

11-13-2015

Combining Storytelling and Web Archives

Yasmin AlNoamany
Old Dominion University

Michele C. Weigle
Old Dominion University, mweigle@odu.edu

Michael L. Nelson
Old Dominion University, mnelson@odu.edu

Follow this and additional works at: https://digitalcommons.odu.edu/computerscience_presentations

 Part of the [Archival Science Commons](#)

Recommended Citation

AlNoamany, Yasmin; Weigle, Michele C.; and Nelson, Michael L., "Combining Storytelling and Web Archives" (2015). *Computer Science Presentations*. 5.
https://digitalcommons.odu.edu/computerscience_presentations/5

This Book is brought to you for free and open access by the Computer Science at ODU Digital Commons. It has been accepted for inclusion in Computer Science Presentations by an authorized administrator of ODU Digital Commons. For more information, please contact digitalcommons@odu.edu.

Combining Storytelling and Web Archives

Yasmin AlNoamany
Michele C. Weigle
Michael L. Nelson

Old Dominion University
Web Science and Digital Libraries Group
ws-dl.cs.odu.edu
@WebSciDL

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

Old Dominion University ECE Department Colloquium
2015-11-13

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

The Web is important for
our cultural heritage.
How to preserve the
Web??!!!

What are web archives?



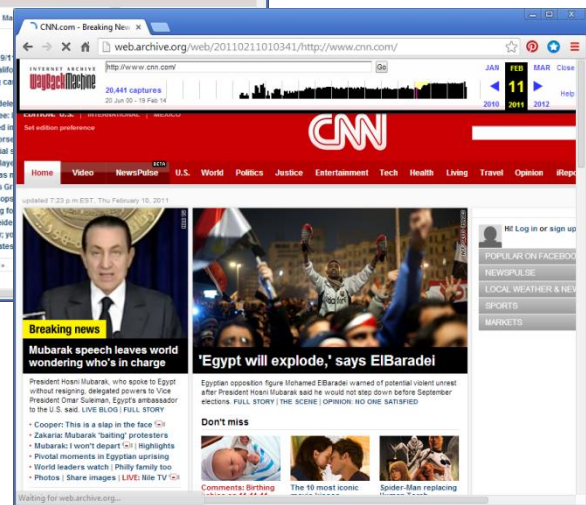
2002



2005



2009



2011

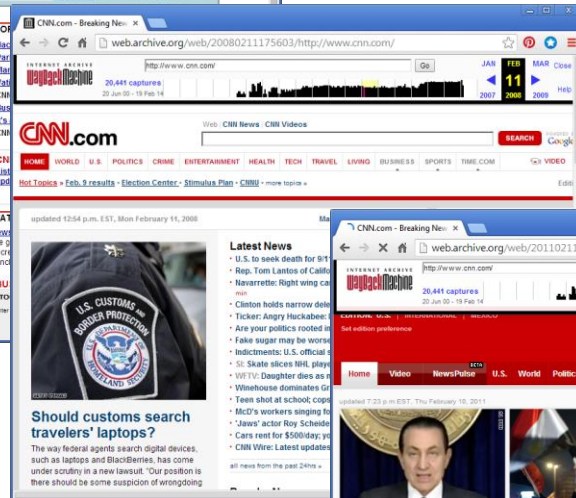
Memento: an archived snapshot as it appeared in the past



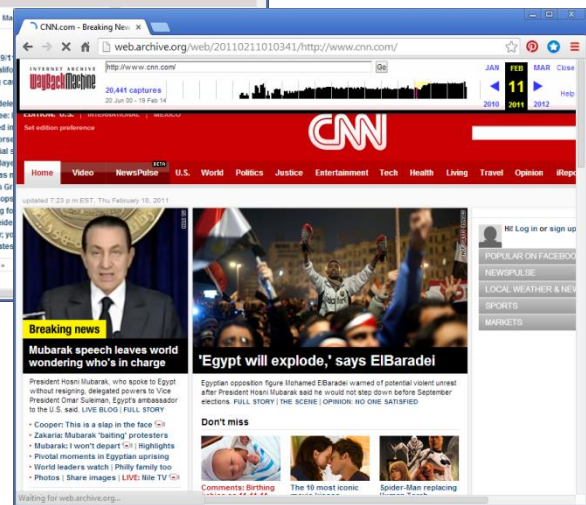
2002



2005



2009



2011

What are some web archives?

- The entire web
- On-demand free services
- National libraries
- ▶ • Theme-based collections



Archive.is
webpage capture



For more information: https://en.wikipedia.org/wiki/Web_archiving
<http://mementoweb.org/>

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

> 3,000 collections

~340 institutions

> 10B archived pages

The screenshot shows the Archive-It website in a browser window. The address bar displays <https://archive-it.org>. The page features a navigation menu with 'HOME', 'EXPLORE', 'LEARN MORE', and 'CONTACT US'. A central banner reads 'A web archiving service to harvest and preserve digital collections a service of the Internet Archive'. Below this, an orange banner promotes a webinar: 'Welcome to Archive-It! Attend a live informational webinar and demo to learn more about the service'. It lists two sessions: 'Jan 07 2014, 11:30 AM PST' and 'Jan 21 2014, 11:30 AM PST'. A search bar is labeled 'Explore Collections' with the placeholder 'Find a Collection by Name' and a 'Search' button. Three featured collections are displayed in a grid:

- Climate change and environmental policy**
By Stanford University, Social Sciences Resource Group
Stanford University's Social Science Resource Group's collection on Intergovernmental and Non-governmental Organizations that focus on the environmental policy of...
- Columbia University Human Rights Web Archive**
By Columbia University Libraries
This collection is dedicated to preserving web-based information of importance to the global community of human rights researchers, students, policymakers and advocates
- Government in Alaska Web Archive**
By Alaska State Library
Curated by the Alaska State Library is this rich collection of state and regional government agency websites.

The screenshot shows the Internet Archive interface for a collection titled "2013 Boston Marathon Bombing". The page includes a header with the Internet Archive logo and collection title, a metadata section with details like "Collected by: Internet Archive Global Events" and "Archived since: Apr, 2013", a search box, and a list of seed URIs. Annotations with arrows point to specific elements: the collection title, metadata, a search box, a list of categories (Boston Sports Teams Response, Institutional Responses, News Coverage, Social Media), a seed URI, the timespan of a resource, and the group it belongs to.

Collection title

Metadata about the collection

Text search box

Collection categorization based on the curator

Seed URI

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

List of the seed URIs

The group that the resource belongs to

What is the problem with the archived collections?

There is more than one collection about “Egyptian Revolution”

The image displays three overlapping screenshots of the Archive-It website, each showing a different collection related to the Egyptian Revolution. The top-left screenshot shows the "2010-2011 Arab Spring" collection, collected by the University of South Florida. The middle screenshot shows the "North Africa & the Middle East 2011-2013" collection, collected by Internet Archive Global Events. The bottom-right screenshot shows the "Egypt Revolution and Politics" collection, collected by the American University in Cairo. 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>

Collection understanding and
collection summarization are not
supported currently

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

Center for
Human Rights
Documentation
& Research

Human Rights

Collected by: [Columbia University Libraries](#)

Archived since: May, 2008

Description: An initiative of CUL's Center for Human Rights Documentation and Research, the Human Rights Web Archive is an effort to preserve and ensure access to freely available human rights resources created mainly by non-governmental organizations, national human rights institutions, and individuals.

Subject: [Society & Culture](#), [Human rights](#), [Non-governmental organizations](#), [Human rights workers](#), [National human rights institutions](#), [Web archives](#)

[More](#) ▼

Narrow Your Results

Sites 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.

Group

- [Amnesty International sections \(48\)](#)
- [Blogs by individuals \(6\)](#)
- [National human rights institutions \(75\)](#)
- [Non-governmental organizations \(359\)](#)
- [Truth commissions, tribunals, and courts \(12\)](#)

Subject

- [Human rights \(403\)](#)
- [Human rights advocacy \(192\)](#)
- [National human rights institutions \(74\)](#)
- [Civil rights \(47\)](#)

Page 1 of 6 (507 Total Results)

Sort By: [Title \(A-Z\)](#) | [Title \(Z-A\)](#) | [URL \(A-Z\)](#) | [URL \(Z-A\)](#)

Title: [Anti Caste Discrimination Alliance, ACDA](#)

URL: <http://acda.co/>

Requested: 1996

2011 01/29/2014

Archive-It - Human Rights

Human rights advocacy (196)
National human rights institutions (88)
Civil rights (52)
Democracy (43)

More ▾

Creator

Sort By: Count | (A-Z)

Asociación Paz con Dignidad (4)
Comité de Familiares de Detenidos-
Desaparecidos en Honduras (3)
Voix des sans-voix pour les droits de l'homme
(Organization) (3)
Amnesty International Hong Kong (2)
Anti Caste Discrimination Alliance (2)

More ▾

Language

Sort By: Count | (A-Z)

English (420)
Spanish (109)
Arabic (91)
French (80)
Russian (32)

Title: Anti Caste Discrimination Alliance, ACDA**URL:** <http://acda.co/>**Description:** UK-focused organization based in Derby, working to eliminate caste-based discrimination. In English. New site; see www.acdauk.org.uk/ for older site. Captured 10 times between Oct 13, 2011 and Jan 2, 2014**Subject:** Discrimination, Caste, Caste-based discrimination, East Indians**Group:** Blogs by individuals**Creator:** Anti Caste Discrimination Alliance**Language:** English**Coverage:** Great Britain**Collector:** Columbia University Libraries, Center for Human Rights Documentation and Research**Title:** Americans for Democracy and Human Rights in Bahrain**URL:** <http://adhrb.org/>**Description:** "Americans for Democracy and Human Rights in Bahrain (ADHRB) fosters awareness of and support for democracy and human rights in Bahrain by engaging with key actors within the US government to advocate for stronger US policies that support human rights for the Bahraini people." Captured once on Jan 21, 2014**Subject:** Democracy, Human rights**Group:** Non-governmental organizations**Creator:** Americans for Democracy and Human Rights in Bahrain**Language:** English**Coverage:** Bahrain**Collector:** Columbia University Libraries, Center for Human Rights Documentation and Research

Center for
Human Rights
Documentation
& Research

Human Rights Web Archive (Columbia University Libraries)



Enter Web Address: All

Searched for <http://acda.co/>

10 Results [RSS](#)

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

[Proxy Mode Help](#)

* denotes when page was updated

Found 10 Captures between Oct 13, 2011 - Jan 2, 2014

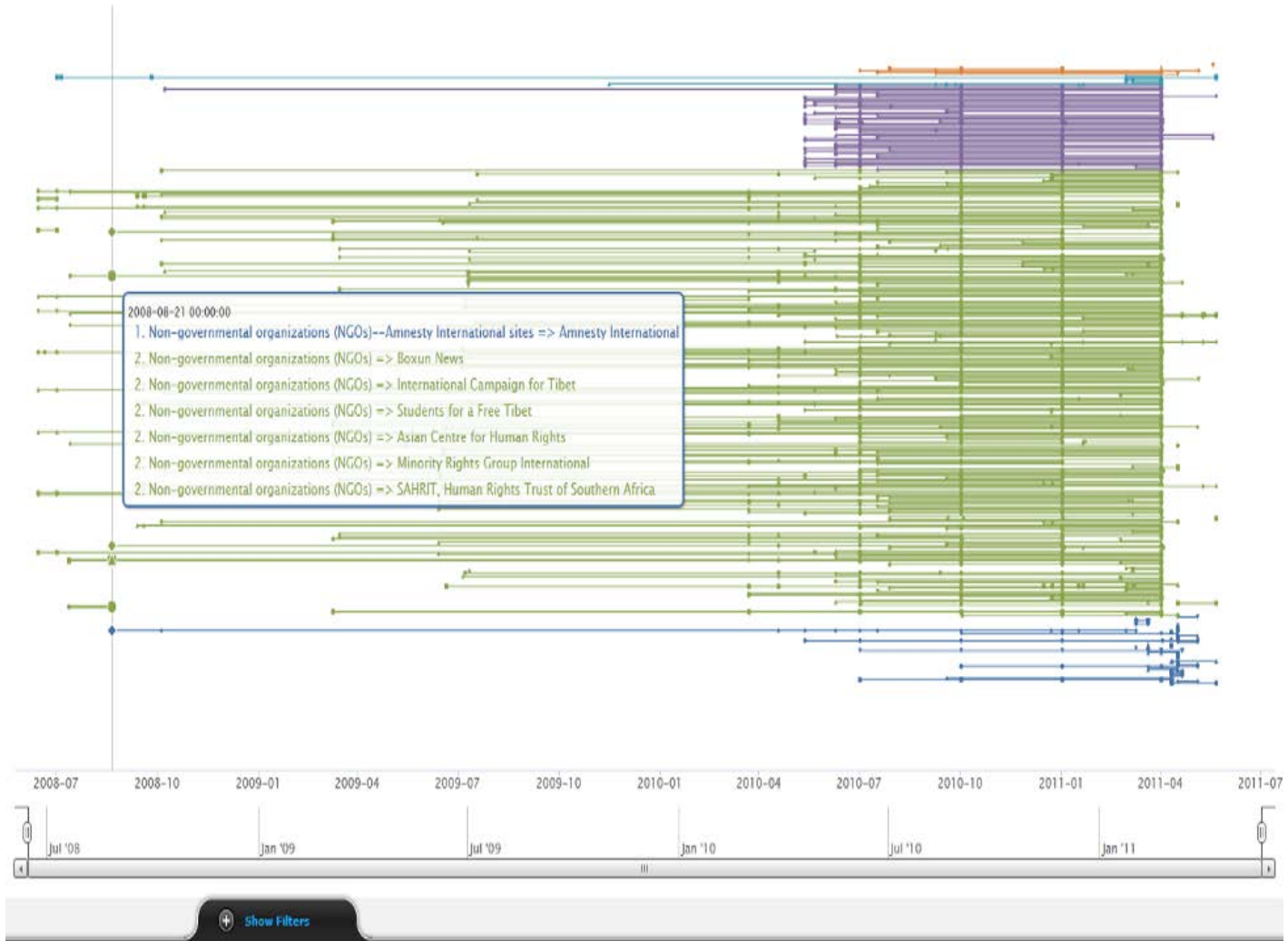
2011	2012	2013	2014
2 pages	4 pages	3 pages	1 page
Oct 13, 2011 *	Mar 2, 2012	Mar 7, 2013	Jan 2, 2014
Dec 2, 2011	Jun 2, 2012	Jun 11, 2013	
	Sep 2, 2012	Oct 1, 2013	
	Dec 2, 2012		

Our Early Attempts at Collection
Understanding Tried To Include
Everything...

<http://ws-dl.blogspot.com/2012/08/2012-08-10-ms-thesis-visualizing.html>

Timelines

COLLECTION: Human Rights



Treemaps

Collection ID: 1068

Visualize!

COLLECTION: Human Rights
 You are viewing 2. Non-governmental organizations (NGOs)



Treemaps

Collection ID: 1068

Visualize!

