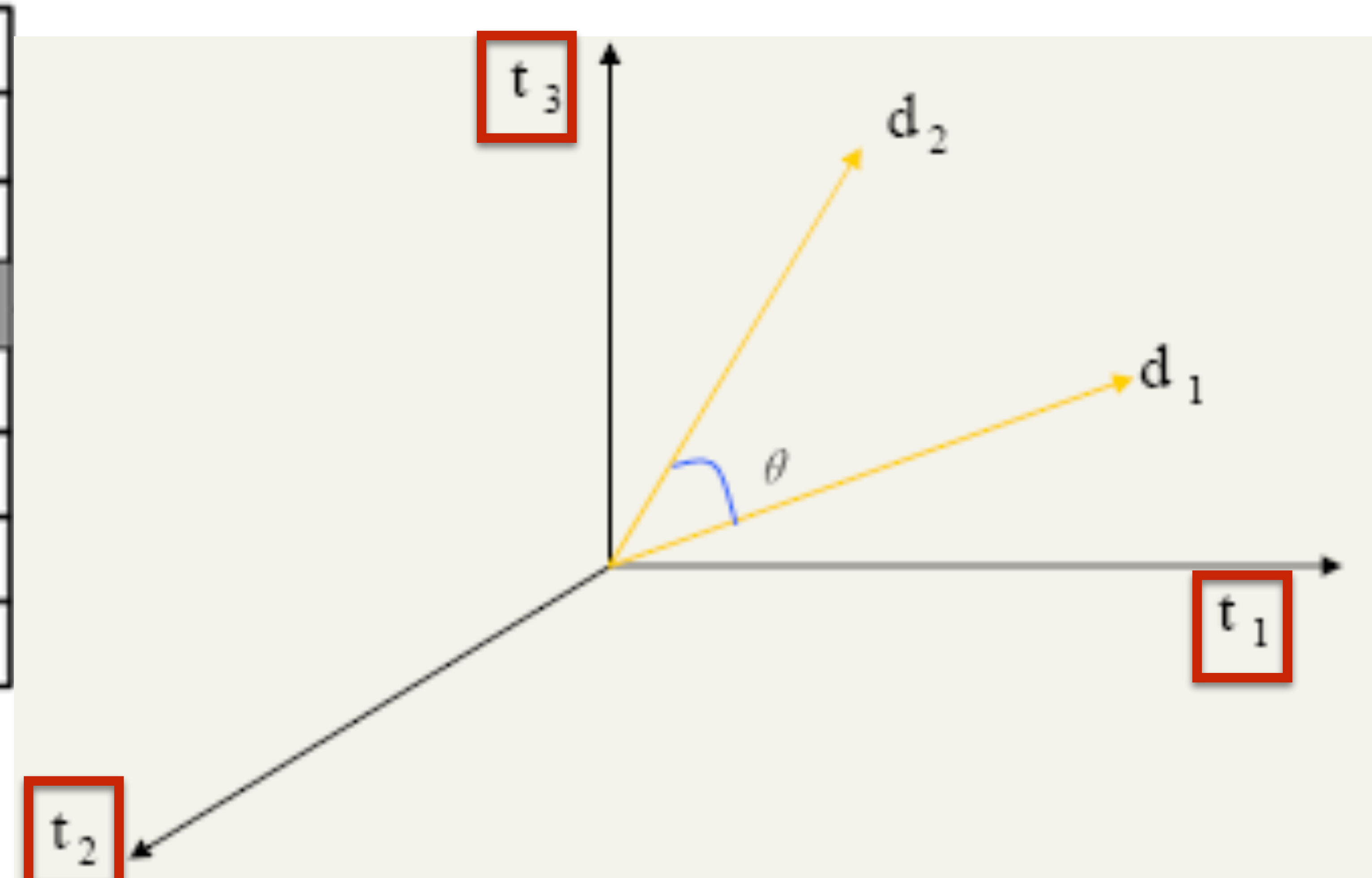
Document Vector



Document Term Matrix

	Document 1	Document 2	Document 3	Document 4	Document 5	Document 6	Document 7	Document 8
Term(s) 1	10	0	1	0	0	0	0	2
Term(s) 2	0	2	0	0	0	18	0	2
Term(s) 3	0	0	0	0	0	0	0	2
Term(s) 4	6	0	0	4	6	0	0	0
Term(s) 5	0	0	0	0	0	0	0	2
Term(s) 6	0	0	1	0	0	1	0	0
Term(s) 7	0	1	8	0	0	0	0	0
Term(s) 8	0	0	0	0	0	3	0	0

Document Vector

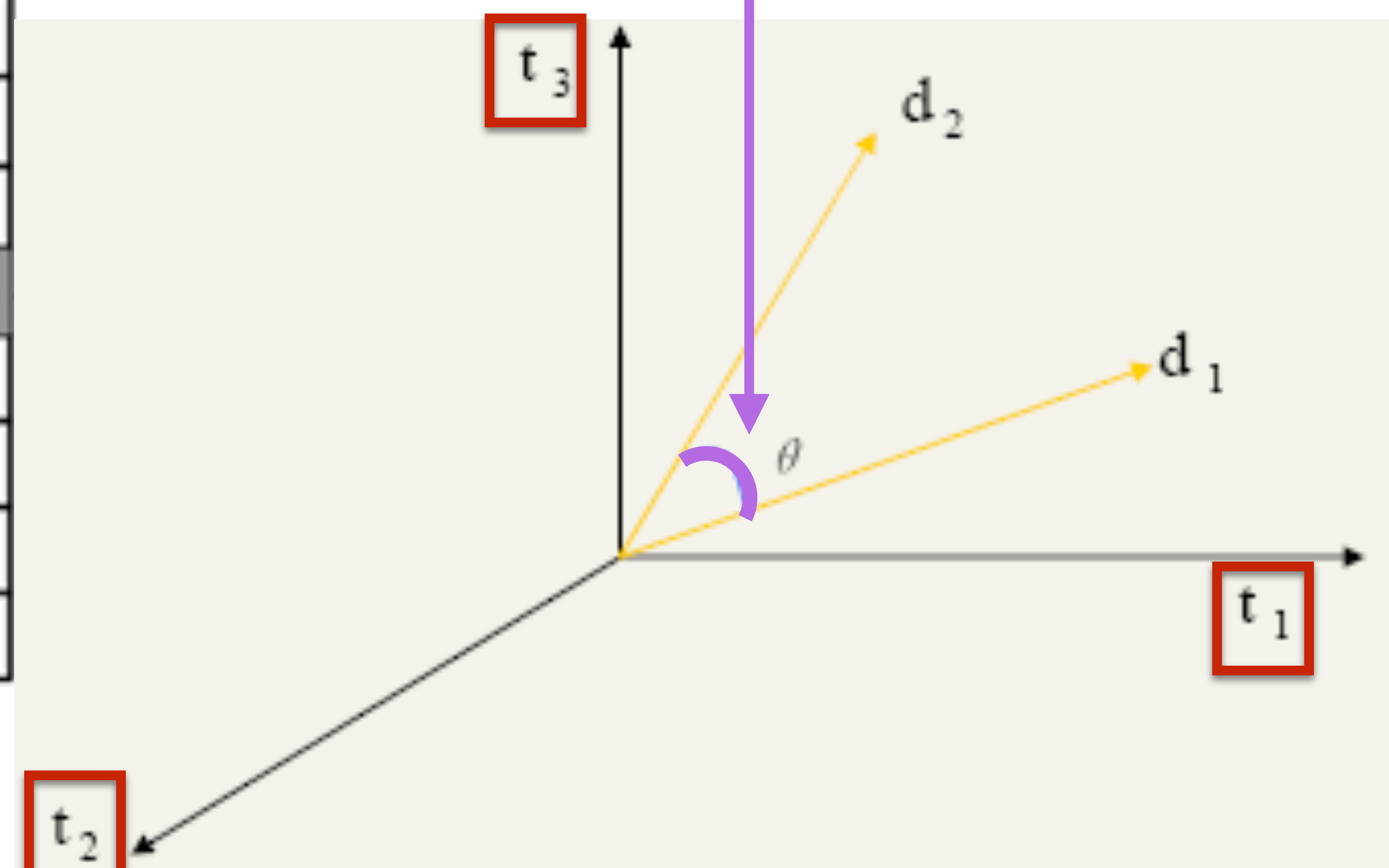


Document Term Matrix

Cosine Similarity = θ (angle)
Distance between Document Vectors

	Document 1	Document 2	Document 3	Document 4	Document 5	Document 6	Document 7	Document 8
Term(s) 1	10	0	1	0	0	0	0	2
Term(s) 2	0	2	0	0	0	18	0	2
Term(s) 3	0	0	0	0	0	0	0	2
Term(s) 4	6	0	0	4	6	0	0	0
Term(s) 5	0	0	0	0	0	0	0	2
Term(s) 6	0	0	1	0	0	1	0	0
Term(s) 7	0	1	8	0	0	0	0	0
Term(s) 8	0	0	0	0	0	3	0	0

Document Vector

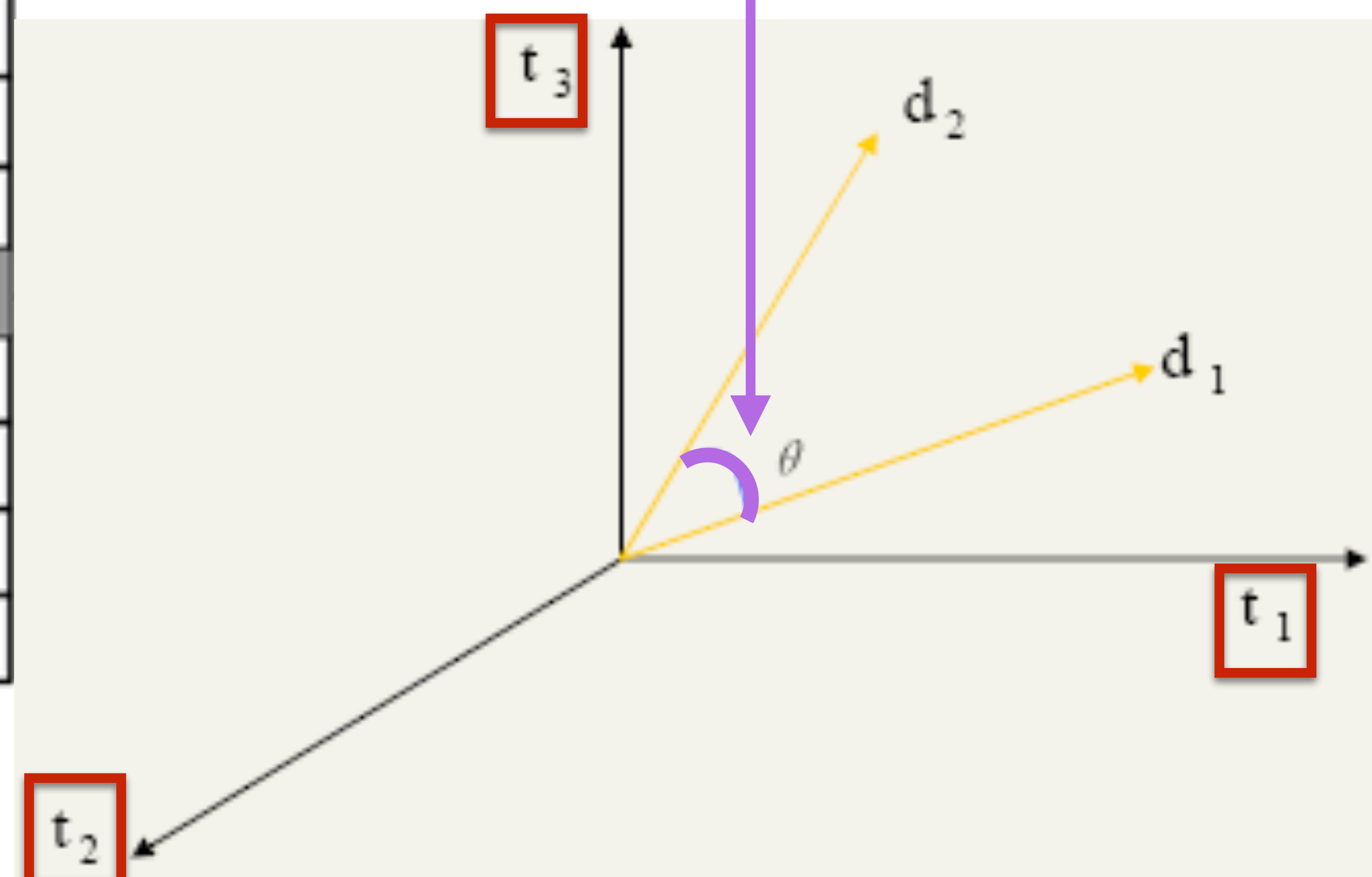


Document Term Matrix

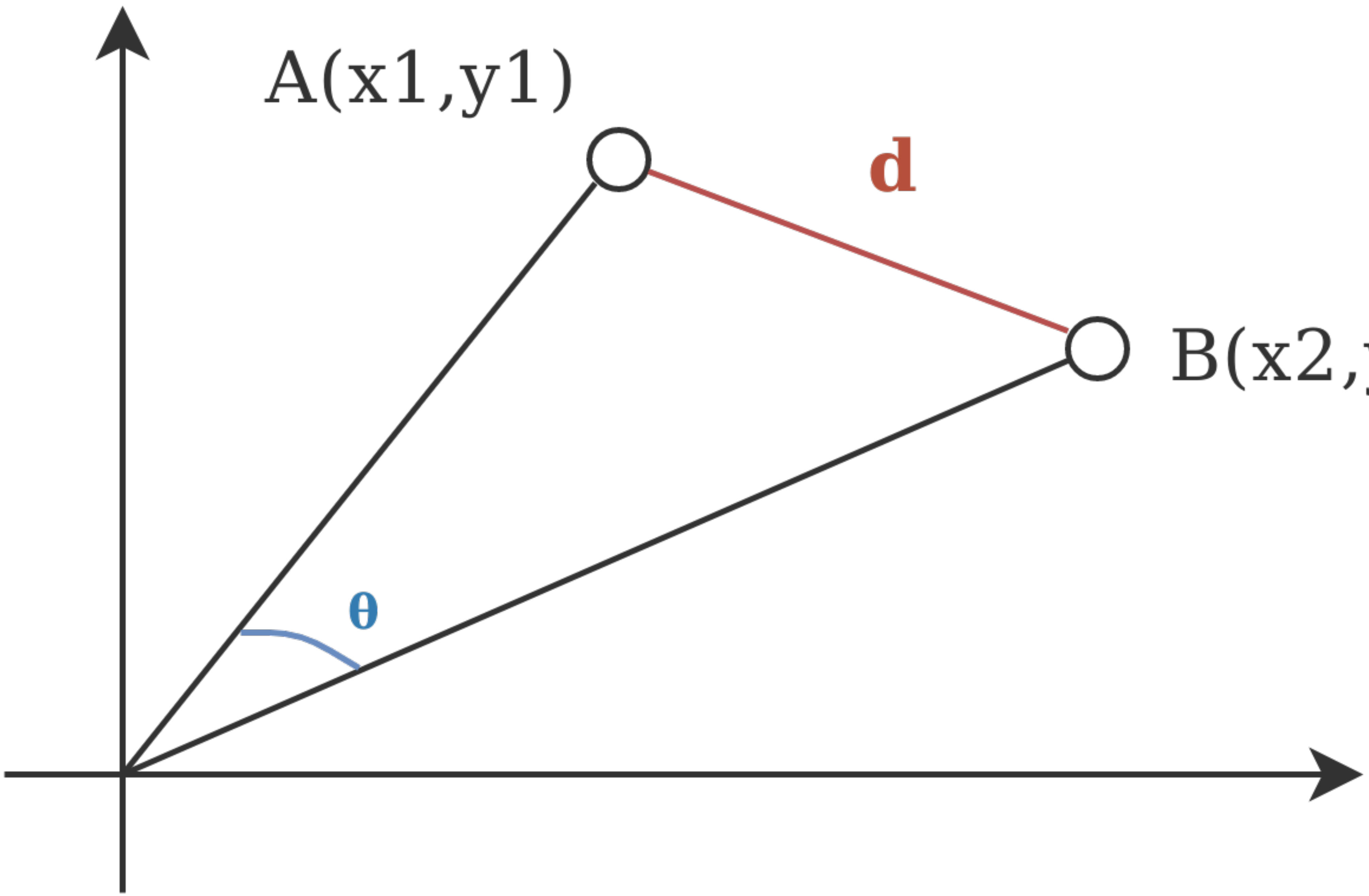
	Document 1	Document 2	Document 3	Document 4	Document 5	Document 6	Document 7	Document 8
Term(s) 1	10	0	1	0	0	0	0	2
Term(s) 2	0	2	0	0	0	18	0	2
Term(s) 3	0	0	0	0	0	0	0	2
Term(s) 4	6	0	0	4	6	0	0	0
Term(s) 5	0	0	0	0	0	0	0	2
Term(s) 6	0	0	1	0	0	1	0	0
Term(s) 7	0	1	8	0	0	0	0	0
Term(s) 8	0	0	0	0	0	3	0	0

Document Vector

Cosine Similarity - Why not using Euclidean Distance (Similarity)?



Document Term Matrix



https://cmry.github.io/notes/euclidean-v-cosine

Is counting good enough?
Ideas for improvement?

Term Frequency - Inverse Document Frequency (TF-IDF)

Term Frequency - Inverse Document Frequency

We introduce a new weight - consisting of two parts!

	Document 1	Document 2	Document 3	Document 4	Document 5	Document 6	Document 7	Document 8
Term(s) 1	10	0	1	0	0	0	0	2
Term(s) 2	0	2	0	0	0	18	0	2
Term(s) 3	0	0	0	0	0	0	0	2
Term(s) 4	6	0	0	4	6	0	0	0
Term(s) 5	0	0	0	0	0	0	0	2
Term(s) 6	0	0	1	0	0	1	0	0
Term(s) 7	0	1	8	0	0	0	0	0
Term(s) 8	0	0	0	0	0	3	0	0

Document Vector

term frequency (how often does the term appear in the current document?) normalised according to document length

document frequency (how many documents contain the term?)

$$\boxed{w_{i,j}} = tf_{i,j} \times \log\left(\frac{N}{df_i}\right)$$

$tf_{i,j}$ = number of occurrences of i in j

df_i = number of documents containing i

N = total number of documents

Term Frequency - Inverse Document Frequency

If a term appears in many documents
(term is not interesting & weight = low)

	Document 1	Document 2	Document 3	Document 4	Document 5	Document 6	Document 7	Document 8
Term(s) 1	10	0	1	0	0	0	0	2
Term(s) 2	0	2	0	0	0	18	0	2
Term(s) 3	0	0	0	0	0	0	0	2
Term(s) 4	6	0	0	4	6	0	0	0
Term(s) 5	0	0	0	0	0	0	0	2
Term(s) 6	0	0	1	0	0	1	0	0
Term(s) 7	0	1	8	0	0	0	0	0
Term(s) 8	0	0	0	0	0	3	0	0

Document Vector

df = high
N/df = low
log(N/df) = low
weight = low

df = low
N/df = high
log(N/df) = high
weight = high

$$\boxed{w_{i,j}} = tf_{i,j} \times \log\left(\frac{N}{df_i}\right)$$

$tf_{i,j}$ = number of occurrences of i in j
df_i = number of documents containing i
N = total number of documents

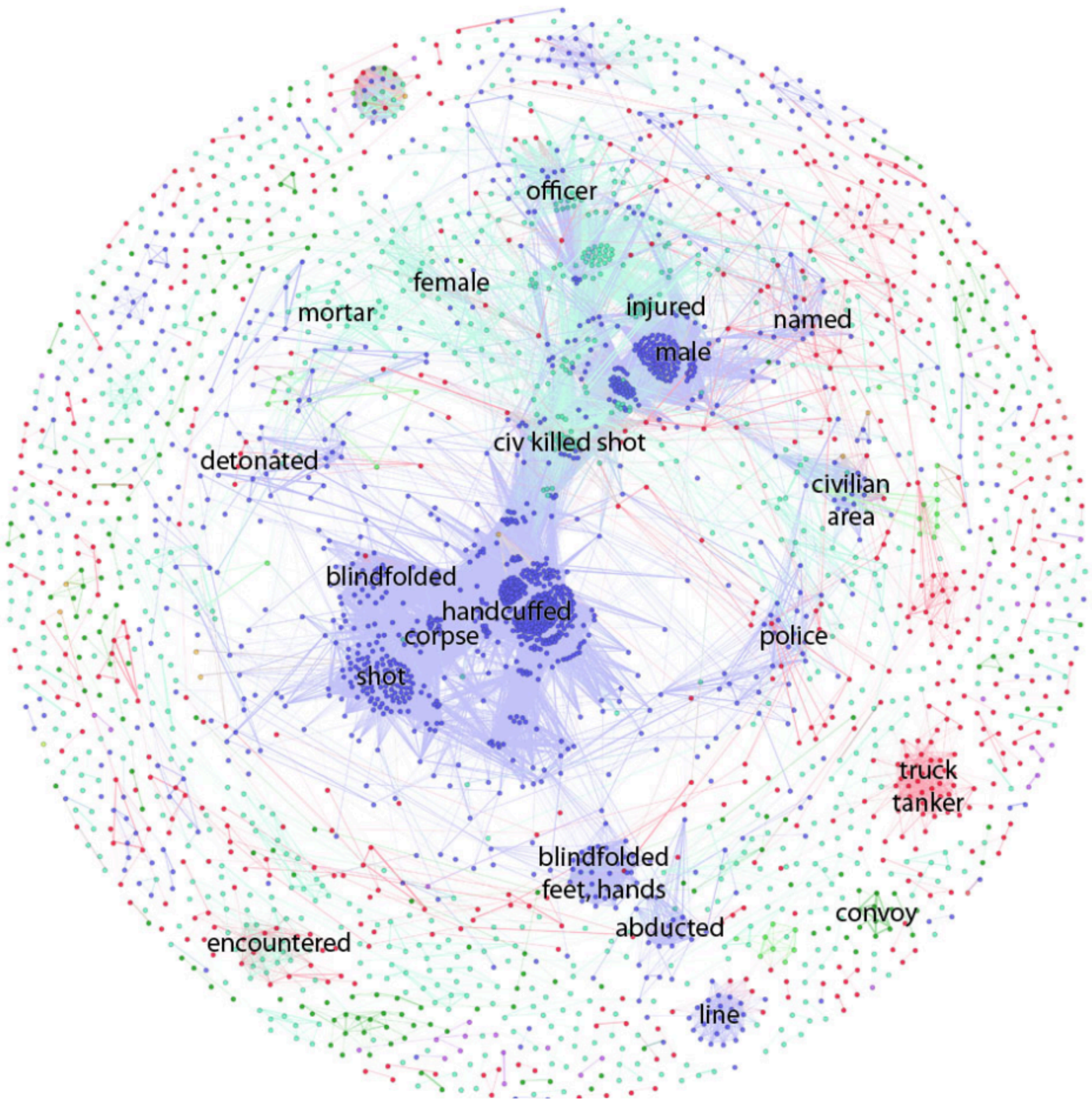
Document Term Matrix Visualization:

WikiLeaks Iraq War Logs

universität wien

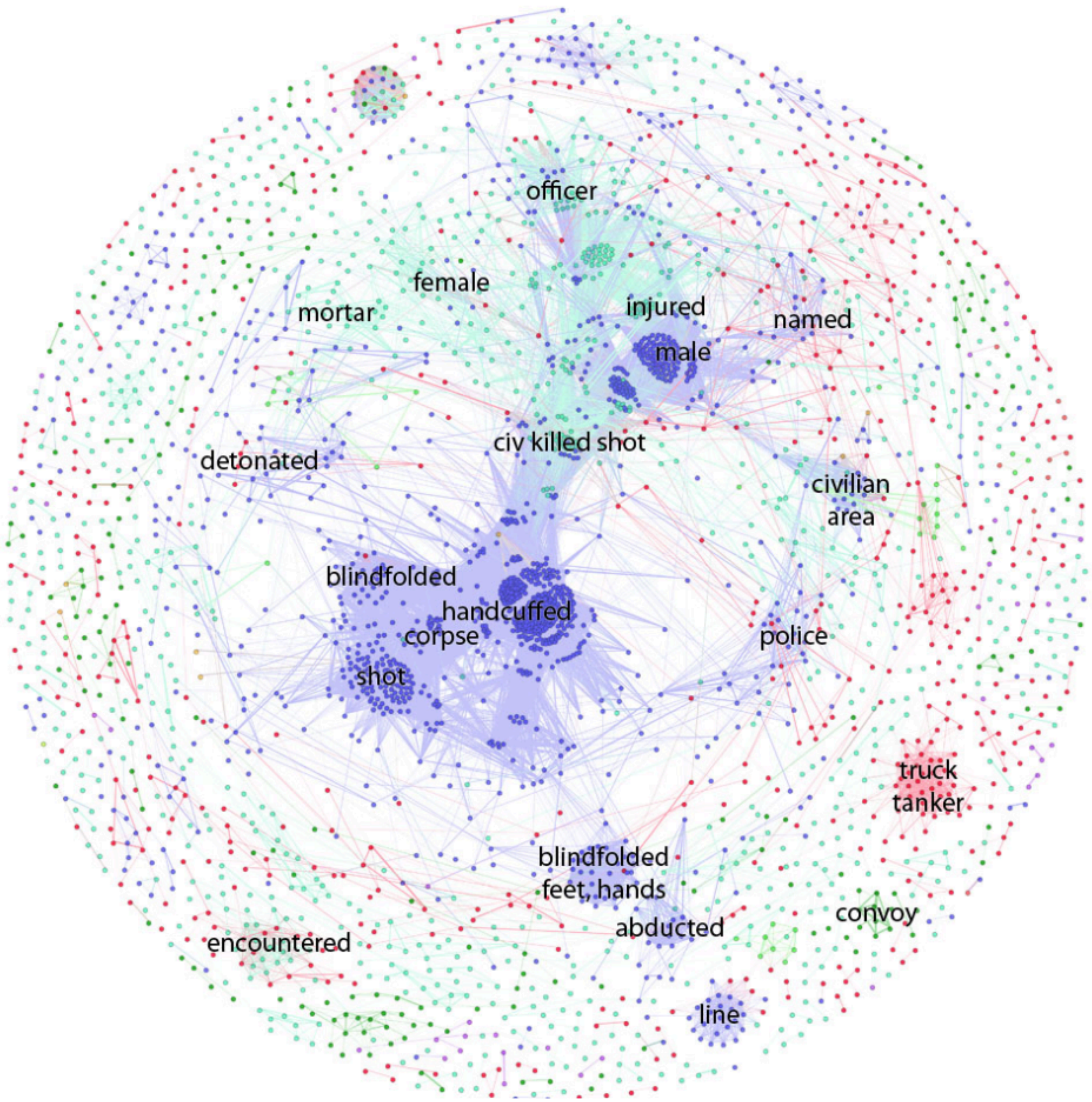
How could we visualize such a matrix?

