# Storytelling for Summarizing Collections in Web Archives

Yasmin AlNoamany
Michele C. Weigle
**Michael L. Nelson**

Old Dominion University
Web Science and Digital Libraries Group
@WebSciDL

This work is supported in part by IMLS LG-71-15-0077

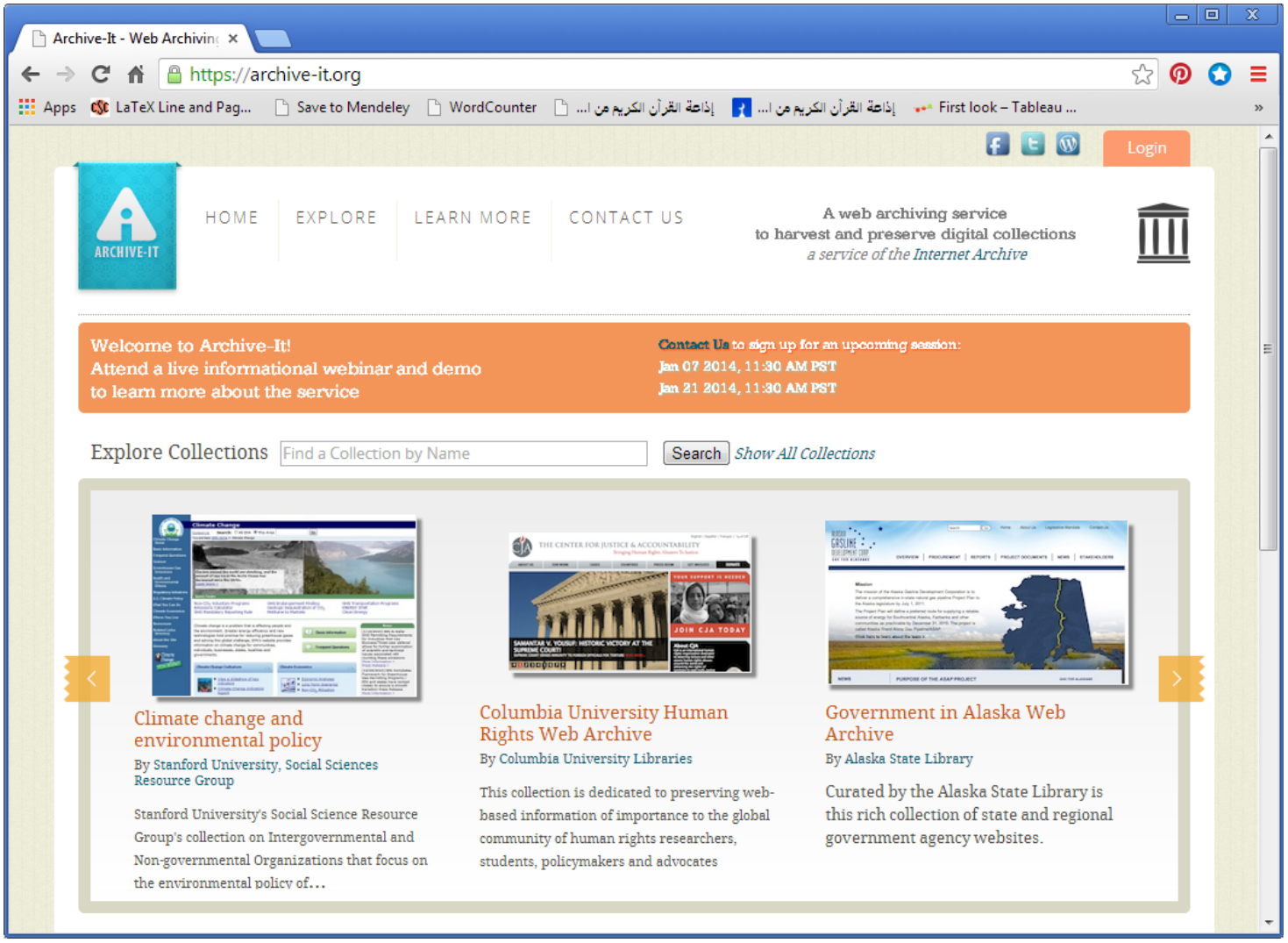
CNI Spring 2016
2016-04-05

# IMLS-Funded Research

1. Use small “stories” to summarize much larger collections of archived web pages
   - big → small
2. Generate web archive collections by mining user-generated stories for seed URIs
   - small → big

http://ws-dl.blogspot.com/2015/10/2015-10-07-imls-and-nsf-fund-web.html

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




> 3,000 collections

> 400 partners

> 10B archived pages

Collection title

Collection categorization according to the curator

Seed URI

Timespan of the resource and the number of times it has been captured

The group that the resource belongs to

Metadata about the collection

Text search box

List of the seed URIs

# Problem: Collection understanding and collection summarization are not currently supported

Not easy to answer “what’s in that collection?”

# There is more than one collection about the Egyptian Revolution

- “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

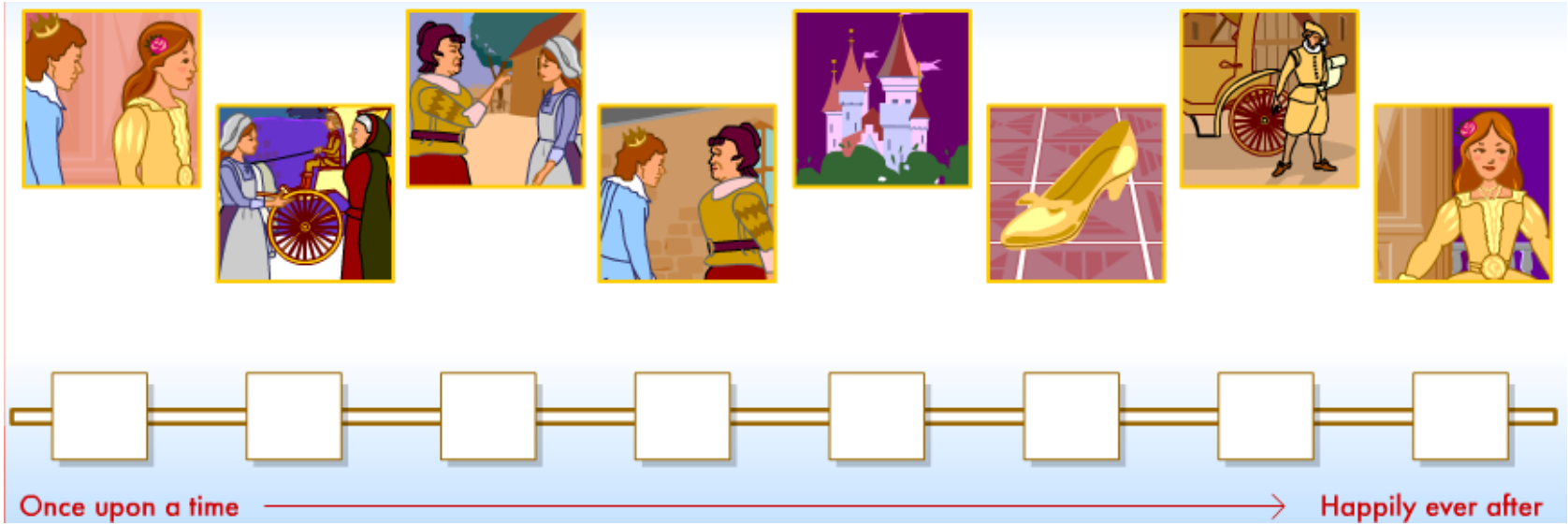
(1000s of Seeds X 1000s of Mementos)
+ Dimension of Time ==
*Conventional Vis Methods*
*Not Applicable*

Using Timelines, Treemaps, etc.:
http://ws-dl.blogspot.com/2012/08/2012-08-10-ms-thesis-visualizing.html