COLLECTION: Human Rights You are viewing COLLECTION: Human Rights

2. Non-governmental organizations (NGOs) (277)

3. National human rights institutions (NHRIs) (55)

1. Non-governmental organizations (NGOs)--Amnesty International sites (42)

4. Blogs by individuals (6)

5. Truth Commissions, Tribunals, and Courts (6)

Contains 6 articles

Days old

1 day
< 1 week
1 - 2 weeks
2 - 4 weeks
1 - 2 months
2 - 4 months
4 - 8 months
8 - 12 months
> 1 year

Snapshot

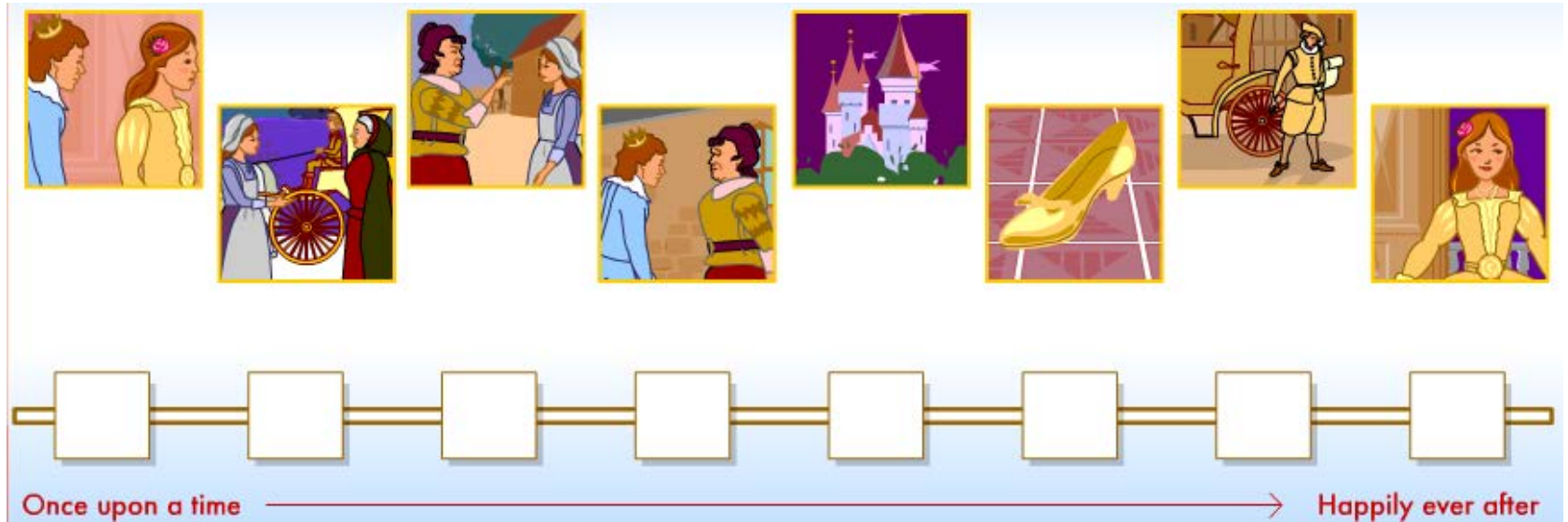
5. Truth Commissions, Tribunals, and Courts (6)

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

Idea: Storytelling

Stories in Literature

Once upon a time

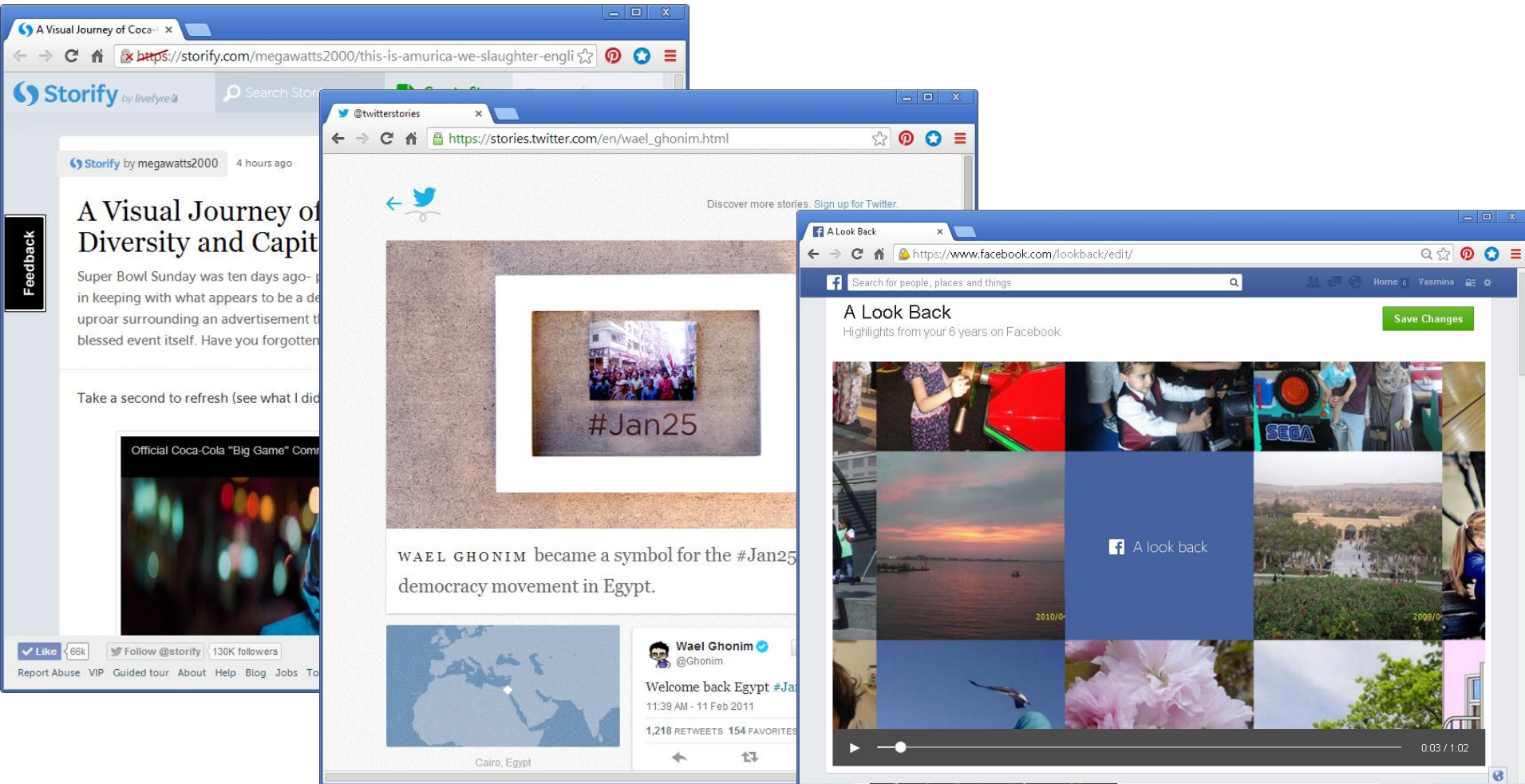


Story elements: setting, characters, sequence, exposition, conflict, climax, resolution

<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 loose context of the conventional definition in Literature

“storytelling” is becoming a popular technique in social media



Facebook
Timeline



Bundlr



Pinterest



paper.li



There is interest in
choosing k items from N
where $k \ll N$

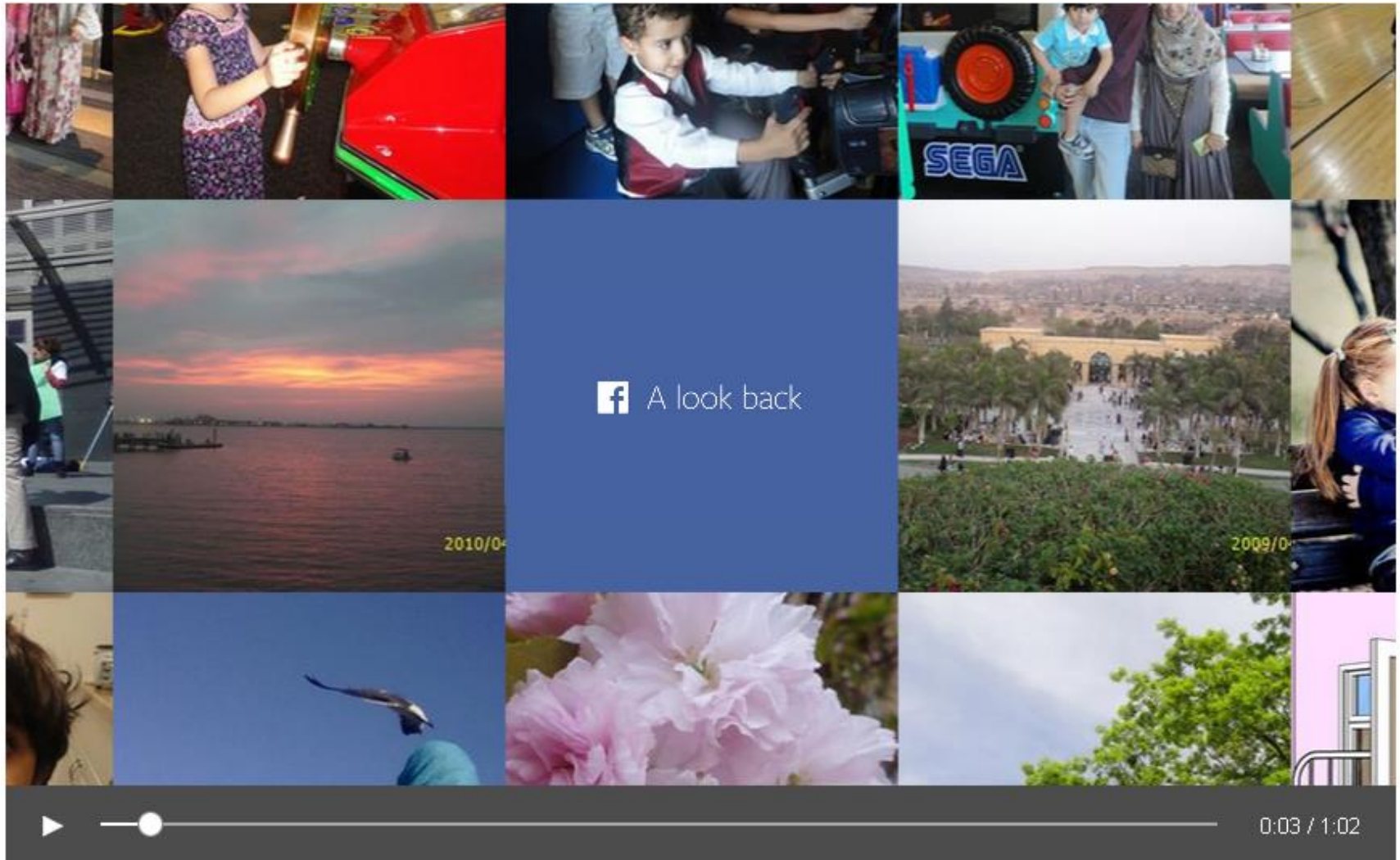
I have 1000s of posts, images, friends' posts, etc. in Facebook

The screenshot shows a Facebook profile for Yasmina Anwar. The browser address bar displays <https://www.facebook.com/yasmina.anwar>. The profile header includes the name "Yasmina Anwar" and navigation options like "Home" and "Yasmina". The main content area shows a timeline of posts from January 2011:

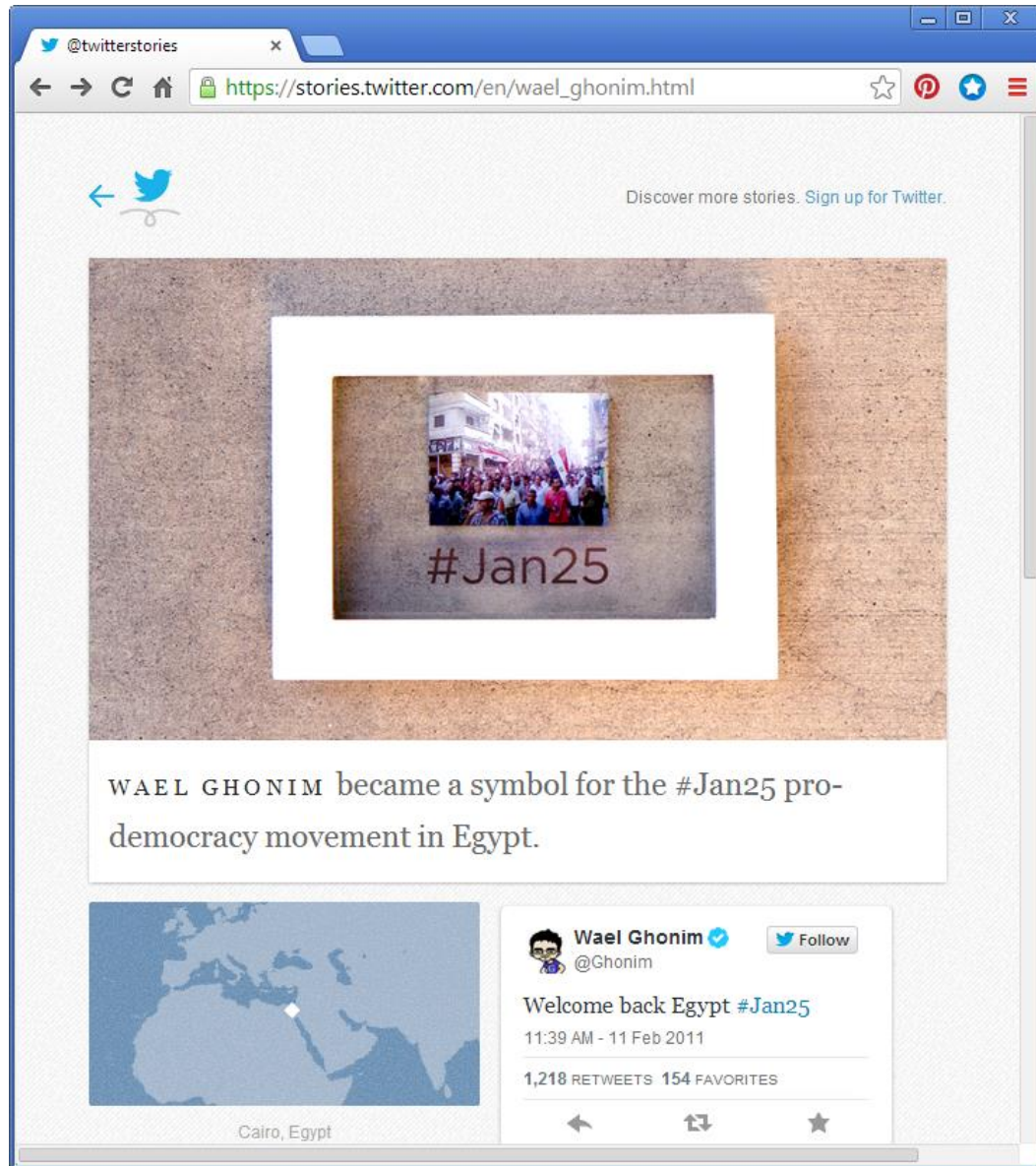
- Post 1:** A text update from Yasmina Anwar dated January 31, 2011: "Hard Day, yarab I wish I could hear a good new !!! Really, w8ing is Difficult !!!". It has 6 likes and 4 comments.
- Post 2:** A text update from Yasmina Anwar dated January 31, 2011: "Those are Egyptians !!!!! Please share!!!". It includes a video link titled "Egypt: A New Spirit of National Pride" with the text: "While Egyptians battle for their political future, social change is already sweeping the country." It has 5 likes and 3 comments.
- Post 3:** A text update from Yasmina Anwar dated January 31, 2011: "100 بوابة نسيمج // لللمرة الاولى منذ اكتوبر 2008 سعر برميل النفط الاثنى في لندن تجاوز الـ دولار مدفوعا بمخاوف مستمرة من الوضع السائد في مصر ومخاطر اضطراب حركة الملاحة في قناة السويس في قناة السويس a7san 3ashan yeqololo emshy". It has 0 likes and 0 comments.

On the left side, there are sections for "Friends" (listing Ahmed AlSum, Mohamed Hammad, Abdel Eugei, Raheena Odu), "Likes" (showing RNN NEWS and HESPRESS with a +12 count), and "Events" (listing "جمعة الغضب للثورة على الفساد والظلم والبطالة والتعذيب" and "Rabiu Family Potluck Going Away Party!"). A "Sponsored" section for QuickBooks is also visible on the right. The bottom of the page shows a "Chat (Off)" button and a URL: <https://www.facebook.com/a.php?u=http%3A%2F%2Fbn.xp...>

Facebook reduces this to 1 minute video (Facebook Look Back)



Interest from Twitter in storytelling



The screenshot shows a Twitter story interface. At the top, the browser address bar displays the URL https://stories.twitter.com/en/wael_ghonim.html. The story content features a central image of a crowd with the hashtag **#Jan25** overlaid. Below the image, the text reads: "WAEEL GHONIM became a symbol for the #Jan25 pro-democracy movement in Egypt." To the left of the text is a map of Egypt with a location pin in Cairo. To the right is the user profile for Wael Ghonim (@Ghonim), including a "Follow" button and a tweet: "Welcome back Egypt #Jan25" dated 11:39 AM - 11 Feb 2011, with 1,218 retweets and 154 favorites.

Discover more stories. [Sign up for Twitter.](#)

#Jan25

WAEEL GHONIM became a symbol for the #Jan25 pro-democracy movement in Egypt.

Cairo, Egypt

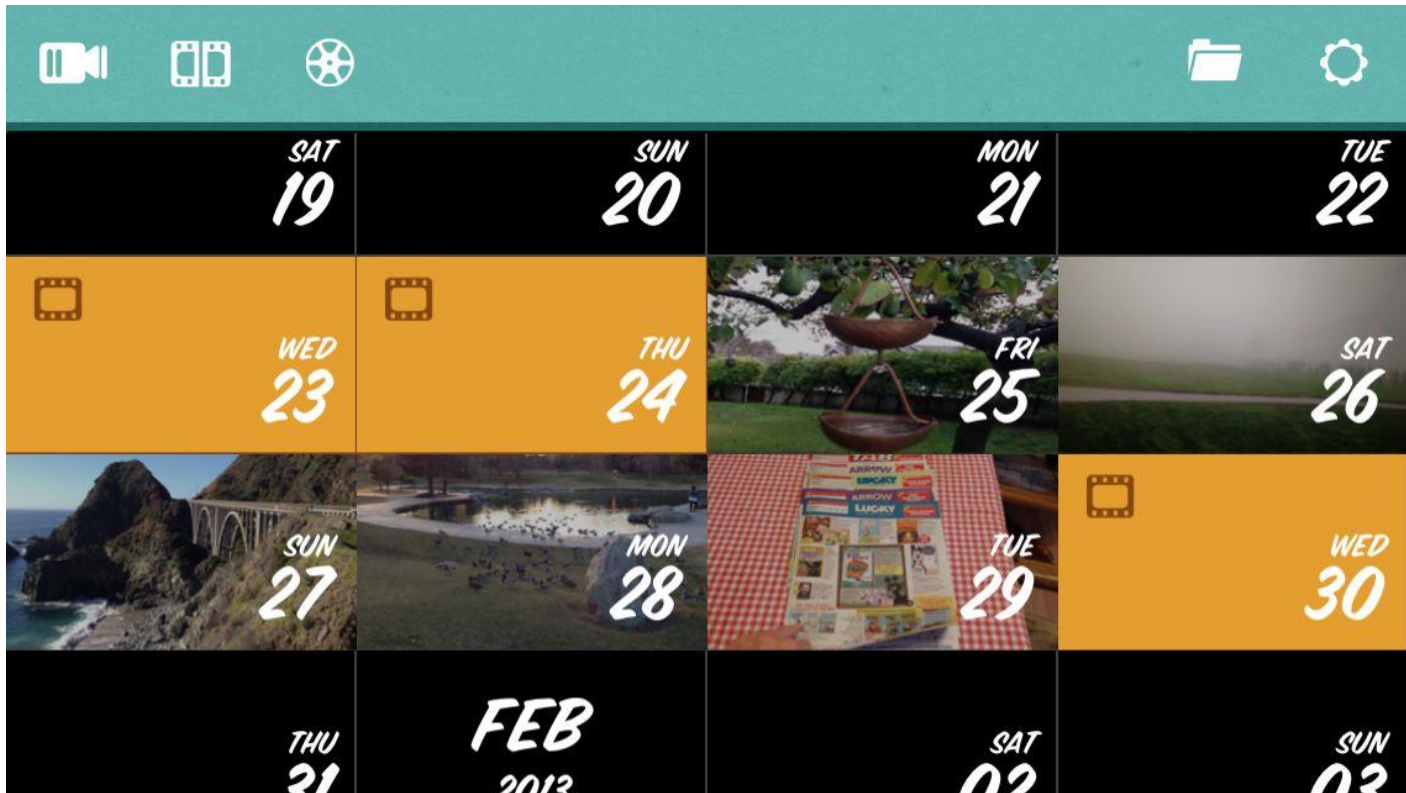
Wael Ghonim [@Ghonim](#) [Follow](#)

Welcome back Egypt #Jan25

11:39 AM - 11 Feb 2011

1,218 RETWEETS 154 FAVORITES

1 Second Everyday



What are the limitations of
storytelling services?