Overview Project: WikiLeaks Iraq War Logs



http://jonathanstray.com/a-full-text-visualization-of-the-iraq-war-logs

Overview Project: WikiLeaks Iraq War Logs

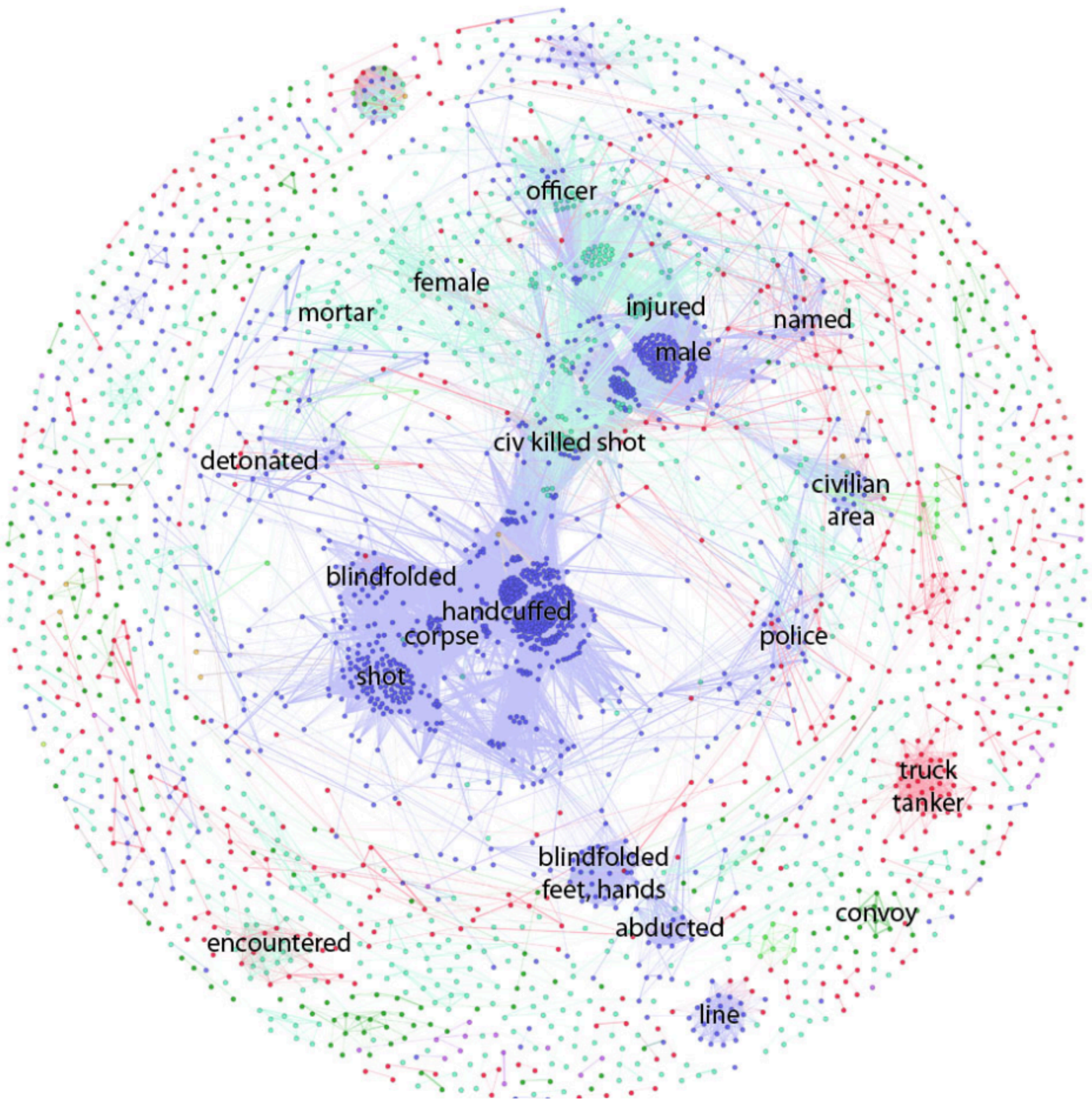


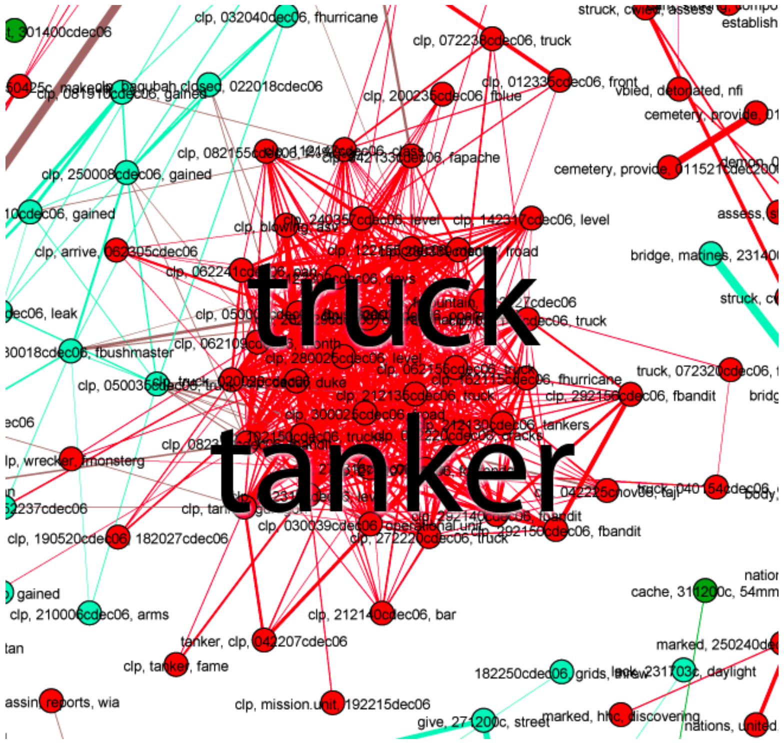
~ 400,000 war reports

visualized as keyword clusters according to TF-IDF (cosine similarity)

http://jonathanstray.com/a-full-text-visualization-of-the-iraq-war-logs

Overview Project: WikiLeaks Iraq War Logs





Red signifies that the military coded these reports as "explosive hazard"

Each point is a incident

http://jonathanstray.com/a-full-text-visualization-of-the-iraq-war-logs

Can you think of any Limitations of the TF-IDF Modeling Approach?

Limitations of TF-IDF

- **no syntactic or semantic relationships** of words or passages
- **words** are treated independently and are **not comparable**
- **topics** (word concepts) are **not reflected** very well
- stop word removal needed (or other **preprocessing** steps) in order **to achieve meaningful results**

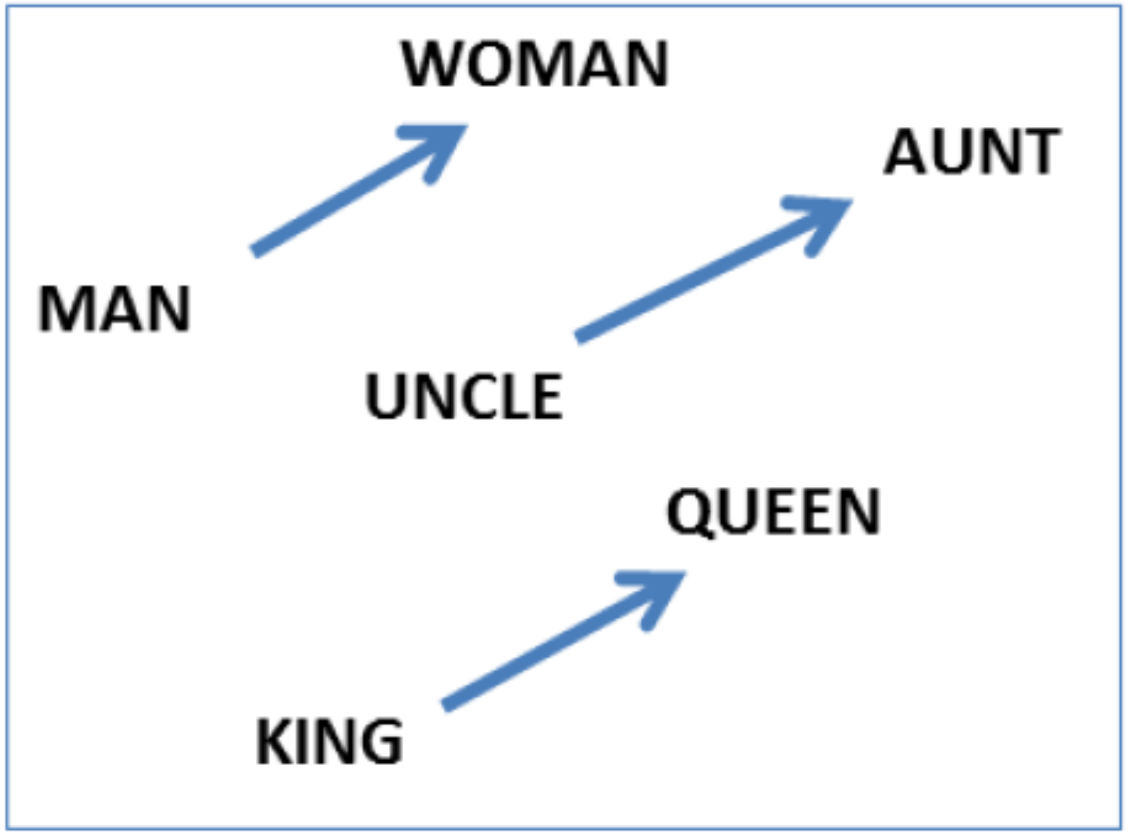
Distributed Word Embeddings

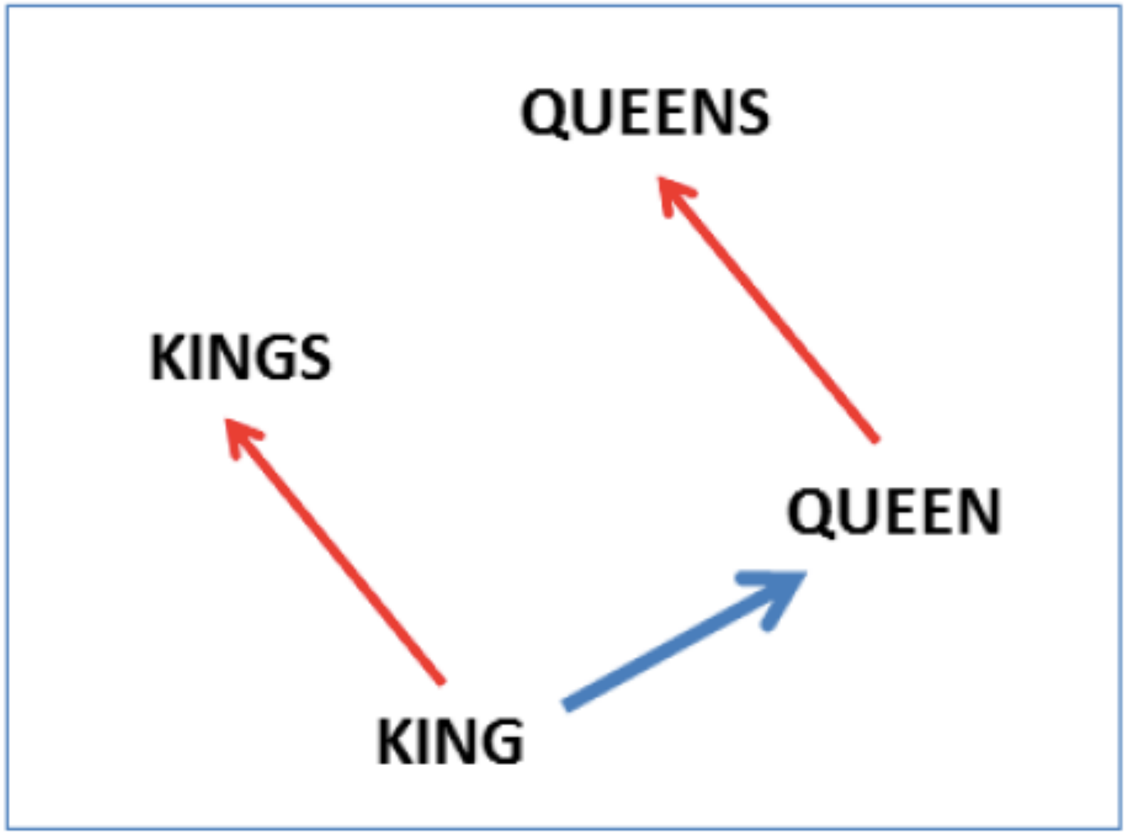
some versions:
word2vec/glove/fasttext/bert/elmo/GPT-2

Words as distributions according to their context

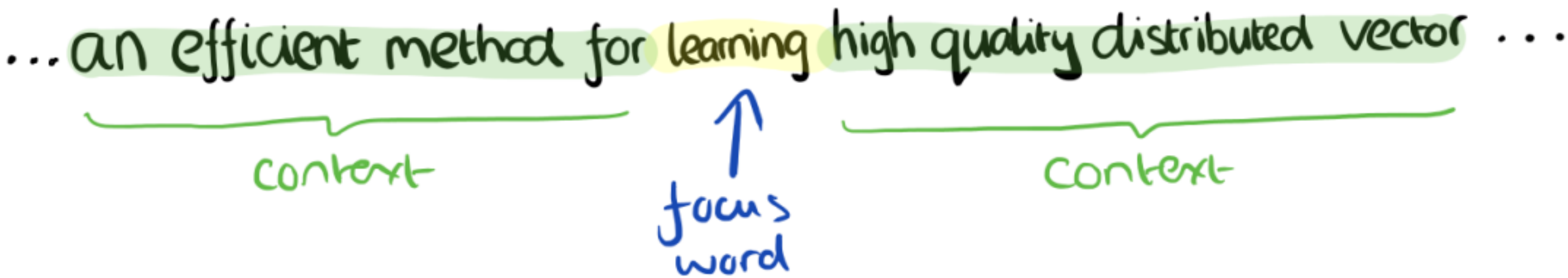
- a word is a representation of all its context windows within a corpus (**context = N-gram** of which a word is part)
- similar **words** tend to have similar representations —> **comparable**
- **syntax** as well as **semantics** is reflected
- no document vectors are created, only word vectors

Efficient estimation of word representations in vector space
Tomas Mikolov et al., ICLR Workshop, 2013

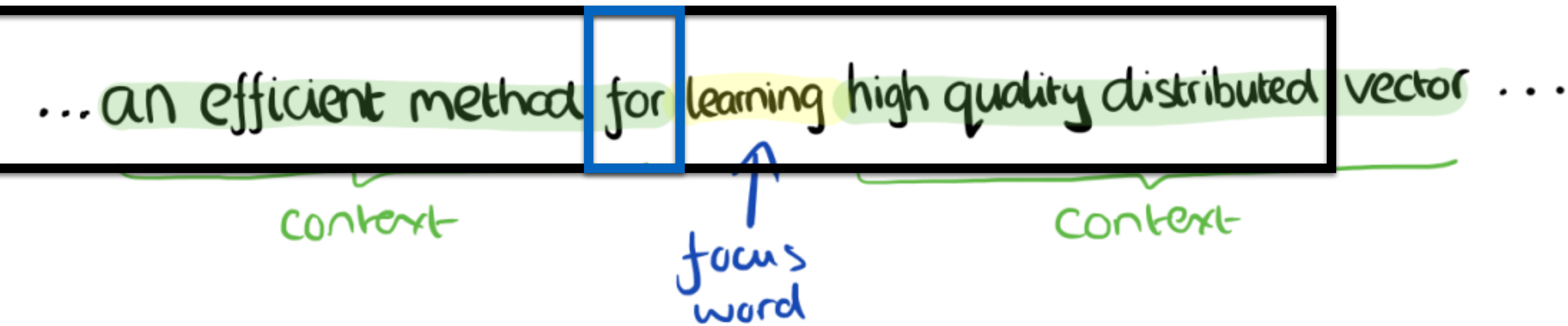




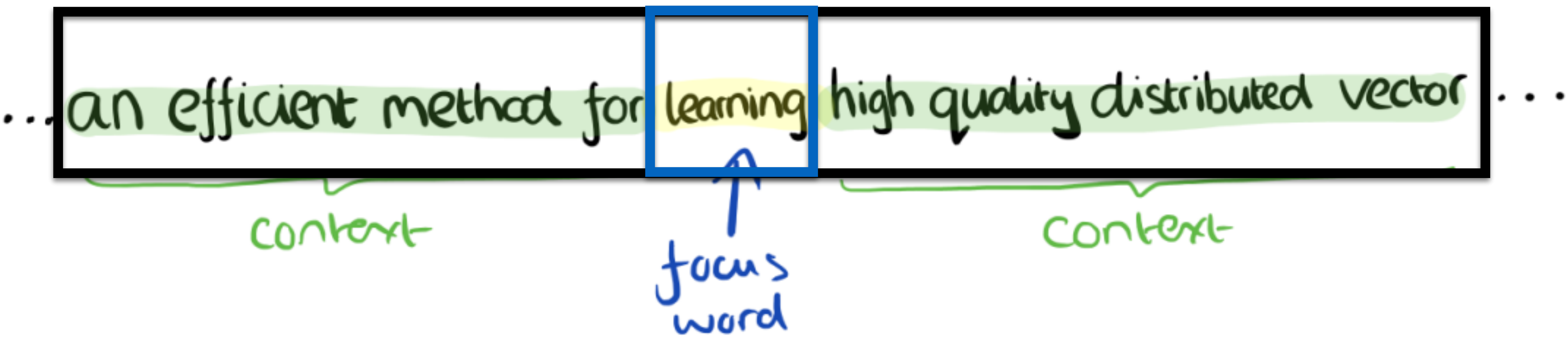
word2vec



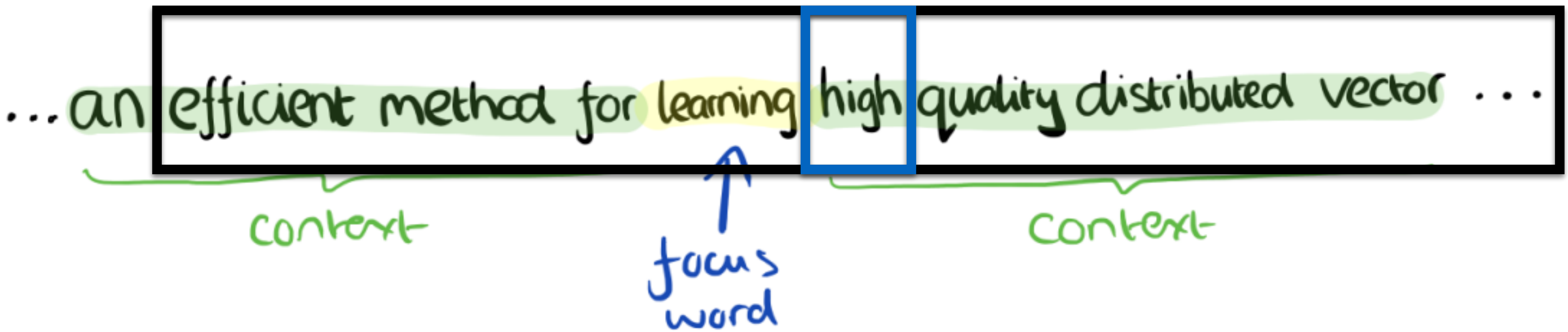
https://blog.acolyer.org/2016/04/21/the-amazing-power-of-word-vectors/



https://blog.acolyer.org/2016/04/21/the-amazing-power-of-word-vectors/

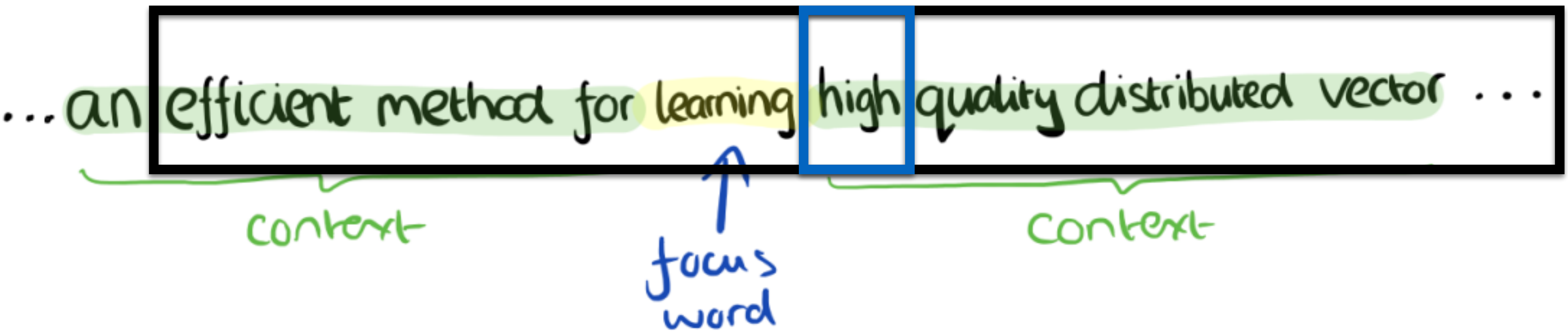


https://blog.acolyer.org/2016/04/21/the-amazing-power-of-word-vectors/



https://blog.acolyer.org/2016/04/21/the-amazing-power-of-word-vectors/

We try to predict the current focus word with a Neural Network



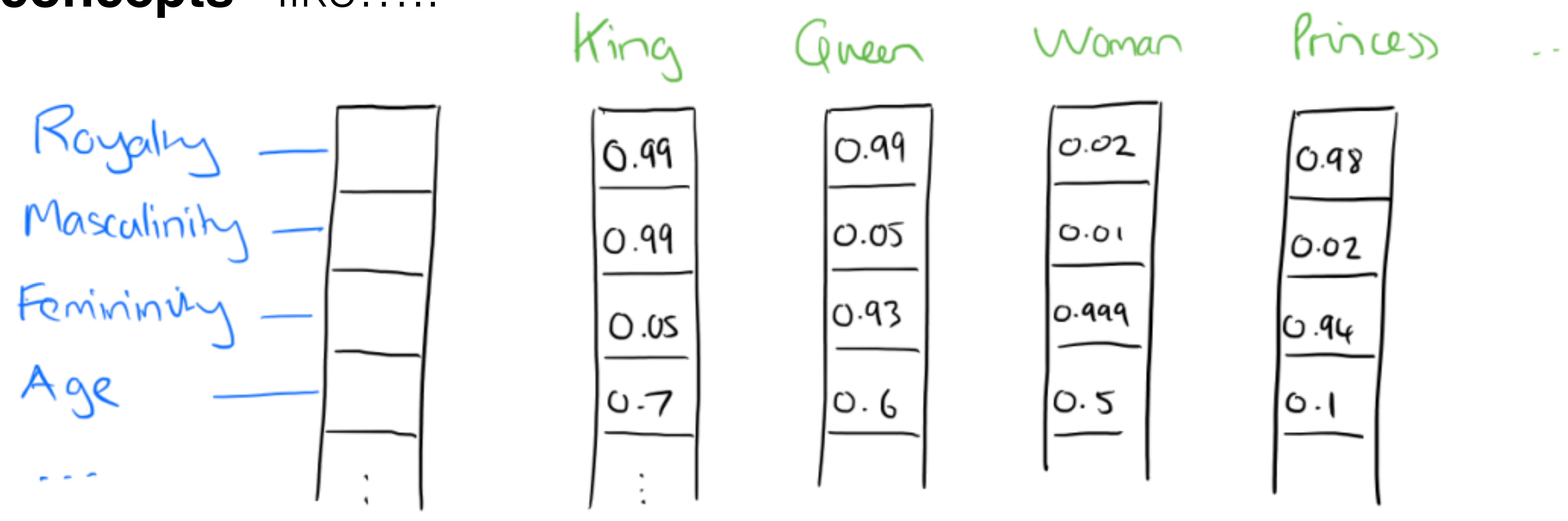
https://blog.acolyer.org/2016/04/21/the-amazing-power-of-word-vectors/

First: Our **output reflects** the relation between all words within the vocabulary (**Co-Occurrences**).

After a dimensionality reduction (from maybe 10,000 words in vocabulary to 100 dimensions) we receive something related to **“concepts”** like…..

First: Our **output reflects** the relation between all words within the vocabulary (**Co-Occurrences**).

