

# Web Data Visualization

Yasmin AlNoamany

L&S 88-2

University of California, Berkeley



All what you need to know about me

# Ph.D. in Computer Science



## Storytelling services



## Archived collections



## Archived enriched stories



# What I currently do

- Postdoc in Software & Data Curation
- BIDS Data Science Fellow
- Research interest:
  - Information Visualization
  - Web Science
  - Web Archiving
  - Reproducibility and Data Management



**Berkeley** Research Data Management







# Today

- Course overview
- Course goals
- Resources
- Introductions
- Today's lecture

# Course Overview



# Waitlist

- There are 25 registered and 10 on waitlist
  - Wow!
- It is highly likely that all who want to take can be accommodated
  - people are still shopping around for classes

# Class format

- Each lecture will be:
  - announcements
  - a lecture
  - a demo/activity/discussion
- Experts from around the campus and from industry
- Structure thus slightly tentative

# Administrivia

- Important Info
  - Announcements, lecture notes, syllabus, schedule, useful links on bCourses
- Communication:
  - posting grades
  - submitting assignments on bCourses, we may shift to Github later
- Bring a laptop to the class. You can borrow one for the semester from the library if you don't have your own.

# Grading

- Participation and attendance 10%
- Class activity 15%
- Assignments 40%
- Final project 35%
- Possible extra credit

# Participation

- Discussions in class
- Questions for required reading
  - written questions on reading
  - two for longer draft book chapters
  - one for shorter papers
  - due 12:00pm Tuesdays
    - Starting next week!
- Due before class: submit questions/comments on the readings of the previous week



# Attendance

- Tell me in advance if you know you'll miss class
- Question credit still possible if submitted in advance

# Class activity

- Bring a laptop to the class
- In-class design exercises
- This won't be every week!

# Assignments

- 5 assignments plus assignment 0!
- Assignment 1 will posted directly after the lecture

# Semester Project

- Teams of 2 or 3 people
- Grading
  - Milestones - 15 %
    - Milestone 1 (dataset and task definitions) - 4 %
    - Milestone 3 (project proposal) – 7 %
    - Milestone 4 (project status update) - 4 %
  - Demo and Presentation - 10 %
  - Paper - 10 %
- Presentations
  - Last weeks of class

There will be no final examination in  
this course!



# Class policy

- You can use existing code, consult with others, etc.
  - but you MUST document / reference where you adapted the code, advice, etc. of others. Give credit where credit is due!
  - use without attribution is plagiarism!
- Assignments are due at beginning of class:
  - upload code, report, data files, etc. to bCourses

# Academic Integrity

- Any evidence of cheating or plagiarism will result in a 0 grade for the assignment/exam, and the incident will be submitted to the department for further review

# Seeking Help

- Course site is your first reference
- Come to office hours
  - Tuesdays 3:00-4:00pm, starting next week
    - Or by appointment
- Email me at [yasminal@berkeley.edu](mailto:yasminal@berkeley.edu)
  - Put “LS 88-002 WDV” in the subject of your email

# Collaboration

- Discuss all questions with each other
- Submit homework individually (unless something else is mentioned), but feel free to discuss
- Project will be teams of 2 or 3
- The Limits of collaboration
  - Don't share solutions with each other (except project partners)
  - Copying solutions will result in failing the course

# We Want You to Succeed!

- If you are feeling overwhelmed, visit our office hours and talk with me



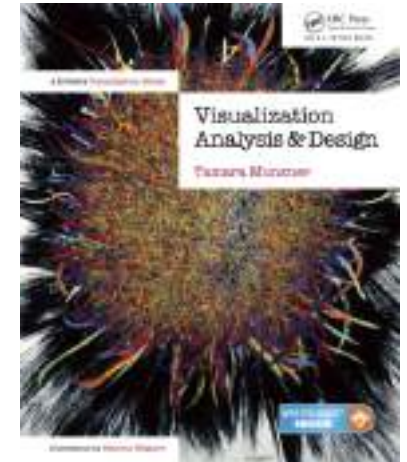
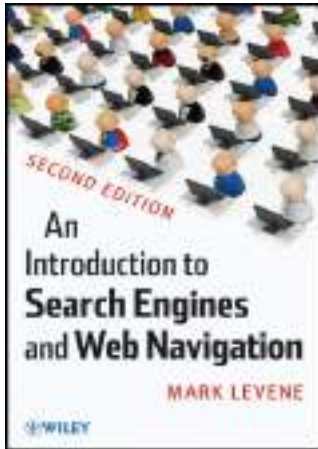
# Course topics

- Intro to the Web
- Intro to Web mining
- Types and formats of Web data
- Scrapping Web data
- Systems and toolkits for visualization
- Getting insight from Web data
  - Web content mining
  - Mining and visualizing social networks
  - Mining and visualizing Web usage
- Designing effective visualization
- What should we avoid in visualization

# Course goals

- Explain how the Web works and how it is used
- Differentiate among the different types of Web data
- Collect, preprocess, and analyze Web data
- Extract data from social network APIs
- Analyze and visualize social network data
- Design and implement different visualization techniques
- Critique existing visualization and suggest improvements
- Use different visualization tools (e.g., R and D3) to generate static and interactive visualizations

# Resources



- Textbooks: chapters are posted on the readings section
- Papers: links posted on course page

# Introductions

- Name
- Major
- Something interesting about yourself

# Today's lecture

# Intro to Web Data Visualization

- Why it is important
- The Web != internet
- Web evolution

Web

- importance of studying Web data
- Types of Web data
- Data formats

Data

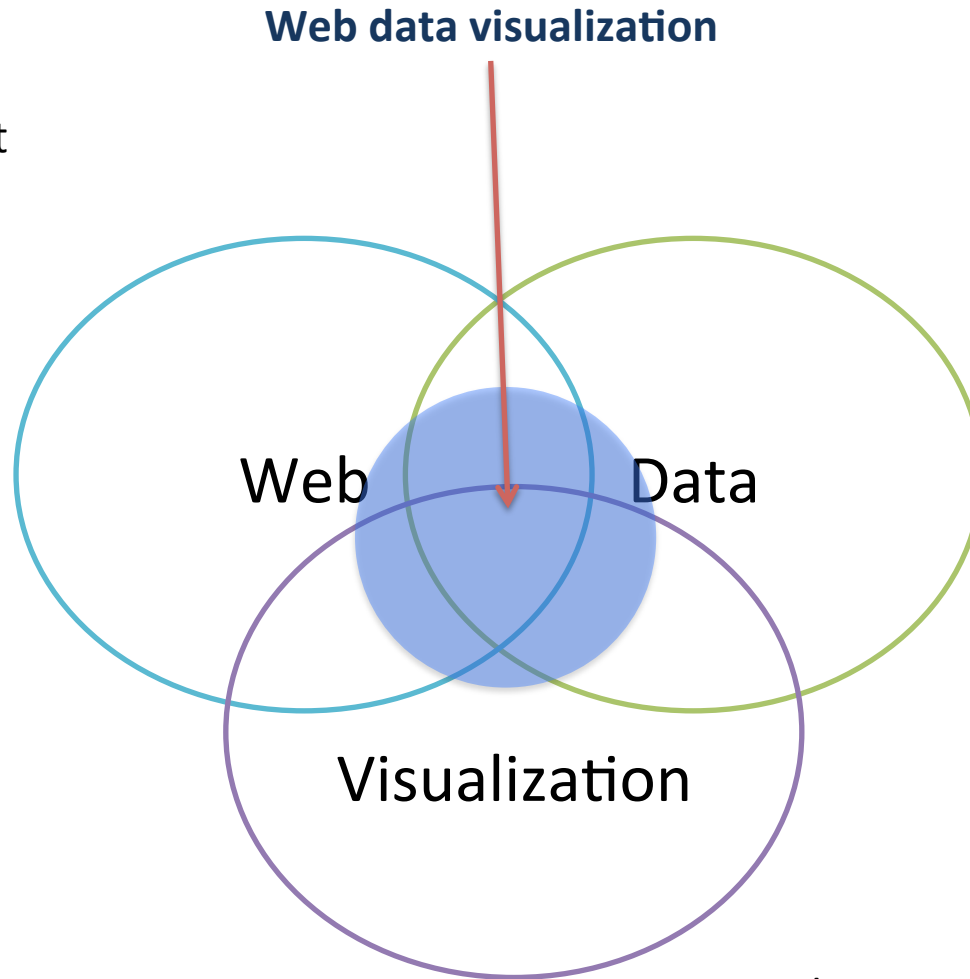
Visualization

- Why we care?
- Definitions
- History of Visualization

# Intro to Web Data Visualization

- Why it is important
- The Web != internet
- Web evolution

- importance of studying Web data
- Types of Web data
- Data formats



- Why we care?
- Definitions
- History of Visualization

# Resources

- (Required) History of the Web - article
- (Required) Evolution of the World Wide - article
- (Required) Ch 1 of Visualization Analysis and Design by Tamara Munzner – till section 1.5
- Ch 1 of Computational and Inferential Thinking: The Foundations of Data Science
- Architecture of the World Wide Web, Volume One, W3C Recommendation
- The Value of Visualization - video

For links and pdfs, check week 1 modules on bCourses



# Why we study the Web

# The Web is shaping our lives

- How we communicate



Image source: <http://marketingland.com/library/channel/social-media-marketing>

# The Web is shaping our lives

- How we shop



# The Web is shaping our lives

- Our understanding



Every non-hyperbolic tweet is from iPhone (his staff).

Every hyperbolic tweet is from Android (from him).



12:20 PM - 6 Aug 2016

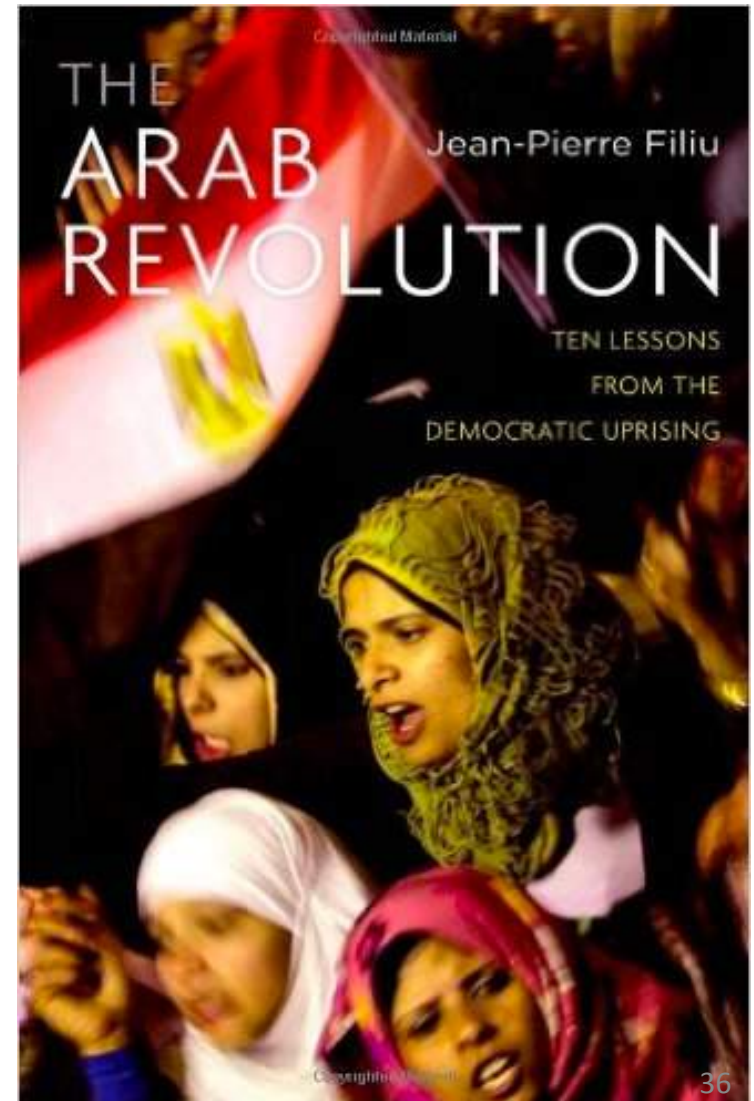
10,349 Retweets 14,504 Likes



265 10K 15K

# The Web is shaping our lives

- Politics - Revolutions





# Not only us!

## CRITTER CHAT

by ALISON SMITH



### SOUTHERN TWO-TOED SLOTH

**Slothful**  
Lives in: Northern South America  
Likes: Snoozing and snacking

**Fishbowl**  
CUBBER  
LittleBirdie

**LOWLAND TURTLE**  
TinyTurtle

**YUKIAMI GLASS FROG**  
FindTheFrog

**7 a.m.**

**Slothful**  
Sorry, guys. I'm going to be a little late. Aaaaahhh.

**TinyTurtle**  
Hurry up! You only come down from the tree canopy once a week to use the bathroom, so this is the only time we get to see you face-to-face for trunk-to-trunk!

**Slothful**  
I'm having trouble getting out of my bed, so I can't see you guys about it. Just a minute. By the way, take me outside to get some.

**FindTheFrog**  
Get a move on, Slothful! The moon is rising and the frogs are singing!

**Slothful**  
Shouldn't the birds be singing?

**LittleBirdie**  
I'm too busy tweeting! #LBJ

**4 p.m.**

**Slothful**  
Hey, FindTheFrog! I'm almost as green as you, but guess the birds are less noisy due to a heavy rain.

**FindTheFrog**  
Nice! My greenish skin through rain helps me hide in the leaves. I guess it works like that for you, too, huh?

**Slothful**  
Hey, TinyTurtle. Have you ever aquatic plants like seaweed in a little green pond?

# History of the Web



Sir Tim Berners-Lee invented the World Wide Web in 1989.

# Web Science

Web Science is the interdisciplinary study of the Web as an entity and phenomenon. It includes studies of the Web's properties, protocols, algorithms, and societal effects.

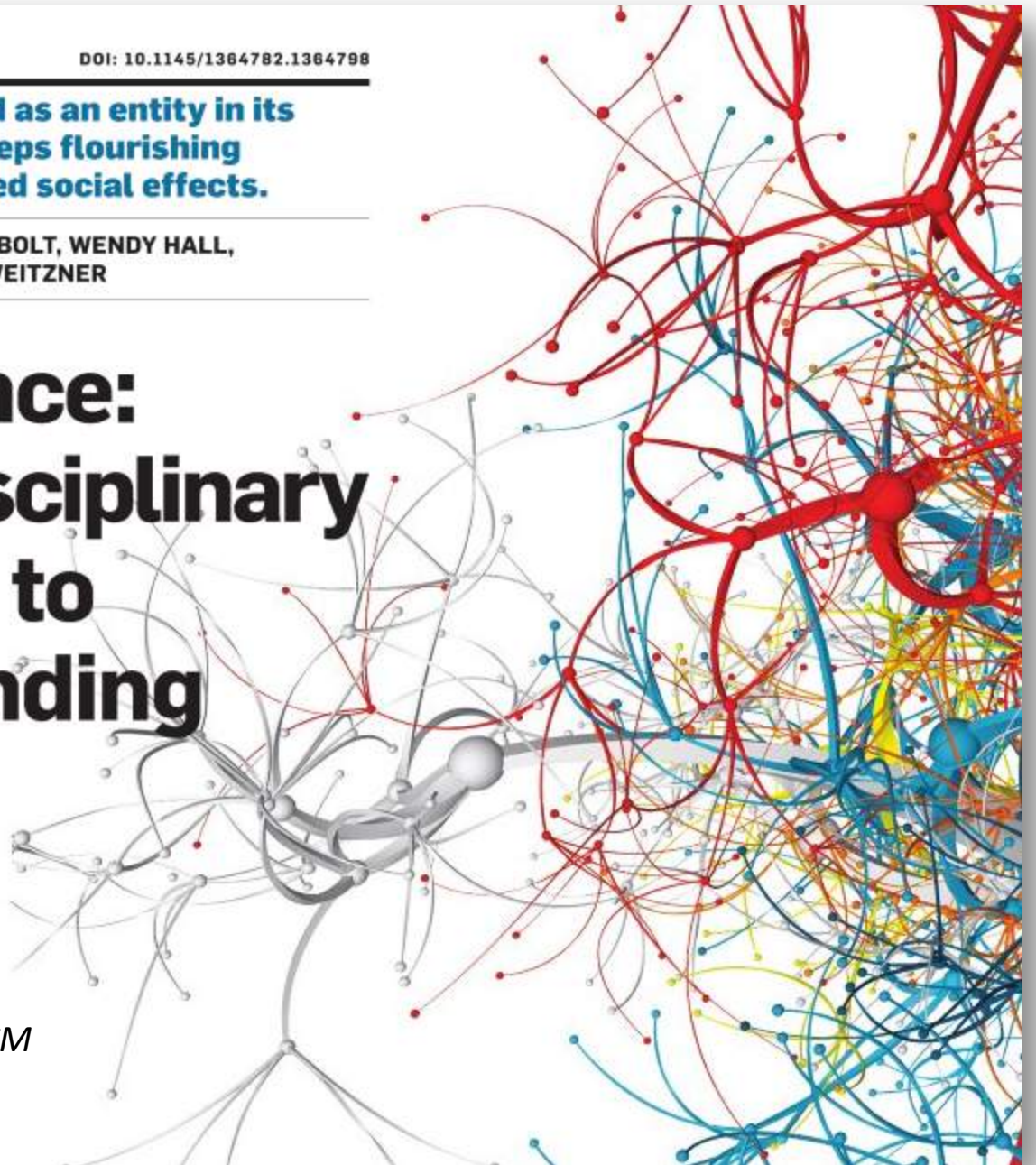


**The Web must be studied as an entity in its own right to ensure it keeps flourishing and prevent unanticipated social effects.**

