

## **ALEXANDRIA**

# Temporal Retrieval, Exploration and Analytics in Web Archives

ERC AdG 2014 - 2019



Wolfgang Nejdl

L3S Research Center Hannover, Germany



## Why Libraries [Still] Matter (Jonathan Zittrain, Sep 2014)

## Valuable Books for Law Library

NO WRITER ATTRIBUTED May 17, 1912

At a recent dinner held by members of the courses Government 4 and Government 23, course in international law, G. G. Wilson, professor of international law, spoke on "The Olivart Collection of International Law." This splendid collection of books has been recently purchased and now forms a part of the Harvard Law Library in Langdell Hall. It is considered the most valuable collection of its kind in the world and contains a number of rare and limited editions. It was gathered together by the Marquis de Olivart, a Spanish writer on international law, who is the vice-president of the Institute of International Law and was formerly a professor at the University of Madrid.

Want to keep up with breaking news? Subscribe to our email newsletter.



http://www.thecrimson.com/article/1912/5/17/valuable-books-for-law-library-pat/



## Why Libraries [Still] Matter (Jonathan Zittrain, Sep 2014)

Librarians apprenticed to degrees in information science to know how to find things, and they coupled that skill with a professional commitment to neutrality, or at least absence of bias.

Fast forward past the year 2000, and both carefully accruing collections and expertly guiding people through them might seem quaint. The Web contains information about nearly anything, and search engines effortlessly index the Web. The only thing left is mopping up: retroactively digitizing materials that are sitting quietly in archives and depositories.

But: Nearly three out of four links found within all *Harvard Law Review* articles are dead. Half of the links in U.S. Supreme Court opinions are dead.

(Jonathan Zittrain, Kendra Albert & Lawrence Lessig. Perma: Scoping and Addressing the Problem of Link and Reference Rot in Legal Citations.

Harvard Law Review, March 2014)

http://blogs.harvard.edu/futureoftheinternet/2014/09/10/why-libraries-matter/





## On the importance of public libraries

John Palfrey: Bibliotech: Why Libraries Matter More Than Ever in the Age of Google (2015)

"If we don't maintain physical libraries, we will lose essential public, intellectual spaces in our communities, places where people can meet face-to-face, and if we don't build digital libraries connected to them, those physical spaces may become



obsolete as big companies such as Google and Amazon increasingly meet our need for knowledge. Physical and digital libraries are interdependent: each can make the other more effective and valuable."



Web Science - Investigating the Fu

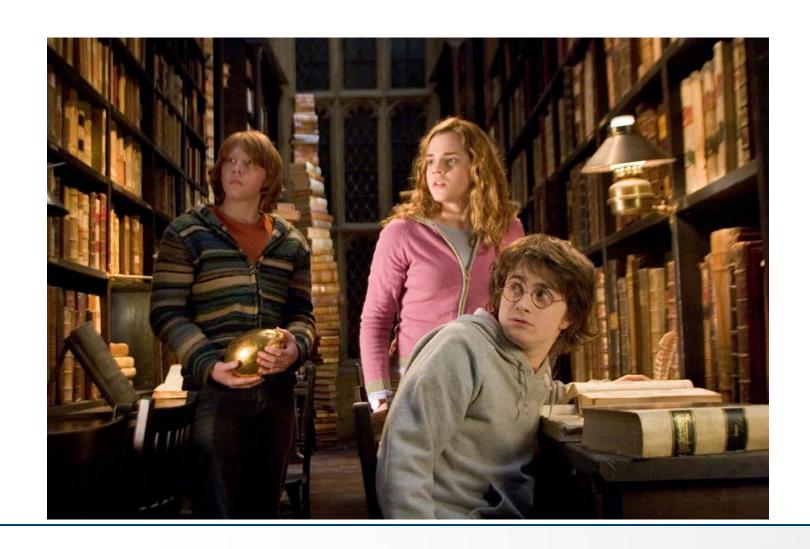
Collections



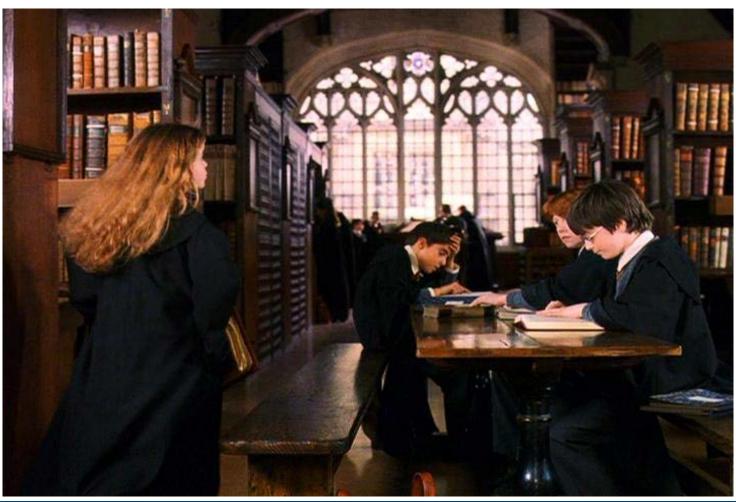


## Libraries and schools

## In Hogwarts and elsewhere









## Public libraries offer

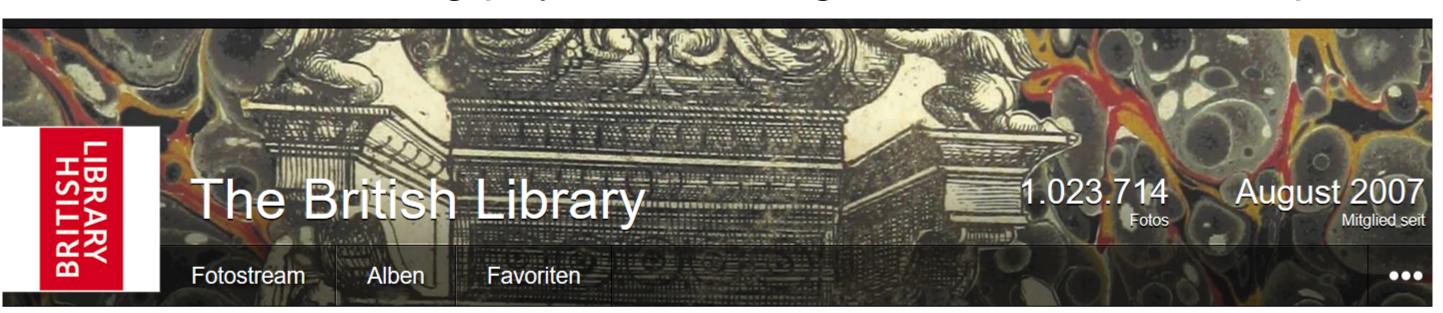
## not only

■ Books, magazines, films, music

## but also

- Spaces and meeting opportunities
- Digital information of any kind
- (Digital) services and (learning) opportunities

free to all, connecting physical and digital information and spaces



The British Library's collections on Flickr Commons offer access to millions of public domain images, which we encourage you to explore and re-use. The release of these collections into the public domain represent the Library's desire to improve knowledge of and about them, to enable novel and unexpected ways of using them, and to begin working with researchers to explore and interpret large scale digital collections.

Verbunden: August 2007

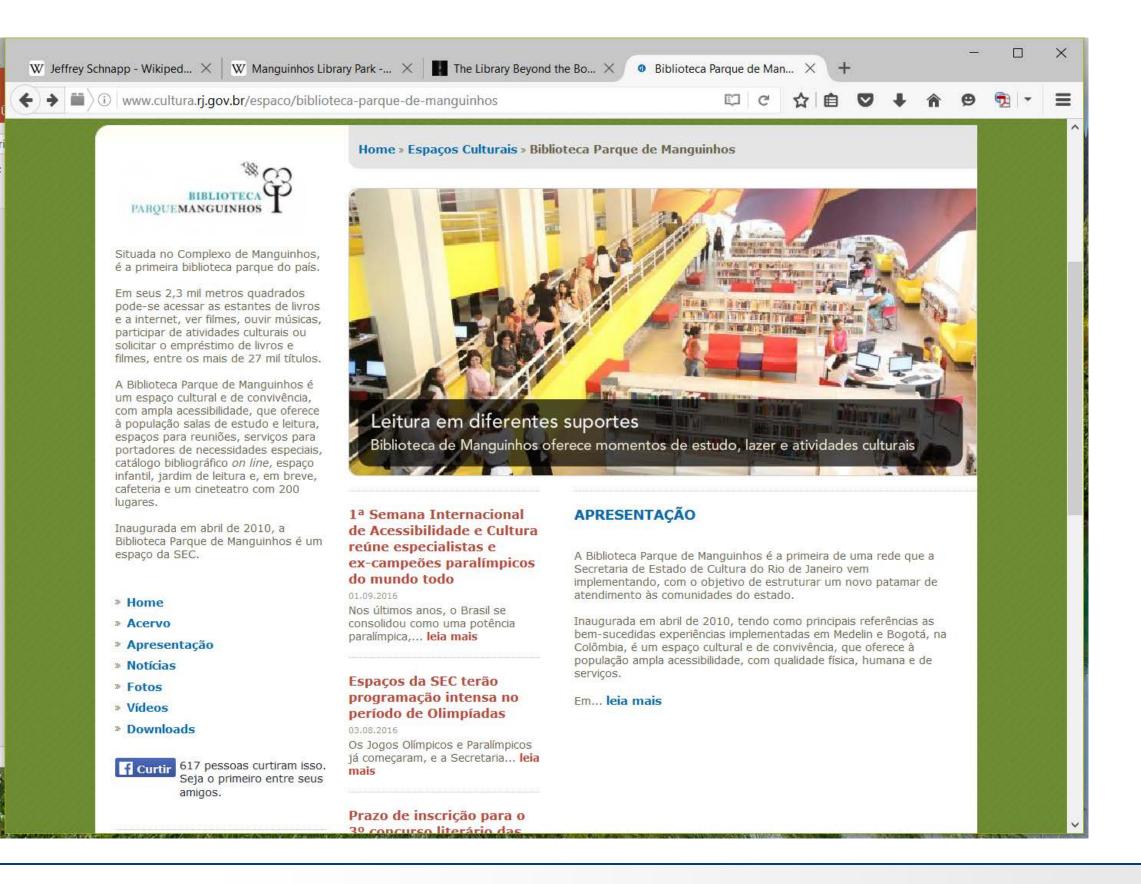
Website:

British Library





# Manguinhos Library Park in Rio de Janeiro



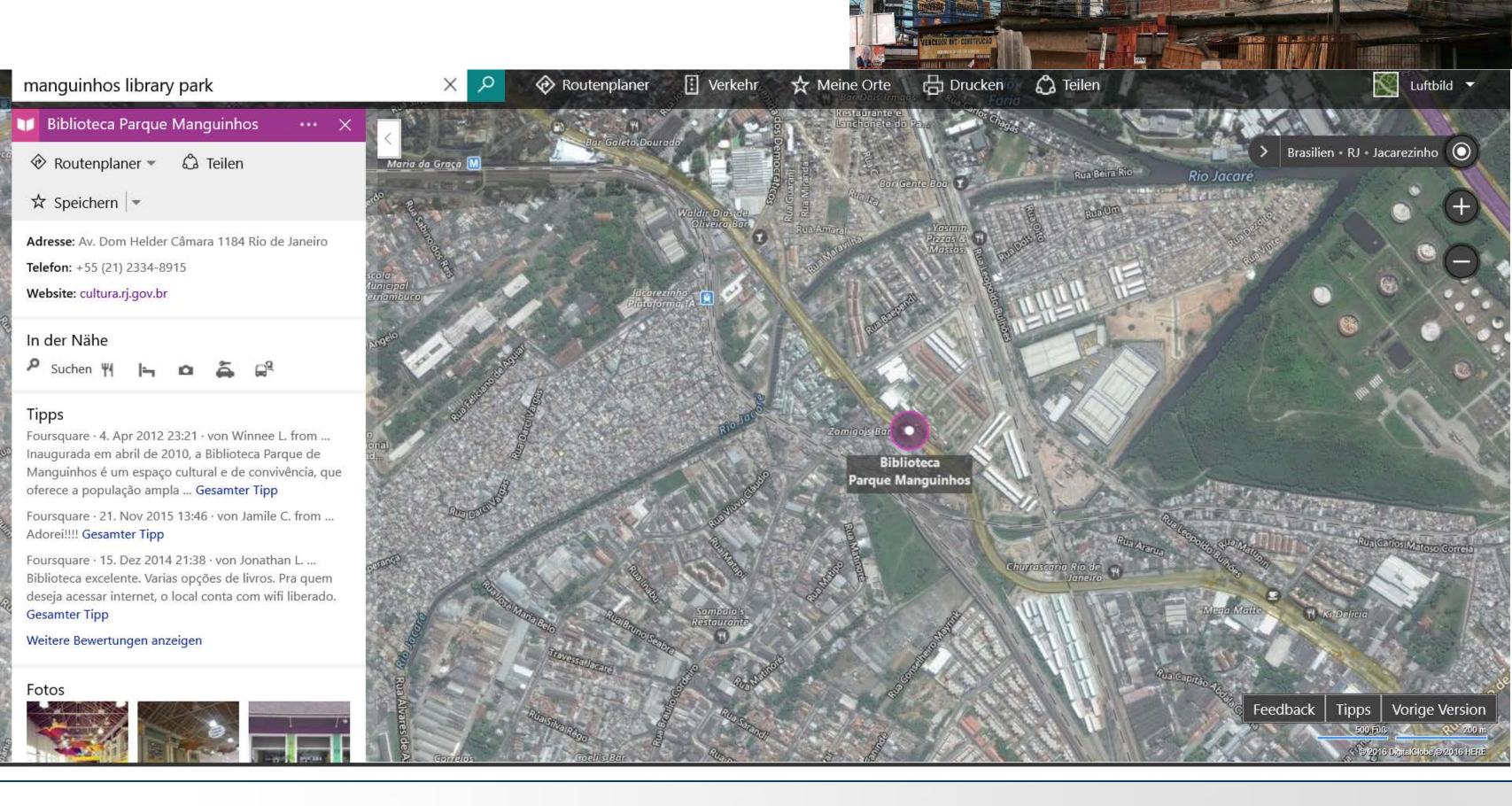
Opened in April, 2012

25000 books900 DVDs3 million songs

Theatre for 200 people
Music room
Classrooms
My neighbourhood
room ("Meu Bairro")

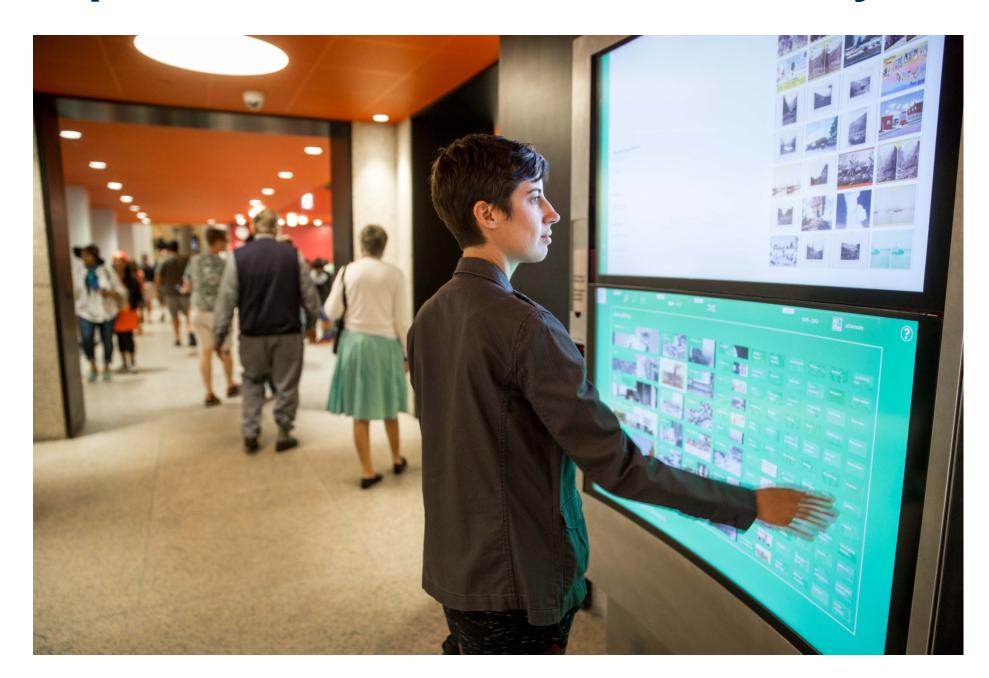


# Manguinhos Favela





## **Explore the Boston Public Library's Digital Collections**



http://www.bpl.org/press/2016/06/06/central-library-renovation-reopening-explore-the-bpls-digital-collections/

https://www.digitalcommonwealth.org/



## Web Science - Investigating the Future of Information and Communication

## **Boston Marathon bombing**

Coordinates: 42°20′59.2″N 71°4′44.1″W

From Wikipedia, the free encyclopedia

"Boston bomber" redirects here. For the Douglas Boston bomber aircraft, see Douglas A-20 Havoc.

The **Boston Marathon bombing** was a terrorist attack, followed by subsequent related shootings, that occurred when two pressure cooker bombs exploded during the Boston Marathon on April 15, 2013. The bombs exploded about 12 seconds and 210 yards (190 m) apart at 2:49 pm EDT, near the marathon's finish line on Boylston Street. The explosion killed 3 civilians and injured an estimated 264 others.<sup>[2][5][10]</sup>

The Federal Bureau of Investigation (FBI) took over the investigation and, on April 18, released photographs and a surveillance video of two suspects. [11][12][13]
The suspects were identified later that day as Chechen brothers Dzhokhar Tsarnaev and Tamerlan Tsarnaev. Shortly after the FBI released identifying images publicly, the suspects killed an MIT policeman, carjacked a civilian SUV, and initiated an exchange of gunfire with the police in nearby Watertown. During the firefight, a Massachusetts Bay Transportation Authority Police officer was injured but survived with severe blood loss. A Boston Police Department officer was also injured and died from his wounds nearly a year later. Tamerlan Tsarnaev was shot several times in the firefight and his brother subsequently ran him over with the stolen SUV in his escape. Tamerlan died shortly after arriving at Boston's Beth Israel Hospital.

