

# Memento - Time Travel for the Web



**Herbert Van de Sompel**

<http://mementoweb.org/>

Memento is funded by  
The Library of Congress



Memento  
Herbert Van de Sompel



Memento wants to make it Easy  
to navigate the Web of the Past

Technical Specification

<https://datatracker.ietf.org/doc/draft-vandesompel-memento/>



Memento  
Herbert Van de Sompel

2



Cornell IS  
Today

Select Date  
May 14 2004

Cornell IS  
May 19 2004

Cornell University - Information Science - Overview

Requested: 1996

2011 12/31/1995 12/31/1995

Information Science

Overview

Information Science is an interdisciplinary department within the Faculty of Computing and Information Science. It brings together faculty, researchers and students who share an interest in studying digital information.

Information Science examines information systems in their social, cultural, economic, historical, legal, and political contexts. Computer science is an important part of the program, but the emphasis is on systems and their use, rather than on the technologies that underlie them.

The diagram symbolizes the three major areas of Information Science and some of the schools, departments and programs that contribute to each area.

**Information Systems** draws from Computer Science and Operations Research; **Human Computer Interaction** from Communication, Psychology, and Cognitive Studies; **Social Studies of Computing** from Science & Technology Studies, Law, and Economics, with many others.

For undergraduates, majors in Information Science are offered in Arts & Sciences (AS) and the College of Agriculture and Life Sciences (CALS). In Engineering, Information Science is one of the specializations in the major in Information Science, Systems, and Technology (ISST). All undergraduate colleges offer minors or concentrations in Information Science.

The Ph.D. program offers opportunity for advanced study and research in all areas of Information Science. The Master of Professional Studies is for students to study Information Science at an advanced level.

The Information Science department is home to several large research groups, including the Human-Computer Interaction group, digital libraries research, and the ePrint arXiv.

Information Science brings together faculty, students and researchers who share an interest in combining computer science with the social sciences of how people and society interact with information.

Arts & Sciences and the College of Agriculture & Life Sciences offer undergraduate majors in Information Science. Engineering students can take the major in Information Science, Systems, and Technology (ISST).

**"I encourage critical analytic thinking about ideas and evidence across any area of content being considered in the class. Students are asked to be skeptical and think about the logic and procedures behind scientific information."**

Jeffrey Hancock  
Associate Professor, Comm. Chair, IS

Cornell University - Information Science - Overview

Requested: 1996

2011 12/31/1995 12/31/1995

Information Science

Overview

Information Science is an interdisciplinary department within the Faculty of Computing and Information Science. It brings together faculty, researchers and students who share an interest in studying digital information.

Information Science examines information systems in their social, cultural, economic, historical, legal, and political contexts. Computer science is an important part of the program, but the emphasis is on systems and their use, rather than on the technologies that underlie them.

The diagram symbolizes the three major areas of Information Science and some of the schools, departments and programs that contribute to each area.

**Information Systems** draws from Computer Science and Operations Research; **Human Computer Interaction** from Communication, Psychology, and Cognitive Studies; **Social Studies of Computing** from Science & Technology Studies, Law, and Economics, with many others.

For undergraduates, majors in Information Science are offered in Arts & Sciences (AS) and the College of Agriculture and Life Sciences (CALS). In Engineering, Information Science is one of the specializations in the major in Information Science, Systems, and Technology (ISST). All undergraduate colleges offer minors or concentrations in Information Science.

The Ph.D. program offers opportunity for advanced study and research in all areas of Information Science. The Master of Professional Studies is for students to study Information Science at an advanced level.

The Information Science department is home to several large research groups, including the Human-Computer Interaction group, digital libraries research, and the ePrint arXiv.

Information Science brings together faculty, students and researchers who share an interest in combining computer science with the social sciences of how people and society interact with information.

Arts & Sciences and the College of Agriculture & Life Sciences offer undergraduate majors in Information Science. Engineering students can take the major in Information Science, Systems, and Technology (ISST).

**"I encourage critical analytic thinking about ideas and evidence across any area of content being considered in the class. Students are asked to be skeptical and think about the logic and procedures behind scientific information."**

Jeffrey Hancock  
Associate Professor, Comm. Chair, IS

Cornell Information Science

Requested: 1996

2011 12/31/1995 12/31/1995

Information Science

About Us

Quick Links

Feature

The HIMALAYA Data Mining Project

News & Events

Johannes Gebrike receives Provost Award for Distinguished Scholarship. (April 4, 2004)

Carl Lagoze wins the American Library Association's Frederick G. Kilgour Award for research in library and information technology. (March 24, 2004)

Geri Gay honored with Stephen H. Weiss Presidential Fellowship teaching award. (March 18, 2004)

©2004 Cornell University

From  
Internet Archive



Memento  
Herbert Van de Sompel



# Information Science

[contact](#) | [cis home](#)

- [home](#)
- [about](#)
- [undergraduate programs](#)
- [graduate program](#)
- [research](#)
- [faculty & researchers](#)

## About Us

- [Overview](#)
- [Seminar Series](#)
- [Related Talks](#)
- [Facilities](#)
- [Contact](#)

## Quick Links

- [Undergraduate Programs](#)
- [Graduate Programs](#)
- [Faculty & Researchers](#)
- [Research Groups](#)
- [Computing & Information Science](#)

## Feature



### The HIMALAYA Data Mining Project

How to extract information from pedabytes of data.

- [details](#)



Information Science brings together faculty, students and researchers who share an interest in combining computer science with the social sciences of how people and society interact with information.

Arts & Sciences and the College of Agriculture & Life Sciences offer undergraduate majors in Information Science. Engineering students can take the major in Information Science, Systems and Technology (ISST). Cornell's application to offer a Ph.D. is awaiting approval from New York State.

Information Science is a program of the [Faculty of Computing and Information Science](#).

- [Program Overview](#)

## News & Events

- [More News](#)
- [Seminar Series](#)

[Johannes Gehrke receives Provost Award for Distinguished Scholarship.](#) (April 4, 2004)

[Carl Lagoze wins the American Library Association's Frederick G. Kilgour Award for research in library and information technology.](#) (March 24, 2004)

[Geri Gay honored with Stephen H. Weiss Presidential Fellowship teaching award.](#) (March 18, 2004)

Memento achieves this by introducing  
a uniform version access capability to  
integrate the past and current Web



# Problem Statement ...

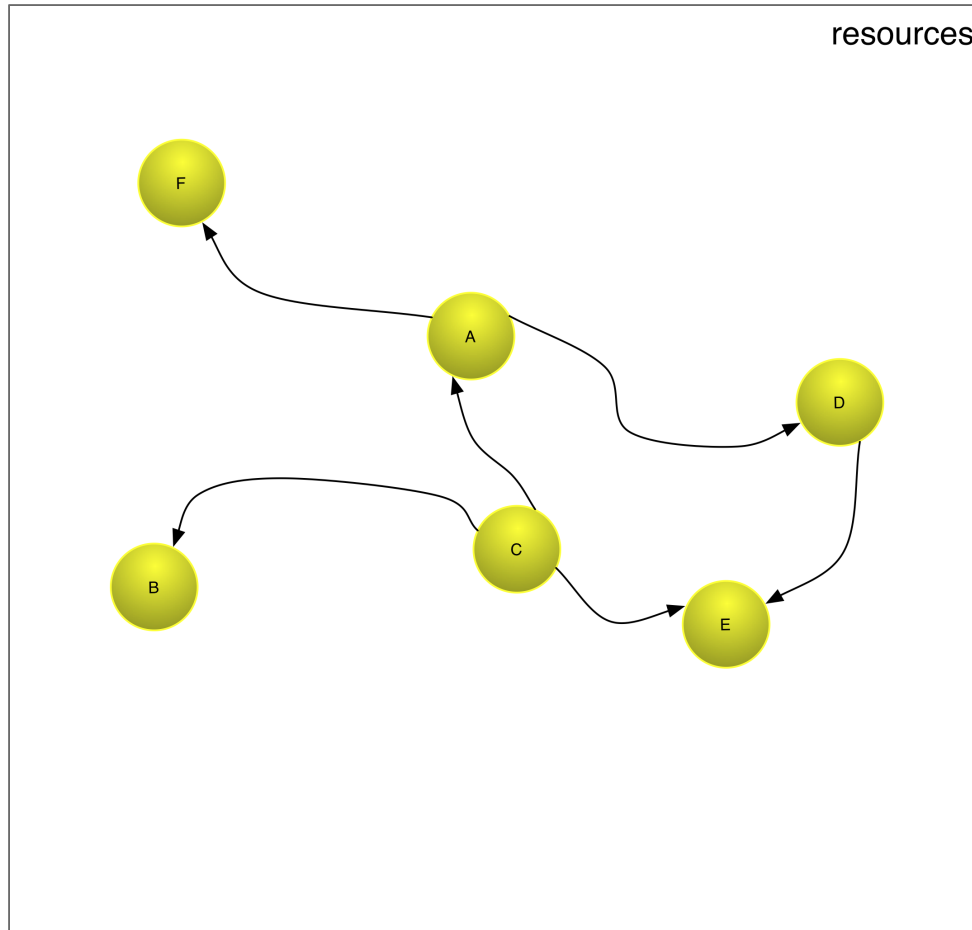


Memento  
Herbert Van de Sompel

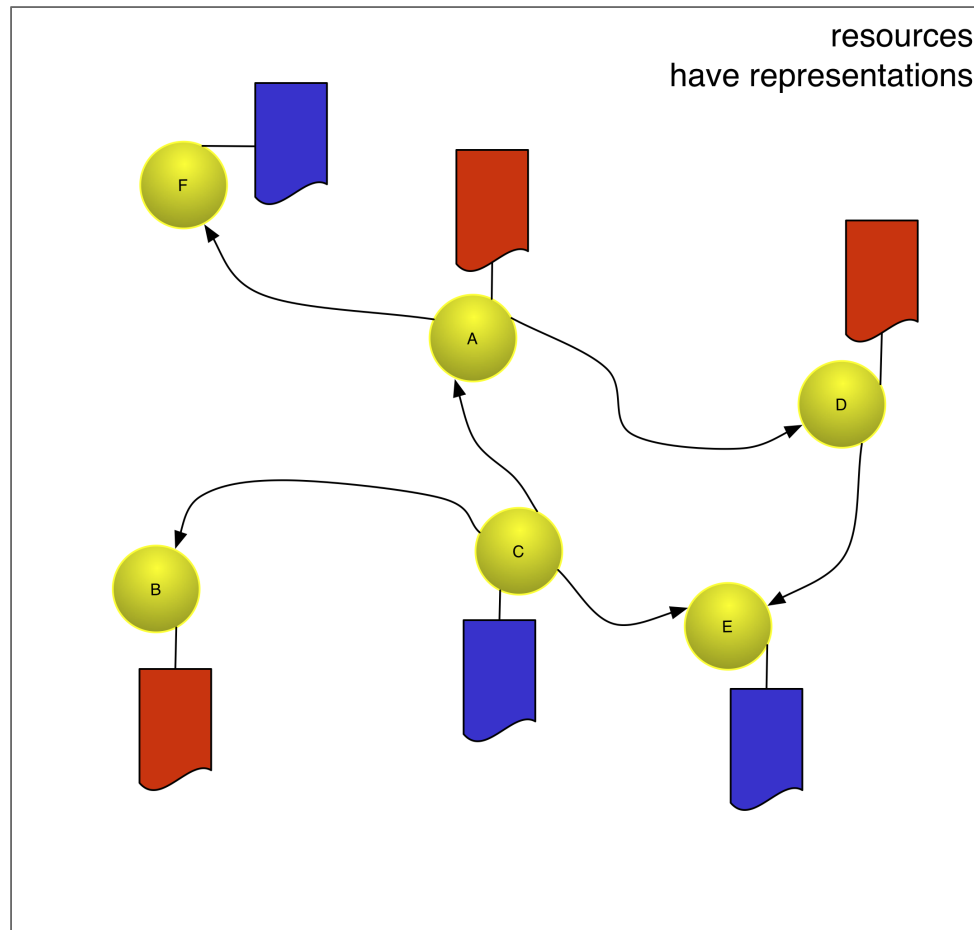
6



# Resources

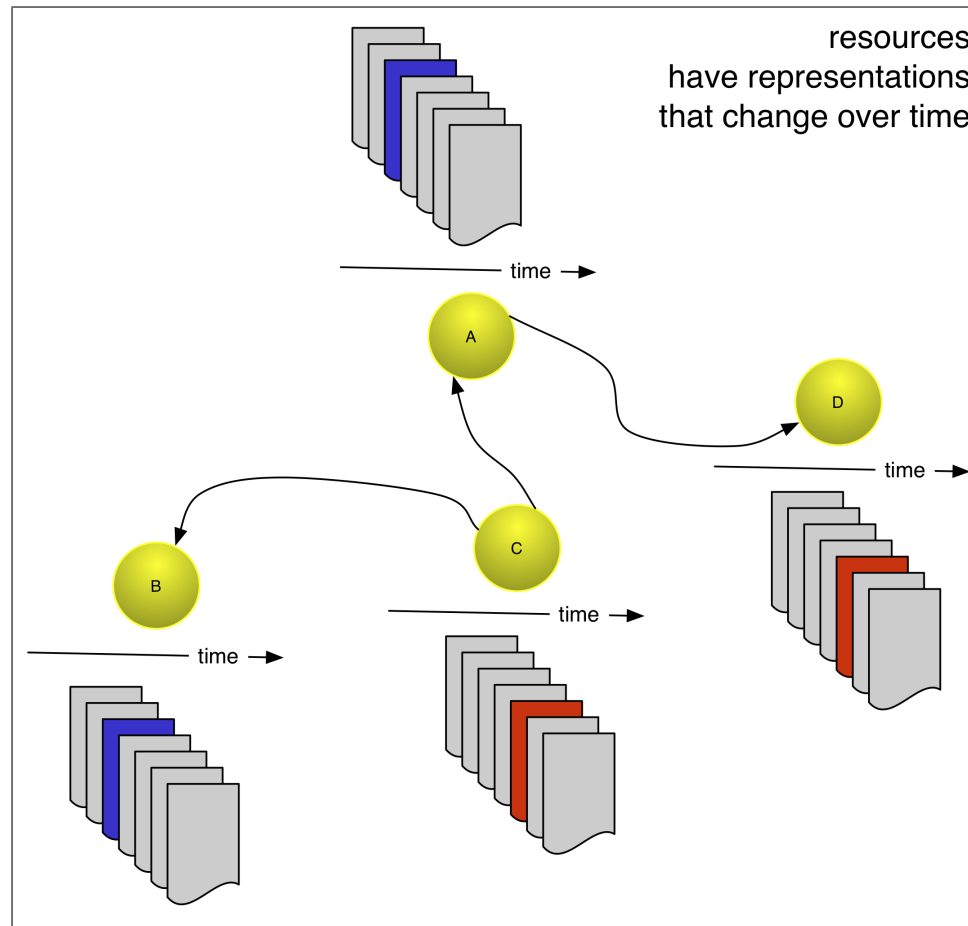


# Resources have Representations

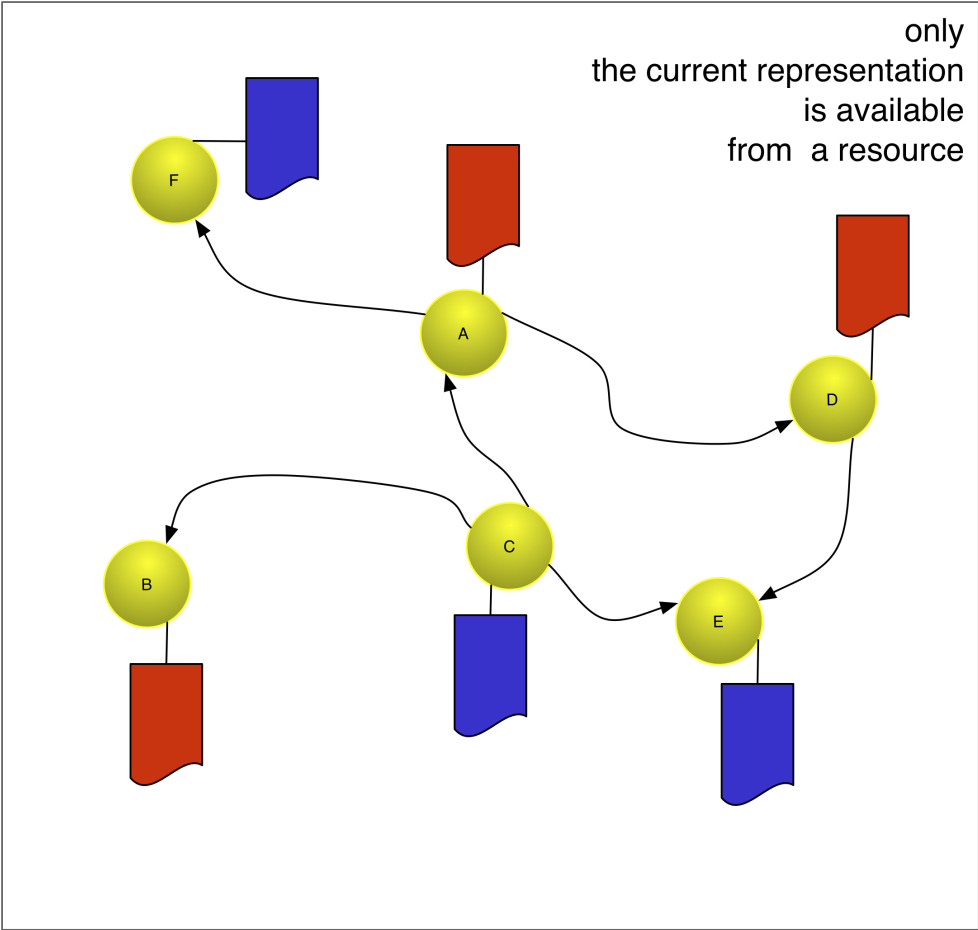




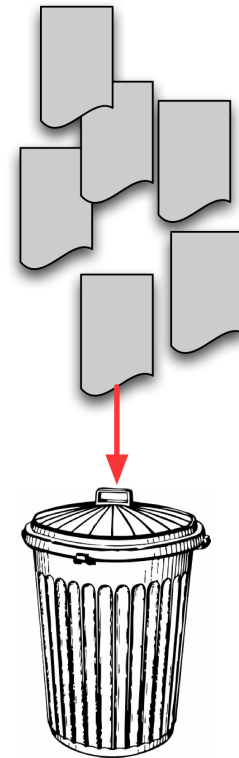
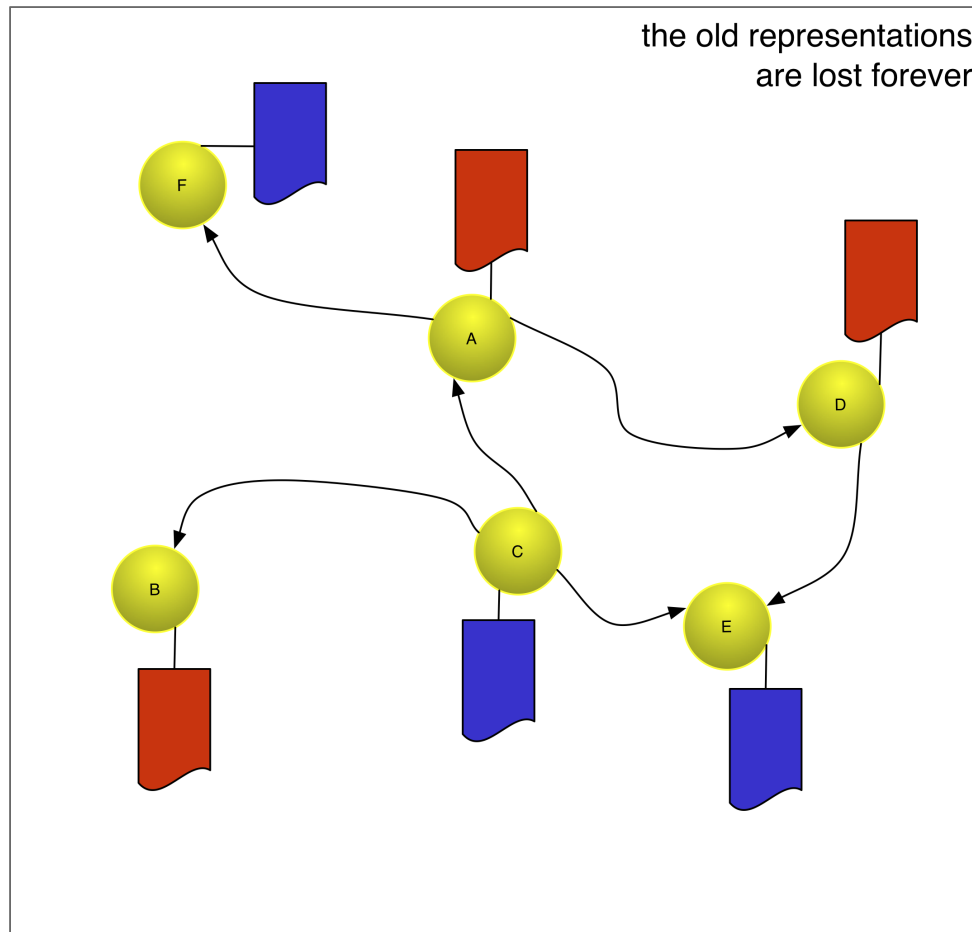
# Resources have Representations that Change over Time