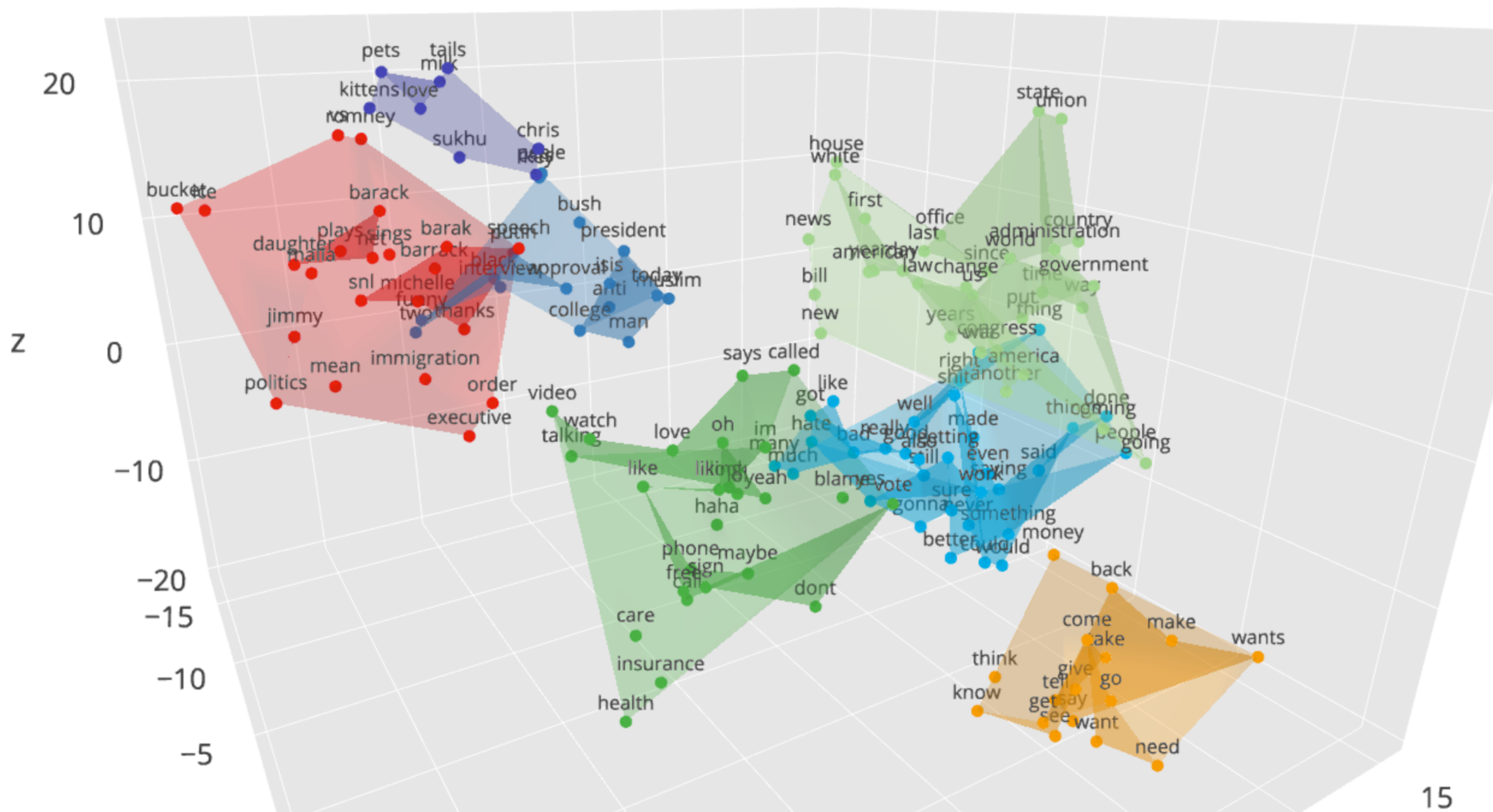
After a dimensionality reduction (from maybe 10,000 words in vocabulary to 100 dimensions) we receive something related to **“concepts”** like…..



https://blog.acolyer.org/2016/04/21/the-amazing-power-of-word-vectors/

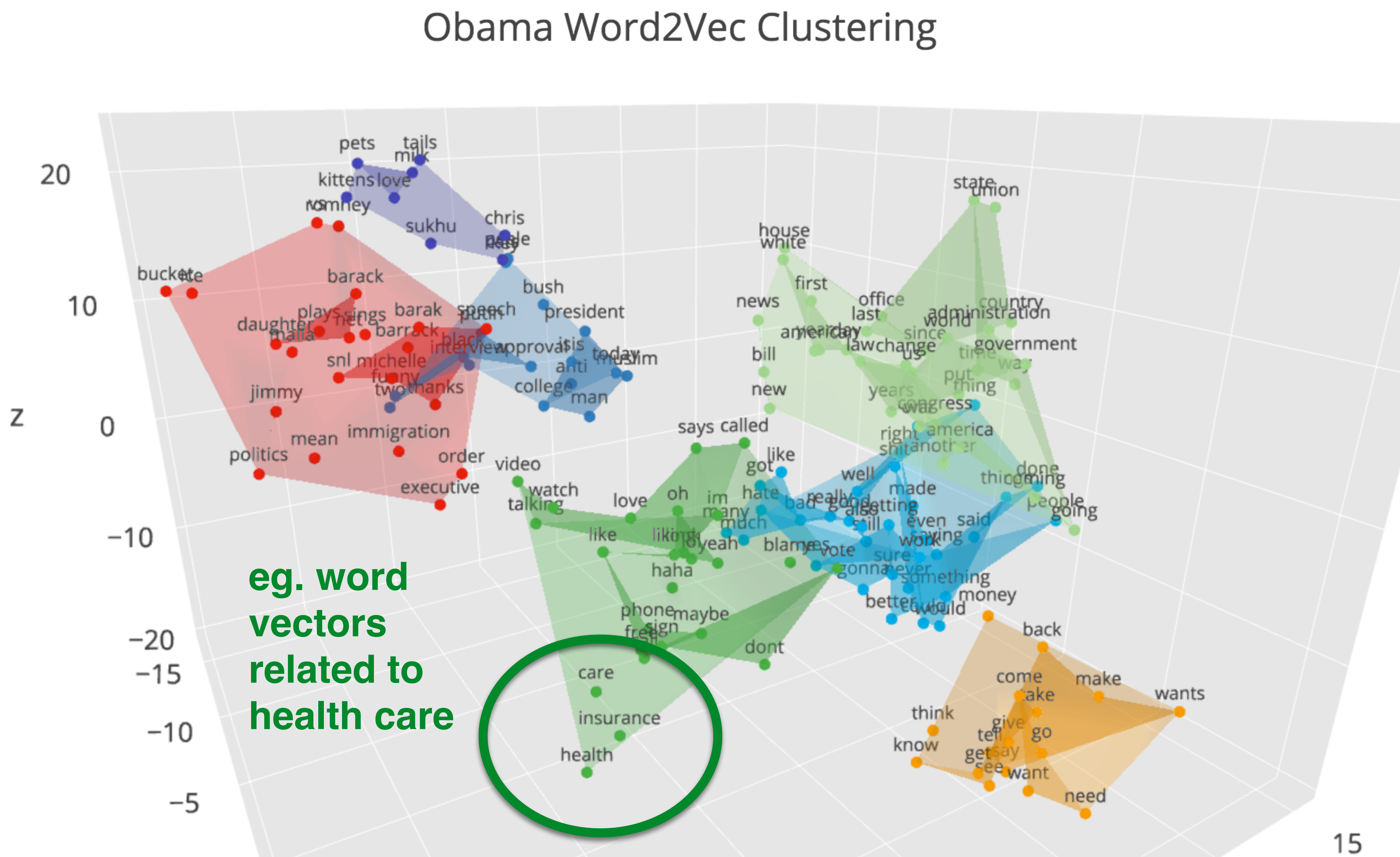
How could we visualize such vector spaces?

Obama Word2Vec Clustering



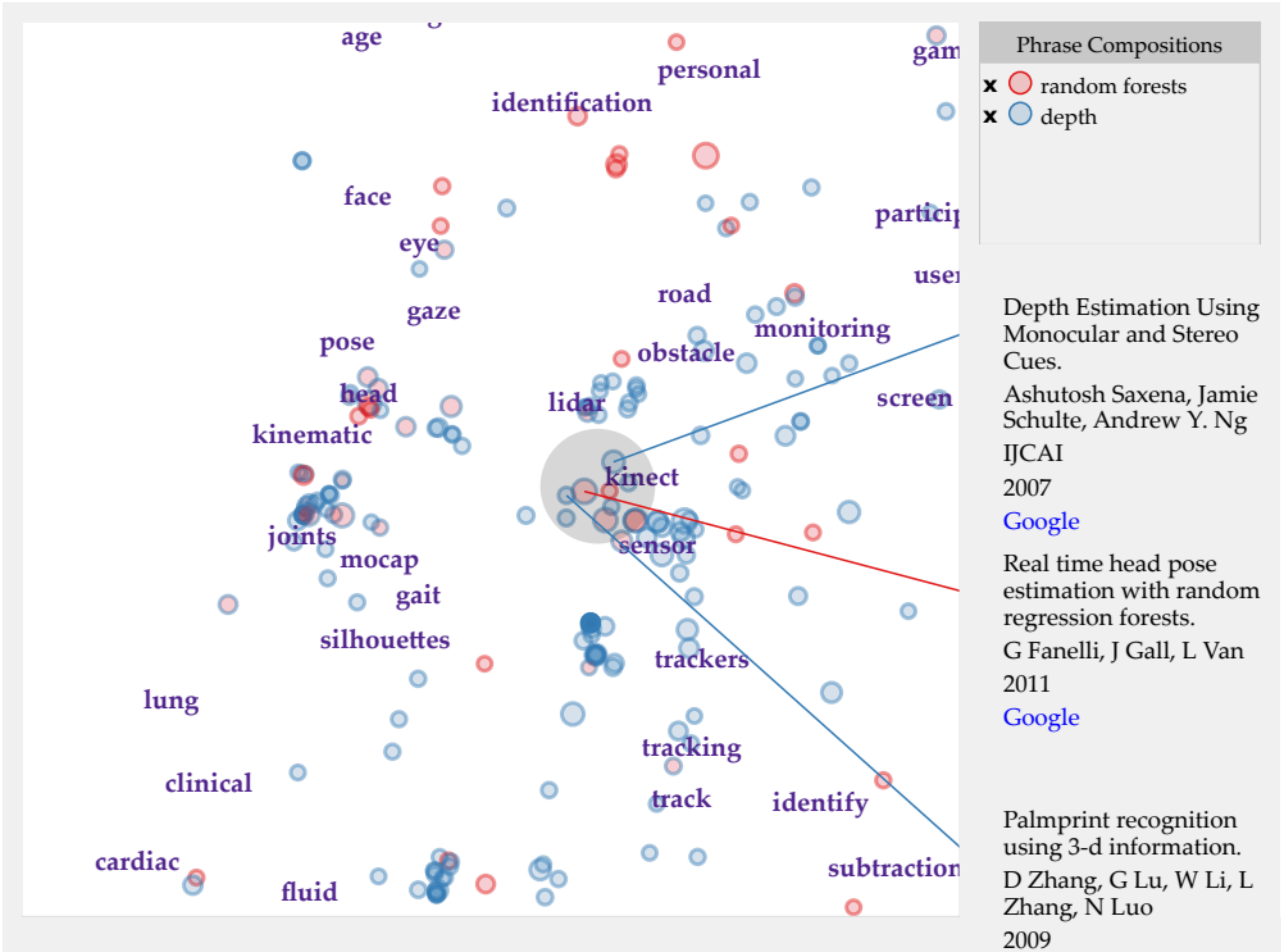
https://plot.ly/~KevinAccount/18.embed

word2vec

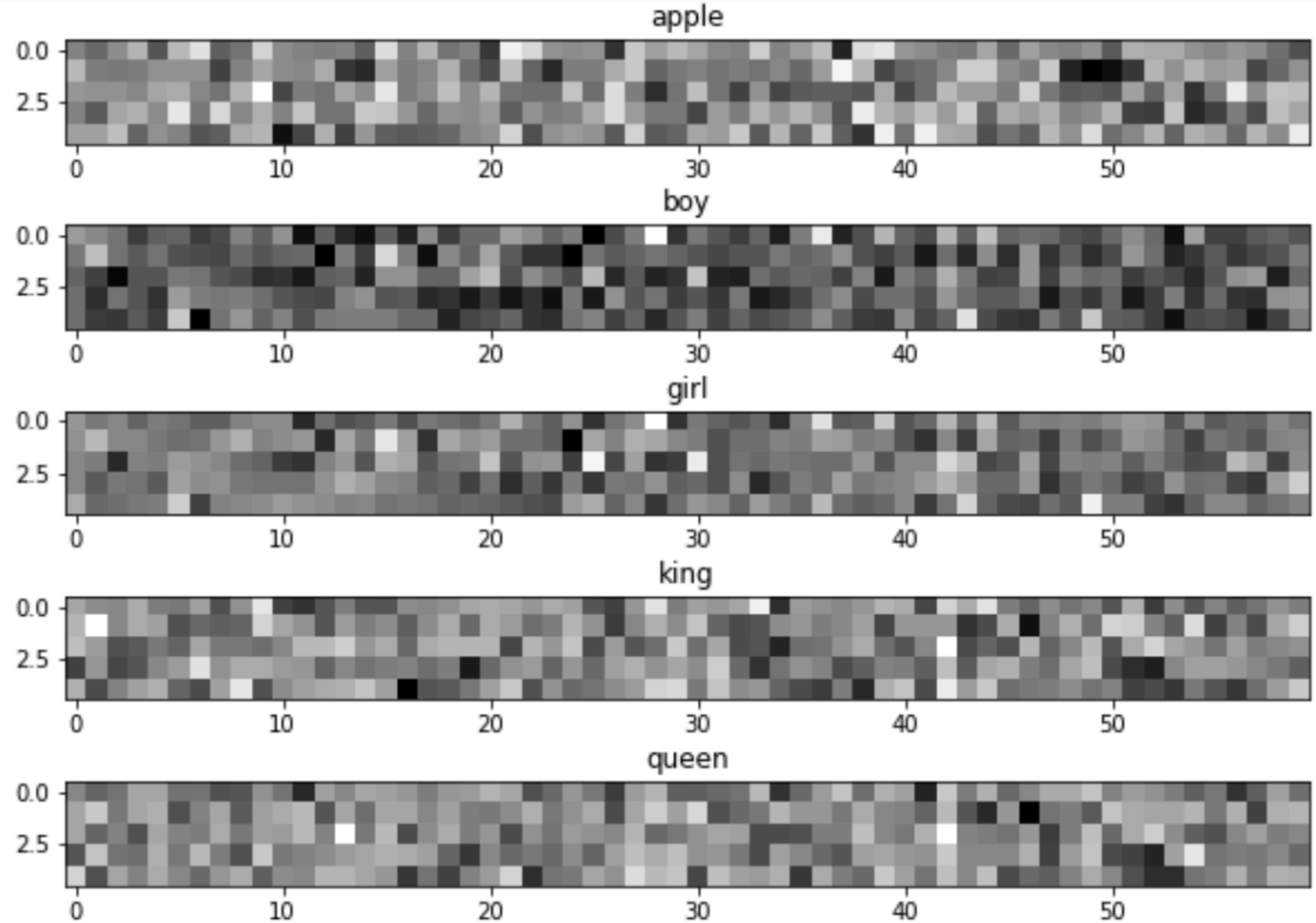


eg. word vectors related to health care

https://plot.ly/~KevinAccount/18.embed



Berger, Matthew, Katherine McDonough, and Lee M. Seversky. "cite2vec: Citation-driven document exploration via word embeddings." *IEEE transactions on visualization and computer graphics* 23.1 (2016): 691-700.



https://towardsdatascience.com/visualisation-of-embedding-relations-word2vec-bert-64d695b7f36

What are possible short comings of word2vec?

Limitations of Word 2 Vec

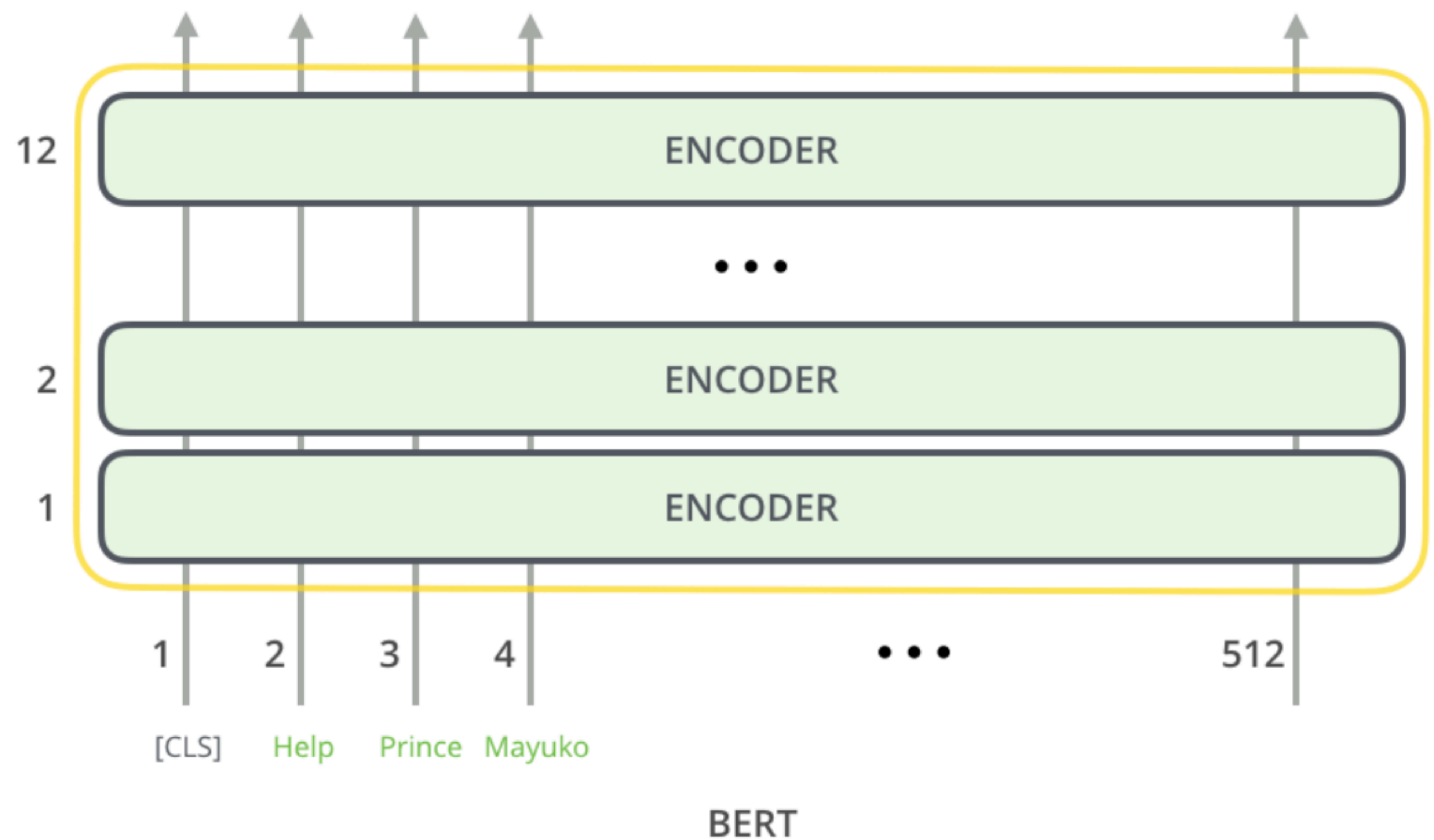
- different contexts get averaged

```
bank vault
bank robber
river bank
```

- how to represent documents? (doc2vec, …)
- cannot handle unknown words

BERT (Bidirectional Encoder Representations from Transformers)

ELMo (Embeddings from Language Models)



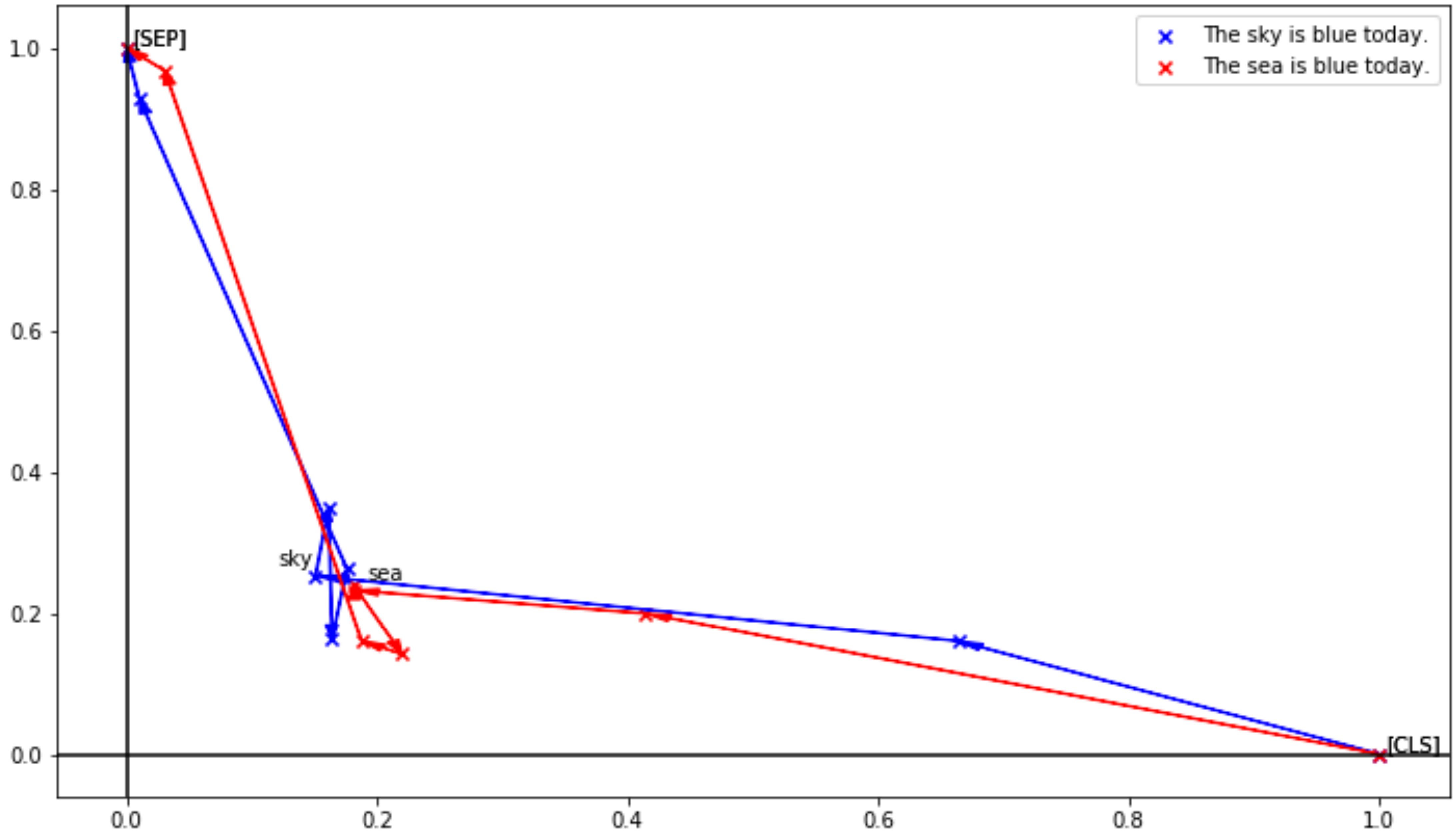
https://jalammar.github.io/illustrated-bert/

```
First 5 values for each meaning of "bank".

bank vault     tensor([ 2.1319, -2.1413, -1.6260,  0.8638,  3.3173])
bank robber    tensor([ 1.1868, -1.5298, -1.3770,  1.0648,  3.1446])
river bank     tensor([ 1.1295, -1.4725, -0.7296, -0.0901,  2.4970])
```

https://mccormickml.com/2019/05/14/BERT-word-embeddings-tutorial/

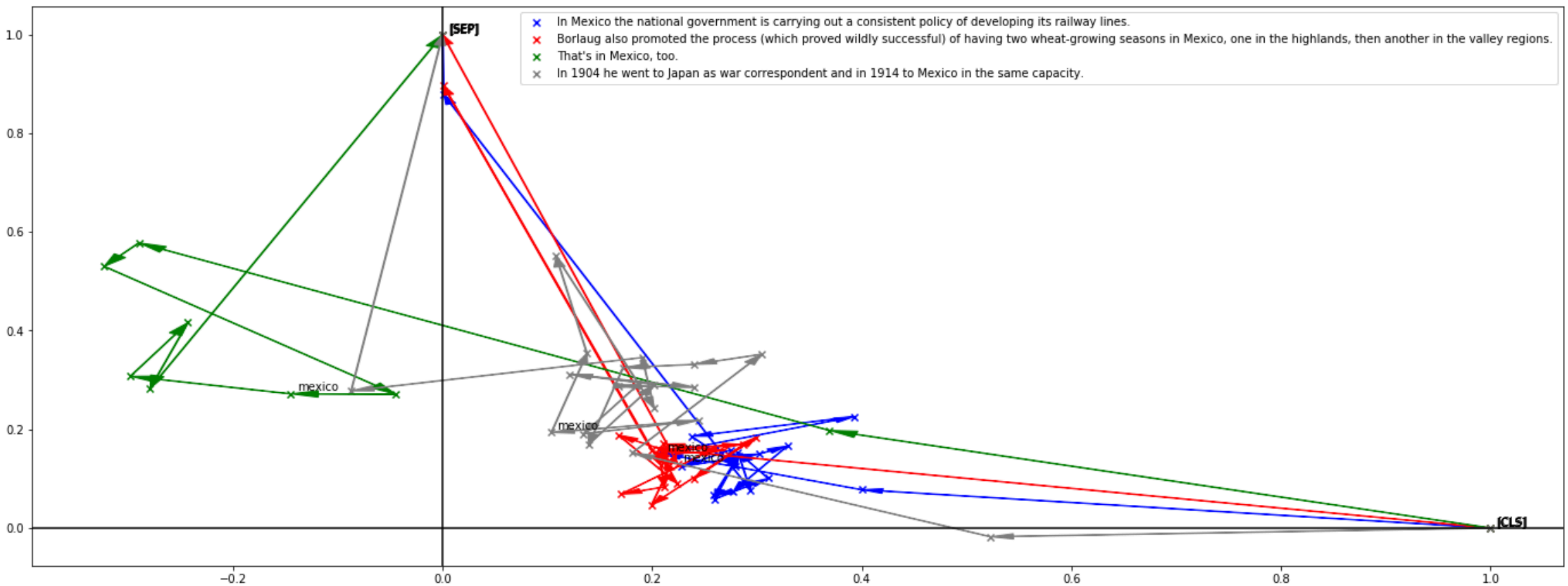
```
[CLS] This is a sentence example. [SEP]
```



https://towardsdatascience.com/visualisation-of-embedding-relations-word2vec-bert-64d695b7f36

bert

```
[CLS] This is a sentence example. [SEP]
```

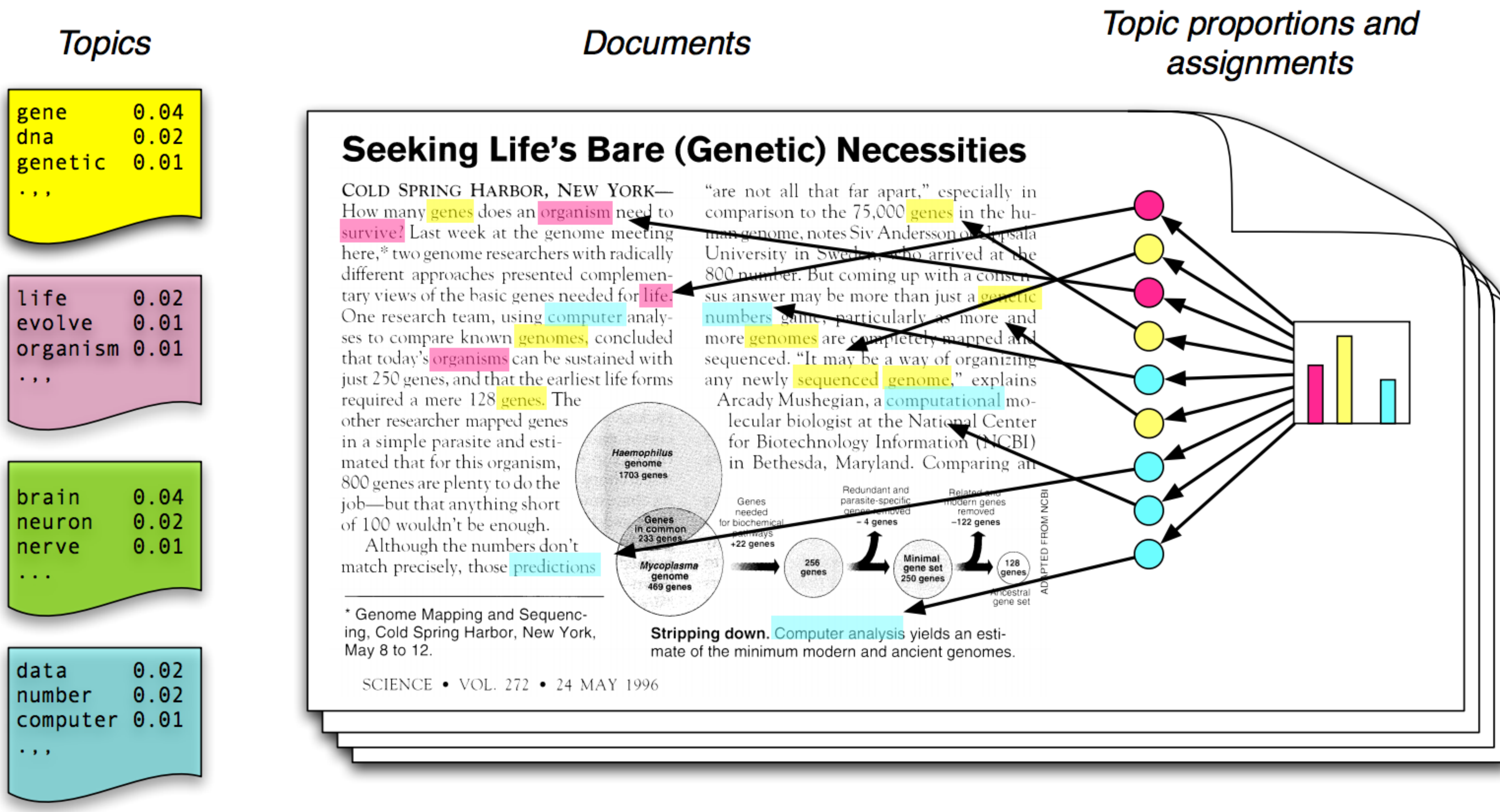


https://towardsdatascience.com/visualisation-of-embedding-relations-word2vec-bert-64d695b7f36

