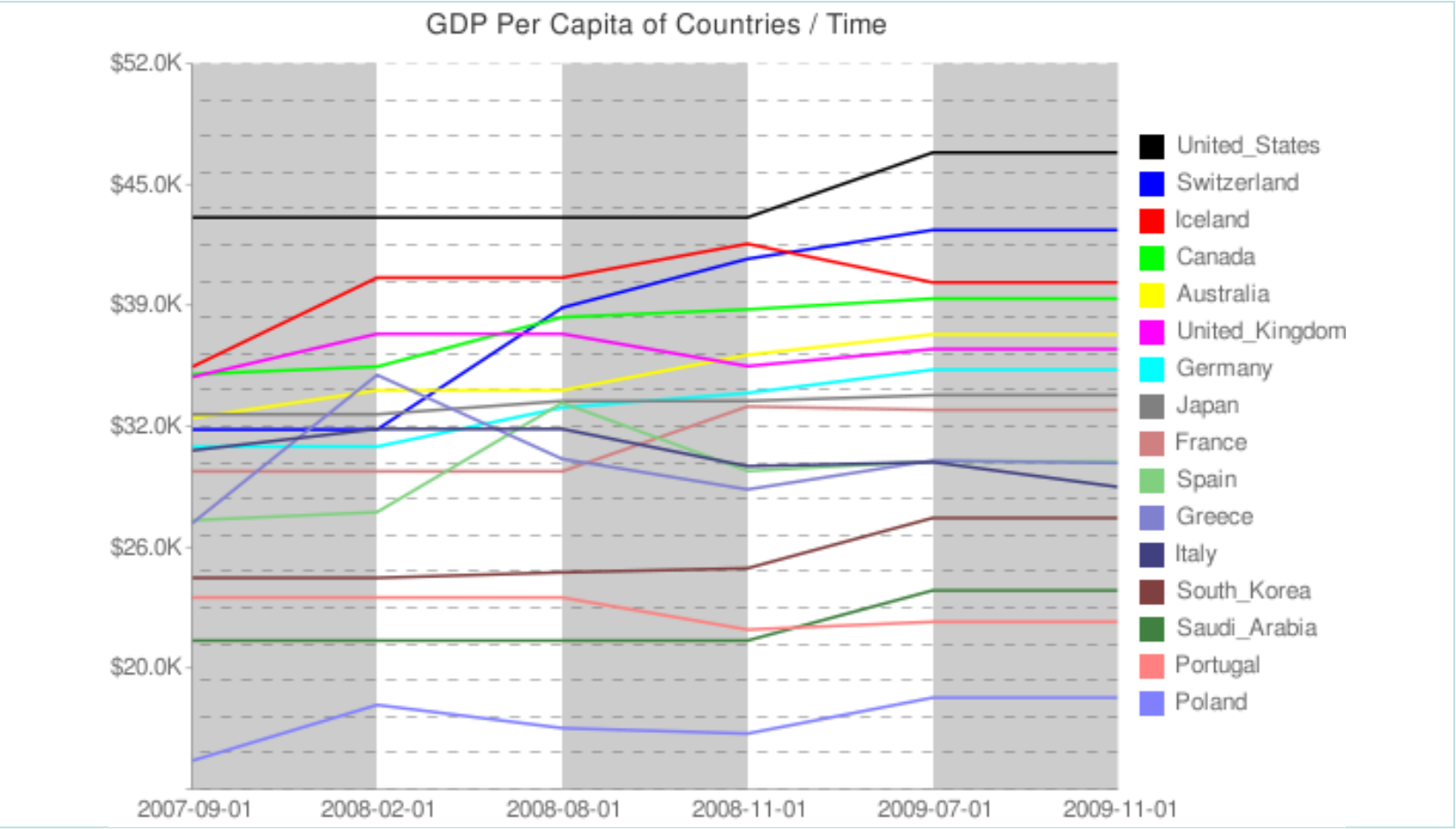
An HTTP-Based Versioning Mechanism for Linked Data

Herbert Van de Sompel
Robert Sanderson
Michael L. Nelson
Lyudmila Balakireva
Harihar Shankar
Scott Ainsworth

Memento is partially funded by the Library of Congress

Presentation at http://bit.ly/ac9GhH

Time-Series Analysis across DBpedia Versions



Data collected through HTTP Navigation

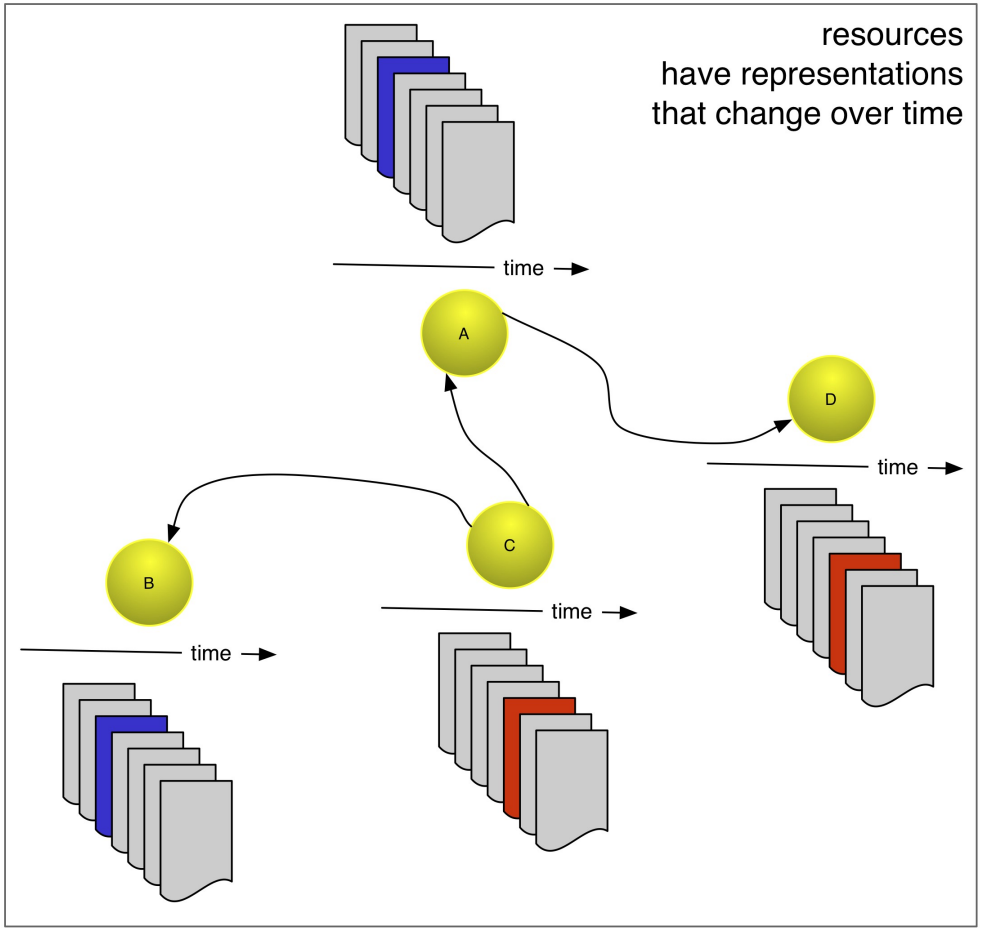
Outline

- Memento - Time Travel for the Web
- Resource Versioning suggested by Memento
- Resource Versioning for Linked Data
- DBpedia Demonstrator

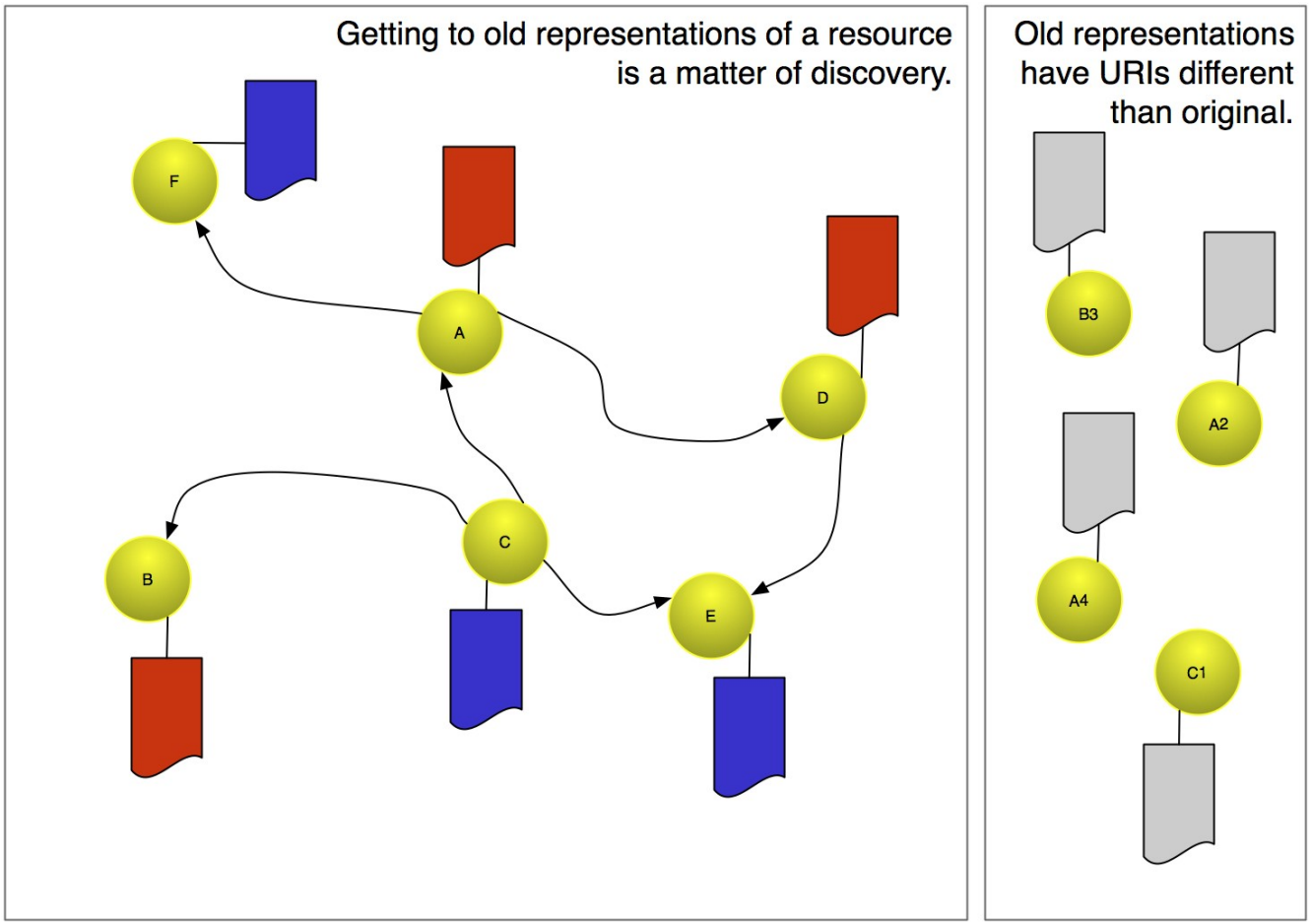
Outline

- Memento - Time Travel for the Web
- Resource Versioning suggested by Memento
- Resource Versioning for Linked Data
- DBpedia Demonstrator

Resources have Representations that Change over Time



Archived Resources serve prior Representations



Sep 11 2001, 20:36:10 UTC

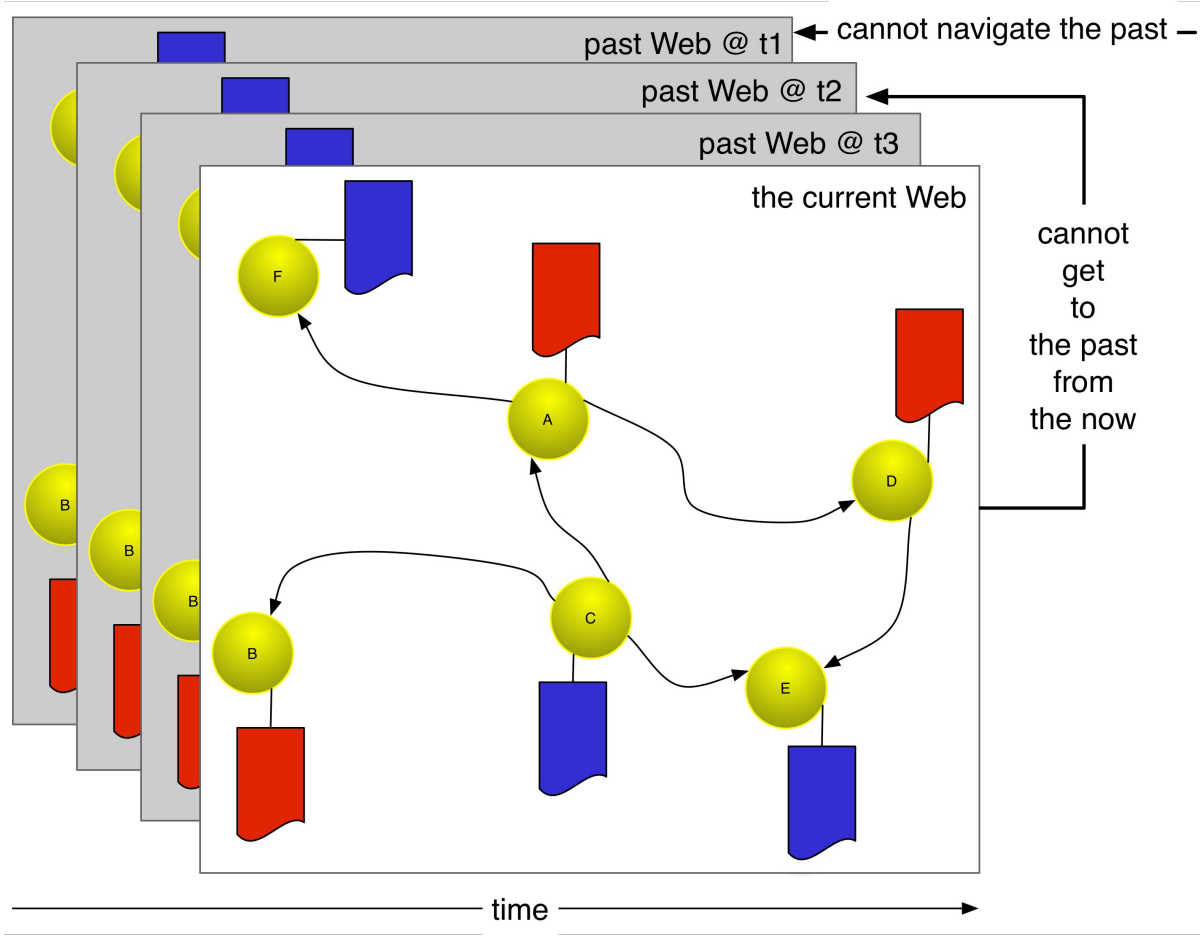
Dec 20 2001, 4:51:00 UTC

Archived Resources

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
archived resource for
http://en.wikipedia.org/wiki/September_11_attacks

Current and Past Web are Not Integrated

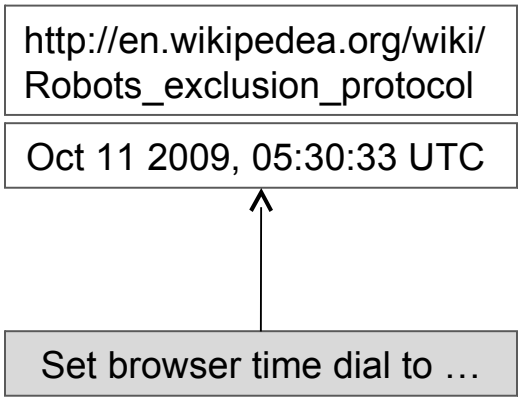


- Current and Past Web based on same technology.
- But, going from Current to Past Web is a matter of (manual) discovery.
- Memento wants to make going from Current to Past Web a (HTTP) protocol matter.
- Memento wants to integrate the Current And Past Web.