BY JAMES HENDLER, NIGEL SHADBOLT, WENDY HALL,  
TIM BERNERS-LEE, AND DANIEL WEITZNER

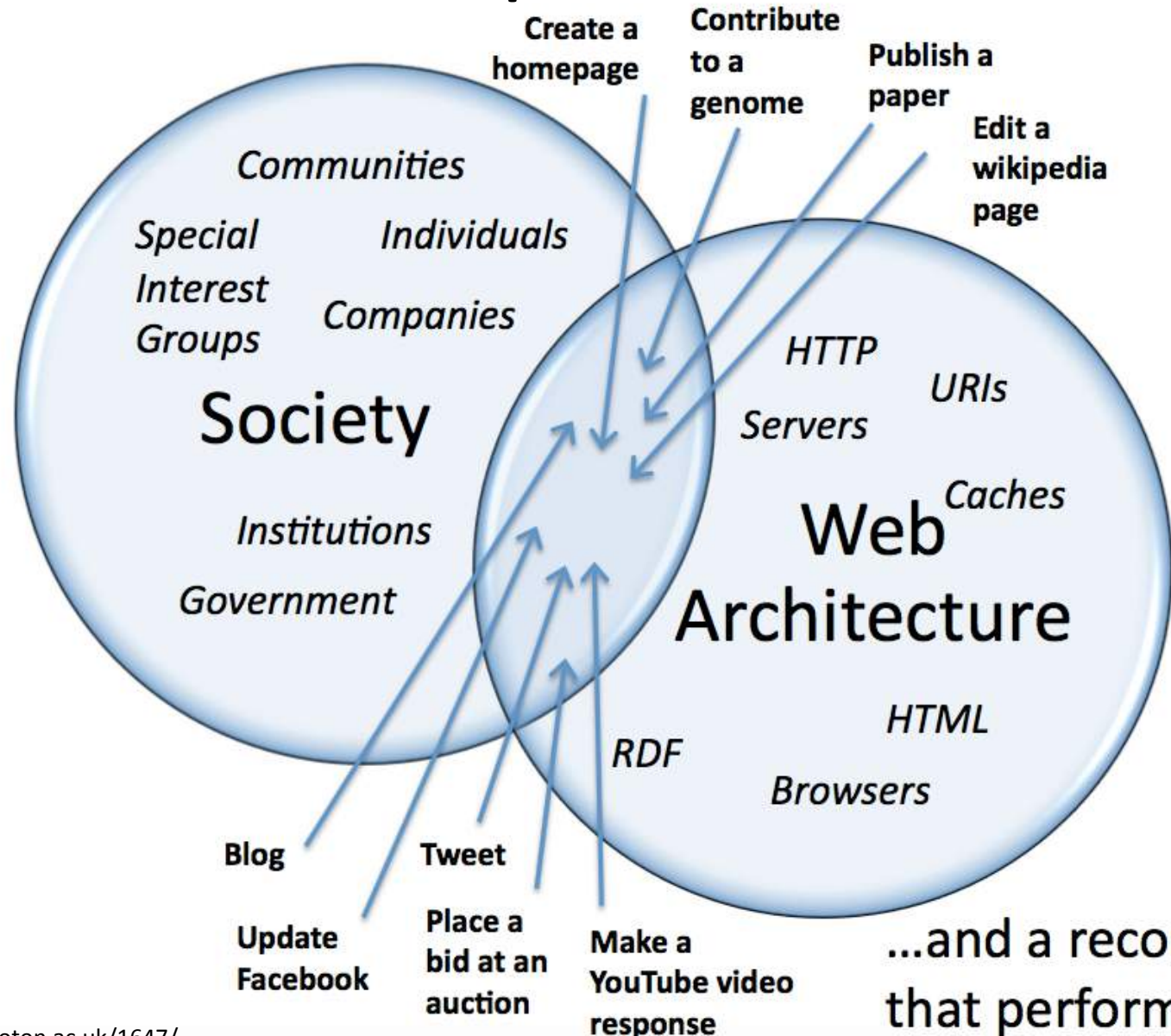
# Web Science: An Interdisciplinary Approach to Understanding the Web

*Communications of the ACM*  
July 2008



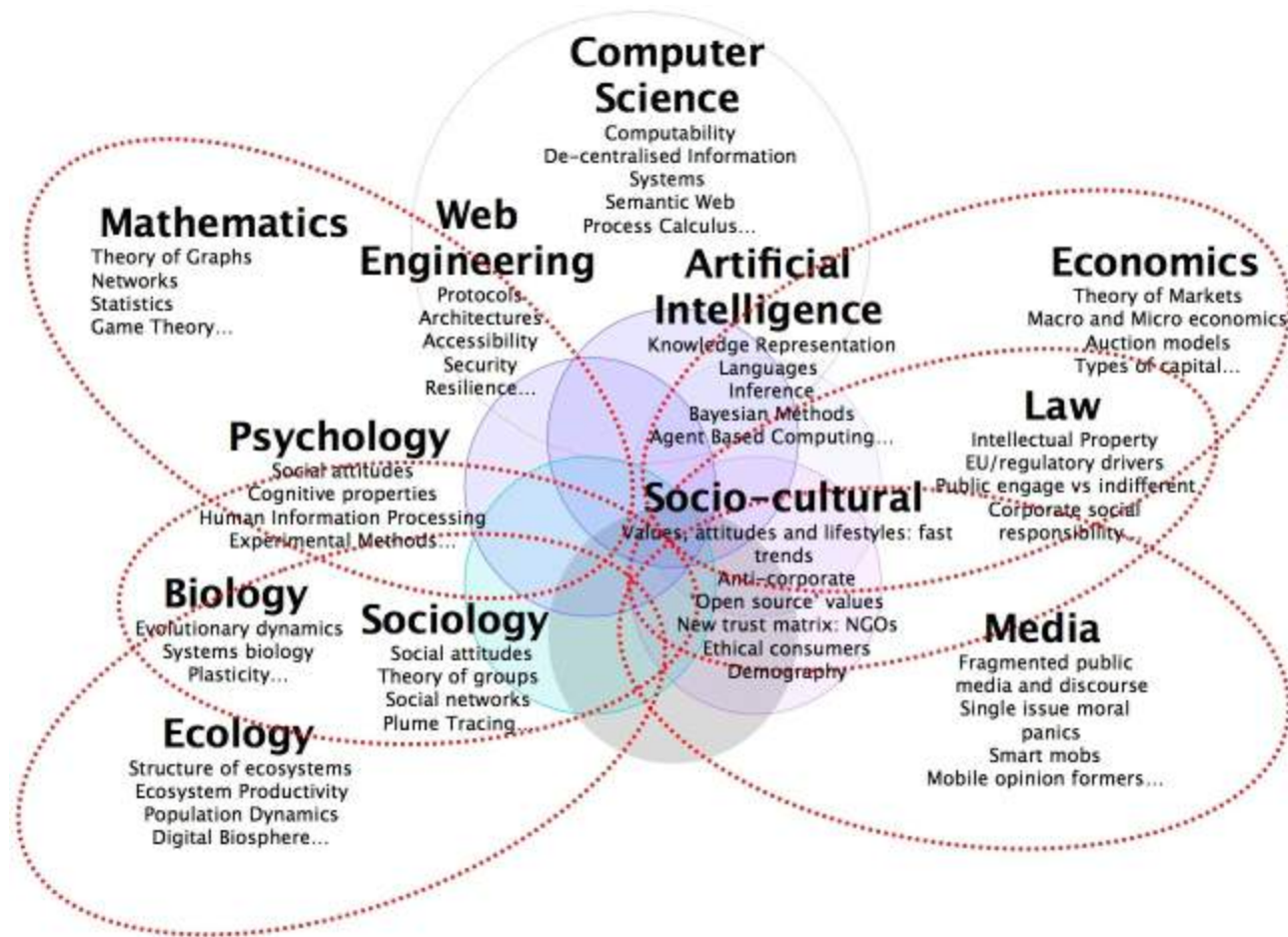
# The Web is a performance...

*The middle space represents the activity of individuals (potentially acting in concert) who create interlinked resources that both reflect and reinforce the interlinkedness of society and social (economic, legal, personal) interaction.*

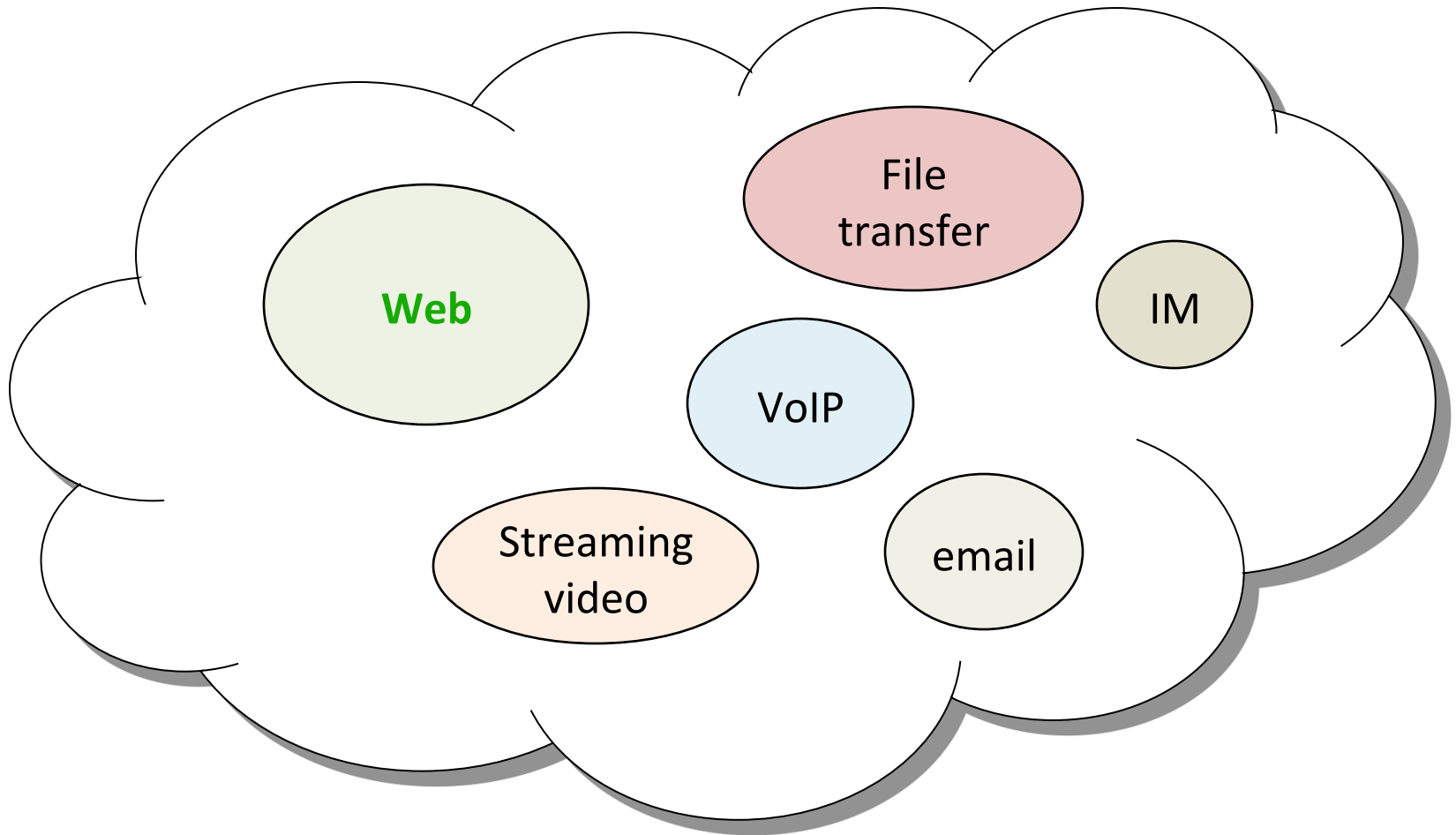




# Web Science is Interdisciplinary



# Internet != Web

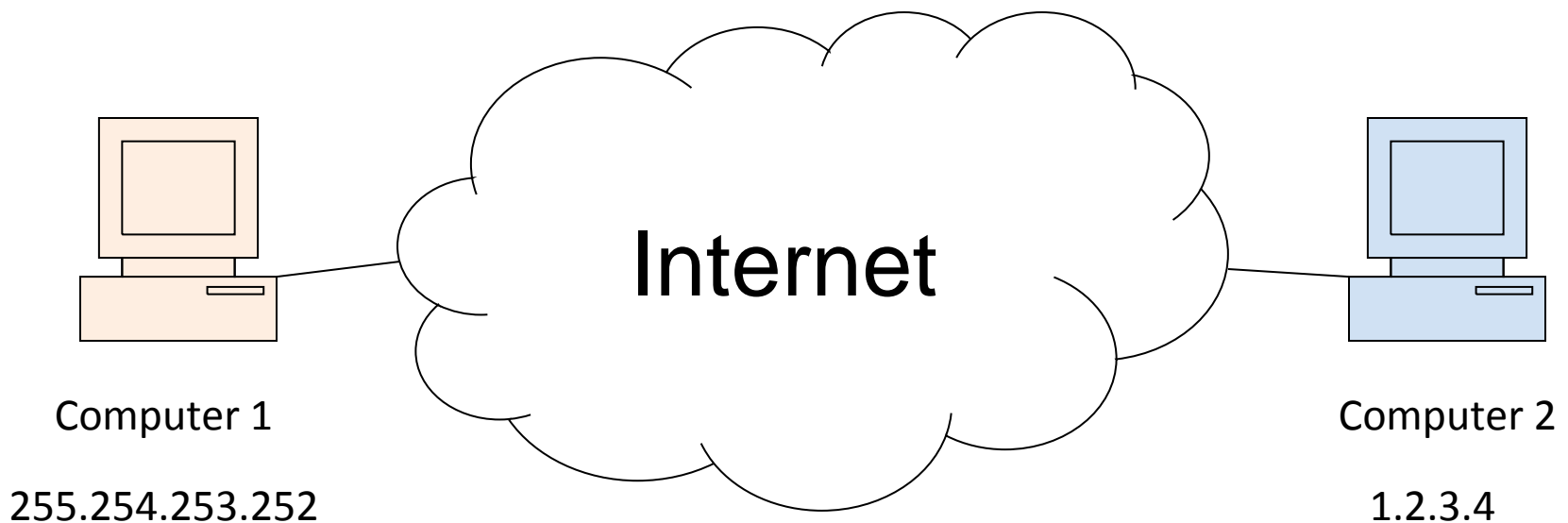


**The Internet**

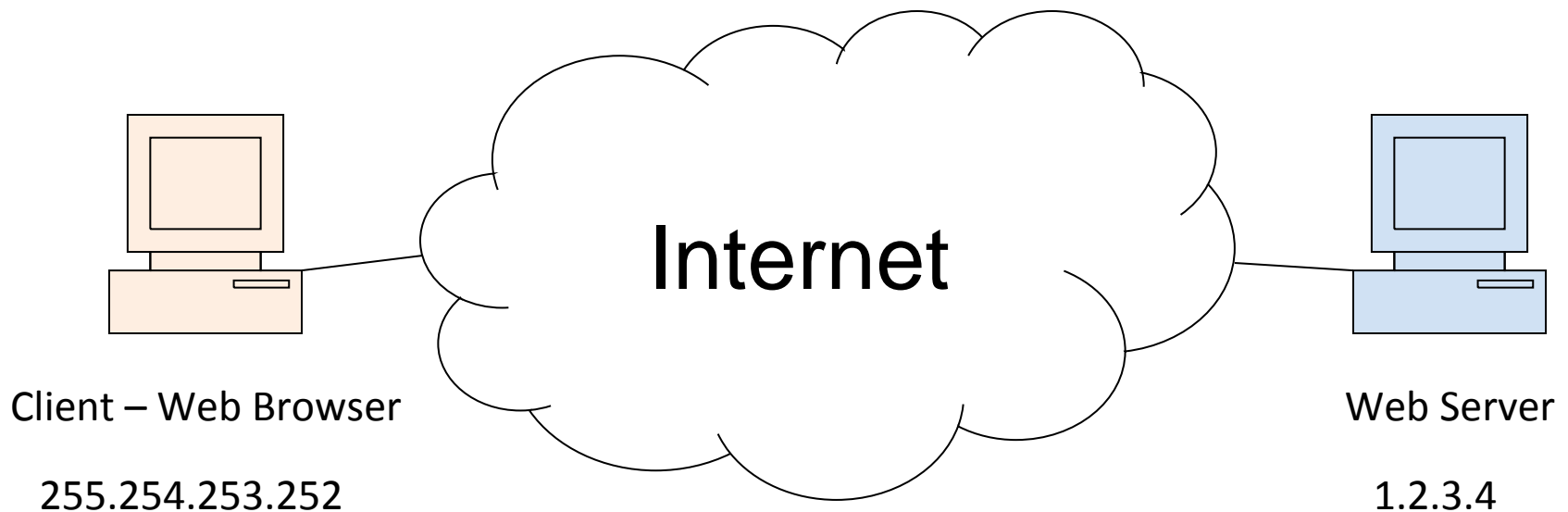


“The **Internet** is a global system of interconnected computer networks that use the standard Internet Protocol Suite (TCP/IP) to serve billions of users worldwide.”