An unprecedented manhunt for Dzhokhar Tsarnaev ensued on April 19, with thousands of law enforcement officers searching a 20-block area of Watertown.<sup>[14]</sup> During the manhunt, authorities asked residents of Watertown and surrounding areas, including Boston, to stay indoors. The public transportation system and most businesses and public institutions were shut down, creating a deserted urban environment of historic size and duration.<sup>[15][16]</sup> Around 6:00 p.m., shortly after the "shelter-in-place" advisory was rescinded, a Watertown resident discovered Dzhokhar hiding in a boat in his back yard. Reports conflict as to whether or not he was armed.<sup>[17]</sup> Located within the boat by thermal imaging, he was shot while in the boat, arrested, and then taken to a hospital shortly thereafter.<sup>[18]</sup>

During an initial interrogation in the hospital, Dzhokhar alleged that Tamerlan was the mastermind. He said they were motivated by extremist Islamist beliefs and the wars in Iraq and Afghanistan, and that they were self-radicalized and unconnected to any outside terrorist groups. According to him, they learned to build explosive devices from an online magazine of the al-Qaeda affiliate in Yemen.<sup>[19]</sup> He said that he and his brother had decided after the Boston bombing to travel to New York City to bomb Times Square. Dzhokhar was indicted on April 22, while still in the hospital, on 30 charges relating to homegrown terrorism, including use of a weapon of mass destruction and malicious destruction of property resulting in death.<sup>[7][20]</sup> He was found guilty on all charges on April 8, 2015, and the following month was sentenced to death.<sup>[21]</sup>

#### Contents [hide]

- 1 Bombing
- 2 Casualties and initial response
- 3 Investigation
- 4 Events of April 18-19
  - 4.1 Identification of suspects: Dzhokhar and Tamerlan Tsarnaev
  - 4.2 MIT shooting and carjacking
  - 4.3 Shootout with police
  - 4.4 Manhunt and capture of Dzhokhar Tsarnaev

#### **Boston Marathon bombing**

Part of terrorism in the United States



The area of the first blast a few minutes after the explosion

Location

**Bombing:** 671–673 and 755 Boylston Street, Boston,

Massachusetts

Shooting: Corner of Vassar Street and Main Street in Cambridge, Massachusetts<sup>[1]</sup>

Firefight and manhunt: Watertown, Massachusetts

Date

**Bombing:** April 15, 2013,

2:49 pm EDT

Shooting: April 18, 2013,

10:48 pm EDT

Firefight and manhunt:

April 19, 2013, 12:30 am – 8:42 pm EDT

Attack type

Bombings,<sup>[2]</sup> shootout, vehicle ramming, lone wolf terrorism<sup>[3]</sup>

Weapons

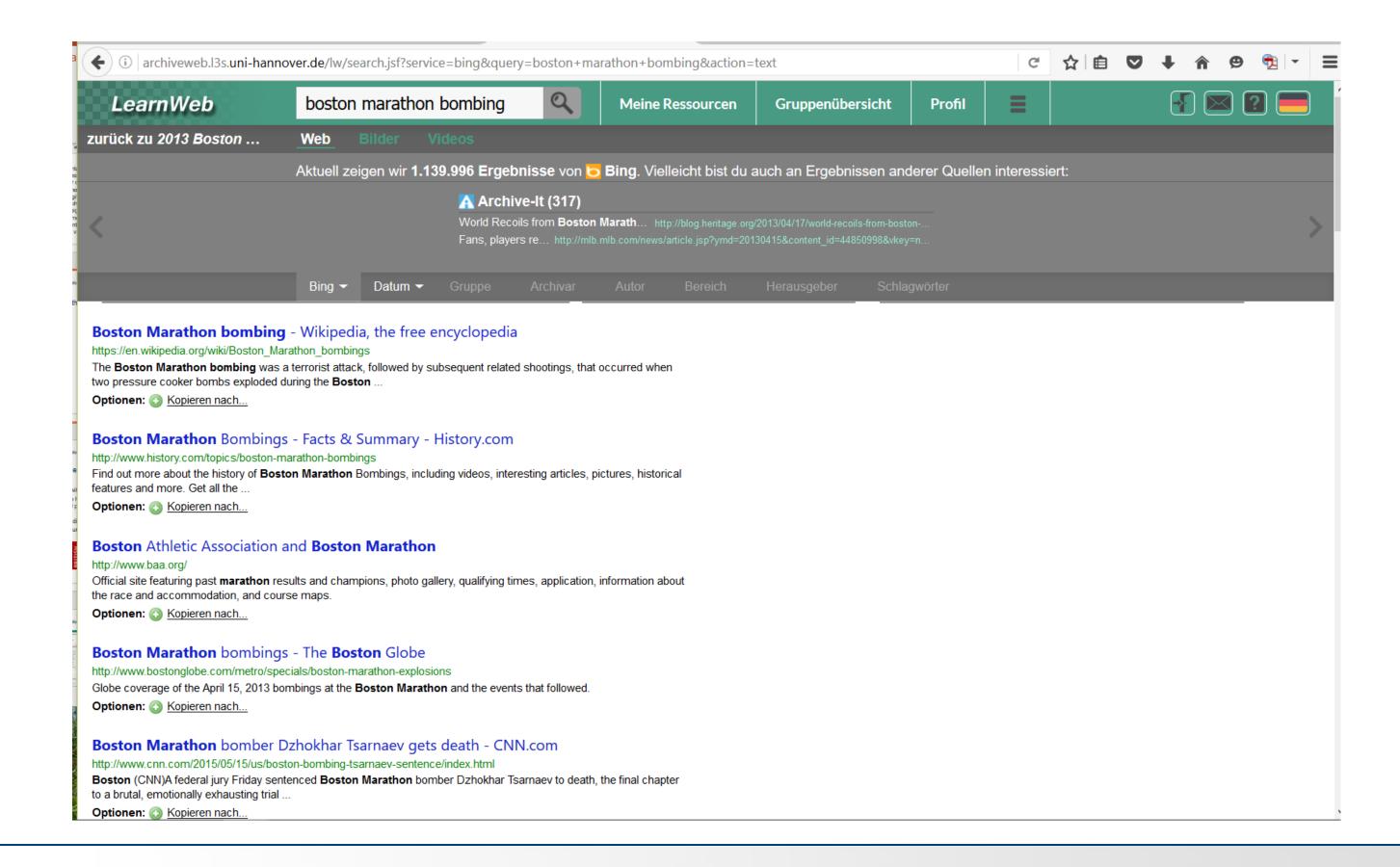
 Bombing: Two pressure cooker bombs

Shooting: 9mm Ruger P95

semi-automatic pistol

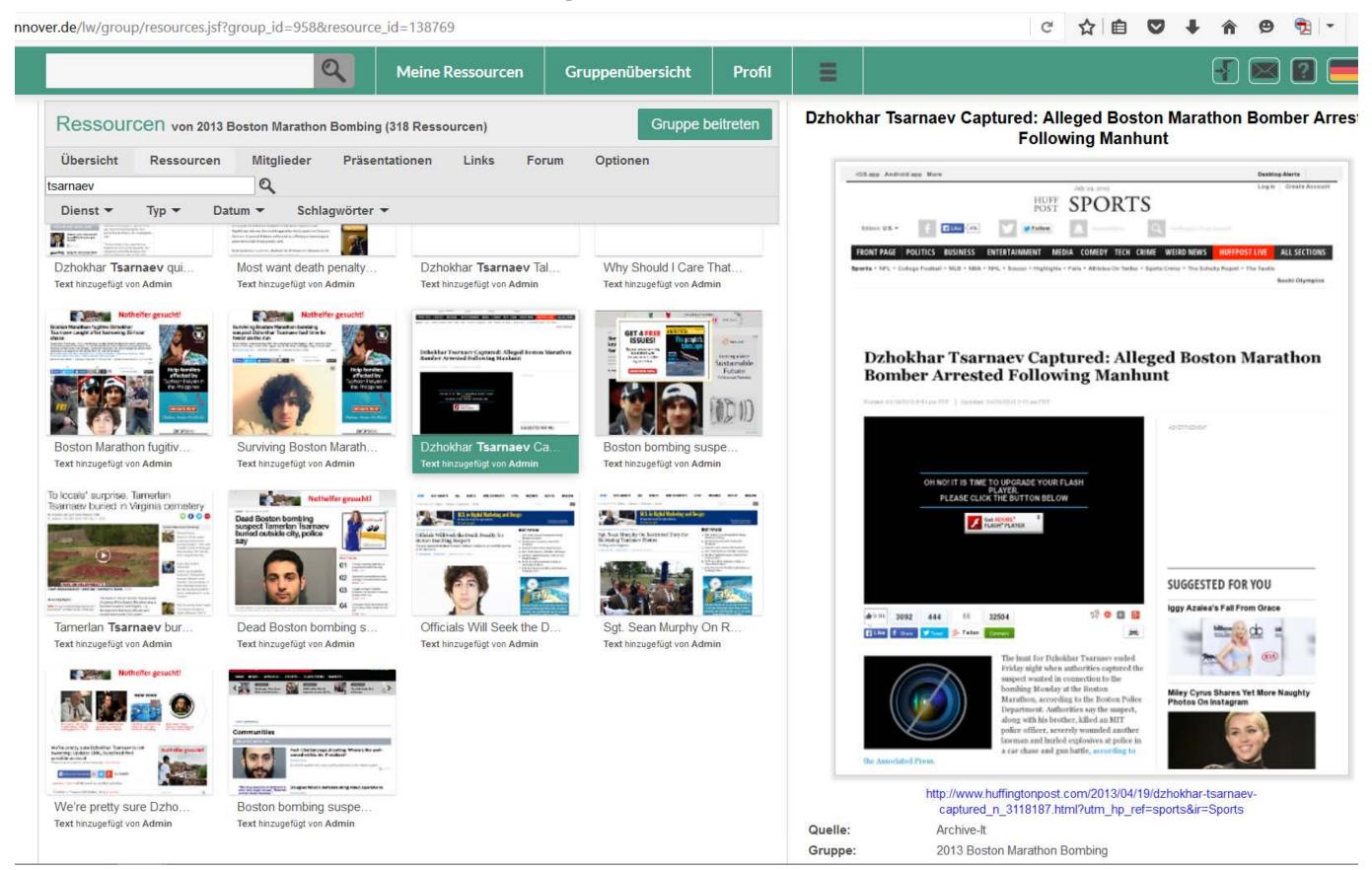


# **Boston Marathon Bombing**





# Boston Marathon Bombing in Archive-IT / ArchiveWeb





# Searching for Dzhokhar Tsarnaev



LearnWeb » http://www.bostonmagazine.com/news/blog/2013/07/23/sgt-sean-murphy-placed-in-restrictive-duty/archiviert zwischen 20.05.14 und 20.05.14

Optionen: (3) Kopieren nach... Gruppe: 2013 Boston Marathon Bombing

#### Tamerlan **Tsarnaev** buried in Virginia cemetary, uncle says

LearnWeb » http://www.cnn.com/2013/05/10/us/virginia-boston-suspect-buria archiviert zwischen 10.05.13 und 10.05.13

Optionen: (3) Kopieren nach... Gruppe: 2013 Boston Marathon Bombing

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 aftereffects.