The Egyptian Revolution on Storify

The screenshot shows a Storify page for the story 'Egypt's Mubarak resigns - the world reacts'. The page is viewed in a browser window with the URL 'storify.com/nzherald/mu'. The Storify logo is in the top left, and navigation links for 'Search Storify', 'Create Story', and 'Login' are in the top right. The story is attributed to 'nzherald' and is 2 years old. The main text of the story reads: 'After 18 days of violent protests in the country and increasing international pressure, Egyptian president Hosni Mubarak has shocked the world with an immediate resignation.' Below the main text are two embedded tweets. The first tweet is from 'TwidderingIdiot' (@TwidderingIdiot) and says: 'Democrats, me thinks thou doth rejoiceth too soon. We know yet not, what will become of our former Middle East confidant, called Egypt.' The second tweet is from 'adelshehadeh' (@adelshehadeh) and says: 'Uninstalling dictator ... 99% complete' with a blacked-out name and the hashtags '#egypt #jan25 #tahrir #mubarak'. On the right side of the page is a profile card for 'nzherald' with 106 stories, 88 followers, and 7 following. Below the profile card is a table of total views for various sources. At the bottom right, there are 'Related stories' including 'Boston Bomber Manhunt: Photos from Instagram' and 'Boston Marathon: Photos of the Manhunt'. A vertical red 'feedback' button is on the right edge of the browser window. The footer of the page includes social media links (Like 43k, Follow @storify 98.9K followers), navigation links (Plans, Guided tour, About, Help, Blog, Jobs, Tools, API, Terms, Privacy), and a copyright notice for 2013 Storify.

Storify by nzherald 2 years ago

Egypt's Mubarak resigns - the world reacts

After 18 days of violent protests in the country and increasing international pressure, Egyptian president Hosni Mubarak has shocked the world with an immediate resignation.

TwidderingIdiot @TwidderingIdiot
Democrats, me thinks thou doth rejoiceth too soon. We know yet not, what will become of our former Middle East confidant, called Egypt.
2 YEARS AGO

0900am: Egypt's military had announced earlier that it had stepped in to "[safeguard](#)" [Egypt](#). Anti-government protesters calling for President Hosni Mubarak to step down were told their demands would be met. ^TRF

adelshehadeh @adelshehadeh
Uninstalling dictator ... 99% complete
#egypt #jan25 #tahrir #mubarak
2 YEARS AGO

nzherald
nzherald
Here to inform you with news and views from nzherald.co.nz via @TroyRF and the editorial team. Follow @nzheraldtv for the best news video.
[+ Follow](#)

106 STORIES 88 FOLLOWERS 7 FOLLOWING

Total views	19,282
nzherald.co.nz	14,796
msn.nzherald.co.nz	3,280
m.nzherald.co.nz	768
storify.com	260
mobile.nzherald.co.nz	170
other	8

Related stories

- People** Boston Bomber Manhunt: Photos from Instagram
BY PEOPLEMAG · 2 MONTHS AGO
404,333 views
- People** Boston Marathon: Photos of the Manhunt
BY PEOPLEMAG · 2 MONTHS AGO
395,744 views

Like 43k Follow @storify 98.9K followers Plans Guided tour About Help Blog Jobs Tools API Terms Privacy © 2013 Storify

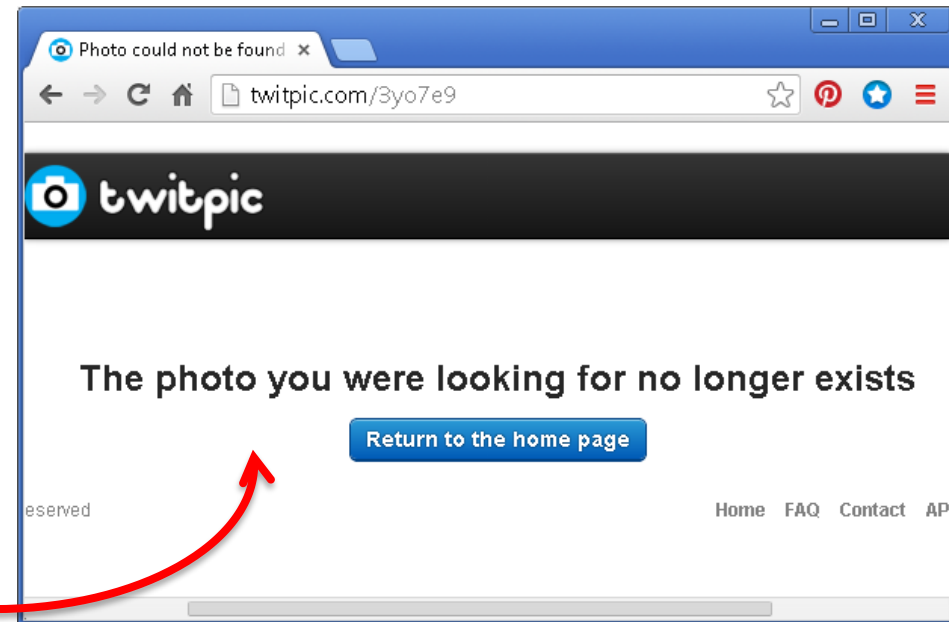
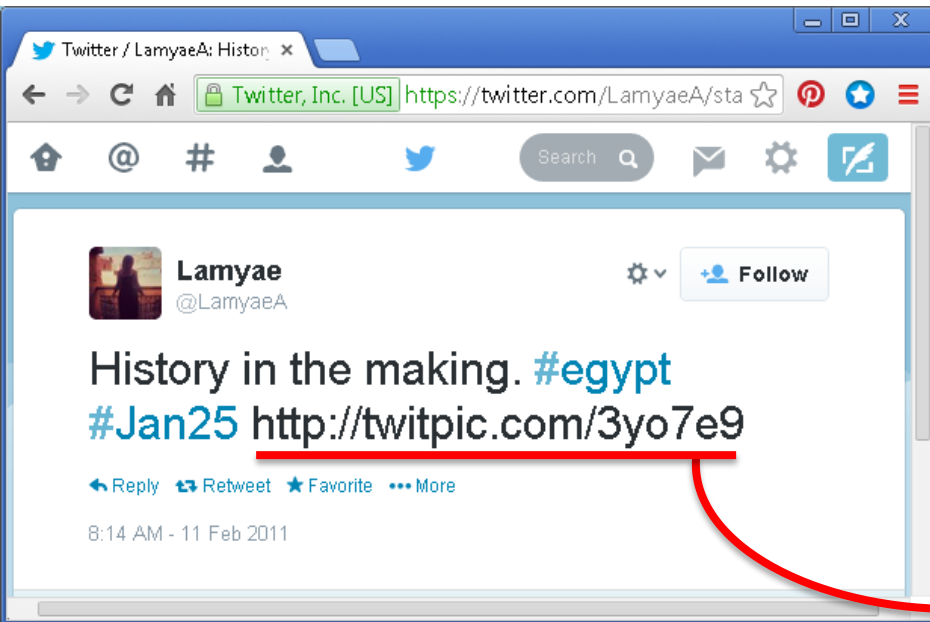
Bookmarking, not preserving!

The image shows two browser windows. The left window is a Storify page titled "Egypt's Mubarak resigns - reacts" by nzherald, posted 2 years ago. It features a tweet from TwidderingIdiot (@TwidderingIdiot) with the text "Democrats, me th... both rejoiceth too s... what will beco... former Middle East con...". Below the tweet is another tweet from adelshehadeh (@adelshehadeh) titled "Uninstalling dictator ... 99% complete" with the text "#egypt #jan25 #tahrir #mubarak". A red box highlights the text "0900am: Egypt's military had announced earlier that it had stepped in to 'safeguard' Egypt." in the Storify page. A blue arrow points from this text to a red box on the right window. The right window is a NZ Herald news article page titled "The New Zealand Herald" with the URL "www.nzherald.co.nz/world/news/article.cfm?c_id=2&objectid=1". The page displays a "404 - Page not found" error message, which is highlighted with a red box. Below the error message, it says "Sorry, the story you were looking for is no longer available." and "Explore other stories by clicking on one of the sections in the navigation bar above or check out any of the stories below." The navigation bar includes sections like National, World, Business, Sport, Technology, Entertainment, and Life & St. The footer of the NZ Herald page includes a table with the following data:

m.nzherald.co.nz	768
storify.com	260
mobile.nzherald.co.nz	170
other	8

Related stories are also listed, including "Boston Bomber Manhunt: Photos from Instagram" and "Boston Marathon: Photos of the Manhunt". The footer of the Storify page includes "Like 43k", "Follow @storify", "98.9K followers", and "© 2013 Storify".

Social media can go off-topic or disappear

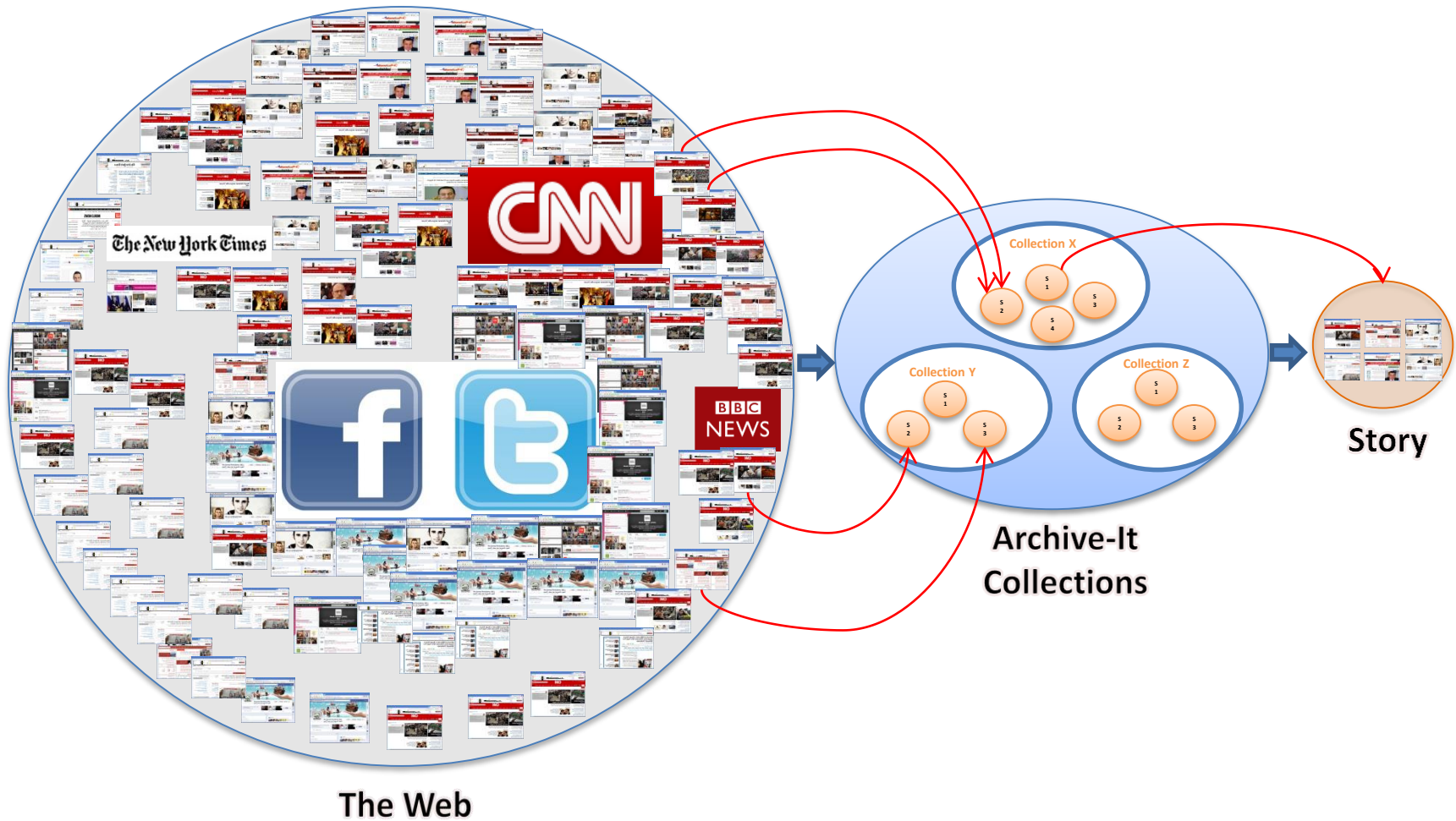


A year after publication, about 11% of content shared on social media will be gone (SalahEldeen, 2012)*

* SalahEldeen, H.M., Nelson, M.L.: Losing my revolution: How many resources shared on social media have been lost? TPD'12.

Despite these limitations, how do we
combine storytelling & archives?

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

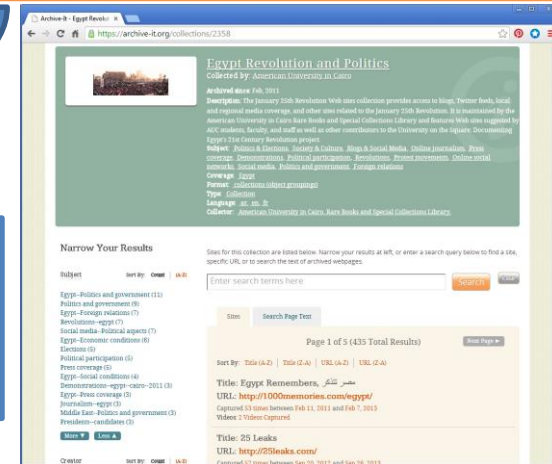


Use interface *people already know how to use* to summarize collections

Storytelling services



Archived collections



Archived enriched stories



Research Question

Can we (semi-)automatically identify, evaluate, and select candidate archived web pages from archived collections for generating stories that summarize the collection?

The archived collection has two dimensions

Archive-It - 2013 Boston | x
 https://archive-it.org/collections/3649
 Apps LaTeX Line and Pag... Save to Mendeley WordCounter | إتاحة القرآن الكريم من... إتاحة القرآن الكريم من... First lock - Tableau...

