

Internet Engineering Task Force  
Internet-Draft  
Intended status: Informational  
Expires: June 22, 2012

H. VandeSompel  
Los Alamos National Laboratory  
M. Nelson  
Old Dominion University  
R. Sanderson  
Los Alamos National Laboratory  
December 20, 2011

**HTTP framework for time-based access to resource states -- Memento  
draft-vandesompel-memento-03**

**Abstract**

The HTTP-based Memento framework bridges the present and past Web by interlinking current resources with resources that encapsulate their past. It facilitates obtaining representations of prior states of a resource, available from archival resources in Web archives or version resources in content management systems, by leveraging the resource's URI and a preferred datetime. To this end, the framework introduces datetime negotiation (a variation on content negotiation), and new Relation Types for the HTTP "Link" header aimed at interlinking resources with their archival/version resources. It also introduces various discovery mechanisms that further support bridging the present and past Web.

**Status of this Memo**

This Internet-Draft is submitted in full conformance with the provisions of [BCP 78](#) and [BCP 79](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on June 22, 2012.

**Copyright Notice**

Copyright (c) 2011 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](#) and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

## Table of Contents

<a href="#">1.</a>	Introduction . . . . .	<a href="#">4</a>
<a href="#">1.1.</a>	Terminology . . . . .	<a href="#">4</a>
<a href="#">1.2.</a>	Purpose . . . . .	<a href="#">4</a>
<a href="#">1.3.</a>	Notational Conventions . . . . .	<a href="#">6</a>
2.	The Memento Framework, Datetime Negotiation component:	
	HTTP headers, HTTP Link Relation Types . . . . .	<a href="#">7</a>
<a href="#">2.1.</a>	HTTP Headers . . . . .	<a href="#">7</a>
<a href="#">2.1.1.</a>	Accept-Datetime, Memento-Datetime . . . . .	<a href="#">7</a>
<a href="#">2.1.1.1.</a>	Values for Accept-Datetime . . . . .	<a href="#">8</a>
<a href="#">2.1.1.2.</a>	Values for Memento-Datetime . . . . .	<a href="#">9</a>
<a href="#">2.1.2.</a>	Vary . . . . .	<a href="#">9</a>
<a href="#">2.1.3.</a>	Location . . . . .	<a href="#">10</a>
<a href="#">2.1.4.</a>	Link . . . . .	<a href="#">10</a>
<a href="#">2.2.</a>	Link Header Relation Types . . . . .	<a href="#">10</a>
<a href="#">2.2.1.</a>	Memento Framework Relation Types . . . . .	<a href="#">10</a>
<a href="#">2.2.1.1.</a>	Relation Type "original" . . . . .	<a href="#">11</a>
<a href="#">2.2.1.2.</a>	Relation Type "timegate" . . . . .	<a href="#">11</a>
<a href="#">2.2.1.3.</a>	Relation Type "timemap" . . . . .	<a href="#">12</a>
<a href="#">2.2.1.4.</a>	Relation Type "memento" . . . . .	<a href="#">12</a>
<a href="#">2.2.2.</a>	Other Relation Types . . . . .	<a href="#">14</a>
3.	The Memento Framework, Datetime Negotiation component:	
	HTTP Interactions . . . . .	<a href="#">15</a>
<a href="#">3.1.</a>	Interactions with an Original Resource . . . . .	<a href="#">16</a>
<a href="#">3.1.1.</a>	Step 1: User Agent Requests an Original Resource . . . . .	<a href="#">16</a>
3.1.2.	Step 2: Server Responds to a Request for an Original Resource . . . . .	<a href="#">17</a>
<a href="#">3.1.2.1.</a>	Original Resource is an Appropriate Memento . . . . .	<a href="#">18</a>
3.1.2.2.	Server Exists and Original Resource Used to Exist . . . . .	<a href="#">19</a>
3.1.2.3.	Missing or Inadequate "timegate" Link in Original Server's Response . . . . .	<a href="#">20</a>
<a href="#">3.2.</a>	Interactions with a TimeGate . . . . .	<a href="#">20</a>
<a href="#">3.2.1.</a>	Step 3: User Agent Negotiates with a TimeGate . . . . .	<a href="#">20</a>
3.2.2.	Step 4: Server Responds to Negotiation with TimeGate . . . . .	<a href="#">21</a>