<http://en.wikipedia.org/wiki/Internet>

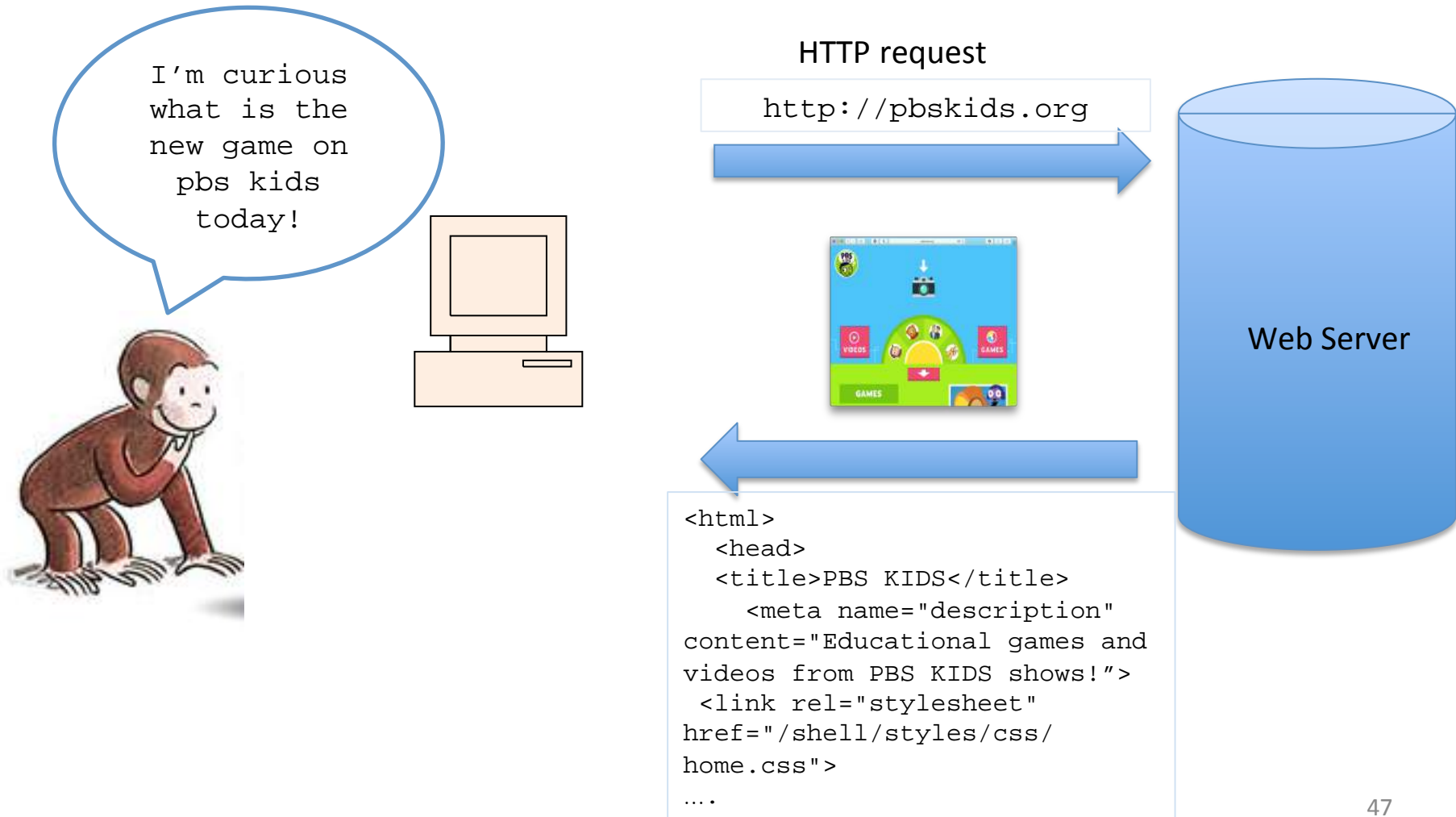


# The Web



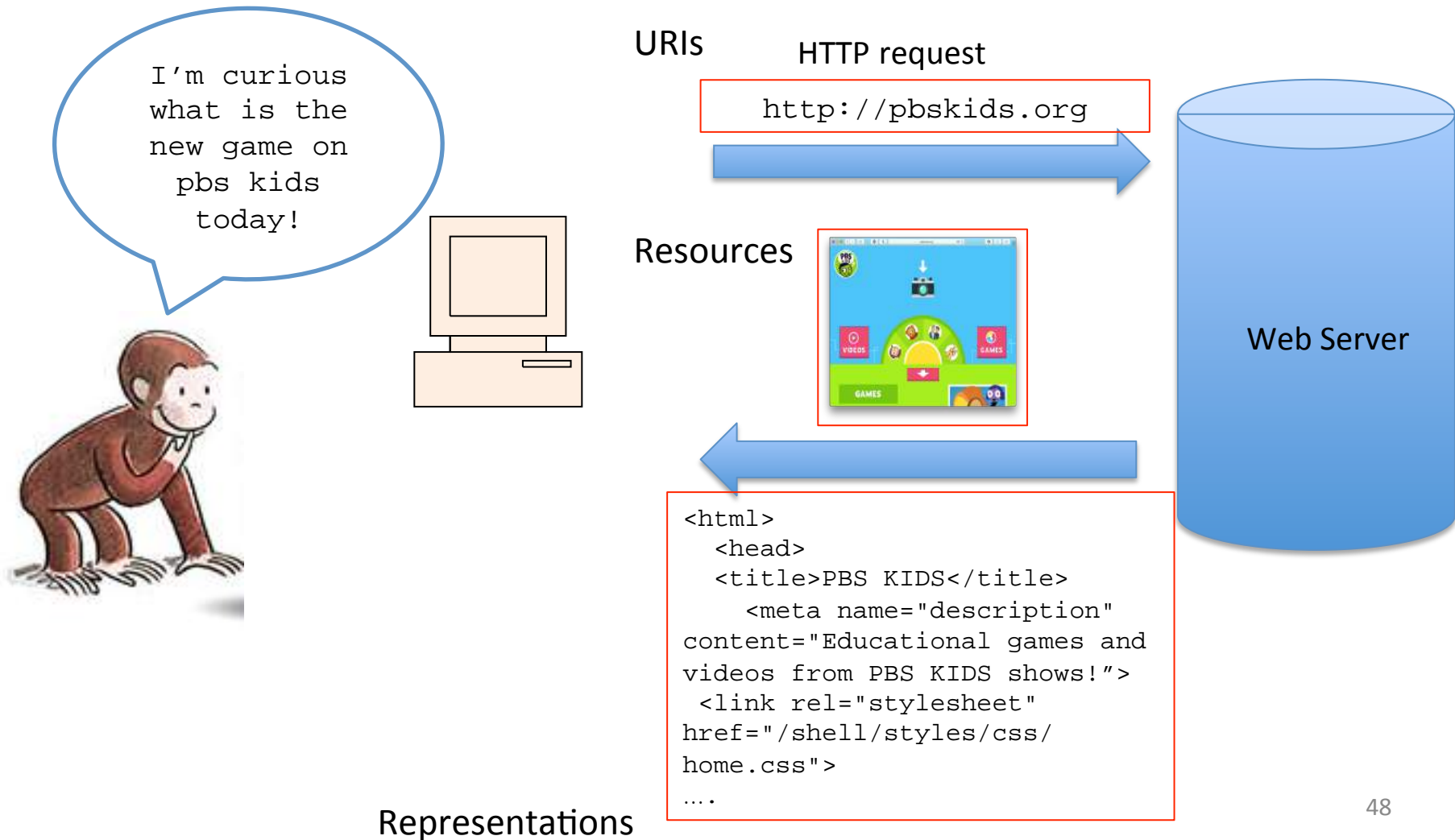
World Wide Web: The system of interlinked hypertext documents accessed over the Internet using the HTTP protocol.

# How the Web works

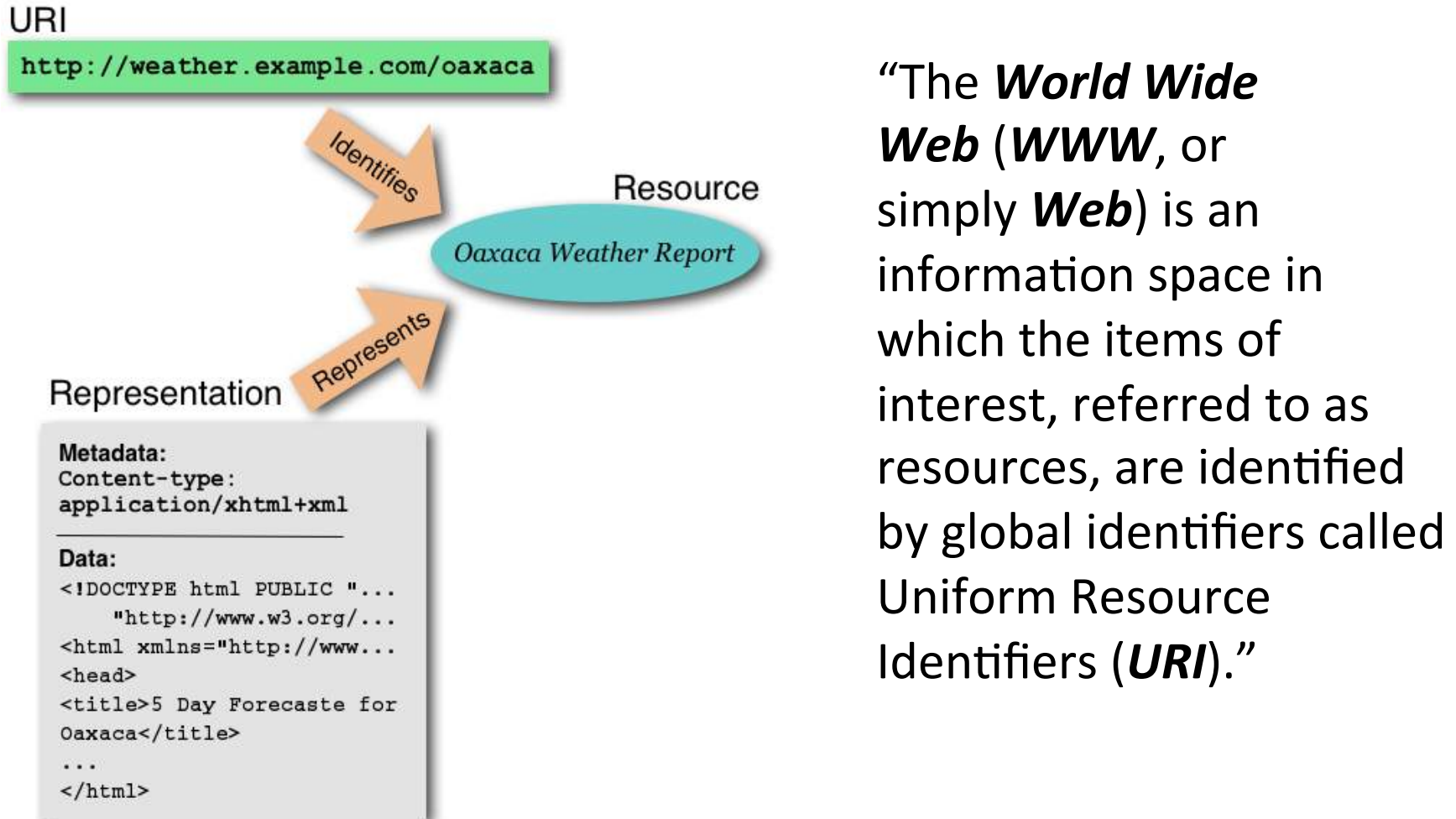




# How the Web works

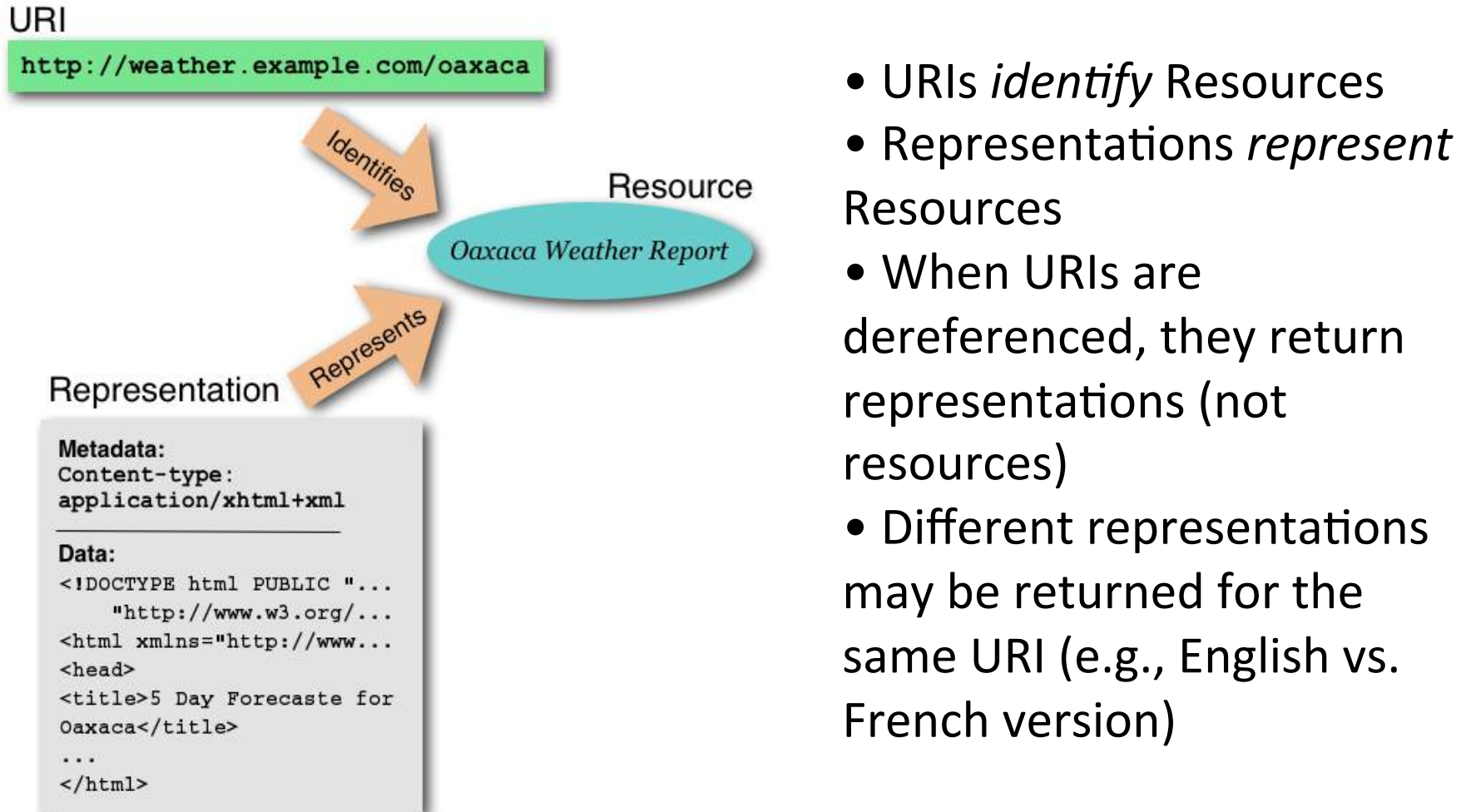


# Web Definition



“The **World Wide Web** (**WWW**, or simply **Web**) is an information space in which the items of interest, referred to as resources, are identified by global identifiers called Uniform Resource Identifiers (**URI**).”