2013 Boston Marathon Bombing
 Collected by: Internet Archive, Global Events
 Archived since: Apr, 2013
 Description: This collection is comprised of contributions from curators and subject matter experts, and includes news articles, blogs, social media sites, and organizational websites related to the 2013 Boston Marathon Bombing and its aftermaths
 Subject: Spontaneous Events

Narrow Your Results

Group Sort By: Count

Boston Sports Teams Response (8)
 Institutional Responses (10)
 News Coverage (148)
 Social Media (49)

Sites Search Page Text

Page 1 of 3 (266 Total Results)

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

URL: <http://abclocal.go.com/ktrk/feature?section=news/local&id=9065470>
 Captured once on Apr 19, 2013

URL: <http://abcnews.go.com/blogs/headlines/2013/04/brothers-who-lost-legs-in-boston-bombing-reunited/>
 Captured 2 times between May 2, 2013 and May 10, 2013

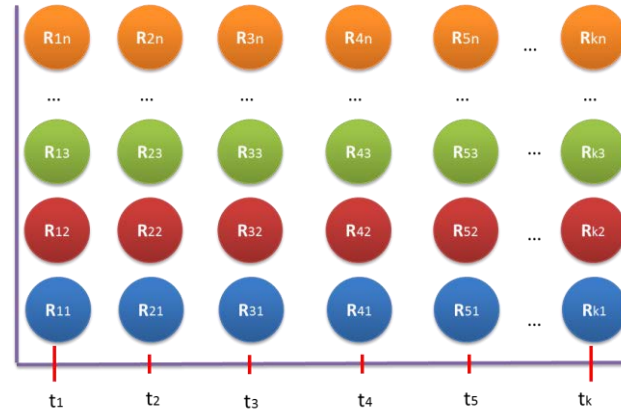
URL: <http://annajcook.blogspot.com/2013/04/a-few-more-thoughts-cats-and-flowers.html>
 Captured once on Apr 22, 2013
 Group: Social Media

URL: <http://annajcook.blogspot.com/2013/04/some-thoughts.html>
 Captured once on Apr 22, 2013
 Group: Social Media

URL: <http://aol.sportingnews.com/nba/story/2013-04-17/boston-marathon-bombing-boston-celtics-honor-victims-jersey-black-stripe/>
 Captured 2 times between Apr 19, 2013 and Apr 24, 2013
 Group: Boston Sports Teams Response

URL: <http://blog.heritage.org/2013/04/17/world-recoils-from-boston->

URI



Time

Internet Archive Wayback | x
 https://wayback.archive-it.org/2358*/http://www.bbc.co.uk/news/world/middle_east/

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

Enter Web Address: All Take Me Back

Searched for http://www.bbc.co.uk/news/world/middle_east/ 627 Results
 Look up URI in general Internet Archive web collection

* denotes when page was updated

Found 627 Captures between Feb 1, 2011 - Feb 6, 2014

2011		2012		2013		2014	
20 pages		317 pages		270 pages		12 pages	
Feb 1, 2011 *	Jan 2, 2012 *	Jan 1, 2013 *	Jan 26, 2014 *				
Feb 1, 2011 *	Jan 9, 2012 *	Jan 2, 2013 *	Jan 27, 2014 *				
Feb 2, 2011 *	Jan 16, 2012 *	Jan 3, 2013 *	Jan 28, 2014 *				
Feb 4, 2011 *	Jan 23, 2012 *	Jan 4, 2013 *	Jan 29, 2014 *				
Feb 5, 2011 *	Jan 30, 2012 *	Jan 5, 2013 *	Jan 30, 2014 *				
Feb 7, 2011 *	Feb 22, 2012 *	Jan 6, 2013 *	Jan 31, 2014 *				
Feb 9, 2011 *	Feb 23, 2012 *	Jan 7, 2013 *	Feb 1, 2014 *				
Feb 11, 2011 *	Feb 24, 2012 *	Jan 8, 2013 *	Feb 2, 2014 *				
Feb 11, 2011 *	Feb 25, 2012 *	Jan 9, 2013 *	Feb 3, 2014 *				
Feb 13, 2011 *	Feb 26, 2012 *	Jan 10, 2013 *	Feb 4, 2014 *				

The Two Dimensions Applied to Stories

Time

same

different

URI

same

Fixed Page – Fixed Time:
differences in GeolP,
mobile, etc. (Kelly, 2013)

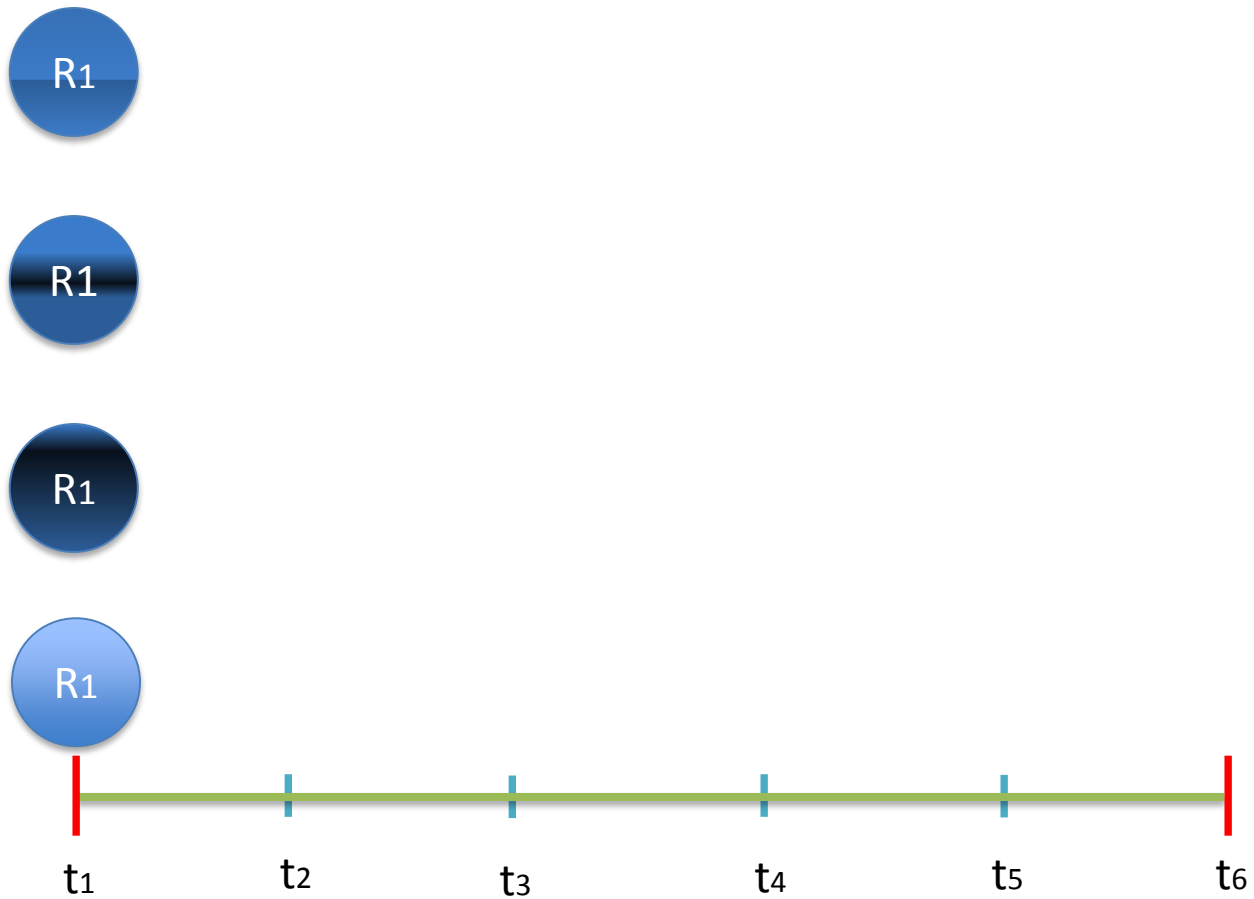
Fixed Page – Sliding Time:
evolution of a single page
(or domain) through time

different

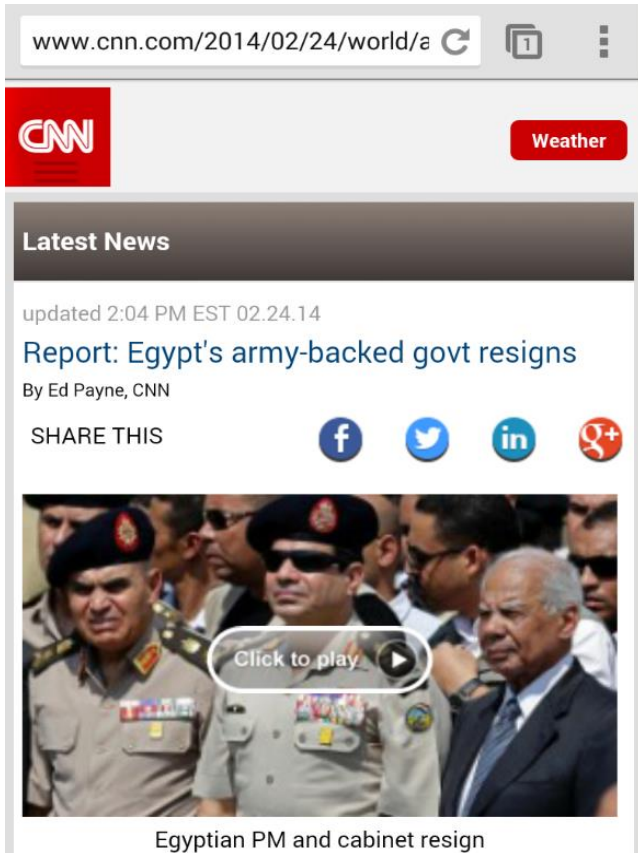
Sliding Page – Fixed Time:
different perspectives on a
point in time

Sliding Page – Sliding Time:
broadest possible coverage
of a collection

Fixed Pages, Fixed Time



Fixed Page, Fixed Time



Android Chrome user-agent

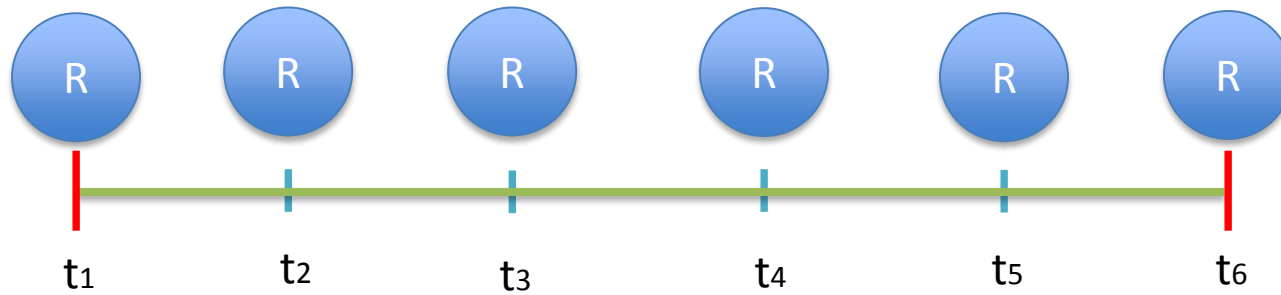
http://www.cnn.com/2014/02/24/world/africa/egypt-politics/index.html?hpt=wo_c2



A desktop Chrome user-agent

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

Fixed Page, Sliding Time



Wayback Machine screenshot of a news article titled "Egypt says it's shutting down mobile phone networks" dated Feb 1, 2011. The article features a photograph of a military tank in the foreground and a crowd of people in the background. The text discusses the government's decision to shut down mobile networks and mentions a gathering in Cairo's Tahrir Square.

Wayback Machine screenshot of a news article titled "Egypt crisis: Mubarak says he'll won't seek re-election" dated Feb 1, 2011. The article features a photograph of President Hosni Mubarak speaking. The text reports his statement that he will not seek re-election and mentions a gathering in front of the Egyptian State TV.

Wayback Machine screenshot of a news article titled "Egypt crisis: 3 dead, 639 injured in Cairo, government officials say" dated Feb 2, 2011. The article features a photograph of a large fire at night. The text reports the death toll and injuries in Cairo, mentioning a gathering in front of the Cairo's Egyptian Museum.

Wayback Machine screenshot of a news article titled "Egypt crisis: Protesters mass in square, at mosque; U.S. plans no flights out" dated Feb 4, 2011. The article features a photograph of protesters holding Egyptian flags. The text reports on the mass protests in Tahrir Square and the U.S. decision to suspend flights to Egypt.

Wayback Machine screenshot of a news article titled "Egypt crisis: Report of assassination attempt not confirmed, diplomat says" dated Feb 5, 2011. The article features a photograph of a large crowd of protesters in a square. The text reports on a suspected assassination attempt and the U.S. State Department's response.

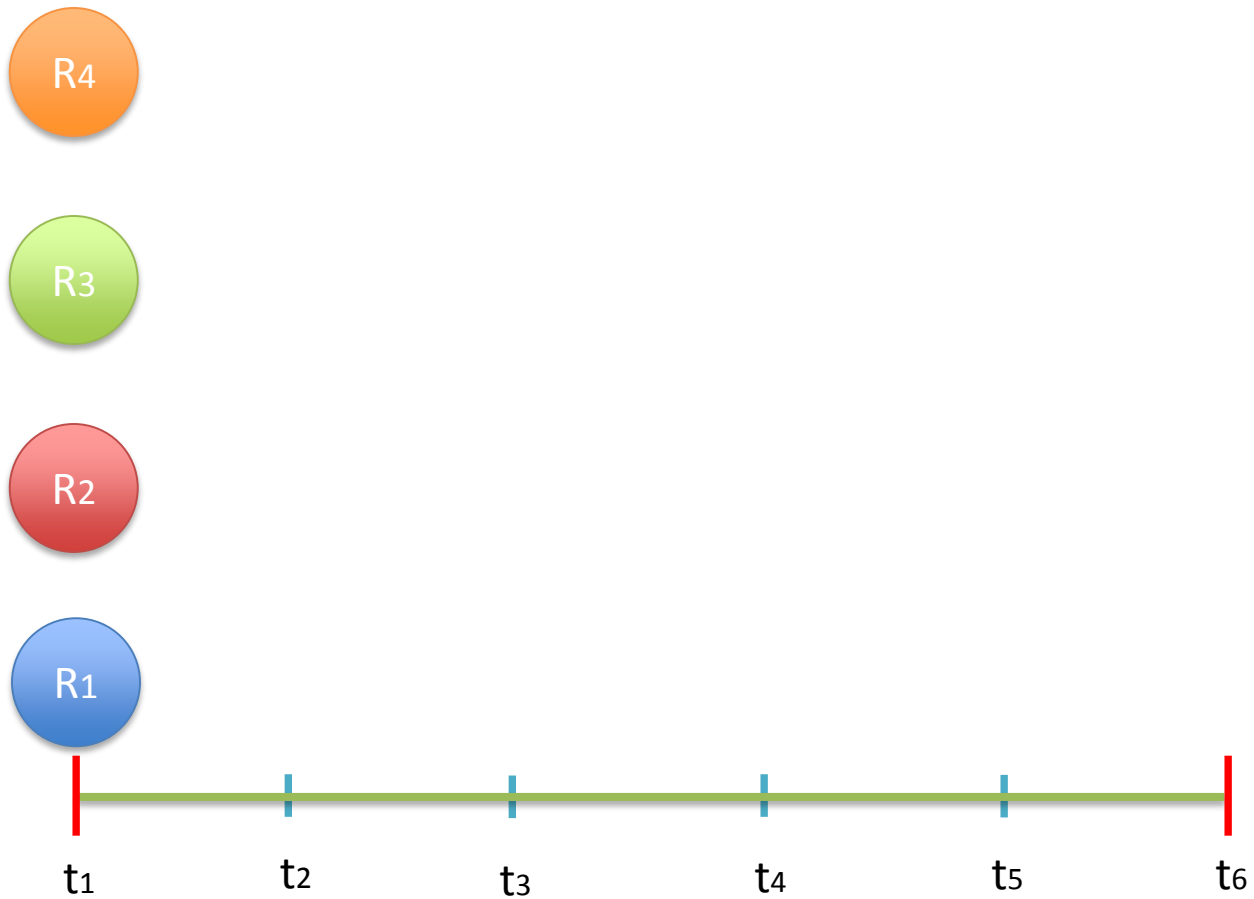
Wayback Machine screenshot of a news article titled "Egypt crisis, Day 14: Google executive released" dated Feb 7, 2011. The article features a photograph of a large crowd of protesters in a square. The text reports on the release of a Google executive and the ongoing protests.

Wayback Machine screenshot of a news article titled "Egypt transitioning, but to what?" dated Feb 9, 2011. The article features a photograph of a woman holding a sign that says "Mubarak, you go out. I go home. The END!". The text discusses the political transition and the role of the military.

Wayback Machine screenshot of a news article titled "Egypt unrest: Protesters begin 18th day of demonstrations" dated Feb 11, 2011. The article features a photograph of a man shouting with his fist raised. The text reports on the 18th day of protests and the government's decision to deploy the military.

Wayback Machine screenshot of a news article titled "Mubarak quits: Switzerland freezes assets; Egyptians didn't talk with Obama" dated Feb 11, 2011. The article features a photograph of a large crowd of protesters celebrating. The text reports on Mubarak's resignation, the freezing of his assets, and the lack of communication with Obama.

Sliding Page, Fixed Time



Wayback Machine screenshot of a CNN article from February 11, 2011, titled "Mubarak's refusal to resign a 'snub to Obama' (VIDEO)". The article reports that Egyptian President Hosni Mubarak refused to resign Thursday, despite international pressure. It includes a video player and social media sharing options.

Wayback Machine screenshot of the CNN homepage from February 11, 2011. The main headline reads "Mubarak is gone, joy in Cairo streets". A large image shows a crowd of people celebrating with Egyptian flags. The page includes a "Breaking news" section and a "Latest news" sidebar.

Wayback Machine screenshot of a CNN article from February 11, 2011, titled "Mubarak quits: Switzerland freezes assets; Egyptians didn't talk with Obama". The article discusses the international reaction to Mubarak's resignation, including Switzerland freezing his assets and the lack of direct communication with Obama. It features a large image of a crowd and a "Top stories" section.