URL: https://archive-it.org/collections/3649

#### Boston bombing suspects: What's known about **Tsarnae**

 $Learn Web * http://www.csmonitor.com/USA/2013/0419/Boston-bombing-suspects-What-s-known-about-Tsarnaev-brothers-so. \\ archiviert zwischen 19.04.13 und 19.04.13$ 

Optionen: (2) Kopieren nach... Gruppe: 2013 Boston Marathon Bombing

#### Dzhokhar **Tsarnaev** Captured: Alleged Boston Marathon Bomber Arrested Following Manhunt

LearnWeb » http://www.huffingtonpost.com/2013/04/19/dzhokhar-tsarnaev-captured\_n\_3118187.html?utm\_hp\_ref=sports&ir=.. archiviert zwischen 20.04.13 und 21.04.13

Optionen: (i) Kopieren nach... Gruppe: 2013 Boston Marathon Bombing

#### Boston bombing suspects admit role in the cover up for **Tsarnaev**

LearnWeb » http://communities.washingtontimes.com/neighborhood/ad-lib/2013/may/1/boston-bombing-suspects-admit-their-r.. archiviert zwischen 02.05.13 und 02.05.13

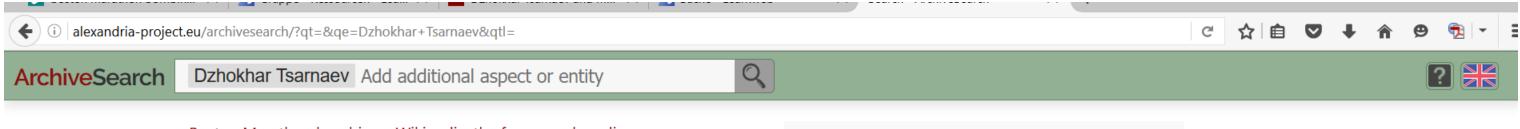
Optionen: (3) Kopieren nach... Gruppe: 2013 Boston Marathon Bombing

archiveweb.l3s.uni-hannover.de/lw/group/resources.jsf?group\_id=958 g; Update: CNN, BuzzFeed find possible



## Web Science - Investigating the Future of Information and Communication

## Searching for Dzhokhar Tsarnaev



#### Boston Marathon bombing - Wikipedia, the free encyclopedia

Wayback » http://en.wikipedia.org/wiki/Dzhokhar\_and\_Tamerlan\_Tsarnaev captured between 4/21/13 and 9/1/16

The Boston Marathon bombing was a terrorist attack, followed by subsequent related shootings, that occurred when two pressure cooker bombs exploded during the Boston ...

#### **Dzhokhar Tsarnaev**'s Claims of Jury Taint 'Overblown ...

Wayback » http://www.nbcnews.com/storyline/boston-bombing-trial/dzhokhar-tsarnaevs-claims-j... captured between 10/2/15 and 1/21/16

Federal prosecutors are urging the judge who presided over the trial of Boston Marathon bomber **Dzhokhar Tsarnaev** to reject his motion for a new trial, calling his ...

#### As Boston marks bombing anniversary, **Dzhokhar Tsarnaev** ...

Wayback » http://www.cnn.com/2014/04/15/us/boston-marathon-bombing-tsarnaev/index.html captured between 4/15/14 and 2/17/16

**Dzhokhar Tsarnaev** sits in a jail cell as dignitaries, survivors and rescuers gather to remember the Boston Marathon bombing.

#### **Dzhokhar Tsarnaev** - Wikipedia, the free encyclopedia

Wayback » https://en.wikipedia.org/wiki/Dzhokhar\_Tsarnaev captured between 4/22/13 and 8/3/16

**Tsarnaev** was the subject of a cover story for an August 2013 issue of Rolling Stone entitled "The Bomber: How a Popular, Promising Student Was Failed by ...

#### Dzhokhar Tsarnaev, Boston Bombing Suspect, Has Life Or ...

Wayback » http://www.huffingtonpost.com/2013/12/01/boston-bombing-death-penalty\_n\_436790.. captured between 12/3/13 and 12/15/13

He is accused of killing three people and injuring hundreds more in a bombing that shook Boston and the nation. Now, the decision whether or not to pursue ...

#### Boston Bomb Suspect **Dzhokhar Tsarnaev** Has Online Following ...

Wayback » http://abcnews.go.com/US/boston-bomb-suspect-dzhokhar-tsarnaev-online/story?id=... captured between 4/24/13 and 5/26/15

Boston has rallied around its police and victims of the marathon bombing, but there are thousands of people who are rallying around **Dzhokhar Tsarnaev**, the man accused ...

#### The arrest of **Dzhokhar Tsarnaev** - CNN.com

Wayback » http://www.cnn.com/2013/07/18/us/gallery/tsarnaev-arrest/index.html

The arrest of DZHOKHAL ISAHHAEV - CIV

#### Dzhokhar Tsarnaev



Dzhokhar Anzorovich "Jahar" Tsarnaev is a Kyrgyzstani-American citizen who was convicted of planting bombs at the Boston Marathon on April 15, 2013, together with his brother Tamerlan Tsarnaev. The bombings killed

three people and injured approximately 280 others. Wikipedia

#### Related entities:

Tamerlan Tsarnaev Boston Marathon bombings

Islamic Society of Boston Shelter in place

Judy Clarke

Capital punishment debate in the United States

Chechen American

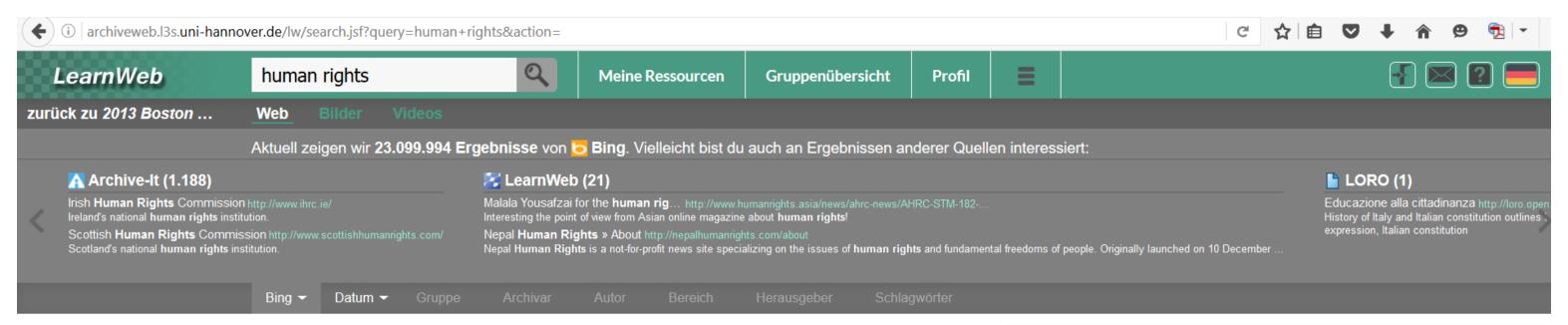
Islamic Center of New England

Capital punishment in Massachusetts

Federal Medical Center, Devens



## Human rights on the web and in archives



#### The Universal Declaration of **Human Rights** | United Nations

http://www.un.org/en/universal-declaration-human-rights/index.html

The Universal Declaration of **Human Rights**. The Universal Declaration of **Human Rights** (UDHR) is a milestone document in the history of **human rights**.

Optionen: (3) Kopieren nach...

#### Human rights - Wikipedia, the free encyclopedia

https://en.wikipedia.org/wiki/Human\_rights

**Human rights** are moral principles or norms, which describe certain standards of **human** behavior, and are regularly protected as legal **rights** in municipal and ...

Optionen: ( ) Kopieren nach...

#### OHCHR - What are Human Rights

http://www.ohchr.org/EN/Issues/Pages/WhatareHumanRights.aspx

**Human rights** are **rights** inherent to all **human** beings, whatever our nationality, place of residence, sex, national or ethnic origin, colour, religion, language, or any ...

Optionen: ( ) Kopieren nach...

#### Human Rights Watch | Defending Human Rights Worldwide

https://www.hrw.org

Defending **Human Rights** Worldwide ... Cluster munition attacks in Syria and Yemen are causing unacceptable civilian suffering and deserve a strong response.

Optionen: (1) Kopieren nach...

