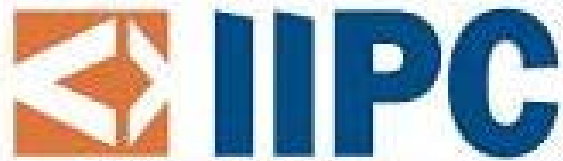# Los Alamos National Laboratory
Research Library

Herbert Van de Sompel
Robert Sanderson
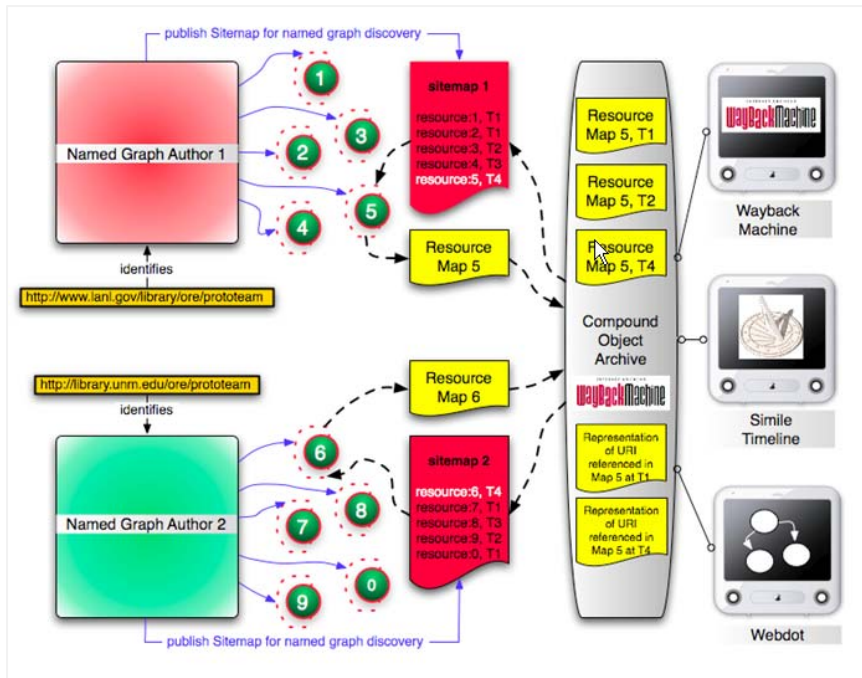
New IIPC Member:
Los Alamos National Laboratory, Research Library

Well, almost …

Digital Preservation Interest Exemplified

http://www.ctwatch.org/quarterly/multimedia/11/ORE_prototype-demo/

- aDORe repository for long term storage & access to the Research Library literature collection

- mod_oai plugin for Apache servers (Tools for a Preservation Ready Web – With Old Dominion U.)

- Increased obsession with (lack of) Time and the Web, cf. OAI-ORE Aggregations

- Memento "Time Travel for the Web"

Los Alamos
NATIONAL LABORATORY

Memento wants to make it easy
to access to Web of the Past

Memento is funded by
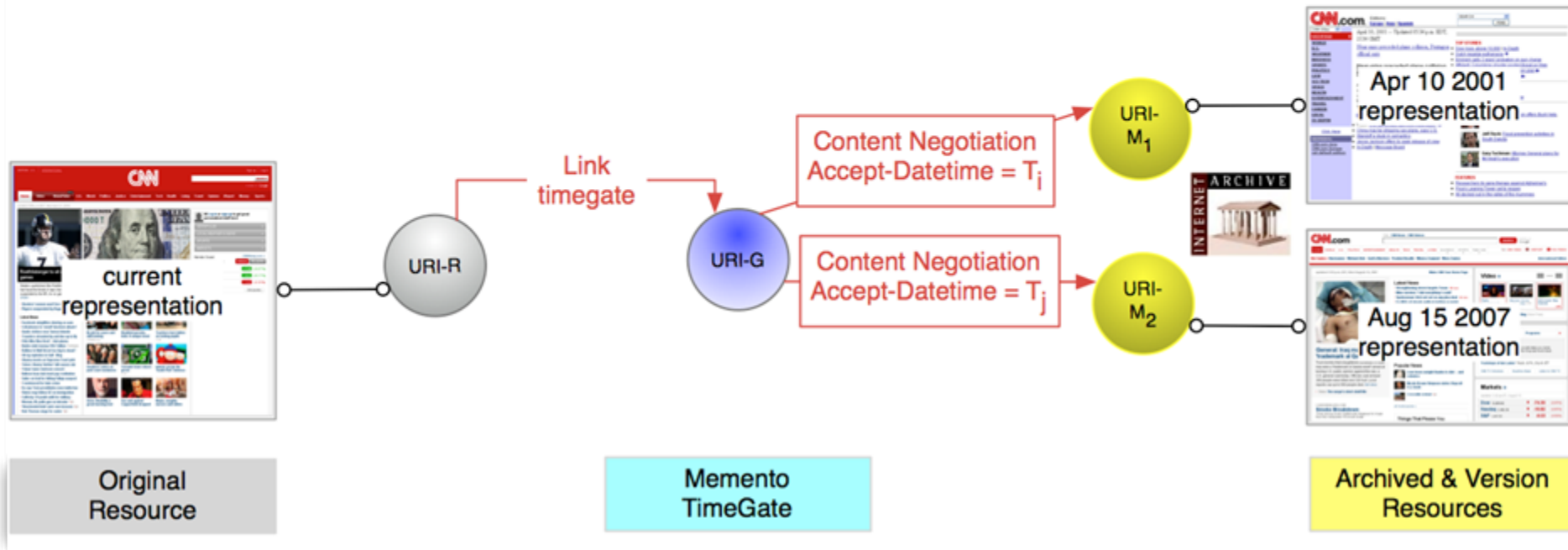The Library of Congress