Vision: Navigate the Web of the Past

http://en.wikipedea.org/wiki/
Robots_exclusion_protocol

Vision: Navigate the Web of the Past



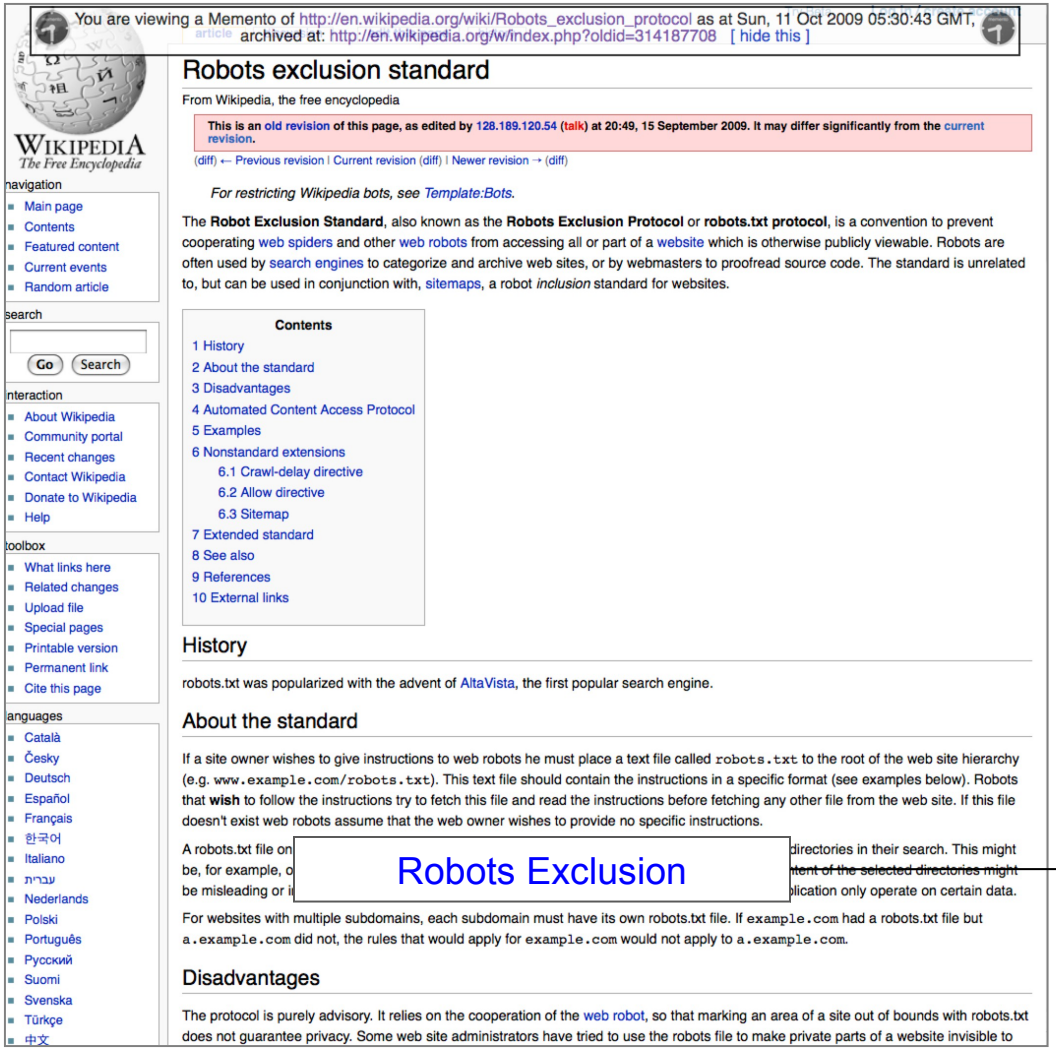
Vision: Navigate the Web of the Past

http://en.wikipedea.org/wiki/Robots_exclusion_protocol

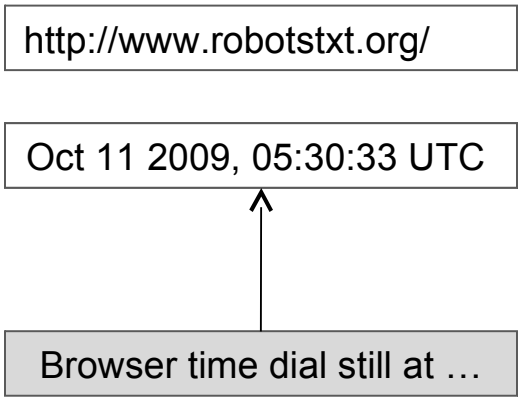
Oct 11 2009, 05:30:33 UTC

From Wikipedia History: Version Sep 15 2009, 20:49:00 UTC

Vision: Navigate the Web of the Past



Vision: Navigate the Web of the Past

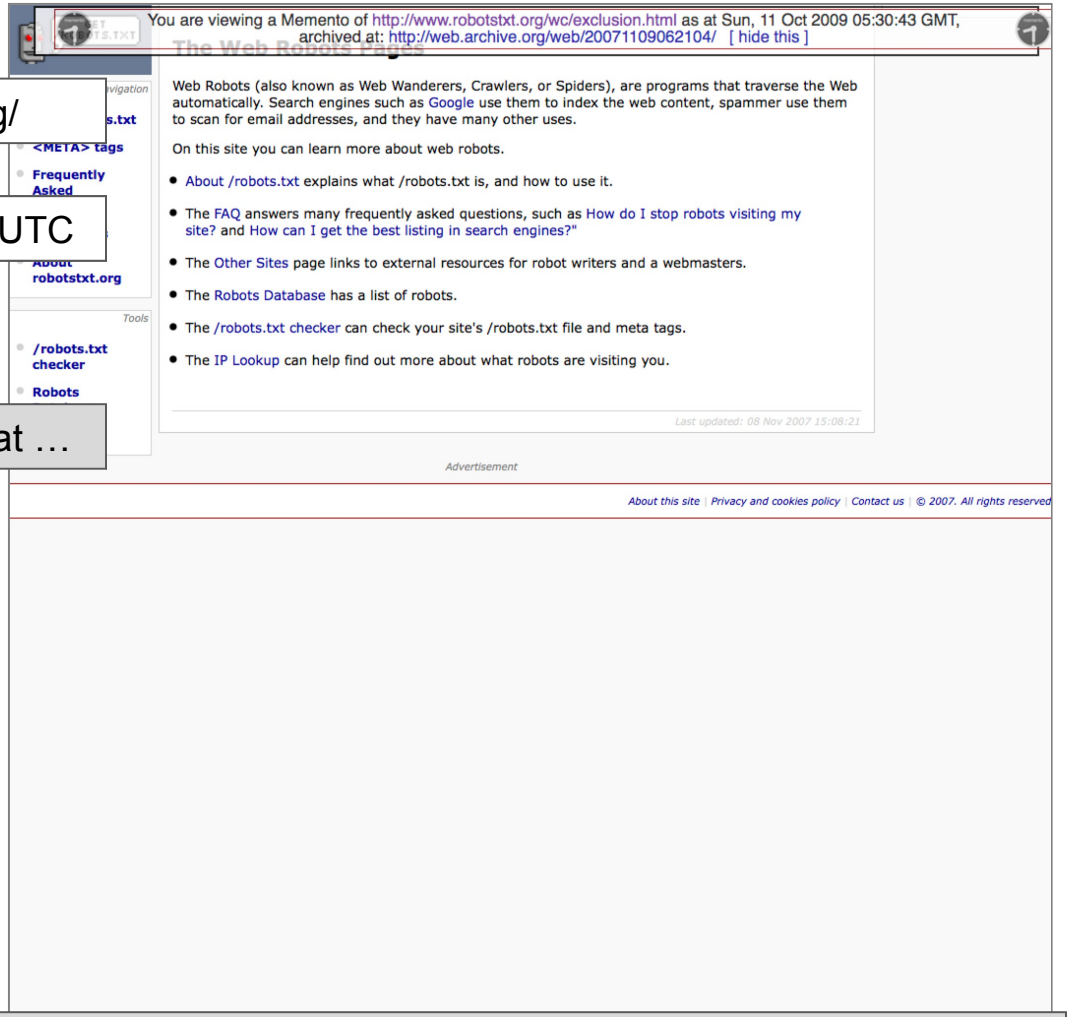


Vision: Navigate the Web of the Past

http://www.robotstxt.org/

Oct 11 2009, 05:30:33 UTC

Browser time dial still at …



From Internet Archive: Version Nov 09 2007, 06:21:04 UTC

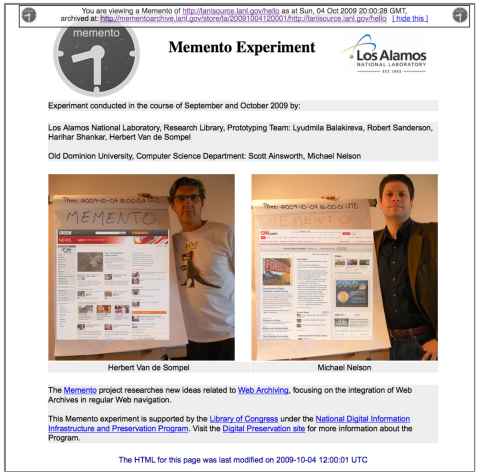
The Memento Approach

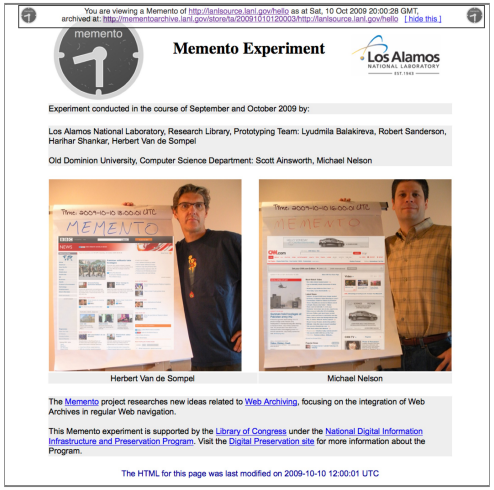
HTTP navigation to an archived resource by leveraging:

- *The original resource;*
- *HTTP datetime content negotiation.*

original resource

Mementos





An HTTP-Based Versioning Mechanism for Linked Data
LDOW 2010, Raleigh, NC

Los Alamos NATIONAL LABORATORY

OLD DOMINION UNIVERSITY

DT-conneg with URI-G to get URI-M

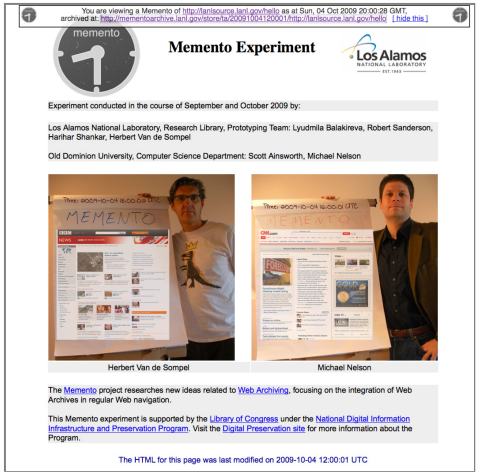
original resource

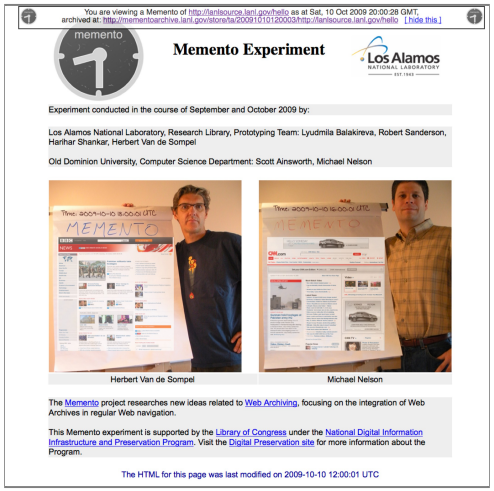
TimeGate

transparently negotiable resource

Mementos

variant resources





original server

archival server

An HTTP-Based Versioning Mechanism for Linked Data
LDOW 2010, Raleigh, NC

Los Alamos NATIONAL LABORATORY

OLD DOMINION UNIVERSITY

HTTP Link timegate

DT-conneg with URI-G to get URI-M

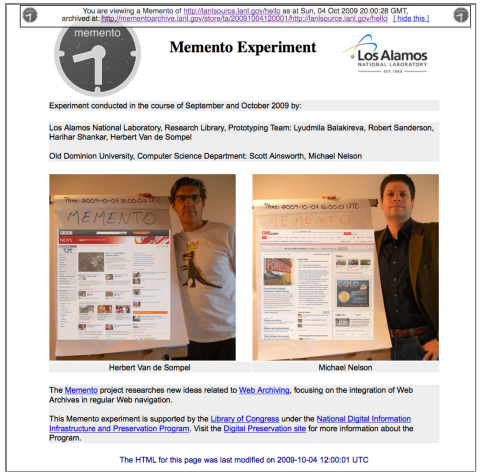
original resource

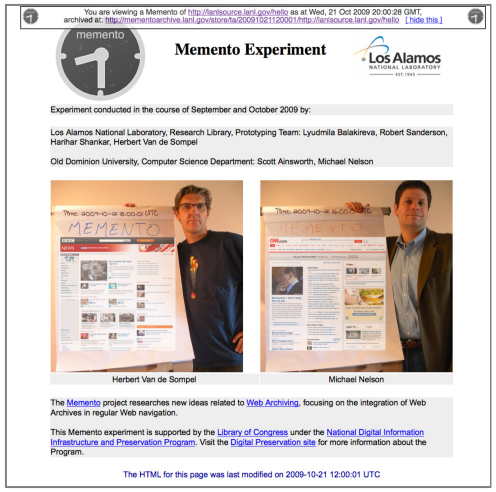
TimeGate

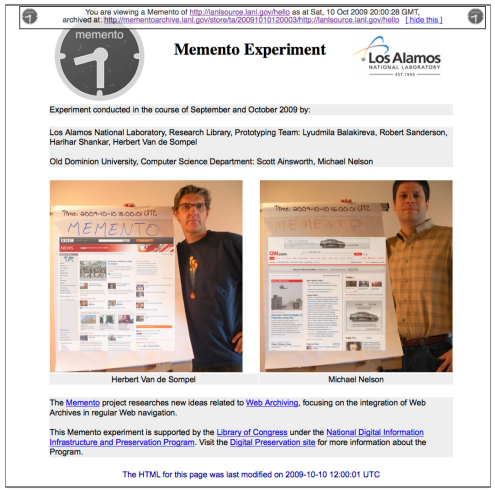
transparently negotiable resource

Mementos

variant resources







original server

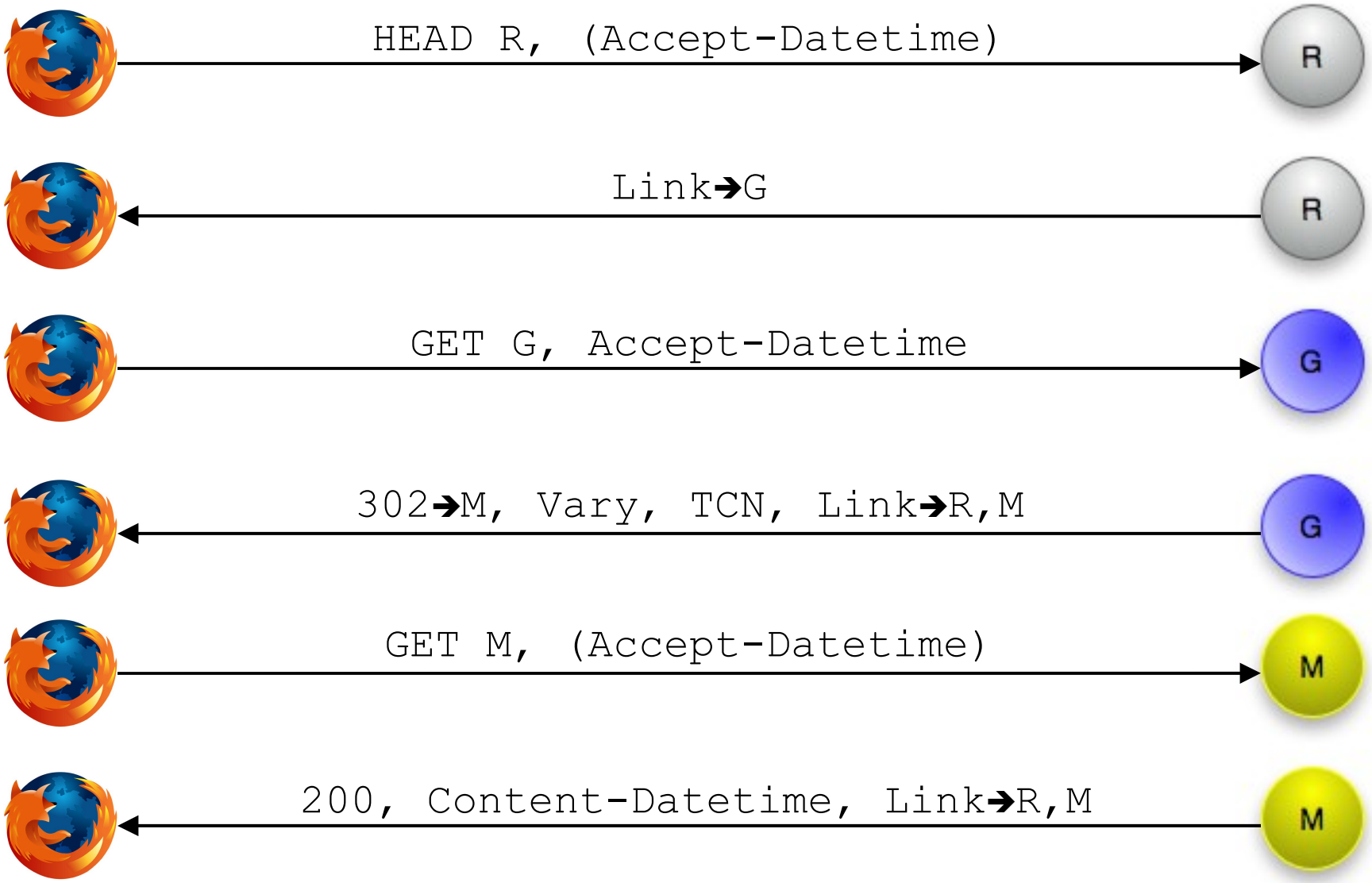
archival server

An HTTP-Based Versioning Mechanism for Linked Data
LDOW 2010, Raleigh, NC

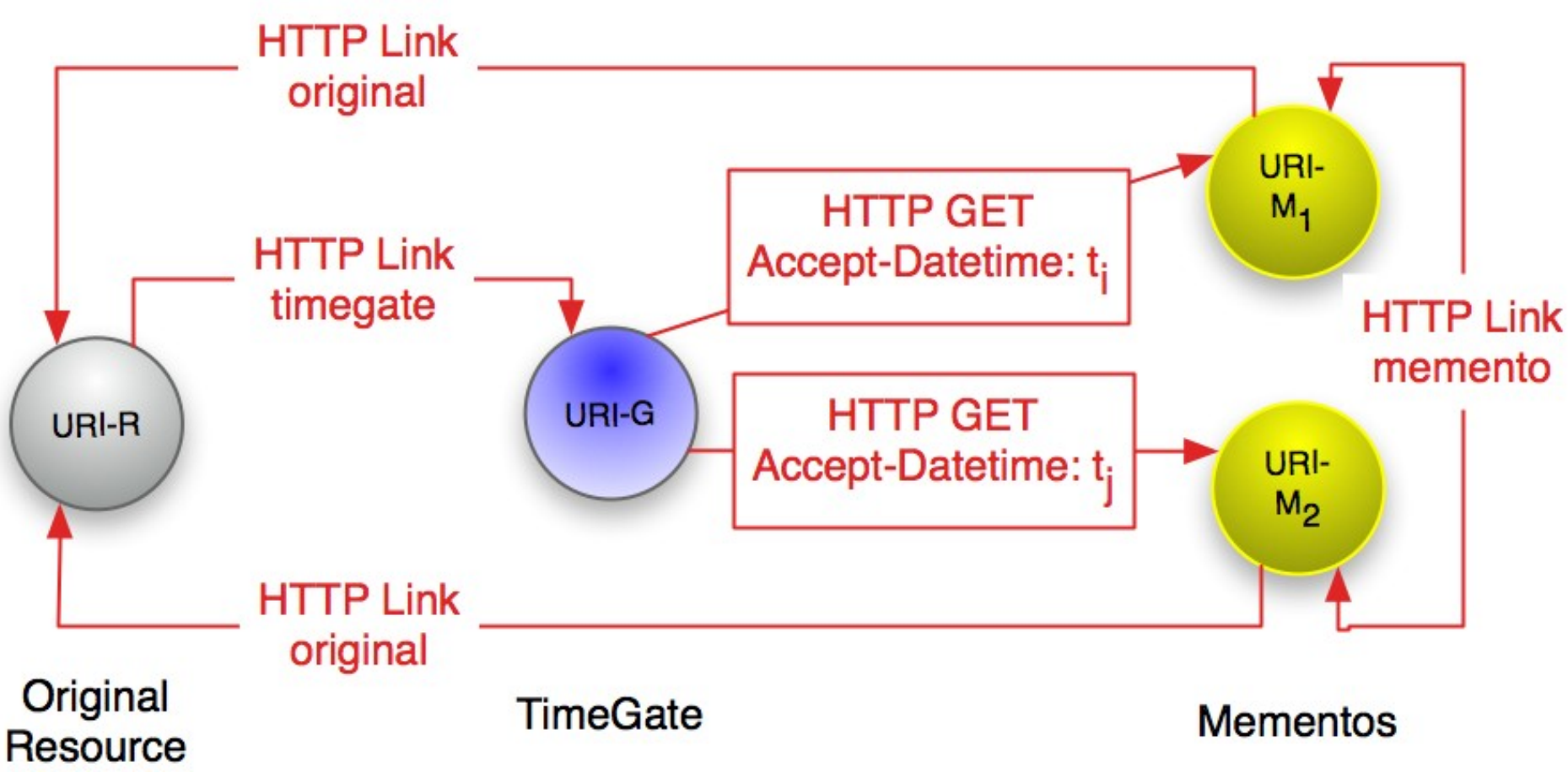
Los Alamos NATIONAL LABORATORY

Old Dominion University

Memento HTTP Flow

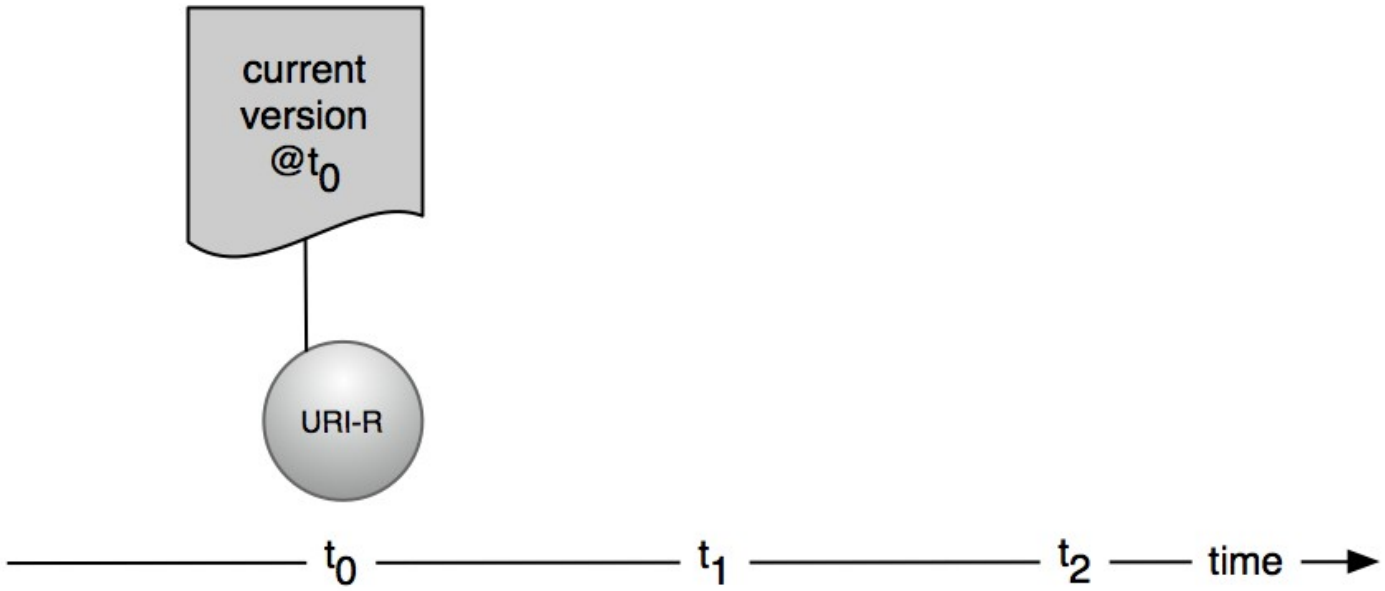


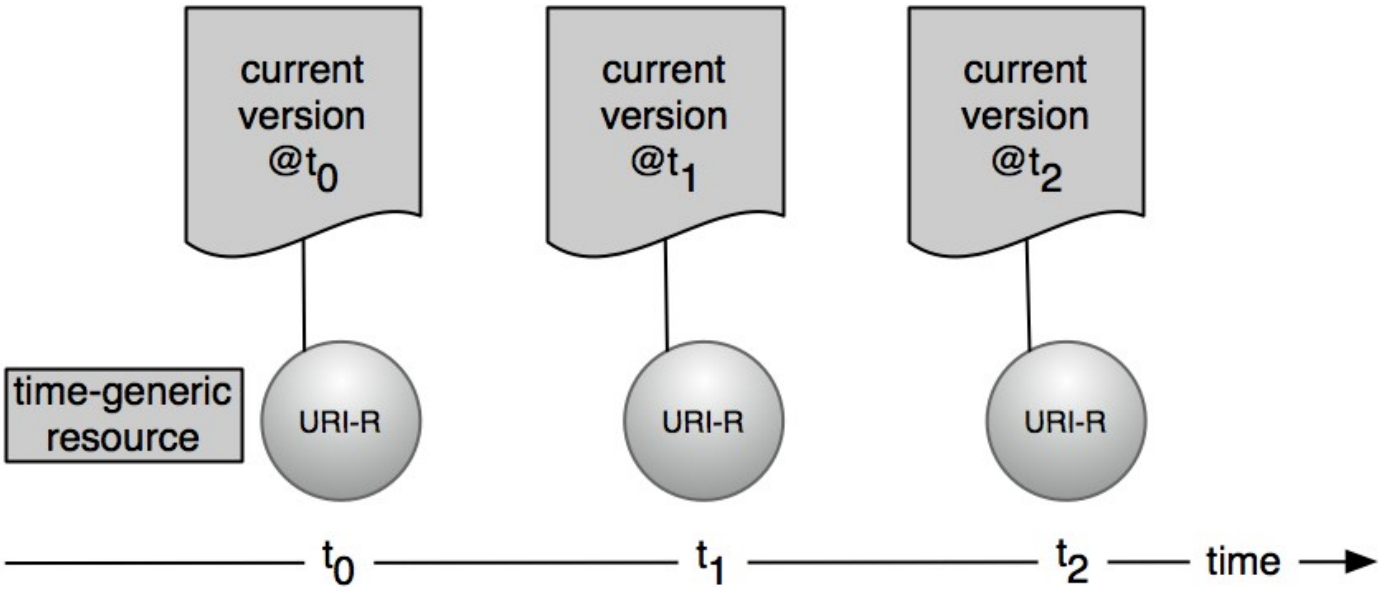
The Memento Framework

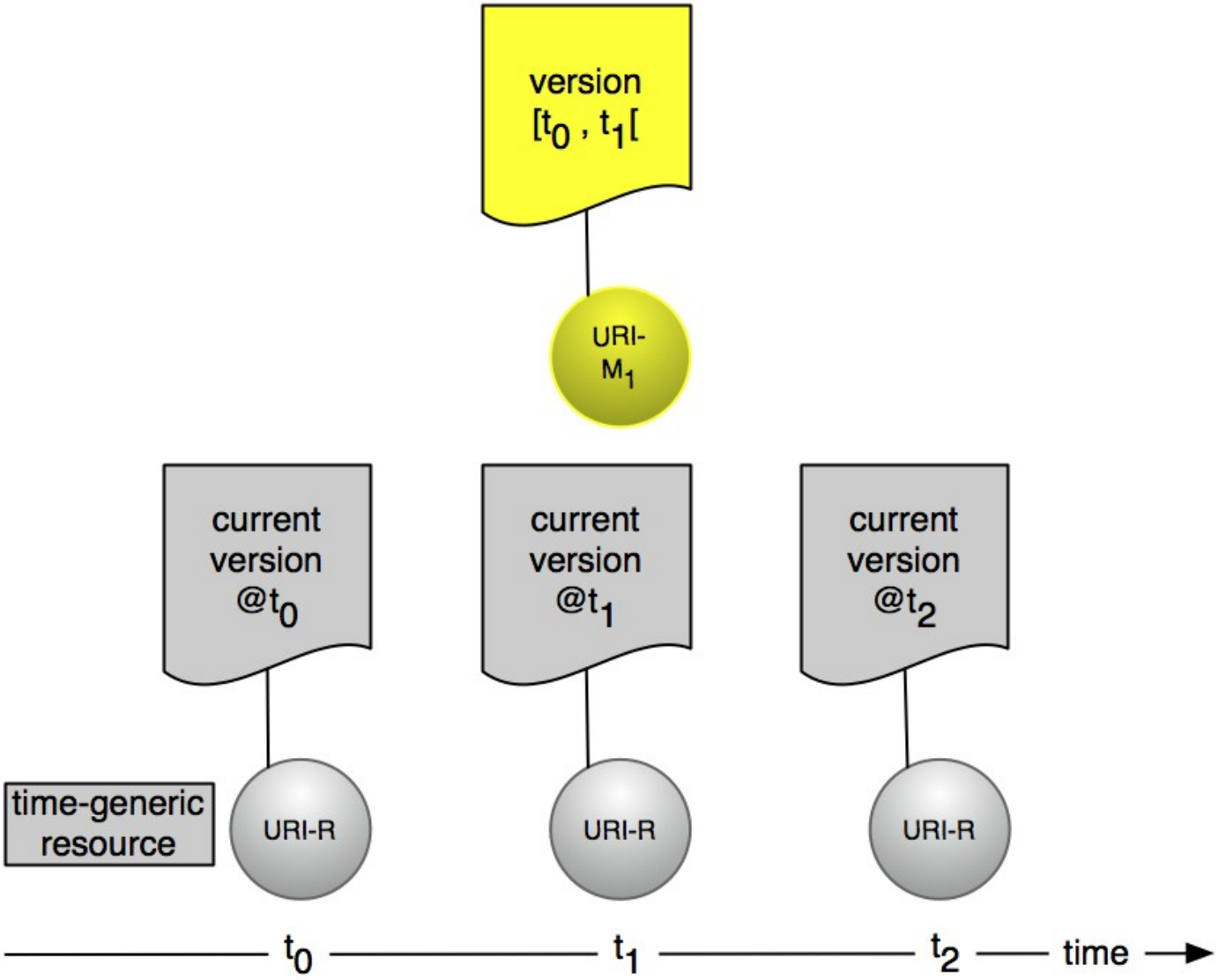


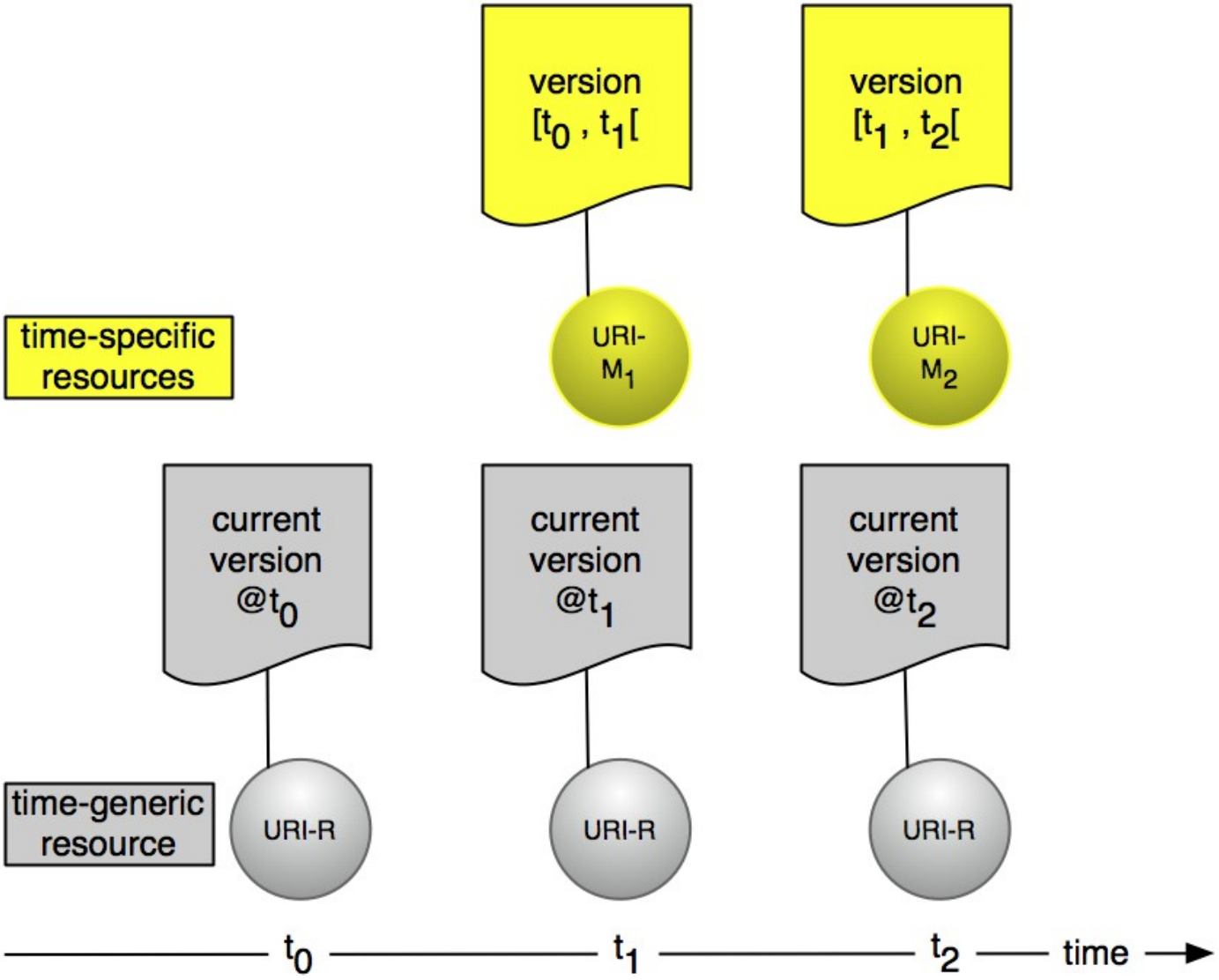
Outline

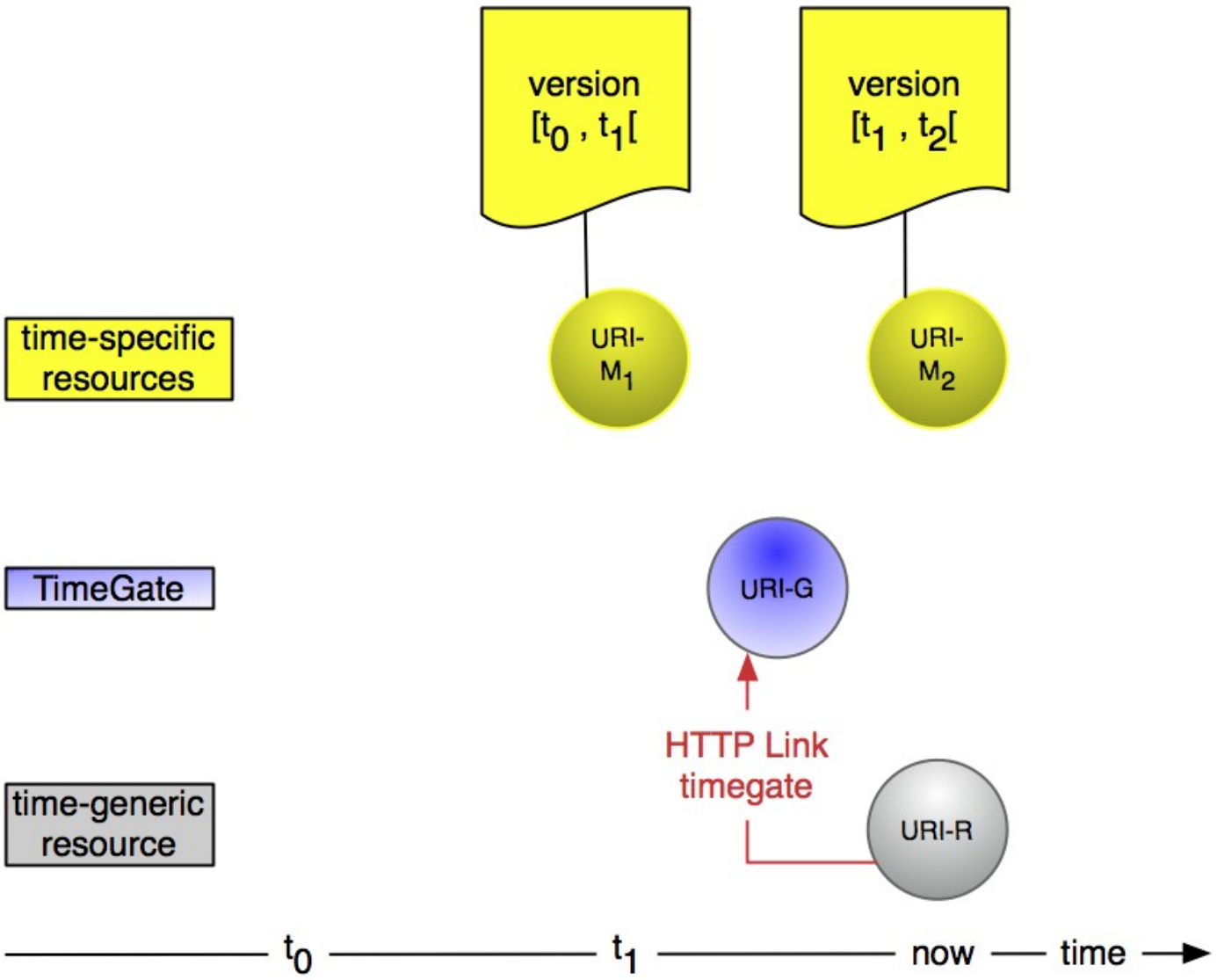