#### Home | HumanRights.gov is the official United States ...

http://www.humanrights.gov/

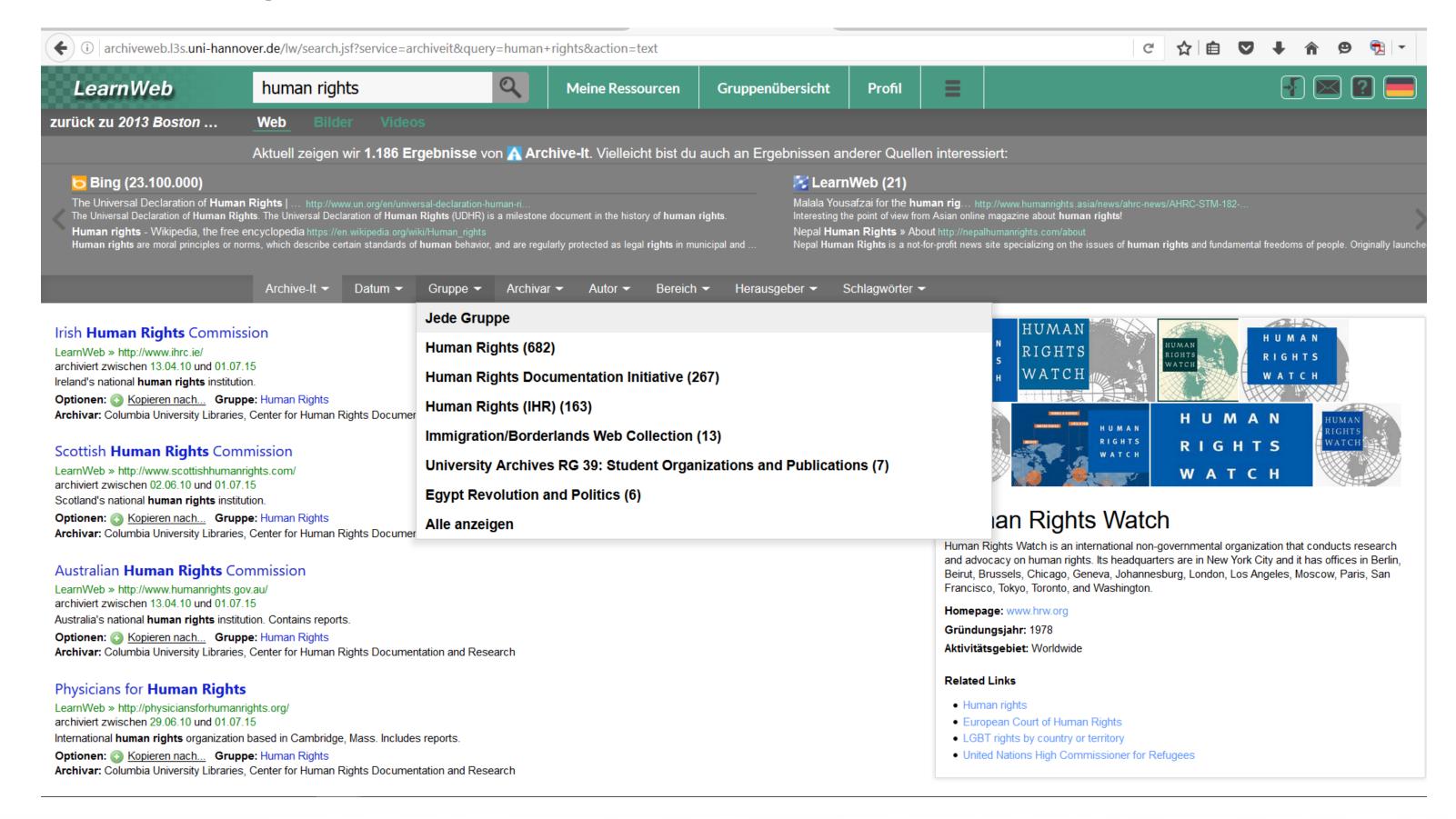
HumanRights.gov is the official United States Government website for international human rights related information



**A** ...



# Human rights on the web and in archives





# ArchiveSearch

**ArchiveSearch** 

Angela Merkel

Barack Obama







#### Barack Obama hails Angela Merkel over handling of refugees ...

Wayback » https://www.theguardian.com/world/2016/apr/24/migration-crisis-obama-and-eu-leade... captured between 4/24/16 and 9/1/16

US president Barack Obama said Angela Merkel was on the right side of history with her management of the refugee crisis and praised the German chancellor as a steady ...

#### Memes of Angela Merkel speaking to Barack Obama go viral ...

Wayback » http://www.dailymail.co.uk/news/article-3116506/The-Sound-Munich-meets-Forrest-G... captured between 6/9/15 and 3/4/16

A photograph of **Angela Merkel** engaged in an expressive conversation with **Barack Obama** in front of a stunning Alpine backdrop has gone viral and inspired dozens of new ...

#### How **Obama**, **Angela Merkel** learned to love one another ...

Wayback » http://www.cnn.com/2016/04/24/politics/barack-obama-angela-merkel-germany-europ... captured between 4/24/16 and 5/1/16

Story highlights. President Barack Obama is in Germany on Sunday; He will lend support to German Chancellor Angela Merkel

#### Obama Joins Angela Merkel in Pushing Trade Deal to a Wary ...

Wayback » http://www.nytimes.com/2016/04/25/world/europe/wary-of-big-business-germans-prot... captured between 4/24/16 and 7/25/16

Standing next to Chancellor **Merkel** in Hanover, President **Obama** said the benefits of a new trans-Atlantic trade deal were "indisputable."

#### Obama In Germany: Merkel Is On 'Right Side Of History' By ...

Wayback » http://www.breitbart.com/london/2016/04/24/obama-merkel-make-case-disputed-ttip-t... captured between 4/25/16 and 4/28/16

US President Barack Obama and German Chancellor Angela Merkel made a joint pitch Sunday for deeper transatlantic trade in the face of mounting opposition, vowing to ...

#### Remarks by President **Obama** and Chancellor **Merkel** in Joint ...

Wayback » https://www.whitehouse.gov/the-press-office/2016/04/24/remarks-president-obama-a...

#### Angela Merkel



Angela Dorothea Merkel is a German stateswoman and former research scientist. Merkel has been the Chancellor of Germany since 2005, and the leader of the Christian Democratic Union since 2000. Wikipedia

#### Barack Obama



Barack Hussein Obama II is the 44th and current President of the United States. He is the first African American to hold the office and the first president born outside of the continental United States. Wikipedia

#### Related entities:

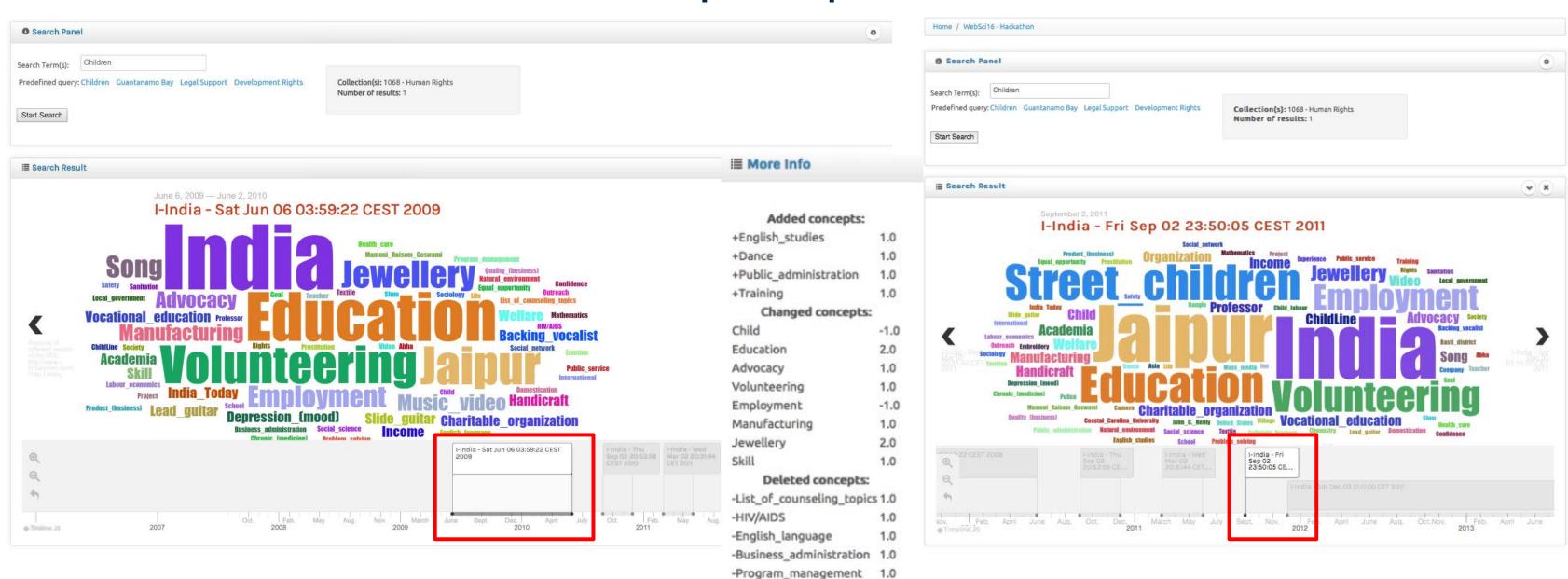
Thomas de Maizière
Hans-Peter Friedrich
José Sócrates
Jens Stoltenberg
Edmund Stoiber
Ilse Aigner
European debt crisis
Annegret Kramp-Karrenbauer
Bundeswehr
Barbara Hendricks (politician)
Richard von Weizsäcker

Federal Ministry of Labour and Social Affairs

http://alexandria-project.eu/archivesearch/

# SaR-Web: cross-language engine results analysis and temporal exploration of Web archive collections

## SaR-Web temporal exploration timeline



SaR-Web contribution

4015chungszentrun

- SaR-Web automatically extracts the most popular concepts from the top ranked search results
- The system enables cross-linguistic and cross-cultural comparison through a semantic based approach
- SaR-Web visualizes search results as concepts over time, thus overcoming the limit of textual descriptions



# Sentiment Analysis in Textual Streams, Eirini Ntoutsi

Huge amounts of opinions are continuously published and are freely available nowadays

Sentiment analysis aims at categorizing these opinions as either positive, negative or neutral.