# Idea:
# Storytelling

# Stories in Literature

## Once upon a time…

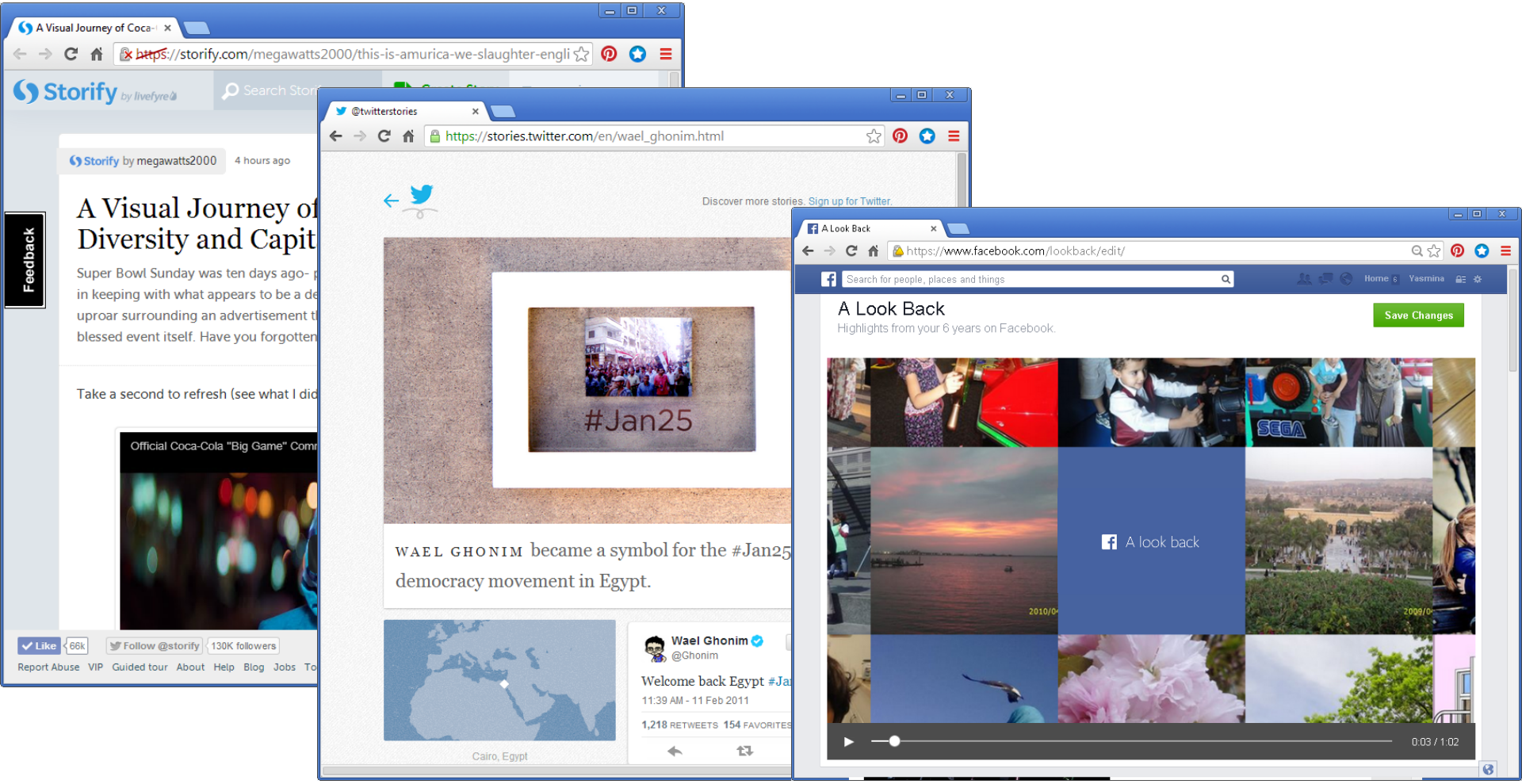



Story elements: setting, characters, sequence, exposition, conflict, climax, resolution

http://www.learner.org/interactives/story/

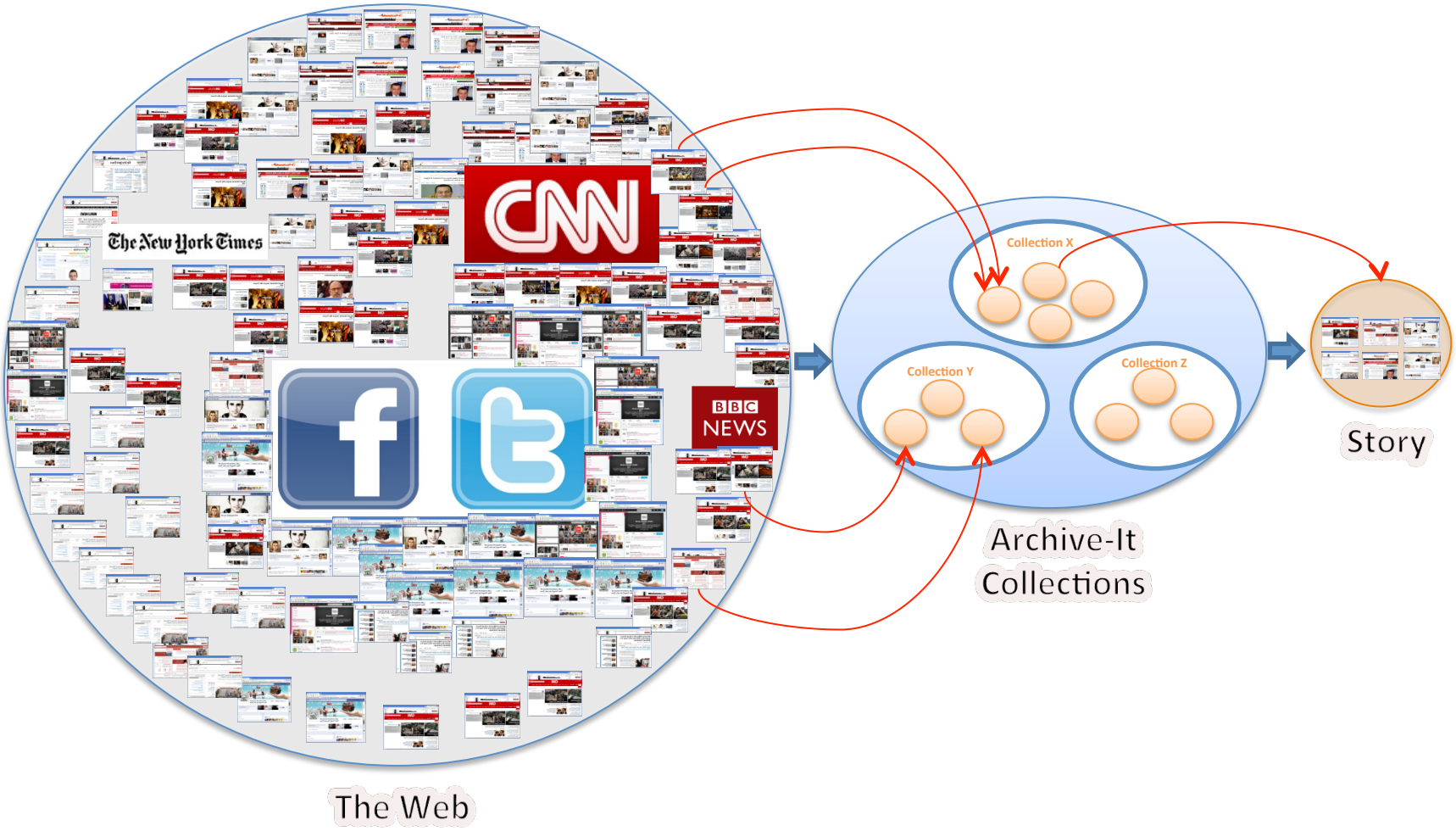
# Stories in social media

"It's hard to define a story, but I know it when I see it" (Alexander, 2008)



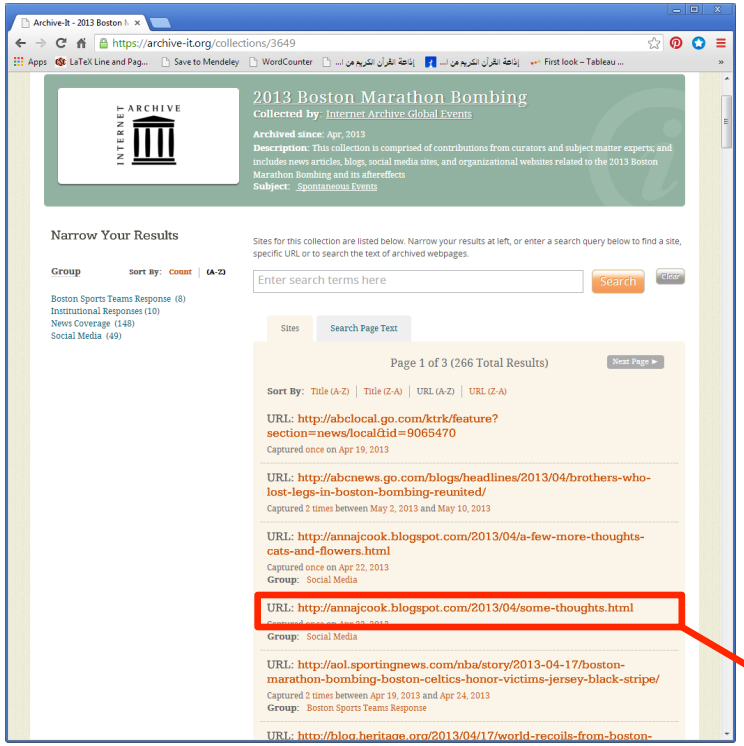

A sampling and arrangement of web resources for summarization.

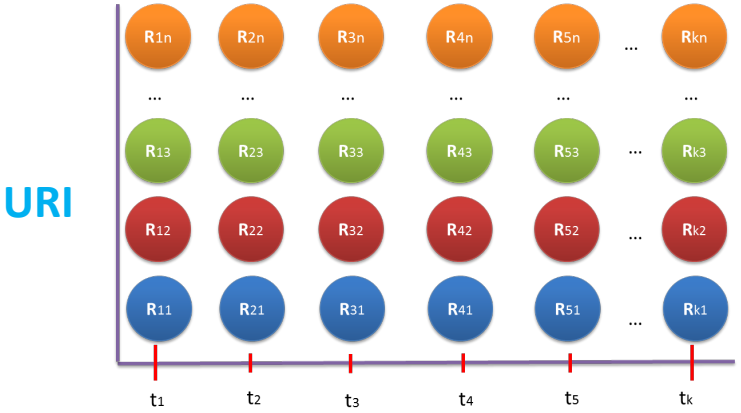
# Collection == thematic sample from the Web
# Story == arranged sample from the collection