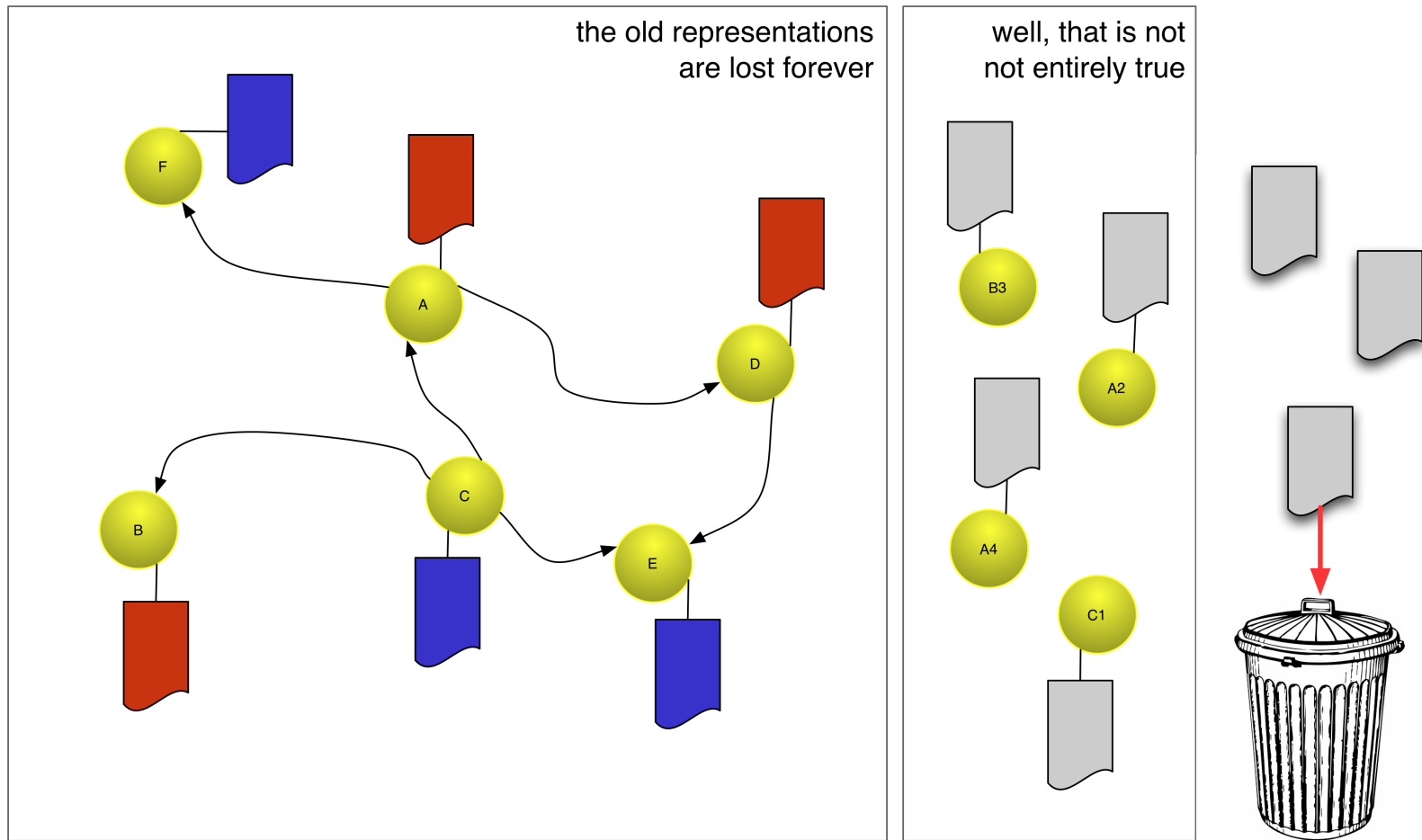
# Only the Current Representation is Available from a Resource



# Old Representations are Lost Forever



# Archived/Version Resources Exist



# Resource Versions on the Web



- Content Management Systems
- Web Archives
- Transactional archives
- Search engine caches
- ...



Sep 11 2001, 20:36:10 UTC

Dec 20 2001, 4:51:00 UTC

# Resource Versions

**CNN.com SPECIAL REPORT**

Updated: 5:42 p.m. EDT -- 11 September 2001

**AMERICA UNDER ATTACK**  
CNN EXCLUSIVE

**Terrorists strike New York, D.C.**  
At 8:45 a.m. EDT, the first of two airliners crashed into the World Trade Center, opening a horrifying and apparently coordinated terrorist attack on the United States, which saw the collapse of the two 110-story towers into surrounding Manhattan streets and a later attack on the Pentagon.

**COMPLETE COVERAGE**

- World Trade Center towers collapse after hit
- Crash destroys part of Pentagon
- White House, Capitol evacuated
- American, United confirm losing planes
- Bush: 'We'll hunt them down' | Statement
- U.S. officials: More attacks can't be ruled out
- Attacks strike financial markets

**U.S. SCENE**

- FAA grounds all U.S. flights until noon Wednesday
- U.S. military on 'high alert'
- Major sports events postponed

**WORLD SCENE**

- Taliban statement: 'Afghanistan feels your pain'
- World in shock | Evacuations amid chaos
- Middle East, Asia condemn attacks

**RESOURCES AND RESCUE**

- Giuliani: 1,500 'walking wounded' in New York
- Emergency contact information

**VIDEO**  
Watch video of second plane crash at World Trade Center  
[More audio/video >>](#)

**PHOTO GALLERY**  
Aftermath of the attacks in New York and Washington

**CHRONOLOGY**  
At 8:45 a.m. EDT, day of disaster begins  
[Map of attacks](#)

**EXTRA INFO**  
[World Trade Center, Pentagon, 767](#)

article | discussion | view source | history

Try Beta | Log in / create account

**September 11 attacks**  
From Wikipedia, the free encyclopedia

This is an old revision of this page, as edited by The Cunctator (talk | contribs) at 04:51, 20 December 2001. It may differ significantly from the current revision.  
(diff) ← Previous revision | Current revision (diff) | Newer revision → (diff)

**In Memoriam, September 11, 2001**

On the morning of September 11, 2001, what might well be the most devastating terrorist attack in the history of the world occurred concurrently in New York City, Washington, D.C. and near Pittsburgh. Four passenger jets were hijacked and then deliberately crashed into the World Trade Center and the Pentagon. Both towers of the World Trade Center subsequently collapsed, and part of the Pentagon was destroyed in the ensuing fire. Casualties are expected to be in the thousands: 265 on the planes; about 3000 people (early estimates ranged as high as 6500 people), including hundreds of firefighters who had rushed in, at the World Trade Center; and 125 at the Pentagon.

Some passengers on the doomed flights were able to make phone calls reporting on events on board. They reported that there were more than one hijacker on each plane and that they took control of the planes using box-cutter knives. It appears that the passengers on the fourth jet tried to overpower the hijackers and that the plane crashed in a sparsely populated area as a result, thereby missing its intended target, which may have been the White House.

The attack had immediate and deep global political effects and economic effects and inspired an international outpouring of memorials and services. The massive undertaking of rescue and recovery, and of providing assistance to the survivors and victims, is ongoing. There will be a great need for donations for a long time.

Though no group has claimed responsibility, the U.S. government immediately launched a full-scale response, stating its intentions to go to war against those responsible. On October 7, a coalition led by the United States launched an attack in Afghanistan. See 2001 U.S. Attack on Afghanistan.

Following the attack, the United States has been on heightened alert for new terrorist attacks. In late September, cases of anthrax started breaking out; although, no connection to the September 11 attack has been found. See 2001 anthrax attack.

<http://web.archive.org/web/20010911203610/http://www.cnn.com/> archived resource for <http://cnn.com>

[http://en.wikipedia.org/w/index.php?title=September\\_11\\_attacks&oldid=282333](http://en.wikipedia.org/w/index.php?title=September_11_attacks&oldid=282333) archived resource for [http://en.wikipedia.org/wiki/September\\_11\\_attacks](http://en.wikipedia.org/wiki/September_11_attacks)



# Versions are Not Integrated with the Web



# Architecture of the Web and Resource Versions



## Architecture of the World Wide Web, Volume One

W3C Recommendation 15 December 2004

**This version:**

<http://www.w3.org/TR/2004/REC-webarch-20041215/>

**Latest version:**

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

**Previous version:**

<http://www.w3.org/TR/2004/PR-webarch-20041105/>

**Editors:**

[Ian Jacobs](#), W3C  
Norman Walsh, Sun Microsystems, Inc.

Resource state may evolve over time. Requiring a URI owner to publish a new URI for each change in resource state would lead to a significant number of broken references. For robustness, Web architecture promotes independence between an identifier and the state of the identified resource.





# Finding Resource Versions

The screenshot shows the Internet Archive website. At the top, there are navigation links for 'Web', 'Moving Images', 'Texts', 'Audio', 'Software', 'Patron Info', 'About IA', and 'Projects'. Below these are 'Forums', 'FAQs', 'Contributions', 'Jobs', and 'Donate'. A search bar is present with the text 'http://cnn.com' and a 'GO!' button. The page also features 'Announcements', 'Web' (150 billion pages), and 'Welcome to the Archive' sections. At the bottom, there are four main categories: 'Moving Images' (222,980 movies), 'Live Music Archive' (70,456 concerts), 'Audio' (430,923 recordings), and 'Texts' (1,732,709 texts). Each category has a 'Browse' button and a 'Curator's Choice' section.

able here 6 months after collection. [See FAQ.](#)

Search Results for Jan 01, 1996 - May 14, 2009							
2000	2001	2002	2003	2004	2005	2006	2007
30 pages	2734 pages	231 pages	156 pages	625 pages	1186 pages	886 pages	953 pages
Jan 20, 2000	Mar 01, 2001	Jan 23, 2002	Jan 29, 2003	Jan 01, 2004	Jan 01, 2005	Jan 01, 2006	Jan 01, 2007
Jan 21, 2000	Mar 02, 2001	Feb 02, 2002	Jan 31, 2003	Jan 02, 2004	Jan 01, 2005	Jan 01, 2006	Jan 01, 2007
Jan 21, 2000	Mar 05, 2001	Feb 09, 2002	Feb 01, 2003	Jan 03, 2004	Jan 02, 2005	Jan 02, 2006	Jan 02, 2007
Jan 21, 2000	Mar 07, 2001	Feb 10, 2002	Feb 01, 2003	Jan 10, 2004	Jan 02, 2005	Jan 02, 2006	Jan 02, 2007
Jan 21, 2000	Mar 09, 2001	Feb 11, 2002	Feb 02, 2003	Jan 11, 2004	Jan 03, 2005	Jan 03, 2006	Jan 02, 2007
Jul 06, 2000	Mar 31, 2001	Feb 13, 2002	Feb 02, 2003	Jan 12, 2004	Jan 04, 2005	Jan 03, 2006	Jan 03, 2007
Jul 07, 2000	Apr 10, 2001	Feb 13, 2002	Feb 03, 2003	Jan 12, 2004	Jan 05, 2005	Jan 03, 2006	Jan 03, 2007
Jul 11, 2000	Apr 13, 2001	Feb 14, 2002	Feb 05, 2003	Jan 13, 2004	Jan 06, 2005	Jan 04, 2006	Jan 04, 2007
Aug 04, 2000	Apr 13, 2001	Feb 15, 2002	Feb 06, 2003	Jan 15, 2004	Jan 06, 2005	Jan 04, 2006	Jan 05, 2007
Aug 15, 2000	Apr 13, 2001	Feb 16, 2002	Feb 07, 2003	Jan 16, 2004	Jan 06, 2005	Jan 04, 2006	Jan 05, 2007
Aug 15, 2000	Apr 14, 2001	Feb 20, 2002	Jan 17, 2003	Jan 17, 2004	Jan 07, 2005	Jan 04, 2006	Jan 06, 2007
Aug 15, 2000	Apr 14, 2001	Feb 22, 2002	Feb 10, 2003	Jan 20, 2004	Jan 07, 2005	Jan 05, 2006	Jan 06, 2007
Aug 15, 2000	Apr 14, 2001	Feb 22, 2002	Feb 11, 2003	Jan 20, 2004	Jan 08, 2005	Jan 05, 2006	Jan 06, 2007
Aug 16, 2000	Apr 18, 2001	Feb 23, 2002	Feb 12, 2003	Jan 20, 2004	Jan 09, 2005	Jan 05, 2006	Jan 07, 2007
Aug 16, 2000	May 03, 2001	Feb 23, 2002	Feb 12, 2003	Jan 24, 2004	Jan 09, 2005	Jan 06, 2006	Jan 07, 2007
Aug 17, 2000	May 03, 2001	Feb 23, 2002	Jan 27, 2003	Jan 27, 2004	Jan 10, 2005	Jan 06, 2006	Jan 07, 2007
Aug 17, 2000	May 03, 2001	Feb 24, 2002	Feb 15, 2003	Jan 28, 2004	Jan 10, 2005	Jan 06, 2006	Jan 07, 2007
Aug 18, 2000	May 03, 2001	Feb 26, 2002	Mar 20, 2003	Jan 30, 2004	Jan 10, 2005	Jan 06, 2006	Jan 07, 2007
Aug 28, 2000	May 03, 2001	Feb 27, 2002	Jan 30, 2003	Jan 30, 2004	Jan 11, 2005	Jan 07, 2006	Jan 08, 2007
Sep 20, 2000	May 03, 2001	Mar 27, 2002	Mar 21, 2003	Jan 31, 2004	Jan 11, 2005	Jan 07, 2006	Jan 08, 2007
Sep 30, 2000	May 03, 2001	Mar 29, 2002	Mar 24, 2003	Feb 02, 2004	Jan 12, 2005	Jan 07, 2006	Jan 08, 2007
Oct 18, 2000	May 04, 2001	May 25, 2002	Mar 25, 2003	Feb 02, 2004	Jan 12, 2005	Jan 07, 2006	Jan 08, 2007
Nov 09, 2000	May 04, 2001	May 27, 2002	Mar 26, 2003	Feb 03, 2004	Jan 13, 2005	Jan 08, 2006	Jan 08, 2007
Nov 10, 2000	May 04, 2001	May 27, 2002	Mar 29, 2003	Feb 04, 2004	Jan 13, 2005	Jan 09, 2006	Jan 08, 2007
Nov 10, 2000	May 04, 2001	May 28, 2002	Mar 29, 2003	Feb 07, 2004	Jan 13, 2005	Jan 09, 2006	Jan 08, 2007
Nov 17, 2000	May 04, 2001	May 29, 2002	Mar 30, 2003	Feb 10, 2004	Jan 14, 2005	Jan 10, 2006	Jan 08, 2007
Nov 18, 2000	May 04, 2001	May 30, 2002	Mar 31, 2003	Feb 10, 2004	Jan 14, 2005	Jan 10, 2006	Jan 09, 2007
Dec 04, 2000	May 04, 2001	May 30, 2002	Apr 01, 2003	Feb 11, 2004	Jan 14, 2005	Jan 10, 2006	Jan 09, 2007
Dec 04, 2000	May 04, 2001	May 30, 2002	Apr 04, 2003	Feb 13, 2004	Jan 15, 2005	Jan 10, 2006	Jan 09, 2007
Dec 04, 2000	May 05, 2001	May 31, 2002	Apr 04, 2003	Feb 13, 2004	Jan 15, 2005	Jan 10, 2006	Jan 09, 2007
	May 05, 2001	Jun 02, 2002	Apr 05, 2003	Feb 14, 2004	Jan 15, 2005	Jan 10, 2006	Jan 10, 2007
	May 05, 2001	Jun 02, 2002	Apr 08, 2003	Feb 14, 2004	Jan 16, 2005	Jan 10, 2006	Jan 10, 2007
	May 05, 2001	Jun 03, 2002	Apr 08, 2003	Feb 17, 2004	Jan 16, 2005	Jan 11, 2006	Jan 11, 2007
	May 05, 2001	Jun 03, 2002	Apr 09, 2003	Feb 18, 2004	Jan 17, 2005	Jan 11, 2006	Jan 11, 2007
	May 05, 2001	Jun 04, 2002	Apr 09, 2003	Feb 20, 2004	Jan 17, 2005	Jan 11, 2006	Jan 12, 2007
	May 05, 2001	Jun 06, 2002	Apr 11, 2003	Feb 21, 2004	Jan 18, 2005	Jan 11, 2006	Jan 12, 2007
	May 05, 2001	Jul 01, 2002	Apr 20, 2003	Feb 28, 2004	Jan 18, 2005	Jan 11, 2006	Jan 12, 2007
	May 05, 2001	Jul 02, 2002	Apr 23, 2003	Mar 02, 2004	Jan 19, 2005	Jan 11, 2006	Jan 12, 2007
	May 05, 2001	Jul 03, 2002	Apr 23, 2003	Mar 07, 2004	Jan 19, 2005	Jan 12, 2006	Jan 13, 2007

Go to <http://www.archive.org/> and search <http://cnn.com>

On [http://web.archive.org/web/\\*/http://cnn.com](http://web.archive.org/web/*/http://cnn.com), select desired datetime



# Finding Resource Versions

The screenshot shows the main article page for "September 11 attacks". The title is "September 11 attacks" with a lock icon. Below the title, it says "From Wikipedia, the free encyclopedia". A note indicates that "9/11" redirects here. The main text describes the September 11 attacks as a series of coordinated suicide attacks by Al-Qaeda on September 11, 2001. An image shows the Twin Towers of the World Trade Center burning. A table below the image provides details: Location (New York City, Arlington County, Virginia; and near Shanksville, Pennsylvania), Date (Tuesday, September 11, 2001, 8:46 am - 10:28 am (UTC-4)), Attack type (Aircraft hijacking, Mass murder, Suicide attack), Death(s) (2,995 including 19 hijackers), Injured (6,000+), and Perpetrator(s) (al-Qaeda led by Osama bin Laden).

The screenshot shows the "Revision history of September 11 attacks" page. It includes a search bar for browsing history, with filters for "From year (and earlier)" and "From month (and earlier)". Below the search bar, there is a list of revisions. Each entry shows the date, time, and user who made the edit, along with the size of the edit and a link to view the revision. The list includes several revisions from 2009, such as one by InnerParty on 9 November 2009 and another by WordyGirl90 on 8 November 2009.



Go to  
[http://en.wikipedia.org/wiki/September\\_11\\_attacks](http://en.wikipedia.org/wiki/September_11_attacks)  
 and click History

Browse History



Memento  
 Herbert Van de Sompel

Dec 20 2001, 4:51:00 UTC

current

# Navigating Resource Versions

September 11 attacks

From Wikipedia, the free encyclopedia

This is an old revision of this page, as edited by The Cunctator (talk | contribs) at 04:51, 20 December 2001. It may differ significantly from the current revision.

**In Memoriam, September 11, 2001**

On the morning of **September 11, 2001**, what might well be the most devastating **terrorist attack** in the history of the world occurred concurrently in **New York City**, **Washington, D.C.** and **Utah**. **Four** **jet** **aircraft** were **hijacked** and then deliberately crashed into the **World Trade Center** subsequently collapsed, and the **Pentagon** in **Arlington, Virginia**. **Casualties** are expected to be in the thousands: 265 on the planes; about 3000 people (early estimates ranged as high as 6500 people), including hundreds of firefighters who had rushed in, at the World Trade Center; and 125 at the Pentagon.

Some passengers on the doomed flights were able to make phone calls reporting on events on board. They reported that there were more than one hijacker on each plane and that they took control of the planes using **box-cutter knives**. It appears that the passengers on the fourth jet tried to overpower the hijackers and that the plane crashed in a sparsely populated area as a result, thereby missing its intended target, which may have been the **White House**.

The attack had immediate and deep global **political effects** and **economic effects** and inspired an international outpouring of **memorials and services**. The massive undertaking of **rescue and recovery**, and of providing **assistance** to the **survivors** and victims, is ongoing. There will be a great need for **donations** for a long time.

Though no group has claimed **responsibility**, the U.S. government immediately launched a full-scale **response**, stating its intentions to go to war against those responsible. On October 7, a coalition led by the United States launched an attack in Afghanistan. See **2001 U.S. Attack on Afghanistan**.

Following the attack, the United States has been on heightened alert for new terrorist attacks. In late September, **cases of anthrax** started breaking out; although, no connection to the September 11 attack has been found. See **2001 anthrax attack**.

The Pentagon

From Wikipedia, the free encyclopedia

*This article is about the United States military building. For the shape, see Pentagon.*

**The Pentagon** is the headquarters of the **United States Department of Defense**, located in **Arlington, Virginia**. As a symbol of the U.S. military, "the Pentagon" is often used **metonymically** to refer to the Department of Defense rather than the building itself.

Designed by the American architect **George Bergstrom** (1876 – 1955), and built by **Philadelphia, Pennsylvania**, general contractor **John McShain**, the building was dedicated on January 15, 1943, after ground was broken for construction on September 11, 1941. General **Brehon Somervell** provided the major motive power behind the project;<sup>[1]</sup> Colonel **Leslie Groves** and Major **Clarence Renshaw** were responsible for overseeing the project for the Army.