A model (sentiment classifier) = algorithms + data In this talk we will focus on

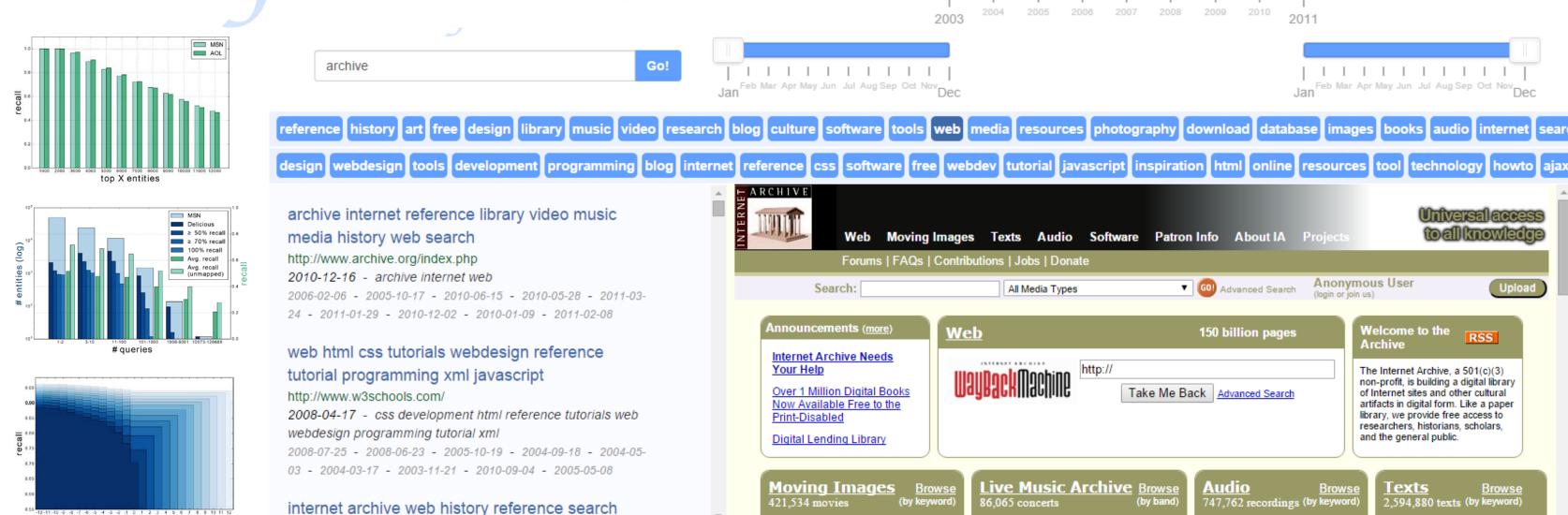
- Algorithms for sentiment learning in a stream environment
- Data labeling for model training



## **Delicious for Temporal Search on Web Archives**

Tempasaa

Try it out: <a href="http://tempas.L3S.de">http://tempas.L3S.de</a>



#months difference to May 2006

- [1] Helge Holzmann, Avishek Anand "Tempas: Temporal Archive Search Based on Tags". WWW 2016.
- [2] Helge Holzmann, Wolfgang Nejdl, Avishek Anand "On the Applicability of Delicious for Temporal Search on Web Archives". SIGIR 2016



# Archiving Software Surrogates on the Web for Future Reference

Linking referenced software in scientific articles to their archived websites at publication time

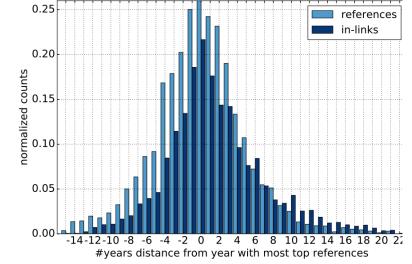
Supports comprehension of the articles and experiments

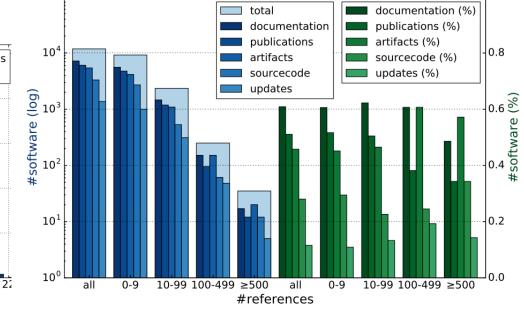
Pariser-Parr-Pople model (2010)

through archived • documentation

• --------

- publications
- artifacts
- source code
- updates

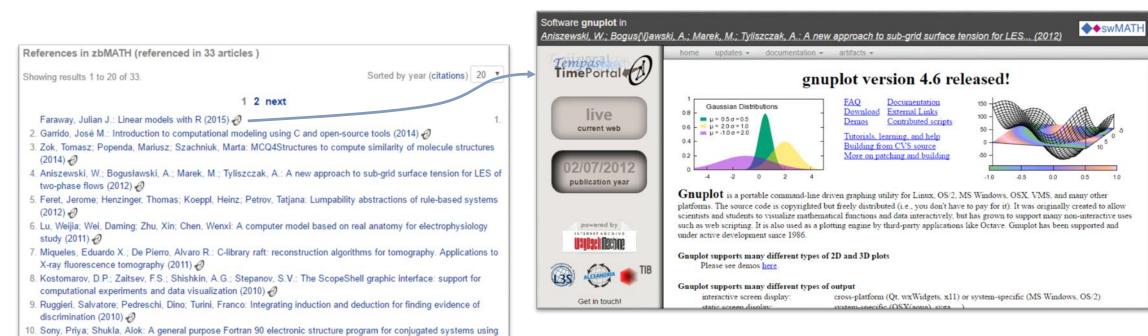












[1] Helge Holzmann, Wolfram Sperber, Mila Runnwerth - "Archiving Software Surrogates on the Web for Future Reference". TPDL 2016.



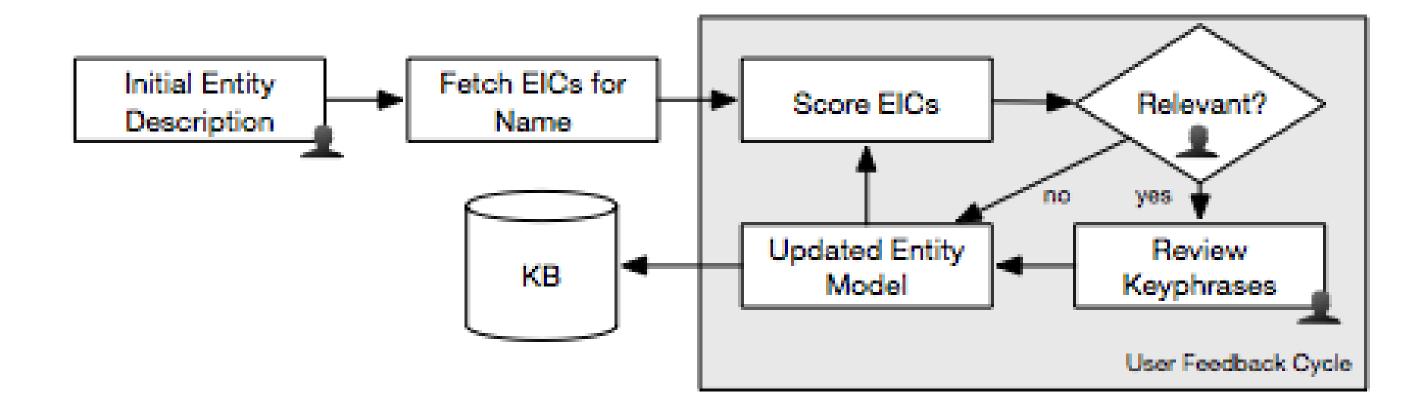
# Towards Archive Exploration

We want to help users better explore documents in a large temporal collection (bit.ly/archive-search -demo, Singh et. al. SIGIR'16)

Retrieval models, like HistDiv (Singh et. al. CHIIR'16), designed for exploratory search need semantic annotations like named entities

Archives contain many long tail entities that may not be present in Knowledge Bases(KB) which NED systems rely on

We propose a human-in-the-loop method to add long tail ambiguous entities to KBs (Singh et. al. CIKM'16)





## **ArchiveSpark: Efficient Web Archive Access**

Try it out: <a href="https://github.com/helgeho/ArchiveSpark">https://github.com/helgeho/ArchiveSpark</a>

## Easy and efficient research corpus building, extraction and derivation from Web archives

```
val rdd = ArchiveSpark.hdfs(cdxPath, warcPath)
val onlineHtml = rdd.filter(r => r.status == 200 && r.mime == "text/html")
val entities = onlineHtml.enrich(Entities)
entities.saveAsJson("entities.gz")
```

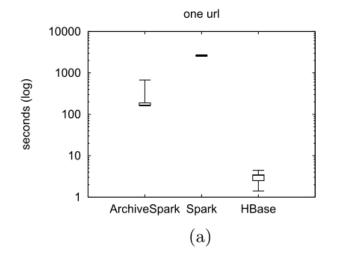


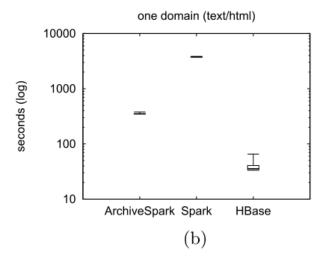
Faster than batch processing of

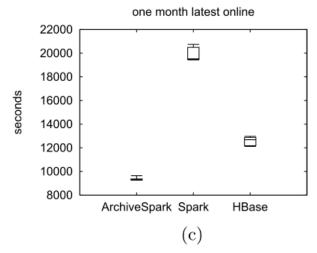
HBase or HDFS (pure Spark):

Web archives either on

- 1. Simple and expressive interface
- 2. Compliance to and reuse of standard formats
- 3. Efficient selection and filtering process
- 4. Easily extensible architecture
- 5. Lineage support
- 6. Output in a standard formats









[1] Helge Holzmann, Vinay Goel, Avishek Anand - "ArchiveSpark: Efficient Web Archive Access, Extraction and Derivation". JCDL 2016.