Digital Preservation Award
2010

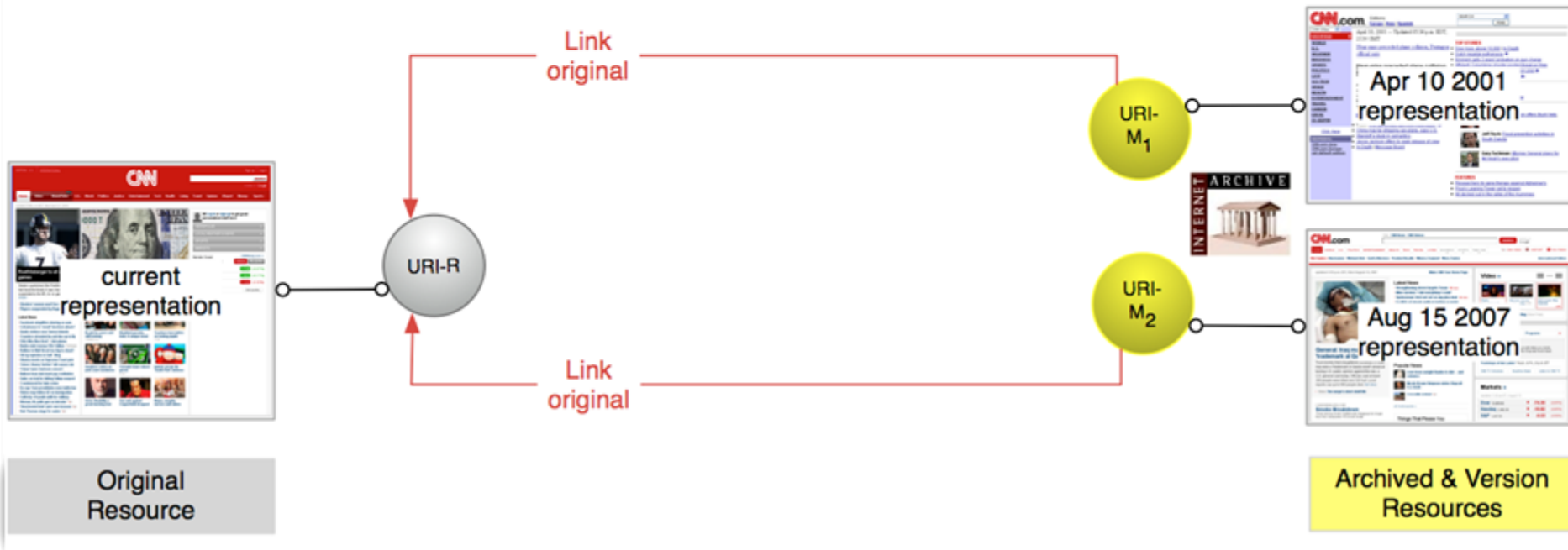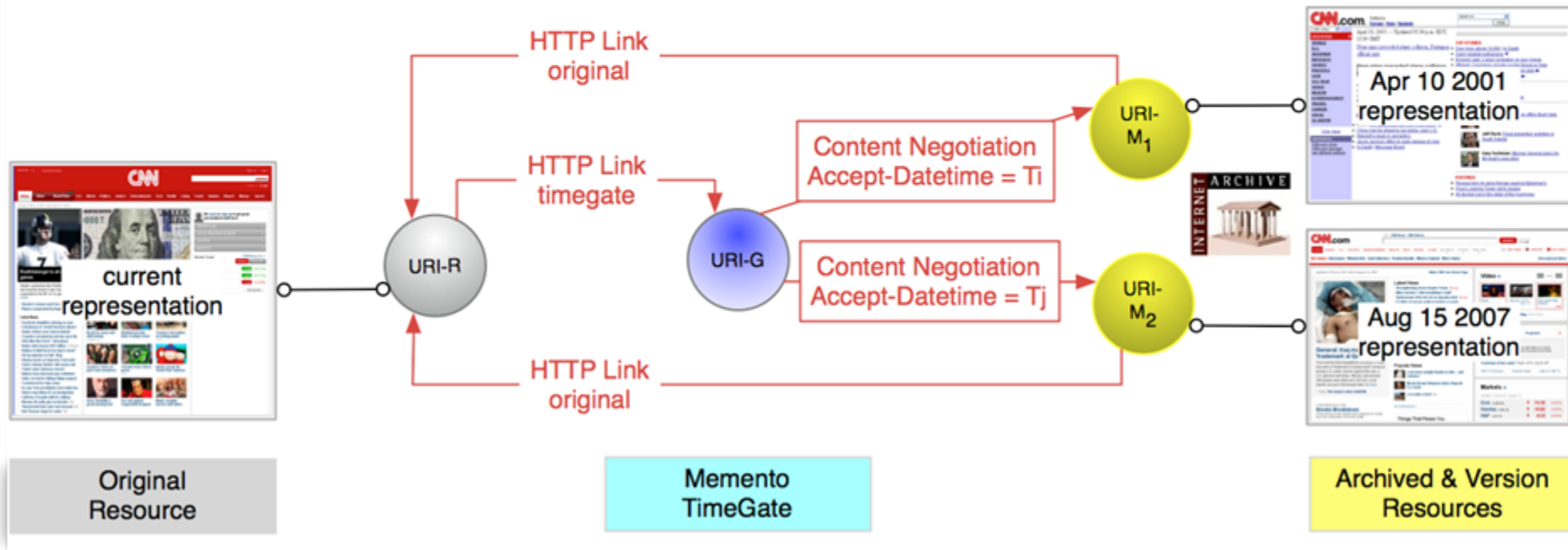# Original Resources and Mementos