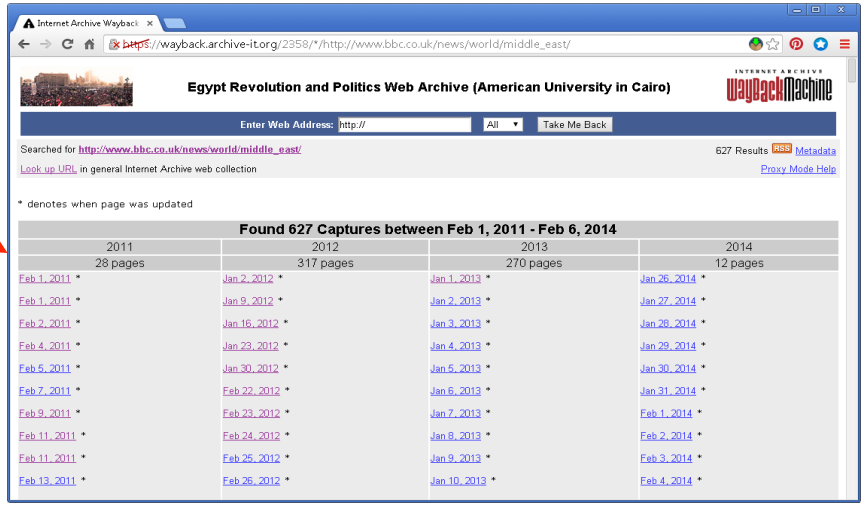



We sample *k* mementos from *N* pages of the collection to create a summary story

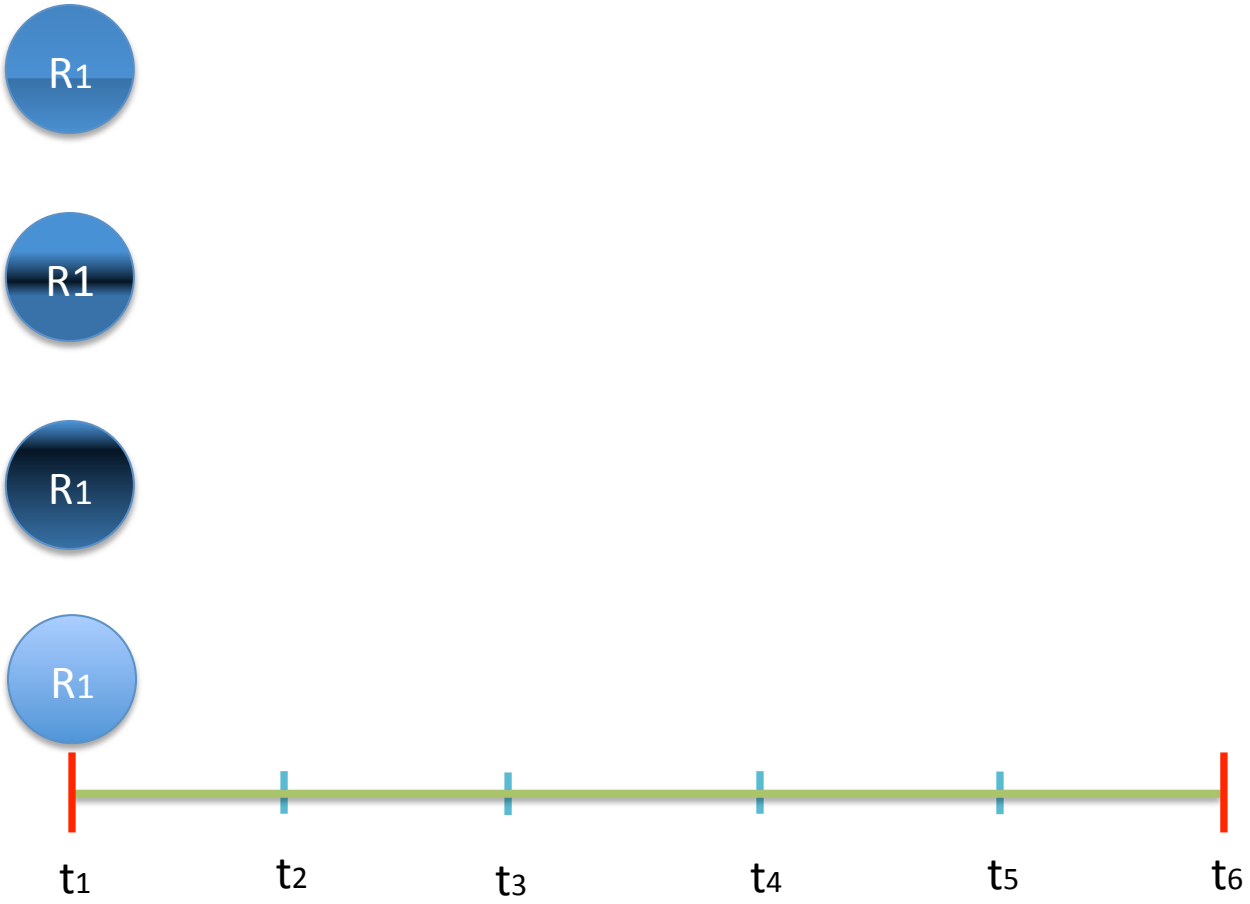
# Collections have two dimensions



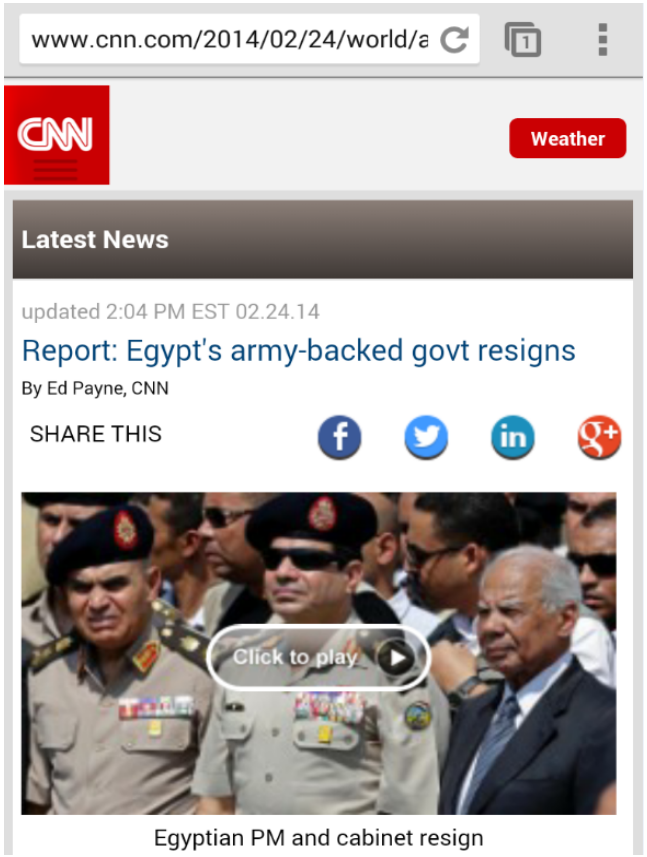



# Fixed Pages, Fixed Time



# Fixed Page, Fixed Time



## Andriod Chrome user-agent

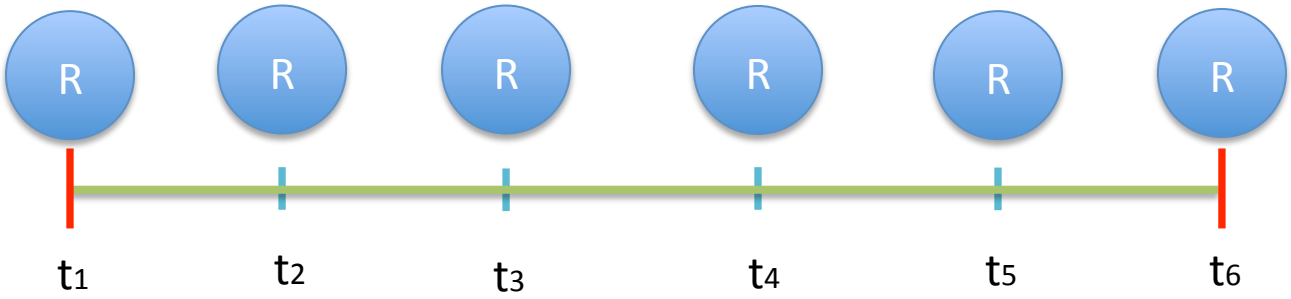
http://www.cnn.com/2014/02/24/world/africa/egypt-politics/index.html?hpt=wo_c2

## A desktop Chrome user-agent

http://www.cnn.com/2014/02/24/world/africa/egypt-politics/index.html?hpt=wo_c2

First Steps in Archiving the Mobile Web: Automated Discovery of Mobile Websites, JCDL 2013: https://www.harding.edu/fmccown/pubs/jcdlsp182-schneider.pdf
A Method for Identifying Personalized Representations in Web Archives, D-Lib Magazine, 2013: http://www.dlib.org/dlib/november13/kelly/11kelly.html

# Fixed Page, Sliding Time



## Egypt says it's shutting down mobile phone networks

Feb 1

Egyptian demonstrators gather in Cairo's Tahrir Square close to a military tank on Monday.

January 31st, 2011
07:40 PM ET
Share
Comments (155 comments)

Read full coverage of the unrest in Egypt updated continually by CNN reporters worldwide. Send your photos and video to iReport and see CNN in Arabic here.

[Update 3:11 a.m. Tuesday in Cairo, 8:11 p.m. ET Monday] CNN's Fareed

Recent Posts

Egypt says it's shutting down mobi networks

Egyptian government tightens grip protests

What's in Mubarak's mind? Former ambassador to Egypt may know

New hires not yet buoying the spiri autoworkers

50 people in 50 days: Girl disappear

## Egypt crisis: Mubarak says he'll won't seek re-election

Feb 1

Egyptian President Hosni Mubarak, in an appearance Egyptian state TV Tuesday night, said he won't seek re-election in September.