The Pentagon is the world's largest office building by floor area<sup>[2][3]</sup>, by a margin of about 35,580 square metres (383,000 sq ft). Approximately 23,000 military and civilian employees<sup>[3]</sup> and about 3,000 non-defense support personnel work in the Pentagon. It has five sides, five floors above ground (plus two basement levels), and five ring corridors per floor with a total of 17.5 miles (28.2 km)<sup>[3]</sup> of corridors. The Pentagon includes a five-acre (20,000 m²) central plaza, which is shaped like a pentagon and informally known as "**ground zero**", a nickname originating during the **Cold War** and based on the presumption that the **Soviet Union** would target one or more nuclear missiles at this central location.

On **September 11, 2001**, hijacked **American Airlines Flight 77** was crashed into the western side of the

<b>Type</b>	Headquarters
<b>Coordinates</b>	<span><span><span><span><span>38°52′15.56″N</span> <span>77°3′21.46″W</span></span></span><span><span>﻿</span> / <span>﻿</span></span><span><span>38.8709722°N 77.0557111°W</span><span><span>﻿</span> / <span>38.8709722; -77.0557111</span></span></span></span></span>
<b>Built</b>	1943 (broke ground on September 11, 1941)
<b>Height</b>	Five floors above ground and Two floors below ground
<b>In use</b>	1943–present

[http://en.wikipedia.org/w/index.php?title=September\\_11\\_attacks&oldid=282333](http://en.wikipedia.org/w/index.php?title=September_11_attacks&oldid=282333) archived resource for [http://en.wikipedia.org/wiki/September\\_11\\_attacks3](http://en.wikipedia.org/wiki/September_11_attacks3)

[http://en.wikipedia.org/wiki/The\\_Pentagon](http://en.wikipedia.org/wiki/The_Pentagon)



Sep 11 2001, 20:36:10 UTC

Sep 11 2001, 21:38:55 UTC

# Navigating Resource Versions

**CNN.com SPECIAL REPORT**

Updated: 5:42 p.m. EDT -- 11 September 2001

**AMERICA UNDER ATTACK**  
CNN EXCLUSIVE

**Terrorists strike New York, D.C.**  
At 8:45 a.m. EDT, the first of two airliners crashed into the World Trade Center, opening a horrifying and apparently coordinated terrorist attack on the United States, which saw the collapse of the two 110-story towers into surrounding Manhattan streets and a later attack on the Pentagon.

**COMPLETE COVERAGE**

- World Trade Center towers collapse after hit
- Crash destroys part of Pentagon
- White House, Capitol evacuated
- American, United confirm losing planes
- Bush: 'We'll hunt them down' | Statement
- U.S. officials: More attacks can't be ruled out
- Attack strikes financial markets

**U.S. SCENE**

- FAA grounds all U.S. flights until noon Wednesday
- U.S. military on 'high alert'
- Major sports events postponed

**WORLD SCENE**

- Taliban statement: 'Afghanistan feels your pain'
- World in shock | Evacuations amid chaos
- Middle East, Asia condemn attacks

**RESOURCES AND RESCUE**

Giuliani: 1,500 'walking wounded' in New York  
[Emergency contact information](#)

**VIDEO**  
Watch video of second plane crash at World Trade Center  
[More audio/video >>](#)

**PHOTO GALLERY**  
Aftermath of the attacks in New York and Washington

**CHRONOLOGY**  
At 8:45 a.m. EDT, day of disaster begins  
[Map of attacks](#)

**EXTRA INFO**  
[World Trade Center, Pentagon, 767](#)

**SPACE**

**CNN.com / SPACE**

Updated: 04:56 PM EDT 20:56 GMT -- 11 Sep 2001

**Cool star chills stellar theories**

A star besieged by unbelievably hot waves of nuclear explosions is ten times cooler than expected, forcing astronomers to rethink their ideas about stellar physics.

[FULL STORY #4](#)

**TOP STORIES**

- NASA shuts down in wake of attacks
- Sleuths try to unmask new spy satellites
- Probe limps to close encounter with comet
- Report: Plane tracking satellites can be jammed

[More News >](#)

**ALPHA EXTRA**  
Space correspondent Miles O'Brien on the space station tour

**SCENES FROM SPACE**  
A NASA telescope is marking a milestone

**RESOURCES:**

**SPACE AGENCY NEWS**  
The latest news releases from worldwide space agencies.  
[Hit the links](#)

**SPACE CALENDARS**  
Calendars of upcoming launches and other events from worldwide space agencies.  
[Calendar list](#)

**INTERACTIVE LIBRARY**  
Your window to the universe -- launch video, 360-degree photo tours and more.

**NASA WEBCAMS**  
Live pictures of NASA projects.  
[Launch it](#)

**SPACE PROBES**  
A guide to active unmanned missions.  
[Probe list](#)

**MORE NEWS**  
Space program subject of 'Astronauts' pilot

<http://web.archive.org/web/20010911203610/http://www.cnn.com/> archived resource for <http://cnn.com>

<http://web.archive.org/web/20010911213855/www.cnn.com/TECH/space/>



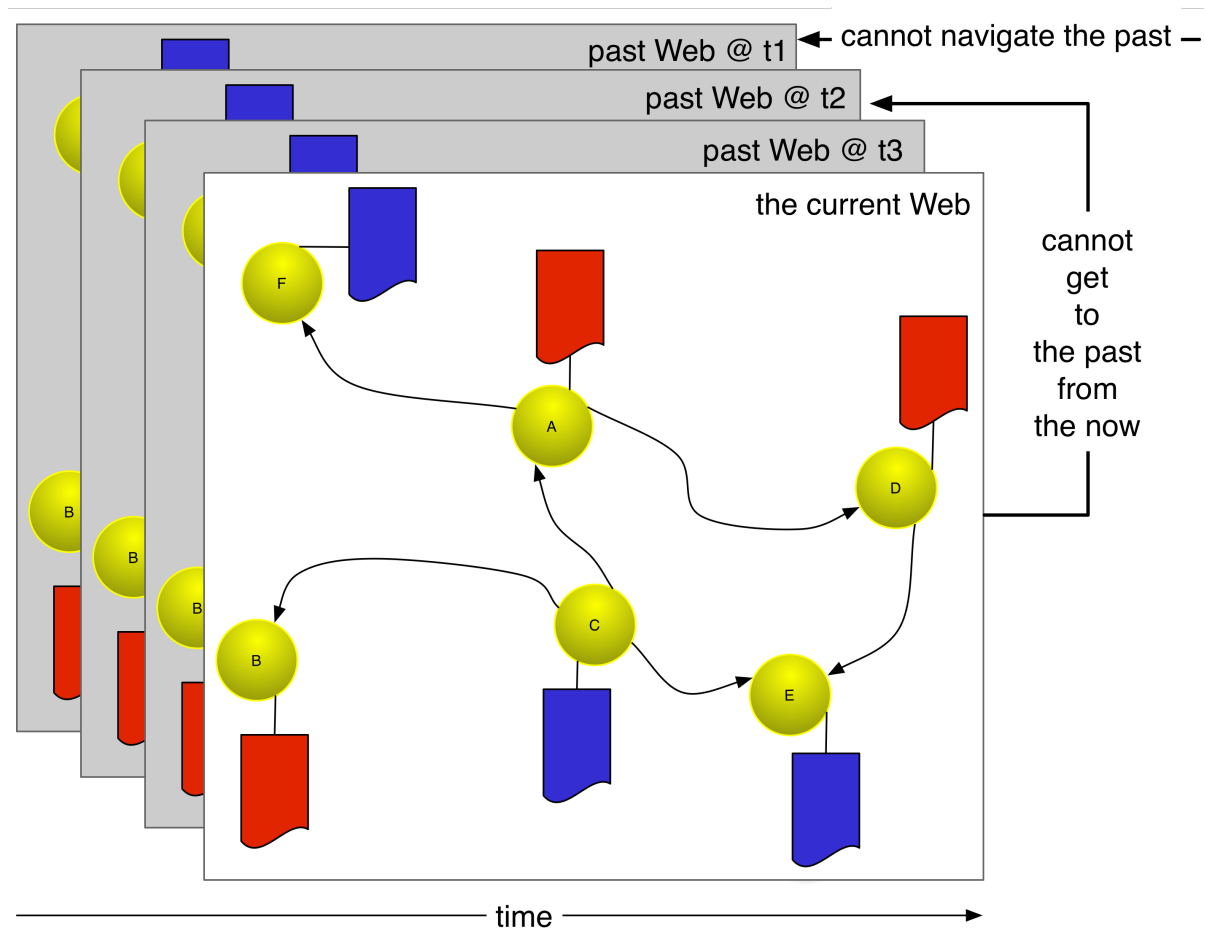
# Versions are Not Integrated with the Web



- Functionally:
  - Need to know about archives
  - Need to search each of them individually
  - Cannot consistently navigate in the past within an archive
  - Cannot navigate across archives
- Architecturally:
  - Cannot talk about a resource as it used to exist
  - Cannot access a prior version knowing the current one
  - Cannot access the current version knowing a prior one



# Memento Wants to Integrate the Past and Current Web



# The Memento Framework:

## Protocol to Integrate Past and Current Web

### Overview



# Memento Framework



- Regards the Web as a big Content Management System
- Introduces a uniform capability to access versions on the Web
- Does not build new archives but leverages existing systems that host versions





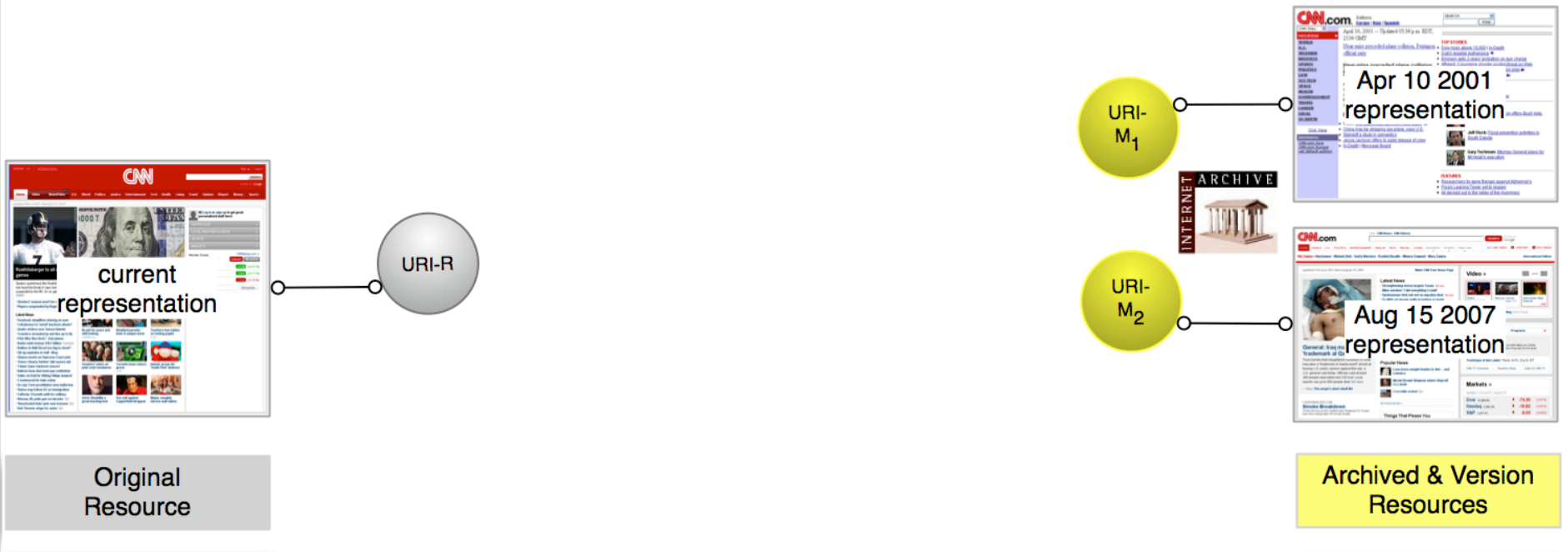
# Memento Framework



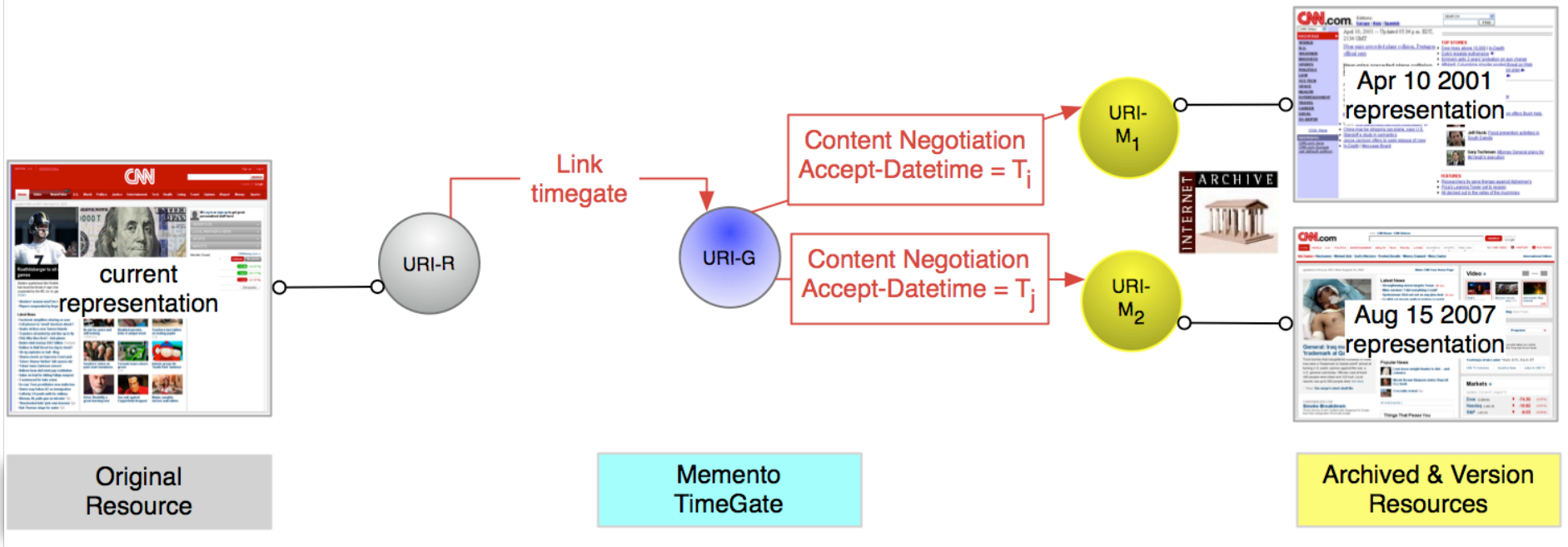
- Is distributed: versions may exist on several servers
- Uses time as a global version indicator
- Is based on the primitives of the Web: resource, resource state, representation, content negotiation, link