<a href="#">3.2.2.1.</a>	Successful Scenario . . . . .	<a href="#">21</a>
3.2.2.2.	Accept-Datetime with Interval Indicator Provided . . . . .	<a href="#">23</a>
<a href="#">3.2.2.3.</a>	Multiple Matching Mementos . . . . .	<a href="#">24</a>
<a href="#">3.2.2.4.</a>	TimeGate Redirects to another TimeGate . . . . .	<a href="#">25</a>
3.2.2.5.	Accept-Datetime and other Accept Headers Provided . . . . .	<a href="#">26</a>
<a href="#">3.2.2.6.</a>	Accept-Datetime Unparseable . . . . .	<a href="#">27</a>
<a href="#">3.2.2.7.</a>	Accept-Datetime Not Provided . . . . .	<a href="#">27</a>
<a href="#">3.2.2.8.</a>	TimeGate Does Not Exist . . . . .	<a href="#">27</a>
<a href="#">3.2.2.9.</a>	HTTP Methods other than HEAD/GET . . . . .	<a href="#">27</a>
<a href="#">3.2.3.</a>	Recognizing a TimeGate . . . . .	<a href="#">28</a>
<a href="#">3.3.</a>	Interactions with a Memento . . . . .	<a href="#">29</a>
<a href="#">3.3.1.</a>	Step 5: User Agent Requests a Memento . . . . .	<a href="#">29</a>
<a href="#">3.3.2.</a>	Step 6: Server Responds to a Request for a Memento . . . . .	<a href="#">29</a>
<a href="#">3.3.2.1.</a>	Common Scenario . . . . .	<a href="#">29</a>
<a href="#">3.3.2.2.</a>	Memento of a 3XX Response . . . . .	<a href="#">31</a>
3.3.2.3.	Memento of Responses with Other HTTP Status Codes . . . . .	<a href="#">33</a>
<a href="#">3.3.2.4.</a>	Mementos Without a TimeGate . . . . .	<a href="#">34</a>
<a href="#">3.3.2.5.</a>	Memento Does not Exist . . . . .	<a href="#">35</a>
<a href="#">3.3.3.</a>	Recognizing a Memento . . . . .	<a href="#">36</a>
<a href="#">3.4.</a>	Interactions with a TimeMap . . . . .	<a href="#">36</a>
<a href="#">3.4.1.</a>	User Agent Requests a TimeMap . . . . .	<a href="#">37</a>
<a href="#">3.4.2.</a>	Server Responds to a Request for a TimeMap . . . . .	<a href="#">37</a>
<a href="#">4.</a>	The Memento Framework, Discovery Component . . . . .	<a href="#">39</a>
<a href="#">4.1.</a>	Discovering TimeGates Via Robots Exclusion Protocol . . . . .	<a href="#">39</a>
<a href="#">4.2.</a>	Discovering Mementos via Robots Exclusion Protocol . . . . .	<a href="#">41</a>
<a href="#">5.</a>	IANA Considerations . . . . .	<a href="#">41</a>
<a href="#">6.</a>	Security Considerations . . . . .	<a href="#">41</a>
<a href="#">7.</a>	Changelog . . . . .	<a href="#">42</a>
<a href="#">8.</a>	Acknowledgements . . . . .	<a href="#">43</a>
<a href="#">9.</a>	References . . . . .	<a href="#">43</a>
<a href="#">9.1.</a>	Normative References . . . . .	<a href="#">43</a>
<a href="#">9.2.</a>	Informative References . . . . .	<a href="#">44</a>
<a href="#">Appendix A.</a>	<a href="#">Appendix B: A Sample, Successful Memento     Request/Response cycle . . . . .</a>	<a href="#">44</a>
Authors' Addresses . . . . .		<a href="#">46</a>



## **1. Introduction**

### **1.1. Terminology**

This specification uses the terms "resource", "request", "response", "entity", "entity-body", "entity-header", "content negotiation", "client", "user agent", "server" as described in [RFC 2616](#) [[RFC2616](#)], and it uses the terms "representation" and "resource state" as described in W3C.REC-aww-20041215 [[W3C.REC-aww-20041215](#)].

In addition, the following terms specific to the Memento framework are introduced:

- o Original Resource: An Original Resource is a resource that exists or used to exist, and for which access to one of its prior states is desired.
- o Memento: A Memento for an Original Resource is a resource that encapsulates a prior state of the Original Resource. A Memento for an Original Resource as it existed at time  $T_j$  is a resource that encapsulates the state that the Original Resource had at time  $T_j$ .
- o TimeGate: A TimeGate for an Original Resource is a resource that is capable of negotiation to allow selective, datetime-based, access to prior states of the Original Resource.
- o TimeMap: A TimeMap for an Original Resource is a resource from which a list of URIs of Mementos of the Original Resource is available.

### **1.2. Purpose**

The state of an Original Resource may change over time. Dereferencing its URI at any specific moment in time during its existence yields a representation of its then current state. Dereferencing its URI at any time past its existence no longer yields a meaningful representation, if any. Still, in both cases, resources may exist that encapsulate prior states of the Original Resource. Each such resource, named a Memento, has its own URI that, when dereferenced, returns a representation of a prior state of the Original Resource. Mementos may, for example, exist in Web archives, Content Management Systems, or Revision Control Systems.