February 1st, 2011
03:37 PM ET
Share

Read full coverage of the unrest in Egypt updated continually by CNN reporters worldwide. Send your photos and video to iReport and see CNN in Arabic here. See also this strong roundup of timely, insightful views on the wave of upheaval

Recent Posts

Egypt crisis: Mubarak says he'll wo election

Woman called 'Jihad Jane' pleads g

Indian warship sunk during 'day at families

50 people in 50 days: Family of four

Gotta Watch: Whoopie pie dispute; extreme skiing

After mass dog slaughter, stresse

## Egypt crisis: 3 dead, 639 injured in Cairo, government officials say

Feb 2

Vehicles burn in front of Cairo's Egyptian Museum early Thursday as people protesting against Egyptian President Hosni Mubarak face off against pro-Mubarak crowds.

February 2nd, 2011
02:25 PM ET
Share
Comments

Read full coverage of the unrest in Egypt updated continually by CNN reporters worldwide. Send your photos and video to iReport and see CNN in Arabic here. See also this strong roundup of timely, insightful views on the wave of upheaval in the Arab world.

Recent Posts

Snowstorm update: 11,000 flights c falling in New York

Shooting reported in California eler

Egypt crisis: 3 dead, 639 injured in C government officials say

SI.com: It's National Signing Day!

Wednesday's intriguing people

Gotta Watch: Google art view; froze redux

## Egypt crisis: Protesters mass in square, at mosque; U.S. plans no flights out

Feb 4

Egyptian anti-government demonstrators gather at Cairo's Tahrir Square on Friday.

February 4th, 2011
01:09 AM ET
Share
Comments (32 comments)

Read full coverage of the unrest in Egypt updated continually by CNN reporters worldwide. Send your photos and video to iReport and see CNN in Arabic here. See also this strong roundup of timely, insightful views on the wave of upheaval in the Arab world.

Recent Posts

Friday's live video events

Egypt crisis: Protesters mass in sq mosque; U.S. plans no flights out

50 people in 50 days: Boy vanished school

The day's most popular stories

Dow rises 20 points

50 people in 50 days: California girl parking lot in '88

## Egypt crisis: Report of assassination attempt not confirmed, diplomat says

Feb 5

Anti-government protesters stand in front of Egyptian army tanks Saturday to prevent them from entering the opposition enclave in Cairo's Tahrir Square.

February 5th, 2011
08:30 AM ET
Share

Read full coverage of the unrest in Egypt updated continually by CNN reporters worldwide. Send your photos and video to iReport and see CNN in Arabic here. See also this strong roundup of timely, insightful views on the wave of upheaval

Recent Posts

Egypt crisis: Report of assassinatio confirmed, diplomat says

Arlington refuses burial of U.S. ally War

The day's most popular stories

Congressman requests visit with suspect

Dollars & Sense: Cashing in on sno steered toward Super Bowl

## Egypt crisis, Day 14: Google executive released

Feb 7

Demonstrators gather Monday in Cairo's Tahrir Square, which has been the center of protests against President Hosni Mubarak.

February 7th, 2011
12:23 PM ET
Share
Comments (18 comments)

Read full coverage and examine a timeline of the unrest in Egypt updated continually by CNN reporters worldwide. Send your photos and video to iReport and see CNN in Arabic here. See also this roundup of timely, insightful views on the wave of upheaval in the Arab world.

Recent Posts

Southern Sudan votes to split from

Egypt crisis, Day 14: Google executi

Monday's intriguing people

Assange lawyers raise specter of

Party ejection led to Ohio frat house police say

Monday quarterbacking: Super Bow fumbles

## Egypt transitioning, but to what?

Feb 9

February 8th, 2011
02:01 PM ET
Share
Comments (23 comments)
Permalink

With protesters still in the streets of Cairo and newly appointed government officials meeting with some opposition leaders, Egypt faces an uncertain future. The outcome will depend largely on the course charted by a transitional government.

Many other countries have been down the path Egypt finds itself on. A popular uprising undermines the authority of a long-standing regime and a period of

Recent Posts

The day's most popular stories

Dallas official under fire for giving

U.S. stocks post solid gains, Obam states with jobless aid

50 people in 50 days: Air Force nurs while on leave

Lindsay Lohan to be charged with

Tunisia now, and where it's going

## Egypt unrest: Protesters begin 18th day of demonstrations

Feb 11

Protesters in Cairo's Tahrir Square erupt in anger after President Hosni Mubarak indicates he will not resign.

February 11th, 2011
02:09 AM ET
Share
Comments (Add comment)

The latest developments, as confirmed by CNN, on the uprising in Egypt. Demonstrators have taken to the streets of Egypt's major cities to demand an end to President Hosni Mubarak's 30-year rule, prompting the government to deploy the military to deal with civil unrest for the first time in a generation. Check

Recent Posts

Egypt unrest: Protesters begin 18th demonstrations

President Obama's statement on E

The day's most popular stories

Arizona countersues feds over bor

Crowd at Egypt's Tahrir Square out Mubarak doesn't resign

Granddaughter wants share of Eins profits

## Mubarak quits: Switzerland freezes assets; Egyptians didn't talk with Obama

Feb 11

Thousands of Egyptians in Cairo celebrate President Hosni Mubarak's resignation Friday night.

February 11th, 2011
01:15 PM ET
Share
Comments (145 comments)

Developments, as confirmed by CNN, on the uprising in Egypt. Click here for up-to-the-minute updates. Check out our full coverage and the latest tweets from CNN correspondents on the ground. Click here to watch developments live from Egypt.

Recent Posts

Mubarak quits: Switzerland freeze Egyptians didn't talk with Obama

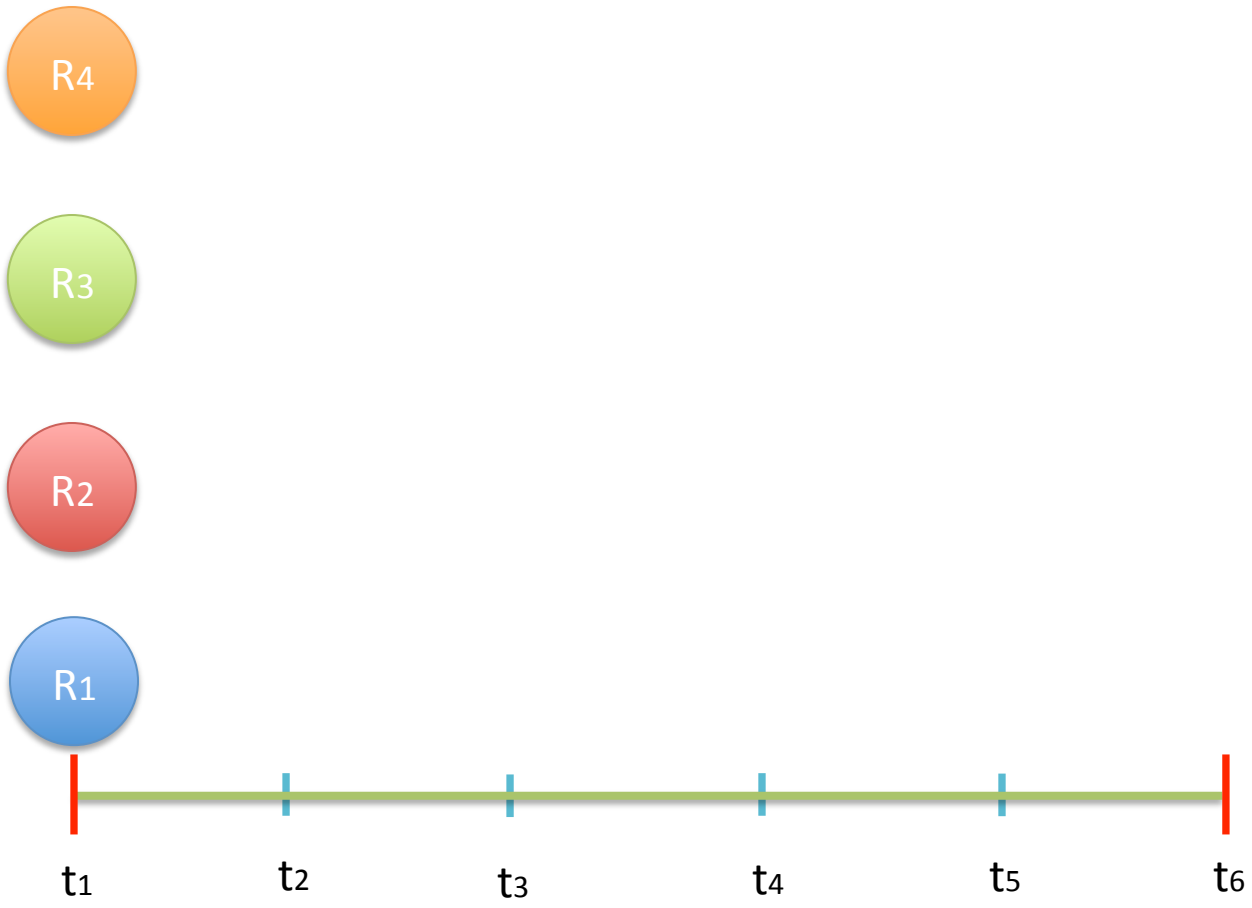
Scoliosis, not collision, likely left w expert says