- Memento - Time Travel for the Web
- Resource Versioning suggested by Memento
- Resource Versioning for Linked Data
- DBpedia Demonstrator

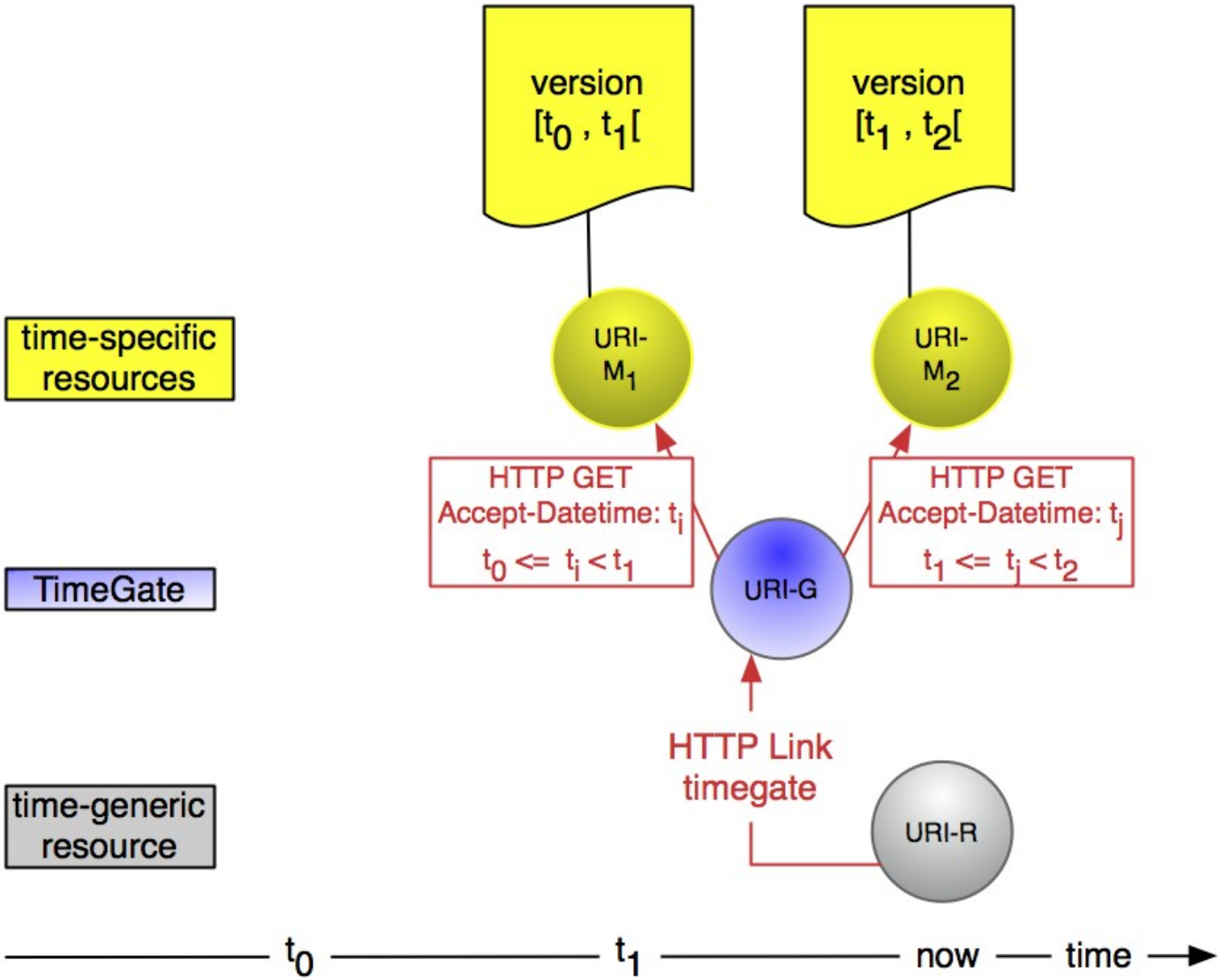

current
version
@t0
URI-R
t0
t1
t2
time

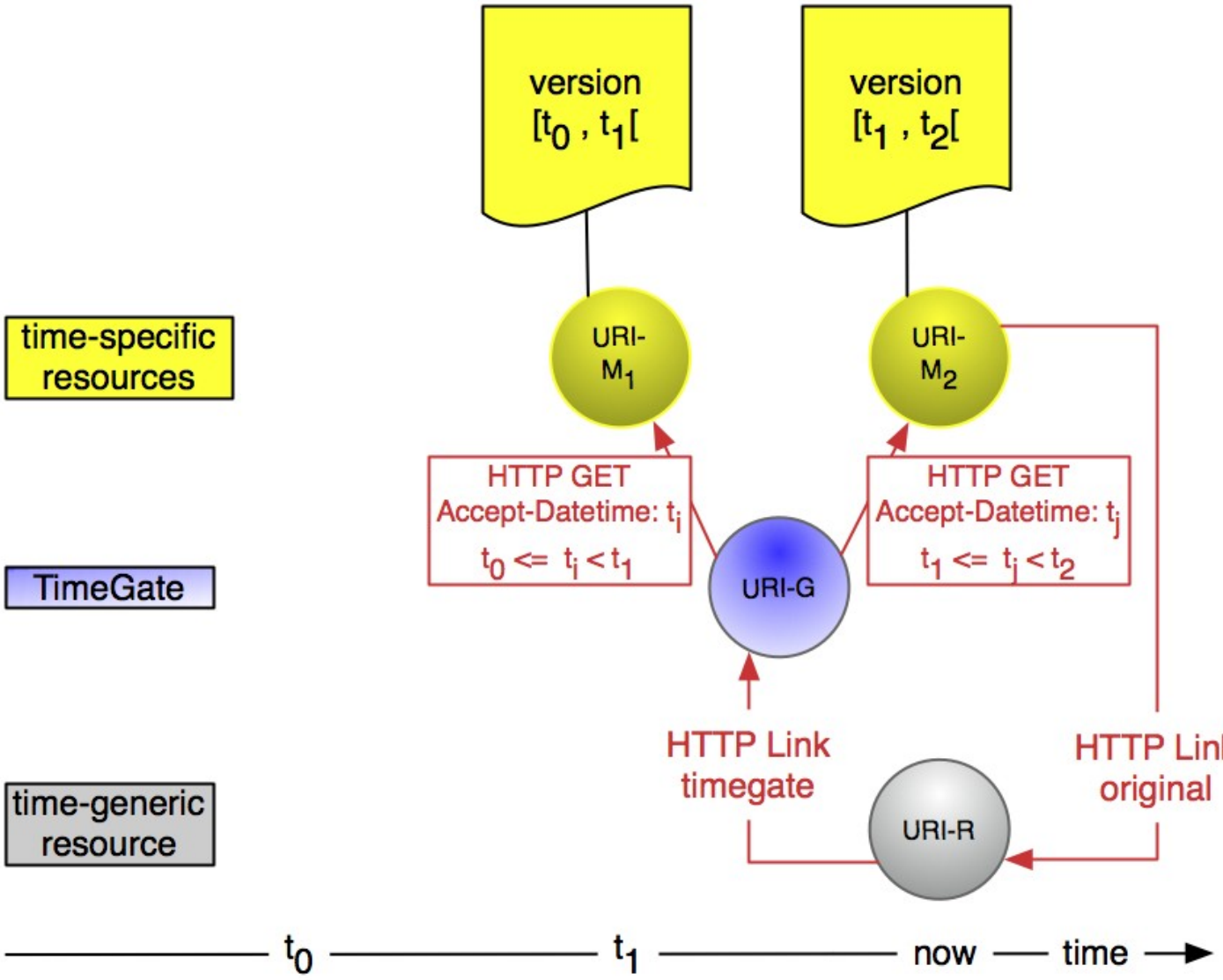

current version @t0
current version @t1
current version @t2
time-generic resource
URI-R
URI-R
URI-R
t0
t1
t2
time


version
[t0 , t1[
URI-
M1
current
version
@t0
current
version
@t1
current
version
@t2
time-generic
resource
URI-R
URI-R
URI-R
t0
t1
t2
time


version
[t0 , t1[
version
[t1 , t2[
time-specific
resources
URI-
M1
URI-
M2
current
version
@t0
current
version
@t1
current
version
@t2
time-generic
resource
URI-R
URI-R
URI-R
t0
t1
t2
time


version
[t0 , t1[
version
[t1 , t2[
time-specific
resources
URI-
M1
URI-
M2
TimeGate
URI-G
HTTP Link
timegate
time-generic
resource
URI-R
t0
t1
now
time


version
[t0 , t1[
version
[t1 , t2[
time-specific resources
URI-M1
URI-M2
HTTP GET
Accept-Datetime: ti
t0 <= ti < t1
HTTP GET
Accept-Datetime: tj
t1 <= tj < t2
TimeGate
URI-G
HTTP Link
timegate
time-generic resource
URI-R
t0
t1
now
time


version
[t0 , t1[
version
[t1 , t2[
time-specific resources
URI-M1
URI-M2
HTTP GET
Accept-Datetime: ti
t0 <= ti < t1
HTTP GET
Accept-Datetime: tj
t1 <= tj < t2
TimeGate
URI-G
HTTP Link
timegate
HTTP Link
original
time-generic resource
URI-R
t0
t1
now
time