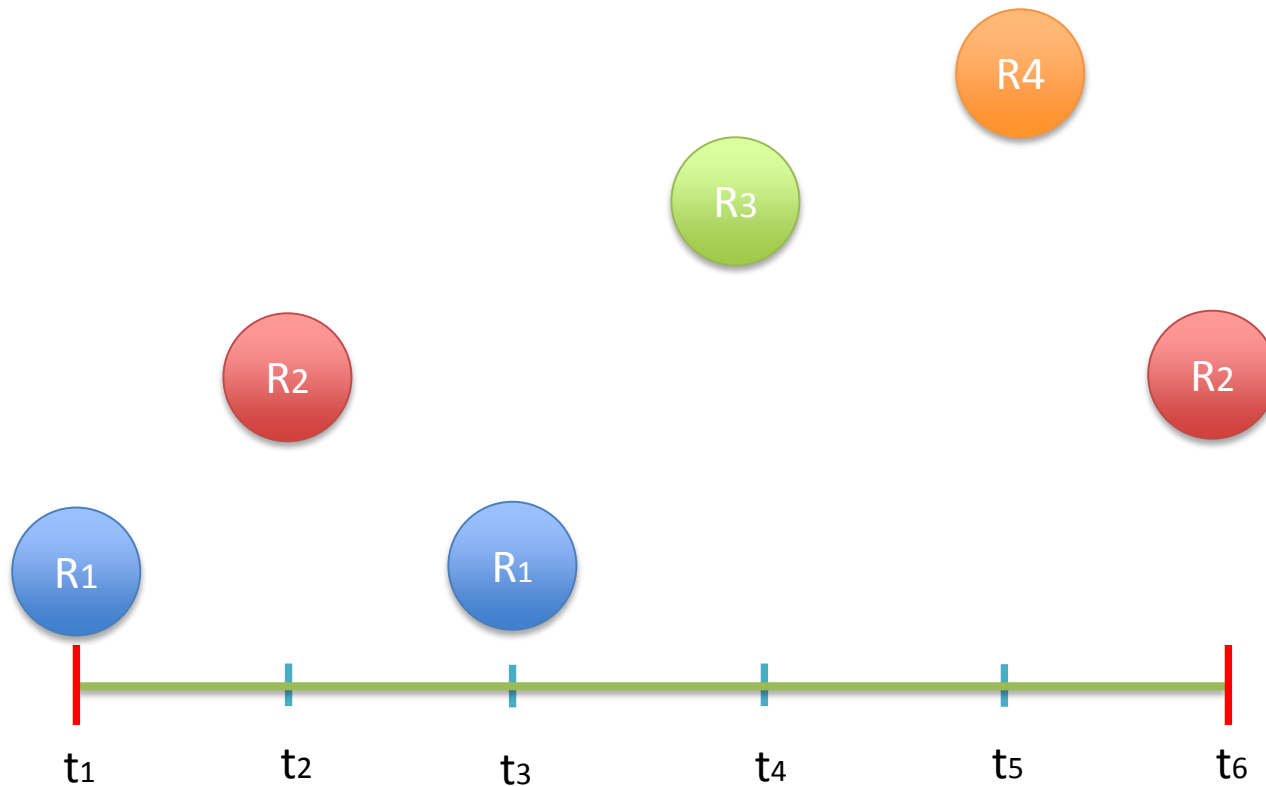
Feb. 11, 2011
Mubarak resigns

BBC News screenshot from February 11, 2011, titled "Egypt's Mubarak resigns as leader". The article states that Hosni Mubarak has stepped down as president of Egypt after weeks of protest. It includes a photo of Mubarak and a "Top Stories" section.

Blog - The Arabist screenshot from February 11, 2011, titled "Tonight we party". The blog features a cartoon illustration of people celebrating and a photo of a crowd celebrating with Egyptian flags. It includes a "Recent Tweets" section and a "Subscribe" button.

Wayback Machine screenshot of a GlobalPost article from February 11, 2011, titled "Mubarak resigns: world reactions (UPDATES) (VIDEO)". The article provides updates on world leaders' reactions to Mubarak's resignation. It includes a video player and a "Top stories" section.

Sliding Page, Sliding Time



Wayback Machine: <http://www.globalvoicesonline.org/2011/01/25/egypt-the-january-25-demonstrations-in-photos/>

Global Voices

English
Subscribe
Global Voices is an international community report on blogs and citizen media from around the world.

Search Countries Topics Contributors

Egypt: The January 25 Demonstrations in Photographs

Posted 25 January 2011 20:51 GMT

This post is part of our special coverage of Egypt Protests 2011.

A picture is worth a thousand tweets, especially when Twitter is blocked in Egypt in order to halt the transfer of information about the ongoing demonstrations in Egypt today.

Egyptian blogger, Zeinobia, decided to take her camera and go down to the streets and take some photographs and upload them to her Flickr account. Here are some of her photographs:

More stories about Middle East & North Africa

Morocco: Online Debates Rage Days Ahead of Parliamentary Elections

Tunisia: Execution of a Tunisian in Iraq Divides Netizens

Israel: Netanyahu Threatens to Replace Intems with Indian Doctors

Jan 25

Wayback Machine: <http://www.cnn.com/2011/01/25/egypt/index.html>

CNN

Egyptians brace for Friday protests as internet, messaging disrupted

by the CNN Wire Staff
January 27, 2011 11:07 p.m. EST

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

Part of available coverage on Unrest in Egypt

The Brotherhood: For good or ill

U.S. regional aid bill

Will Egypt's unrest spread?

STORY HIGHLIGHTS
NEW: Vice President Biden says Egypt's Mubarak should be more

Cairo, Egypt (CNN) — Hours ahead of what are expected to be massive displays of anti-government ferment across the world's most populous Arab nation, the internet went dark in parts of Egypt

Jan 27

Wayback Machine: <http://www.blogs.cnn.com/category/world-egypt-world-tales/>

Egypt says it's shutting down mobile phone networks

Recent Posts

Egypt says it's shutting down mobile phone networks

Egyptian government tightens grip ahead of protests

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

How does not yet buying the spirits of auto-lovers

60 people in 50 days: Girl disappeared on spring break

60 people in 50 days: Teen vanished while visiting sister

Federal judge says key parts of health care reform unconstitutional

Egypt: How did it get to this point?

Jan 31

Wayback Machine: <http://news.blogs.cnn.com/category/world-egypt-world-tales/>

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

February 2nd, 2011

Read full coverage of the unrest in Egypt updated continually by CNN reporters worldwide. Send your photos and video to Report 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 canceling in New York

Shooting reported in California after

3 dead, 639 injured in C government officials say

Siccom: It's National Signing Day!

Wednesday's intriguing people

Gotta Watch: Google art view: Froze redux

Feb 2

BBC News - Middle East

Wayback Machine: <http://www.bbc.com/news/world-middle-east>

'Day of departure' rally in Cairo

4 February 2011 Last updated at 03:20 ET

Tens of thousands of Egyptians gather again in Cairo's Tahrir Square for a 'day of departure' rally aimed at ousting President Hosni Mubarak.

Egypt unrest

Mubarak tells ABC he is 'fed up'

Violence erupts in dark site

Iran connection

Are Arab uprisings linked to 1979 Iranian revolution?

Yemen discontent

Opposition leaders warn that Yemen could be the next Tunisia

Feb 4

Wayback Machine: <http://news.blogs.cnn.com/category/world-egypt-world-tales/>

Egypt crisis, Day 14: Google executive released

February 7th, 2011

Read full coverage and examine a timeline of the unrest in Egypt updated continually by CNN reporters worldwide. Send your photos and video to Report 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 executive released

Monday's intriguing people

Assange lawyers raise specter of

Party election led to Ohio tax house

Monday quarterbacking: Super Bowl fumbles

Feb 7

BBC News - Egypt protests

Wayback Machine: <http://www.bbc.com/news/egypt>

Egypt protests: Hosni Mubarak refuses to step down

10 February 2011 Last updated at 18:22 ET

Egypt's President Hosni Mubarak has said he will stay in office and transfer all power only after September's presidential election.

President Mubarak addressed the nation in a television broadcast

Two Koreas 'at a turning point'

Manhunt closes US national park

Person held over LA arson attacks

Google tests new search home page

Features & Analysis

Science of pol

Mind over mat

Connecting in

Becoming aw

Feb 10

Wayback Machine: <http://news.blogs.cnn.com/category/world-egypt-world-tales/>

Egypt unrest: Protesters begin 18th day of demonstrations

February 11th, 2011

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

Arizona counter-sues fed over bor

Crowd at Egypt's Tahrir Square

Mubarak doesn't resign

Granddaughter wants share of Eln

Feb 11

BBC News - Egypt crisis

Wayback Machine: <http://www.bbc.com/news/egypt>

Egypt's Mubarak resigns as leader

11 February 2011 Last updated at 13:46 ET

Hosni Mubarak has stepped down as president of Egypt, after weeks of protest in Cairo and other cities.

The news was greeted with a huge outburst of joy and celebration by thousands in Cairo's Tahrir Square - the heart of the demonstrations.

Mr Mubarak ruled for 30 years, suppressing dissent and protest, and jailing opponents.

Announcing the resignation, Vice-President Omar Suleiman said the president had handed power to the army.

Mr Suleiman said on state TV that the high command of the armed forces had taken over.

Chandra Levy killer gets 60 years

Nokia and Microsoft form alliance

US man shot pair 'in cold blood'

Father 'admits he killed' twins

Features & Analysis

Day in picture

Cowboy revolt

Prince John W

Pictures of the

Feb 11

Handcrafted Stories in Storify

The screenshot shows a Storify page with the title "The Egyptian Revolution on Archive-It collection". The page features a vertical list of five news snippets, each with a small image and a headline. The snippets are categorized under "Middle East". The first snippet is about Egypt's army vows not to use force against demonstrators. The second is about Wael Nawara, an opposition leader, telling the BBC that Mr. Mubarak's promises did not satisfy protesters. The third is about three people killed and hundreds wounded in the Egyptian capital. The fourth is about tens of thousands of Egyptians gathering in Cairo's Tahrir Square for a "day of departure" rally. The fifth is about Egyptian President Hosni Mubarak holding talks with key ministers on reviving the economy as protesters keep up pressure from Tahrir Square. The page includes a "Feedback" button on the left and a footer with "Like 59k" and "Follow @storify 119K followers".

Sliding Page, Fixed Time

http://storify.com/yasmina_anwar/the-egyptian-revolution-on-archive-it-collection

The screenshot shows a Storify page with the title "The Egyptian Revolution Story from Archive-It: Heterogeneous". The page features a vertical list of news snippets, each with a small image and a headline. The snippets are categorized under "Middle East". The first snippet is about Egyptians bracing for Friday protests as internet and messaging are disrupted. The second is about Egypt saying it's shutting down mobile phone networks. The third is about an Egypt crisis with 3 dead and 639 injured in Cairo, according to government officials. The fourth is about Egyptian Chronicles, mentioning Wael Ghorim as Google's head of marketing in MENA. The fifth is about a "Day of departure" rally in Cairo, where huge crowds have occupied the main square in the Egyptian capital. The page includes a "Feedback" button on the left and a footer with "Like 64k" and "Follow @storify 127K followers".

Sliding Page, Sliding Time

http://storify.com/yasmina_anwar/the-egyptian-revolution-story-from-archive-it-hete

Other Interfaces Possible...

Mozilla Firefox

File Edit View History Bookmarks Tools Help

Replaying Story of Egyptian Revolution

www.cs.odu.edu/~yasmin/explore

Egyptian Revolution

Timeline Flipbook List Map View in Dipity Search

CNN

Egypt braces for anti-government protest
Jan 25, 2011

Egypt's armed forces: 'No violence' ag...
Jan 31, 2011

Protesters, tanks in standoff
Feb 5, 2011

Egypt's Mubarak resigns after 30-year ...
Feb 11, 2011

Obama to Mubarak: Deliver on vows
Jan 28, 2011

Egypt shuts down internet as 'march of ...'
Feb 1, 2011

Men at the barricades: We won't quit
Feb 7, 2011

Mubarak asks government to resign amid
Jan 28, 2011

Clashes in Cairo
Feb 2, 2011

Out of prison, back to Tahrir Square
Feb 9, 2011

Will Egypt follow Tunisia's lead?
Jan 25, 2011

Without police, Cairo 'like Wild West'
Jan 29, 2011

Military steps in between Mubarak toes, ...
Feb 3, 2011

Mubarak speech leaves world wondering
Feb 10, 2011

Americans getting out of Egypt
Jan 31, 2011

'Day of Departure' protest tense but m...
Feb 4, 2011

Powered by: dipity

<http://www.dipity.com/YasminAlnoamany/Jan-25-Egyptian-Revolution/>

Yasmin's Research Status

1. Establishing a baseline
 - a. For archived collections (Visualizing Digital Collections at Archive-It, JCDL 2012)
 - b. For human stories (Characteristics of Social Media Stories, TPD 2015)
2. Calculate the aboutness of the collections
3. Find off-topic URIs (Detecting Off-Topic Pages in Web Archives, TPD 2015)
4. Identify the best k mementos that summarize the collection for desired story type
5. Evaluate & visualize the generated stories

What about using stories to seed web
archive collections?

Creating collections is *hard*.

Requires combination of awareness,
domain knowledge, crowdsourcing, etc.

Ukrainian Themed Collections, Prior to February 2014

Archive-It - Explore Archived Content

https://archive-it.org/explore?q=ukraine

Requested: 1996 2011 01/29/2014

Archive-It - Explore Archived C...

Date	Sort By: Count	(A-Z)
2009-06-03 00:00:00 (1)		

Collector	Sort By: Count	(A-Z)
Georgetown University (3)		

Ukraine fragile sites 2/08

Collected by: Georgetown University
Archived since: Feb, 2008
Description: List of sites about Ukrainian issues, most coming from **Ukraine** itself.
Subject: Society & Culture
Creator: Jenia Ustinova
Collector: Georgetown University

Ukrainian Blog

Collected by: Georgetown University
Archived since: Mar, 2008
Description: Blogs in **Ukraine**
Subject: Society & Culture
Creator: Jenia Ustinova
Collector: Georgetown University

Ukrainian Websites

Collected by: Georgetown University
Archived since: Jun, 2009
Description: A collection of websites illustrating the turbulence of the civil society in **Ukraine**. 2009.
Subject: Society & Culture, Government, Politics & Elections
Creator: Leonid Godunov
Date: 2009-06-03 00:00:00

Ukrainian Conflict Collection Started February 2014

Archive-It - Ukraine Conflict x +

https://archive-it.org/collections/4399

Requested: 1996 2011 01/07/2015

ARCHIVE-IT

HOME EXPLORE LEARN MORE CONTACT US

The leading web archiving service for collecting and accessing cultural heritage on the web
Built at the Internet Archive

Explore >> Internet Archive Global Events >> Ukraine Conflict



Ukraine Conflict

Collected by: [Internet Archive Global Events](#)

Archived since: Feb, 2014

Description: This collection seeks to document conflict in Ukraine as it progresses. Contributions to this collection were made by the Archive-It team and subject matter experts from Stanford University, Global Investigative News Network, Harvard Ukrainian Research Institute, and Library of Congress in the fields of Investigative Journalism, Russian, and Eurasian studies. Content includes news outlets, social media, blogs, and government websites. Sites are in English, Ukrainian, Russian, and other languages.

Subject: [Government](#), [Spontaneous Events](#)

Narrow Your Results

Sites 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.

Contributor Sort By: Count (A-Z)

Crowdsourcing Collection of Seed URIs

Twitter / archiveitorg: RIP Nelson Mandela. We have ...

Requested: 1996 2011 01/29/2014

Twitter / archiveitorg: RIP Nels...

Search

 **Archive-It** @archiveitorg 

RIP Nelson Mandela. We have been archiving sites since July, and are looking for additional URLs. Please nominate!
docs.google.com/forms/d/1oWlnr...

 Reply  Retweet  Favorited  More

8 RETWEETS 2 FAVORITES

4:42 PM - 6 Dec 2013

What about social media, especially
curated social media?

(late January, 2014)

Riots persist in Kyiv (with images, tweets) · TheInkBee · Storify

storify.com/TheInkBee/riots-persist-in-kyiv

Requested: 1996 2011 01/29/2014

Riots persist in Kyiv (with imag... +

Storify by livefyre

Riots persist in Kyiv

What started out as peaceful protests in November of 2013 is now increasingly violent and threatens to tear apart an entire Nation: Ukraine is experiencing political turmoil and it is uncertain whether this conflict will be resolved any time soon.

Feedback

Media student. Writer. Avid Film&TV viewer. Fluent in Spanish, English, and French. ~All laughs and no sleep
YYZ-YUL-YOW
http://storify.com/TheInkBee

+ Follow

1 STORY
5 FOLLOWERS
63 FOLLOWING

Total views 17
other 17

Don't miss

Amid clashes, unclear roadmap for peace in Ukraine
BY STORIFY.COM

Like 64k **Follow @storify** 126K followers

Report Abuse VIP Guided tour About Help Blog Jobs Tools API Terms Privacy © 2014 Storify

<https://storify.com/TheInkBee/riots-persist-in-kyiv>

(late January, 2014)

Protests In Kiev Turn Violent (with images, tweets) · jesseoriordan · Storify

storify.com/jesseoriordan/ukraine-fight-for-their-rights

Requested: 1996 2011 01/29/2014

Protests In Kiev Turn Violent

Ukrainians fight for their rights as Yanukovich and the government try to take away civil liberties with new repressive laws which results in violent clashes between protesters and riot police.

For the last two months Ukraine's capital, Kiev, has been flooded with protesters opposing the current President Viktor Yanukovich choice to not sign with the EU. Thousands of protesters have set up camps in the downtown of the city forming little villages for the protesters to sleep, eat, go to the bathroom and receive medical care after clashes with riot police. Several government buildings have also been taken over by the protesters, including the city hall. The protests began out peaceful, mainly consisting of speeches. But the violence has been escalating until a few days ago when the government tried to rush in new laws, totally disregarding civil liberties and making it jail-able offences to blockade public buildings, wear masks or

Feedback

http://storify.com/jesseoriordan


+ Follow

1	0	1
STORY	FOLLOWERS	FOLLOWING

Total views 46

storify.com	44
other	2

Don't miss

 **Ukraine Live: Police clash with anti-government protesters**
BY STORIFY.COM

Like 64k Follow @storify 126K followers

Report Abuse VIP Guided tour About Help Blog Jobs Tools API Terms Privacy © 2014 Storify

<https://storify.com/jesseoriordan/ukraine-fight-for-their-rights>

(late January, 2014)

The screenshot shows a web browser window displaying a Storify page. The browser's address bar shows the URL storify.com/jesseoriordan/ukraine-fight-for-their-rights. The page title is "Protests In Kiev Turn Violent (with images, tweets) · jesseoriordan · Storify". The Storify interface includes a search bar, "Create Story" and "Login" buttons, and a "Feedback" button on the left. The main content area features a headline "Kiev looks like a scene from an apocalyptic movie." followed by a tweet from Shaun Walker titled "Three protesters killed as Kiev riots turn deadly". Below this is a paragraph: "In the article it talks about how the leaders of the protesters and opposition call for the President to call a snap election to stop the violence." A second tweet from Carl Bildt (@carbildt) is shown, dated 6:14 AM - 28 Jan 2014, with the text: "I warmly welcome the decision in Ukraine Rada some minutes ago to cancel the package of repressive laws. Hope for further steps." A "Don't miss" sidebar on the right lists related articles. At the bottom, there are social media interaction buttons (Like, Follow) and a footer with navigation links and copyright information.