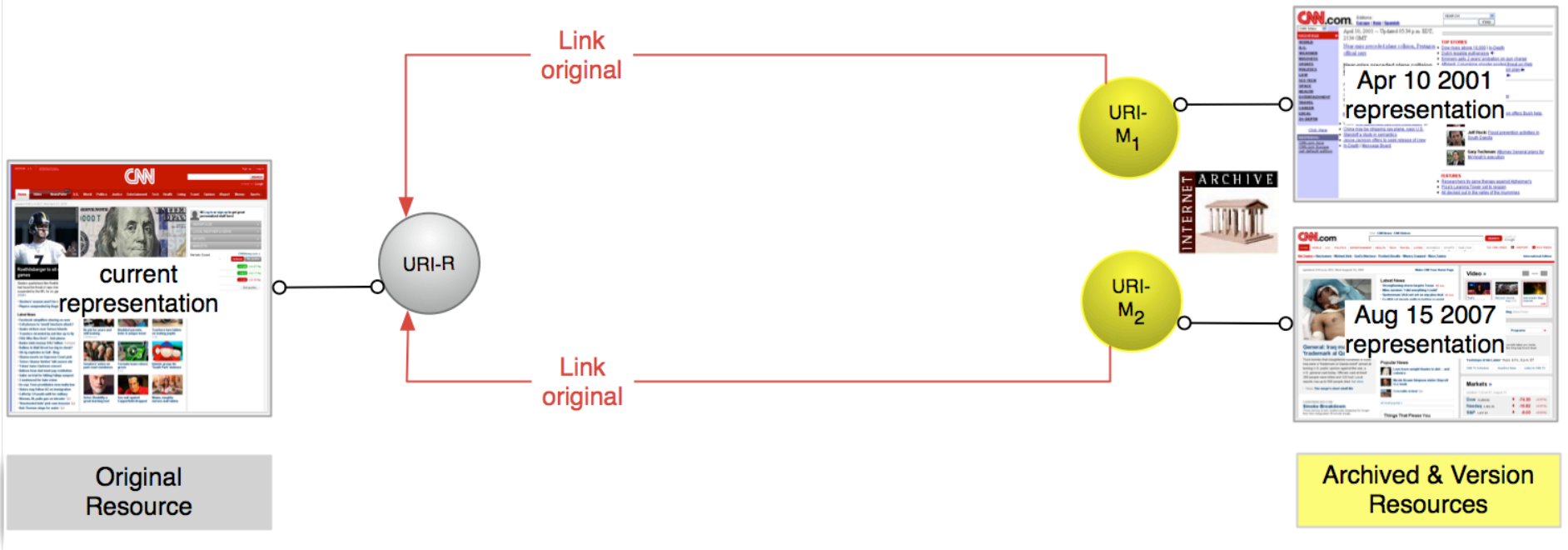
# Original Resources and Mementos



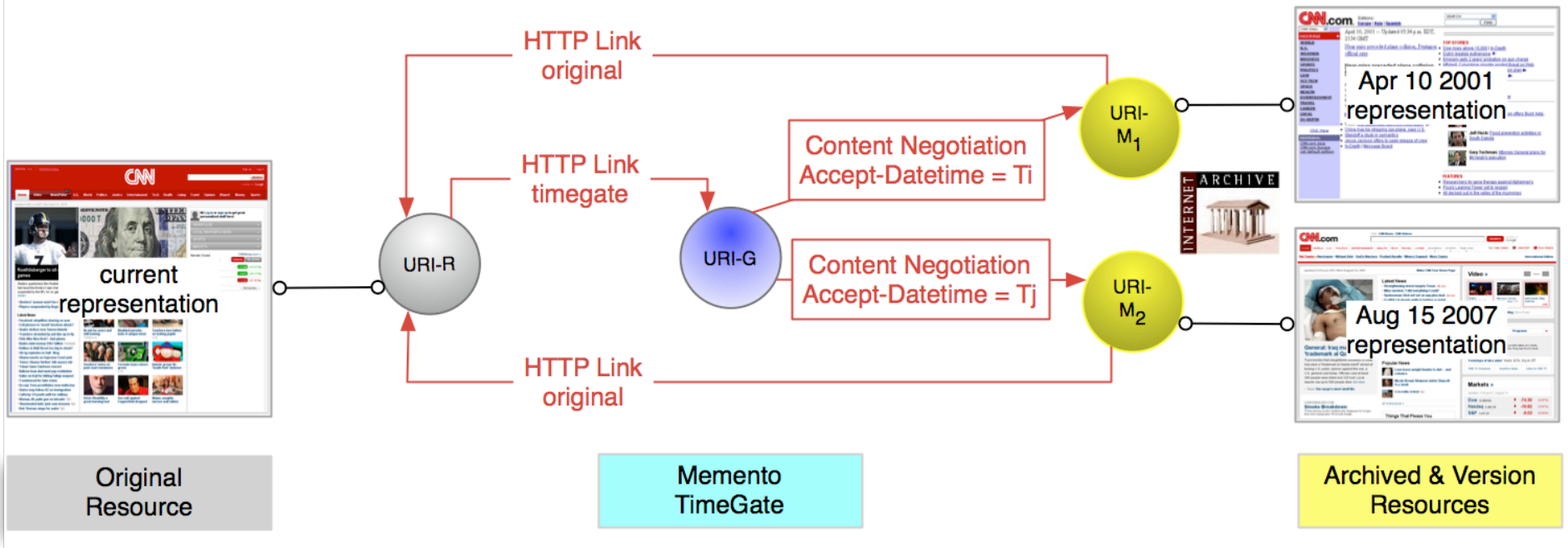
# Bridge from Present to Past



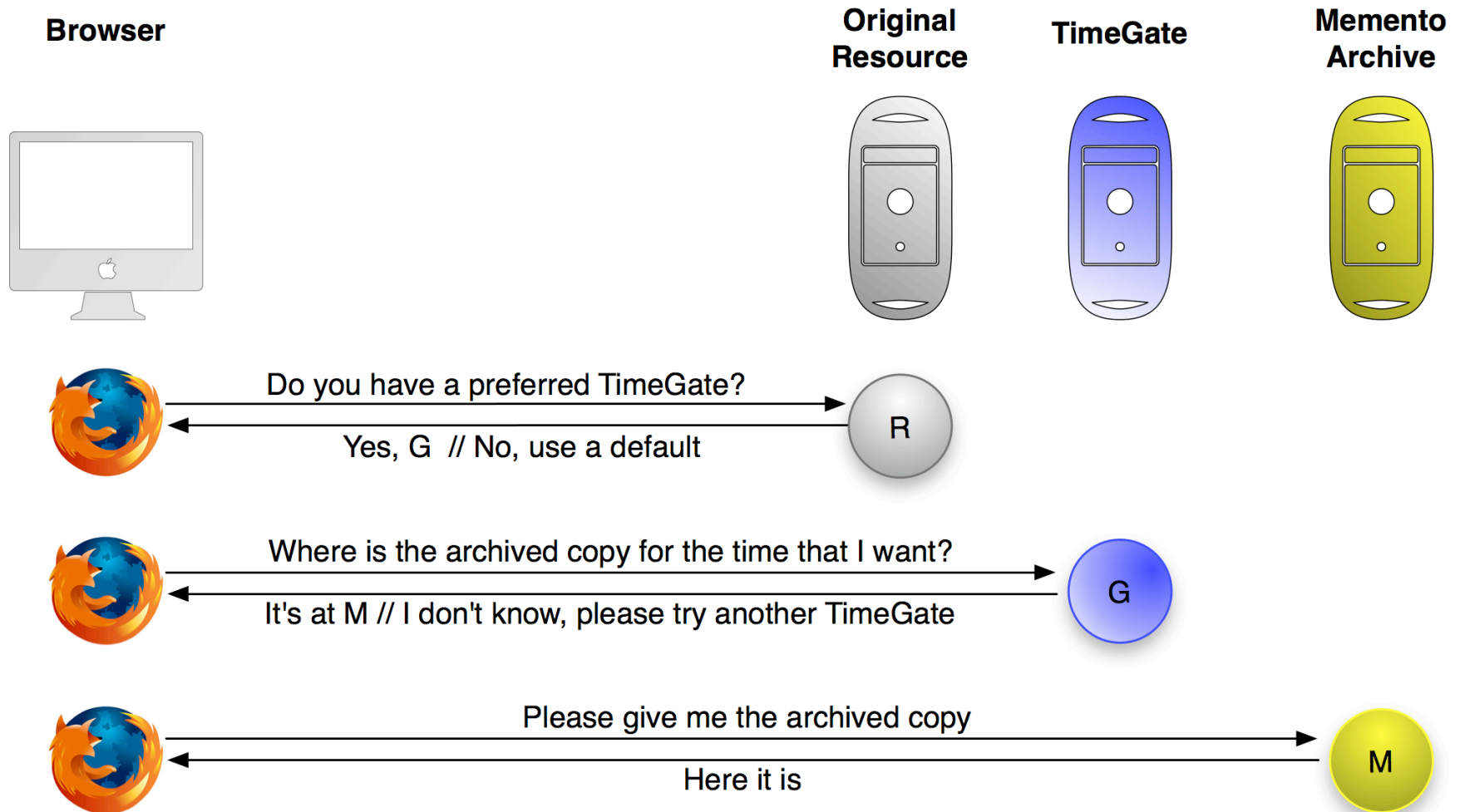
# Bridge from Past to Present



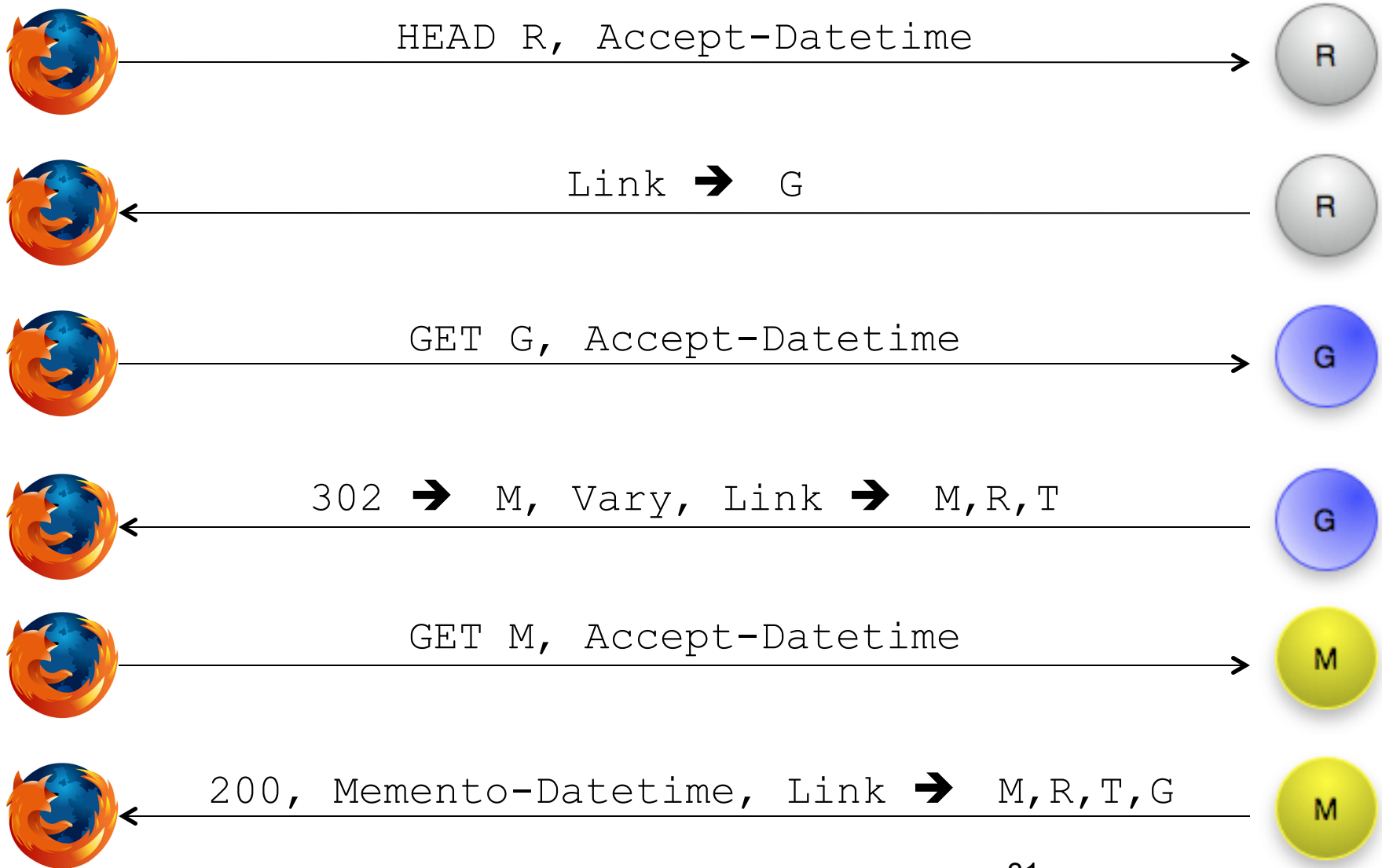
# Memento Framework



# Memento Client Server Interaction



# Memento HTTP Flow



# The Memento Framework:

Protocol to Integrate Past and Current Web

## HTTP Headers





# HTTP Headers used in Memento

- Defines two new headers:
  - request: `Accept-Datetime`
  - response: `Memento-Datetime`
- Introduce new content for two existing headers:
  - response: `Vary ; Link`
- Use one existing headers without modification:
  - response: `Location`



# HTTP Request Headers: Accept-Datetime

- Accept-Datetime
  - Issued against TimeGate, (Original Resource), (Memento)
  - Header value:
    - Desired datetime of Memento (MANDATORY)  
Must be in RFC 1123 format and in GMT
    - Interval indicator to express the client is only interested in Mementos within the interval (OPTIONAL)
      - Expressed as two ISO8601 durations:  
"-P3DT5H;+P2DT6H"

```
Accept-Datetime: Mon, 12 Oct 2009 14:20:33 GMT
```



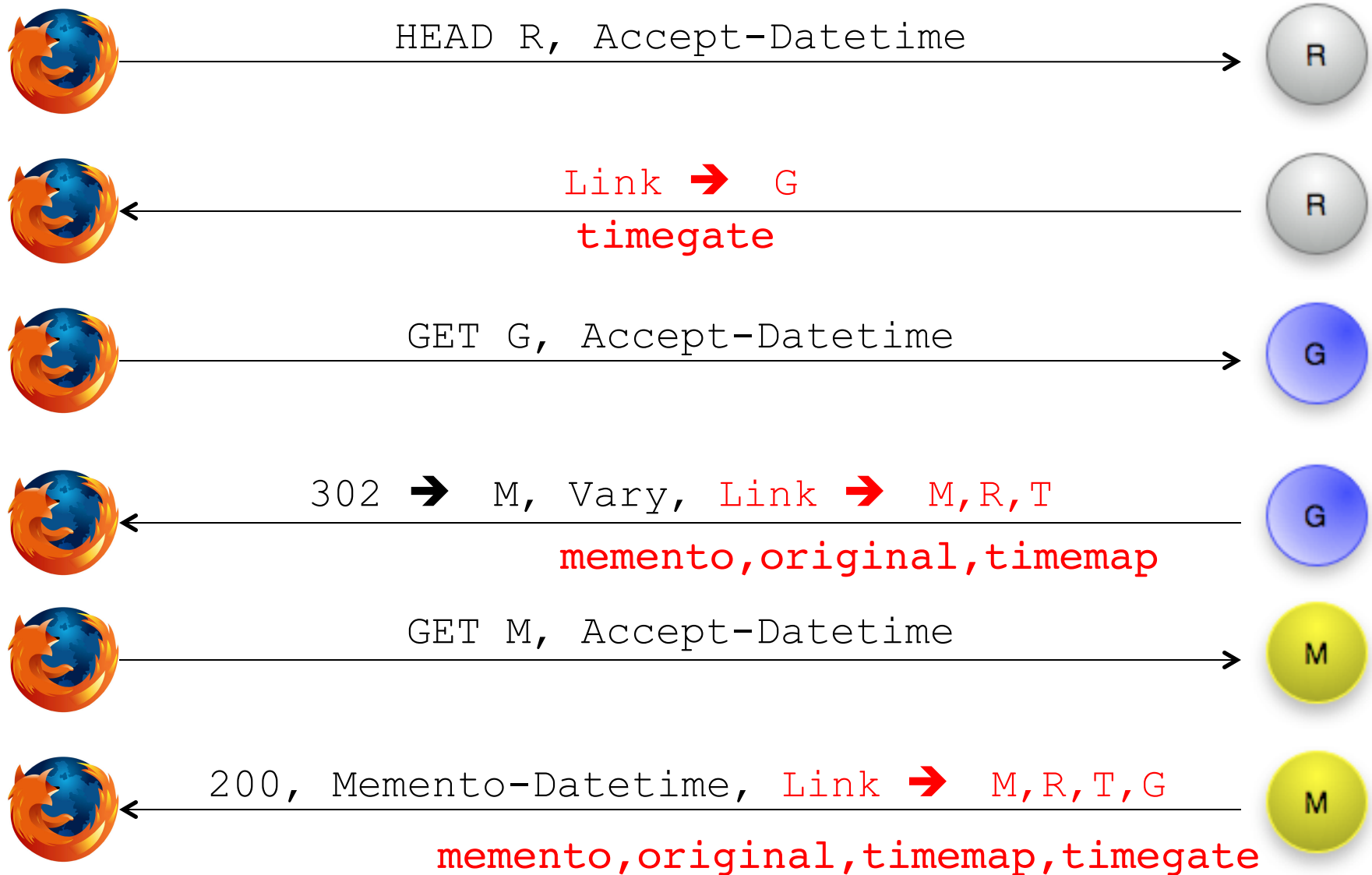
# HTTP Response Headers: Link

- Link
  - Returned by Original Resource, TimeGate and Mementos
  - Various new Relation Types are introduced:
    - "original"
    - "timegate"
    - "memento"
    - "timemap"
  - Rules about usage of these Links for Original Resources, TimeGates, and Mementos
  - HTTP Link Header: RFC 5988  
See: <https://datatracker.ietf.org/doc/rfc5988/>

```
Link: <http://web.archive.org/web/20010911223004/http://  
cnn.com>;rel="memento";datetime="Mon, 11 Sep 2001 22:30:04 GMT"
```



# Memento HTTP Flow



# The Memento Framework:

Protocol to Integrate Past and Current Web

## HTTP Interactions



# Memento HTTP Flow: Step 1



```
HEAD / HTTP/1.1
Host: a.example.org
Accept-Datetime: Tue, 11 Sep 2001 20:35:00 GMT
Connection: close
```



Memento  
Herbert Van de Sompel



## Memento HTTP Flow: Step 2



```
HTTP/1.1 200 OK
Date: Thu, 21 Jan 2010 00:02:12 GMT
Server: Apache
Link: <http://arxiv.example.net/timegate/http://a.example.org>
      ; rel="timegate"
Content-Length: 255
Connection: close
Content-Type: text/html; charset=iso-8859-1
```



Memento  
Herbert Van de Sompel



## Memento HTTP Flow: Step 3



```
GET /timegate/http://a.example.org
HTTP/1.1
Host: arxiv.example.net
Accept-Datetime: Tue, 11 Sep 2001 20:35:00 GMT
Connection: close
```



Memento  
Herbert Van de Sompel





## Memento HTTP Flow: Step 4



```
HTTP/1.1 302 Found
Date: Thu, 21 Jan 2010 00:06:50 GMT
Server: Apache
Vary: negotiate, accept-datetime
Location:
  http://arxiv.example.net/web/20010911203610/http://a.example.org
Link: <http://a.example.org>; rel="original",
  <http://arxiv.example.net/web/20000915112826/http://a.example.org>
  ; rel="first memento"; datetime="Tue, 15 Sep 2000 11:28:26 GMT",
  <http://arxiv.example.net/web/20080708093433/http://a.example.org>
  ; rel="last memento"; datetime="Tue, 08 Jul 2008 09:34:33 GMT",
  <http://arxiv.example.net/timemap/http://a.example.org>
  ; rel="timemap"; type="application/link-format",
  <http://arxiv.example.net/web/20010911203610/http://a.example.org>
  ; rel="memento"; datetime="Tue, 11 Sep 2001 20:36:10 GMT",
  <http://arxiv.example.net/web/20010911203610/http://a.example.org>
  ; rel="prev memento"; datetime="Tue, 11 Sep 2001 20:30:51 GMT",
  <http://arxiv.example.net/web/20010911203610/http://a.example.org>
  ; rel="next memento"; datetime="Tue, 11 Sep 2001 20:47:33 GMT"
Content-Length: 0
Content-Type: text/plain; charset=UTF-8
Connection: close
```

## Memento HTTP Flow: Step 5



```
GET /web/20010911203610/http://a.example.org
HTTP/1.1
Host: arxiv.example.net
Accept-Datetime: Tue, 11 Sep 2001 20:35:00 GMT
Connection: close
```



Memento  
Herbert Van de Sompel



## Memento HTTP Flow: Step 6



```
HTTP/1.1 200 OK
Date: Thu, 21 Jan 2010 00:09:40 GMT
Server: Apache-Coyote/1.1
Memento-Datetime: Tue, 11 Sep 2001 20:36:10 GMT
Link: <http://a.example.org>; rel="original",
      <http://arxiv.example.net/web/20000915112826/http://a.example.org>
      ; rel="first memento"; datetime="Tue, 15 Sep 2000 11:28:26 GMT",
      <http://arxiv.example.net/web/20080708093433/http://a.example.org>
      ; rel="last memento"; datetime="Tue, 08 Jul 2008 09:34:33 GMT",
      <http://arxiv.example.net/timemap/http://a.example.org>
      ; rel="timemap"; type="application/link-format",
      <http://arxiv.example.net/timegate/http://a.example.org>
      ; rel="timegate",
      <http://arxiv.example.net/web/20010911203610/http://a.example.org>
      ; rel="memento"; datetime="Tue, 11 Sep 2001 20:36:10 GMT",
      <http://arxiv.example.net/web/20010911203610/http://a.example.org>
      ; rel="prev memento"; datetime="Tue, 11 Sep 2001 20:30:51 GMT",
      <http://arxiv.example.net/web/20010911203610/http://a.example.org>
      ; rel="next memento"; datetime="Tue, 11 Sep 2001 20:47:33 GMT"
Content-Length: 23364
Content-Type: text/html;charset=utf-8
Connection: close
```

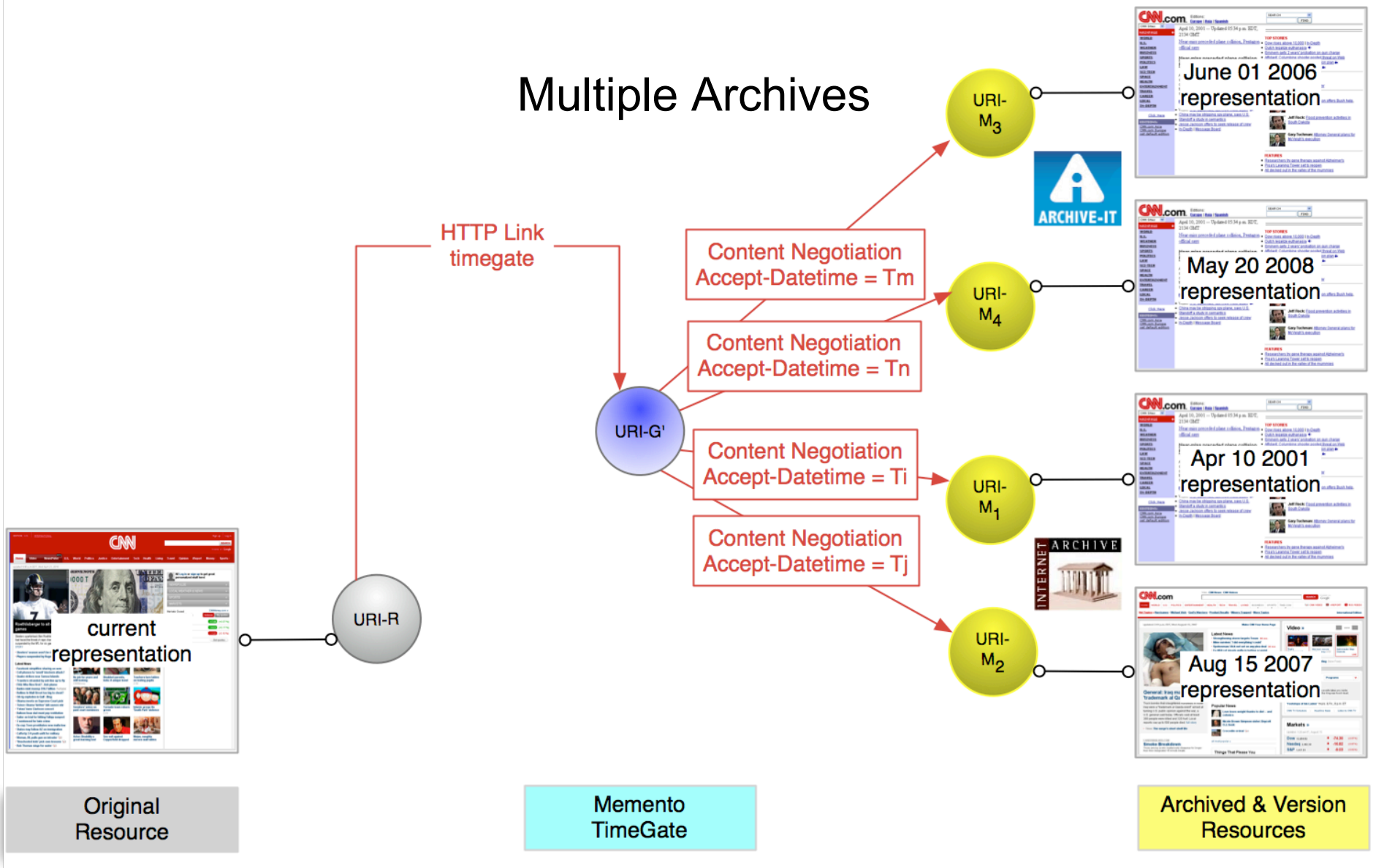
# The Memento Framework:

Protocol to Integrate Past and Current Web

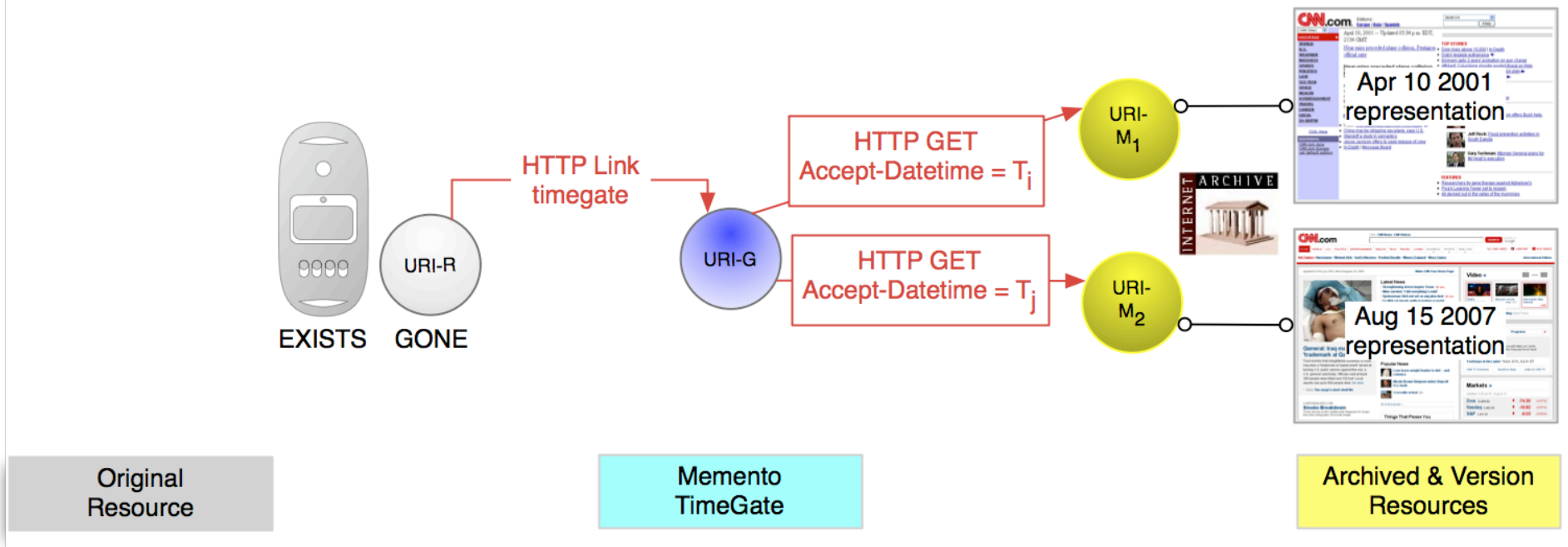
## Interesting Cases



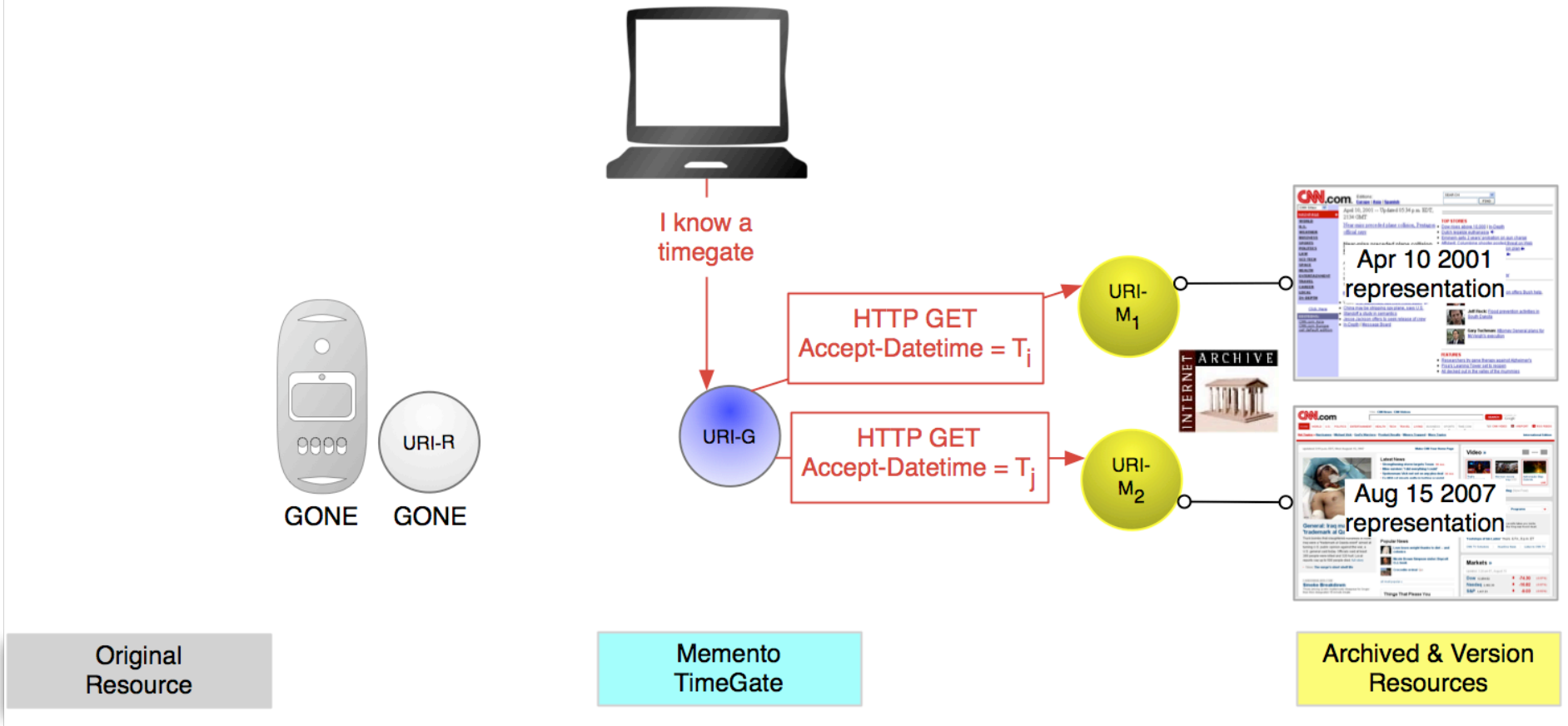
# Multiple Archives



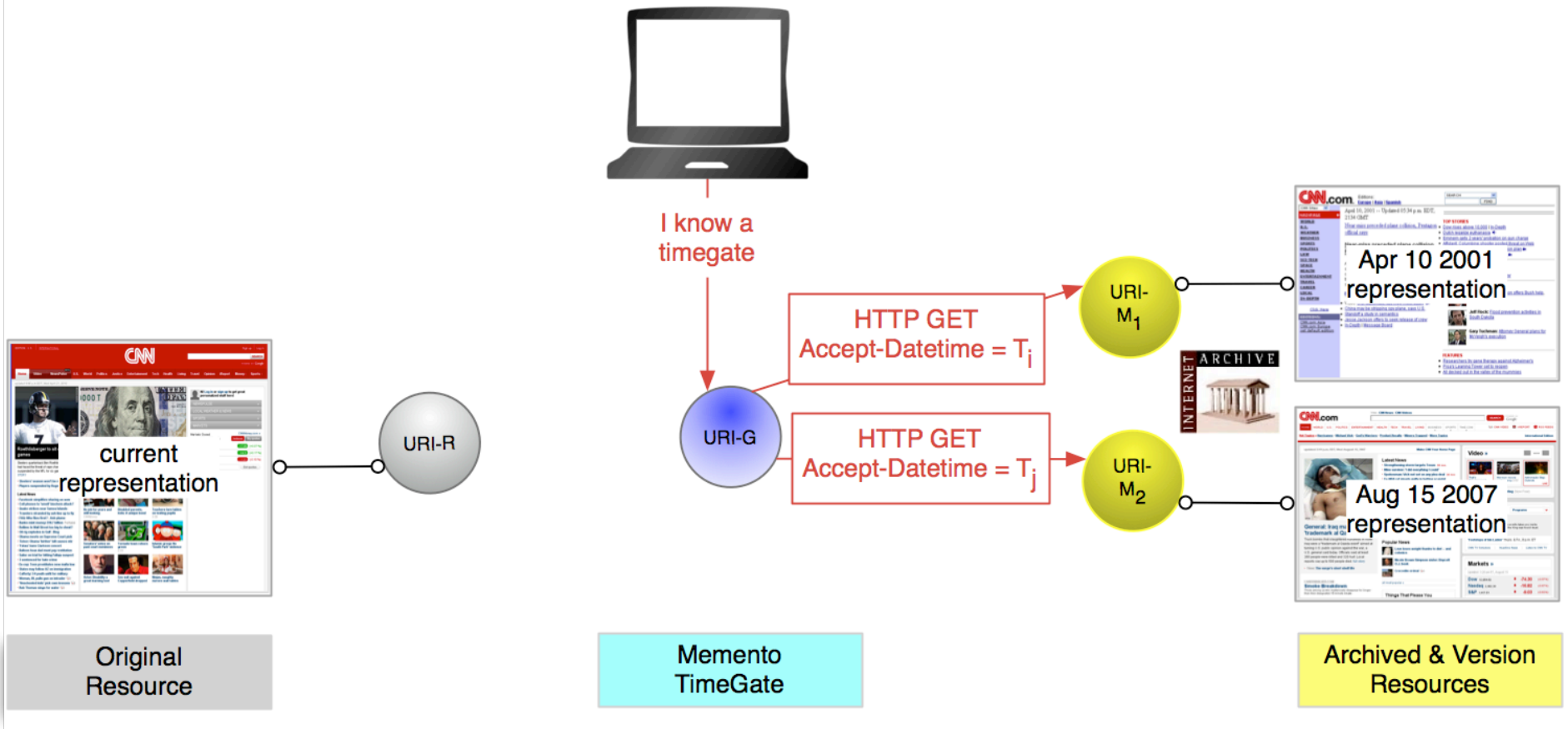
# Original Resource Gone



# Original Resource's Server Gone



# Original Resource Provides no Link

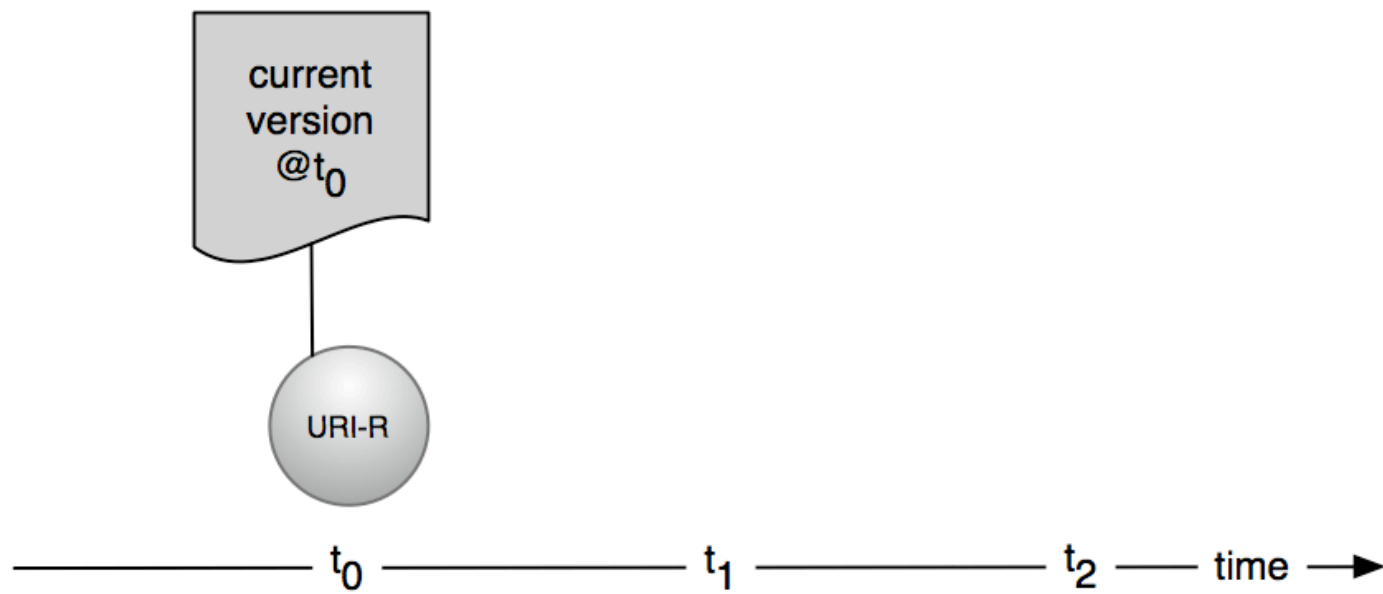


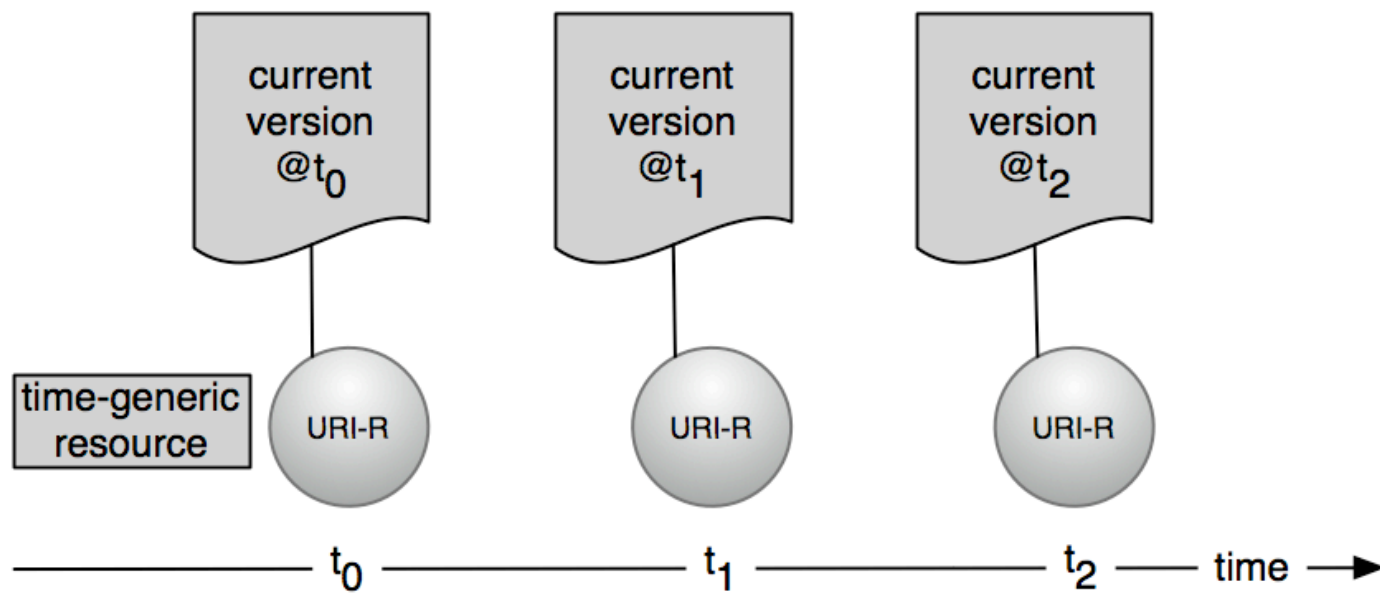


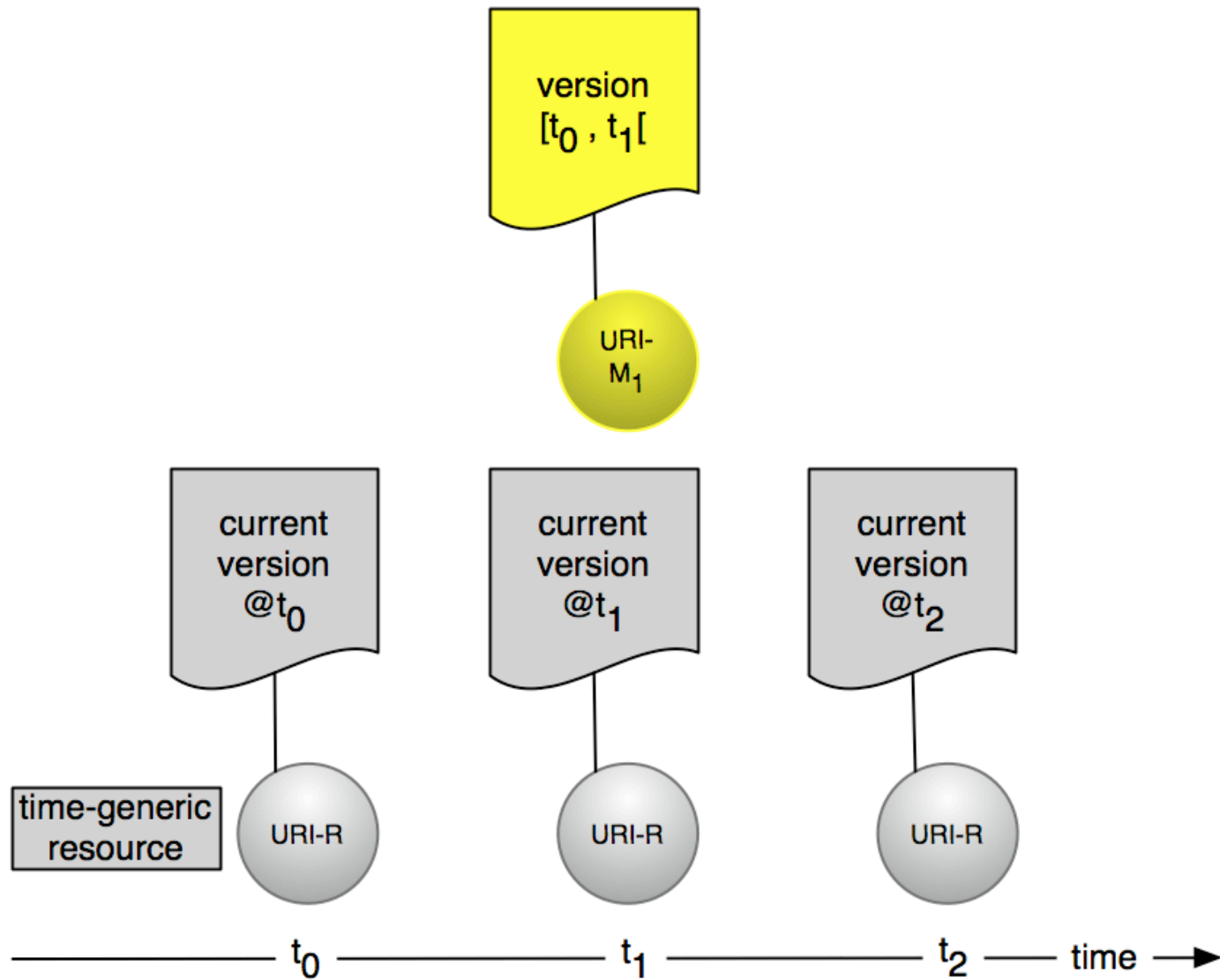
# The Memento Framework:

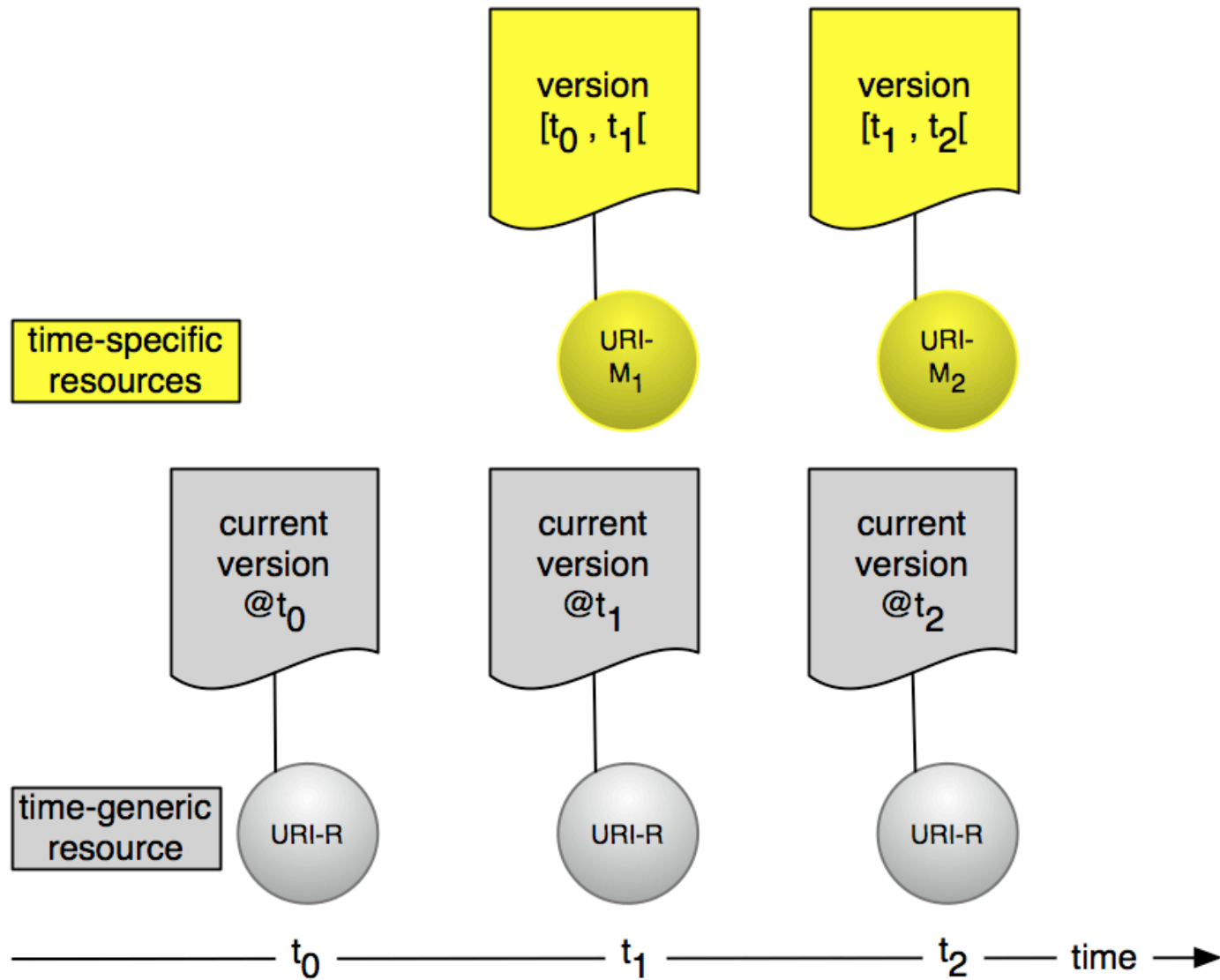
## Resource Versioning

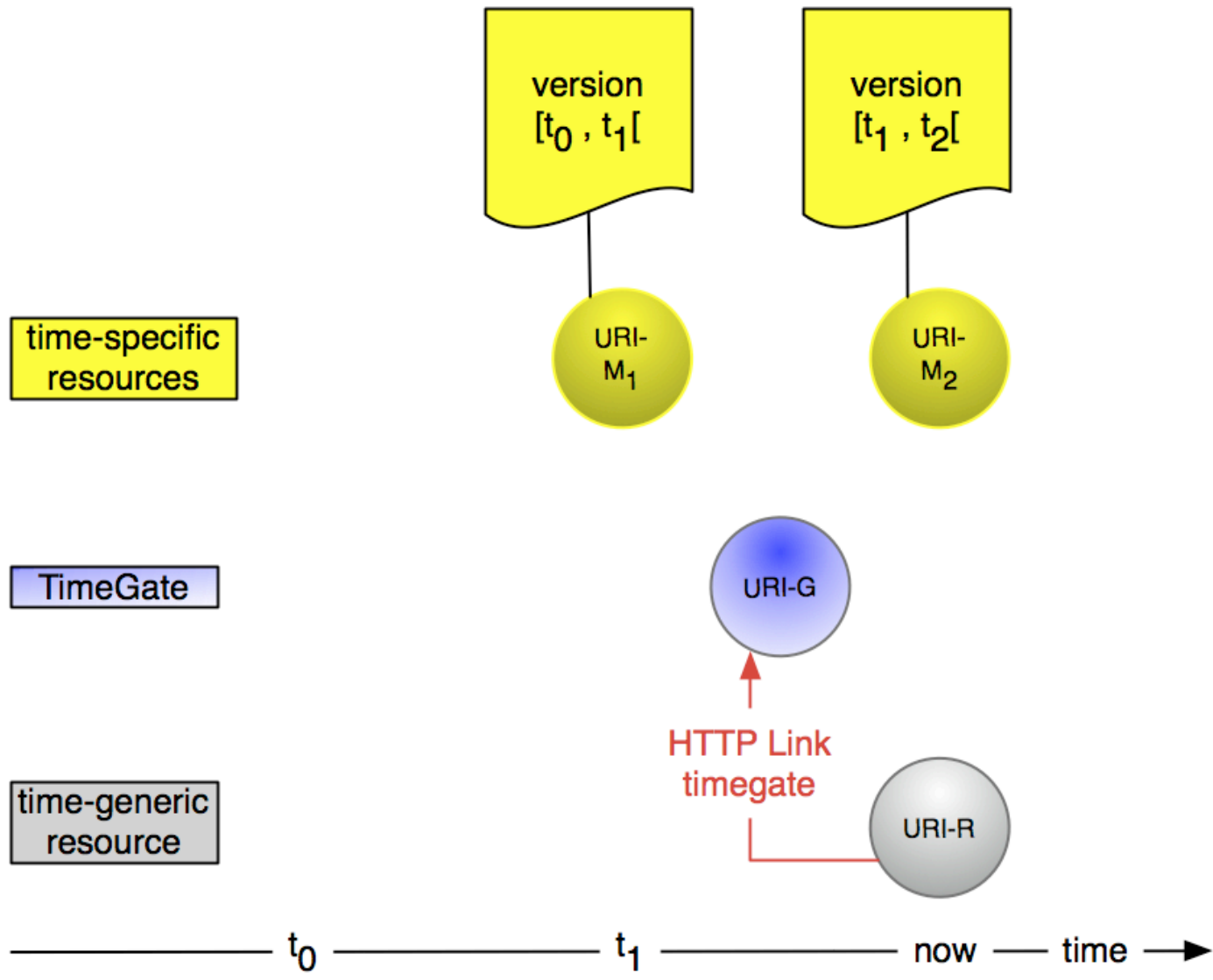


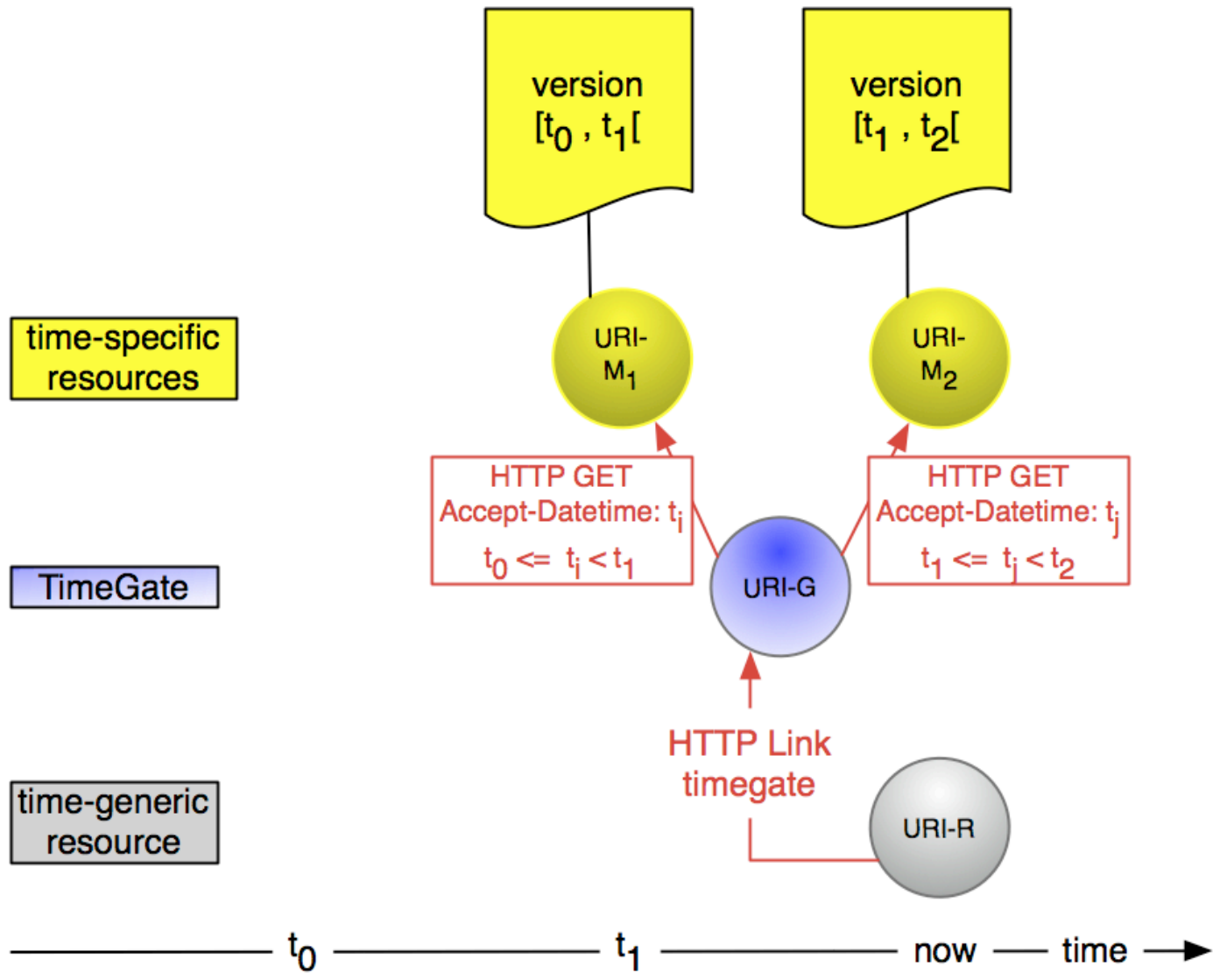


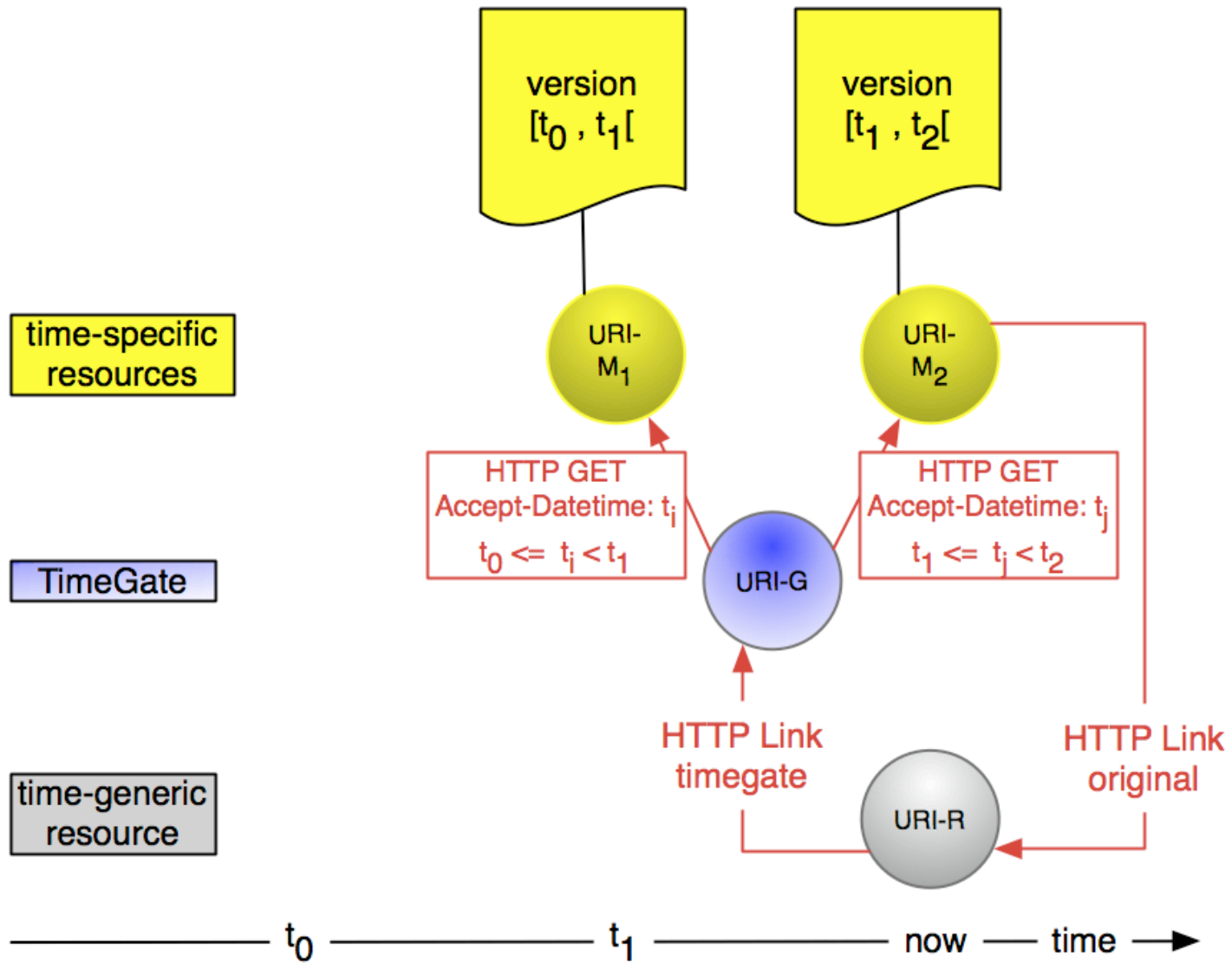








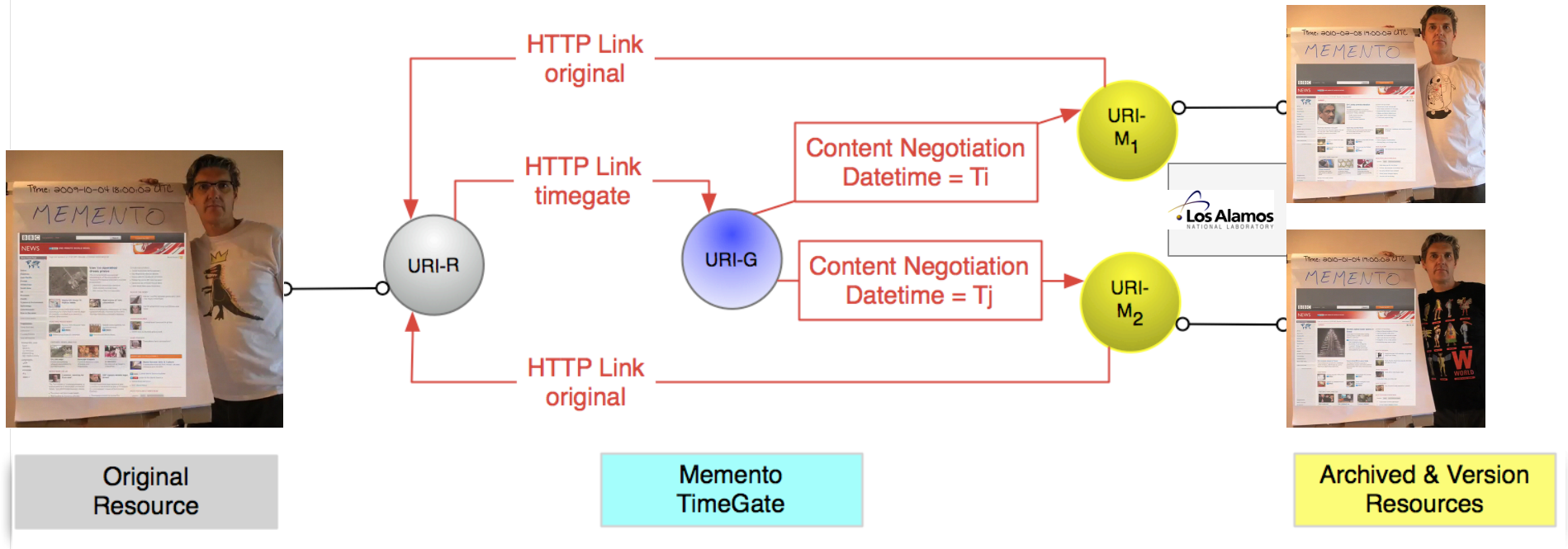






# Memento Framework

Original Resource: <http://lanlsource.lanl.gov/pics/picoftheday.png>



# Time Travel across Versions of a Picture of the Day



Movie at: <http://www.mementoweb.org/demo/picoftheday.mov>



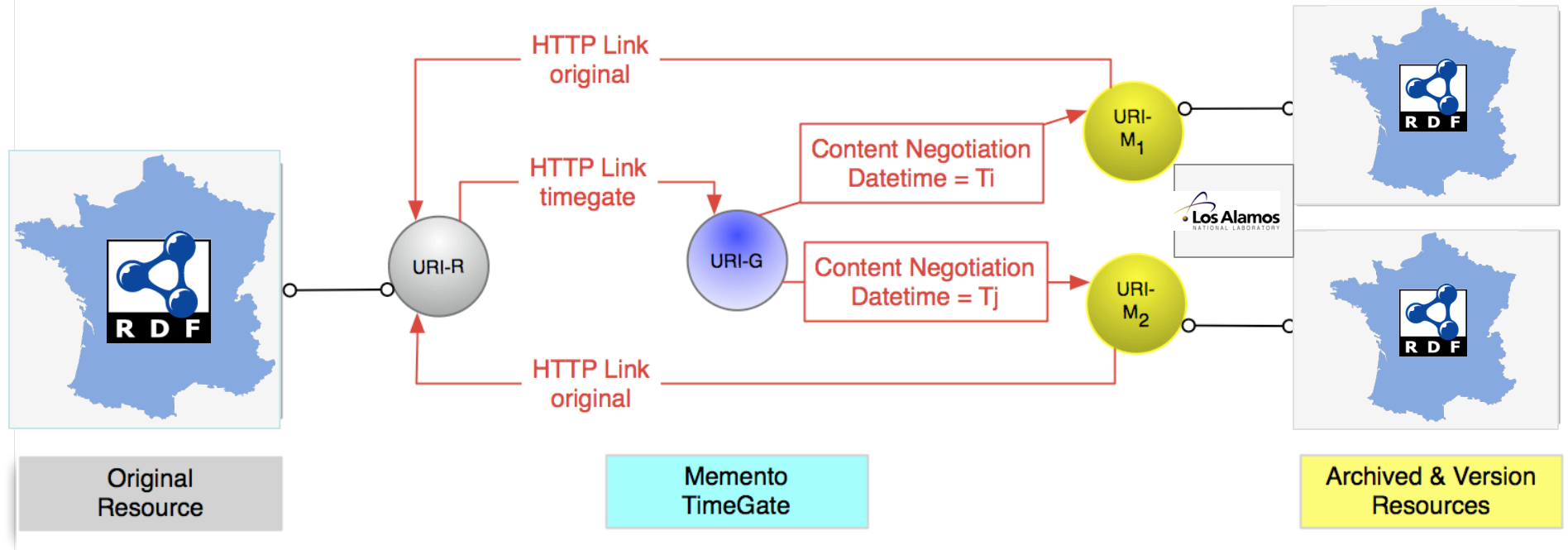
Memento  
Herbert Van de Sompel

58

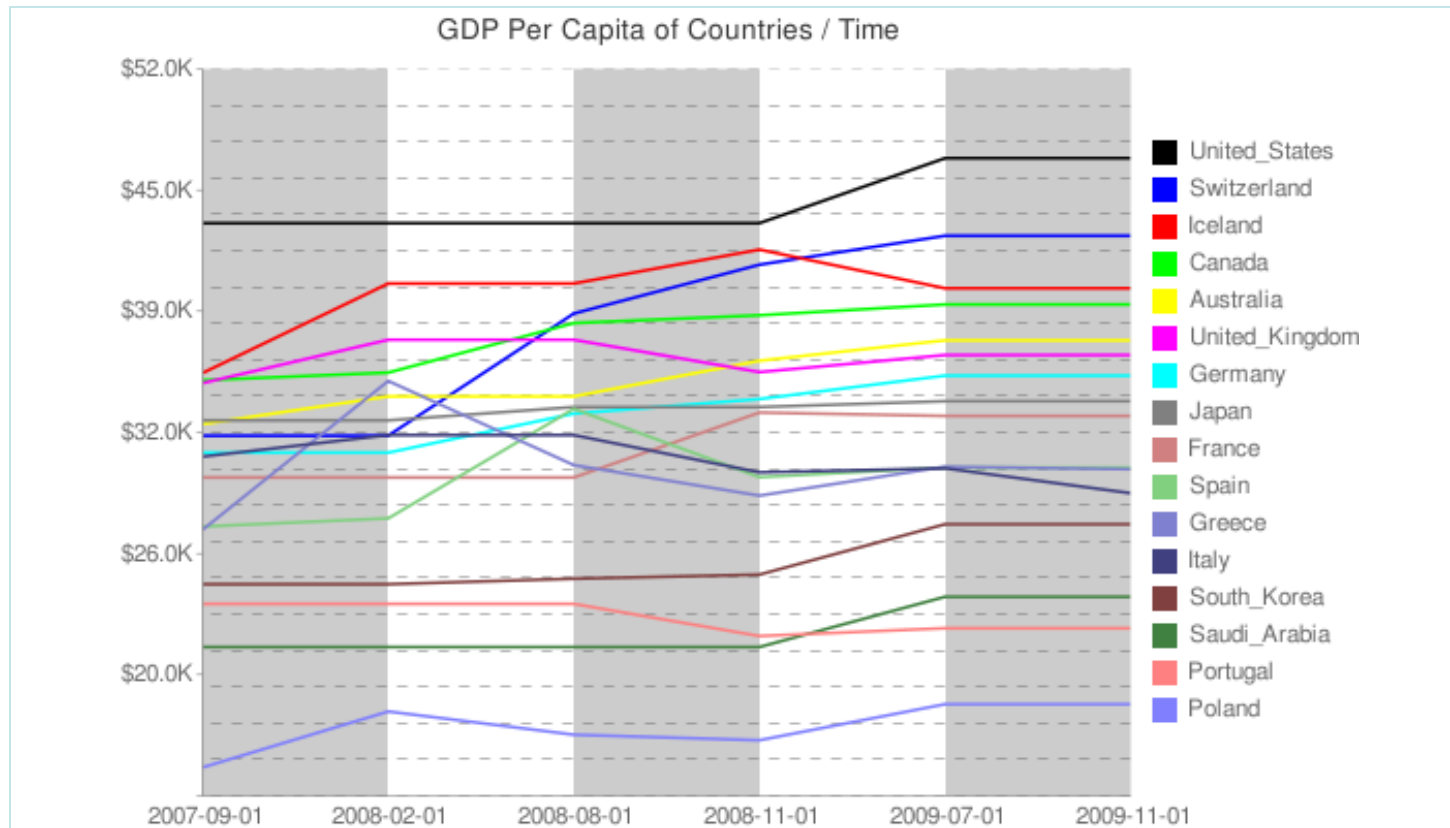


# Memento Framework

Original Resource: <http://dbpedia.org/resource/France>



# Time-Series Analysis across DBpedia Versions



Data collected through HTTP Navigation

paper at <http://arxiv.org/abs/1003.3661>



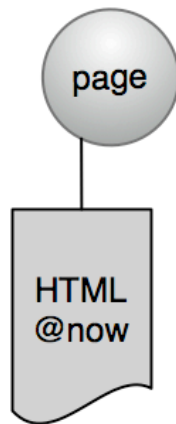
# The Memento Framework:

## The Increasing Value of a URI



# URI as Access Point to a Page

`http://weather.example.com/oxaca`

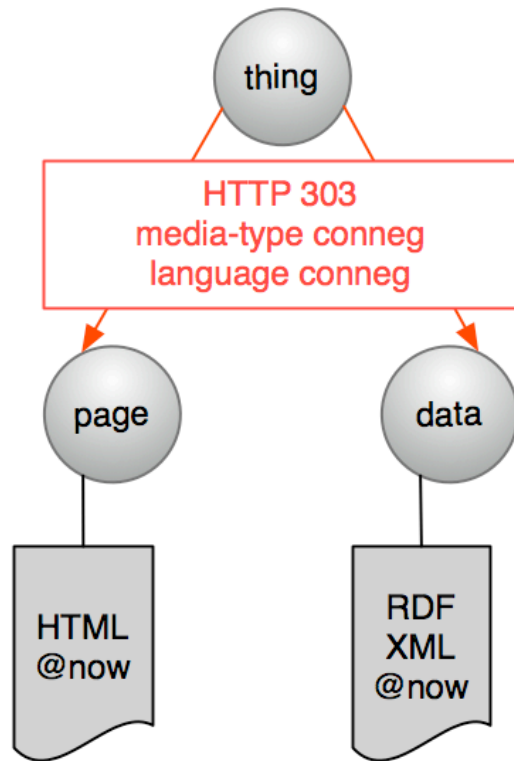


Memento  
Herbert Van de Sompel

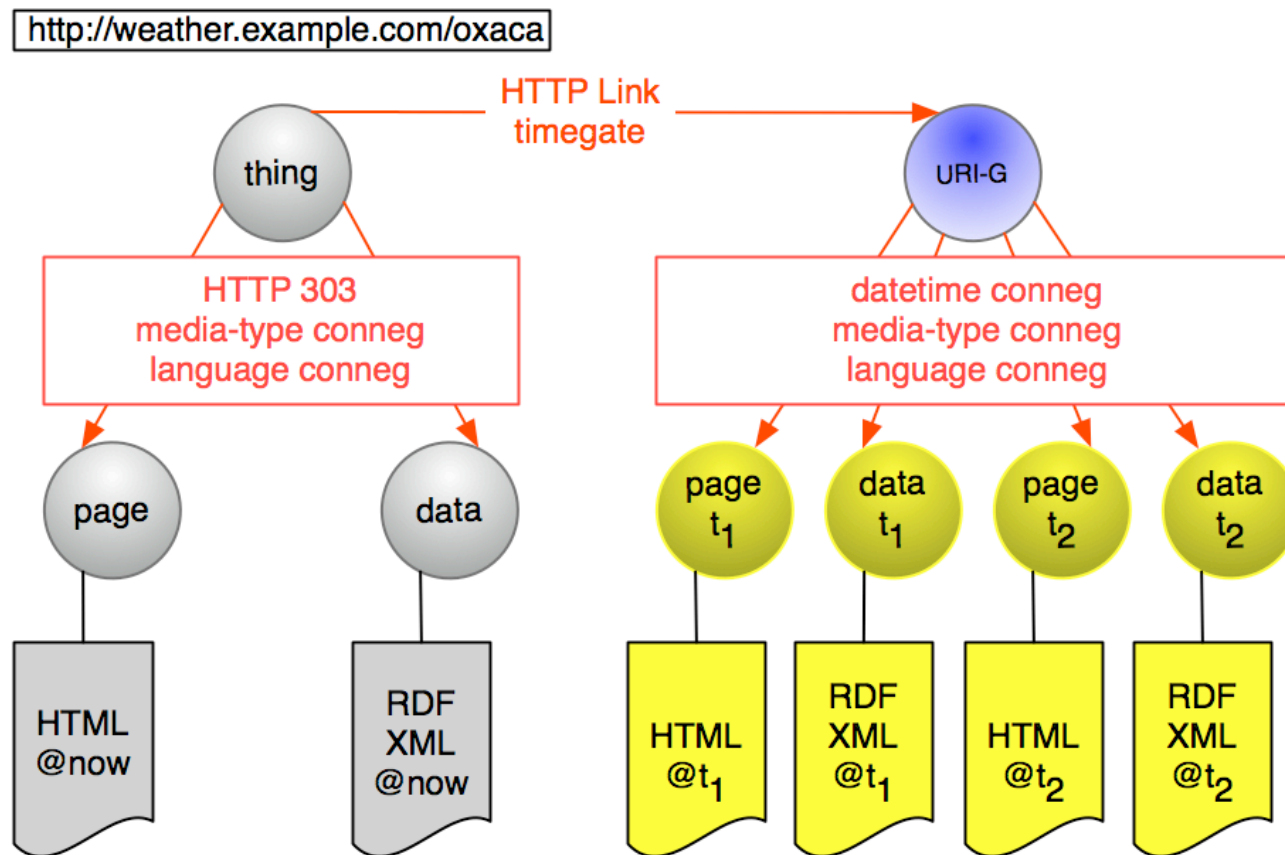


# URI as Access Point to Page and Data

`http://weather.example.com/oxaca`



# URI as Access Point to Current and Past Pages and Data





# The Memento Framework:

## Tools & Adoption



# Memento Client Support



- Several client tools developed by us and others
- Add-ons for FireFox
- Applications for Android
- Paper in Code4Lib Journal <http://journal.code4lib.org/articles/4979>



<https://addons.mozilla.org/en-US/firefox/addon/mementofox/>



# Memento Server Support

Requested Date: 1996 2010 06/03/2008 Displayed Date: 12/03/2007

ESW Wiki + Log in

page discussion view source history

## W3C<sup>®</sup> Main Page

Revision as of 12:53, 3 December 2007 by [DanielLewis](#) (Talk) (diff) ← Older revision | Latest revision (diff) | Newer revision → (diff)

The **W3C** has a formal track for making [standards specs](#). The specs answer a lot of questions, but not all of them. This wiki is for connecting the people who make the specs with the people who build on them.

Pages here have no formal status but may have [WikiConsensus](#). Questions & answers here may be misleading, or just plain wrong. Or, they may be useful.

Notable topics include:

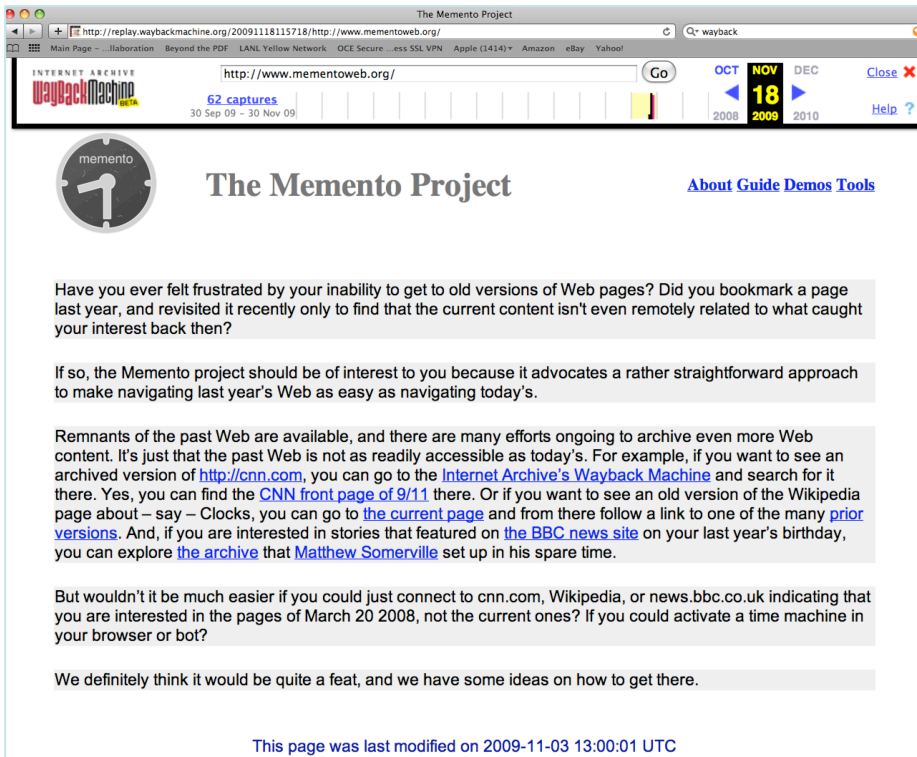
- [Quality Assurance](#), [MarkupValidator](#), [TrackingIssues](#), [UriTesting](#)
- [GoodURIs](#), [HCLSIG BioRDF Subgroup/Tasks/URI Best Practices](#)
- [Semantic Web FAQs](#), [SemanticWebTools](#), [SPARQL](#), [OWIAuthoringTools](#)
  - emerging topics: [LinkedData](#), [RdfAndSql](#), [EmbeddingRDFinHTML](#), [RDFa](#)
- [Security](#), starting around March 2006, for discussion on W3C Security-related specifications and enhancements to [XML-Dsig](#)
- [SchemaComposition](#) and topics related to [XML Schema](#), prompted by the [W3C Workshop on XML Schema 1.0 User Experiences](#)

- Plug-in for MediaWiki

*<http://www.mediawiki.org/wiki/Extension:Memento>*



# Memento Server Support



- Memento-compliant Wayback software:
  - In production at the Internet Archive & UK National Archives
  - Available to Web archives, worldwide



<http://mementoweb.org/tools/wayback/>

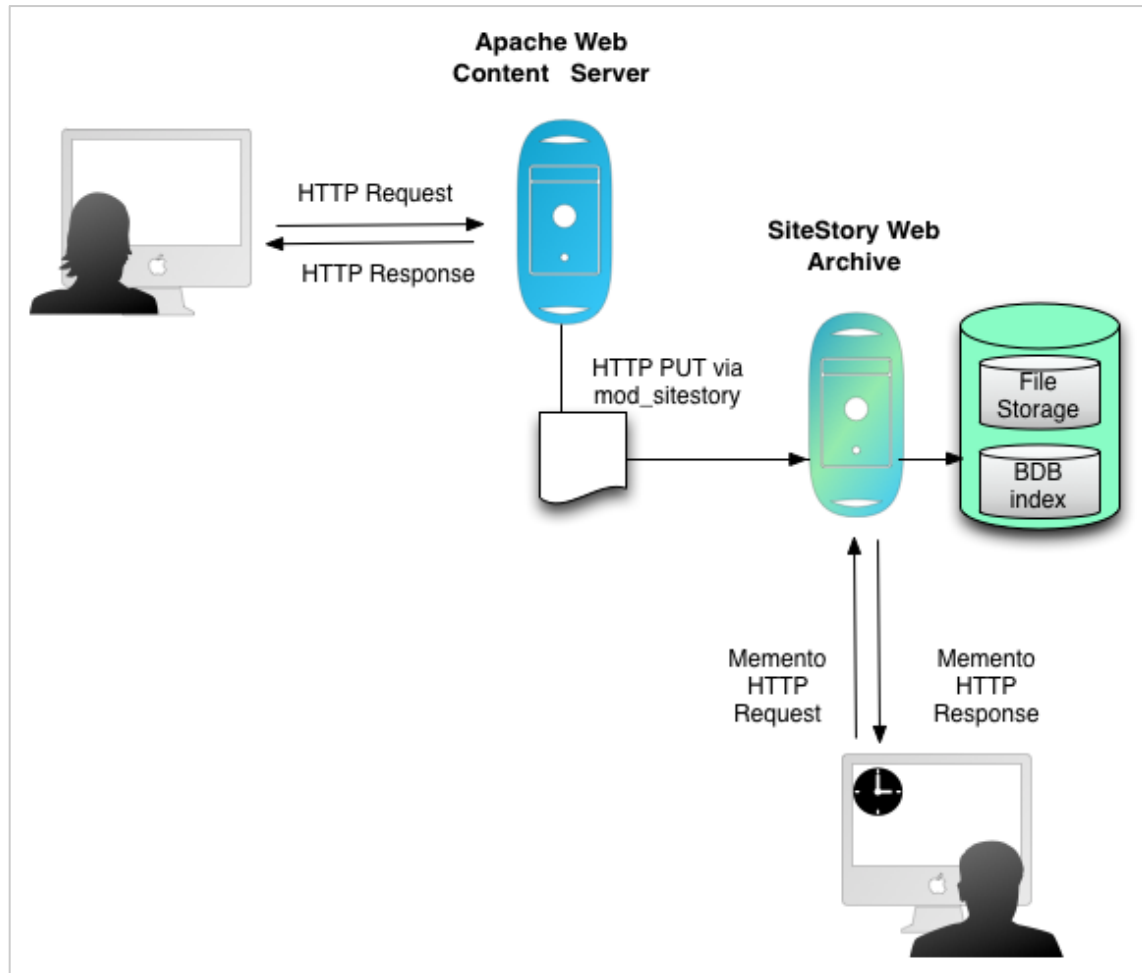


Memento  
Herbert Van de Sompel

68



# SiteStory Transactional Archiving



<http://mementoweb.github.com/SiteStory/>



Memento  
Herbert Van de Sompel



# Memento Server Validator

## Test Details

Tests for TimeGate: <a href="http://memento.waybackmachine.org/memento/timegate/http://www.ietf.org/">http://memento.waybackmachine.org/memento/timegate/http://www.ietf.org/</a>	
Test Result	URI/Date Returned
✓ TimeGate returns status 302	
✓ TimeGate has Location header for Memento	<a href="http://memento.waybackmachine.org/memento/20100402222407/http://www.ietf.org/">http://memento.waybackmachine.org/memento/20100402222407/http://www.ietf.org/</a>
✓ TimeGate acknowledges Accept-Datetime in Vary header	
✓ Link Header has link with rel=original	<a href="http://www.ietf.org/">http://www.ietf.org/</a>
✓ Link Header has link with rel=timebundle	<a href="http://memento.waybackmachine.org/list/timebundle/http://www.ietf.org/">http://memento.waybackmachine.org/list/timebundle/http://www.ietf.org/</a>
✓ Link Header has link with rel=timemap	<a href="http://memento.waybackmachine.org/list/timemap/link/http://www.ietf.org/">http://memento.waybackmachine.org/list/timemap/link/http://www.ietf.org/</a>
✓ TimeMap Link has application/link-format type	
✓ Link Header has one or more links with rel=memento	<a href="http://memento.waybackmachine.org/memento/20100219084310/http://www.ietf.org/">http://memento.waybackmachine.org/memento/20100219084310/http://www.ietf.org/</a>
✓ -- Selected Memento in Link Header with rel=memento	
✓ Memento has parsable datetime attribute	Fri, 19 Feb 2010 08:43:10 GMT
✓ Memento has parsable datetime attribute	Wed, 06 Nov 1996 11:49:54 GMT

- Server side client:
  - Attempts to perform all Memento actions against a given URI
  - Reports success/failure of the interactions and warnings for optional aspects
  - Kept up to date with IETF Internet Draft

<http://mementoweb.org/tools/validator/>



# Memento & IIPC



INTERNATIONAL  
INTERNET  
PRESERVATION  
CONSORTIUM

---

HOME ABOUT IIPC WEB ARCHIVING PROJECTS MEMBER ARCHIVES CALENDAR

---

## MEMENTO

**Status:** Active Project



**IIPC Memento Aggregator**

Goal: To aggregate the metadata of the distributed archives of the IIPC, and

- To provide Memento based access to the holdings of open archives
- To provide knowledge of the holdings of restricted archives
- To provide knowledge to IIPC members of the holdings of totally closed archives

> Initial demo for participants, then IIPC

<http://netpreserve.org/projects/memento>



Memento  
Herbert Van de Sompel



# Memento & Wikipedia, Mediawiki



The screenshot shows the Wikipedia interface for the page "Wikipedia:Requests for comment/Memento". At the top right, there are links for "Create account" and "Log in". Below that, navigation tabs include "Project page", "Talk", "Read", "Edit", and "View history". A search box is located to the right of these tabs. A banner at the top right encourages participation in a photo competition to improve Wikipedia. The main heading is "Wikipedia:Requests for comment/Memento". Below the heading, it states "From Wikipedia, the free encyclopedia" and provides a link to "< Wikipedia:Requests for comment". The main text begins with "This is a **request for comment** (RFC) regarding turning on the **Memento** extension." It continues: "This is a preliminary RFC to assess community interest among English Wikipedia users for this functionality. No significant commitment of Wikimedia Foundation engineering resources has been made yet. An early pilot would likely run on the English Wikipedia; hence the initial poll is taking place here." A "Contents" section is visible, listing four items: "1 What is Memento?", "2 How does it work?", "3 How does it impact the editing community?", and "4 Why support it?". The left sidebar contains various navigation links such as "Main page", "Contents", "Featured content", "Current events", "Random article", "Donate to Wikipedia", "Interaction", "Help", "About Wikipedia", "Community portal", "Recent changes", "Contact Wikipedia", "Toolbox", and "Print/export".

[http://en.wikipedia.org/wiki/Wikipedia:Requests\\_for\\_comment/Memento](http://en.wikipedia.org/wiki/Wikipedia:Requests_for_comment/Memento)



Memento  
Herbert Van de Sompel





# Memento & DBpedia



## DBpedia Archive

[About](#) [Demos](#) [Guide](#) [Tools](#) [Depot](#)

This document provides information about the Memento compliant DBpedia Archive provided by the Memento Depot.

Questions and comments are welcome on [memento-dev@googlegroups.com](mailto:memento-dev@googlegroups.com).

## DBpedia Archive @ mementodepot.org

### Archive Content

This archive contains prior versions of [DBpedia](#). It is based on data available from the [DBpedia Downloads page](#), and covers the following DBpedia dumps:

- DBpedia 3.7, DBpedia 3.6, DBpedia 3.5.1, DBpedia 3.5, DBpedia 3.4, DBpedia 3.3, DBpedia 3.2, DBpedia 3.1, DBpedia 3.0, DBpedia 3.0RC, DBpedia 2.0
- The English (en) version of all Core Datasets. No other languages are available.

<http://mementoweb.org/depot/native/dbpedia/>



Memento  
Herbert Van de Sompel



# The Memento Framework:

## Persistent Web Annotations



# Web-Centric Annotation: No Persistence

The image shows a screenshot of a BBC News website with a user annotation overlaid on the left side. The annotation, by Paul Murray on March 8, 2010, criticizes the BBC's lead story about an Oscar triumph for 'The Hurt Locker' in light of a magnitude six earthquake and a domestic civil servants strike. The user suggests that entertainment awards should not trump real news stories and notes that BBC Radio 2 ran the story at the end of its bulletin. The annotation includes a 'Useful?' section with 'Yes (0)' and 'No (0)' counts, and options to 'Report abuse' and 'Share'.

**Paul Murray** - Mar 8, 2010

**Lead story this morning.**

With a magnitude six earthquake, attacks around the world killing hundreds and a domestic civil servants strike, I find it very odd that the BBC website should have chosen to lead with "Oscar triumph for The Hurt Locker".

In my opinion entertainment awards barely scrape the definition of news let alone trump the many real stories available today. BBC Radio 2 ran the story at the very end of its bulletin.

Useful? [Yes](#) (0) [No](#) (0)  
[Report abuse](#) [Share](#) ▼

The background shows the BBC News homepage as of June 14, 2010. The navigation bar includes 'me', 'News', 'Sport', 'Weather', 'TV', and 'Radio'. A red banner promotes 'ONE-MINUTE WORLD NEWS'. The page was last updated at 21:59 GMT, Monday, 14 June 2010 22:59 UK. The latest news item is 'Fourth Colombian hostage found alive after three fellow captives'. Below this is a section titled 'DEEPWATER HORIZON RESPONSE' with a photo of Barack Obama and the headline 'Obama calls for clean energy push'. The text states that US President Obama asks supporters to back a campaign for clean energy as he visits areas affected by the BP oil spill. A list of links includes 'Mapping the oil slick', 'In graphics: Tackling the leak', and 'Q&A: Scale of BP oil spill'.

Google Sidewiki Annotation on <http://news.bbc.co.uk/> as of 2010-06-14



# Web-Centric Annotation: No Persistence


The screenshot shows a web page with a sidebar on the left and a main content area on the right. The sidebar is titled "Google Sidewiki" and features a user profile for "Robert Sanderson" with a small profile picture and a link to "Edit your name". Below the profile, there is a text area for writing an entry about the page, with instructions to highlight specific parts. A box below the text area states "This page has no Annotations :(". At the bottom of the sidebar, there is a link for "Helpful information about 'BBC NEWS | News Front Page'". The main content area has a red header bar with the text "Page last updated at 07:04 GMT, Monday, 8 March 2010". Below this is a "LATEST" section with a link to "FBI releases some of its secret files on the late Senator Edwa". The main headline is "Oscar joy for The Hurt Locker", and below it is a photograph of three people (two men and one woman) holding multiple Oscar statuettes on a red carpet. The background of the photo includes "abc" logos.

Archived page from:

<http://www.dracos.co.uk/work/bbc-news-archive/2010/03/08/07.05.html>



# Web-Centric Annotation: Desired Persistence

[Paul Murray](#) - Mar 8, 2010 

**Lead story this morning.**

With a magnitude six earthquake, attacks around the world killing hundreds and a domestic civil servants strike, I find it very odd that the BBC website should have chosen to lead with "Oscar triumph for The Hurt Locker".