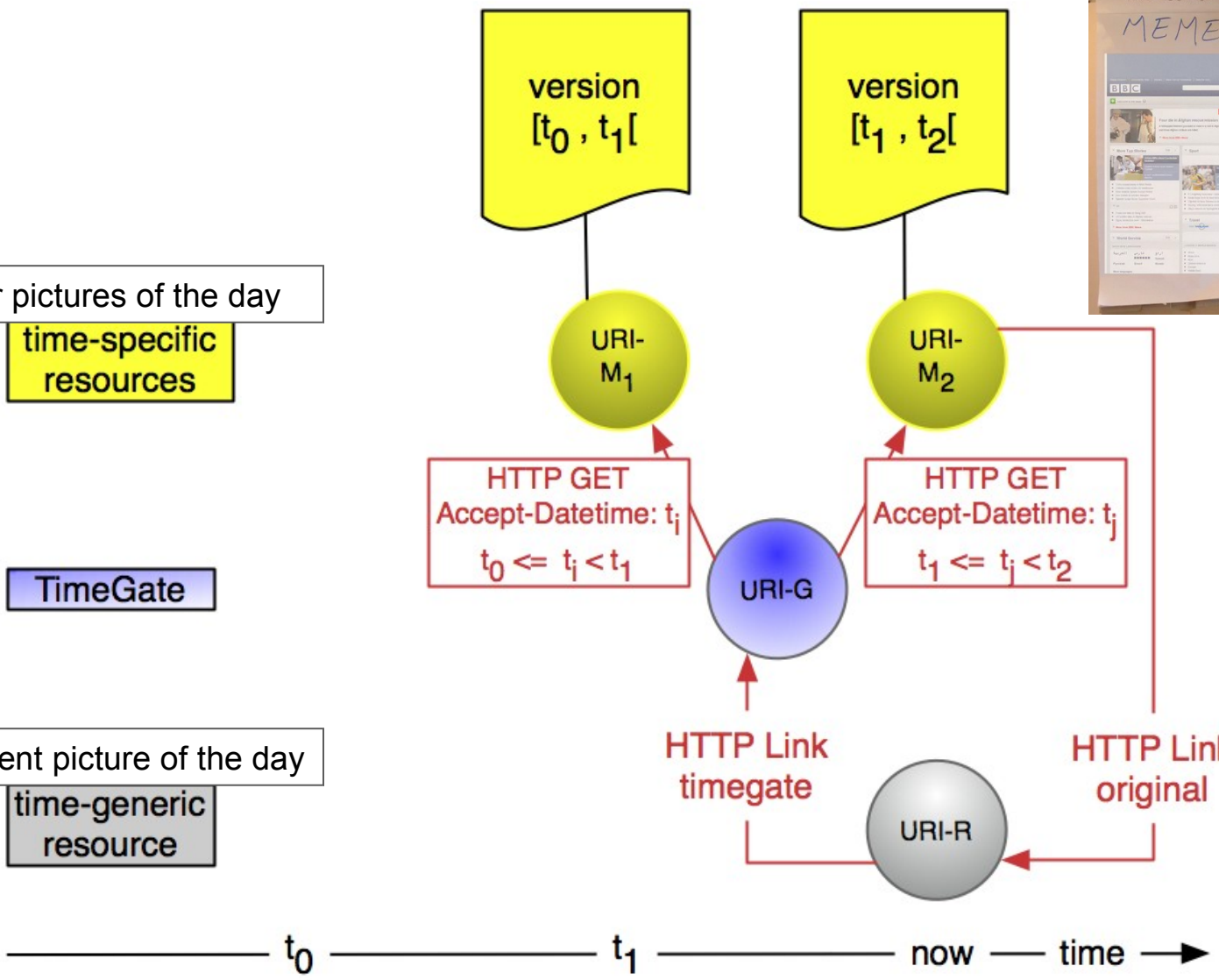
Prior pictures of the day

time-specific resources

TimeGate

Current picture of the day

time-generic resource

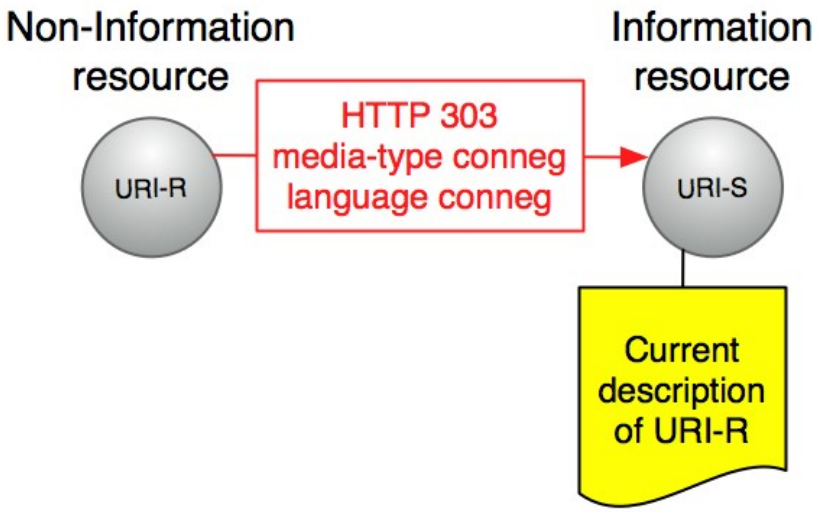


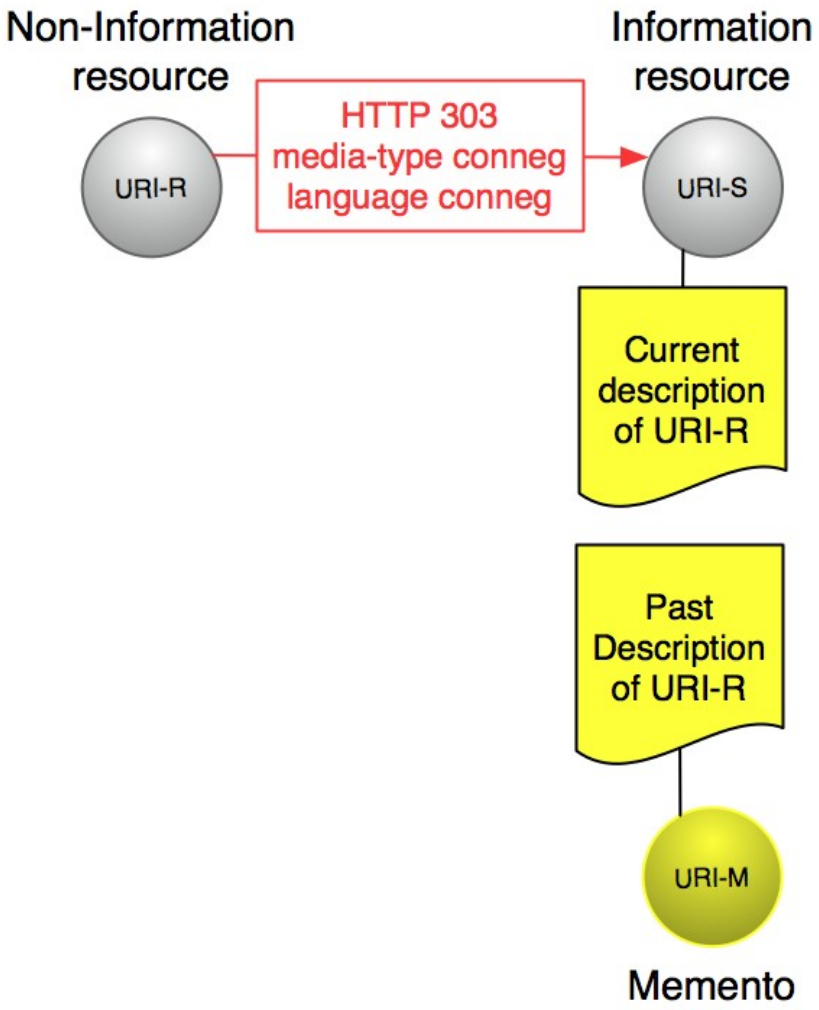
Time Travel across Versions of a *Picture of the Day*

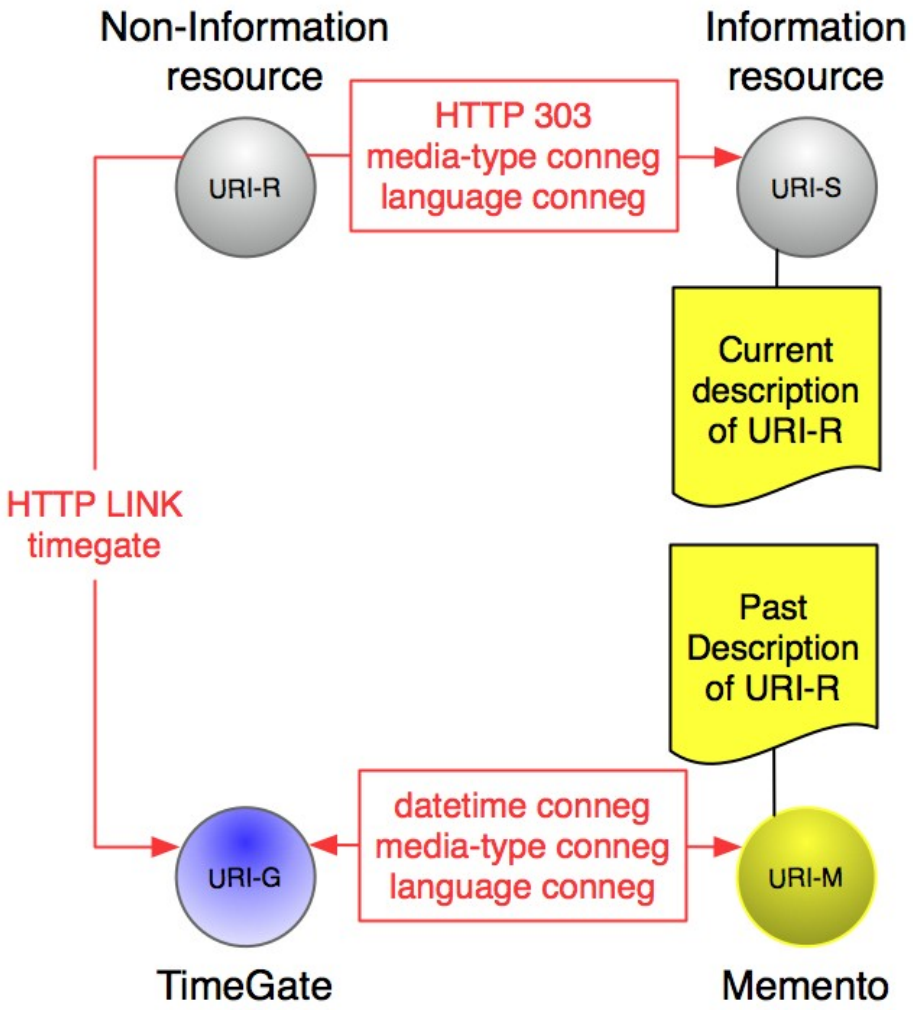
Data collected through HTTP Navigation

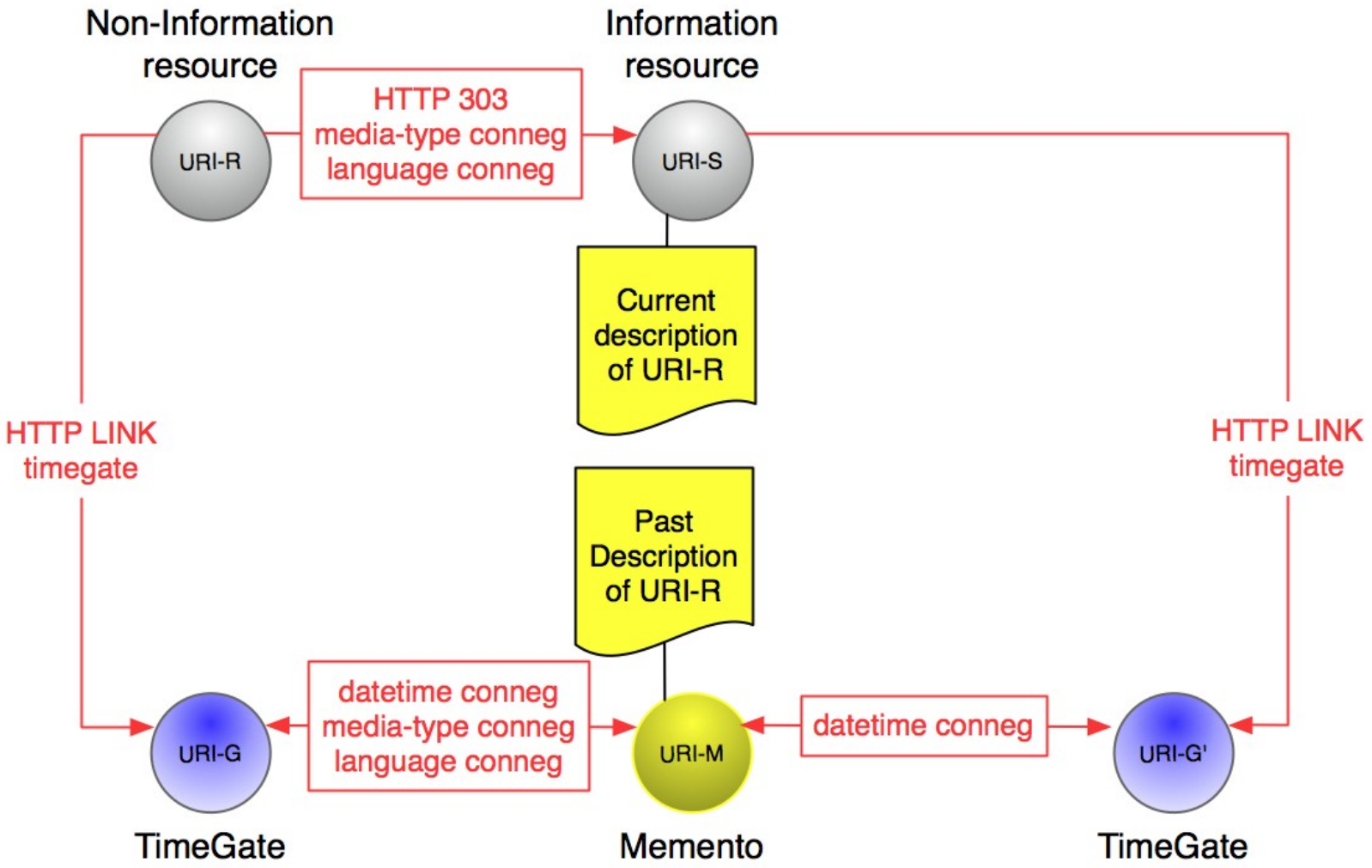
Outline

- Memento - Time Travel for the Web
- Resource Versioning suggested by Memento
- Resource Versioning for Linked Data
- DBpedia Demonstrator