# URIs, Resources, and Representations



- URIs *identify* Resources
- Representations *represent* Resources
- When URIs are dereferenced, they return representations (not resources)
- Different representations may be returned for the same URI (e.g., English vs. French version)

# Remember Three Things

## URIs

`http://pbskids.org`

## Representations

```
<html>
  <head>
    <title>PBS KIDS</title>
    <meta name="description"
content="Educational games and videos
from PBS KIDS shows!">

    <link rel="stylesheet" href="/shell/
styles/css/home.css">
... .
```

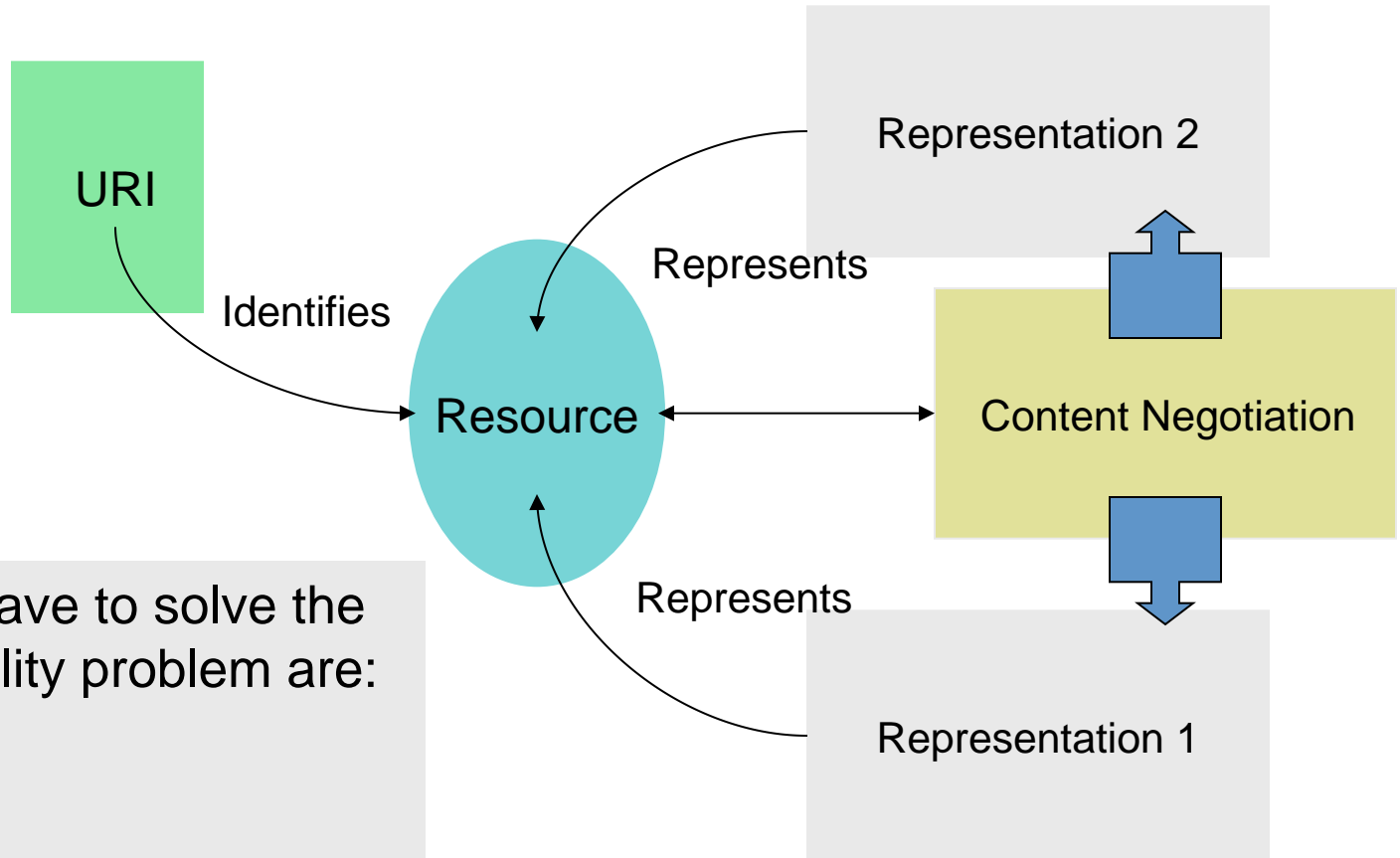
Identify

Represent

## Resources



# W3C Web Architecture



The tools we have to solve the interoperability problem are:

- URI
- Resource
- Representation

# Request methods

- GET
  - The GET method requests a representation of the specified resource.
- HEAD
  - The HEAD method asks for a response identical to that of a GET request, but without the response body.
- POST
  - The POST method requests that the server accept the entity enclosed in the request as a new subordinate of the web resource identified by the URI.

# Response Codes

- **1xx**: Informational - Request received, continuing process
- **2xx**: Success - The action was successfully received, understood, and accepted
- **3xx**: Redirection - Further action must be taken in order to complete the request
- **4xx**: Client Error - The request contains bad syntax or cannot be fulfilled
- **5xx**: Server Error - The server failed to fulfill an apparently valid request

# Response Codes

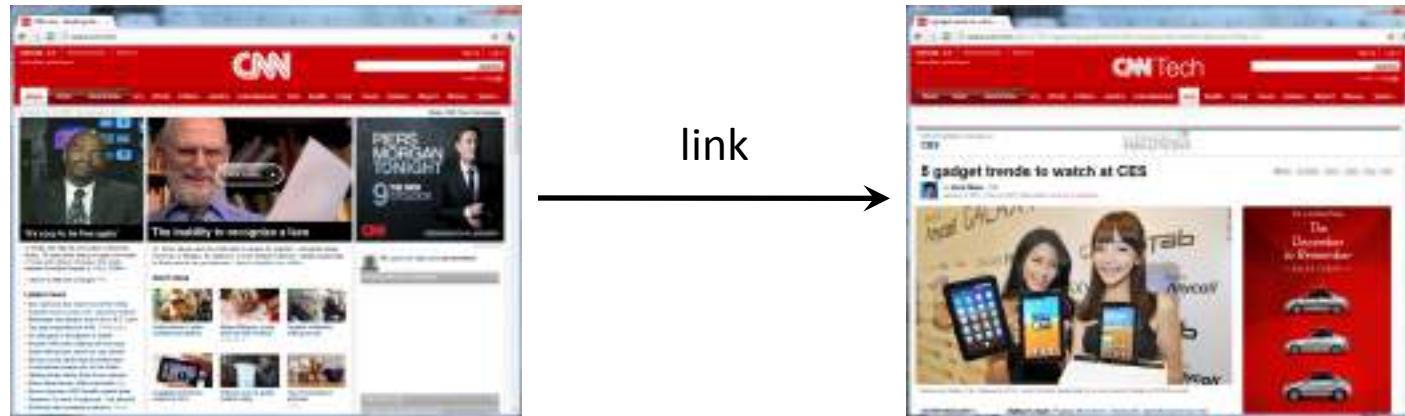
- **1xx**: Informational - Request received, continuing process
- **2xx**: Success - The action was successfully received, understood, and accepted
- **3xx**: Redirection - Further action must be taken in order to complete the request
- **4xx**: Client Error - The request contains bad syntax or cannot be fulfilled
- **5xx**: Server Error - The server failed to fulfill an apparently valid request



For more information about principles, constraints, good practice in dealing with the WWW, check out the W3C Recommendations:

<http://www.w3.org/TR/webarch/>

# How is the Web structured?



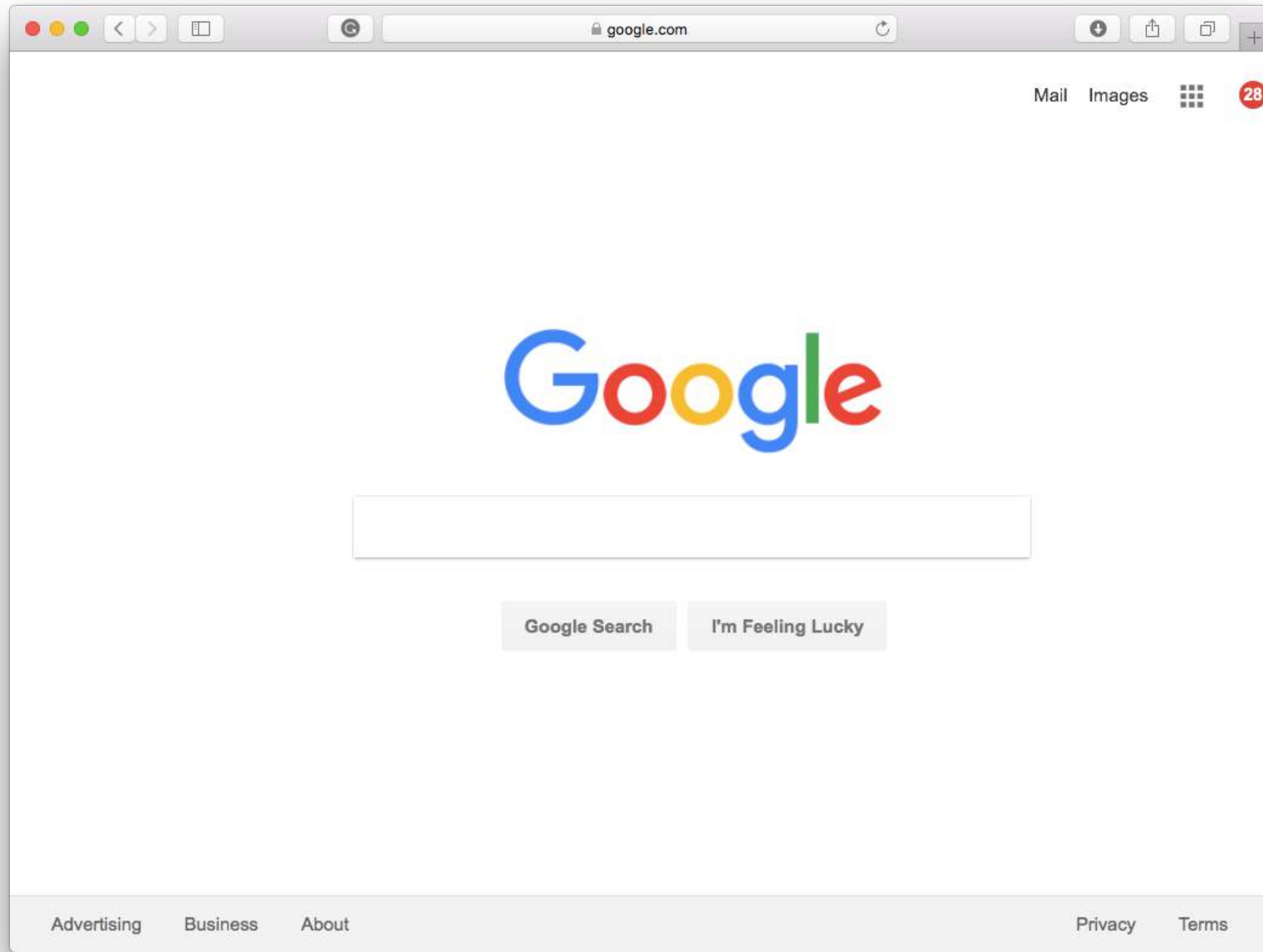
Graph Theory: Pages are nodes &  
links are directed edges

# The Evolution of the Web

Web 1.0 → Web 2.0 → Web 3.0

Read → Read-Write → Semantics

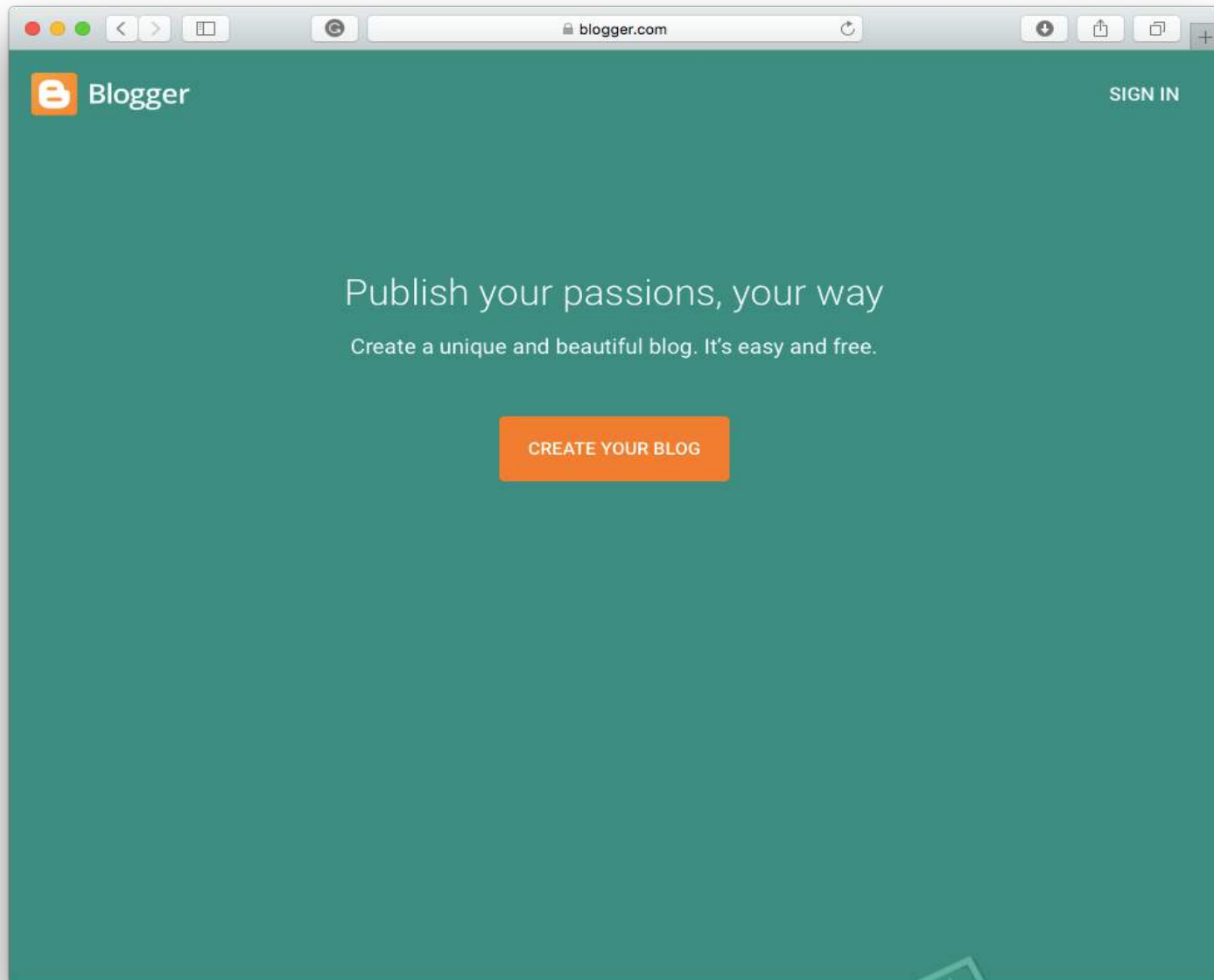
# Web 1



# Web 1



# Web 2



# Web 2





# Web 2

reddit.com

MY SUBREDDITS ▼ POPULAR - ALL - RANDOM | ASKREDDIT - FUNNY - PICS - TODAYILEARNED - NEWS - WORLDNEWS - GAMING - VIDEOS - AWW - GIFS - SHOWERT | MORE »

reddit

hot new rising controversial top gilded wiki pr Want to join? Log in or sign up in seconds. | English

trending subreddits r/DeepFriedMemes r/bonehurtingjuice r/instant\_regret r/surrealmemes r/dontdeadopeninside 96 comments

1 25.0k Fire Emblem got to real for me (i.imgur.com) submitted 4 hours ago by Luxyzinho to r/gaming 285 comments share report

2 23.4k Female chief in Malawi broke up 850 child marriages and sent girls back to school. Not all heroes wear capes. backstory (i.redd.it) submitted 4 hours ago by GallowBoob to r/pics 370 comments share report

3 23.2k Dude that "gets caught" is actually married to the lady in the gif [gifs] (np.reddit.com) submitted 4 hours ago by GFY\_EH to r/bestof 902 comments share report

4 9103 This newspaper headline is scarring VS Pie eating contest (i.imgur.com) submitted 4 hours ago by GallowBoob to r/CrappyDesign 176 comments share report

5 19.3k HIV/AIDS is no longer the leading cause of death in Africa (businessinsider.com) submitted 5 hours ago by TrumpDumper to r/UpliftingNews 707 comments share report

6 7243 Clocktower penthouse in Brooklyn [3600x2400] (i.redd.it) submitted 4 hours ago by jackalooz to r/RoomPorn 273 comments share report

7 25.2k I put my coffee cup in the microwave and it scorched right on the 'Do Not Microwave' warning (i.redd.it)

search

username password

☐ remember me reset password login

MOKSHA WELLNESS

Look like an expert

Strengthen your brand with a custom domain, business email, and website from Squarespace.