Original Resource

URI-R

current representation

URI-M₁

Apr 10 2001 representation

URI-M₂

Aug 15 2007 representation

Archived & Version Resources

Los Alamos
NATIONAL LABORATORY

Bridge from Present to Past

Original Resource — Memento TimeGate — Archived & Version Resources

- URI-R — current representation (CNN)
- Link timegate → URI-G
- Content Negotiation Accept-Datetime $= T_i$ → URI-M_1 → Apr 10 2001 representation
- Content Negotiation Accept-Datetime $= T_j$ → URI-M_2 → Aug 15 2007 representation
- ARCHIVE INTERNET

Bridge from Past to Present

Memento Framework

Memento Progress

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.

Memento Protocol Information

The archive is accessible via the Memento "Time Travel for the Web" protocol. Hence, it supports datetime negotiation to access a version of an RDF description as it existed at a specific moment in time. Check out the brief introduction to Memento if you are not familiar with it.

Presuming the Original Resource (current version) is http://dbpedia.org/data/Oaxaca, then the resources that are relevant for the Memento protocol are:

- TimeGate: http://dbpedia.mementodepot.org/timegate/document/http://dbpedia.org/data/Oaxaca
- Memento: http://dbpedia.mementodepot.org/memento/20080103/http://dbpedia.org/data/Oaxaca ; http://dbpedia.mementodepot.org/memento/20110701/http://dbpedia.org/data/Oaxaca ; etc.
- TimeMap: http://dbpedia.mementodepot.org/timemap/link/http://dbpedia.org/data/Oaxaca
- TimeBundle: http://dbpedia.mementodepot.org/timebundle/http://dbpedia.org/data/Oaxaca
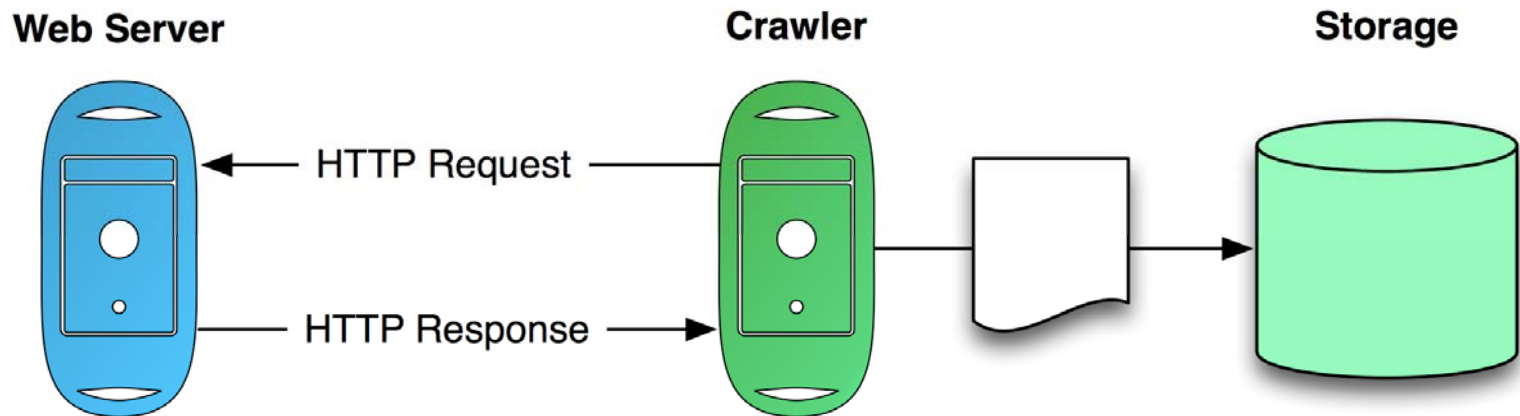
http://mementoweb.org/depot/