Topic Modeling

Probabilistic topic models

David Blei, Communications of the ACM 2012



Input

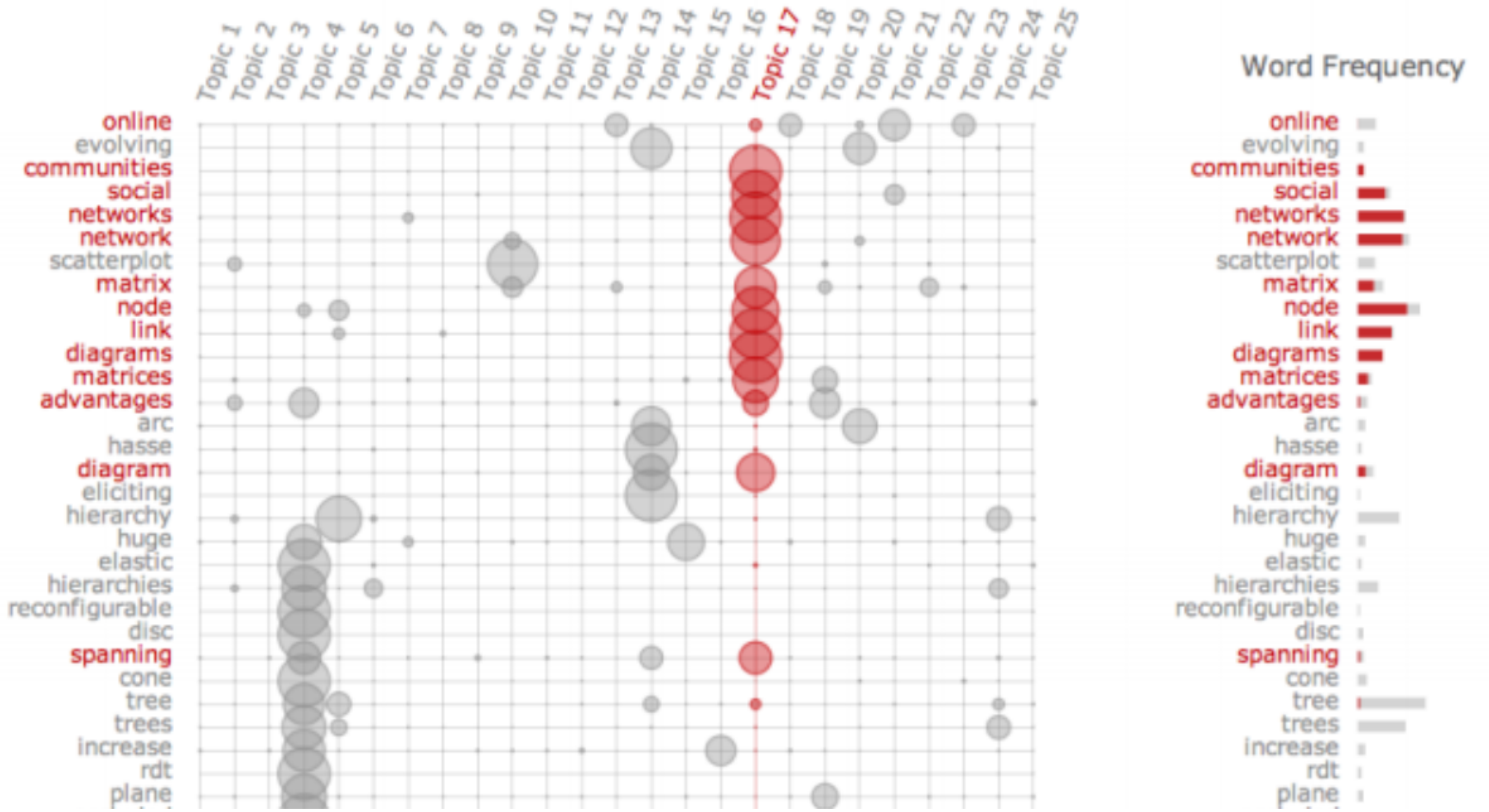
- Corpus = Text Sources (amount=M)
- Number of Topics (N) you want the model to cover

Output: two different kinds of distributions

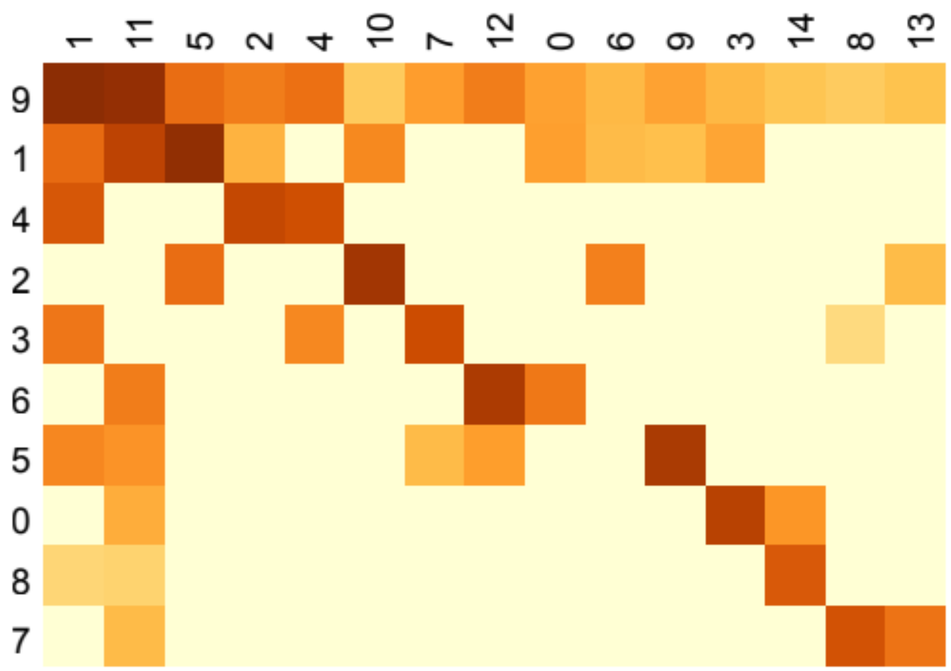
- N Topic Distributions with probabilities for each word of the vocabulary
- M Text Distributions with probabilities for each of the N topics

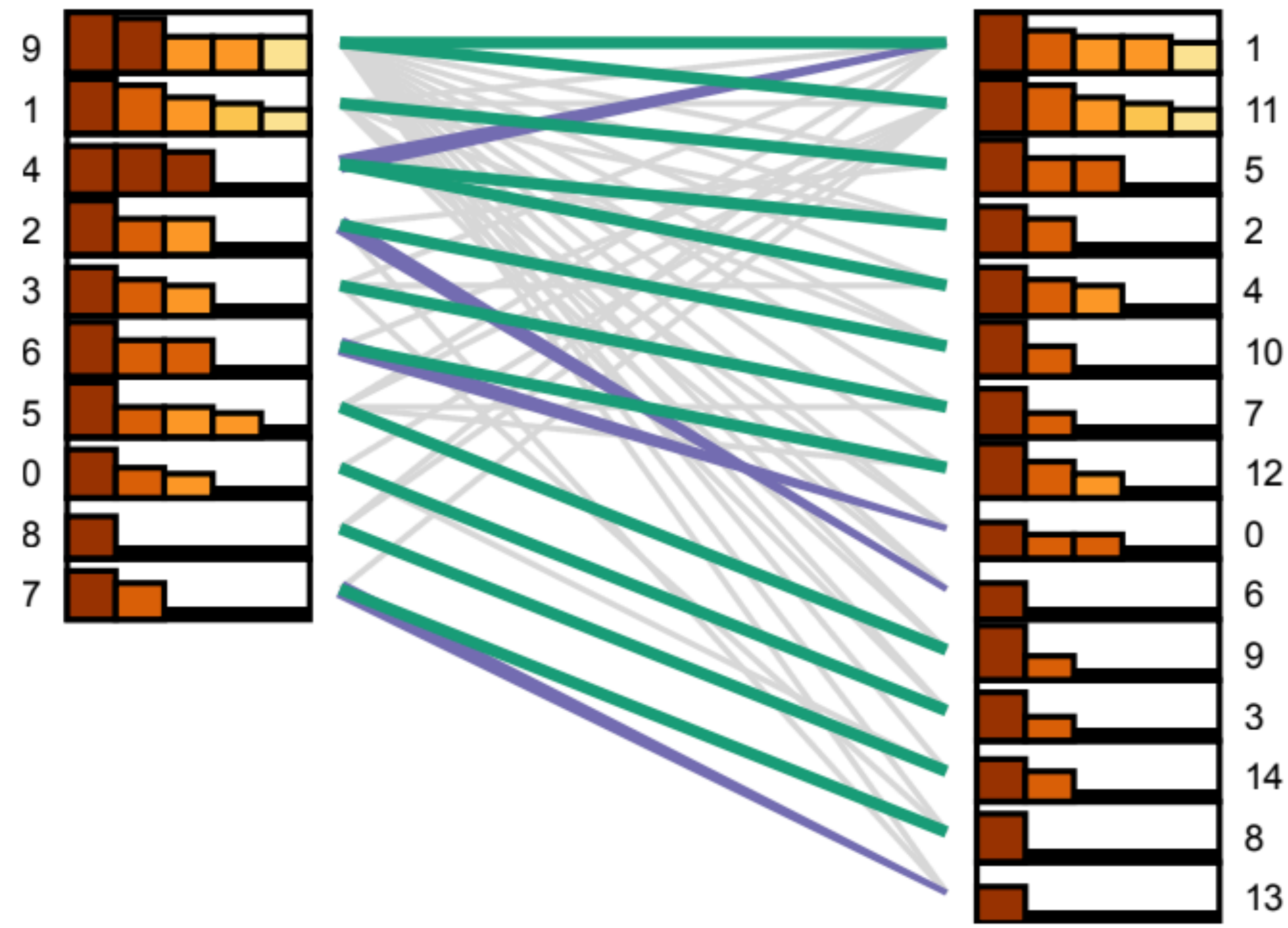
Termite: Visualization techniques for assessing textual topic models

Jason Chuang, Christopher D. Manning, Jeffrey Herr; Proceedings of the International Working Conference on Advanced Visual Interfaces 2012

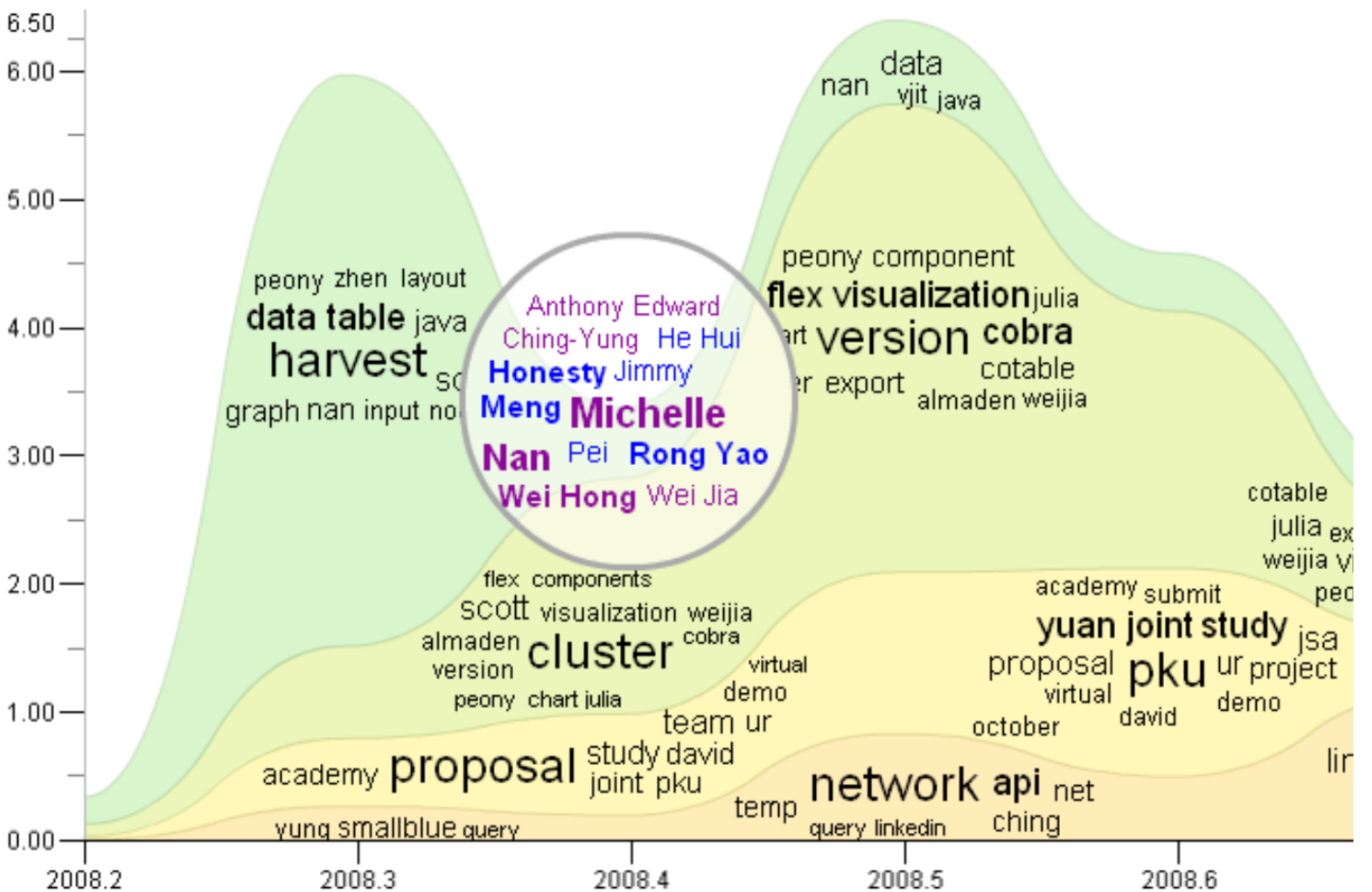


Serendip: Topic model-driven visual exploration of text corpora.
Alexander, E., Kohlmann, J., Valenza, R., Witmore, M., & Gleicher, M. (2014, October). In Visual Analytics Science and Technology (VAST), 2014



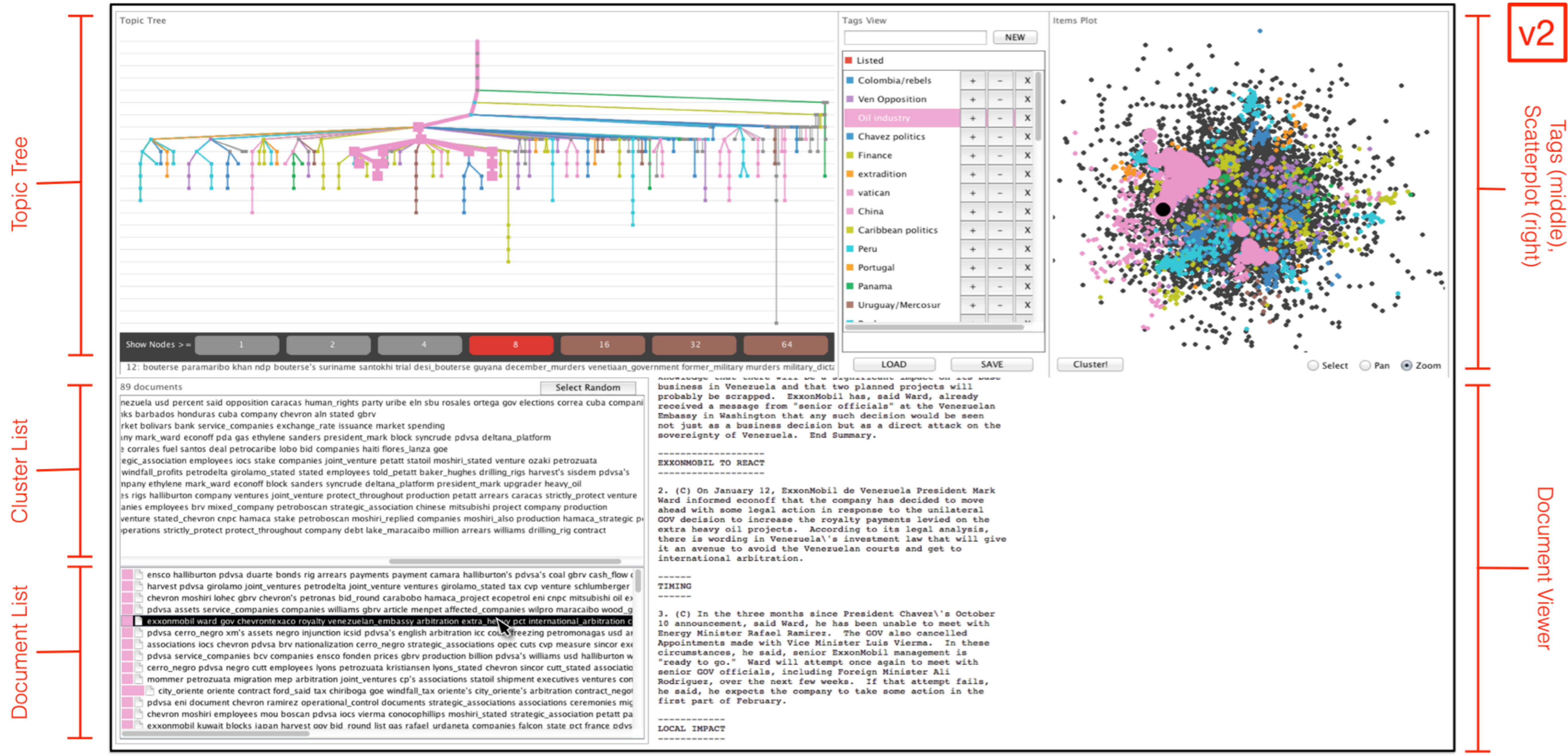


TIARA: A Visual Exploratory Text Analytic System, Wei et. al, KDD 2010)



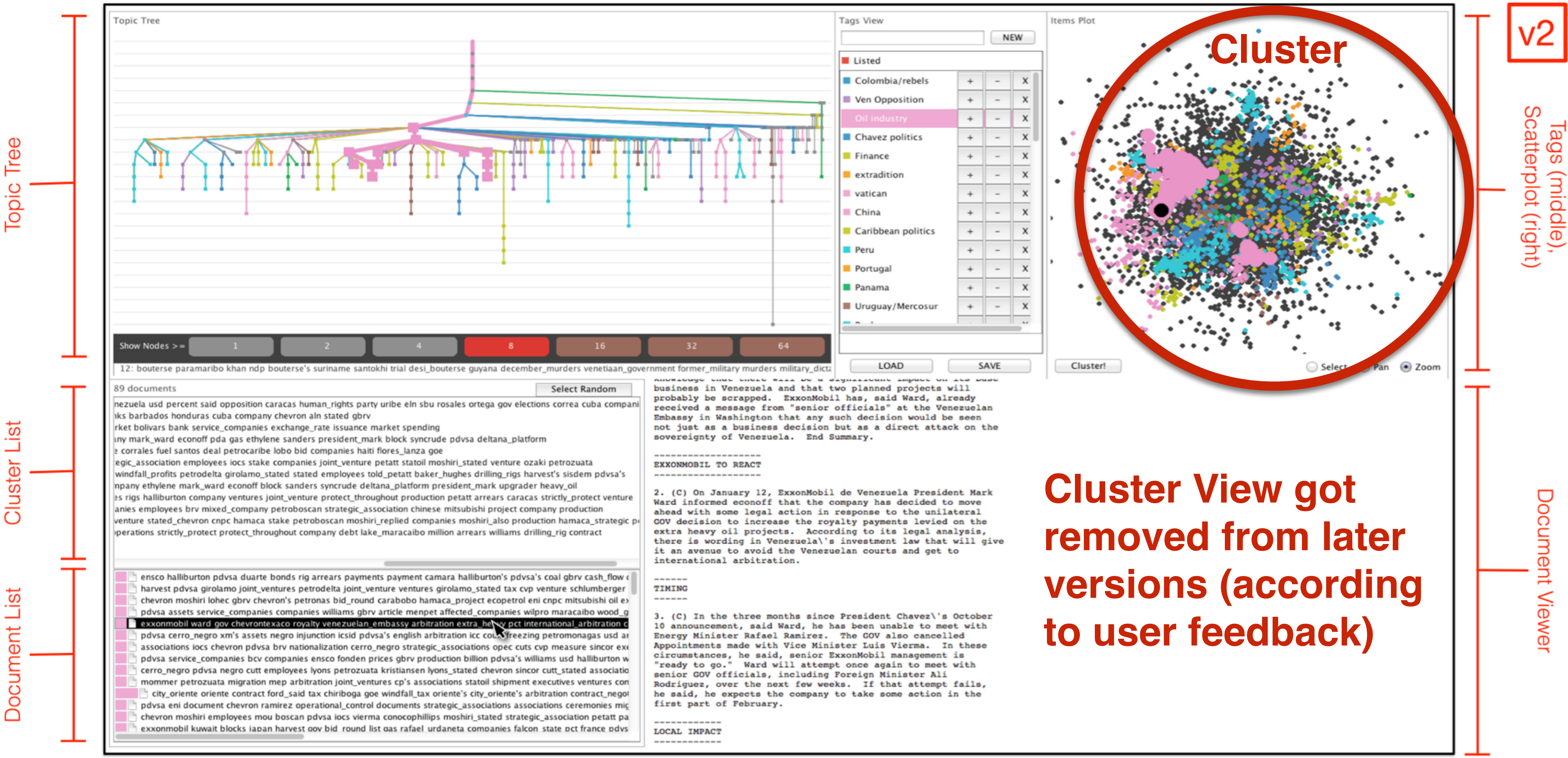
Overview: The Design, Adoption, and Analysis of a Visual Document Mining Tool For Investigative Journalists (Brehmer et al., IEEE InfoVis 2014)

old version of the software:



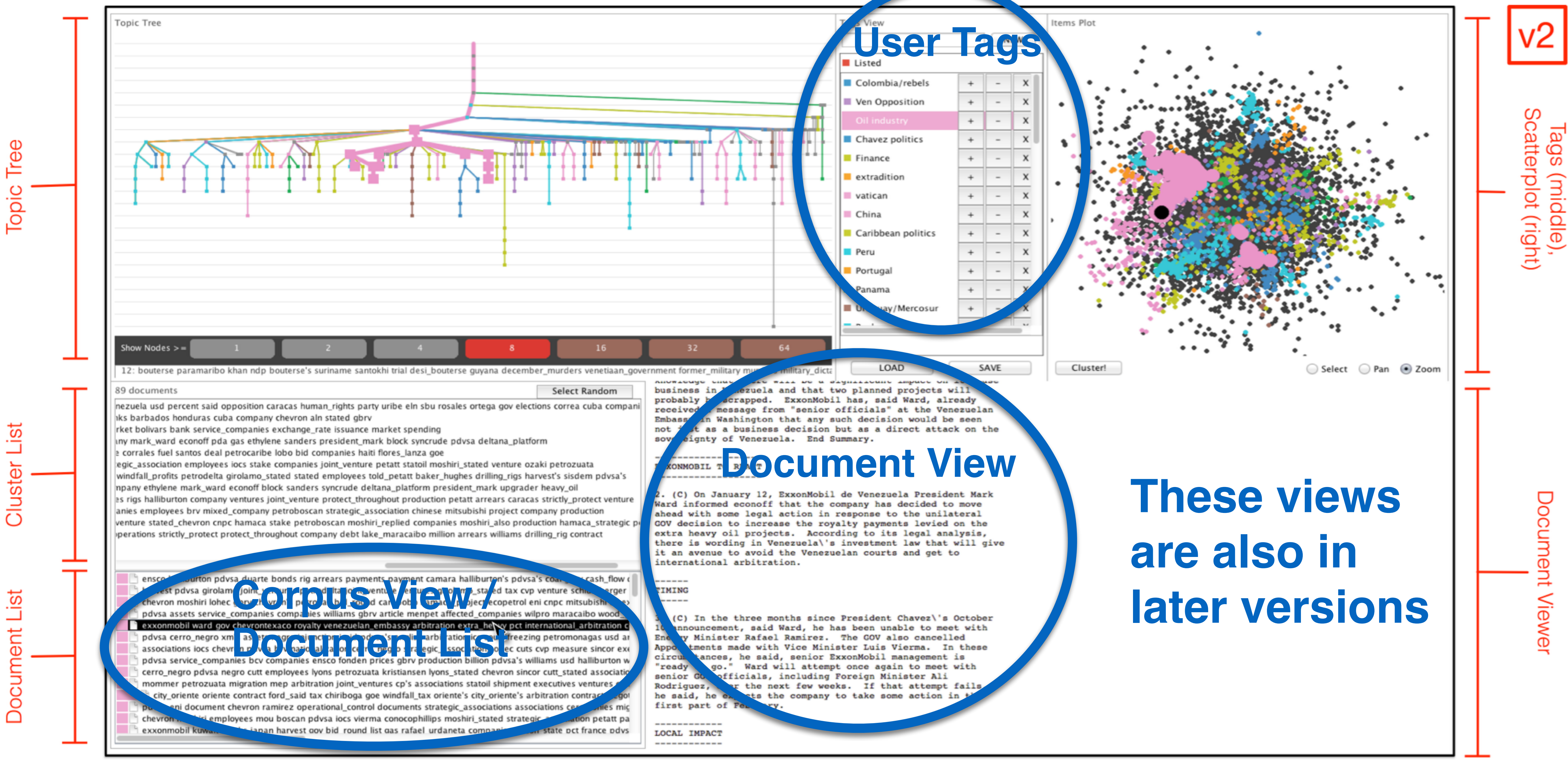
Overview: The Design, Adoption, and Analysis of a Visual Document Mining Tool For Investigative Journalists (Brehmer et al., IEEE InfoVis 2014)

old version of the software:



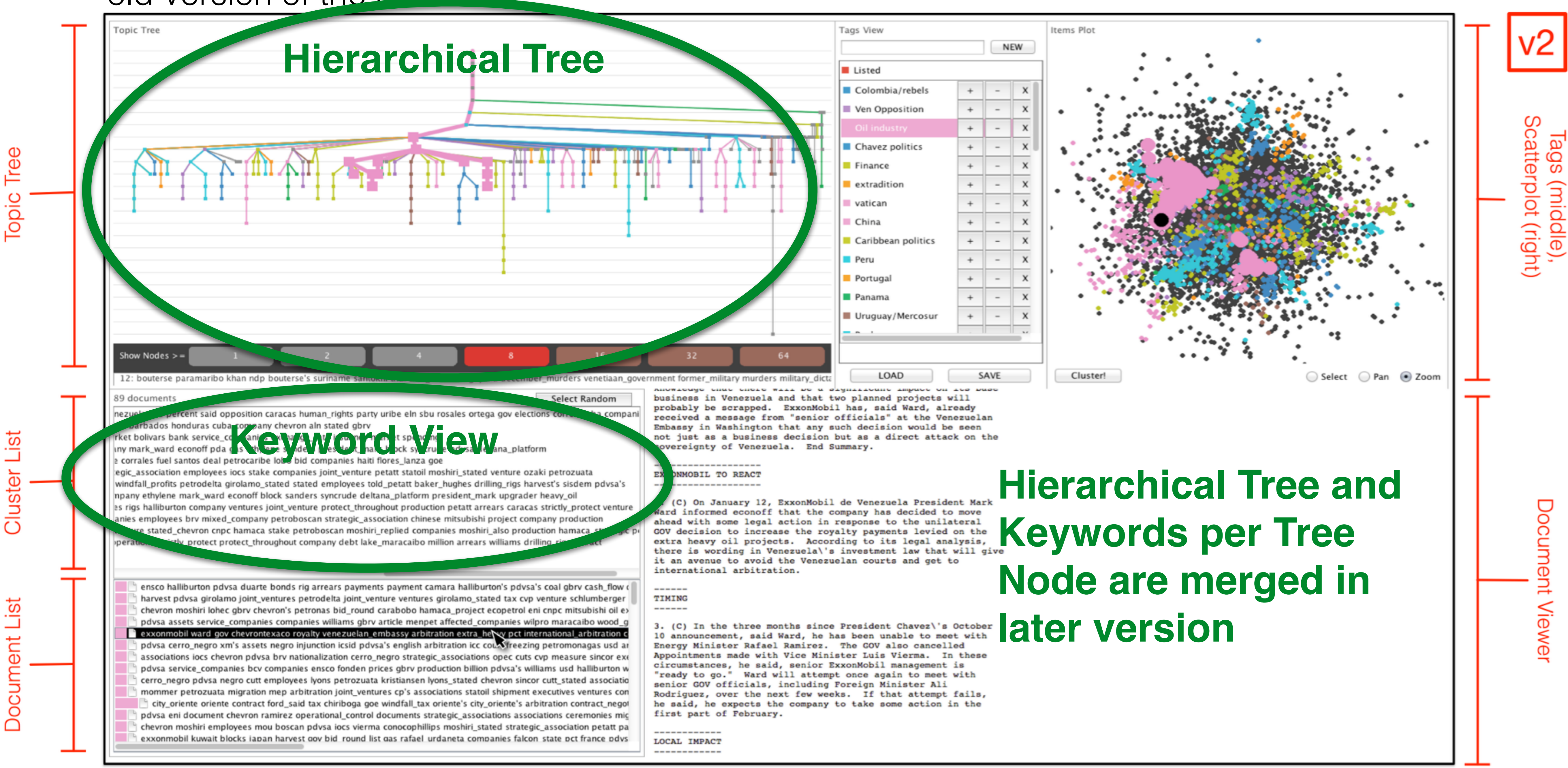
Overview: The Design, Adoption, and Analysis of a Visual Document Mining Tool For Investigative Journalists (Brehmer et al., IEEE InfoVis 2014)

old version of the software:



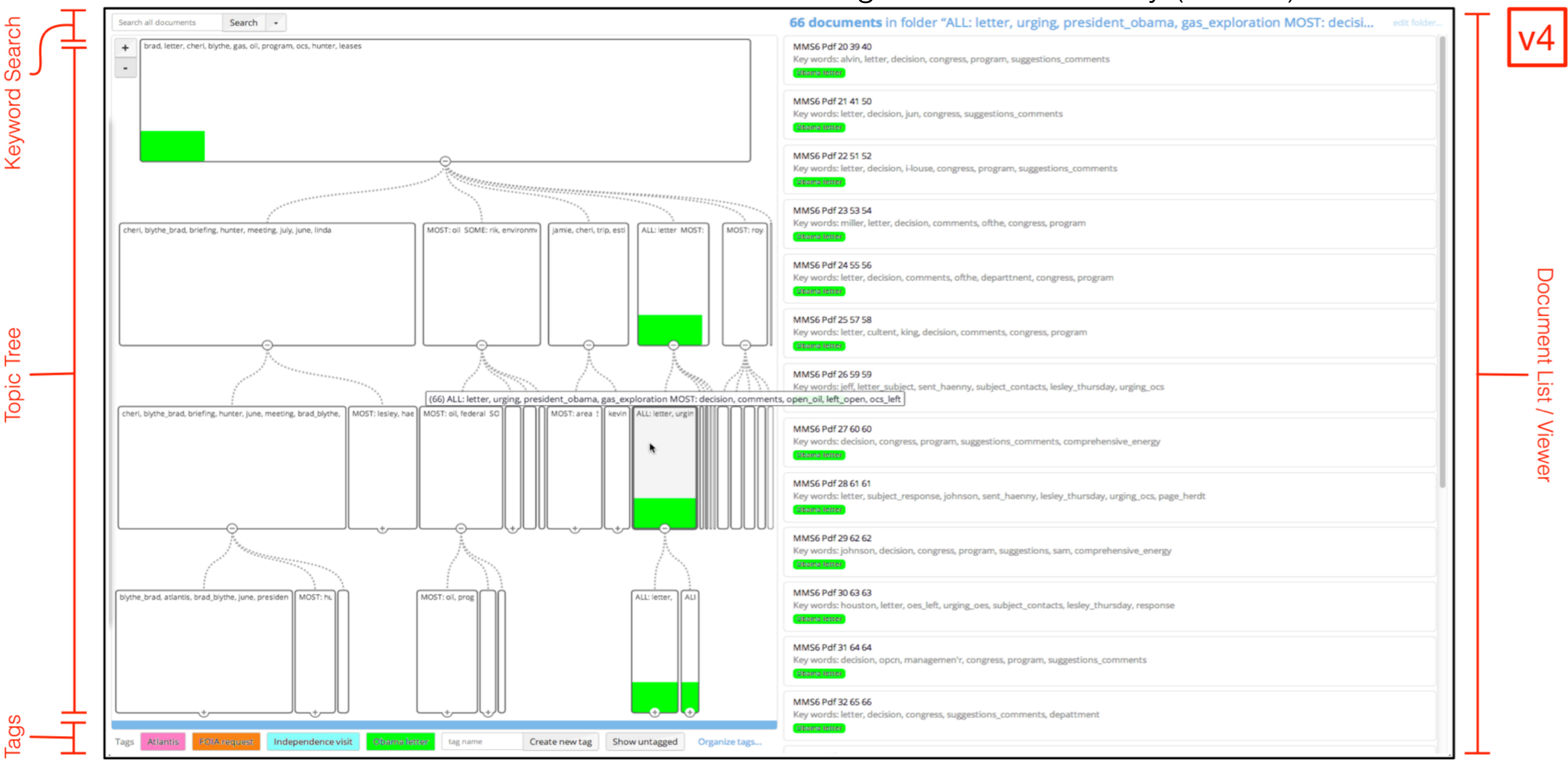
Overview: The Design, Adoption, and Analysis of a Visual Document Mining Tool For Investigative Journalists (Brehmer et al., IEEE InfoVis 2014)

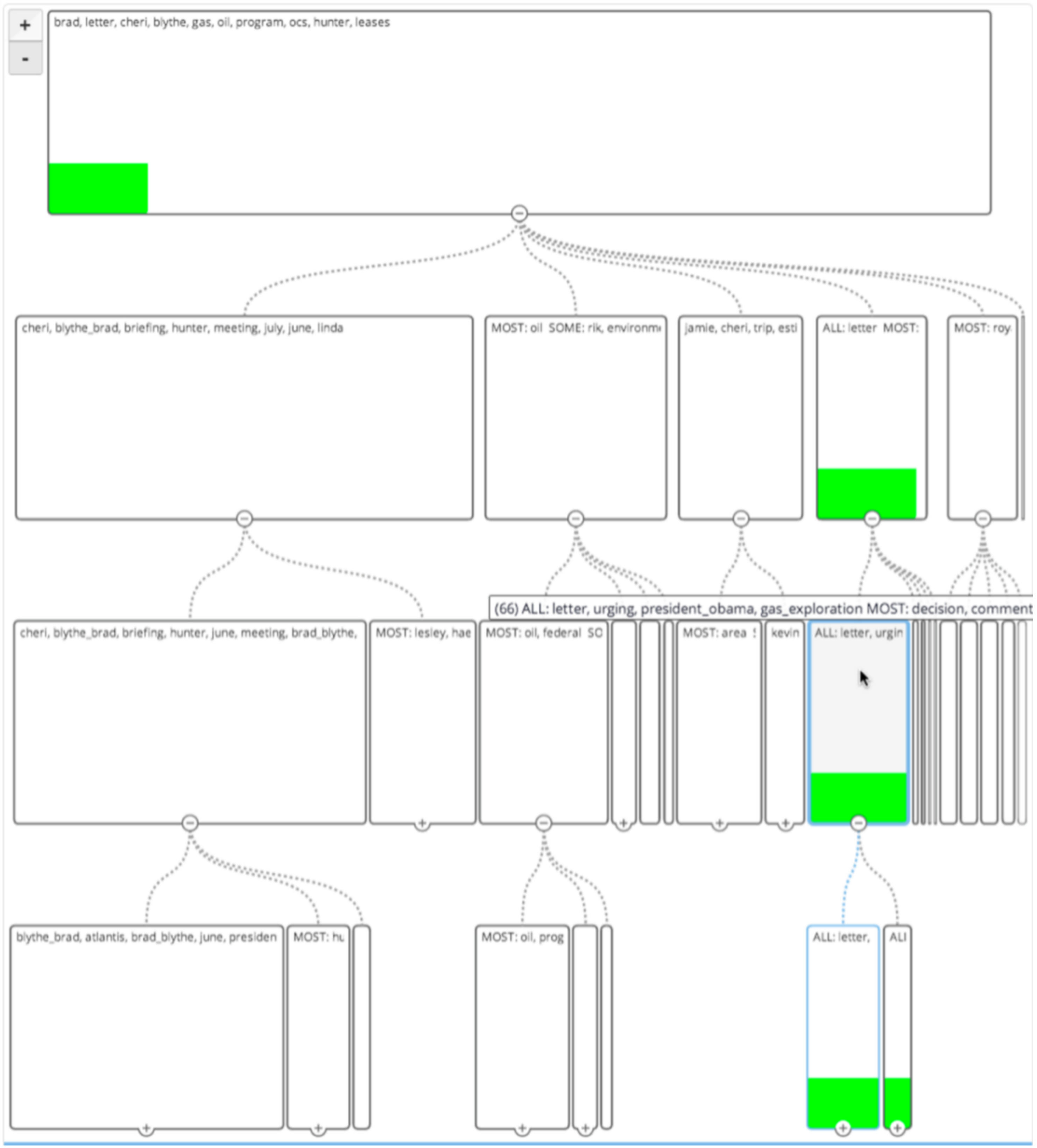
old version of the software:



Overview: The Design, Adoption, and Analysis of a Visual Document Mining Tool For Investigative Journalists (Brehmer et al., IEEE InfoVis 2014)

new version focuses on hierarchical tree according to cosine similarity (TF-IDF) of documents:





- **Size of nodes** related to number of documents
- **Keywords located** within the corresponding nodes
- **Tags color-coded** according to percentage of fill level within node
- **Tree expandable**

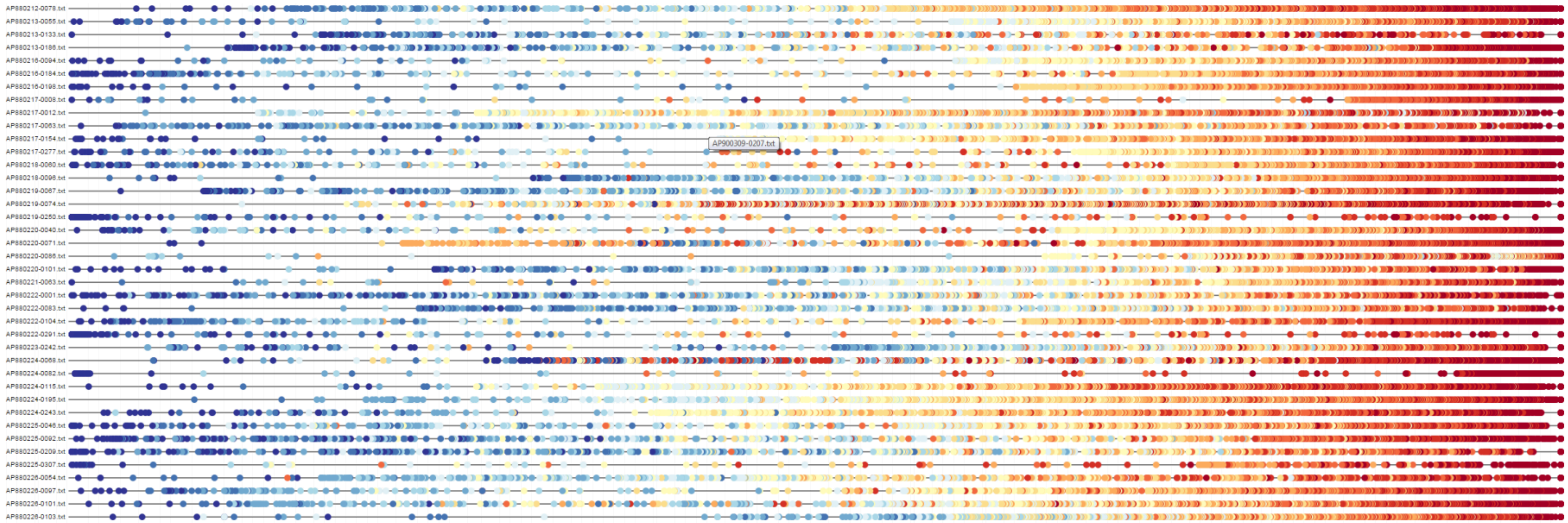


Fig. 1: Buddy plots show consistency of document relationships across topic models by encoding similarity with respect to individual documents. In this figure, each row represents a document, with the rest of the corpus encoded as circular glyphs along the row. Distance from the row's document in one model is encoded using horizontal position, while distance in a second model is encoded using color. This combination of encodings lets us see similarities from two models within one row of glyphs. Deviations in similarity between the two models can be identified as breaks from a smooth gradient. Though the two models seem to correlate well with documents at either extreme (blue documents to the left, red documents to the right), we see dramatic shifts between different classifications for documents in between, identified by breaks in the blue-to-red gradient structure.

Alexander, Eric, and Michael Gleicher. "Task-driven comparison of topic models." *IEEE transactions on visualization and computer graphics* 22.1 (2015): 320-329.

Topic Modeling

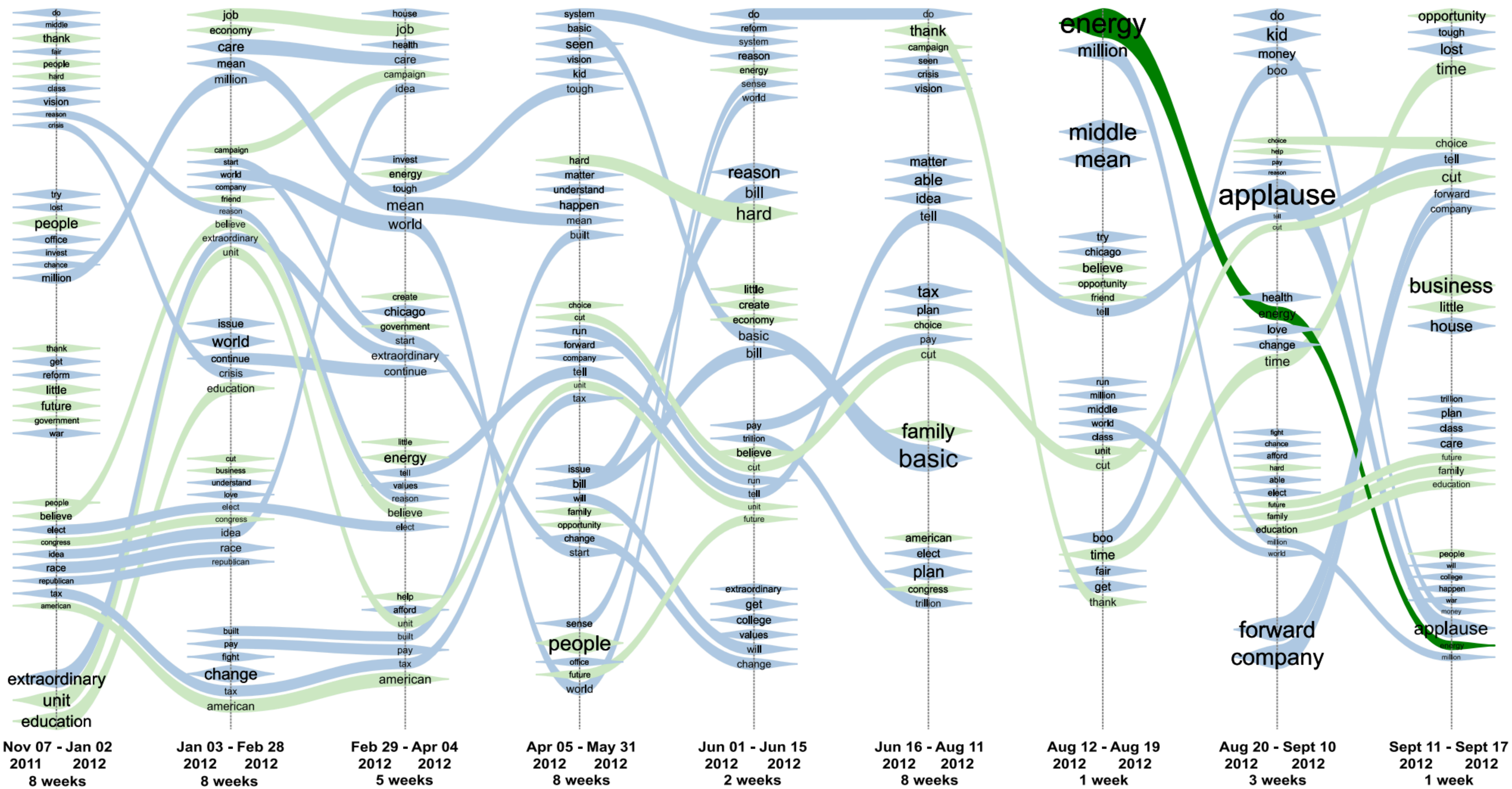
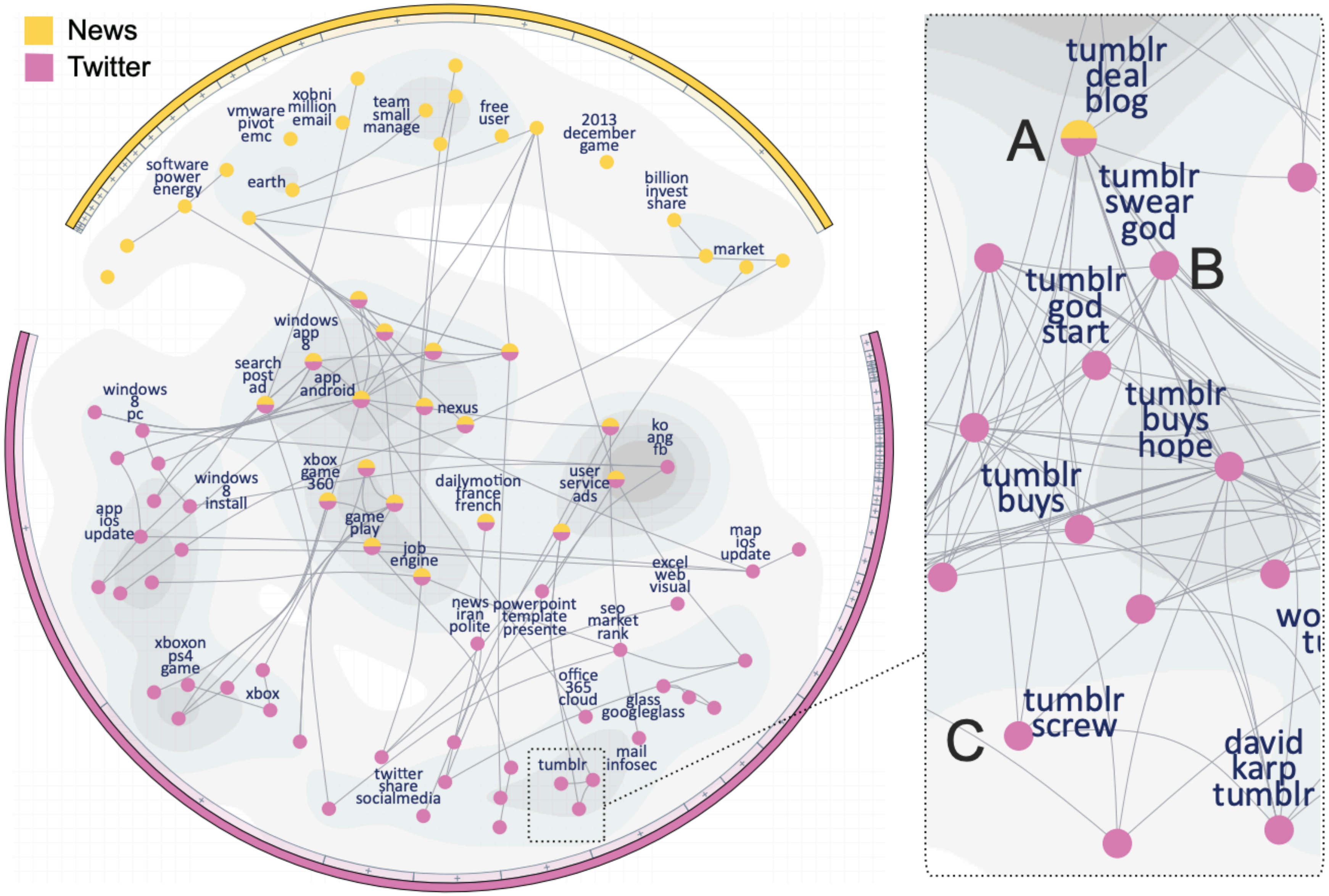


Fig. 1. ThemeDelta visualization for Barack Obama' campaign speeches during the U.S. 2012 presidential election (until September 10, 2012). Green lines are shared terms between Obama and Romney. Data from the "The American Presidency Project" at UCSB (http://www.presidency.ucsb.edu/).

Cui, Weiwei, et al. "Textflow: Towards better understanding of evolving topics in text." *IEEE transactions on visualization and computer graphics* 17.12 (2011): 2412-2421.

Topic Modeling



(a) (b)

Figure 12. Matching the news corpus with the Twitter corpus: (a) overview; (b) comparison of Tumblr related topics.

Liu, Shixia, et al. "Topicpanorama: A full picture of relevant topics." *2014 IEEE Conference on Visual Analytics Science and Technology (VAST)*. IEEE, 2014.