Submit a new link

Submit a new text post

daily reddit gold goal 66%

help support reddit



**yelp**

Find tacos, cheap dinner, Max's Near Fremont, CA, US

Restaurants Nightlife Home Services Write a Review Events Talk

## Sprouts Farmers Market ? Unclaimed

★★★★★ 24 reviews Details

Write a Review Add Photo

\$\$ · Grocery, Fruits & Veggies, Health Markets Edit

6399 Jarvis Ave  
Newark, CA 94560  
b/t Dumbarton Ct & Lido Blvd

Get Directions  
(510) 789-0361  
sprouts.com  
Send to your Phone

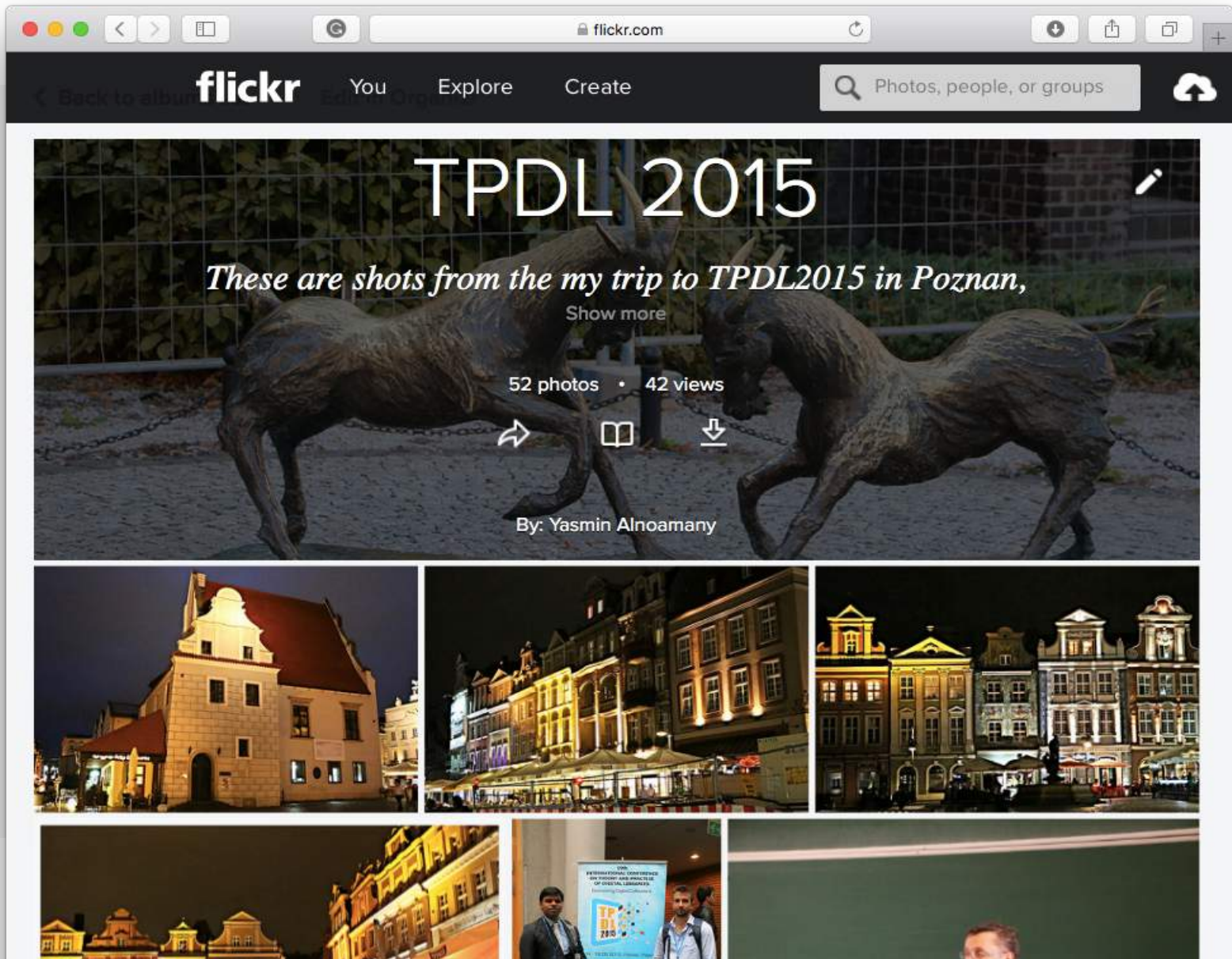
Photo of Sprouts Farmers Market - Newark, CA, United States

"This new location has better parking than the **Fremont location**, its very clean, easy to navigate and has a pretty decent selection of whatever you're looking for." in 4 reviews

"This is a great store-staff are friendly, much easier to park compare to the sprout on **Mowry.**" in 4 reviews

Today 7:00 am  
Price range Mo  
Work here?

# Web 2



# Web 2





# Web 2

The image is a screenshot of a web browser displaying the Twitter profile of Yasmina Anwar. The browser's address bar shows "Twitter, Inc.". The profile header features a circular profile picture of a woman in a red headscarf and a banner image of purple and yellow flowers. Below the header, statistics are listed: 2,190 Tweets, 405 Following, 441 Followers, 711 Likes, 3 Lists, and 1 Moments. An "Edit profile" button is on the right. The bio identifies her as a Postdoc at UC Berkeley, a Mom, Wife, Traveler, and Photographer, and mentions her involvement with ArabWIC and Systems. Her location is Norfolk, VA, and her website is cs.odu.edu/~yasmin. She joined in November 2010 and was born on July 23, 1985. A section for "254 Photos and videos" is partially visible at the bottom left. The main content area shows a tweet from Yasmina Anwar (@yasmina\_anwar) dated August 21, which includes a retweet of a tweet from PHD Comics (@PHDcomics) about an eclipse. Below this, a retweet by V Ikeshoji-Orlati (@vikeshojiorlati) is also visible.

Twitter, Inc.

Search Twitter

**Yasmina Anwar**  
@yasmina\_anwar

Postdoc @UCBerkeley, @WebSciDL,  
Mom, Wife, Traveler, Photographer,  
Working for a more inclusive tech  
community, @ArabWIC @Systems

Norfolk, VA

cs.odu.edu/~yasmin

Joined November 2010

Born on July 23, 1985

254 Photos and videos

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

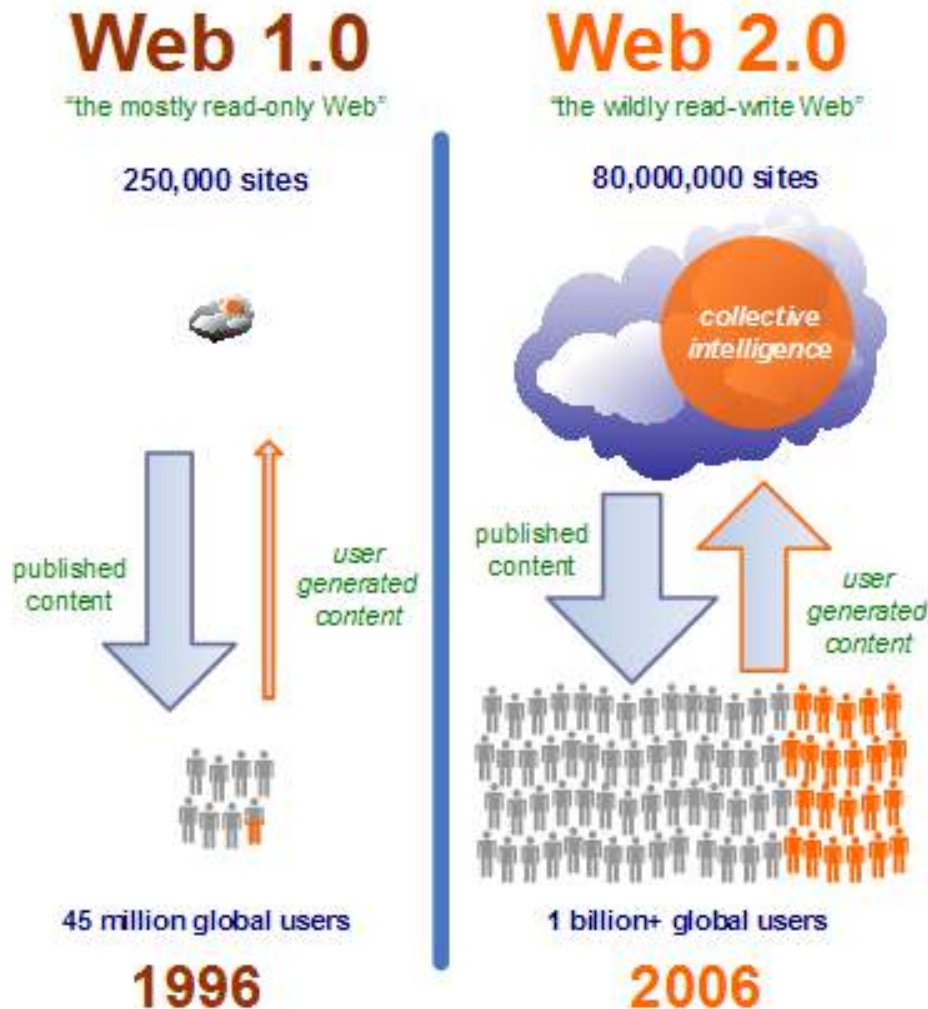
**Yasmina Anwar** @yasmina\_anwar · Aug 21  
:D

**PHD Comics** @PHDcomics  
#Eclipse phdcomics.com/comics.php?f=1...

You Retweeted

**V Ikeshoji-Orlati** @vikeshojiorlati · Aug 21  
#Eclipse2017 maps to sate your data viz appetite for the day.

# Web 1.0 versus Web 2.0

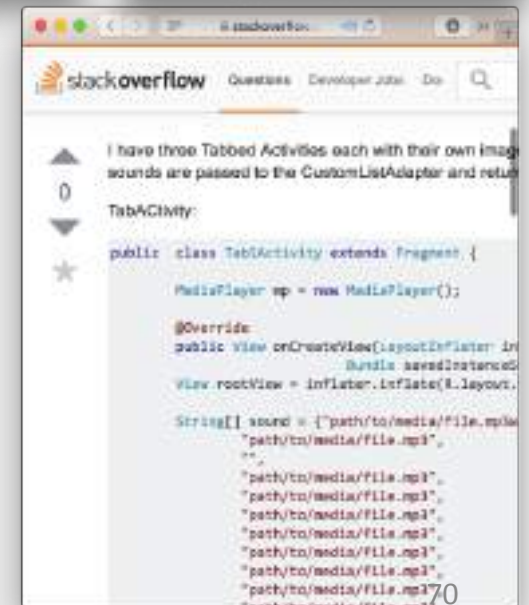
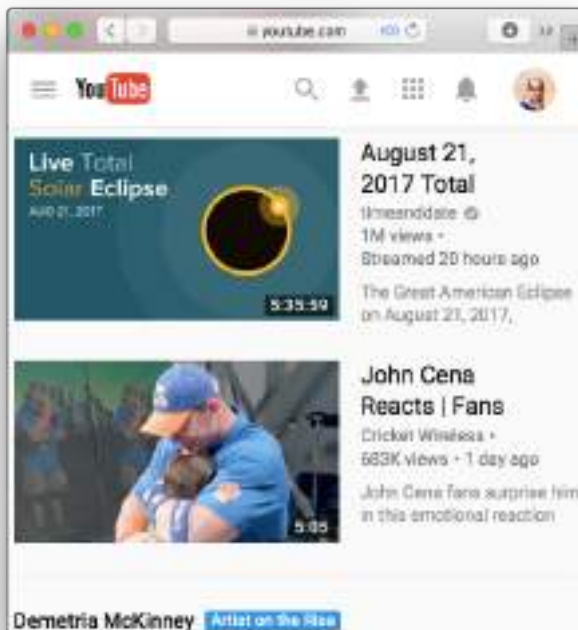
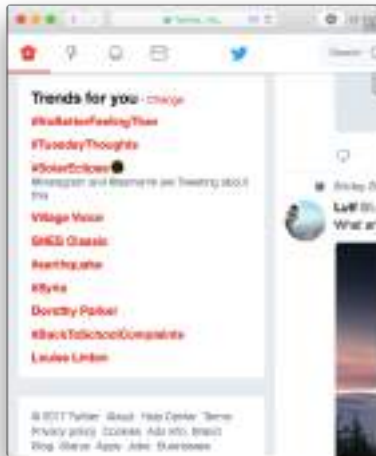


# Web 3

- From interlinks between documents to interlinks between things



# Examples of web data



# Web data format

- Delimited Text
  - tabbed delimited
  - comma delimited (CSV)
- Extensible Markup Language (XML)
  - looks a bit like HTML
  - user-defined tags to identify data
- JavaScript Object Notation (JSON)
  - collection of name/value pairs
  - smaller than XML
  - easier to parse



# Examples

## Table

Date	temp
20090101	26
20090102	34
20090103	27

## CSV

```
Date,temp
20090101,26
20090102,34
20090103,27
```

## JSON

```
{ "observations":
[
  { "date": "20090101", "max_temp": 26 },
  { "date": "20090102", "max_temp": 34 },
  { "date": "20090103", "max_temp": 27 }
]
}
```

## XML

```
<weather_data>
  <observation>
    <date>20090101</date>
    <max_temp>26</max_temp>
  </observation>
  <observation>
    <date>20090102</date>
    <max_temp>34</max_temp>
  </observation>
  <observation>
    <date>20090103</date>
    <max_temp>27</max_temp>
  </observation>
</weather_data>
```

# How to convert between data formats?

- Mr. Data Converter
  - <http://shancarter.github.io/mr-data-converter/>
  - developed by a graphics editor at *The New York Times*
  - input: CSV or tab-delimited data
  - output: HTML table, JSON, MySQL, Python, PHP, XML, ...
- Other tools
  - Data format converter: <http://www.convertcsv.com/xml-to-csv.htm>
  - search Google for "csv to json", "csv to xml", "xml to json"

# More About Web Science

- Video: Nigel Shadbolt on Web Science (2008)  
<http://webscience.org/professor-nigel-shadbolt-explains-web-science/>
- Slides: "What is Web Science?" by Carr, Pope, Hall, Shadbolt (2008)  
<http://www.slideshare.net/lescarr/what-is-web-science>