Protests In Kiev Turn Violent (with images, tweets) · jesseoriordan · Storify

storify.com/jesseoriordan/ukraine-fight-for-their-rights

Requested: 1996 2011 01/29/2014

Protests In Kiev Turn Violent (w... +

Storify by livefyre

Search Storify Create Story Login

Kiev looks like a scene from an apocalyptic movie.

Three protesters killed as Kiev riots turn deadly

At least three people died in a day of violence in Kiev on Wednesday, as an opposition leader said he was willing to face "a bullet in th...

SHAUN WALKER

In the article it talks about how the leaders of the protesters and opposition call for the President to call a snap election to stop the violence.

Carl Bildt @carbildt

I warmly welcome the decision in Ukraine Rada some minutes ago to cancel the package of repressive laws. Hope for further steps.

6:14 AM - 28 Jan 2014

Don't miss

Ukraine Live: Police clash with anti-government protesters BY STORIFY.COM

Amid clashes, unclear roadmap for peace in Ukraine BY STORIFY.COM

Pro-E.U. protesters clash with police in Ukraine BY WASHINGTONPOST.COM

Like 64k Follow @storify 126K followers

Report Abuse VIP Guided tour About Help Blog Jobs Tools API Terms Privacy © 2014 Storify

Compelling narratives are created by knowledgeable (and biased) users arranging news stories, images, social media, commentary, etc.

Intuitively, this is a "good story". How do we quantify this?

<https://storify.com/jesseoriordan/ukraine-fight-for-their-rights>

Characteristics of Social Media Stories

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

Old Dominion University
Web Science and Digital Libraries Group
<http://ws-dl.cs.odu.edu/>, @WebSciDL

**TPDL 2015
16.09.2015, Poznań, Poland**

<http://www.cs.odu.edu/~mln/pubs/tpdl-2015/tpdl-2015-stories.pdf>

“Storytelling” is becoming a popular technique in social media



Facebook
Timeline



Bundlr



Pinterest



Story creation interface on Storify

Title

Egyptian Revolution "The beginning of the story"

Text Element

Khaled Saeed (6 June 2010) :Khaled Saeed was beaten and tortured to death, by Egyptian police.



"We Are All Khaled Said": Will the Revolution Come to Egypt?

Khaled Said, a small businessman in the historic Egyptian city of Alexandria, was dragged from an Internet café by police and beaten to d...

MIKE GIGLIO



Mina Fayek @minafayek

Khaled Said: The face that launched a revolution
english.ahram.org.eg/NewsContent/1/...

WED, MAR 04 2015 04:15:14

Web Elements



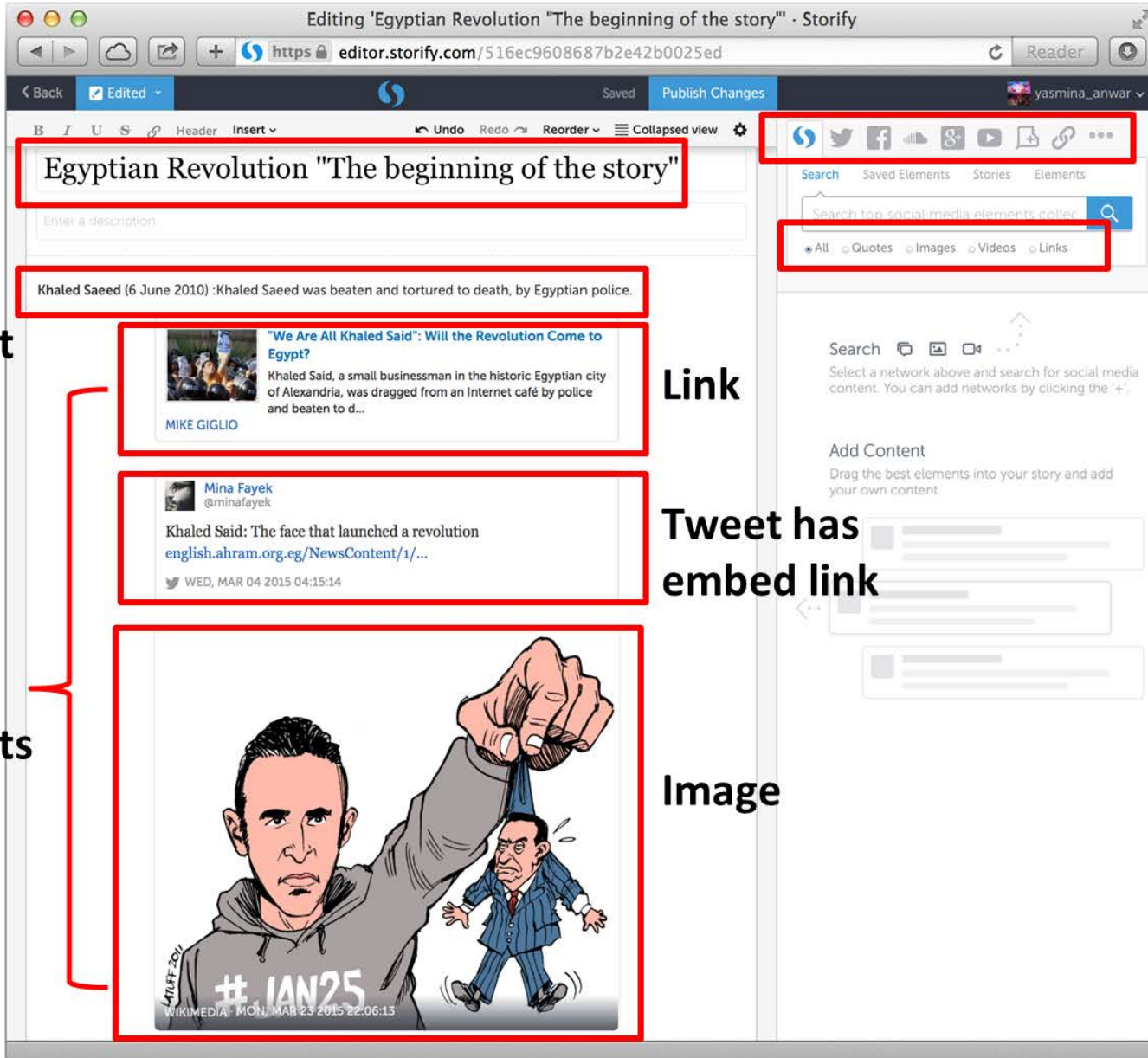
Link

Tweet has embed link

Image

Artifacts to be searched

Types of elements



Research Question

We'd like to automatically generate stories, but we don't yet know what constitutes a "good" story.

Q: What are the structural characteristics of popular (i.e., receiving the most views) human-generated stories?

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

- This story has 31 resources

The screenshot shows a Storify page for the event 'Boston Marathon Bombing'. The page is titled 'Boston Marathon Bombing' and is a collection of tweets. Three tweets are highlighted with red boxes and numbered 1, 2, and 3. The first tweet is from The Boston Globe (@BostonGlobe) retweeting a tweet from @northeastern about rumors of suspicious packages. The second tweet is from Matthew Keys (@MatthewKeysLive) reporting that the Boston chapter of the American Red Cross says no need for blood donations at this time. The third tweet is from The Associated Press (@AP) reporting that the Navy sends a bomb team to Boston to assist with an explosion investigation. The page also shows a 'Feedback' button on the left, a 'Follow' button for the Storify page, and a sidebar with 'More from yahoo news' and a table of views.

Source	Views
storify.com	1,946
theandd.com	1,626
24saltake.com	131
other	38

Article Title	By	Time Ago	Views	Likes	Comments
Metro North train details in NYC: 4 dead	BY YAHOO NEWS	17 DAYS AGO	42,429	1	like
Thanksgiving liveblog	BY YAHOO NEWS	18 DAYS AGO	20,236	2	likes - 1 comment
Los Angeles airport shooting: TSA agent down	BY YAHOO NEWS	2 MONTHS AGO	63,018	4	likes - 3 comments
Halloween live feed	BY YAHOO NEWS	2 MONTHS AGO	50,315	7	likes

What are the types of resources that compose a story?

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

The image shows a screenshot of a Storify story page titled "Egyptian Revolution of 2011". The page features a header with the Storify logo and a search bar. Below the header, there is a section titled "#Mursi is president" with a timestamp of "1 YEAR AGO". The main content area contains several paragraphs of text. A red box highlights a quote: "Since yesterday, protesters have been cleaning & re-painting Tahrir square streets & surrounding areas. This photo shows a female disabled protester painting part of the Tahrir square." A blue box labeled "Quotes" has an arrow pointing to this highlighted text. Below the text, there is a video player showing a man speaking. A red box highlights the video player, and a blue box labeled "Video" has an arrow pointing to it. The video player shows a play button and a progress bar. At the bottom of the page, there are social media sharing options and a footer with the year 2013.

What are the most frequently used domains?

- 90% twitter.com
- 7% instagram.com
- 3% facebook.com

The screenshot shows a Storify page titled "Boston Marathon Bombing" by Yahoo News. The page features three tweets highlighted with red boxes, each labeled with an orange "Twitter.com" sticker. The first tweet is from The Boston Globe (@BostonGlobe) retweeting @northeastern, stating that rumors about suspicious packages are false. The second tweet is from Matthew Keys (@MatthewKeysLive) reporting that the Boston chapter of the American Red Cross says no need for blood donations at this time. The third tweet is from The Associated Press (@AP) reporting that the Navy sends a bomb team to Boston to assist with an explosion investigation. The page also shows a "Follow" button for Yahoo News, statistics for stories, followers, and following, and a "More from yahoo news" section with links to other news items.

The popularity of the most used domains in the stories

The screenshot shows a Storify page titled "@RichardGrenell tweets al". The URL is storify.com/RuBegonia/richardgrenell-tweets-about-cnn-as-story-unfolds. The page features three tweets from Richard Grenell (@RichardGrenell) dated October 8, 2013. The first tweet, posted at 9:08 PM, discusses CNN's report on aid to Egypt and mentions @jimsciutto. The second tweet, at 9:24 PM, questions the White House's stance and CNN's reporting. The third tweet, at 9:37 PM, notes a change in CNN's headline regarding military aid. The Storify interface includes a search bar, a "Create Story" button, and a "Feedback" button on the left.

The screenshot shows a Storify page titled "Do You Hear the People...". The URL is storify.com/OwenCobb/do-you-hear-the-people-type-1. The page features a story about the French Revolution, including a painting of Jesus Christ and a reference to the American Revolution. The story text reads: "Revolution Facebook and YouTube campaign, created by the 3rd Party, forming events to Storm the Bastille and raid Versailles. Instead, the French revolutionaries had to settle for a secret meeting at a tennis court." Below the painting is a link to "French Revolution - Tennis Court Oath, Storming the Bastille" with a description: "The French Revolution - The Tennis Court Oath, storming the Bastille... the National Assembly, but was locked out of their regular me...". The Storify interface includes a search bar, a "Create Story" button, and a "Feedback" button on the left.

Twitter.com is ranked 9 by Alexa global rank

mr_sedivy.tripod.com is ranked 3,084,064 by Alexa global rank

What is the timespan (editing time) of the stories?

- This story covers ~4 hours
(from 3:12 PM – 7:27 PM)
- Is there a relation between the timespan and the features of the story?

The screenshot shows a Storify page for the 'Boston Marathon Bombing' event. The page is titled 'Boston Marathon Bombing' and is a collection of tweets from various sources. The tweets are arranged in a vertical timeline. The first tweet is from 'The Boston Globe' (@BostonGlobe) at 7:27 PM on April 15, 2013, stating that rumors about suspicious packages are false. The second tweet is from 'Matthew Keys' (@MatthewKeysLive) at 7:27 PM on April 15, 2013, reporting that the Boston chapter of the American Red Cross says no need for blood donations at this time. The third tweet is from 'The Associated Press' (@AP) at 7:21 PM on April 15, 2013, reporting that the Navy sends a bomb team to Boston to assist with an explosion investigation. The Storify page also includes a 'Feedback' button, a search bar, and a 'Create Story' button. On the right side, there is a 'yahoonews' profile card with statistics: 51 stories, 570 followers, and 13 following. Below the profile card, there are several news items from Yahoo News, including 'Metro North train derail in NYC: 4 dead', 'Thanksgiving liveblog', 'Los Angeles airport shooting: TSA agent down', and 'Halloween live feed'.

<http://storify.com/yahoonews/boston-marathon-bombing>

What is the “HTTP 404” rate of the resources?

The image shows two overlapping browser windows. The background window is a Storify page titled "Egypt's Mubarak resigns - reacts". The foreground window is a NZ Herald News page showing a "404 - Page not found" error. A blue arrow points from the Storify page to the error page. A red box highlights the "404 - Page not found" message in the foreground window.

404 - Page not found

Sorry, the story you were looking for is no longer available.

Explore other stories by clicking on one of the sections in the navigation bar above or check out any of the stories below.

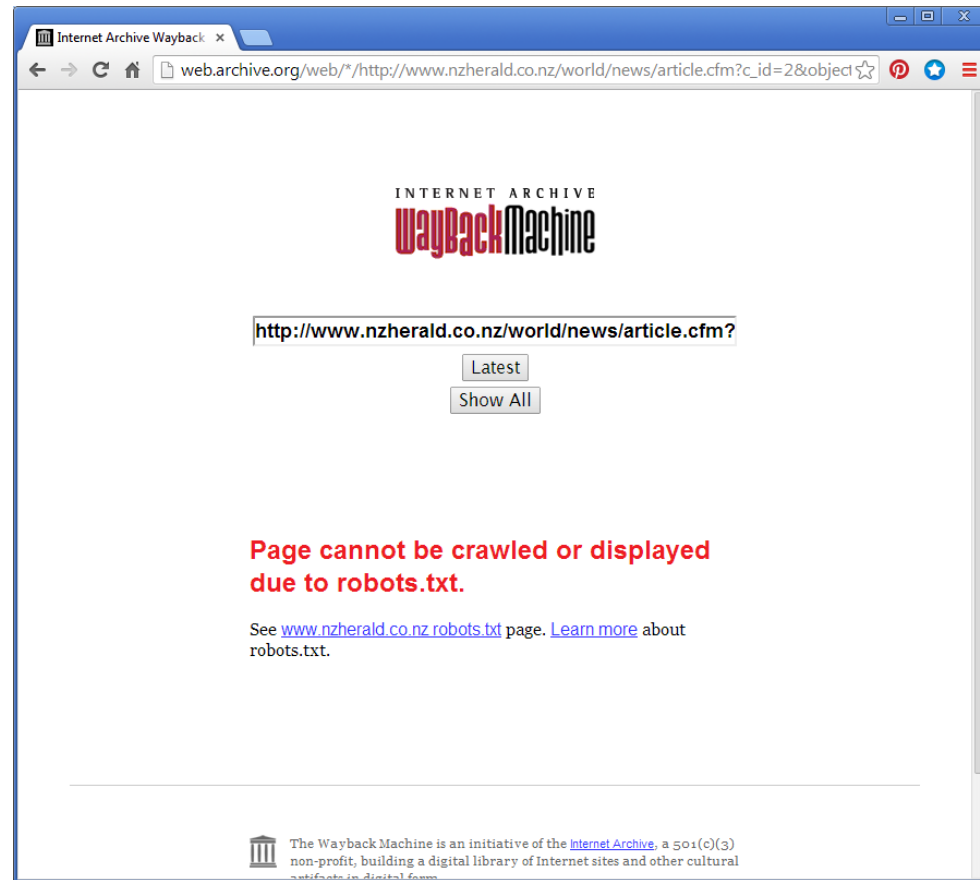
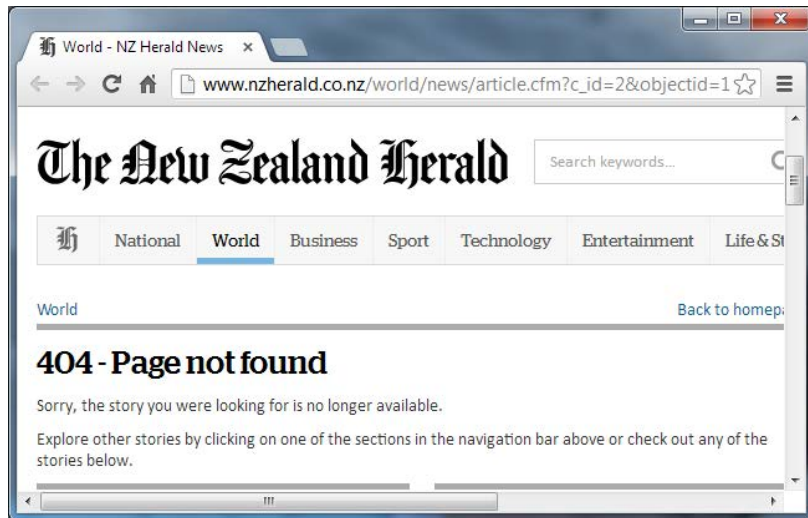
0900am: Egypt's military had announced earlier that it had stepped in to "safeguard" Egypt. Anti-government protesters calling for President Hosni Mubarak to step down were told their demands would be met. ^TRF

m.nzherald.co.nz	768
storify.com	260
mobile.nzherald.co.nz	170
other	8

Related stories

- People** Boston Bomber Manhunt: Photos from Instagram
BY PEOPLEMAG · 2 MONTHS AGO
404,333 views
- People** Boston Marathon: Photos of the Manhunt
BY PEOPLEMAG · 2 MONTHS AGO
395,744 views

Can we find these missing resources in the archives?



http://www.nzherald.co.nz/world/news/article.cfm?c_id=2&objectid=10705523

What differentiates a popular story?