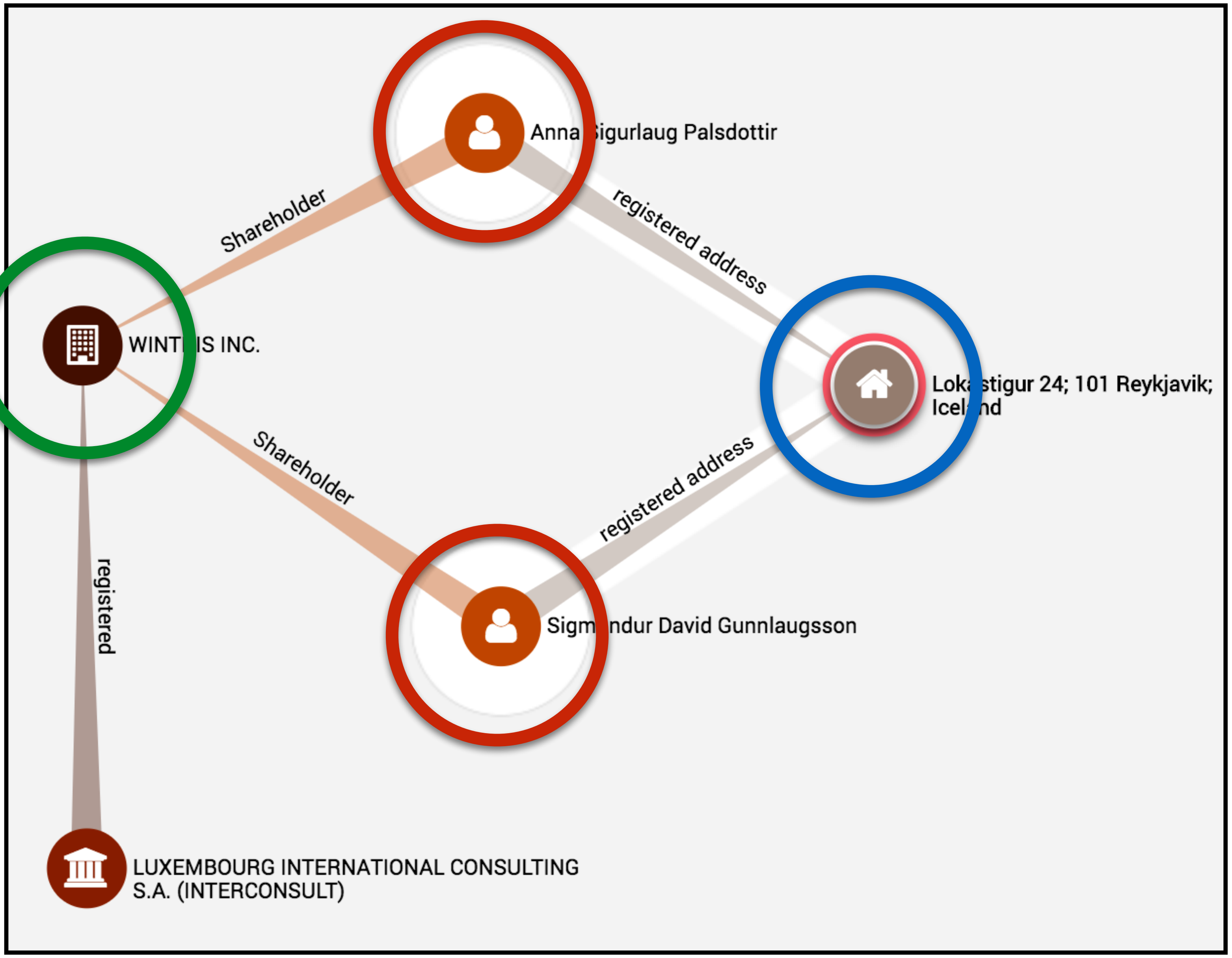
Named Entities

Named Entities are:

- People
- Organizations/Companies
- Locations
- Time Events

Panama Papers Networks use different visual encodings for:

People
Organizations
Locations



What is possible?

⟨Google⟩$_1$, headquartered in ⟨Mountain View⟩$_6$, unveiled the new ⟨Android⟩$_4$

⟨phone⟩$_3$ at the ⟨Consumer Electronic Show⟩$_7$. ⟨Sundar Pichai⟩$_5$ said in his

1. Google — ORGANIZATION
Sentiment: Score 0 Magnitude 0
Wikipedia Article
Salience: 0.26

2. users — PERSON
Sentiment: Score 0.4 Magnitude 0.9
Salience: 0.15

3. phone — CONSUMER GOOD
Sentiment: Score 0 Magnitude 0
Salience: 0.13

4. Android — CONSUMER GOOD
Sentiment: Score 0.1 Magnitude 0.2
Wikipedia Article
Salience: 0.12

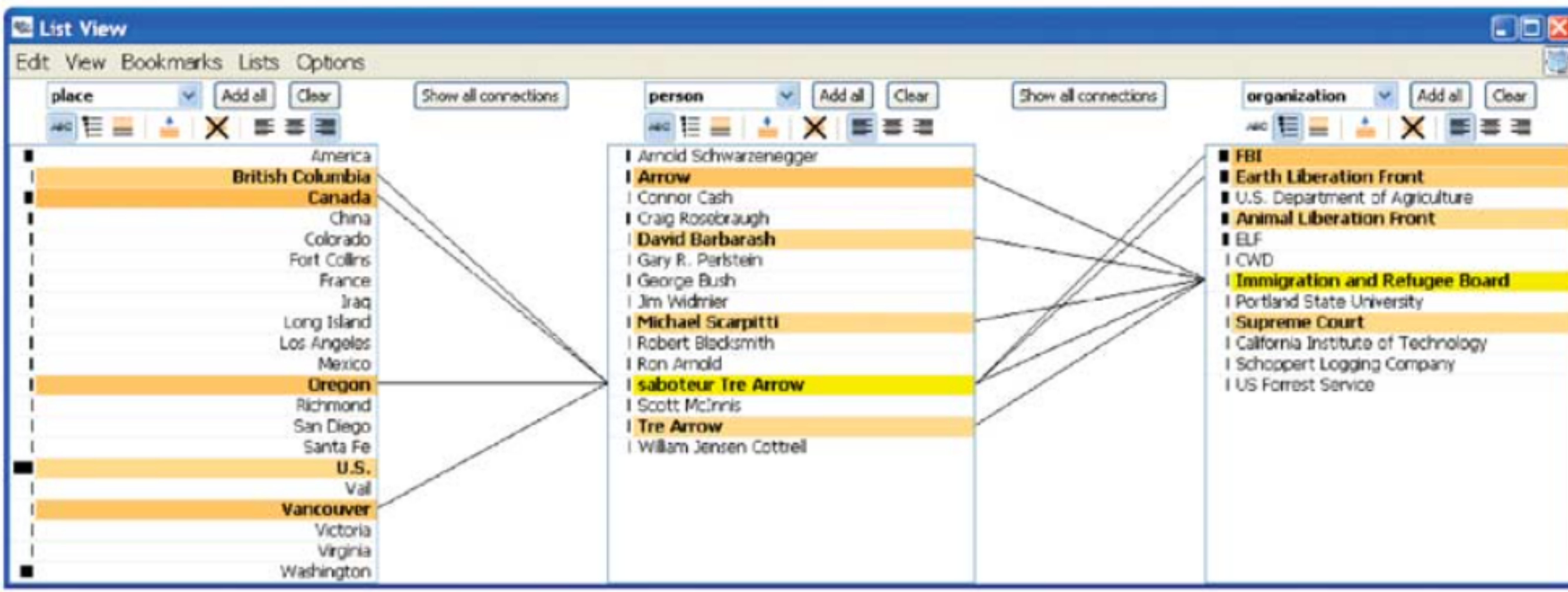
5. Sundar Pichai — PERSON
Sentiment: Score 0 Magnitude 0.1
Wikipedia Article
Salience: 0.11

6. Mountain View — LOCATION
Sentiment: Score 0 Magnitude 0
Wikipedia Article
Salience: 0.10

https://cloud.google.com/natural-language/

Named Entities

Jigsaw: Supporting investigative analysis through interactive visualization, Stasko et. al, Information Visualization, 2008



Sentiment Analysis

What is Sentiment Analysis?

https://cloud.google.com/natural-language/

What is Sentiment Analysis?

Sentiment analysis refers to the use of natural language processing, text analysis, computational linguistics, and biometrics to systematically identify, extract, quantify, and study affective states and subjective information.

https://monkeylearn.com/sentiment-analysis/

Example Sentence:
Google, headquartered in Mountain View, unveiled the new Android phone at the Consumer Electronic Show. Sundar Pichai said in his keynote that users love their new Android phones.

https://cloud.google.com/natural-language/

Sentiment Analysis

1. **Google** ORGANIZATION
Sentiment: Score 0 Magnitude 0

2. **users** PERSON
Sentiment: Score 0.6 Magnitude 1.3

3. **phone** CONSUMER GOOD
Sentiment: Score 0 Magnitude 0

4. **Android** CONSUMER GOOD
Sentiment: Score 0.5 Magnitude 1

5. **Sundar Pichai** PERSON
Sentiment: Score 0.1 Magnitude 0.2

6. **Mountain View** LOCATION
Sentiment: Score 0 Magnitude 0

7. **Consumer Electroni...** EVENT
Sentiment: Score 0.1 Magnitude 0.1

8. **phones** CONSUMER GOOD
Sentiment: Score 0.7 Magnitude 0.7

9. **keynote** OTHER
Sentiment: Score 0.2 Magnitude 0.2

https://cloud.google.com/natural-language/

Updated Sentence:
Google, headquartered in Mountain View, unveiled the new **and very expensive** Android phone at the Consumer Electronic Show. Sundar Pichai said in his keynote that users love their new Android phones.

https://cloud.google.com/natural-language/

Sentiment Analysis

1. **Google** ORGANIZATION
Sentiment: Score 0 Magnitude 0

2. **users** PERSON
Sentiment: Score 0.6 Magnitude 1.3

3. **phone** CONSUMER GOOD
Sentiment: Score -0.8 Magnitude 0.8

4. **Android** CONSUMER GOOD
Sentiment: Score 0 Magnitude 1.7

5. **Mountain View** LOCATION
Sentiment: Score 0 Magnitude 0

6. **Consumer Electroni...** EVENT
Sentiment: Score 0 Magnitude 0

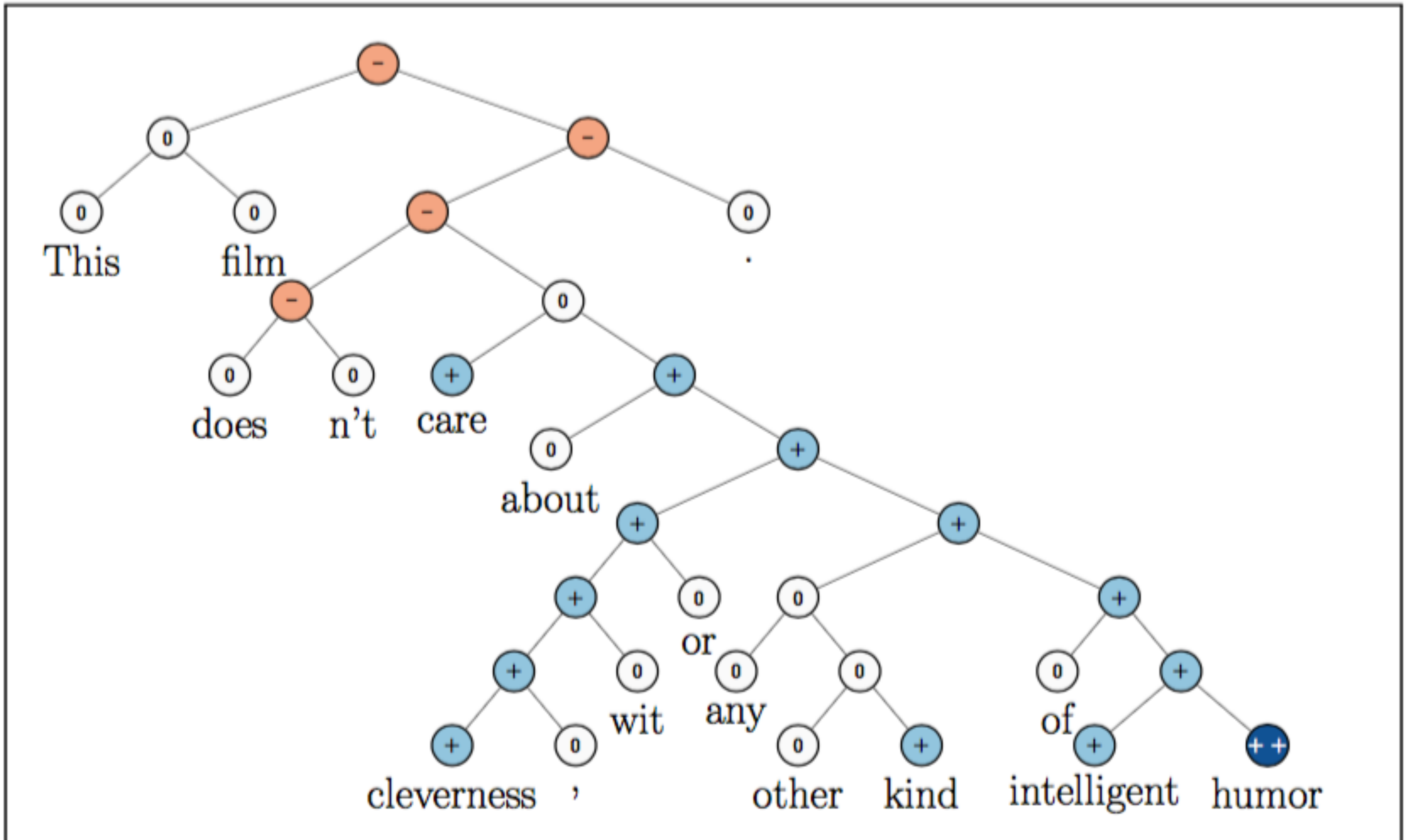
7. **Sundar Pichai** PERSON
Sentiment: Score 0.1 Magnitude 0.2

8. **phones** CONSUMER GOOD
Sentiment: Score 0.7 Magnitude 0.7

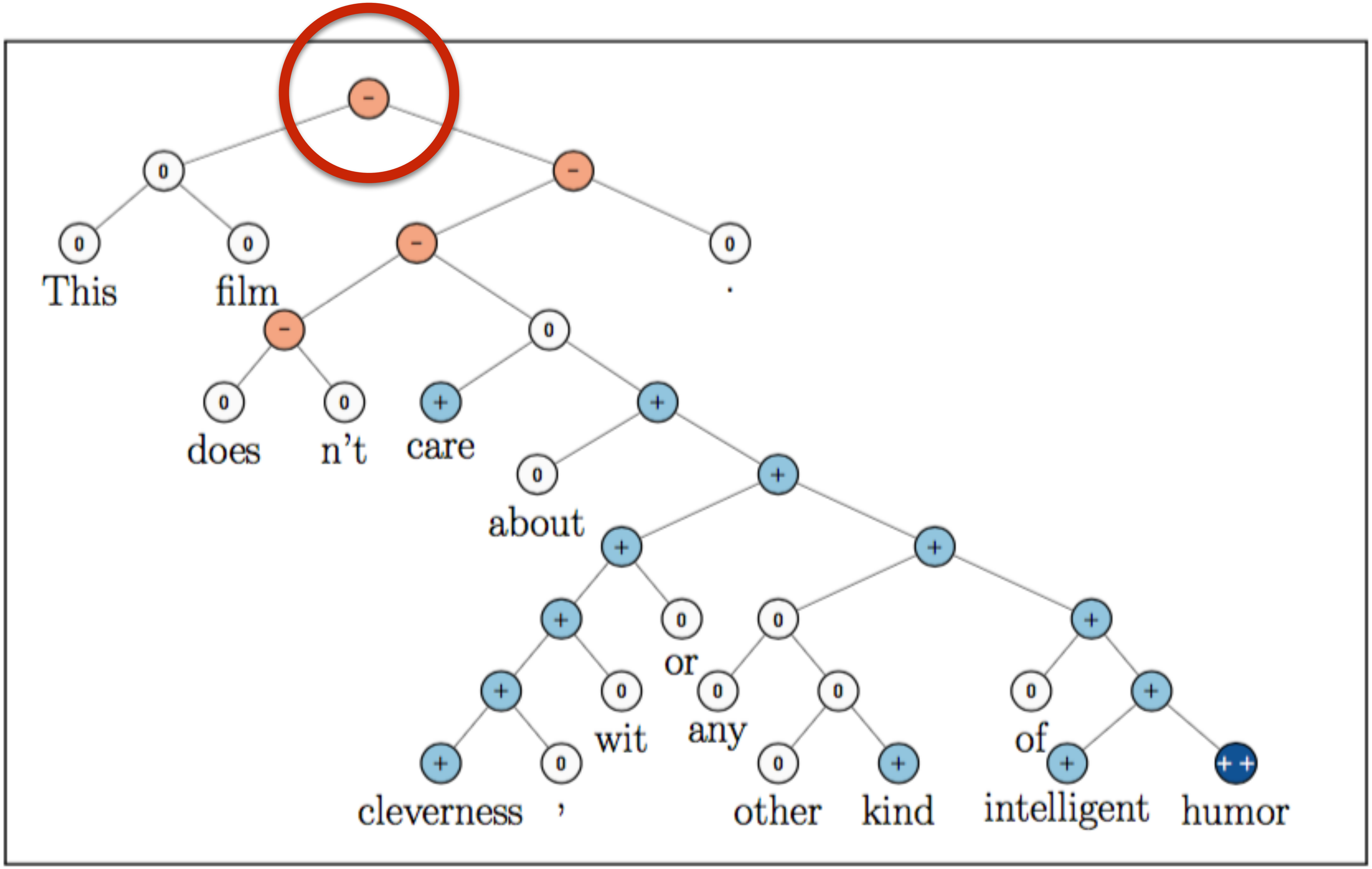
9. **keynote** OTHER
Sentiment: Score 0.2 Magnitude 0.2

https://cloud.google.com/natural-language/

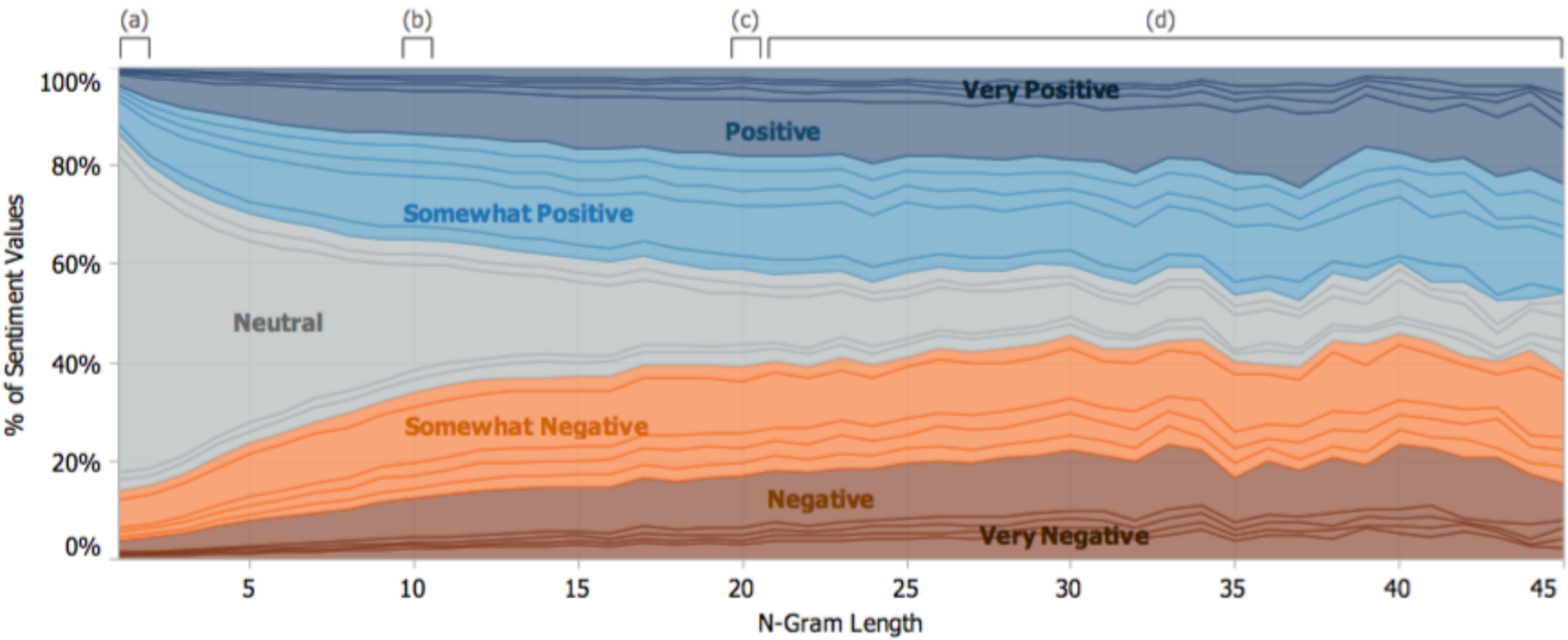
Recursive Deep Models for Semantic Compositionality Over a Sentiment Treebank, Socher et al., EMNLP 2013



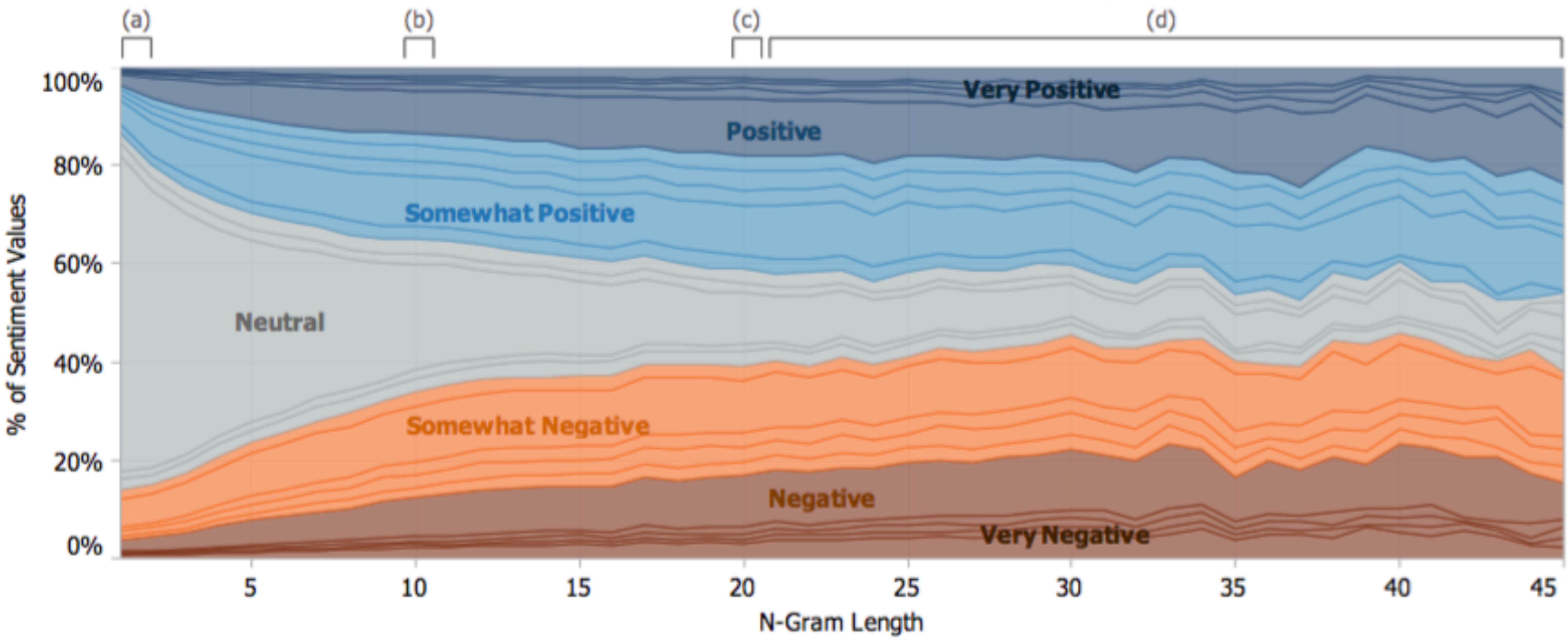
Although more nodes within this sentence have a positive meaning, the sentence is labeled negative!



Recursive Deep Models for Semantic Compositionality Over a Sentiment Treebank, Socher et al., EMNLP 2013



The longer a phrase gets, the more likely it is either positive or negative!