In my opinion entertainment awards barely scrape the definition of news let alone trump the many real stories available today. BBC Radio 2 ran the story at the very end of its bulletin.

Useful? [Yes](#) (0) [No](#) (0)  
[Report abuse](#) [Share](#) ▼

Page last updated at 07:04 GMT, Monday, 8 March 2010

**LATEST: FBI releases some of its secret files on the late Senator**

## Oscar joy for The Hurt Locker



# Open Annotation: Dealing with Web Time

- As regular Web resources, Body and Target of an Annotation have representations that can change over time.
- Body and Target can change independently of each other.
- If an Annotation involves resources as they existed at a particular point in time, this needs to be recorded.
- The Open Annotation data model provides hooks for doing so via the notion of State.



# Create an Annotation

The screenshot shows a Wikipedia portal page titled "Portal:Current events/2010 January 23". The page includes navigation tabs for "portal", "discussion", "edit this page", and "history". The main content area displays "23 January 2010 (Saturday)" and "Categories: 2010 by day". A purple rectangular box is overlaid on the page, highlighting the date and category information. In the bottom right corner, an annotation form is visible, containing fields for "Title:", "Comment:", and "Type:" (set to "General"). Below these fields are checkboxes for "Comment is URL:" and "About Whole Image:", and a "Post To:" section with radio buttons for "Rob's Annotations" (selected) and "Contextualification". At the bottom of the form are links for "+Segment", "-Segment", "Create", "Logout", and "Close", along with two blue icons.



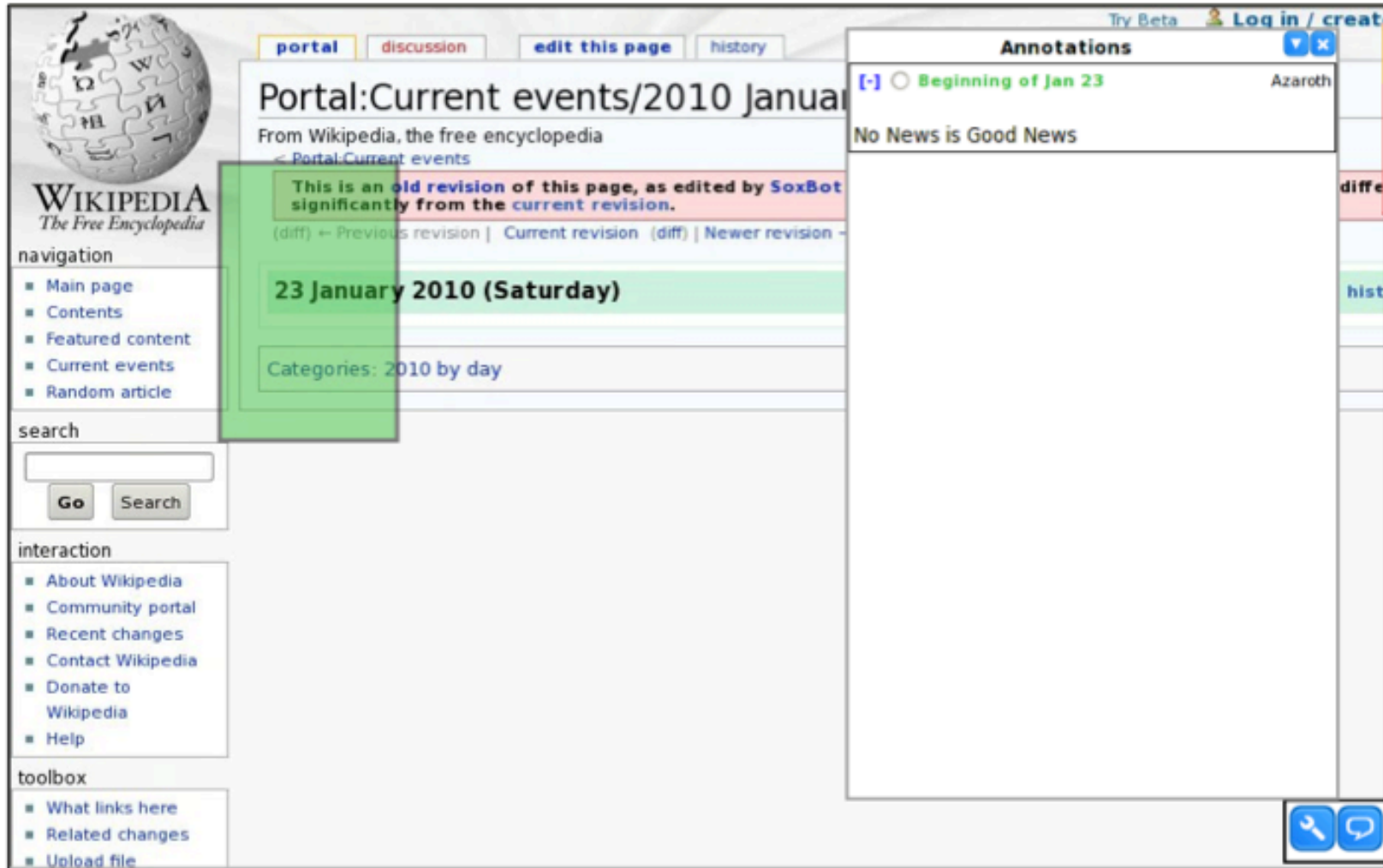
# Reconstruct the Annotation without Memento

The screenshot shows a Wikipedia portal page titled "Portal:Current events/2010 January". The page content includes a navigation sidebar on the left with sections for "navigation", "search", "interaction", and "toolbox". The main content area lists current events for "23 January 2010 (Saturday)", such as the aftermath of the 2010 Haiti earthquake and Joe Biden's visit to Iraq. An "Annotations" window is overlaid on the right side of the page, showing a green annotation titled "Beginning of Jan 23" by user "Azaroth" with the text "No News is Good News". The annotation window has a close button (X) and a refresh button ([-]).





# Reconstruct the Annotation with Memento



paper at <http://arxiv.org/abs/1003.2643>



# The Memento Framework:

## Citation Rot



# Citation Rot Studies

- Link rot studies for scholarly literature focus on HTTP URIs used in the reference sections of papers:
  1. Does the referenced resource still exist?
  2. Do archived versions of the referenced resource exist?
  3. Does an archived version of the referenced resource exist with an archival datetime close to the date of publication of the citing paper?
- Can one really reconstruct the scholarly record as it was at the time of publication of the citing paper?
- Studies so far have been really small scale because answering (2) and (3) involved tedious manual work.
- But with Memento these can be automated.

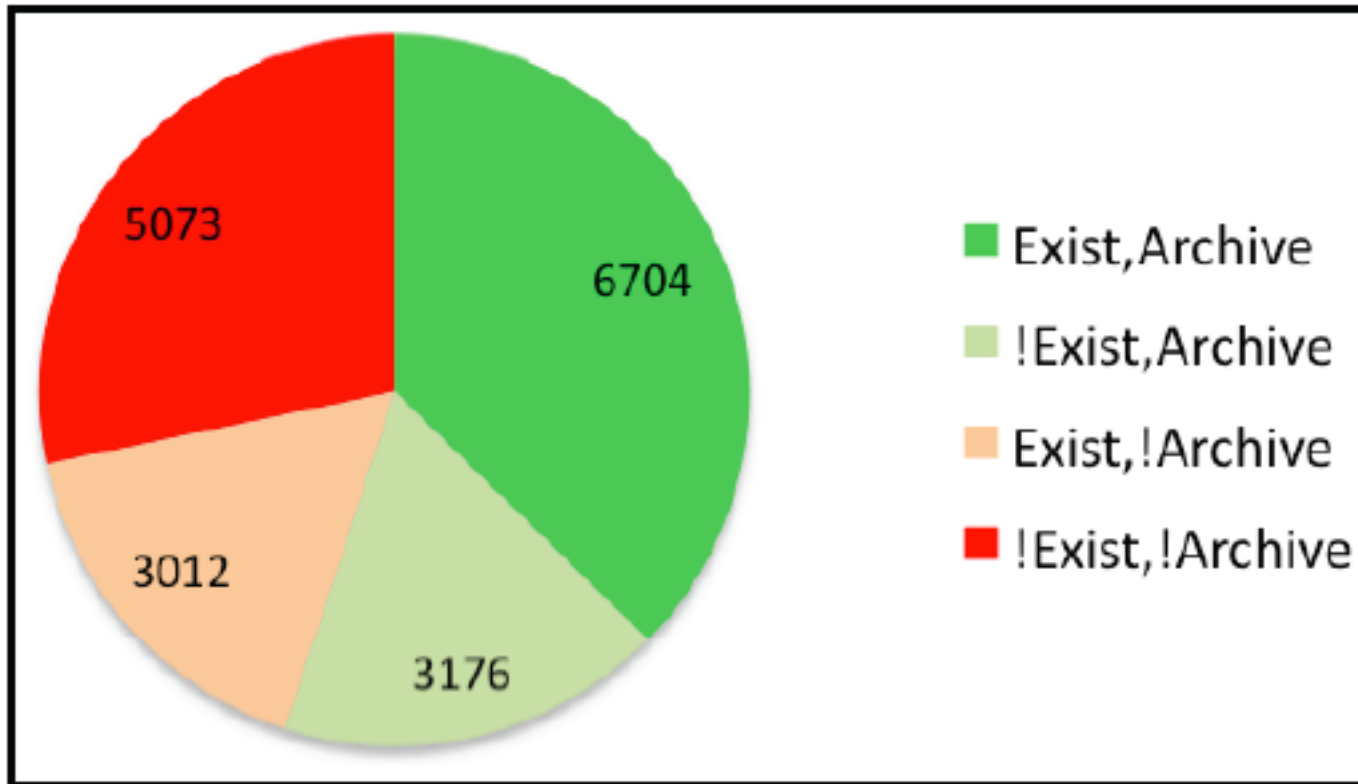


# Citation Rot Studies at Scale with Memento

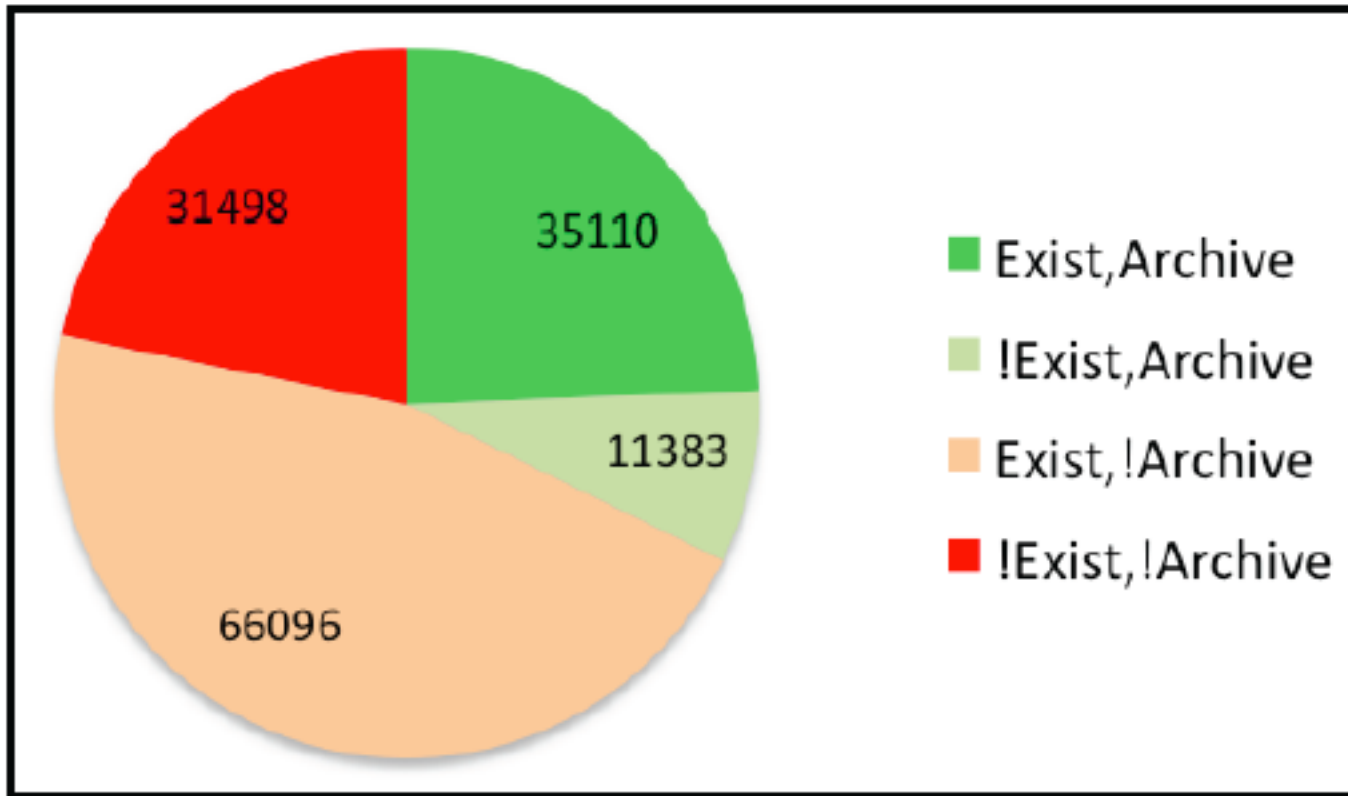
- Pilot study:
  - Papers from arXiv: 400,000 paper => 144,000 unique URIs
  - Thesis from UNT ETD repository: 3,600 papers => 18,000 URIs
  - URIs of known scholarly repositories removed (e.g. <http://dx.doi.org>), i.e. focusing in on the *periphery* of the scholarly record.



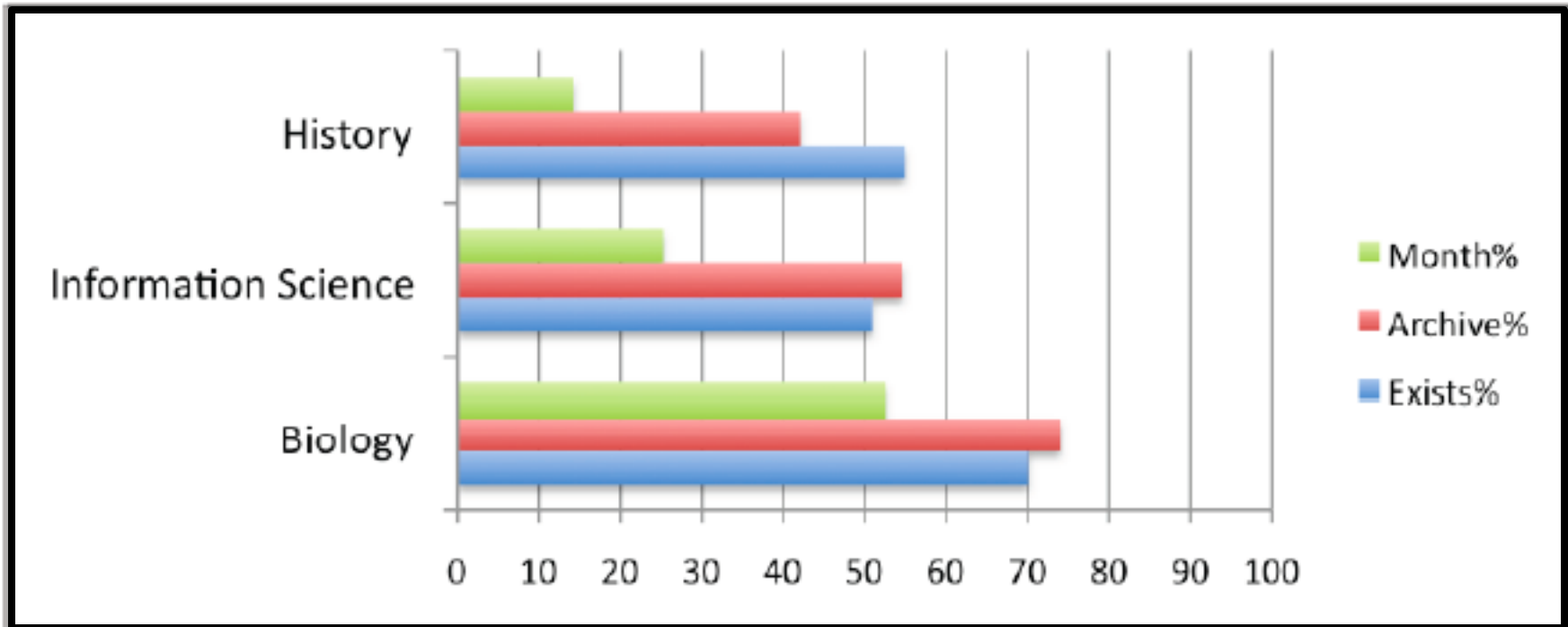
# UNT



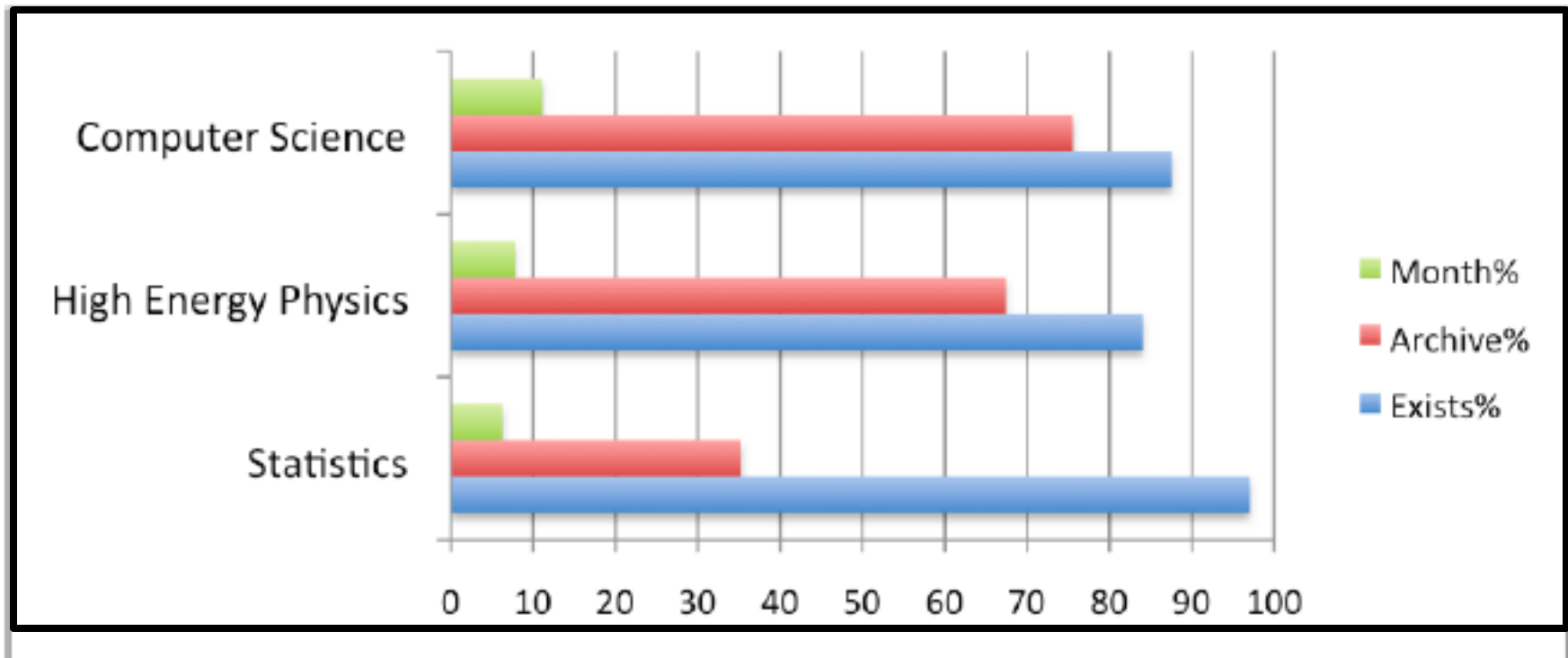
# arXiv



# UNT



# arXiv





## References

- Van de Sompel, H., Nelson, M.L., Sanderson, R., Balakireva, L., Ainsworth, S., Shankar, H. (2009) Memento. <http://arxiv.org/abs/0911.1112>
- Van de Sompel, H., Sanderson, R., Nelson, M.L., Balakireva, L., Ainsworth, S., Shankar, H. (2010) An HTTP-Based Versioning Mechanism for Linked Data. Proceedings of the 3rd Workshop on Linked Data on the Web. <http://arxiv.org/abs/1003.3661>
- Sanderson, R., and Van de Sompel, H. (2010) Making Web Annotations Persistent over Time. Proceedings of the 10th ACM/IEEE-CS Joint Conference on Digital libraries. <http://arxiv.org/abs/1003.2643>
- Sanderson, R., Phillips, M. and Van de Sompel, H. (2011) Analyzing the persistence of referenced Web resources with Memento. OR 2011. <http://arxiv.org/abs/1105.3459>
- Van de Sompel, H., Nelson, M.L., Sanderson, R. (2011) HTTP based framework for time-based access to resource states - Memento <http://www.mementoweb.org/guide/rfc/ID/>