## Web Science - Investigating the Future of Information and Communication

## **Boston Marathon bombing**

Coordinates: 42°20′59.2″N 71°4′44.1″W

From Wikipedia, the free encyclopedia

"Boston bomber" redirects here. For the Douglas Boston bomber aircraft, see Douglas A-20 Havoc.

The **Boston Marathon bombing** was a terrorist attack, followed by subsequent related shootings, that occurred when two pressure cooker bombs exploded during the Boston Marathon on April 15, 2013. The bombs exploded about 12 seconds and 210 yards (190 m) apart at 2:49 pm EDT, near the marathon's finish line on Boylston Street. The explosion killed 3 civilians and injured an estimated 264 others.<sup>[2][5][10]</sup>

The Federal Bureau of Investigation (FBI) took over the investigation and, on April 18, released photographs and a surveillance video of two suspects. [11][12][13]
The suspects were identified later that day as Chechen brothers Dzhokhar Tsarnaev and Tamerlan Tsarnaev. Shortly after the FBI released identifying images publicly, the suspects killed an MIT policeman, carjacked a civilian SUV, and initiated an exchange of gunfire with the police in nearby Watertown. During the firefight, a Massachusetts Bay Transportation Authority Police officer was injured but survived with severe blood loss. A Boston Police Department officer was also injured and died from his wounds nearly a year later. Tamerlan Tsarnaev was shot several times in the firefight and his brother subsequently ran him over with the stolen SUV in his escape. Tamerlan died shortly after arriving at Boston's Beth Israel Hospital.

An unprecedented manhunt for Dzhokhar Tsarnaev ensued on April 19, with thousands of law enforcement officers searching a 20-block area of Watertown.<sup>[14]</sup> During the manhunt, authorities asked residents of Watertown and surrounding areas, including Boston, to stay indoors. The public transportation system and most businesses and public institutions were shut down, creating a deserted urban environment of historic size and duration.<sup>[15][16]</sup> Around 6:00 p.m., shortly after the "shelter-in-place" advisory was rescinded, a Watertown resident discovered Dzhokhar hiding in a boat in his back yard. Reports conflict as to whether or not he was armed.<sup>[17]</sup> Located within the boat by thermal imaging, he was shot while in the boat, arrested, and then taken to a hospital shortly thereafter.<sup>[18]</sup>

During an initial interrogation in the hospital, Dzhokhar alleged that Tamerlan was the mastermind. He said they were motivated by extremist Islamist beliefs and the wars in Iraq and Afghanistan, and that they were self-radicalized and unconnected to any outside terrorist groups. According to him, they learned to build explosive devices from an online magazine of the al-Qaeda affiliate in Yemen.<sup>[19]</sup> He said that he and his brother had decided after the Boston bombing to travel to New York City to bomb Times Square. Dzhokhar was indicted on April 22, while still in the hospital, on 30 charges relating to homegrown terrorism, including use of a weapon of mass destruction and malicious destruction of property resulting in death.<sup>[7][20]</sup> He was found guilty on all charges on April 8, 2015, and the following month was sentenced to death.<sup>[21]</sup>

#### Contents [hide]

- 1 Bombing
- 2 Casualties and initial response
- 3 Investigation
- 4 Events of April 18-19
  - 4.1 Identification of suspects: Dzhokhar and Tamerlan Tsarnaev
  - 4.2 MIT shooting and carjacking
  - 4.3 Shootout with police
  - 4.4 Manhunt and capture of Dzhokhar Tsarnaev

#### **Boston Marathon bombing**

Part of terrorism in the United States



The area of the first blast a few minutes after the explosion

Location

**Bombing:** 671–673 and 755 Boylston Street, Boston,

Massachusetts

Shooting: Corner of Vassar Street and Main Street in Cambridge, Massachusetts<sup>[1]</sup>

Firefight and manhunt: Watertown, Massachusetts

Date

**Bombing:** April 15, 2013,

2:49 pm EDT

Shooting: April 18, 2013,

10:48 pm EDT

Firefight and manhunt:

April 19, 2013, 12:30 am – 8:42 pm EDT

Attack type

Bombings,<sup>[2]</sup> shootout, vehicle ramming, lone wolf terrorism<sup>[3]</sup>

Weapons

 Bombing: Two pressure cooker bombs

Shooting: 9mm Ruger P95

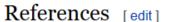
semi-automatic pistol





# Wikipedia references





https://en.wikipedia.org/wiki/Boston\_Marathon\_bombing

- 1. A "He loved us, and we loved him" . MIT. Retrieved April 11, 2015.
- A a b Straw, Joseph; Ford, Bev; McShane, Lawrence (April 17, 2013). "Police narrow in on two suspects in Boston Marathon bombings" D. The Daily News. New York. Retrieved May 15, 2013.
- 3. ^ Sherman, Pat (April 21, 2013). "UCSD professor says Boston Marathon was 'lone wolf' terrorism" ☑. La Jolla Light. Retrieved June 17, 2014.
- 4. <sup>A</sup> <sup>a</sup> <sup>b</sup> Ransom, Jan (May 28, 2015). "Death benefit given to family of officer wounded in Tsarnaev shootout" ₽. The Boston Globe. Retrieved May 28, 2015.
- 5. A a b c d Kotz, Deborah (April 24, 2013). "Injury toll from Marathon bombs reduced to 264" D. The Boston Globe. Retrieved April 29, 2013. "Boston public health officials said Tuesday that they have revised downward their estimate of the number of people injured in the Marathon attacks, to 264."
- 6. ^ a b c d Carter, Chelsea J.; Botelho, Gregory (April 20, 2013).

  "'Captured!!!' Boston police announce Marathon bombing suspect in custody"

  □ CNN.
- a:"Richard H. Donohue Jr., 33,... was shot and wounded in the incident... Another 15 police officers were treated for minor injuries sustained during the explosions and shootout".
- 7. ^ a b c "United States vs. Dzhokhar Tsarnaev, Case 1:13-mj-02106-MBB" (PDF). United States Department of Justice. April 21, 2013. Retrieved April 22, 2013. (Archive 1)
- 8. ^ a b Wilson, Scott (April 23, 2013). "Boston bombing suspect cites U.S. wars as motivation, officials say" . The Washington Post. et al. Retrieved April 23, 2013.
- 9. ^ Cooper, Michael; Schmidt, Michael S.; Schmitt, Eric (April 23, 2013). "Boston Suspects Are Seen as Self-Taught and Fueled by Web" . The New York Times. Retrieved April 11, 2015.
- 10. ^ a b "What we know about the Boston bombing and its aftermath" ☑. CNN. April 19, 2013. Retrieved April 19, 2013.
- 11. <sup>A a b</sup> Estes, Adam Clark; Abad-Santos, Alexander; Sullivan, Matt (April 15, 2013). "Explosions at Boston Marathon Kill 3 Now, a 'Potential Terrorist Investigation' "

  ☐ The Atlantic Wire. Retrieved April 17, 2013

- 113. A Rawlings, Nate (April 19, 2013). "Was Boston Actually on Lockdown?" 
  Time. Retrieved April 29, 2013.
- 114. ^ "City of Somerville Safety Advisory" ₽. Somerville News. April 19, 2013. Retrieved April 11, 2015.
- 115. ↑ "Gunfire heard in search for Boston Marathon bomb suspect" ☑. Reuters. April 19, 2013. Retrieved April 19, 2013.
- 116. ^ "Boston Marathon Manhunt: Search for bombing suspect is law enforcement's first major test of post-9/11 training" ₽. CBS News. April 19, 2013. Retrieved April 19, 2013.
- 117. ^ "Boston police" ☑. Twitter. April 19, 2013. Retrieved April 11, 2015. "Taxi service in the City of Boston has been restored."
- 118. <sup>A a b c</sup> Mutzabaugh, Ben (April 19, 2013). "Boston flights operating, airlines waive change fees" ☑. USA Today. Retrieved April 19, 2013.
- 119. \* U.S. News (April 8, 2015). "'We got him!': Boston bombing suspect captured alive" R. NBC News. Retrieved April 8, 2015.
- 120. ^ "Boston Bomb Suspect's Dad Tells Him to Surrender, Warns 'Hell Will Break Loose' if Son Dies" ₽. ABC News. April 19, 2013. Retrieved April 19, 2013.
- 121. ^ "Boston Marathon bombers: suspect Dzhozkar Tsarnaev's uncle Ruslan Tsarni pleads 'turn yourself in' "