As seen on [pennlive.com](#)

Va. reporter, cameraman killed during live broadcast

Alison Parker Adam Ward

WDBJ7's photographer Adam Ward and reporter Alison Parker were killed Aug. 26, 2015. Updates from social media that morning:

by [PennLive.com](#) a day ago **19,795 Views**

ABC News @ABC
NEWS
Follow

NEW: VA State Police: Alleged suspect in #WDBJ shooting transported to hospital with life-threatening injuries.
[abcn.ws/1U6gBt3](#)
9:35 AM - 26 Aug 2015

ABC News

19,795 views

Here's my Facebook movie. Find yours at <https://facebook.com/lookback/> #FacebookIs10

lookback in Facebook's 10th Birthday

Reactions and coverage of lookback...

by [Yasmina Anwar](#) 2 years ago 1 Like **64 Views**

Your Facebook life in 62 seconds - CNN.com
5 hours ago ... Facebook posted the clips, titled "A Look Back," without fanfare Tuesday to commemorate the 10th anniversary o...
BRANDON GRIGGS, CNN

Take a Look Back Into Your Facebook History - Mashable
15 hours ago ... Facebook is celebrating 10 years of existence today, and to mark the occasion, the company created a special "Look...
MASHABLE

Facebook's New 'A Look Back' Tool Highlights Your 20...

64 views

Data set construction

1. Queried Storify Search API with the top 1000 English keywords as seen by Yahoo
2. Downloaded 145,682 stories in JSON on Feb. 2015
3. Considered stories authored in 2014 or earlier, resulting in 37,486 stories
4. Eliminated stories with only zero or one elements or zero views, resulting in:
 - a. **14,568** unique stories
 - b. **10,199** unique users
 - c. **1,251,160** web and text elements

Characteristics of human-generated stories

Features	Views	Web elements	Text elements	Subscribers	Timespan in hours
25 th percentile	14	10	0	0	0.18
50 th percentile	51	23	1	4	3
75 th percentile	268	69	9	21	120
90 th percentile	1949	210	19	85	1747
Maximum	11,284,896	2,216	559	1,726,143	36,111

Characteristics of human-generated stories

Features	Views	Web elements	Text elements	Subscribers	Timespan in hours
25 th percentile	14	10	0	0	0.18
50 th percentile	51	23	1	4	3
75 th percentile	268	69	9	21	120
90 th percentile	1949	210	19	85	1747
Maximum	11,284,896	2,216	559	1,726,143	36,111

44 % of the stories have no text elements at all

What is the Average Timespan
for Stories?

What is the Average Timespan for Stories?

Intervals	Percentage	Median web elements	Median text elements	Median views (normalized by time)
0-60 seconds	14.0%	15	0	23
1-60 minutes	26.7%	19	0	53
1-24 hours	23.4%	25	5	110
1-7 days	13.5%	26	7	78
1-4 weeks	8.4%	26	9	80
1-12 months	10.9%	38	2	129
1-4 years	3.1%	56	15	156

There is nearly linear relation between the time length of the story and the number of elements

What is the Average Timespan for Stories?

Intervals	Percentage	Median web elements	Median text elements	Median views (normalized by time)
0-60 seconds	14.0%	15	0	23
1-60 minutes	26.7%	19	0	53
1-24 hours	23.4%	25	5	110
1-7 days	13.5%	26	7	78
1-4 weeks	8.4%	26	9	80
1-12 months	10.9%	38	2	129
1-4 years	3.1%	56	15	156

The story with the longest timespan in our data set covers more than 4 years and with more than 13,000 views. It had only 33 web elements and 51 total elements.

The distribution of the elements (1,251,160)

Resource	Proportion
links	70.8%
images	18.4%
text	8.1%
video	2.0%
quotes	0.7%

Text elements are relatively rare, meaning that few users choose to annotate the web elements in their story.

What are the most frequently used domains?

The most frequent domains in the stories

- Domain Canonicalization
 - `www.cnn.com` → `cnn.com`
- Dereference all the shortened URIs
 - For example: `t.co`, `bit.ly`
- **25,947** unique domains

The most frequent domains in the stories represent ~92%

Host	Frequency	Percentage	Alexa Global Rank as of 2015-03	Category
twitter.com	943,859	82.05%	8	Social media
instagram.com	45,188	3.93%	25	Photos
youtube.com	22,076	1.92%	3	Videos
facebook.com	13,930	1.21%	2	Social media
flickr.com	7,317	0.64%	126	Photos
patch.com	5,783	0.50%	2,096	News
plus.google.com	3,413	0.30%	1	Social media
tumblr.com	3,066	0.27%	31	Blogs
blogspot.com	1,857	0.16%	18	Blogs
imgur.com	1,756	0.15%	36	Photos
coolpile.com	1,706	0.15%	149,281	Entertainment
wordpress.com	1,615	0.14%	33	Blogs
giphy.com	1,055	0.09%	1,604	Photos
bbc.com	966	0.08%	156	News
lastampa.it	927	0.08%	2,440	News
pinterest.com	892	0.08%	32	Photos
softandapps.info	861	0.07%	160,980	News
photobucket.com	768	0.07%	341	Photos
nytimes.com	744	0.06%	97	News
soundcloud.com	736	0.06%	167	Audio

The most frequent domains in the stories represent ~92%

Host	Frequency	Percentage	Alexa Global Rank as of 2015-03	Category
twitter.com	943,859	82.05%	8	Social media
instagram.com	45,188	3.93%	25	Photos
youtube.com	22,076	1.92%	3	Videos
facebook.com	13,930	1.21%	2	Social media
flickr.com	7,317	0.64%	126	Photos
patch.com	5,783	0.50%	2,096	News
plus.google.com	3,413	0.30%	1	Social media
tumblr.com	3,066	0.27%	31	Blogs
blogspot.com	1,857	0.16%	18	Blogs
imgur.com	1,756	0.15%	36	Photos
coolpile.com	1,706	0.15%	149,281	Entertainment
wordpress.com	1,615	0.14%	33	Blogs
giphy.com	1,055	0.09%	1,604	Photos
bbc.com	966	0.08%	156	News
lastampa.it	927	0.08%	2,440	News
pinterest.com	892	0.08%	32	Photos
softandapps.info	861	0.07%	160,980	News
photobucket.com	768	0.07%	341	Photos
nytimes.com	744	0.06%	97	News
soundcloud.com	736	0.06%	167	Audio

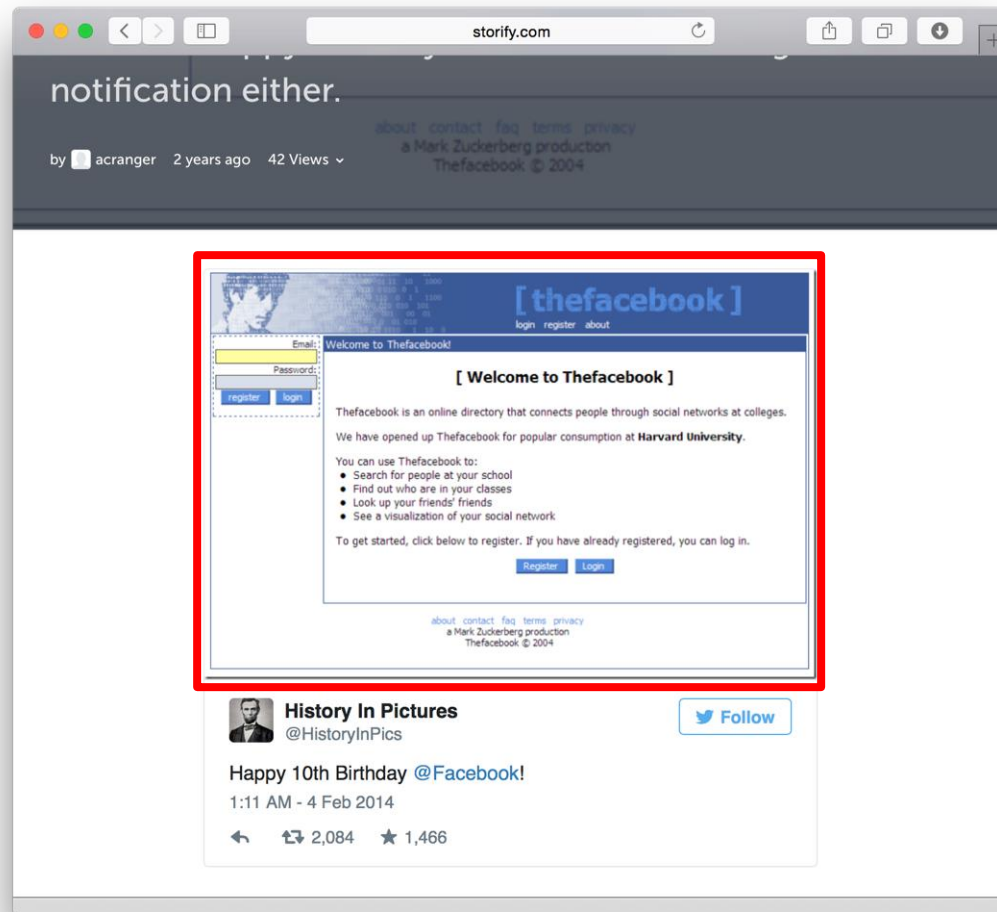
The most frequent domains in the stories represent ~92%

Host	Frequency	Percentage	Alexa Global Rank as of 2015-03	Category
twitter.com	943,859	82.05%	8	Social media
instagram.com	45,188	3.93%	25	Photos
youtube.com	22,076	1.92%	3	Videos
facebook.com	13,930	1.21%	2	Social media
flickr.com	7,317	0.64%	126	Photos
patch.com	5,783	0.50%	2,096	News
plus.google.com	3,413	0.30%	1	Social media
tumblr.com	3,066	0.27%	31	Blogs
blogspot.com	1,857	0.16%	18	Blogs
imgur.com	1,756	0.15%	36	Photos
coolpile.com	1,706	0.15%	149,281	Entertainment
wordpress.com	1,615	0.14%	33	Blogs
giphy.com	1,055	0.09%	1,604	Photos
bbc.com	966	0.08%	156	News
lastampa.it	927	0.08%	2,440	News
pinterest.com	892	0.08%	32	Photos
softandapps.info	861	0.07%	160,980	News
photobucket.com	768	0.07%	341	Photos
nytimes.com	744	0.06%	97	News
soundcloud.com	736	0.06%	167	Audio

The most frequent domains in the stories represent ~92%

Host	Frequency	Percentage	Alexa Global Rank as of 2015-03	Category
twitter.com	943,859	82.05%	8	Social media
instagram.com	45,188	3.93%	25	Photos
youtube.com	22,076	1.92%	3	Videos
facebook.com	13,930	1.21%	2	Social media
flickr.com	7,317	0.64%	126	Photos
patch.com	5,783	0.50%	2,096	News
plus.google.com	3,413	0.30%	1	Social media
tumblr.com	3,066	0.27%	31	Blogs
blogspot.com	1,857	0.16%	18	Blogs
imgur.com	1,756	0.15%	36	Photos
coolpile.com	1,706	0.15%	149,281	Entertainment
wordpress.com	1,615	0.14%	33	Blogs
giphy.com	1,055	0.09%	1,604	Photos
bbc.com	966	0.08%	156	News
lastampa.it	927	0.08%	2,440	News
pinterest.com	892	0.08%	32	Photos
softandapps.info	861	0.07%	160,980	News
photobucket.com	768	0.07%	341	Photos
nytimes.com	744	0.06%	97	News
soundcloud.com	736	0.06%	167	Audio

The embedded resources of the tweets



<https://pbs.twimg.com/media/BfnnY-0CcAA56oy.png>

The embedded resources of the tweets

- We sampled 5% of
47,512 tweets:
15,217
- The unique URIs:
14,616

Domain	Percentage	Category
twimg.com	46.17%	Images
instagram.com	4.28%	Images
youtube.com	2.82%	Videos
linkis.com	2.04%	Media sharing
facebook.com	1.40%	Social Media
wordpress.com	0.61%	Blogs
vine.co	0.53%	Videos
blogspot.com	0.52%	Blogs
storify.com	0.49%	Social Network
bbc.com	0.44%	News

The embedded resources of the tweets

- We sampled 5% of 47,512 tweets:
15,217
- The unique URIs:
14,616
- 46% are photos from twitter.com

Domain	Percentage	Category
twimg.com	46.17%	Images
instagram.com	4.28%	Images
youtube.com	2.82%	Videos
linkis.com	2.04%	Media sharing
facebook.com	1.40%	Social Media
wordpress.com	0.61%	Blogs
vine.co	0.53%	Videos
blogspot.com	0.52%	Blogs
storify.com	0.49%	Social Network
bbc.com	0.44%	News

The embedded resources of the tweets

- We sampled 5% of 47,512 tweets: 15,217
- The unique URIs: 14,616
- 0.49% of the stories point to other stories in Storify

Domain	Percentage	Category
twimg.com	46.17%	Images
instagram.com	4.28%	Images
youtube.com	2.82%	Videos
linkis.com	2.04%	Media sharing
facebook.com	1.40%	Social Media
wordpress.com	0.61%	Blogs
vine.co	0.53%	Videos
blogspot.com	0.52%	Blogs
storify.com	0.49%	Social Network
bbc.com	0.44%	News

Is there a correlation between
Alexa global rank and rank
within Storify?

Most of the time, the highly ranked resources correlate the most used resources in human-generated stories

n	10	15	25	50	100
Kendal	0.1555	0.4476	0.3372	0.3194	0.2485

Most of the time, the highly ranked resources correlate the most used resources in human-generated stories

n	10	15	25	50	100
Kendal	0.1555	0.4476	0.3372	0.3194	0.2485

This is in contrast to the usage of resources in Pinterest [1]

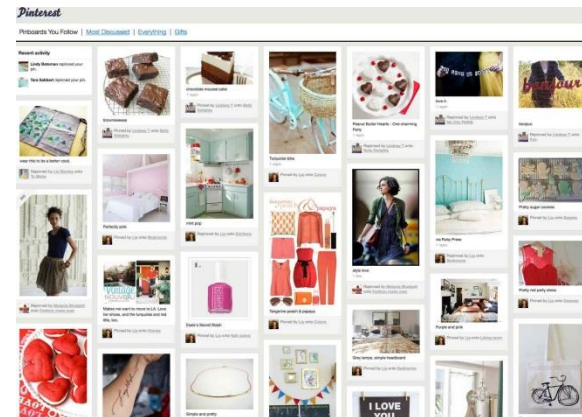


Image source: <http://workmonk.in/testblog/wp-content/uploads/2014/01/pinterest.jpg>

[1] Zhong, C., Shah, S., Sundaravadivelan, K., Sastry, N.: Sharing the Loves: Understanding the How and Why of Online Content Curation. ICWSM 2013

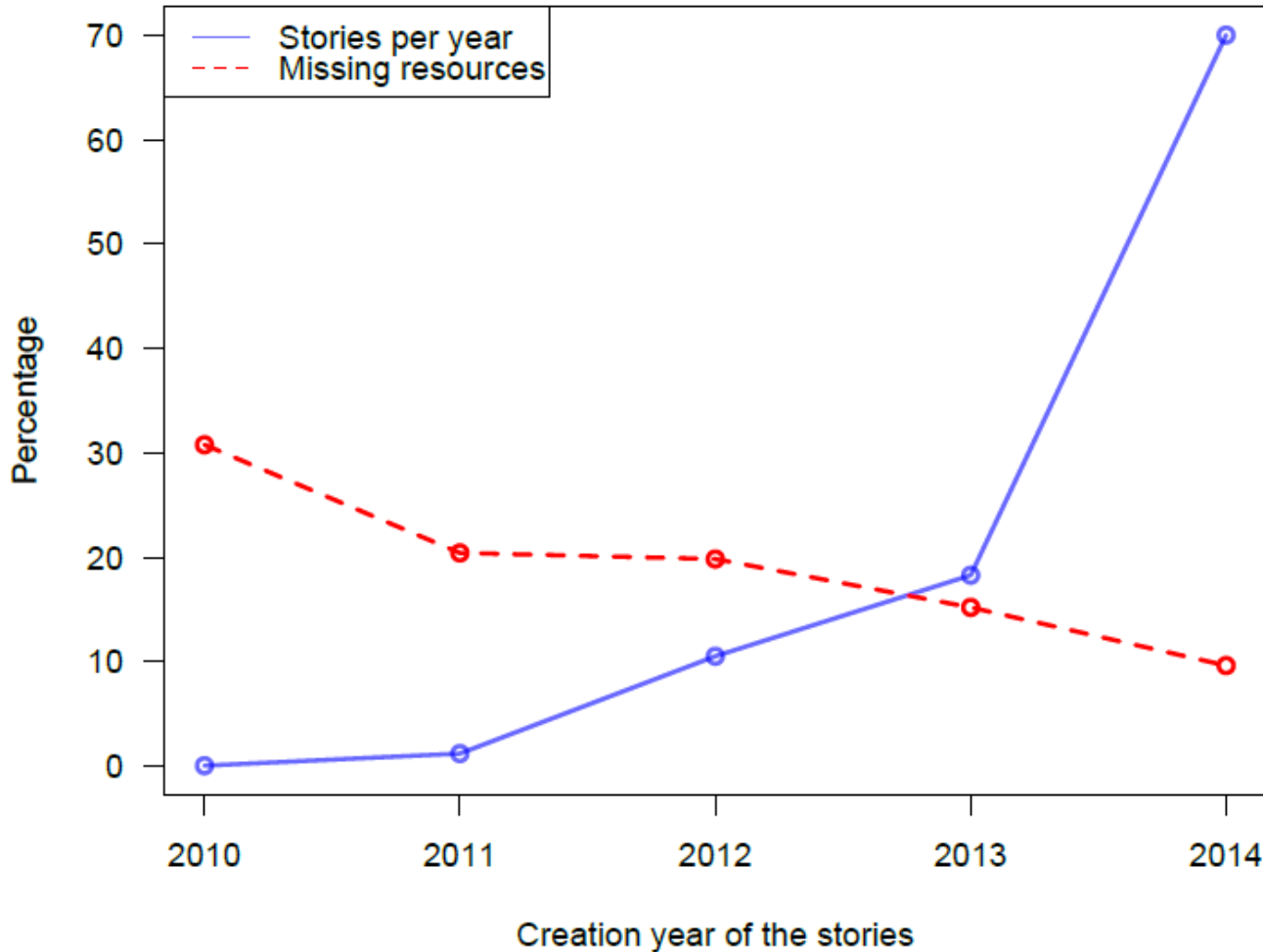
How many of the resources in
these stories disappear
every year?