Non-Information resource
Information resource
HTTP 303
media-type conneg
language conneg
URI-R
URI-S
Current description of URI-R

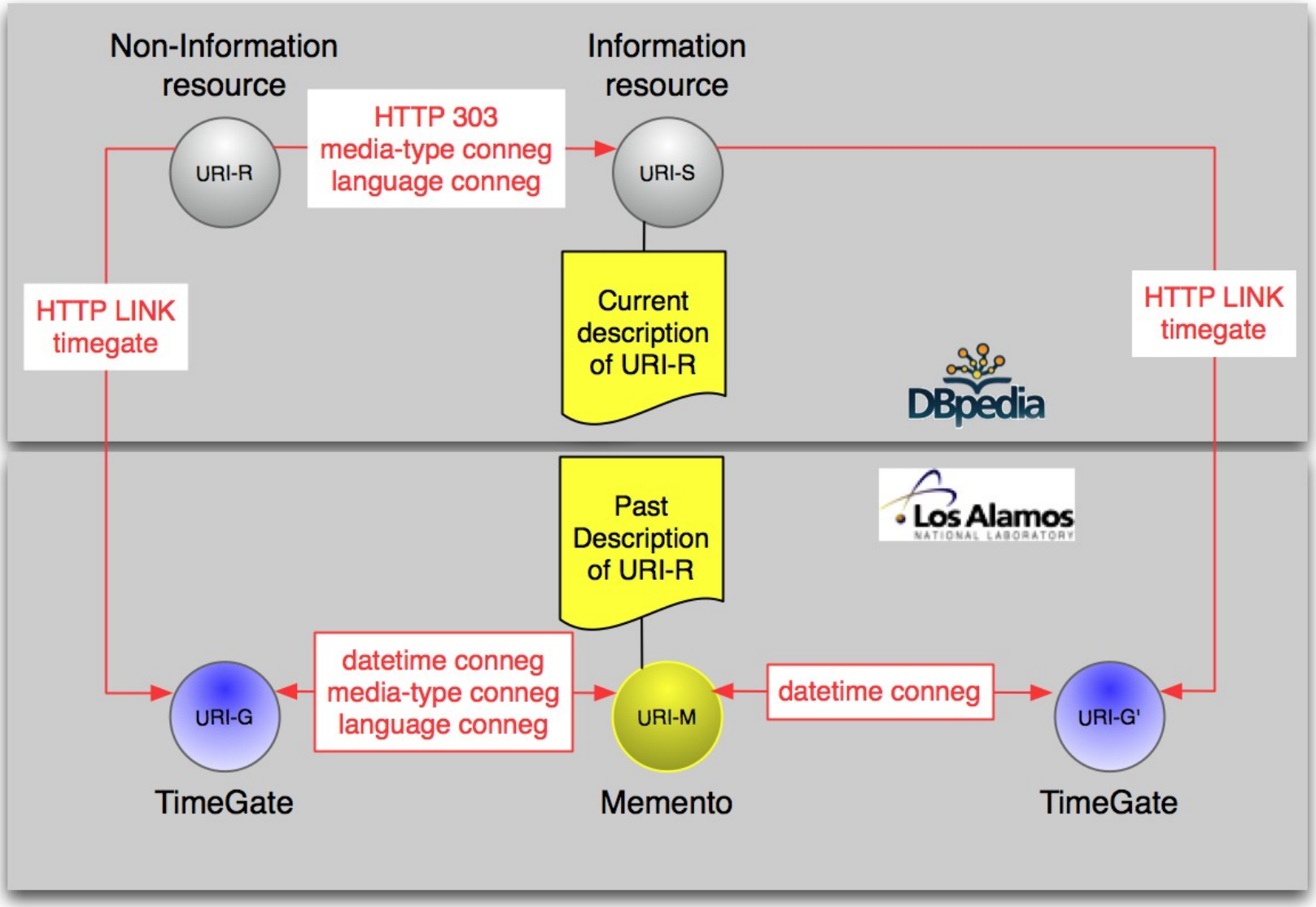

Non-Information resource
Information resource
HTTP 303
media-type conneg
language conneg
URI-R
URI-S
Current description of URI-R
Past Description of URI-R
URI-M
Memento


Non-Information resource
Information resource
HTTP 303
media-type conneg
language conneg
URI-R
URI-S
Current description of URI-R
HTTP LINK
timegate
Past Description of URI-R
datetime conneg
media-type conneg
language conneg
URI-G
URI-M
TimeGate
Memento

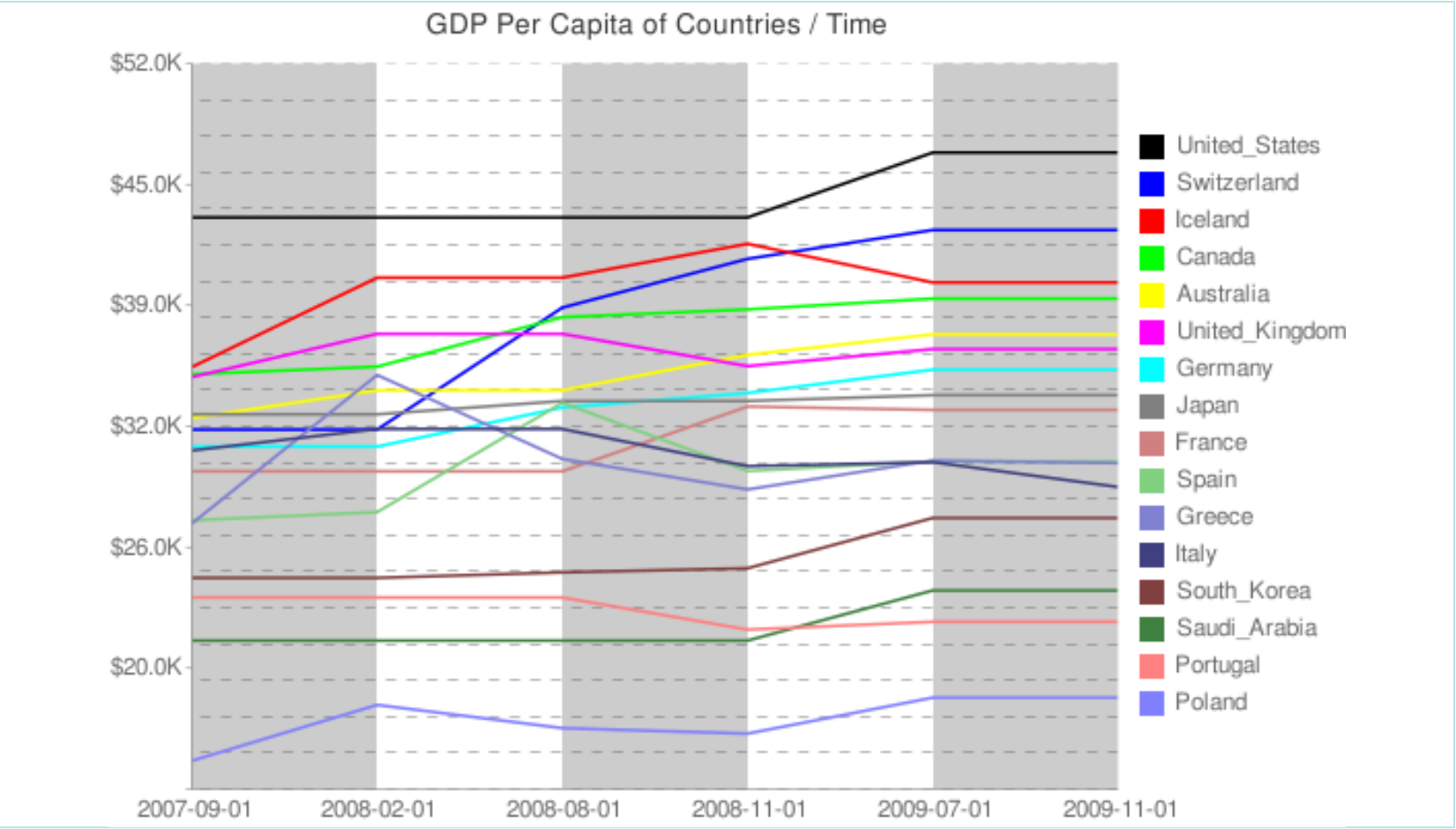

Non-Information resource
Information resource
HTTP 303
media-type conneg
language conneg
URI-R
URI-S
Current description of URI-R
HTTP LINK timegate
HTTP LINK timegate
Past Description of URI-R
datetime conneg
media-type conneg
language conneg
datetime conneg
URI-G
URI-M
URI-G'
TimeGate
Memento
TimeGate

Outline

- Memento - Time Travel for the Web
- Resource Versioning suggested by Memento
- Resource Versioning for Linked Data
- DBpedia Demonstrator

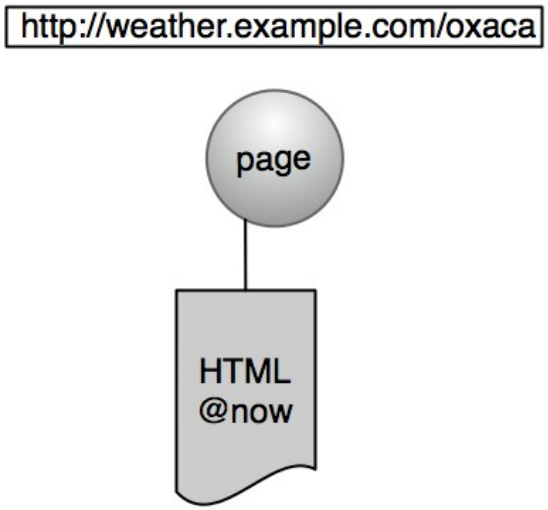

Non-Information resource
Information resource
HTTP 303
media-type conneg
language conneg
URI-R
URI-S
HTTP LINK
timegate
Current description of URI-R
HTTP LINK
timegate
DBpedia
Past Description of URI-R
Los Alamos
NATIONAL LABORATORY
datetime conneg
media-type conneg
language conneg
datetime conneg
URI-G
URI-M
URI-G'
TimeGate
Memento
TimeGate