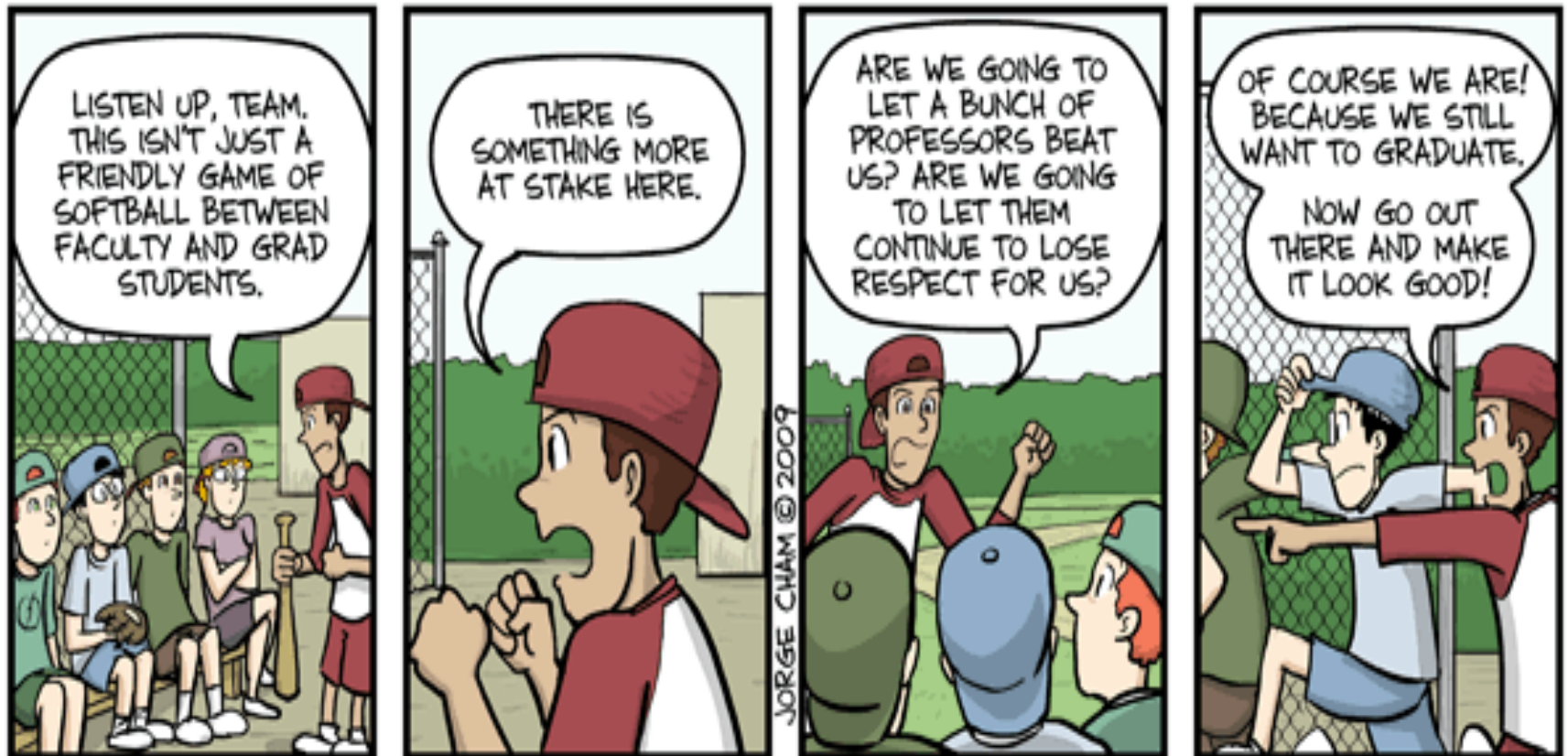
# Visualization

# “A picture is worth a thousand words”



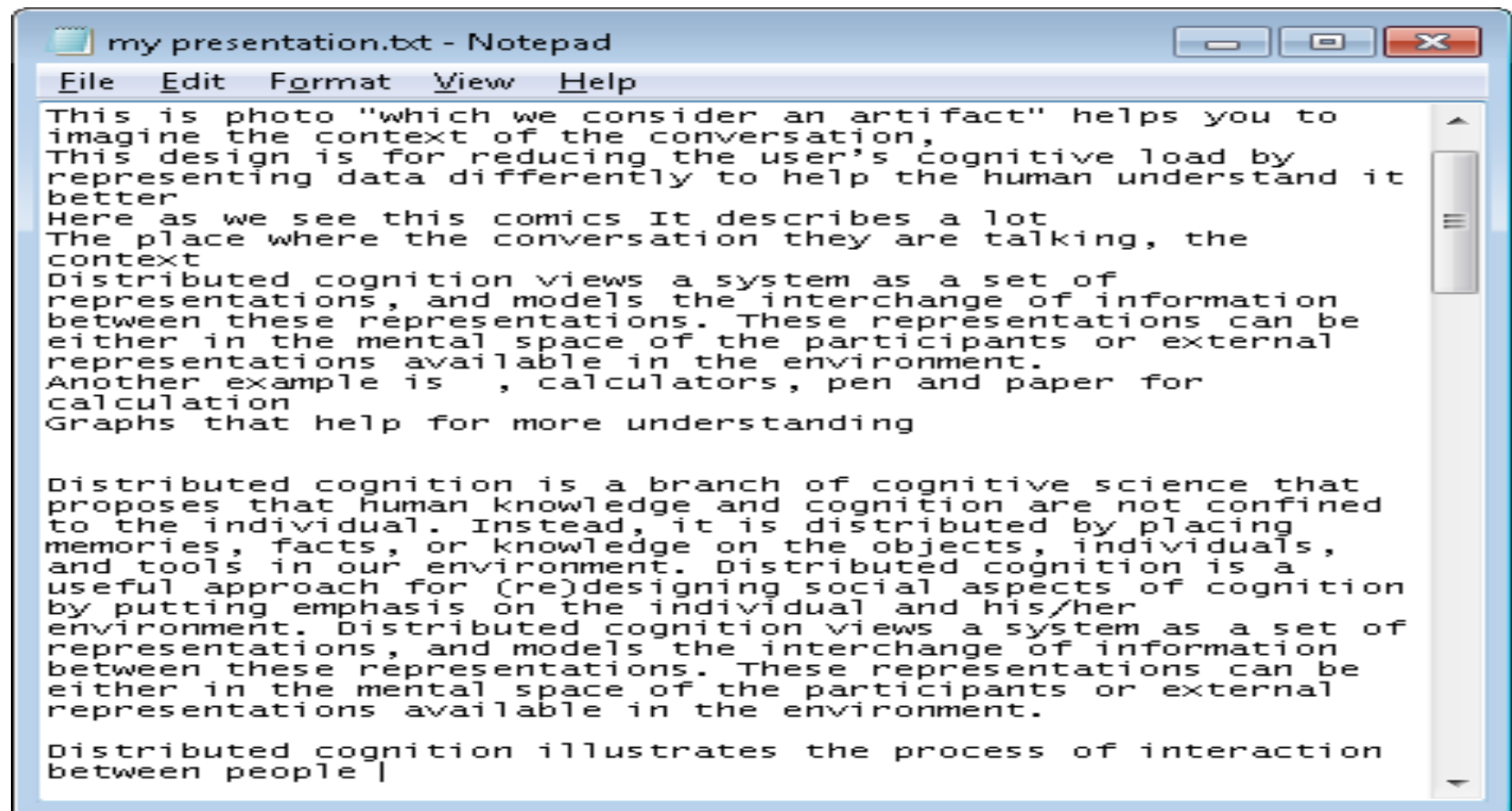
Creamy, delicious, yummy,  
fudge ice cream, smooth,  
chocolate-chip mint ice  
cream, strawberry ice  
cream with real chunks of  
strawberry, colored sugar  
sprinkles, waffle sugar  
cone, sweet, wonderful,  
tastes great, cold, nice to  
eat, dessert, good, yummy  
toppings, chocolate  
sprinkles, comforting,  
good, fun, dripping, terrific,

# Example



[WWW.PHDCOMICS.COM](http://WWW.PHDCOMICS.COM)

# Example



# How Many 3s?

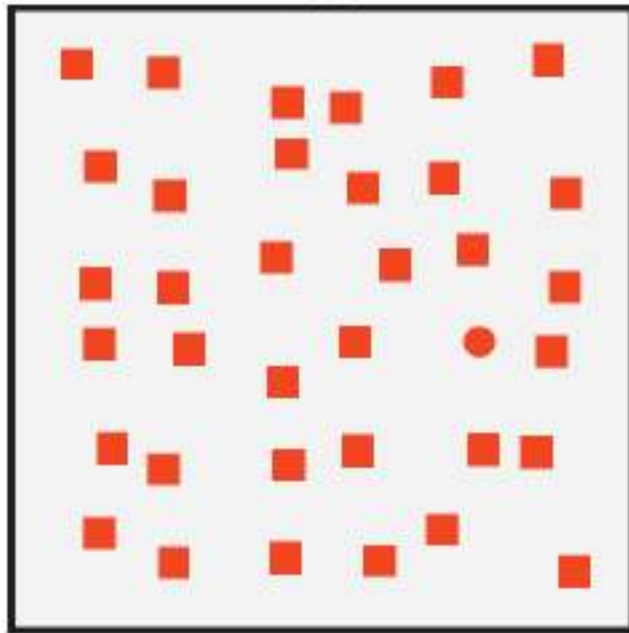
128176875613897654698450698560498282676  
298098584582245098564589450984509809435  
859091030209905959595772564675050678904  
567884578980982167765487636490856091294  
9686



# How Many 3s?

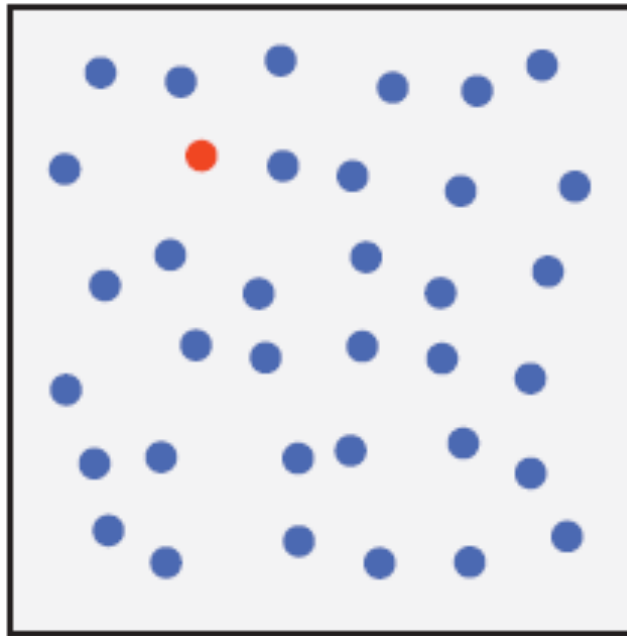
12817687561**3**897654698450698560498282676  
2980985845822450985645894509845098094**3**5  
8590910**3**0209905959595772564675050678904  
5678845789809821677654876**3**6490856091294  
9686

# Can you find the red dot?



Source: <https://www.csc2.ncsu.edu/faculty/healey/PP/index.html>

# Can you find the red dot?



Source: <https://www.csc2.ncsu.edu/faculty/healey/PP/index.html>

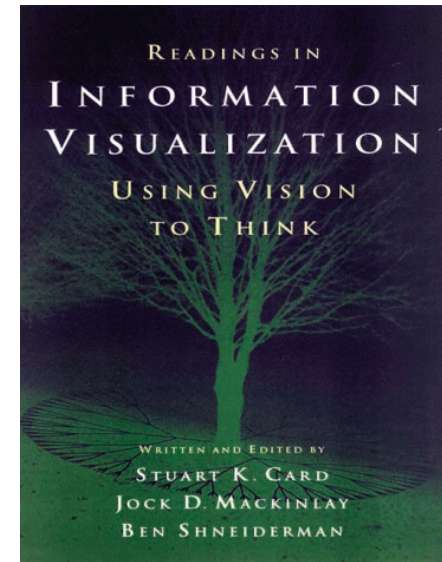
# Visualization Definition

“The use of computer-supported, interactive, visual representations of abstract data to amplify cognition and generate insight.” [Card, Mackinlay, & Shneiderman 1999]

“... finding the artificial memory that best supports our natural means of perception.” [Bertin 1967]

"The purpose of visualization is insight, not pictures." [Ben Shneiderman]

“Computer-based visualization systems provide visual representations of datasets designed to help people carry out tasks more effectively.” [Munzner 2014]



# Why human in the loop?

- Computer-based visualization systems provide visual representations of datasets designed to help **people** carry out tasks more effectively
  - Visualization is suitable when there is a need to augment human capabilities rather than replace people with computational decision-making methods
  - don't need vis when fully automatic solution exists and is trusted
  - Humans must validate and verify results before deployment of an algorithm
  - many analysis problems ill-specified
    - don't know exactly what questions to ask in advance

# Why computer in the loop?

- **Computer-based** visualization systems provide visual representations of datasets designed to help people carry out tasks more effectively
  - beyond human patience: scale to large datasets, support interactivity
  - Time limit
  - Datasets can be dynamic

# Why depend on vision?

- Computer-based visualization systems provide **visual** representations of datasets designed to help people carry out tasks more effectively
  - human visual system is high-bandwidth channel to brain
  - overview possible due to background processing
    - subjective experience of seeing everything simultaneously
    - significant processing occurs in parallel and pre-attentively

# Why use an external representation?

- Computer-based visualization systems provide visual **representations** of datasets designed to help people carry out tasks more effectively
  - Visualization allows people to offload internal cognition and memory usage to the perceptual system, using carefully designed images as a form of external representations
  - external representation: replace cognition with perception



# Why show the data in detail?

- Seeing the dataset structure in detail is better than seeing only a brief summary of it
- Statistical characterization of datasets is a very powerful approach, but has the intrinsic limitation of losing information through summarization

# Why show the data in detail?

	Set I		Set II		Set III		Set IV	
	x	y	x	y	x	y	x	y
	10	8.04	10	9.14	10	7.46	8	6.58
	8	6.95	8	8.14	8	6.77	8	5.76
	13	7.58	13	8.74	13	12.74	8	7.71
	9	8.81	9	8.77	9	7.11	8	8.84
	11	8.33	11	9.26	11	7.81	8	8.47
	14	9.96	14	8.1	14	8.84	8	7.04
	6	7.24	6	6.13	6	6.08	8	5.25
	4	4.26	4	3.1	4	5.39	19	12.5
	12	10.84	12	9.13	12	8.15	8	5.56
	7	4.82	7	7.26	7	6.42	8	7.91
	5	5.68	5	4.74	5	5.73	8	6.89
Mean	9	7.5	9	7.5	9	7.5	9	7.5
Variance	10	3.75	10	3.75	10	3.75	10	3.75

Anscombe's Quartet: Raw Data

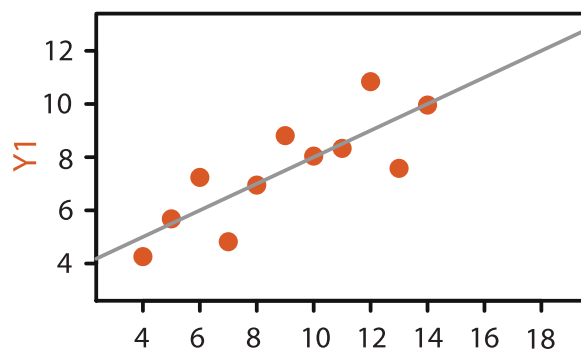
# Why show the data in detail?

## Anscombe's Quartet

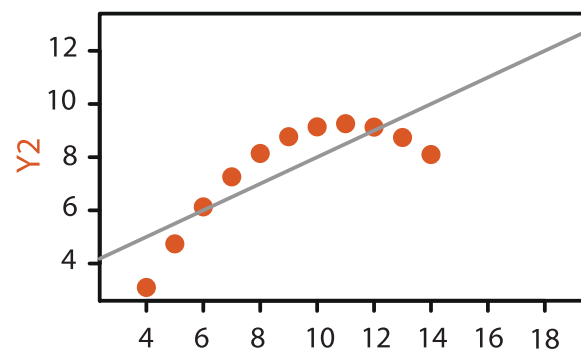
### Identical statistics

x mean	9
x variance	10
y mean	7.5
y variance	3.75
x/y correlation	0.816

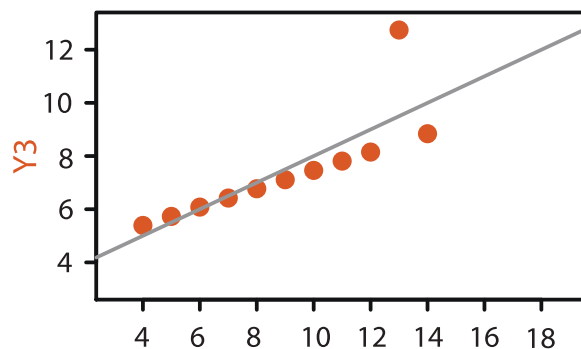
- summaries lose information
  - confirm expected and find unexpected patterns
  - assess validity of statistical model



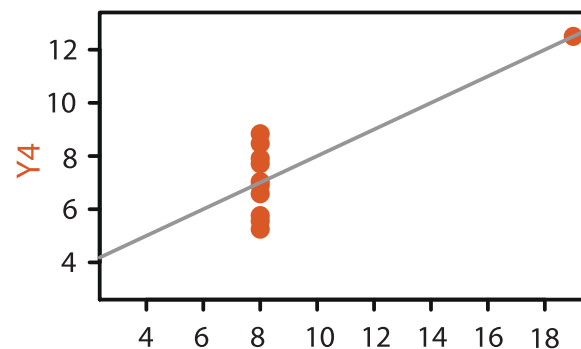
X1



X2



X3



X4

# Is Visualization New?

- Egyptian Paintings



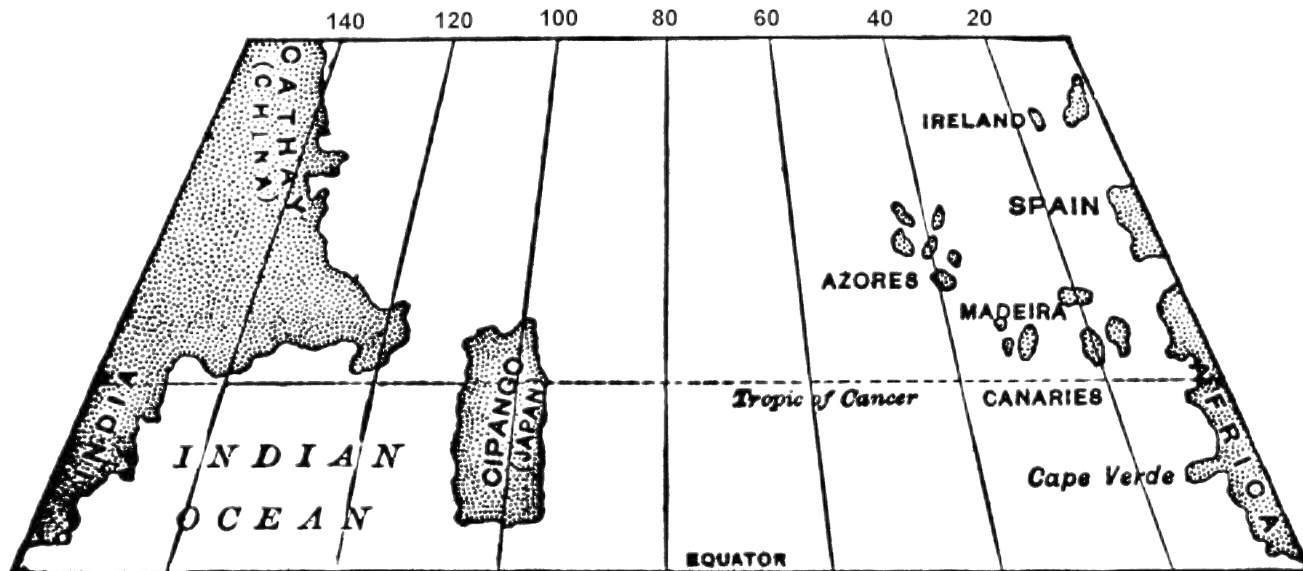
# Is Visualization New?