Can we find these missing
resources in the archives?

The existence of the resources

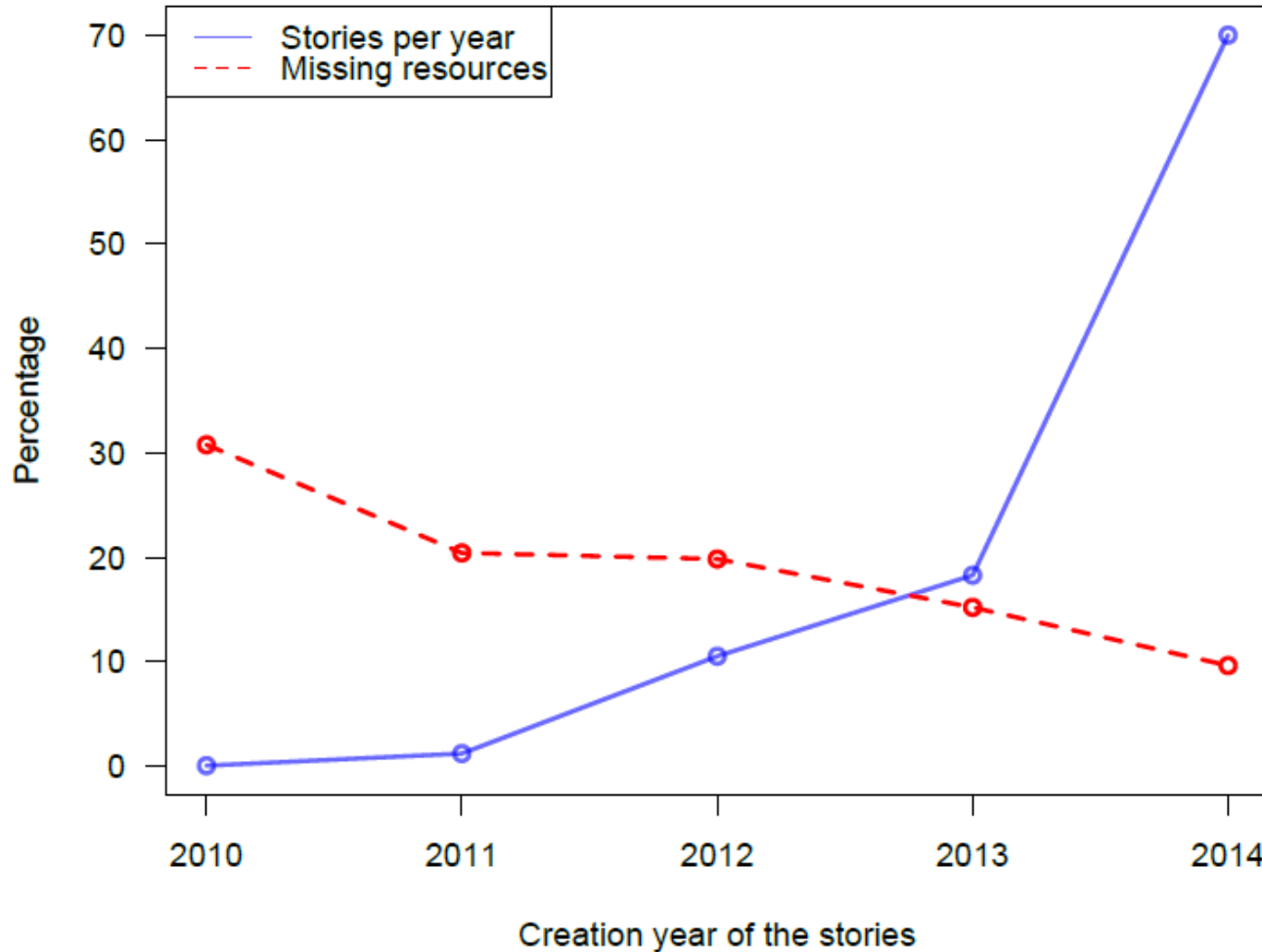
- We checked the live web and public web archives for **265,181** URIs
 - 202,452 URIs from story web elements
 - 47,512 randomly sampled tweet URIs
 - 15,217 URIs of embedded resources in those tweets
- The unique URIs are **253,978**
- We examined the results of the five most frequent domains in the stories
 - twitter.com, instagram.com, youtube.com, facebook.com, flickr.com

The decay rate of the resources in the stories



40.8% of the stories
contain missing
resources with an
average value of
10.3% per story

The decay rate of the resources in the stories



There is a nearly linear decay rate of resources through time. This matches the finding by SalahEldeen [2].

Existence on the live web and in the archives

Resources	Existence on live web			Found in archives		
	Available	Missing	Total	Of the available	Of the missing	Total
Twitter	95.5%	4.5%	47,385	0.9%	3.4%	477
Instagram	86.6%	13.4%	43,396	0.3%	0.07%	103
Youtube	99.3%	0.7%	19,809	16.0%	0.75%	3,140
Facebook	95.2%	4.8%	12,793	0.6%	0.49%	80
Flickr	95.6%	4.4%	6,859	0.4%	0.0%	25
others	82.1%	17.9%	109,120	26.8%	15.5%	27,033
Twitter resources	90.1%	9.9%	14,616	8.0%	14.1%	1,257

Of all the unique URIs, 11.8% are missing on the live web.

Existence on the live web and in the archives

Resources	Existence on live web			Found in archives		
	Available	Missing	Total	Of the available	Of the missing	Total
Twitter	95.5%	4.5%	47,385	0.9%	3.4%	477
Instagram	86.6%	13.4%	43,396	0.3%	0.07%	103
Youtube	99.3%	0.7%	19,809	16.0%	0.75%	3,140
Facebook	95.2%	4.8%	12,793	0.6%	0.49%	80
Flickr	95.6%	4.4%	6,859	0.4%	0.0%	25
others	82.1%	17.9%	109,120	26.8%	15.5%	27,033
Twitter resources	90.1%	9.9%	14,616	8.0%	14.1%	1,257

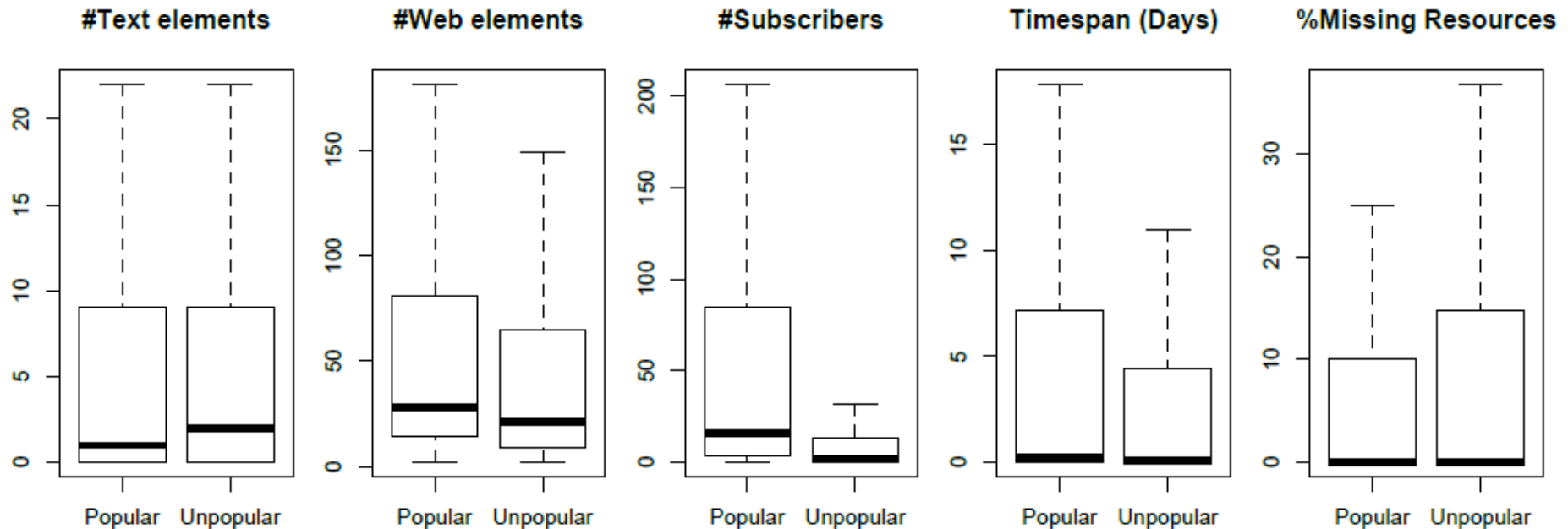
- The social media is not well-archived like the regular web
- Facebook uses robots.txt to block web archiving by the Internet Archive

What differentiates a
popular story?

Specifying the popular and the unpopular stories

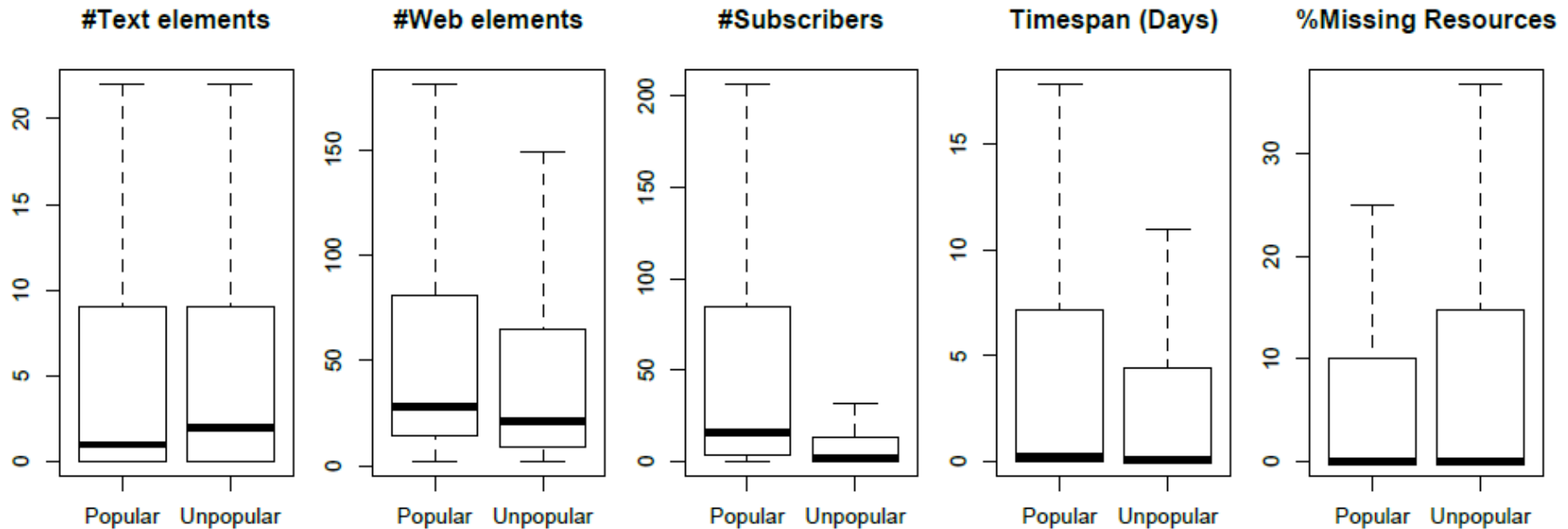
- The stories are divided based on the number of their views, then normalized by the amount of time they were available on the web.
- Popular stories
 - the top 25% of stories that have the most views
 - 3,642 stories

The distributions for the features of the stories



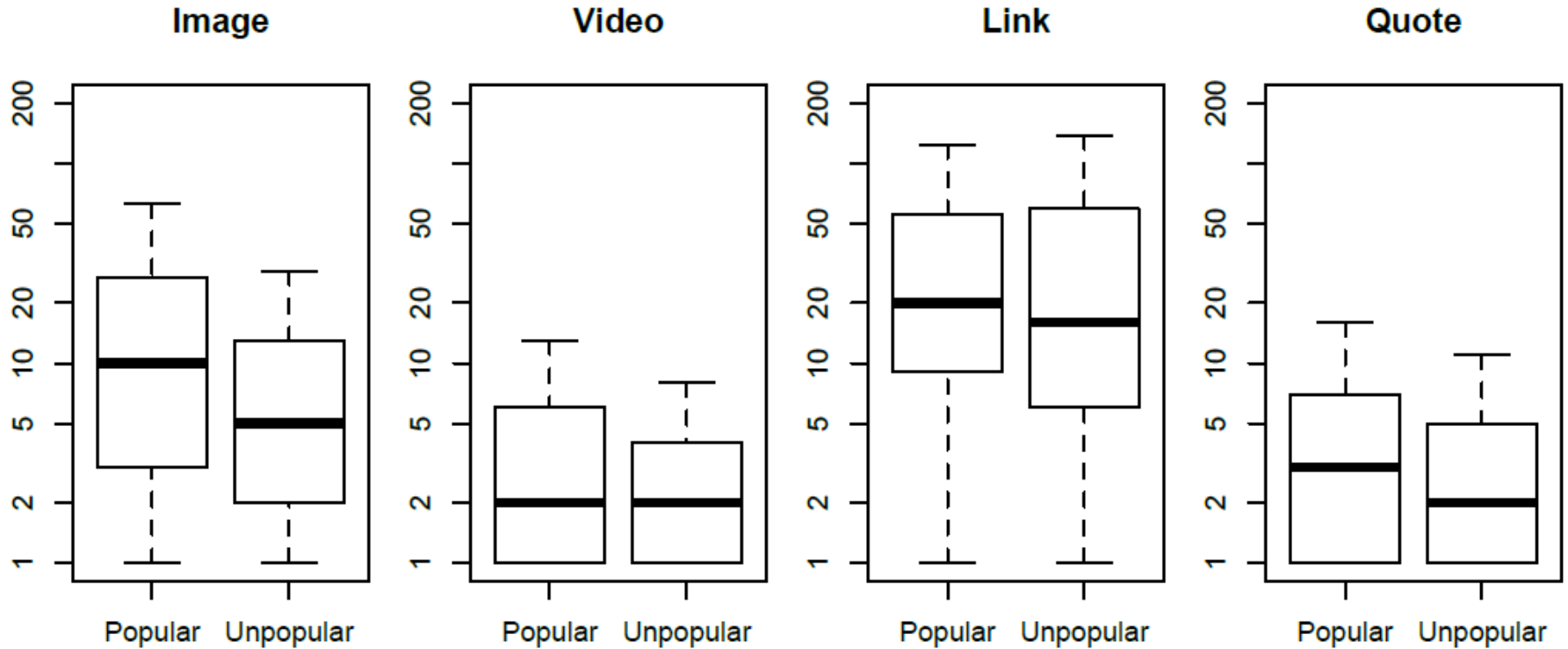
- Based on Kruskal-Wallis test, at the $p \leq 0.05$ significance level, the popular and the unpopular stories are different in terms of most of the features
- Popular stories tend to have:
 - more web elements (medians of 28 vs. 21)
 - longer timespan (5 hours vs. 2 hours) than the unpopular stories

Do Popular Stories have a Lower Decay Rate?



The 75th percentile of decay rate per popular story is 10% of the resources, while it is 15% in the unpopular stories

The distributions for the elements of the stories



Conclusions

- We analyzed 14,568 stories from Storify comprising 1,251,160 elements.
- Popular stories have a min/median/max value of 2/28/1950 elements, with the unpopular stories having 2/21/2216
- Popular stories have a median of 12 multimedia resources (the unpopular stories have a median of 7)
- Of the popular stories, 38% receive continuing edits (as opposed to 35%)
- In popular stories, only 11% of web elements are missing on the live web (as opposed to 13%)
- The percentage of the missing resources is proportional with the age of the stories.