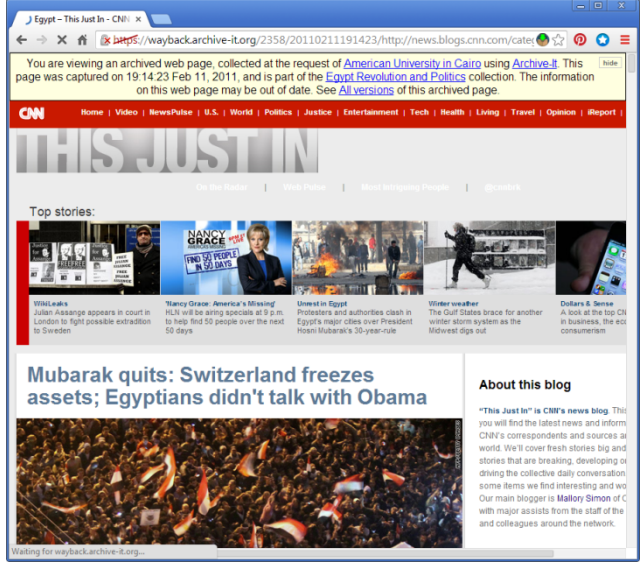
3 teenage boys among 7 Somalis fa Malaysia dive case