- Egyptian Paintings
- Cave guys, hunting



# Is Visualization New?

- Egyptian Paintings
- Cave guys, prehistory, hunting
- Directions and maps

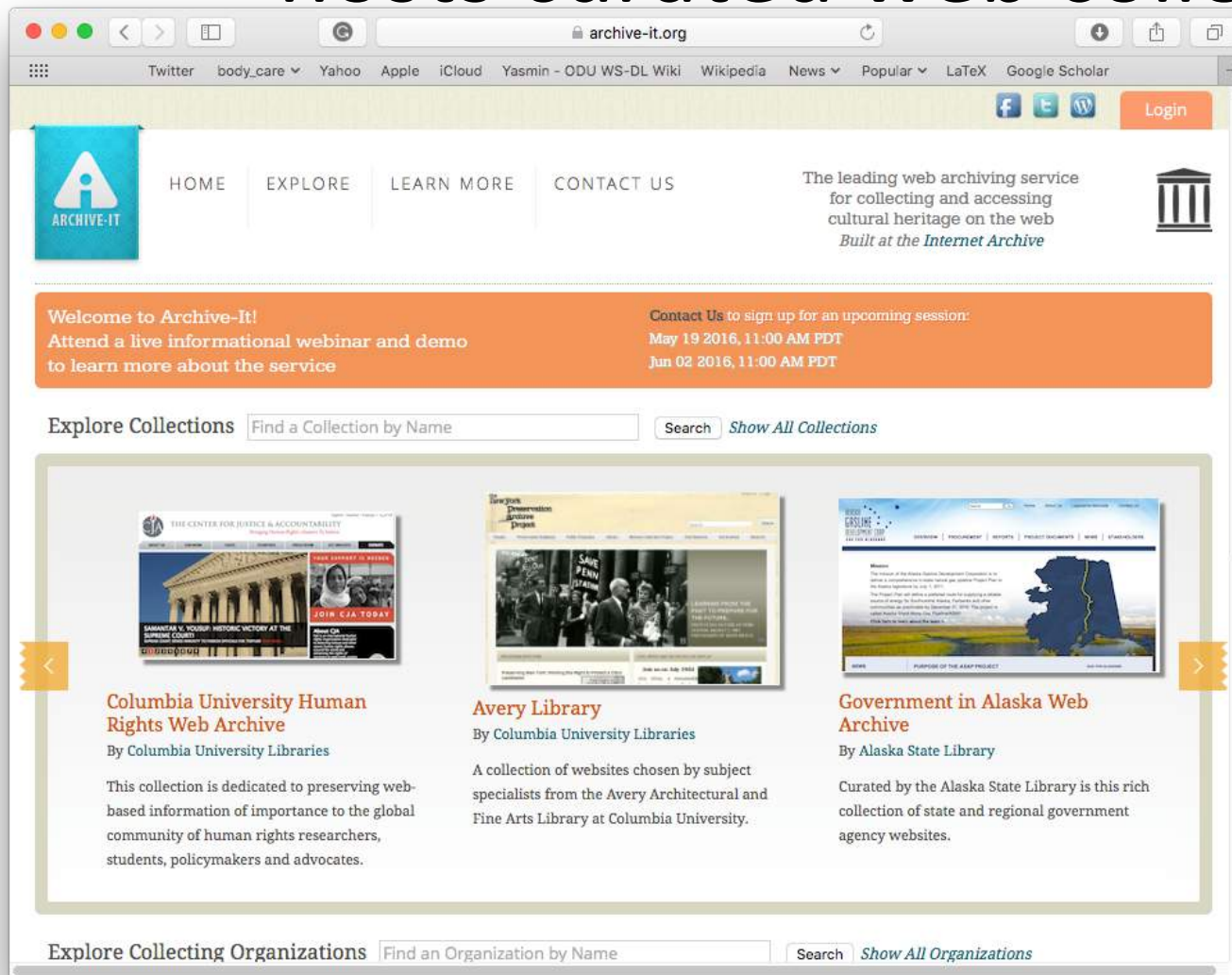


# Why Web Data Visualization?

# Understanding the past



# Archive-It, a subscription-based service, hosts curated web collections

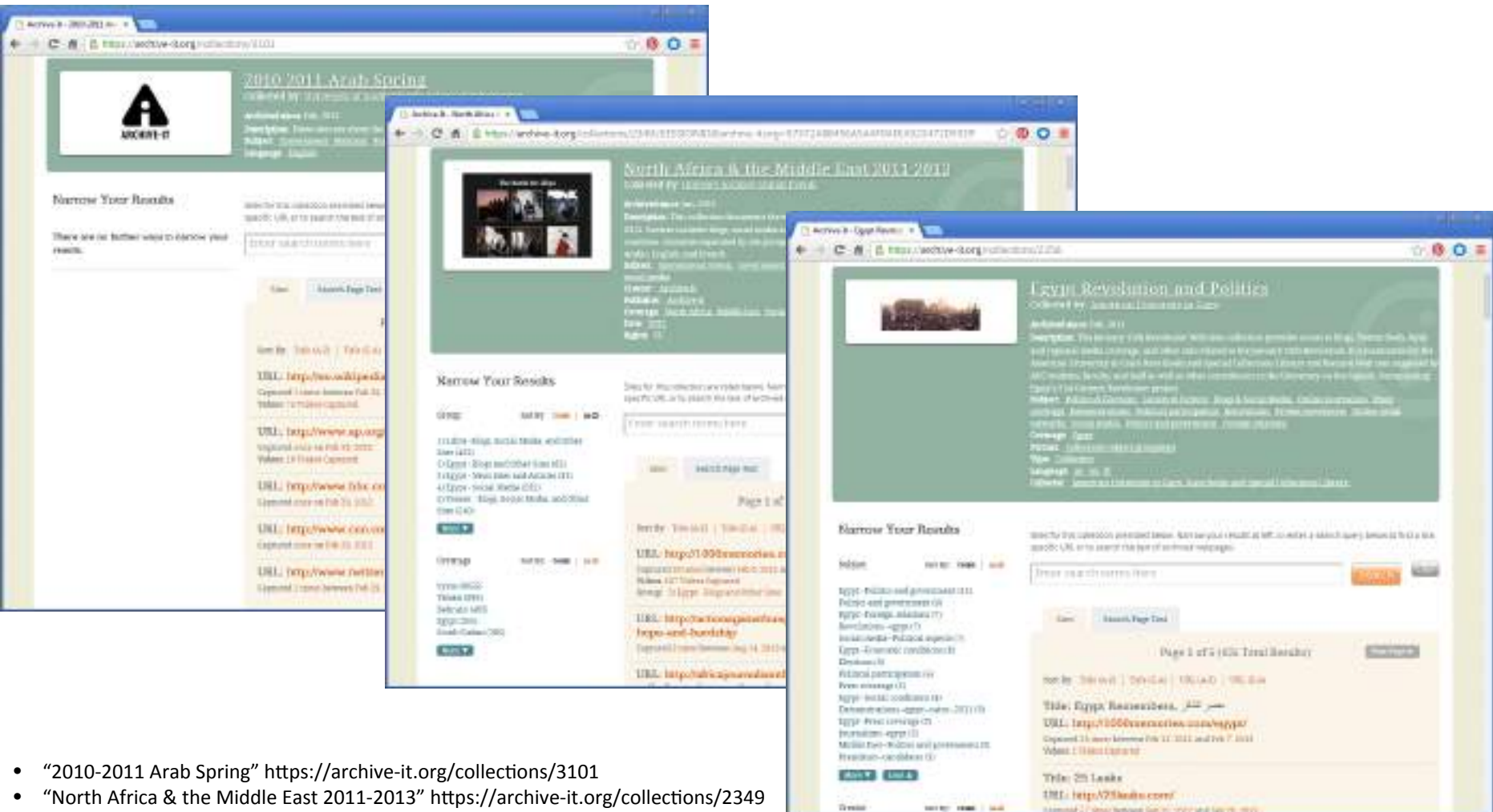


> 3,500  
collections

~340  
institutions

> 10B archived  
pages

# There is more than one collection about the Egyptian Revolution



The image displays three overlapping screenshots of the Archive-It website, illustrating different digital collections related to the Egyptian Revolution. The top-left screenshot shows the '2010-2011 Arab Spring' collection, featuring a search bar and a list of results. The middle screenshot shows the 'North Africa & the Middle East 2011-2013' collection, with a search bar and a list of results. The bottom-right screenshot shows the 'Egypt Revolution and Politics' collection, with a search bar and a list of results. Each screenshot includes a search bar, a list of results, and a 'Narrow Your Results' section.

- “2010-2011 Arab Spring” <https://archive-it.org/collections/3101>
- “North Africa & the Middle East 2011-2013” <https://archive-it.org/collections/2349>
- “Egypt Revolution and Politics” <https://archive-it.org/collections/2358>

# Current browsing and searching services for the “Egypt Revolution and Politics” collection

Archive-It - Egypt Revolution and Politics

https://archive-it.org/collections/2358

## Egypt Revolution and Politics

Collected by: American University in Cairo

Archived since: Feb. 2011

Description: The January 25th Revolution Web site collection provides access to blogs, Twitter feeds, local and regional media coverage, and other sites related to the January 25th Revolution. It is maintained by the American University in Cairo Rare Books and Special Collections Library and Reserves. Web sites suggested by AUC students, faculty, and staff as well as other contributors to the University on the Square, documenting Egypt's 21st Century Revolution project.

Subject: Politics & Government, Society & Culture, Blogs & Social Media, Online Journalism - Press coverage, Demonstrations, Political participation, Revolutions, Protest movements, Online social networks, Social media, Politics and government, Foreign relations

Coverage: Egypt

Format: Collections (object groupings)

Type: Collection

Language: ar, en, fr

Collector: American University in Cairo, Rare Books and Special Collections Library

### Narrow Your Results

Subject:  Sort By: Count |

Egypt-Politics and government (11)  
Politics and government (9)  
Egypt-Foreign relations (7)  
Revolutions-Egypt (7)  
Social media-Political aspects (7)  
Egypt-Economic conditions (6)  
Elections (5)  
Political participation (5)  
Press coverage (5)  
Egypt-Social conditions (4)  
Demonstrations-Egypt-Cairo-2011 (3)  
Egypt-Press coverage (3)  
Journalism-Egypt (3)  
Middle East-Politics and government (2)  
Presidents-Candidates (2)

Start for this collection are listed below. Narrow your results at left, or enter a search query below to find a site, specific URL, or to search the text of archived webpages.

Enter search terms here

Page 1 of 5 (435 Total Results)

Sort By: Title (A-Z) | Title (Z-A) | URL (A-Z) | URL (Z-A)

**Title: Egypt Remembers, مصر تذكّر**

Created: August 1, 2011  
Captured 53 times between Feb 11, 2011 and Feb 1, 2012  
Videos: 2 Videos Captured

# Current browsing and searching services for the “Egypt Revolution

Internet Archive Wayback Machine

https://wayback.archive-it.org/2358/\*http://1000memories.com/egypt/

Egypt Revolution and Politics Web Archive (American University in Cairo)

INTERNET ARCHIVE Wayback Machine

Enter Web Address:  All

Searched for <http://1000memories.com/egypt/> 26 Results [RSS](#) [Metadata](#)

[Look up URL](#) in general Internet Archive web collection [Proxy Mode Help](#)

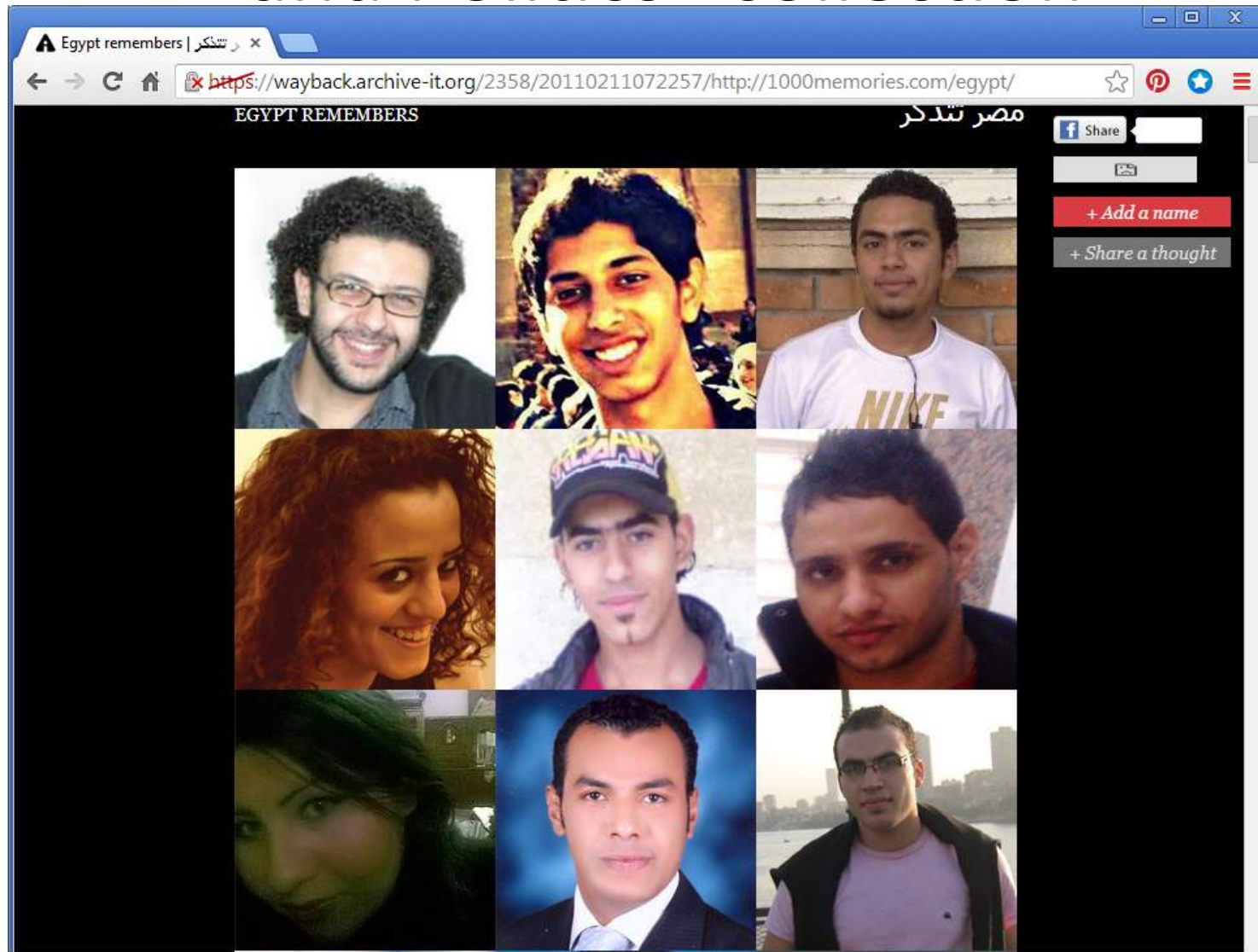
\* denotes when page was updated

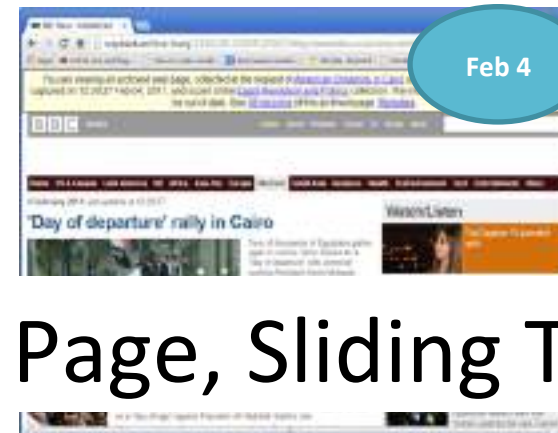
Found 26 Captures between Feb 6, 2011 - Jan 20, 2014