- Collaboration with IA on improved Memento integration in Wayback

- Collaboration with Old Dominion University on improved MementoFox FireFox add-on

- Linked Data archives: DBpedia & Live DBpedia

- Work with WikiPedia aimed at native Memento support

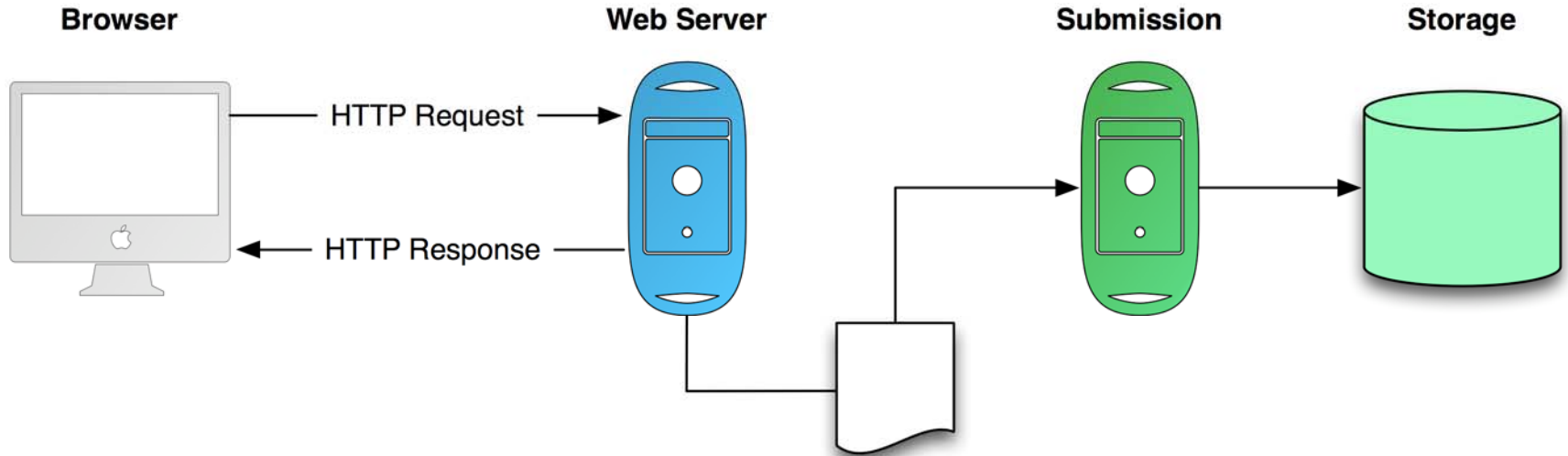- Collaboration with IIPC on Memento Aggregator

Transactional Web Archive:
Another Approach to
Archiving the Web

Los Alamos
NATIONAL LABORATORY

Crawl-Based Web Archiving

For example: Heritrix crawler for Internet Archive

Los Alamos
NATIONAL LABORATORY

Server-Side Transactional Web Archiving

For example: TTApache, PageVault, Vignette Web Capture

LANL Transactional Web Archive Software

- Open source release upcoming
  - 1+ year development
  - Tested in various environments
  - Memento compliant
  - Deduplication optimizations
  - Offload to WARC files

- Of interest to institutions that want to pro-actively archive their Web presence

- Of interest regarding archiving dynamic Web content (see Thursday's Workshop)