On the Radar widget in Egypt, CPAC Levy case, Gibbs' farewell

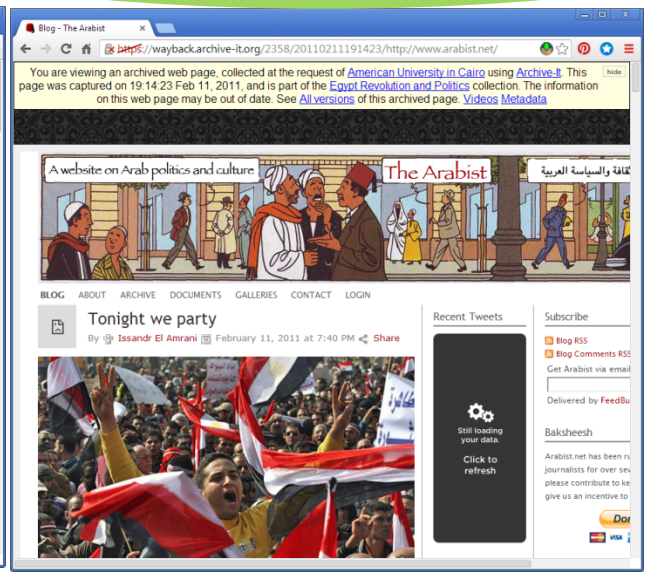
Friday's live video events

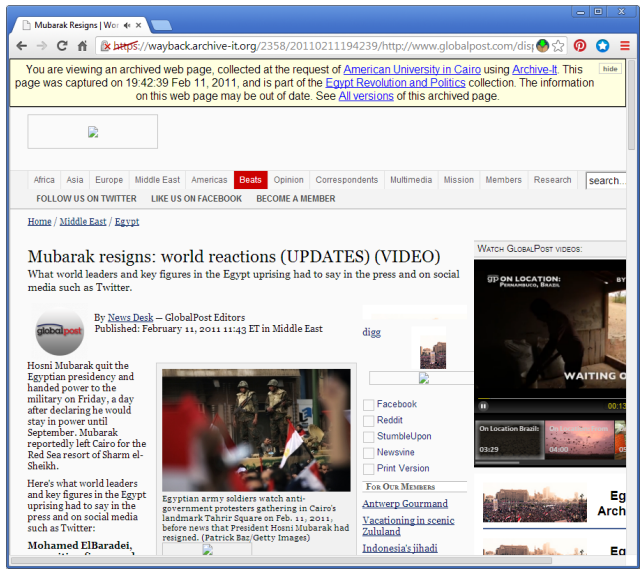
# Sliding Page, Fixed Time



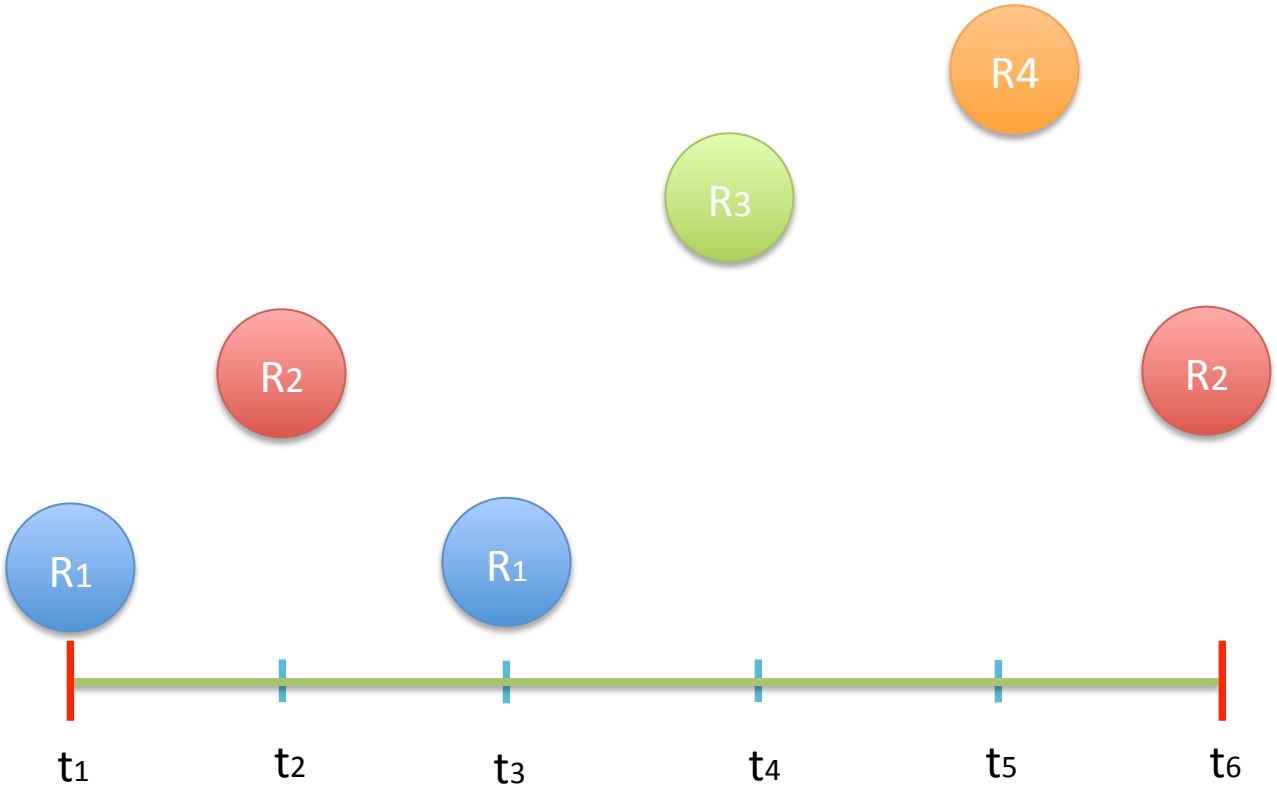



**Feb. 11, 2011**
**Mubarak resigns**

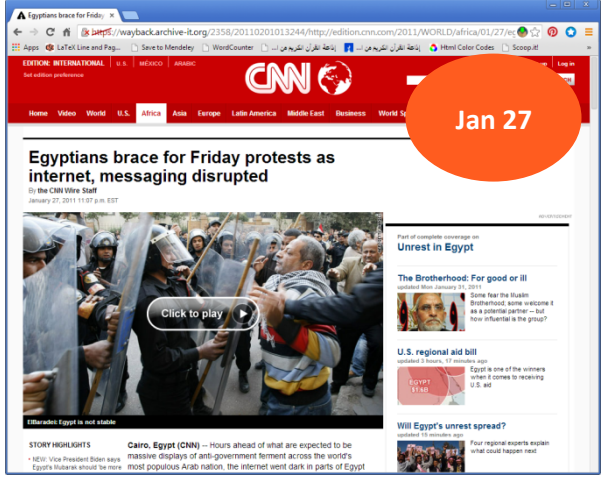


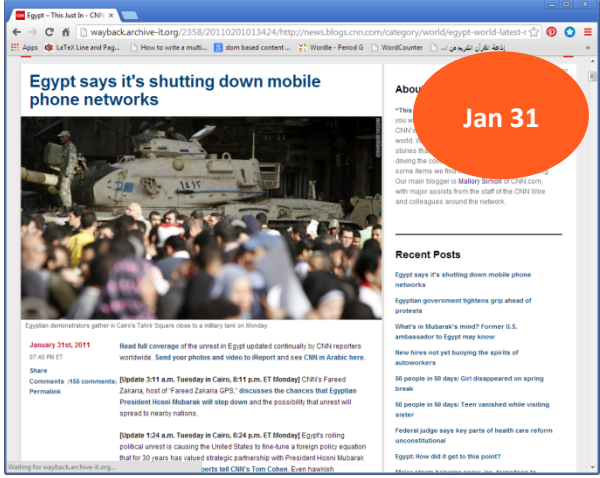



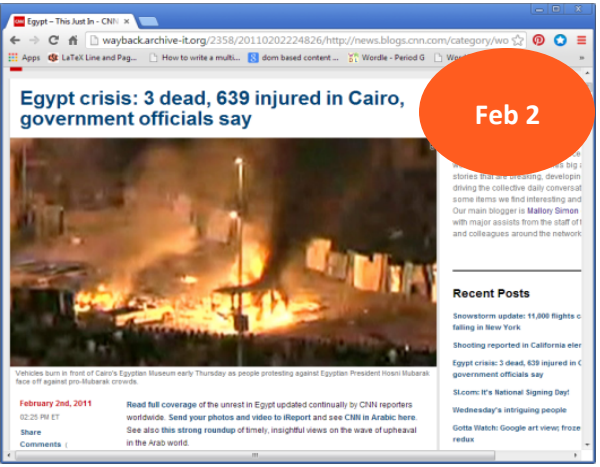
# Sliding Page, Sliding Time



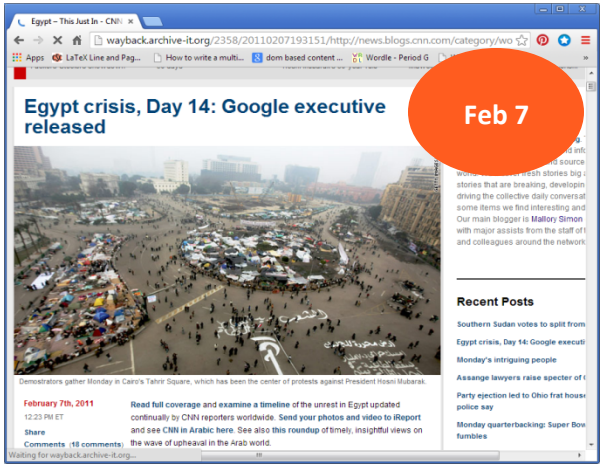
Jan 25
Egypt: The January 25 Demonstrations in Photographs

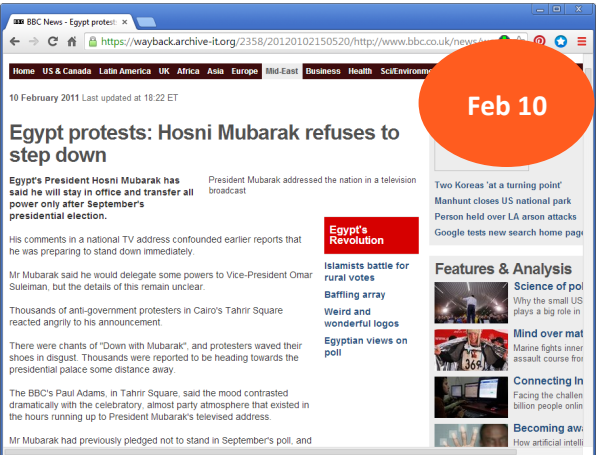

Jan 27
Egyptians brace for Friday protests as internet, messaging disrupted

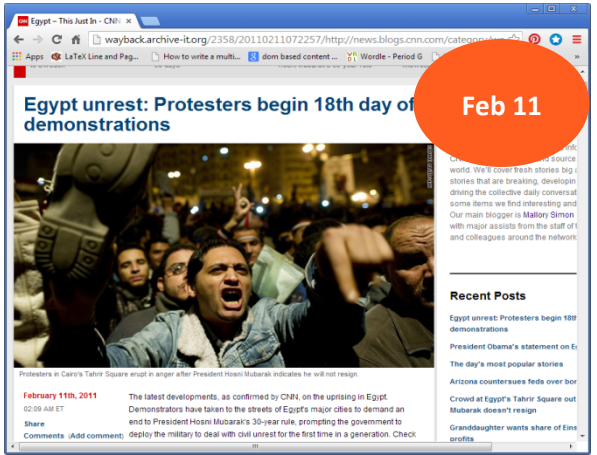

Jan 31
Egypt says it's shutting down mobile phone networks


Feb 2
Egypt crisis: 3 dead, 639 injured in Cairo, government officials say

Feb 4
'Day of departure' rally in Cairo


Feb 7
Egypt crisis, Day 14: Google executive released


Feb 10
Egypt protests: Hosni Mubarak refuses to step down


Feb 11
Egypt unrest: Protesters begin 18th day of demonstrations

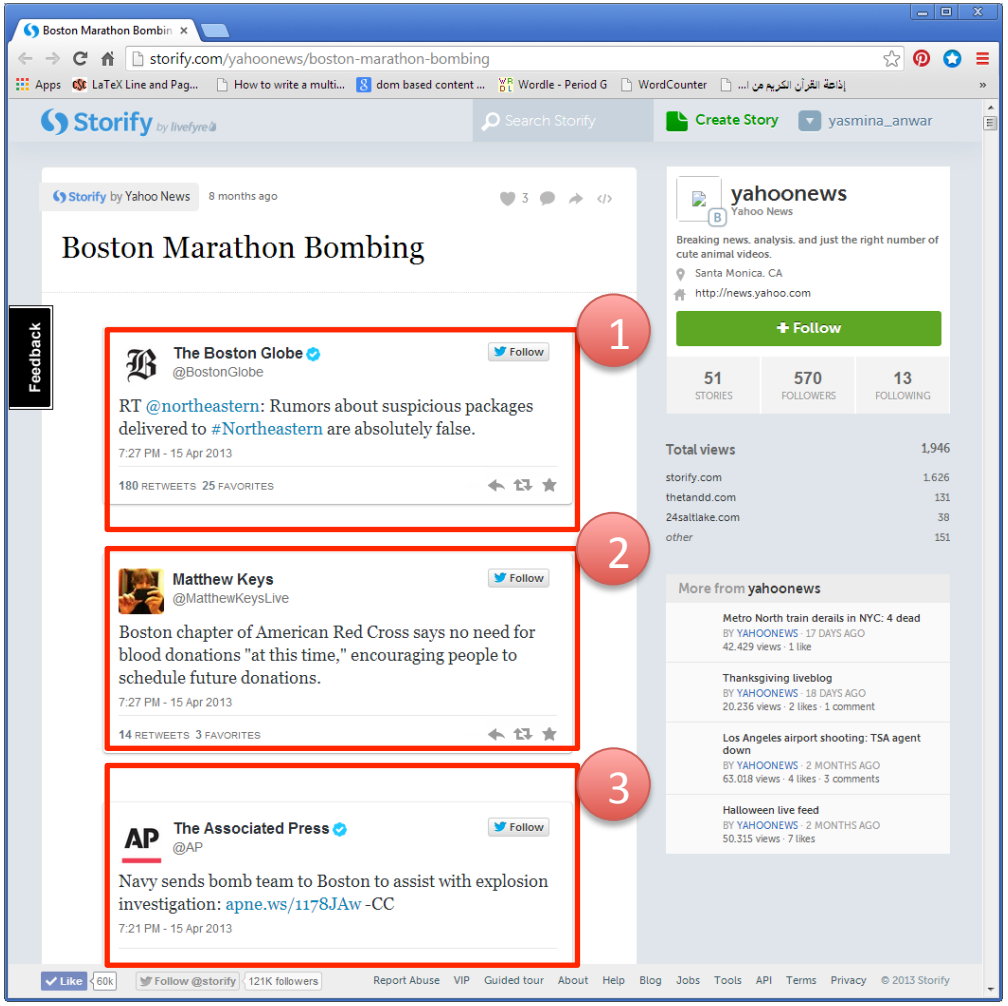
Feb 11
Egypt's Mubarak resigns as leader

# What do stories in Storify look like?

"Characteristics of Social Media Stories", TPDL 2015
http://www.cs.odu.edu/~mln/pubs/tpdl-2015/tpdl-2015-stories.pdf

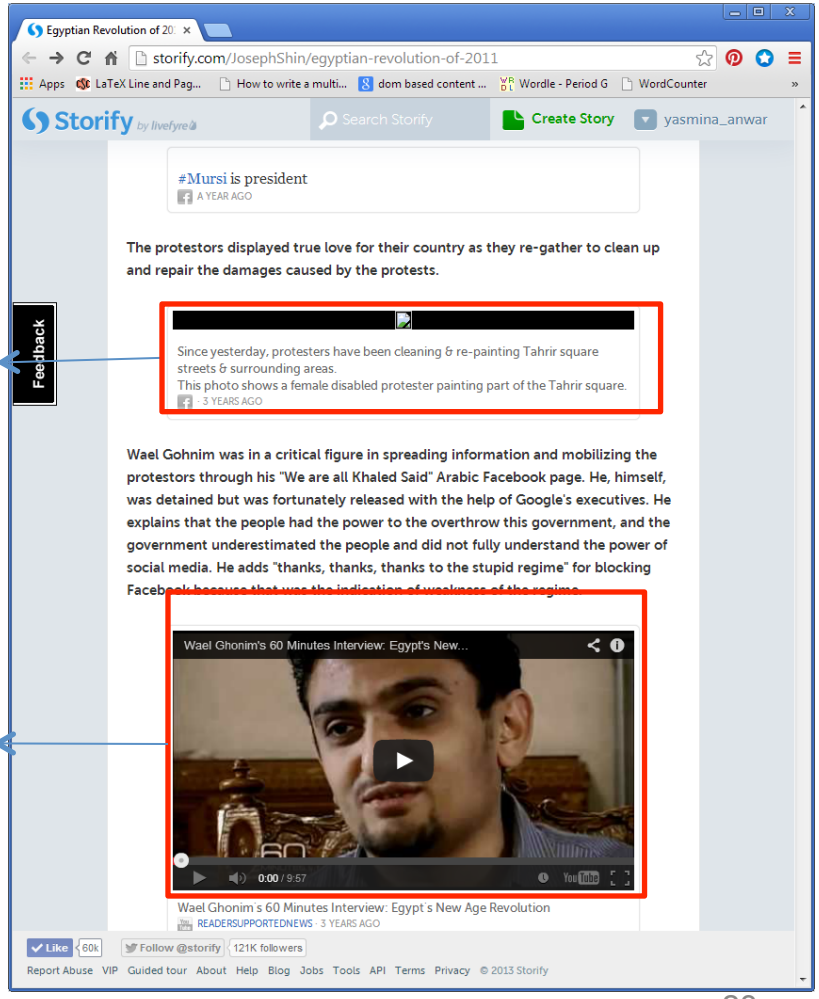
# What is the length of a story (the number of resources per story)?