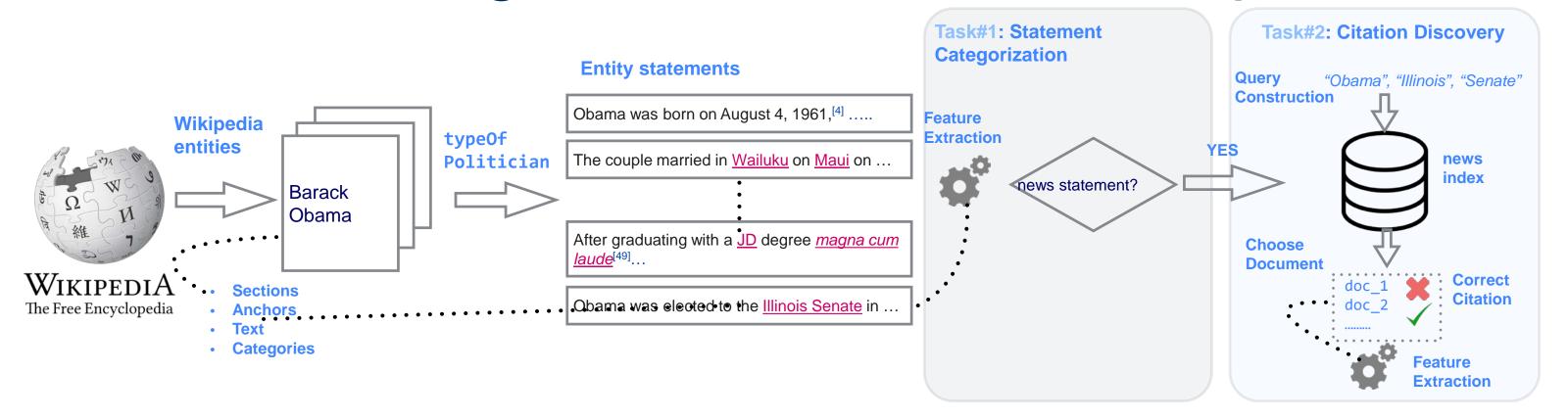
  ☐. The Telegraph. London, UK. Associated Press. April 19, 2013. Retrieved April 19, 2013.
- 122. ^ "Second Boston Marathon bombing suspect in custody" ☑. WCBV. April 20, 2013. Retrieved April 20, 2013.
- 123. <sup>▲</sup> "Boston police explain how bombing suspect was caught" ☑. UK: BBC News. April 20, 2013. Retrieved April 29, 2013.
- 124. ^ Anderson, Derek J. (April 19, 2013). "Watertown family finds alleged marathon bomber in boat" . Boston Globe. Retrieved April 19, 2013.
- 125. ^ Thomas, Linda (April 21, 2013). "Northwest technology helped find Boston bombing suspect" ₽. MyNorthwest.com. Retrieved April 21, 2013.
- 126. ^ By Evan Allen (April 23, 2013). "Boston police superintendent recounts officers' long search, tense final confrontation" ☑. Boston Globe. Retrieved May 8, 2013.
- 197 A Doctor latest name: Americana take to the atreats in colebration

- 234. ^ a b Peterson, Hayley (April 20, 2013). "The moment one woman and two men were taken into custody over Boston bombings near Dartmouth campus where suspect studied" Daily Mail. London. Retrieved April 11, 2015.
- 235. <sup>A a b</sup> Ellement, John R.; Wallack, Todd; Sacchetti, Maria; Carroll, Matt; Finucane, Martin (May 1, 2013). "Two Kazakh men, Cambridge man, face charges in disposal of backpack owned by Boston Marathon bombing suspect" □. Boston Globe. Retrieved May 2, 2013.
- 236. ^ "FBI: Friends tried to cover bombing suspect's tracks" ☑. USA today (Facebook post). May 2, 2013. Retrieved May 8, 2013.
- 237. ^ Ellement, John R.; Wallack, Todd; Sacchetti, Maria; Carroll, Matt; Finucane, Martin (May 1, 2013). "Three additional people taken into custody in Boston Marathon bombings" ☑. Boston Globe. Retrieved May 2, 2013.
- 238. ^ Vincent, H. Michael (April 21, 2013). "More arrests in Boston bombings, FBI eyes terror cell" ₽. Rockford Record. Retrieved April 11, 2015.
- 239. \* DeQuattro, Dee (April 22, 2013). "Two men in New Bedford arrested on immigration charges" & ABC 6.
- 240. ^ Taylor, Adam (May 1, 2013). "Here's The Criminal Complaint Against The New Suspects in the Boston Bombing Case" ₽. Business Insider. Retrieved May 2, 2013.
- 241. ^ Kubin, Jacquie (May 1, 2013). "Boston Update: Azamat Tazhayakov, Dias Kadyrbayev arrested in bombings" ₽. Retrieved May 2, 2013.
- 242. ^ "2 friends of Boston bombing suspect indicted" ☑. USA Today. August 8, 2013. Retrieved August 9, 2013.
- 243. A "News" . msn.com. Retrieved April 8, 2015.
- 244. ^ "Dzhokhar Tsarnaev's College Friend Pleads Guilty" ☑. Boston Magazine. August 22, 2014. Retrieved April 8, 2015.
- 245. ^ "Tsarnaev friends' sentencing delayed" №. WCBV 5 (ABC). November 7, 2014. Retrieved April 8, 2015.
- 246. A a b McMahon, Shannon; Sargent, Hilary (June 5, 2015).

  "Tsarnaev friend Tazhayakov sentenced to 42 months for obstruction of justice" R. Boston Globe.

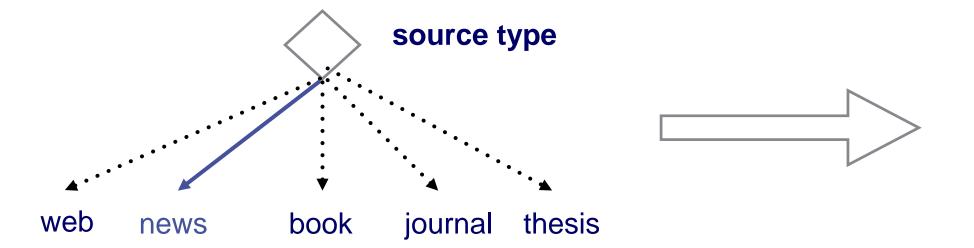


# Finding News Citation for Wikipedia



## **Task#1: Statement Categorization**

Obama was elected to the Illinois Senate in ...[2]



## **Task#2: Citation Discovery**



news collection

- [1] Illinois Sen. Barack Obama's Announcement Speech Washington Post, 2007.
- [2] Jackson, David; Ray Long (April 3, 2007). "Obama Knows His Way Around a Ballot". Chicago Tribune.

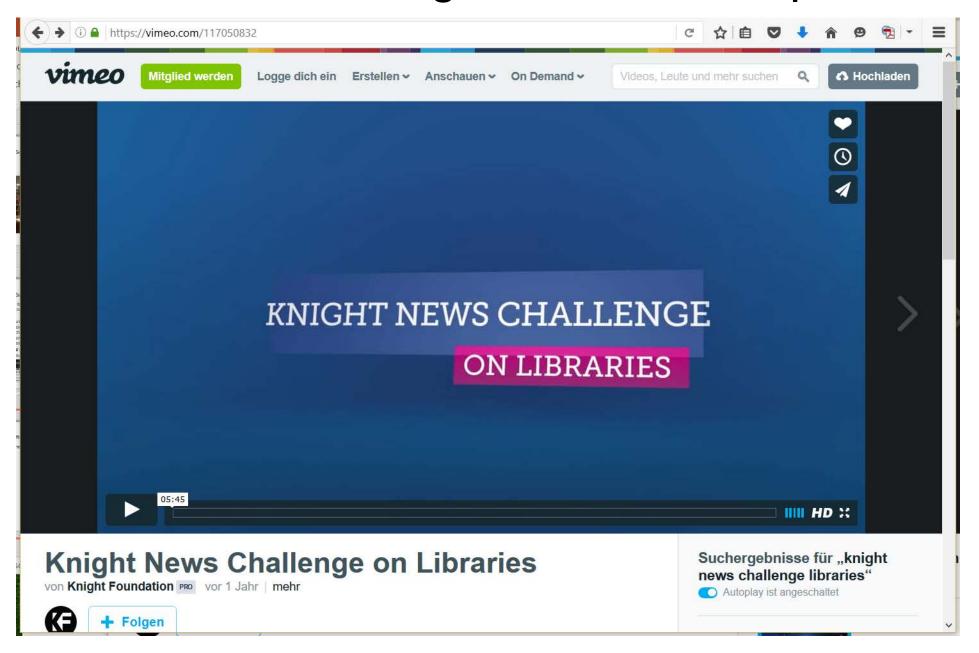
[10] In Illinois, Obama Proved Pragmatic and Shrewd - The New York Times, 2007.

Fetahu, B.; Markert, K.; Nejdl, W. & Anand, A. (2016), Finding News Citations for Wikipedia, In the Proceedings of the 25th CIKM, Indianapolis, USA, ACM.



## How might libraries serve 21st century information needs?

Knight Foundation News Challenge: Libraries, Sep 2014 – June 2016



https://www.newschallenge.org/challenge/libraries/winners

https://www.newschallenge.org/challenge/how-might-libraries-serve-

21st-century-information-needs/winning



## How can we provide better access to Web archives?

## **ALEXANDRIA** Vision and Research Questions