Time-Series Analysis across DBpedia Versions



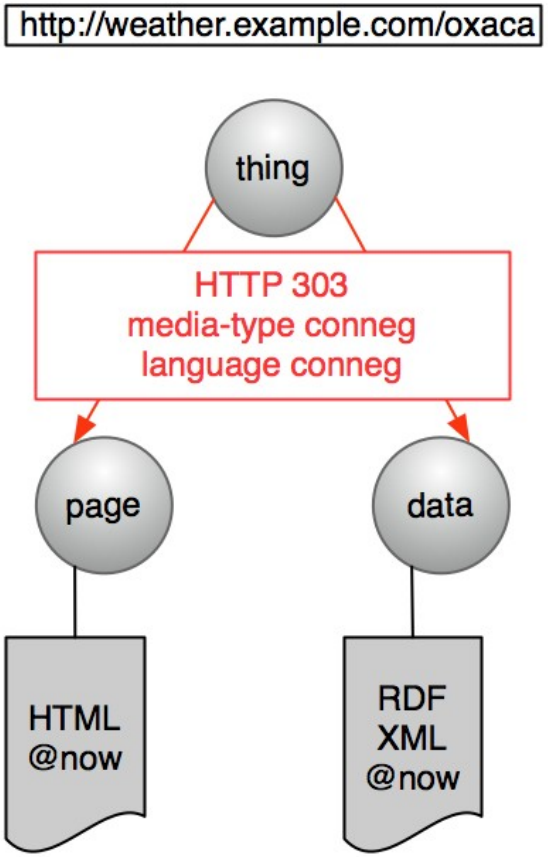
Data collected through HTTP Navigation

Conclusions



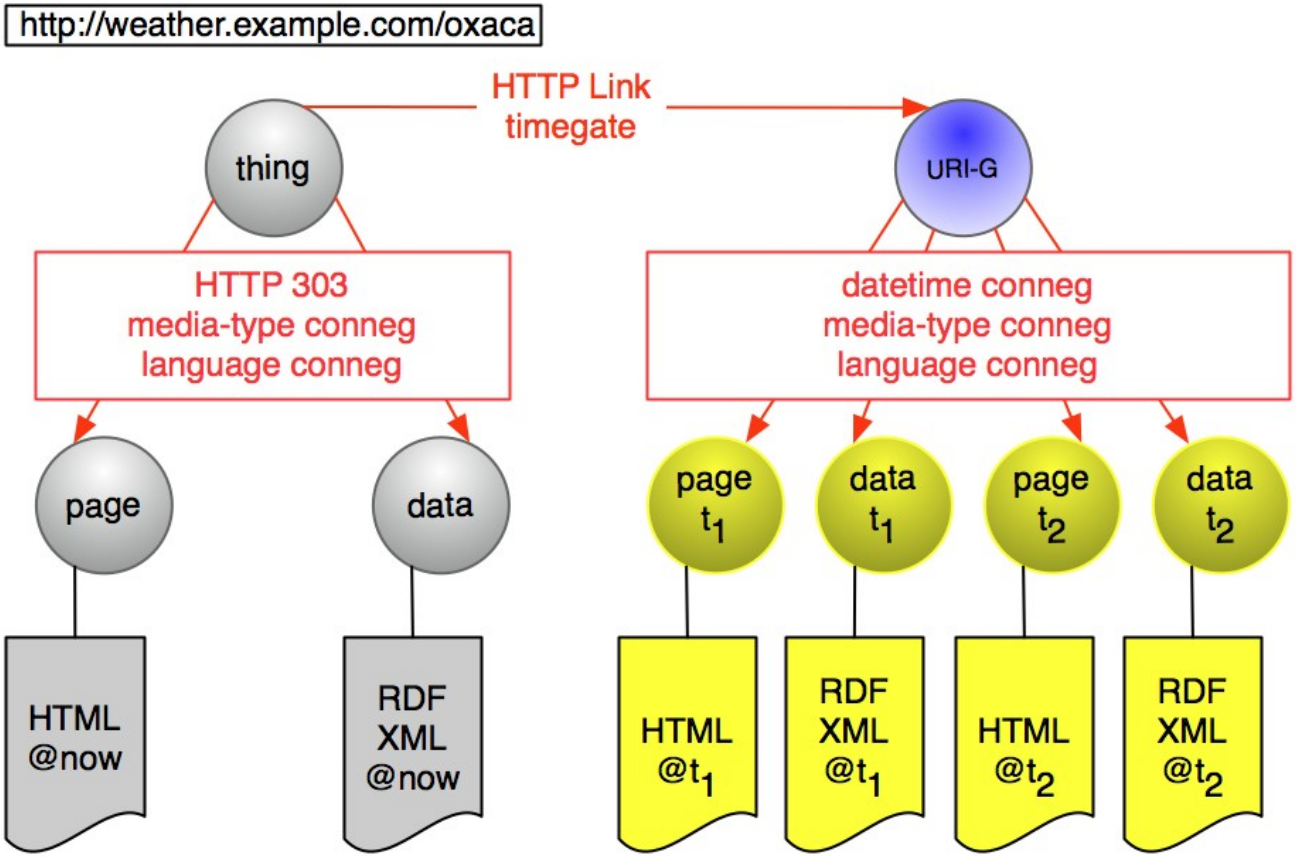
URI as access point to page

Conclusions



Increased value: URI as access point to page and data

Conclusions



Increased value: URI as access point to current & historical page and data

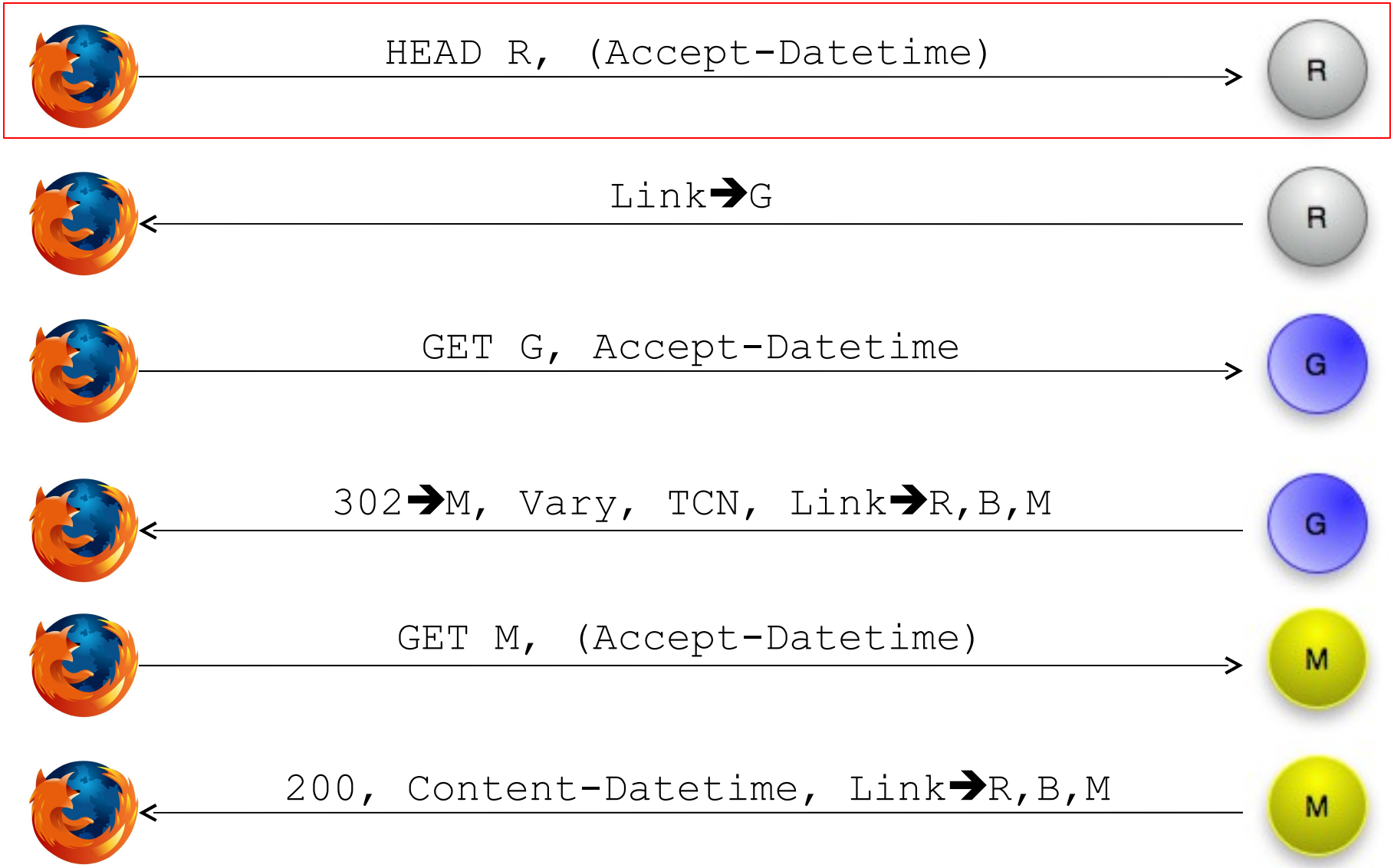
Memento wants to make navigating the Web's Past Easy

http://www.mementoweb.org
http://groups.google.com/group/memento-dev

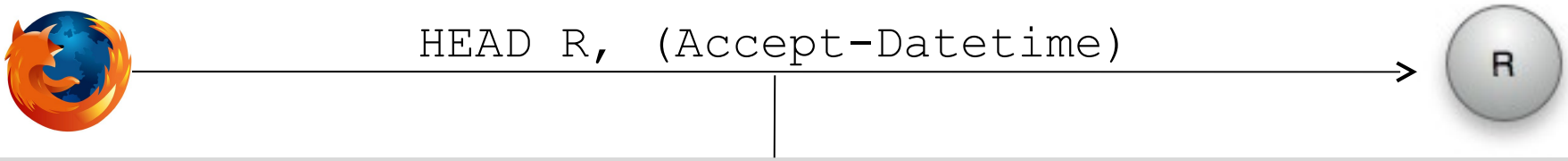
References

- Tim Berners-Lee (1996,2000) Generic Resources.
  http://www.w3.org/DesignIssues/Generic.html
- 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
- Van de Sompel, H., Nelson, M.L., Sanderson, R., Balakireva, L., Ainsworth, S., Shankar, H. (2009) Memento: Time Travel for the Web.
  http://arxiv.org/abs/0911.1112

Memento HTTP Flow

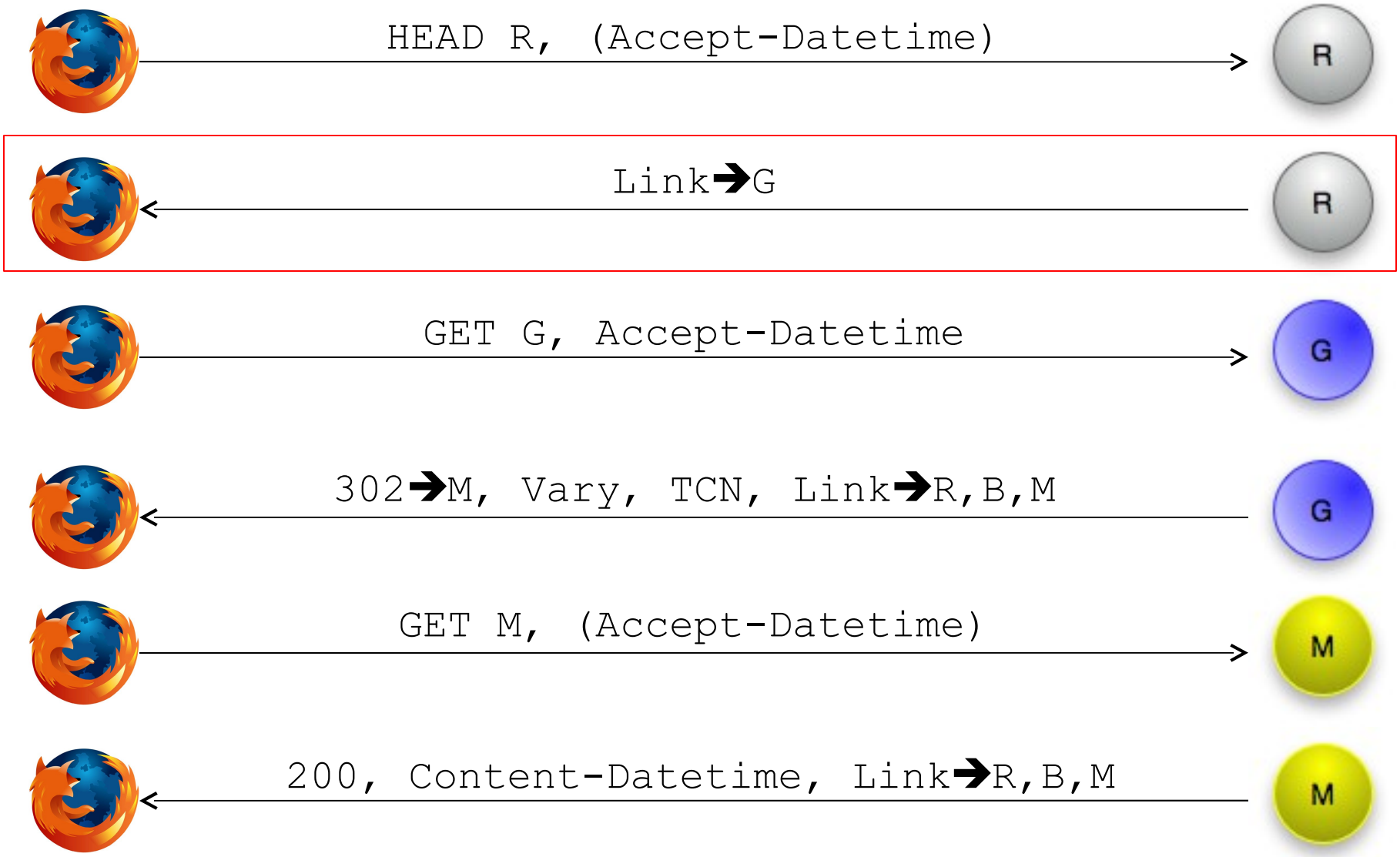


Memento HTTP Flow: URI-R

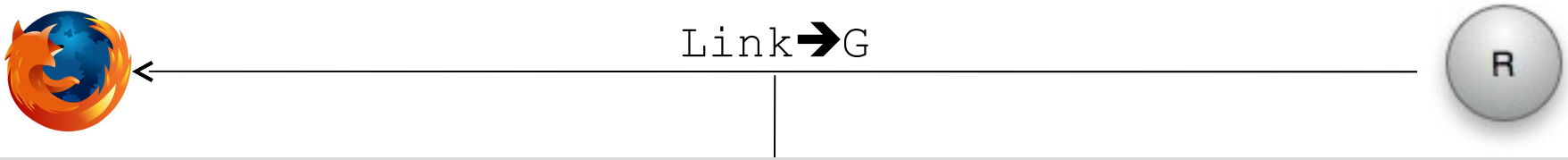



```
HEAD /resource/France HTTP/1.1
Host: dbpedia.org
Accept: text/html,application/xhtml+xml,application/xml;q=0.9,*/*;q=0.8
Accept-Language: en-us,en;q=0.5
Accept-Encoding: gzip,deflate
Accept-Charset: ISO-8859-1,utf-8;q=0.7,*;q=0.7
```

Memento HTTP Flow

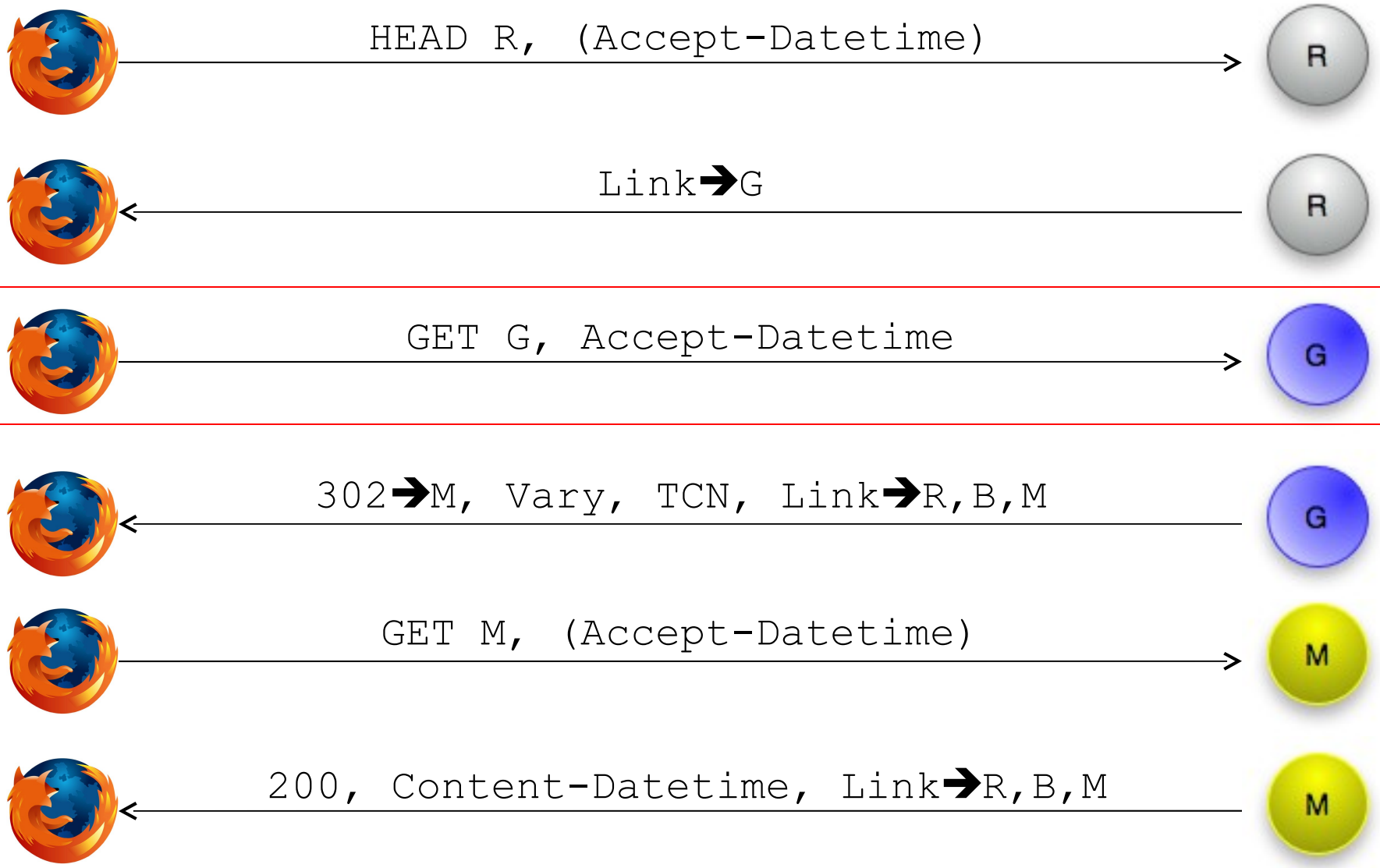


Memento HTTP Flow: Success – URI-R

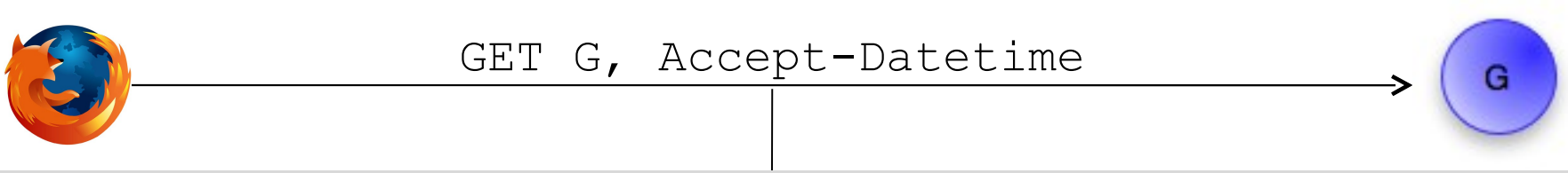



```
HTTP/1.1 303 See Other
Server: Virtuoso/06.01.3127 (Solaris) x86_64-sun-solaris2.10-64  VDB
Connection: close
Content-Type: text/html; charset=UTF-8
Date: Tue, 20 Apr 2010 16:48:51 GMT
Accept-Ranges: bytes
Location: http://dbpedia.org/page/France
Transfer-Encoding: chunked
Content-Encoding: gzip
Link:
<http://mementoarchive.lanl.gov/dbpedia/timegate/http://dbpedia.org/resource/France>
; rel="timegate"
```