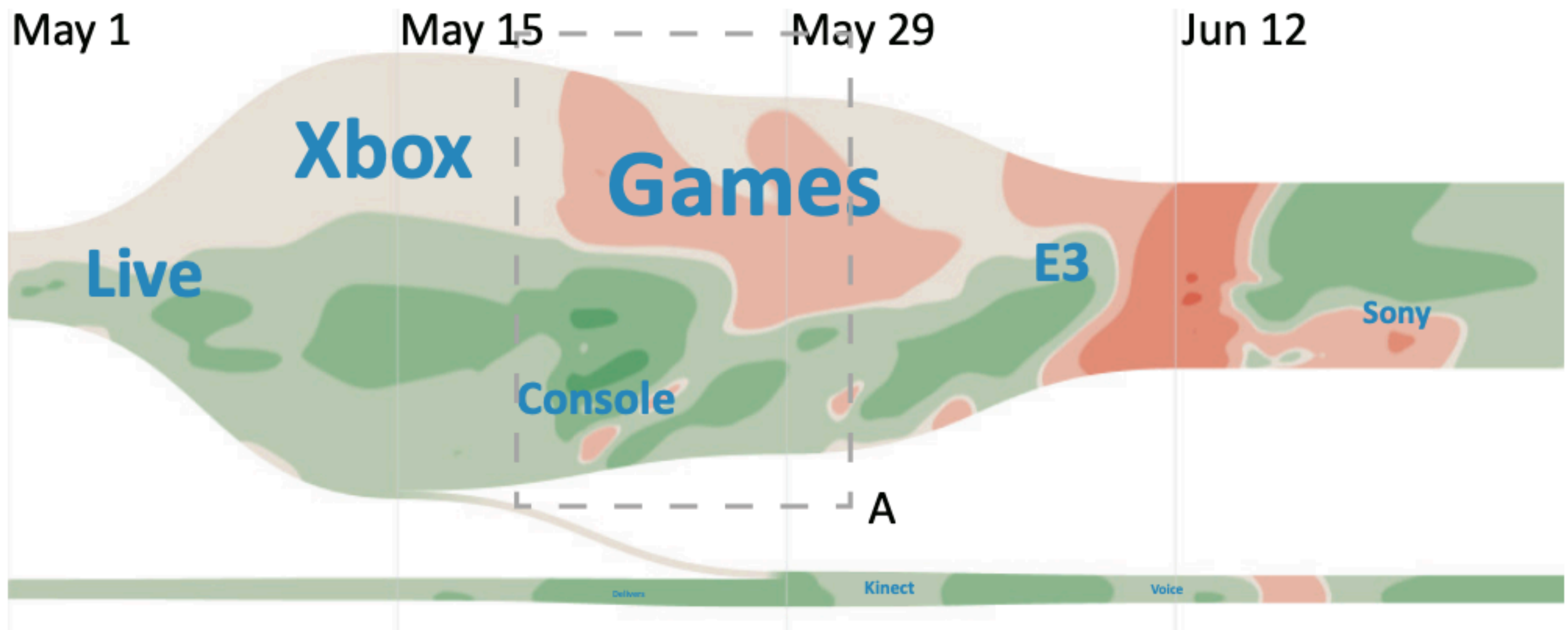
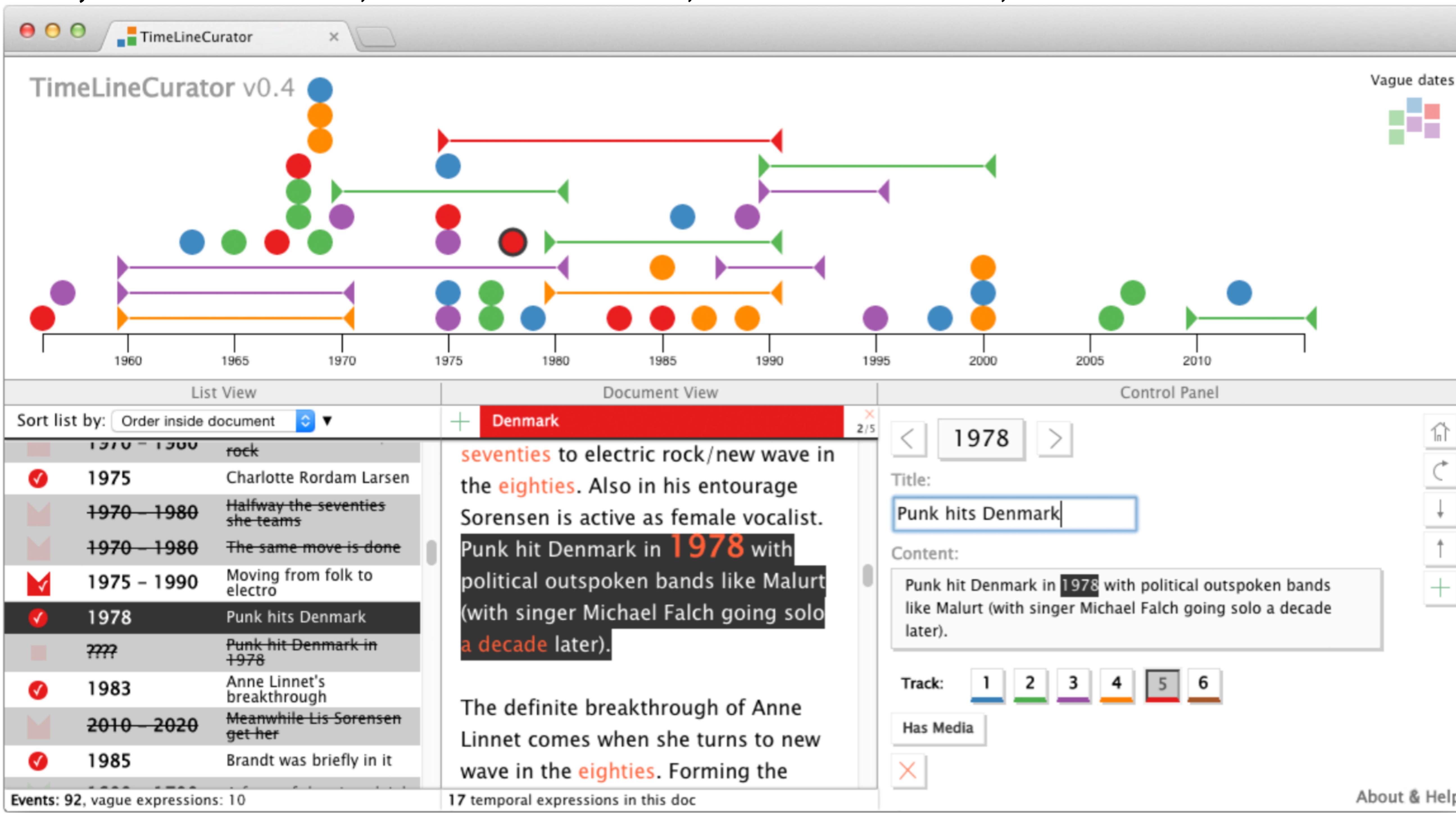


Fig. 7. Opinion diffusion on the Xbox topic from the period between May 15 to 29 when Xbox One was announced.

Wu, Yingcai, et al. "Opinionflow: Visual analysis of opinion diffusion on social media." *IEEE transactions on visualization and computer graphics* 20.12 (2014): 1763-1772.

Temporal Events

TimeLineCurator: Interactive Authoring of Visual Timelines from Unstructured Text; Johanna Fulda, Matthew Brehmer, Tamara Munzner; IEEE VAST 2015



Time Events get extracted from an input text source

Time Events get extracted from an input text source and visualised as Timeline

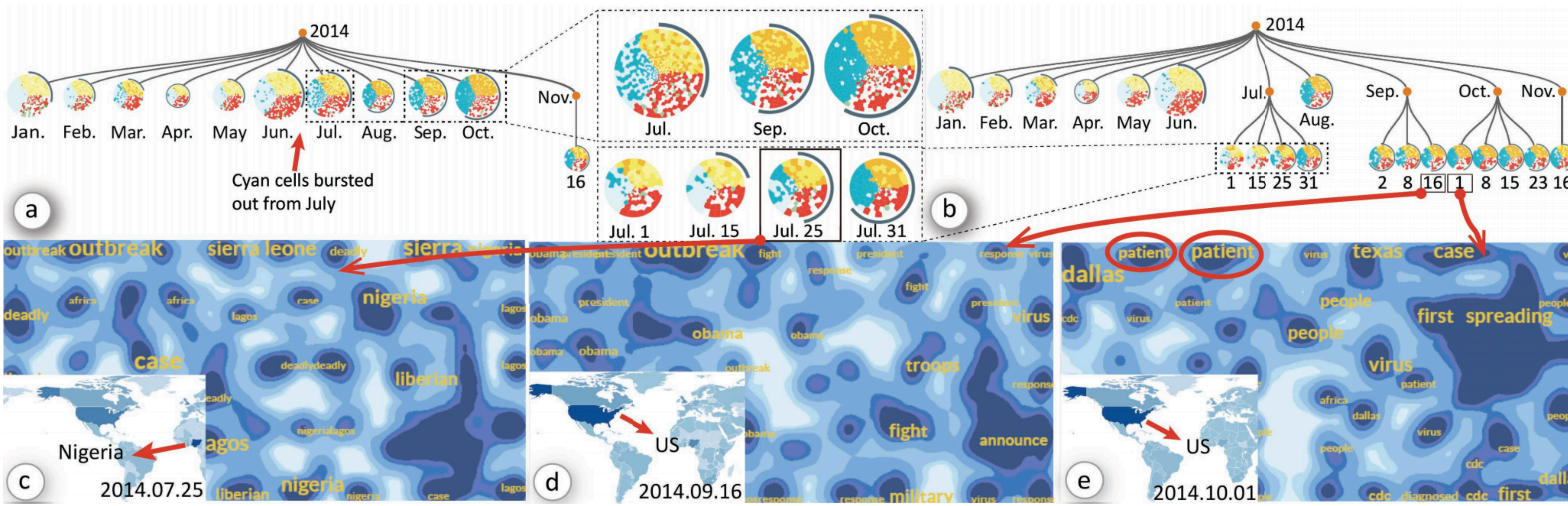


Fig. 8. Case study of the Ebola outbreak: (a) visual tree of the past subevents with respect to the Ebola outbreak in 2014; (b) visual tree created by expanding the nodes of July, September, and October in the original tree shown in (a); (c)-(e) geographic maps showing the geospatial changes of three subevents highlighted in (b) as well as the topic maps showing the topical changes of the three same subevents.

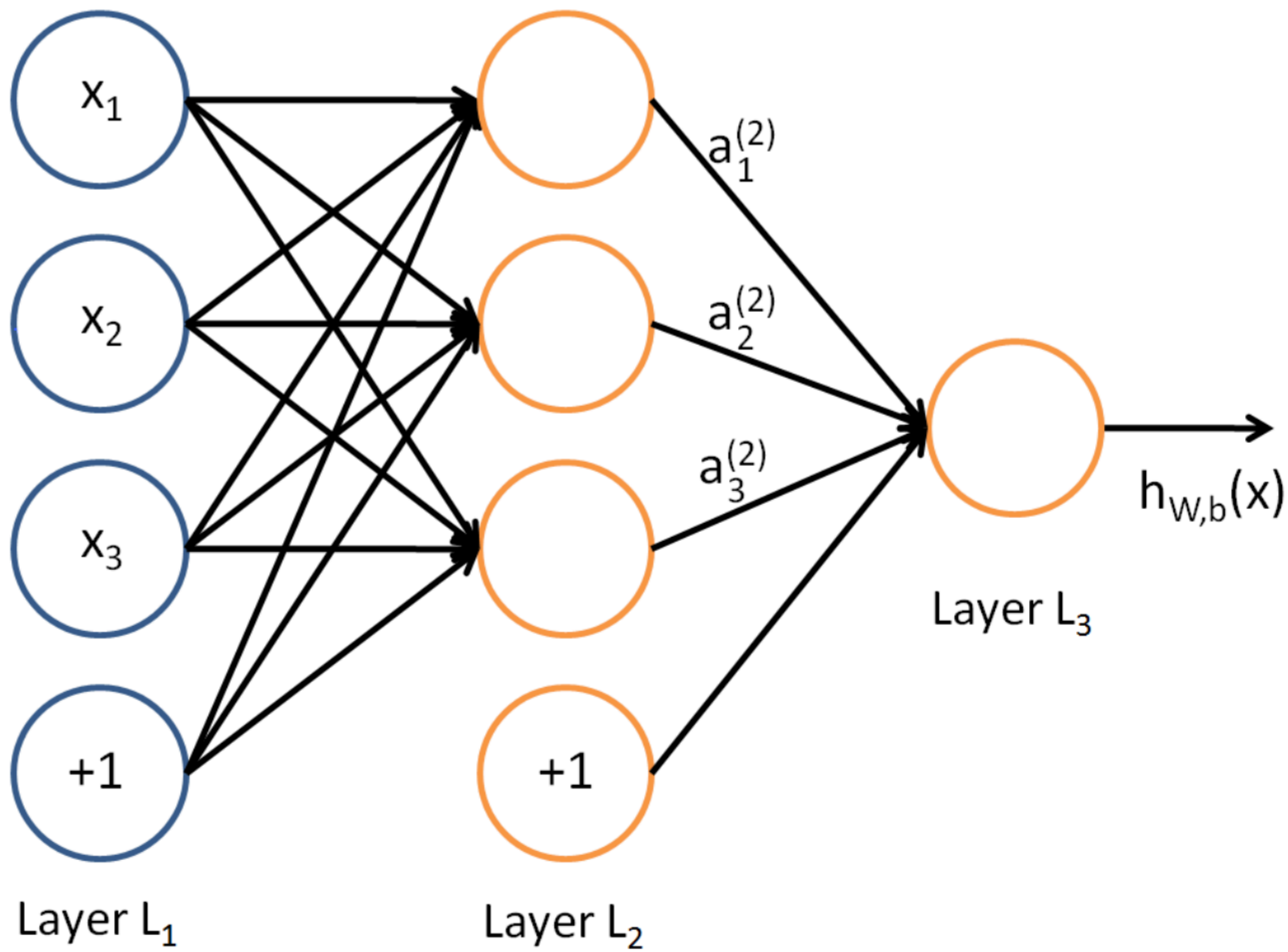
Wu, Yingcai, et al. "StreamExplorer: A multi-stage system for visually exploring events in social streams." *IEEE transactions on visualization and computer graphics* 24.10 (2018): 2758-2772.

Deep Learning

What is deep learning?

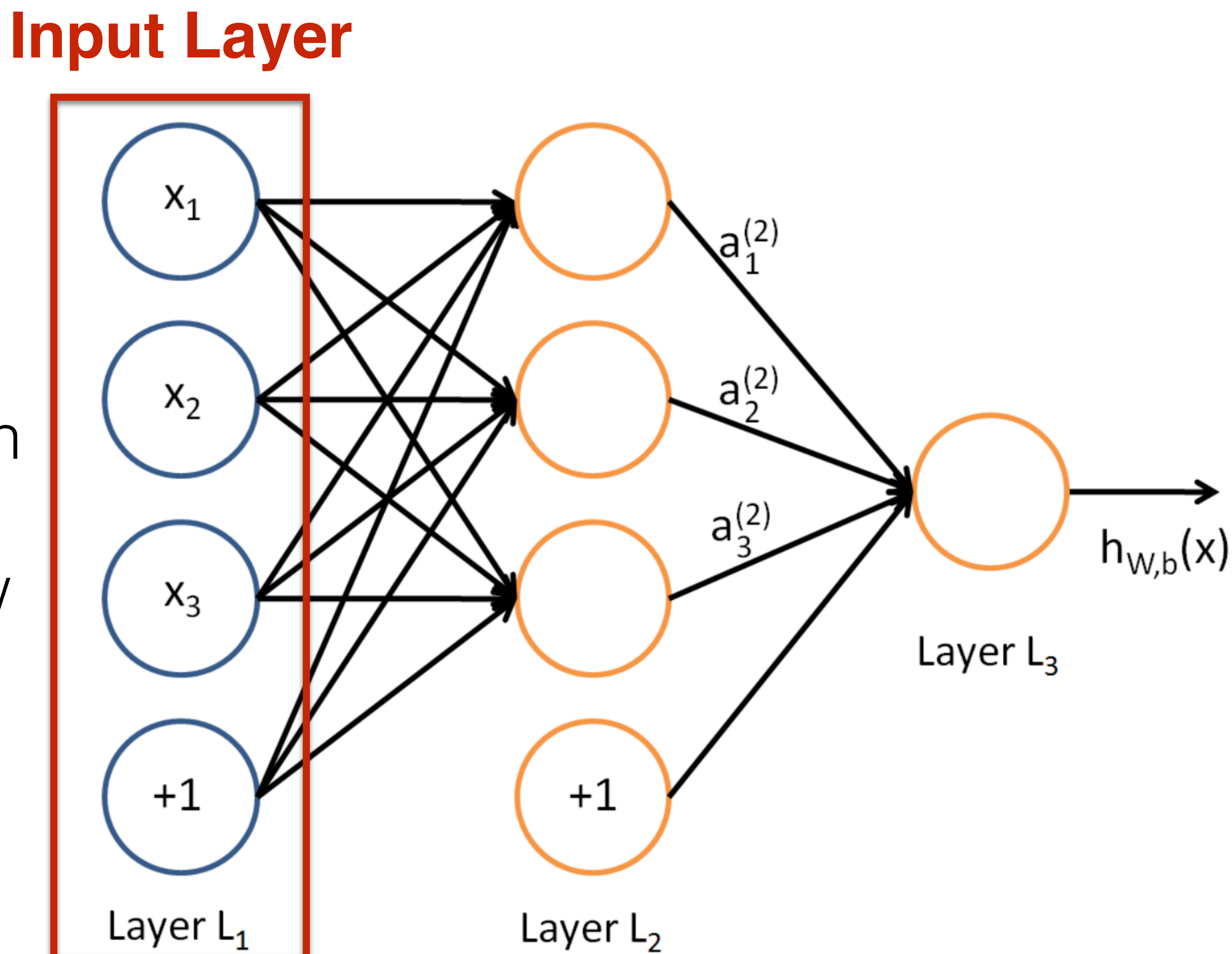
https://cloud.google.com/natural-language/

Starting from a Neural Network (NN):
Google, headquartered in Mountain View, unveiled the new



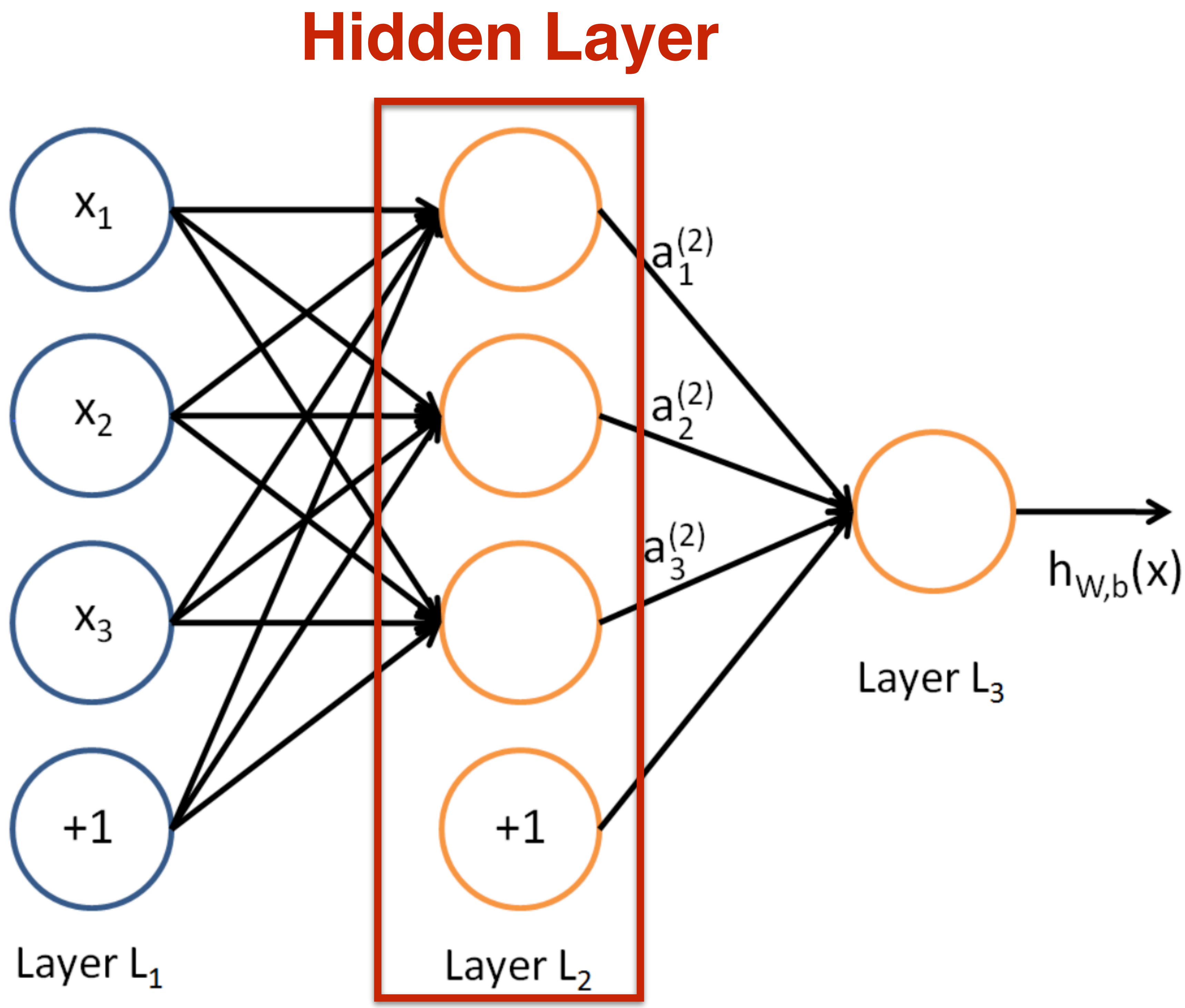
https://towardsdatascience.com/understanding-neural-networks-from-neuron-to-rnn-cnn-and-deep-learning-cd88e90e0a90

Starting from a Neural Network (NN):
Google, headquartered in Mountain View, unveiled the new



https://towardsdatascience.com/understanding-neural-networks-from-neuron-to-rnn-cnn-and-deep-learning-cd88e90e0a90

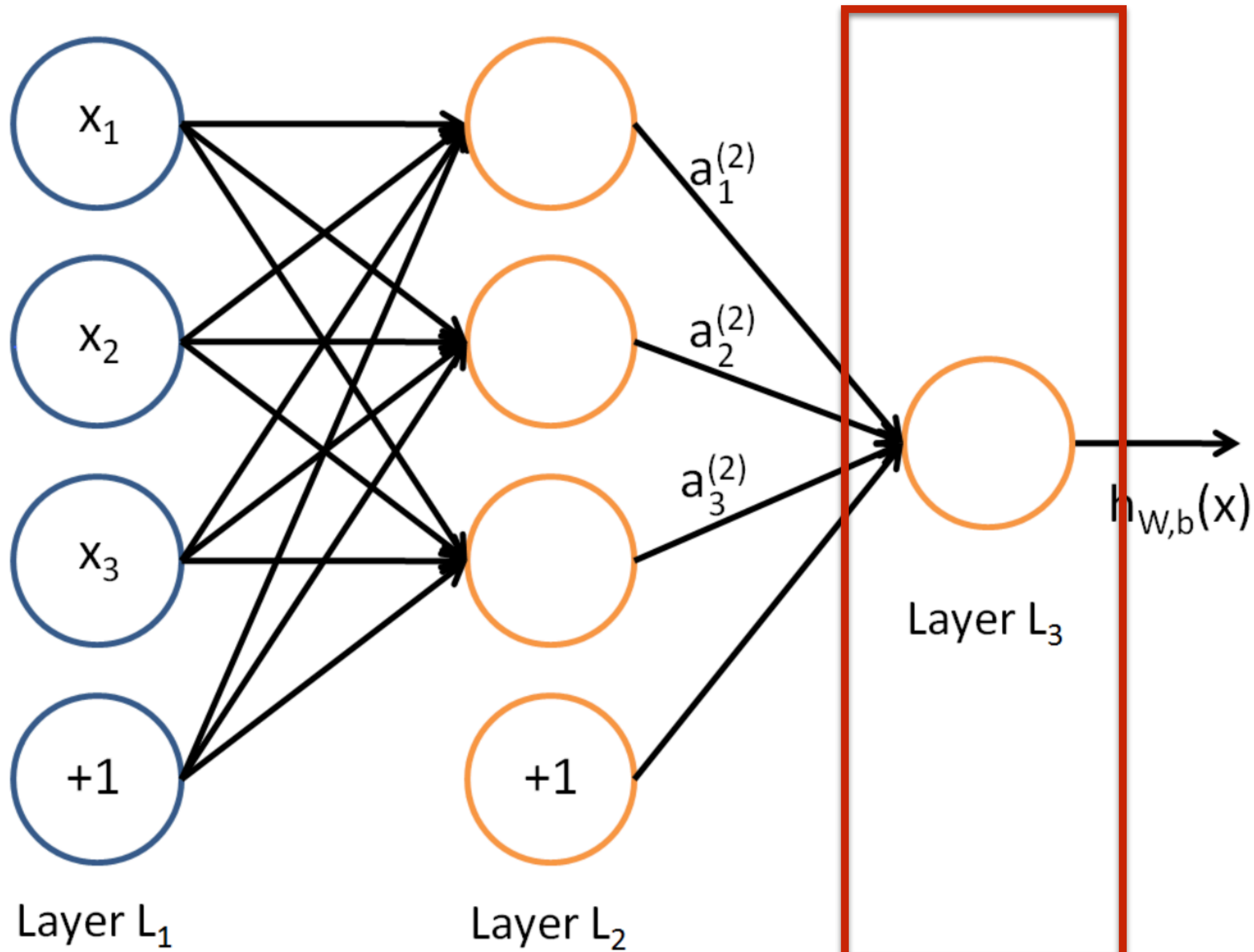
Starting from a Neural Network (NN):
Google, headquartered in Mountain View, unveiled the new



https://towardsdatascience.com/understanding-neural-networks-from-neuron-to-rnn-cnn-and-deep-learning-cd88e90e0a90

Output Layer

Starting from a Neural Network (NN):
Google, headquartered in Mountain View, unveiled the new



https://towardsdatascience.com/understanding-neural-networks-from-neuron-to-rnn-cnn-and-deep-learning-cd88e90e0a90

Deep Neural Network:
More than one hidden layer.

https://medium.com/dair-ai/deep-learning-for-nlp-an-overview-of-recent-trends-d0d8f40a776d

Deep Neural Network:
More than one hidden layer.

Deep Learning:
Representation learning instead of feature engineering.

https://medium.com/dair-ai/deep-learning-for-nlp-an-overview-of-recent-trends-d0d8f40a776d

Deep Neural Network:
More than one hidden layer.

Deep Learning:
Representation learning instead of feature engineering.

Types:
Convolutional (CNN), Recurrent (RNN), Variational Autoencoders (VAE), Generative Adversarial Networks (GAN), …

https://medium.com/dair-ai/deep-learning-for-nlp-an-overview-of-recent-trends-d0d8f40a776d

Tasks:
POS tagging
Parsing
Named-Entity Recognition
Semantic Role Labeling
Sentiment Classification
Machine translation
Question answering
Dialogue systems
Contextual Embeddings
…

Young, Tom, et al. "Recent trends in deep learning based natural language processing." *ieee Computational intelligenCe magazine* 13.3 (2018): 55-75.

TABLE VII: **Machine translation** (Numbers are BLEU scores)

Paper	Model	WMT2014 English2German	WMT2014 English2French
Cho et al. [82]	Phrase table with neural features		34.50
Sutskever et al. [74]	Reranking phrase-based SMT best list with LSTM seq2seq		36.5
Wu et al. [162]	Residual LSTM seq2seq + Reinforcement learning refining	26.30	41.16
Gehring et al. [163]	seq2seq with CNN	26.36	41.29
Vaswani et al. [113]	Attention mechanism	28.4	41.0

BLEU Score:

"the closer a machine translation is to a professional human translation, the better it is" – this is the central idea behind BLEU.

https://en.wikipedia.org/wiki/BLEU

Young, Tom, et al. "Recent trends in deep learning based natural language processing." *ieee Computational intelligenCe magazine* 13.3 (2018): 55-75.