- This story has 31 resources



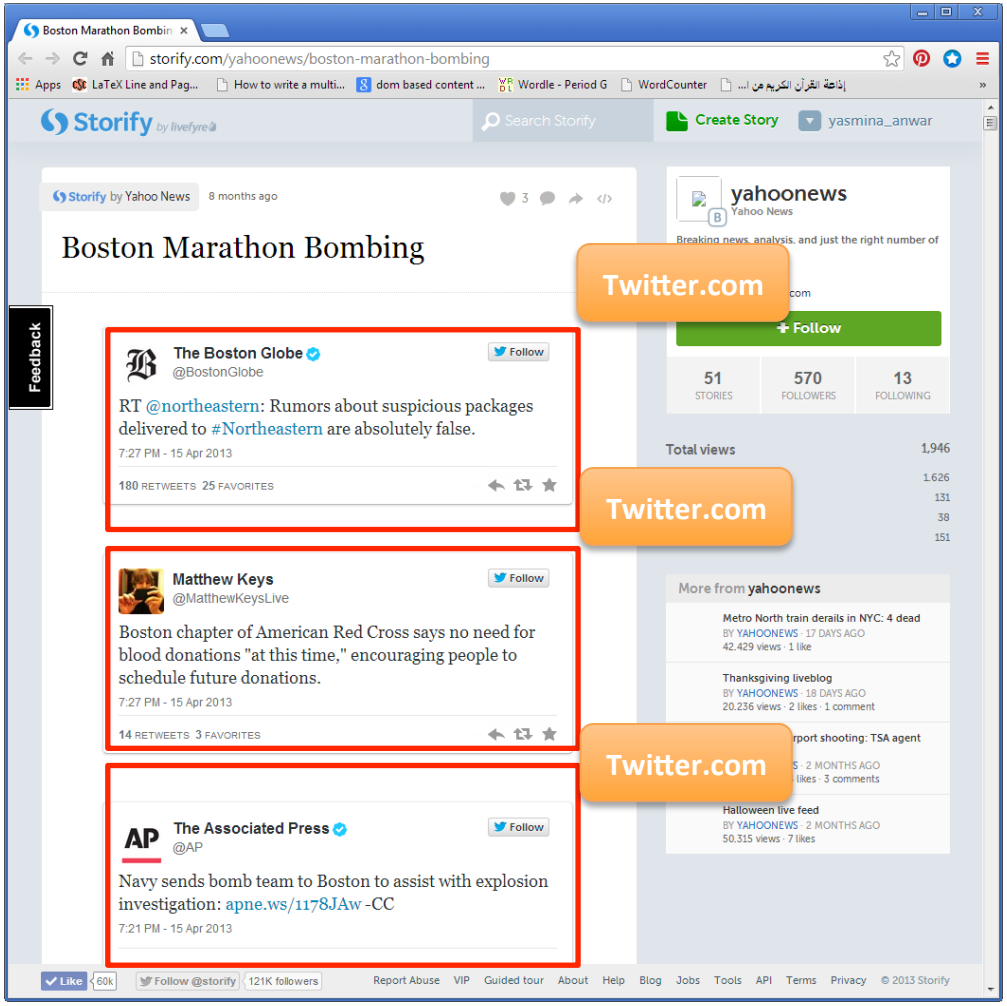
# What are the types of resources that compose a story?

- This story has
  - 19 quotes
  - 8 images
  - 4 videos



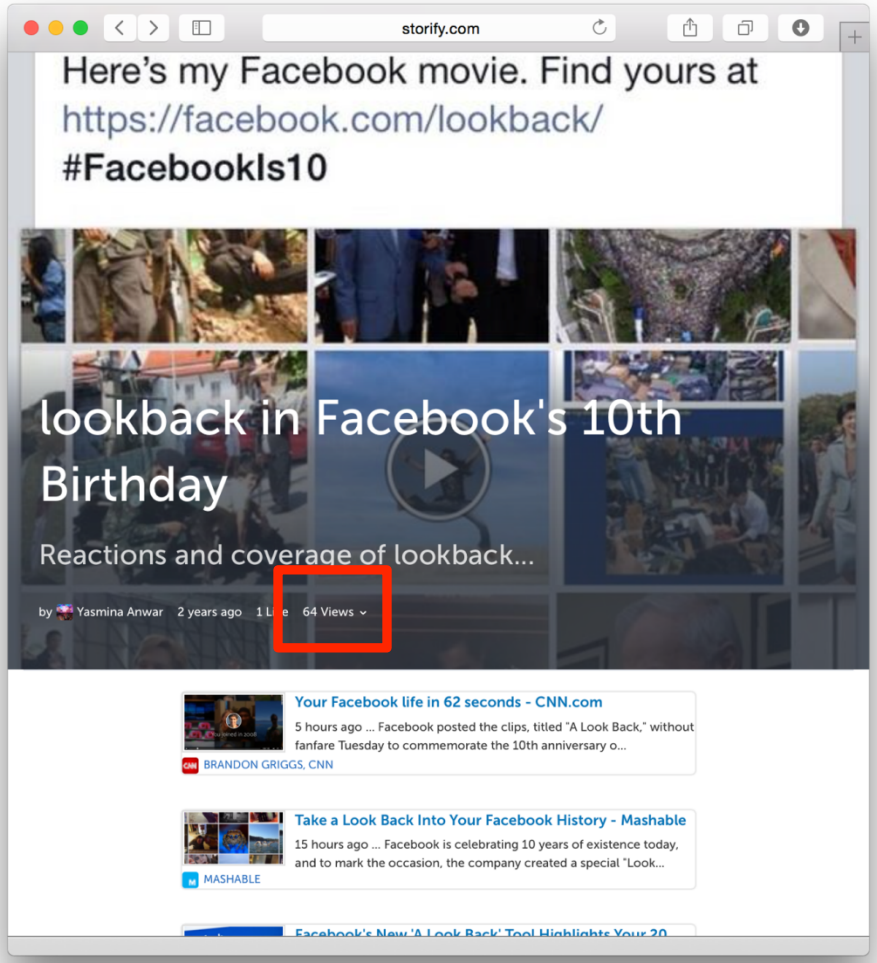
# What are the most frequently used domains?

- This story uses:
  - 90% twitter.com
  - 7% instagram.com
  - 3% facebook.com



# What differentiates a popular story?

**19,795 views**



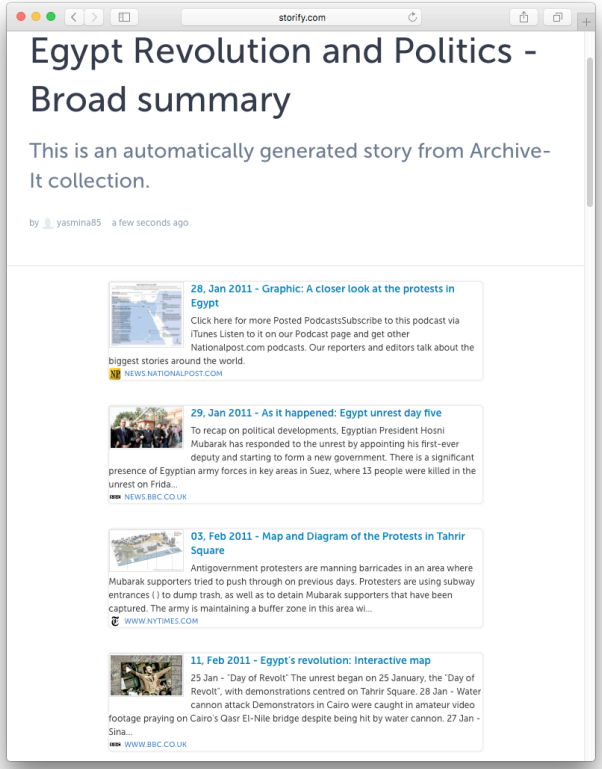

**64 views**

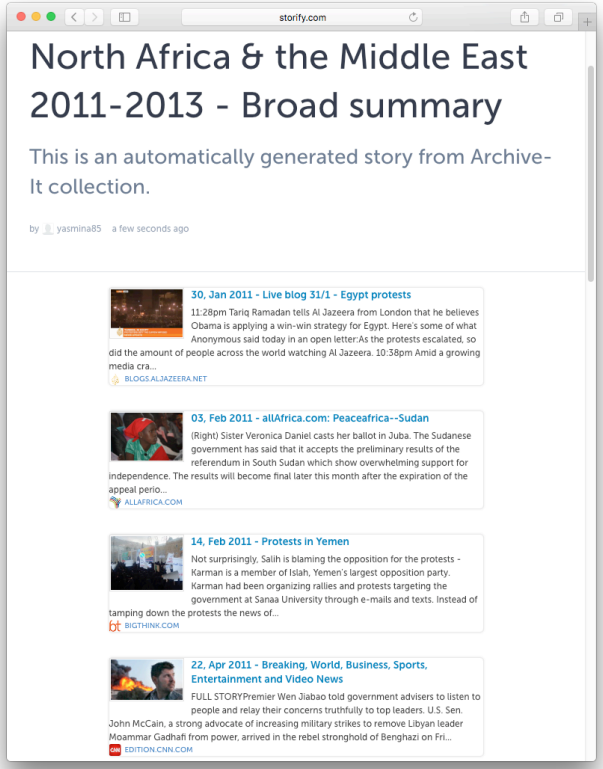
(skipping many details,
see TPDL 2015 paper)