Memento HTTP Flow

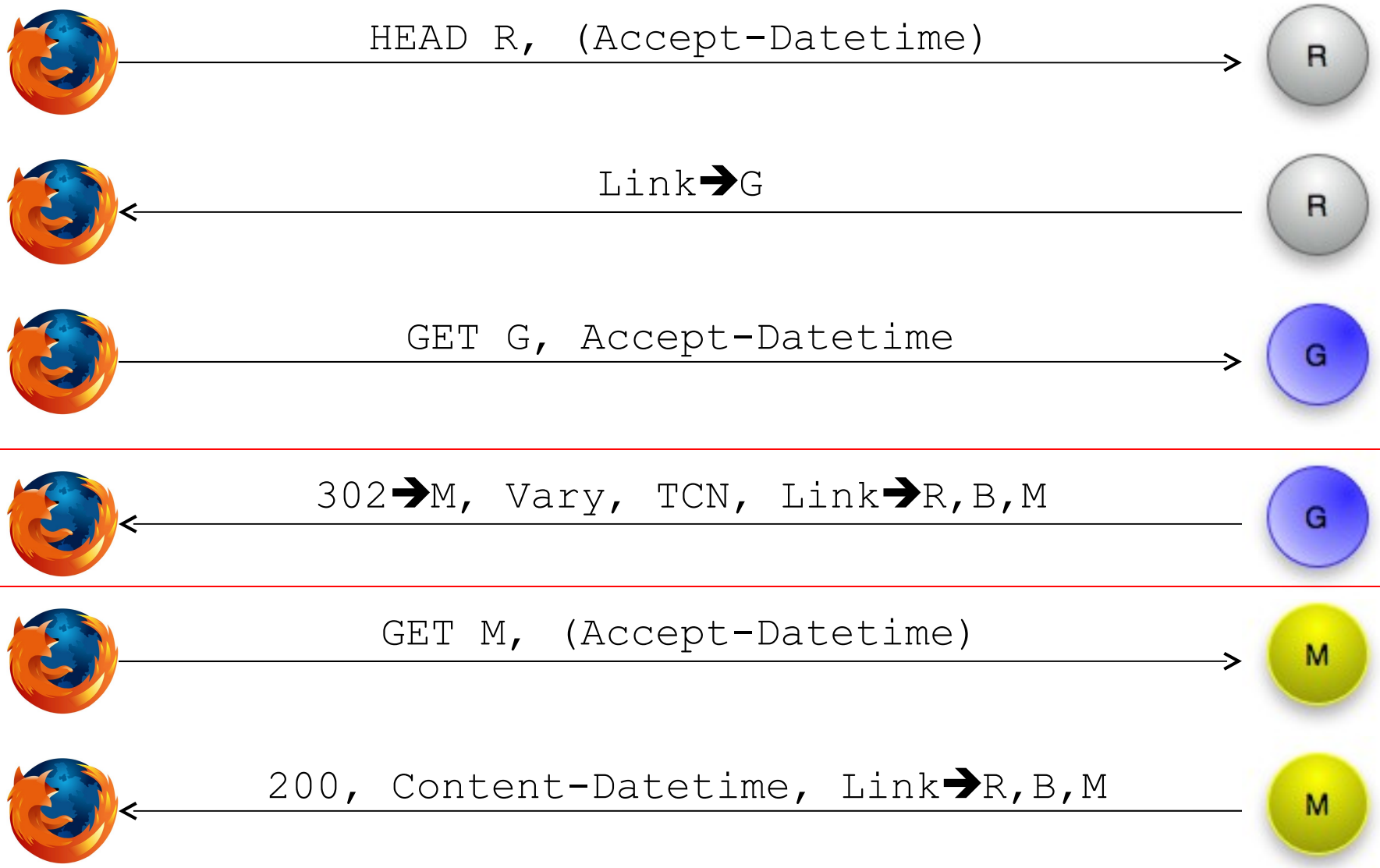


Memento HTTP Flow: URI-G

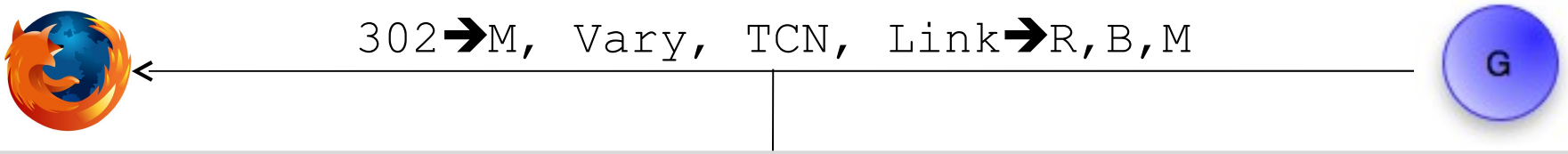



```
GET /dbpedia/timegate/http://dbpedia.org/resource/France HTTP/1.1
Host: mementoarchive.lanl.gov
Accept: text/html,application/xhtml+xml,application/xml;q=0.9,*/*;q=0.8
Accept-Language: en-us,en;q=0.5
Accept-Encoding: gzip,deflate
Accept-Charset: ISO-8859-1,utf-8;q=0.7,*;q=0.7
Accept-Datetime: Wed, 08 Jul 2009 06:00:00 GMT
```

Memento HTTP Flow

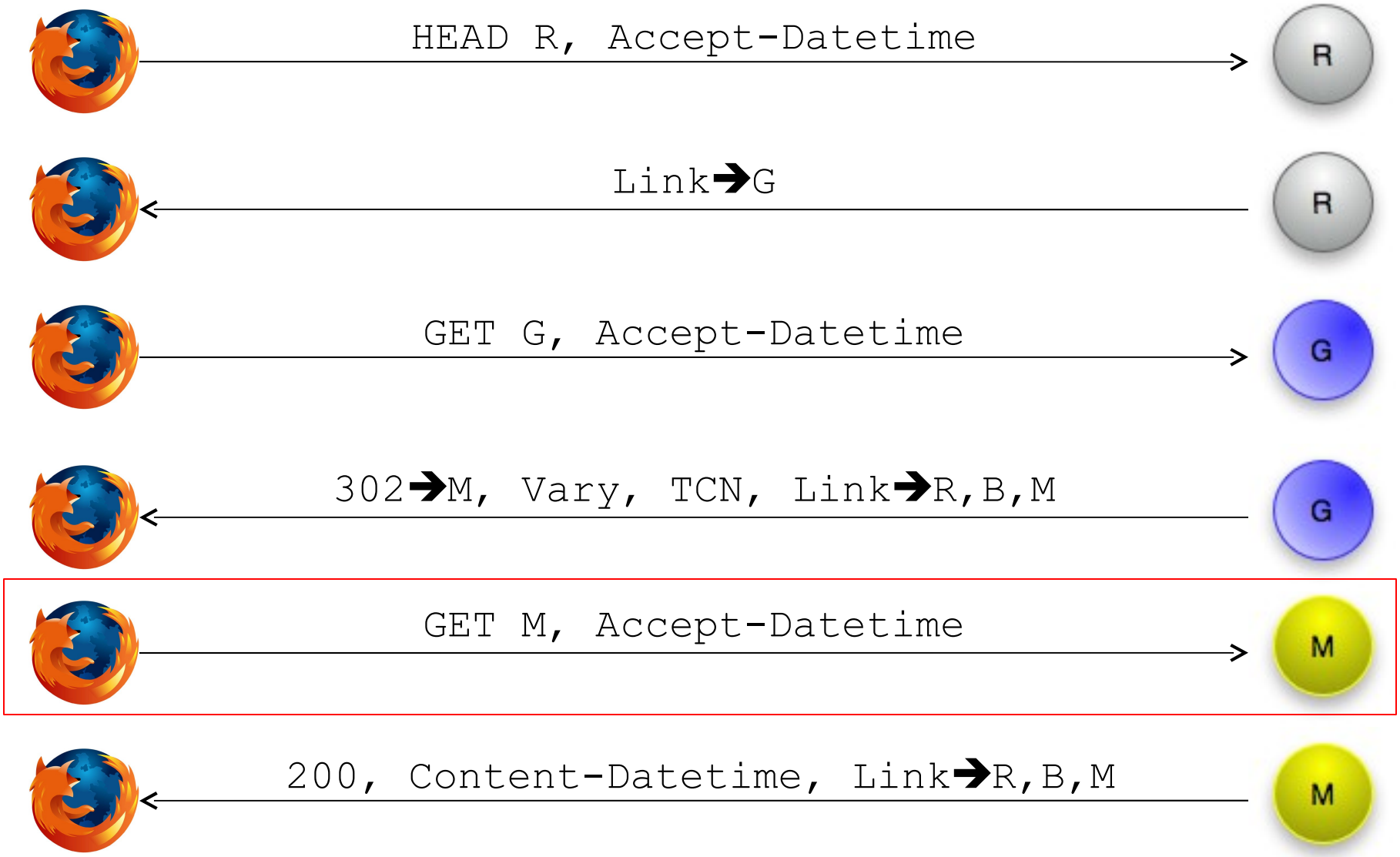


Memento HTTP Flow: Success – URI-G

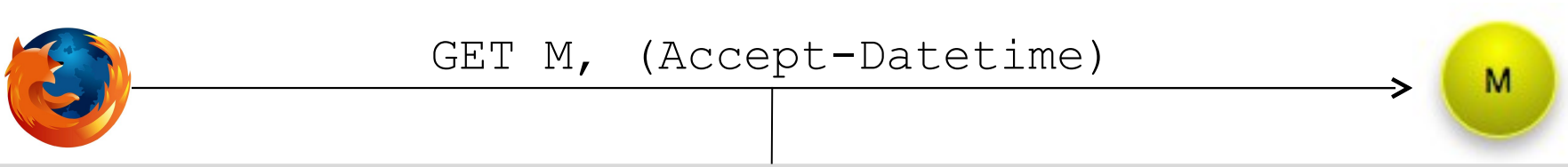



```
HTTP/1.1 302 Found
Date: Tue, 20 Apr 2010 17:51:00 GMT
Server: Apache
TCN: choice
Vary: negotiate, Accept-Datetime
Location:
http://mementoarchive.lanl.gov/dbpedia/memento/20090701/http://dbpedia.org/page/Fran
ce.html
Link:
 <http://dbpedia.org/resource/France>;rel="original",
 <http://mementoarchive.lanl.gov/dbpedia/memento/20070901/http://dbpedia.org/page/
  France.html>;rel="first-memento";datetime="Sat, 01 Sep 2007 00:00:00 GMT",
 <http://mementoarchive.lanl.gov/dbpedia/memento/20091101/http://dbpedia.org/page/
  France.html>;rel="last-memento next-memento";datetime="Sun, 01 Nov 2009 00:00:00
  GMT",
 <http://mementoarchive.lanl.gov/dbpedia/memento/20081101/http://dbpedia.org/page/
  France.html>;rel="prev-memento";datetime="Sat, 01 Nov 2008 00:00:00 GMT",
 <http://mementoarchive.lanl.gov/dbpedia/memento/20090701/http://dbpedia.org/page/
  France.html>;rel="memento";datetime="Wed, 01 Jul 2009 00:00:00 GMT”,
 <http://mementoarchive.lanl.gov/dbpedia/timebundle/http://dbpedia.org/page/
  France>;rel="timebundle”
Transfer-Encoding: chunked
```

Memento HTTP Flow

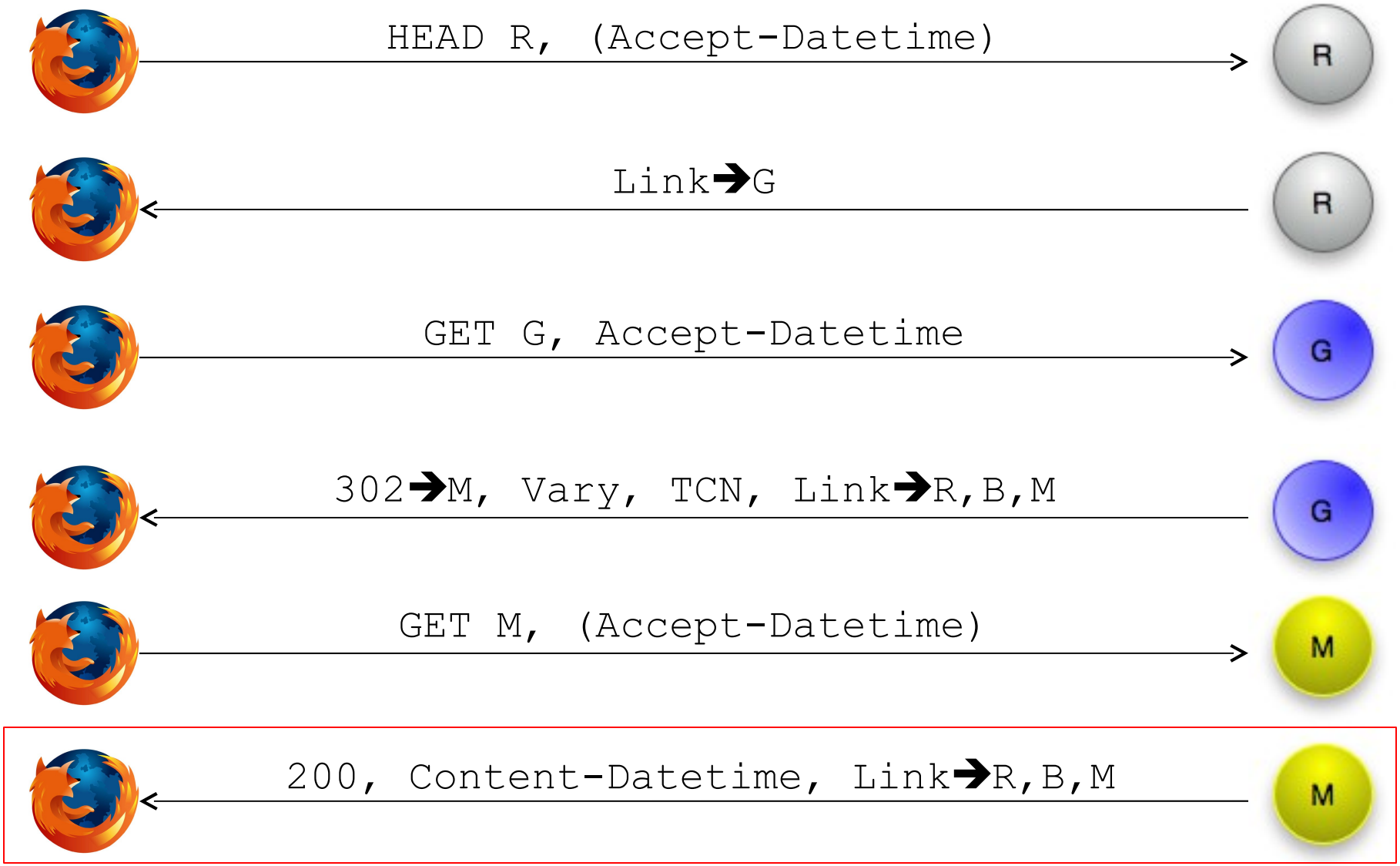


Memento HTTP Flow: URI-M




```
GET /dbpedia/memento/20090701/http://dbpedia.org/page/France.html HTTP/1.1
Host: mementoarchive.lanl.gov
Accept: text/html,application/xhtml+xml,application/xml;q=0.9,*/*;q=0.8
Accept-Language: en-us,en;q=0.5
Accept-Encoding: gzip,deflate
Accept-Charset: ISO-8859-1,utf-8;q=0.7,*;q=0.7
Connection: close
```

Memento HTTP Flow



Memento HTTP Flow: Success – URI-M

200, Content-Datetime, Link➔R,B,M


```
HTTP/1.1 200 OK
Date: Tue, 20 Apr 2010 17:51:00 GMT
Server: Apache
Content-Datetime: Wed, 01 Jul 2009 00:00:00 GMT
Link:
 <http://mementoarchive.lanl.gov/dbpedia/timegate/http://dbpedia.org/page/
  France.html>;rel="timegate",
 <http://dbpedia.org/page/France.html>;rel="original",
 <http://mementoarchive.lanl.gov/dbpedia/memento/20070901/http://dbpedia.org/page/
  France.html>;rel="first-memento";datetime="Sat, 01 Sep 2007 00:00:00 GMT",
 <http://mementoarchive.lanl.gov/dbpedia/memento/20091101/http://dbpedia.org/page/
  France.html>;rel="last-memento next-memento";datetime="Sun, 01 Nov 2009 00:00:00
  GMT",
 <http://mementoarchive.lanl.gov/dbpedia/memento/20081101/http://dbpedia.org/page/
  France.html>;rel="prev-memento";datetime="Sat, 01 Nov 2008 00:00:00
 <http://mementoarchive.lanl.gov/dbpedia/timebundle/http://dbpedia.org/page/
  France.html>;rel="timebundle"
Connection: close
Transfer-Encoding: chunked
Content-Type: text/html; charset=UTF-8
```