TABLE X: Comparison of ELMo + Baseline with the previous state of the art (SOTA) on various NLP tasks. The table has been adapted from [41]. SOTA results have been taken from [41]; SQUAD [166]: QA task; SNLI [178]: Stanford Natural Language Inference task; SRL [153]: Semantic Role Labelling; Coref [179]: Coreference Resolution; NER [180]: Named Entity Recognition; SST-5 [4]: Stanford Sentiment Treebank 5-class classification;

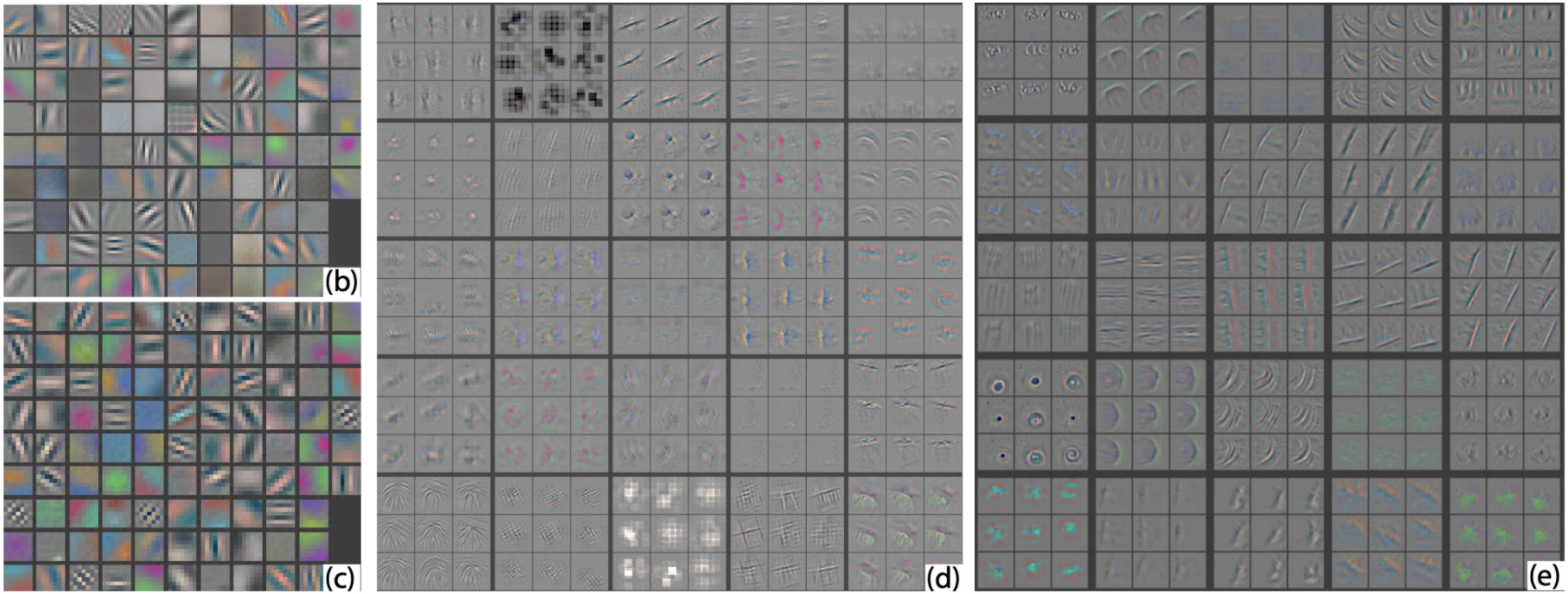
Task	Previous SOTA	Previous SOTA Results	Baseline	ELMo + Baseline	Increase (Absolute/Relative)
SQuAD	Liu et al. [181]	84.4	81.1	85.8	4.7 / 24.9%
SNLI	Qian et al. [182]	88.6	88.0	88.70 $\pm$ 0.17	0.7 / 5.8%
SRL	Luheng et al. [183]	81.7	81.4	84.6	3.2 / 17.2%
Coref	Kenton et al. [184]	67.2	67.2	70.4	3.2 / 9.8%
NER	Matthew et al. [185]	91.93 $\pm$ 0.19	90.15	92.22 $\pm$ 0.10	2.06 / 21%
SST-5	Bryan et al. [186]	53.7	51.4	54.7 0.5	3.3 / 6.8%

Task	BiLSTM+ ELMo+Attn	BERT
QNLI	79.9	91.1
SST-2	90.9	94.9
STS-B	73.3	86.5
RTE	56.8	70.1
SQuAD	85.8	91.1
NER	92.2	92.8

TABLE XI: QNLI [187]: Question Natural Language Inference task; SST-2 [4]: Stanford Sentiment Treebank binary classification; STS-B [188]: Semantic Textual Similarity Benchmark; RTE [189]: Recognizing Textual Entailment; SQUAD [166]: QA task; NER [180]: Named Entity Recognition.

Young, Tom, et al. "Recent trends in deep learning based natural language processing." *ieee Computational intelligenCe magazine* 13.3 (2018): 55-75.

Deep Learning

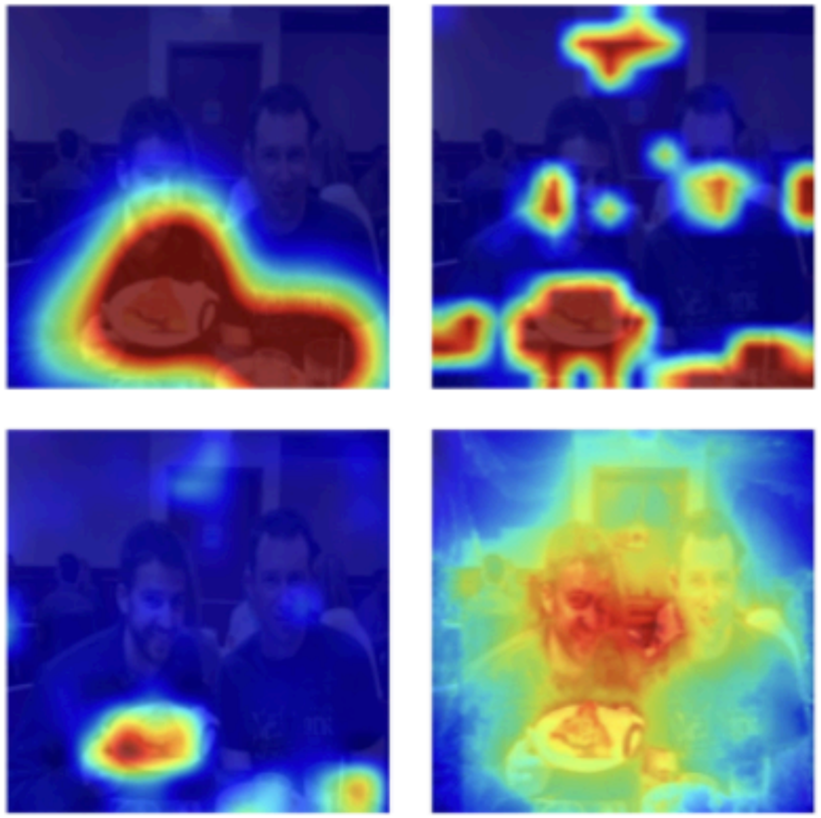


A figure from "Visualizing and Understanding Convolutional Networks" by Zeiler and Fergus, 2013, shows early results for a technique called **feature visualization** that visualizes the learned features in intermediate hidden layers of a deep learning model.

https://medium.com/multiple-views-visualization-research-explained/visualization-in-deep-learning-b29f0ec4f136

Attention

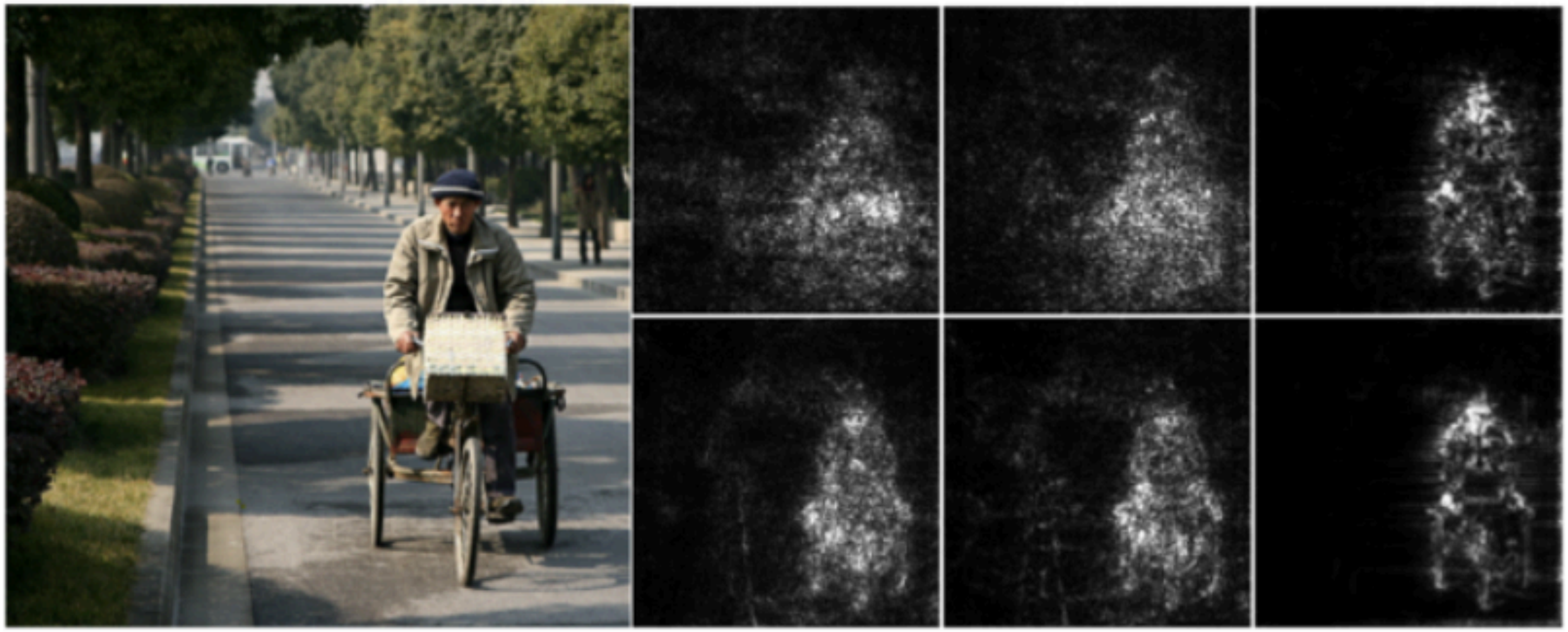
Das, Agrawal, et al. 2016



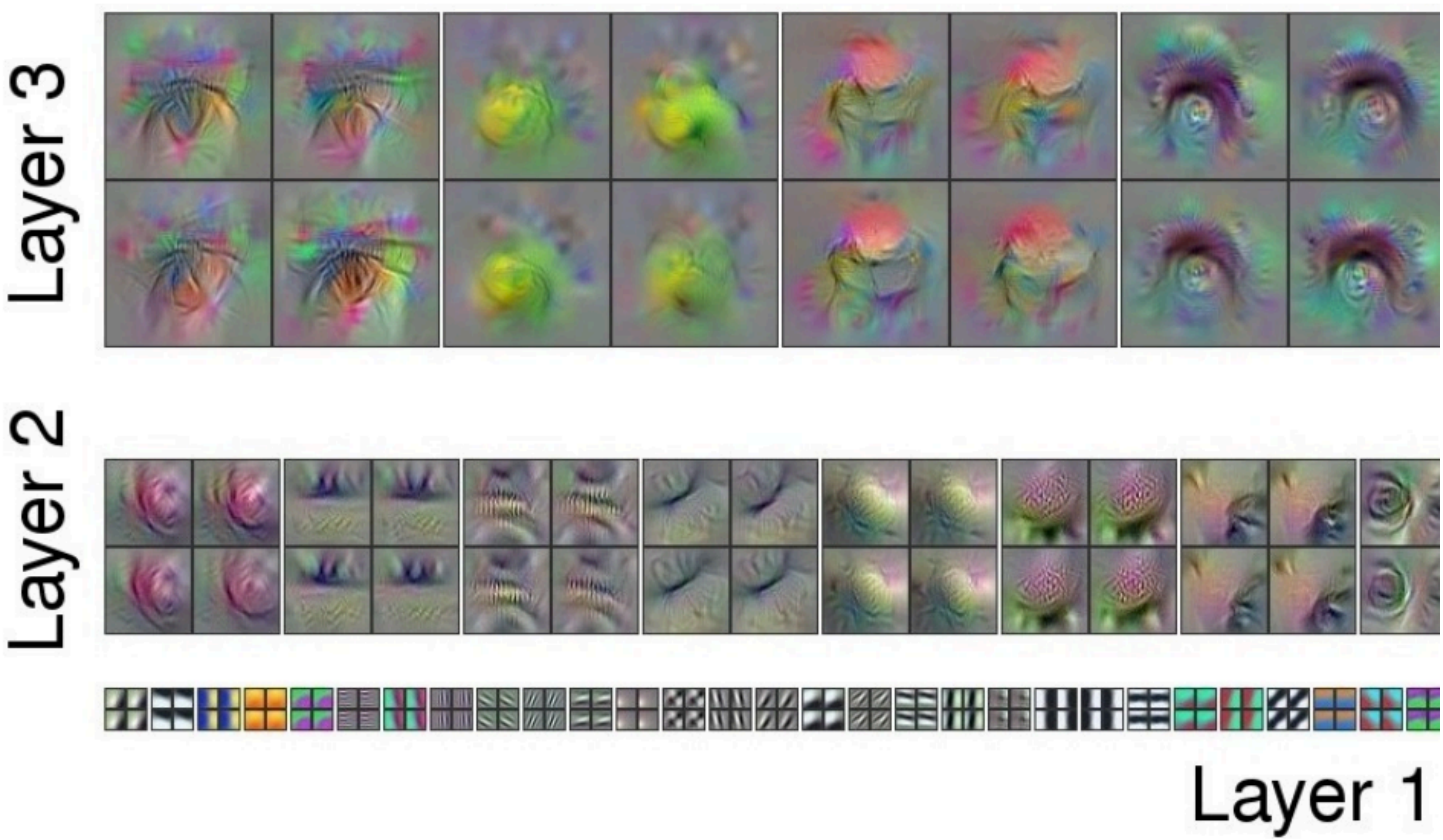
https://medium.com/multiple-views-visualization-research-explained/visualization-in-deep-learning-b29f0ec4f136

Saliency

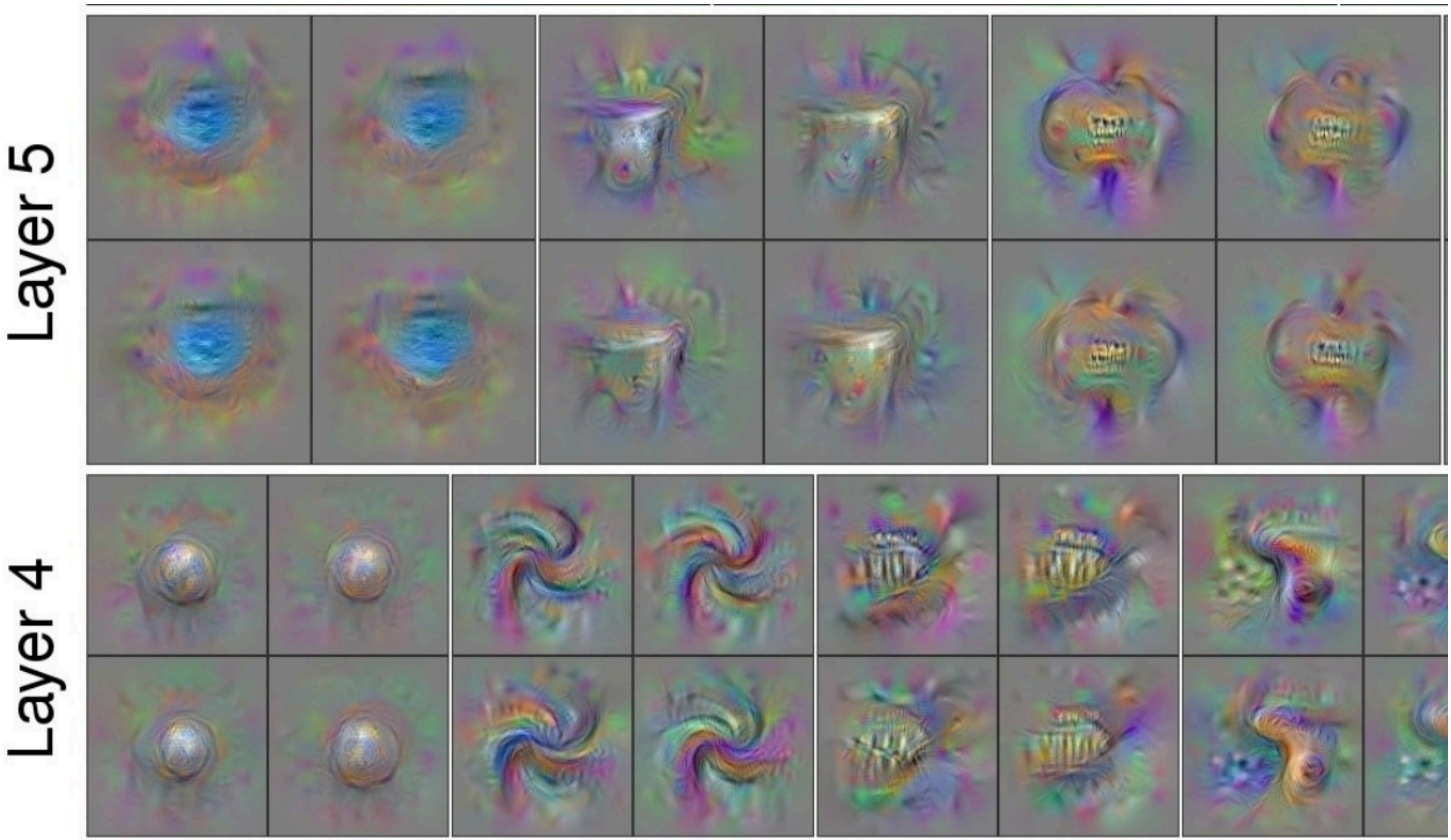
Smilkov, et al. 2017



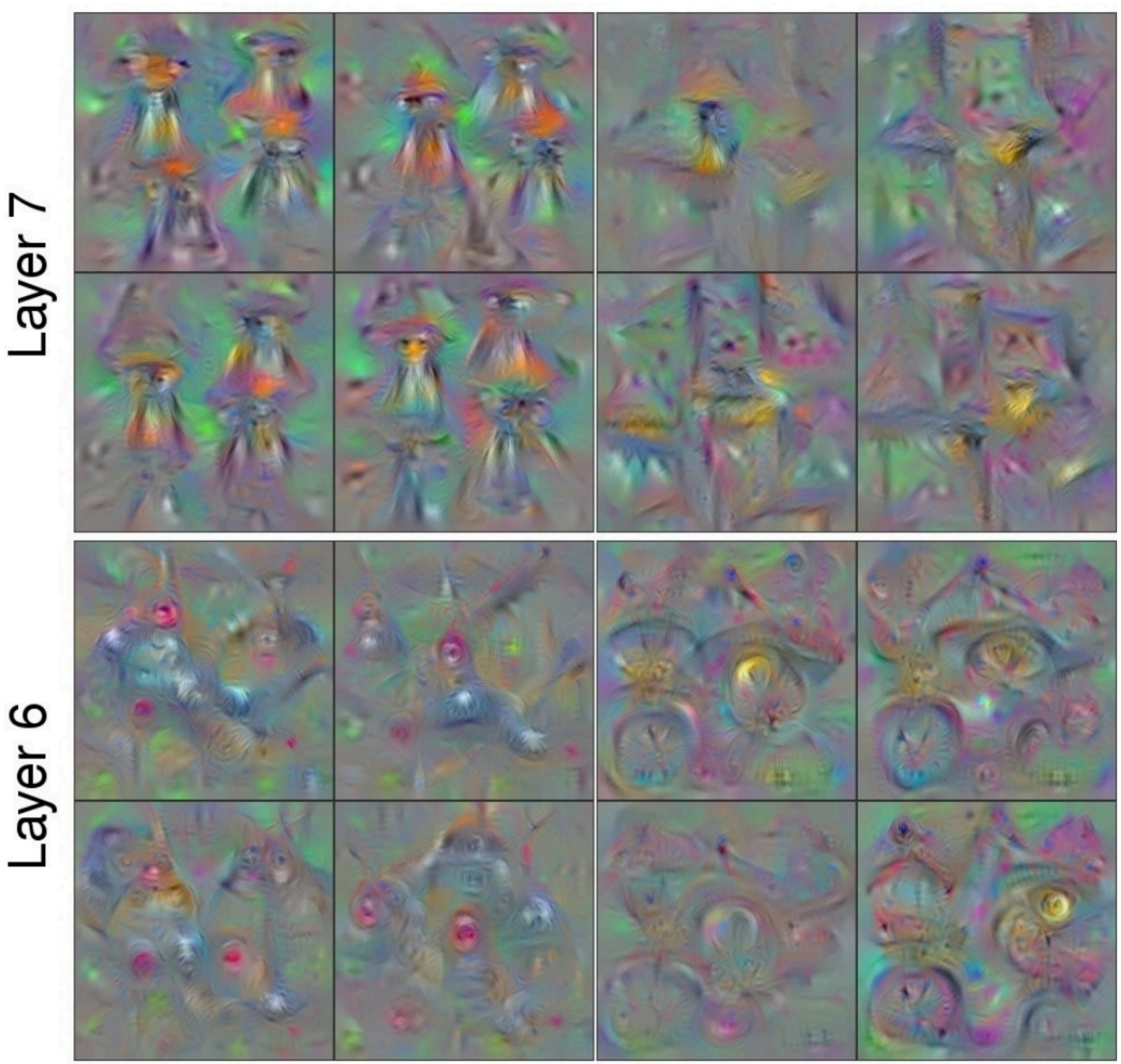
https://medium.com/multiple-views-visualization-research-explained/visualization-in-deep-learning-b29f0ec4f136



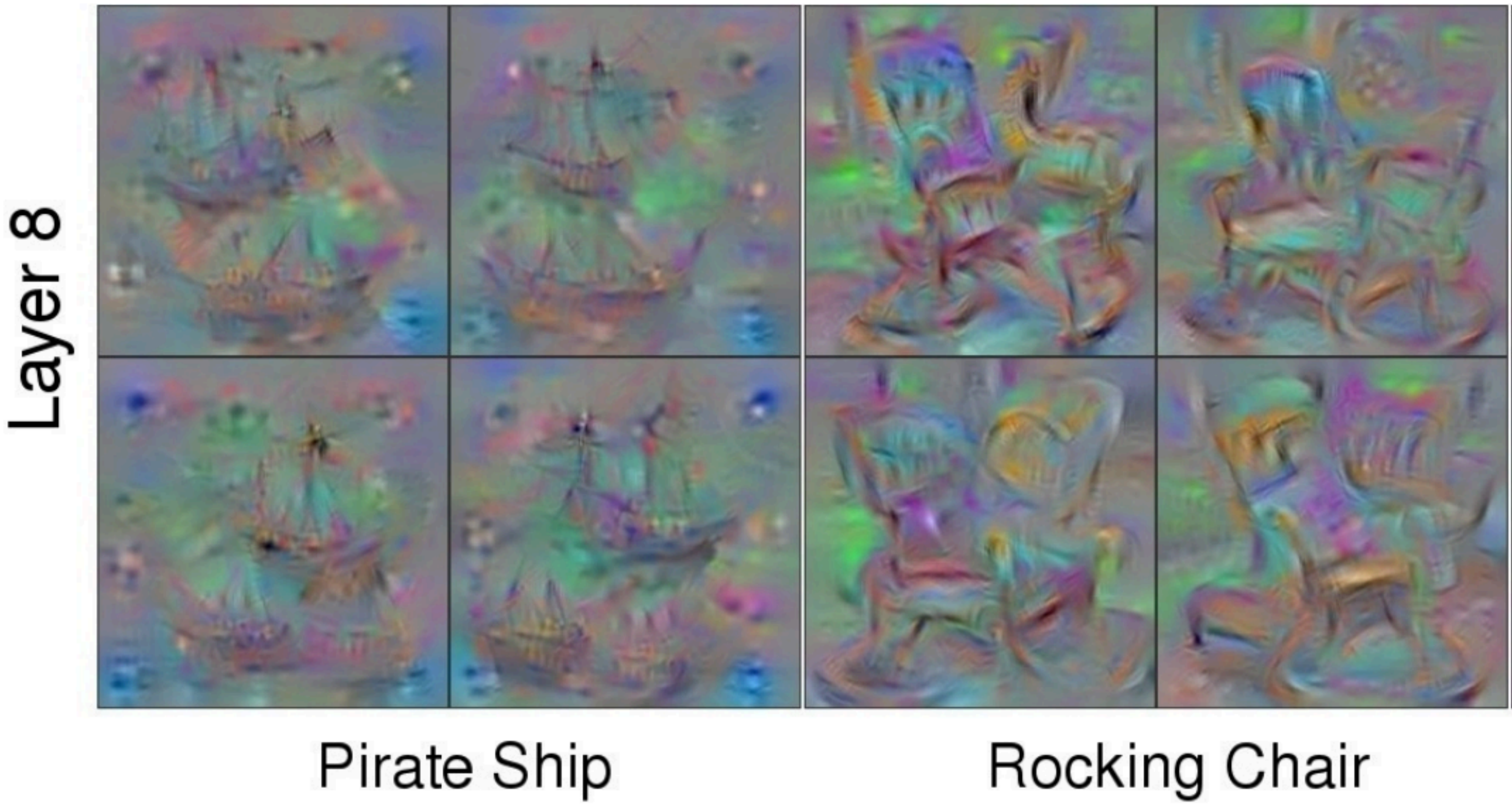
Yosinski, Jason, et al. "Understanding neural networks through deep visualization." *arXiv preprint arXiv:1506.06579* (2015).



Yosinski, Jason, et al. "Understanding neural networks through deep visualization." *arXiv preprint arXiv:1506.06579* (2015).



Yosinski, Jason, et al. "Understanding neural networks through deep visualization." *arXiv preprint arXiv:1506.06579* (2015).



Yosinski, Jason, et al. "Understanding neural networks through deep visualization." *arXiv preprint arXiv:1506.06579* (2015).

Projects