2011 11 pages	2012 4 pages	2013 9 pages	2014 2 pages
<a href="#">Feb 6, 2011</a> *	<a href="#">May 18, 2012</a> *	<a href="#">Jan 18, 2013</a>	<a href="#">Jan 19, 2014</a>
<a href="#">Feb 6, 2011</a> *	<a href="#">Jul 18, 2012</a>	<a href="#">Mar 19, 2013</a>	<a href="#">Jan 20, 2014</a>
<a href="#">Feb 7, 2011</a> *	<a href="#">Sep 18, 2012</a>	<a href="#">May 19, 2013</a>	
<a href="#">Feb 7, 2011</a> *	<a href="#">Nov 18, 2012</a>	<a href="#">May 20, 2013</a>	
<a href="#">Feb 8, 2011</a> *		<a href="#">Jul 19, 2013</a>	
<a href="#">Mar 9, 2011</a> *		<a href="#">Jul 20, 2013</a>	
<a href="#">Mar 11, 2011</a> *		<a href="#">Sep 19, 2013</a>	
<a href="#">May 18, 2011</a> *		<a href="#">Nov 19, 2013</a> *	
<a href="#">Jul 18, 2011</a> *		<a href="#">Nov 20, 2013</a>	
<a href="#">Sep 18, 2011</a> *			
<a href="#">Nov 18, 2011</a> *			



# Current browsing and searching services for the “Egypt Revolution and Politics” collection





Sliding Page, Sliding Time




storify.com

# Egypt Revolution and Politics - Broad summary

This is an automatically generated story from Archive-It collection.

by [jennaa83](#) · a few seconds ago


---



**28 Jan 2011 - Graphic: A closer look at the protests in Egypt**

Click here for more Related PodcastsSubscribe to this podcast via iTunes Listen to it on our Podcast page and get other National post.com podcasts. Our reporters and editors talk about the biggest stories around the world.


[NEWS.NATIONALPOST.COM](#)



**29 Jan 2011 - As it happened: Egypt unrest day five**

To recap on political developments, Egyptian President Hosni Mubarak has resigned to the unrest by appointing his first-ever deputy and starting to form a new government. There is a significant presence of Egyptian army forces in key areas in Suez, where 15 people were killed in the unrest on Friday.


[NEWS.BBC.CO.UK](#)



**03 Feb 2011 - Map and Diagram of the Protests in Tahrir Square**

Antigovernment protesters are manning barricades in an area where Mubarak supporters tried to push through on previous days. Protesters are using subway entrances (1) to dump trash, as well as to detain Mubarak supporters that have been captured. The army is maintaining a buffer zone in this area will.

[WWW.BYTHEWEEK.COM](#)



**11 Feb 2011 - Egypt's revolution: Interactive map**

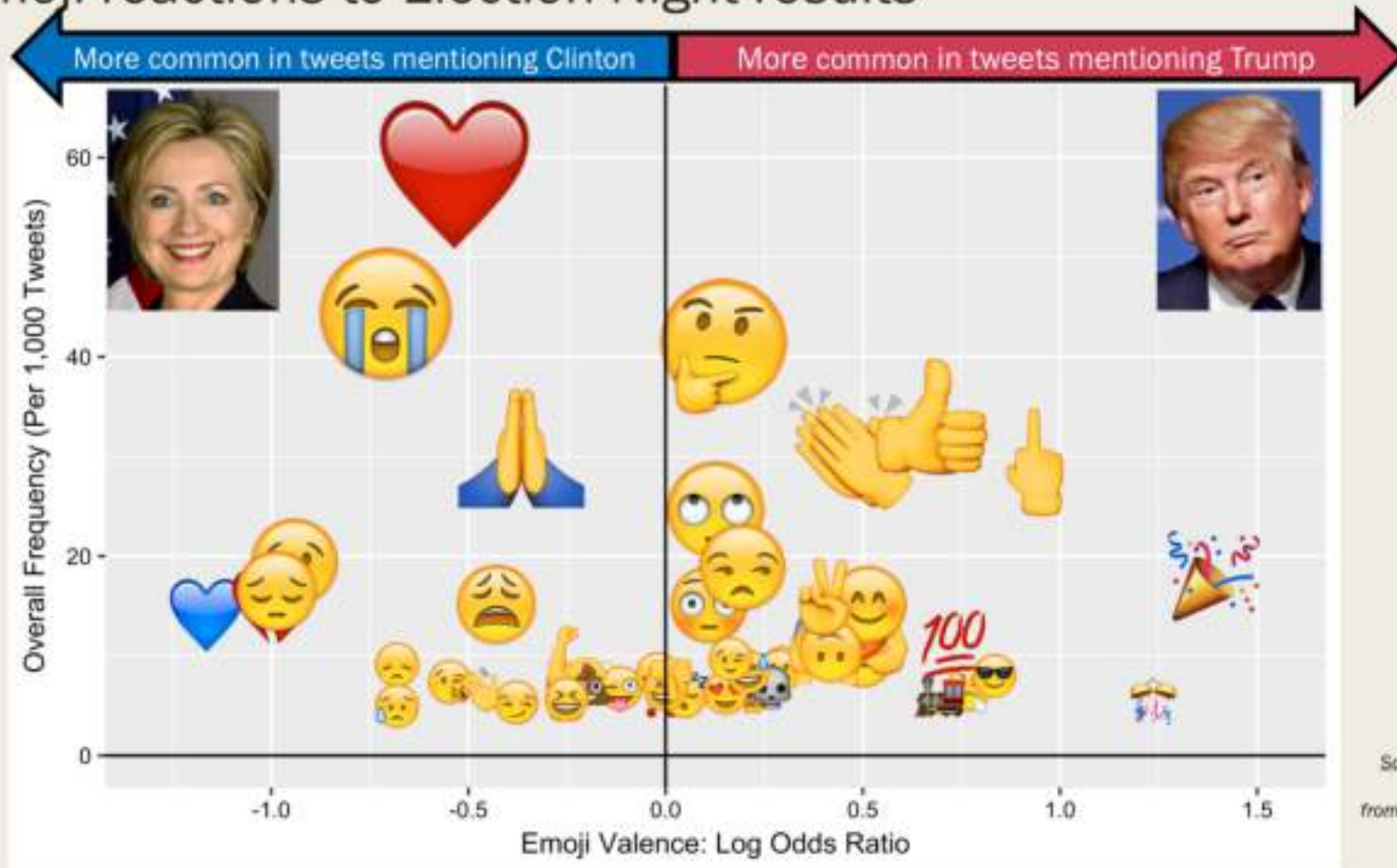
25 Jan - "Day of Revolt" The unrest began on 25 January, the "Day of Revolt", with demonstrations centred on Tahrir Square. 26 Jan - Violent cannon attack Demonstrators in Cairo were caught in amateur video footage praying on Cairo's Qasr El-Nila bridge despite being hit by water cannon. 27 Jan - Sina.

[WWW.BBC.CO.UK](#)



# The top emojis of Election Day 2016

## Emoji reactions to Election Night results



Source: Analysis of sampled Election Night tweets, from Nov. 9, 2016 12:00am – 3:30am EST. (PRISMOJI)



# Word clouds for three selected emojis the 🍑, the 🍑, and the 🤔 in tweets after midnight

## The divided emoji states of America



Yeah wow @realDonaldTrump you real worked hard AWESOME JOB!! THANK YOU !! 👍👍  
Please tell Obama ....YOUR FIRED...!! 🍑



I'm not scared let's just get these four years over with. Fuck trump, fuck the people who voted him and the people who didn't vote at all 🍑

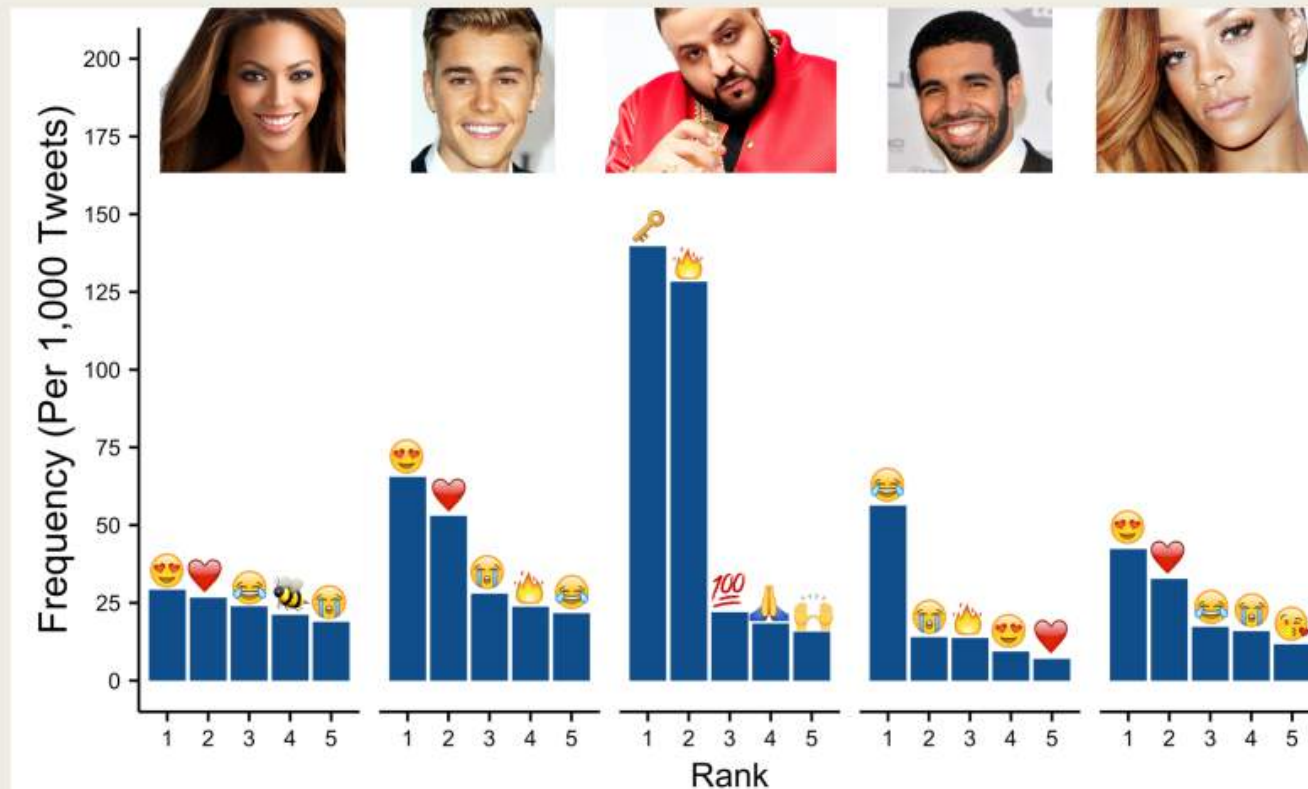


How can I go to work in the morning knowing the people I work with voted for trump. 🤔 I just can't.

Source: Analysis of sampled Election Night tweets, from Nov. 9, 2016 12:00am – 3:30am EST. Word clouds shown for tweets mentioning Hillary Clinton or Donald Trump as well as the given emoji above. Randomly selected tweet shown for each emoji.

# Humans may have a common emoji language of #fanlove

DJ Khaled is in a league of his own



Source: Analysis of 250,000 public tweets mentioning @beyonce, @justinbieber, @djkhaleid, @drake, and @rihanna from Aug. 1-4, 2016. (PRISMOJI)

David Robinson: “I saw a hypothesis about Donald Trump’s twitter account that simply begged to be investigated with data.”



Follow

Every non-hyperbolic tweet is from iPhone  
(his staff).

Every hyperbolic tweet is from Android (from  
him).



12:20 PM - 6 Aug 2016

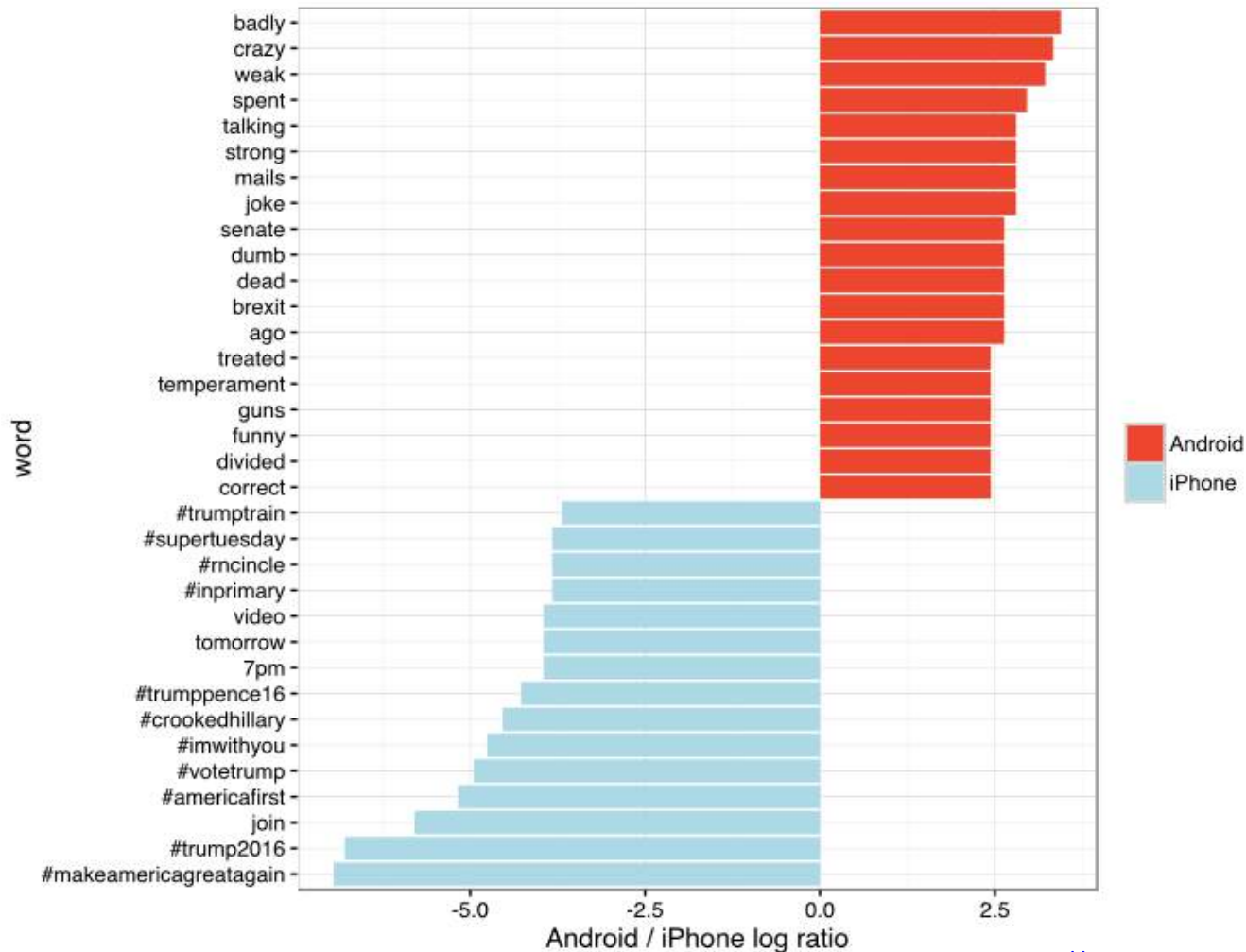
10,349 Retweets 14,504 Likes



265 10K 15K

Source: <http://varianceexplained.org/r/trump-tweets/>

# Text analysis of Trump's tweets confirms he writes only the (angrier) Android half



# Next Time

- Assignment 1 due
- Questions/comments on readings of Week 1
- Topics
  - Types of Web data
  - Getting data from the Web
    - Introduction to Web APIs
    - Retrieve data from different APIs
- Readings
  - (Required) Visualize This, Chapter 2: Handling Data
  - (Required) Web Mining : Accomplishments & Future Directions
  - Data Preparation for Mining World Wide Web Browsing Patterns
  - Can I text mine from databases or websites?

# Assignment 1

## Data Exploration and Analysis