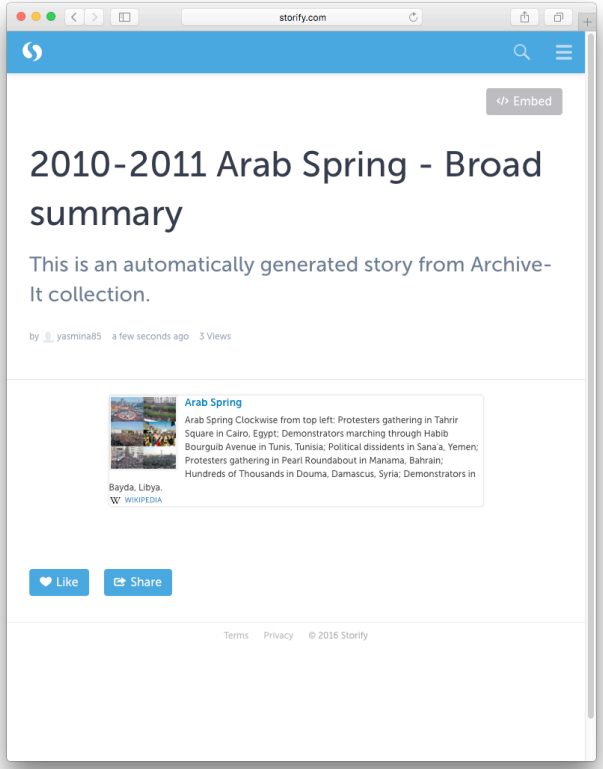
# We should create stories with:

- ~28 pages
- moar images!
- where possible, select pages from social media, news, blogs
- additional dimensions of quality:
  - are well archived (e.g., not missing images, stylesheets)
  - generate nice summaries in the Storify interface

# Stories from collections about the Egyptian Revolution

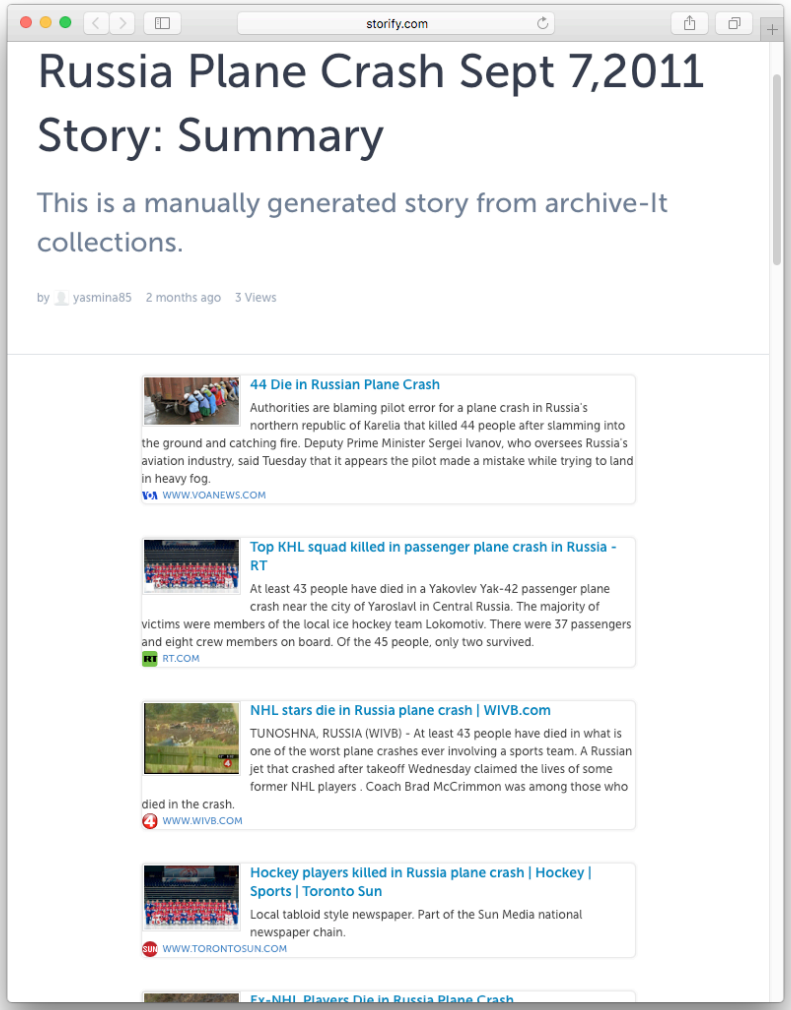


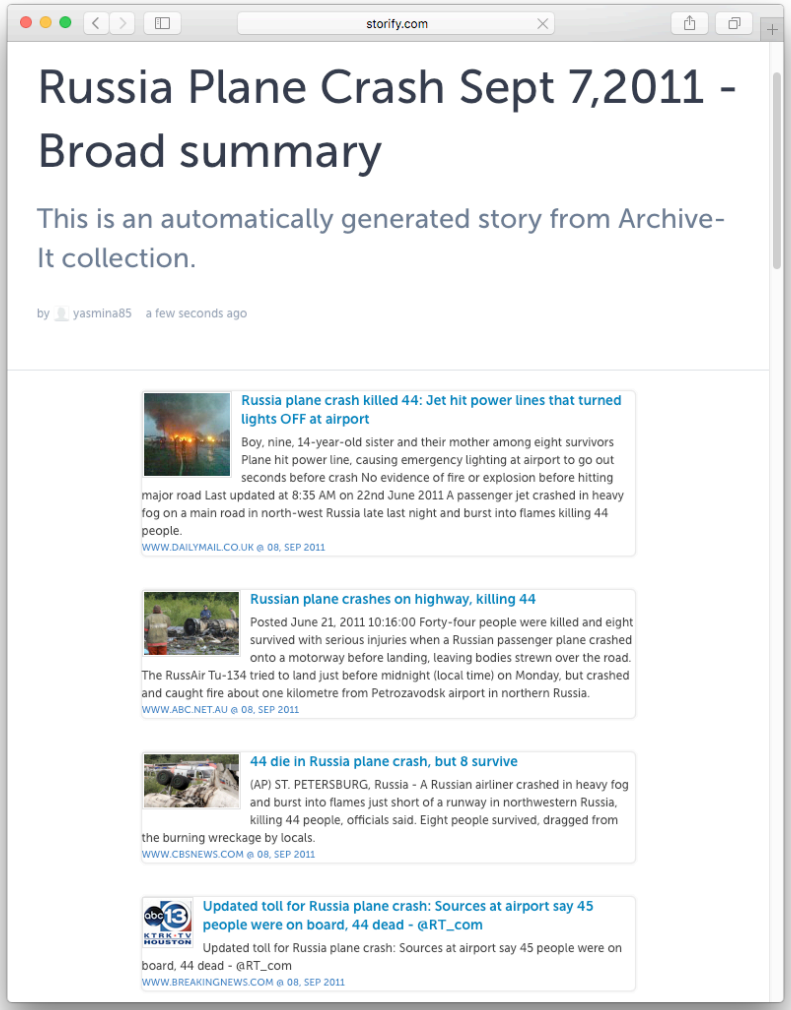




https://storify.com/yasmina85/auto-stories-from-archived-collections-56fbc3d1b8d27c6f6571c647
https://storify.com/yasmina85/auto-stories-from-archived-collections-5702ff8f228eede273d49c21
https://storify.com/yasmina85/auto-stories-from-archived-collections-5702c7f1228eede273d48ddf

# Evaluation: can humans tell human generated stories from machine generated?



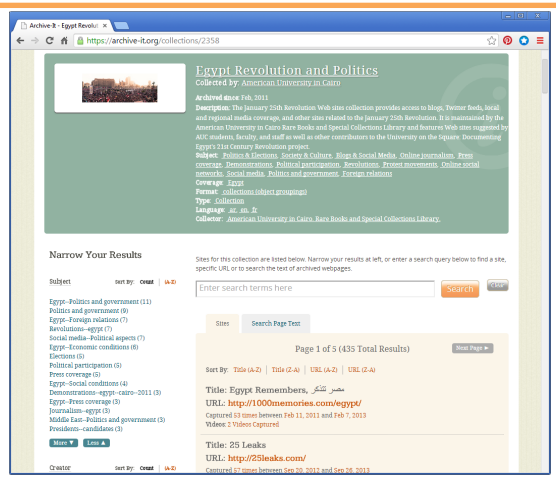

https://storify.com/yasmina85/this-is-manually-generated-story-from-archive-it-c-56b25ae72c0664474ee34f13
https://storify.com/yasmina85/auto-stories-from-archived-collections-56f1cfd36bc660f47f1b9f5e

# Use an interface *people already know how to use* to summarize collections

**Storytelling services**

curata
Storify
pearltrees
Facebook Timeline
Scoop.it!
TweetDeck by Twitter
Bundlr
roojoom
Pinterest
curate.us
paper.li
memolane
Snip.it

**Archived collections**




**Archived enriched stories**

more info:
https://github.com/yasmina85/OffTopic-Detection
http://ws-dl.blogspot.com/2015/09/2015-09-28-tpdl-2015-in-poznan-poland.html
http://ws-dl.blogspot.com/2015/08/2015-08-20-odu-l3s-stanford-and.html