Examples are:

Mementos for Original Resource <http://www.ietf.org/> :



- o <http://web.archive.org/web/19970107171109/http://www.ietf.org/>
- o <http://webarchive.nationalarchives.gov.uk/20080906200044/http://www.ietf.org/>

Mementos for Original Resource

[http://en.wikipedia.org/wiki/Hypertext\\_Transfer\\_Protocol](http://en.wikipedia.org/wiki/Hypertext_Transfer_Protocol) :

- o [http://en.wikipedia.org/w/index.php?title=Hypertext\\_Transfer\\_Protocol&oldid=366806574](http://en.wikipedia.org/w/index.php?title=Hypertext_Transfer_Protocol&oldid=366806574)
- o [http://en.wikipedia.org/w/index.php?title=Hypertext\\_Transfer\\_Protocol&oldid=33912](http://en.wikipedia.org/w/index.php?title=Hypertext_Transfer_Protocol&oldid=33912)
- o [http://web.archive.org/web/20071011153017/http://en.wikipedia.org/wiki/Hypertext\\_Transfer\\_Protocol](http://web.archive.org/web/20071011153017/http://en.wikipedia.org/wiki/Hypertext_Transfer_Protocol)

Mementos for Original Resource <http://www.w3.org/TR/webarch/> :

- o <http://www.w3.org/TR/2004/PR-webarch-20041105/>
- o <http://www.w3.org/TR/2002/WD-webarch-20020830/>
- o <http://webarchive.nationalarchives.gov.uk/20100304163140/http://www.w3.org/TR/webarch/>

In the abstract, Memento introduces a mechanism to access versions of Web resources that:

- o Is fully distributed in the sense that resource versions may reside on multiple hosts, and that any such host is likely only aware of the versions it holds;
- o Uses the global notion of datetime as a resource version indicator and access key;
- o Leverages the following primitives of W3C.REC-aww-20041215 [[W3C.REC-aww-20041215](#)]: resource, resource state, representation, content negotiation, and link.

The core components of Memento's mechanism to access resource versions are:

1. The abstract notion of the state of a resource identified by URI-R as it existed at some time Tj. Note the relationship with the ability to identify a the state of a resource at some datetime Tj by means of a URI as intended by the proposed Dated URI scheme I-D.masinter-dated-uri [[I-D.masinter-dated-uri](#)].



2. A bridge from the present to the past, consisting of:
  - o An appropriately typed link from a resource identified by URI-R to an associated TimeGate identified by URI-G, which is aware of (at least part of the) version history of the resource identified by URI-R;
  - o The ability to content negotiate in the datetime dimension with the TimeGate identified by URI-G, as a means to obtain a representation of the state that the resource identified by URI-R had at some datetime Tj.
3. A bridge from the past to the present, consisting of an appropriately typed link from a resource identified by URI-M, which encapsulates the state a resource identified by URI-R had at some datetime Tj, to the resource identified by URI-R.

[Section 2](#) and [Section 3](#) of this document are concerned with specifying an instantiation of these abstractions for resources that are identified by HTTP(S) URIs, whereas [Section 4](#) details approaches to discover TimeGates, TimeMaps, and Mementos on the HTTP(S) Web by other means than typed links.

### **[1.3.](#) Notational Conventions**

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC 2119](#) [[RFC2119](#)].

When needed for extra clarity, the following conventions are used:

- o URI-R is used to denote the URI of an Original Resource.
- o URI-G is used to denote the URI of a TimeGate.
- o URI-M is used to denote the URI of a Memento.
- o URI-T is used to denote the URI of a TimeMap.
- o When scenarios are described that involve multiple Mementos, URI-M<sub>0</sub> denotes the URI of the first Memento known to the responding server, URI-M<sub>n</sub> denotes the URI of the most recent known Memento, URI-M<sub>j</sub> denotes the URI of the selected Memento, URI-M<sub>i</sub> denotes the URI of the Memento that is temporally previous to the selected Memento, and URI-M<sub>k</sub> denotes the URI of the Memento that is temporally after the selected Memento. The respective datetimes for these Mementos are T<sub>0</sub>, T<sub>n</sub>, T<sub>j</sub>, T<sub>i</sub>, and T<sub>k</sub>; it holds that T<sub>0</sub> ≤ T<sub>i</sub> ≤ T<sub>j</sub> ≤ T<sub>k</sub> ≤ T<sub>n</sub>.



## **2. The Memento Framework, Datetime Negotiation component: HTTP headers, HTTP Link Relation Types**

The Memento framework is concerned with Original Resources, TimeGates, Mementos, and TimeMaps that are identified by HTTP or HTTPS URIs. Details are only provided for resources identified by HTTP URIs but apply similarly to those with HTTPS URIs.

### **2.1. HTTP Headers**

The Memento framework operates at the level of HTTP request and response headers. It introduces two new headers ("Accept-Datetime", "Memento-Datetime"), introduces new values for two existing headers ("Vary", "Link"), and uses an existing header ("Location") without modification. All these headers are described below. Other HTTP headers are present or absent in Memento response/request cycles as specified by [RFC 2616](#) [[RFC2616](#)].

#### **2.1.1. Accept-Datetime, Memento-Datetime**

The "Accept-Datetime" request header is used by a user agent to indicate it wants to retrieve a representation of a Memento that encapsulates a past state of an Original Resource. To that end, the "Accept-Datetime" header is conveyed in an HTTP GET/HEAD request issued against a TimeGate for an Original Resource, and its value indicates the datetime of the desired past state of the Original Resource. The "Accept-Datetime" request header has no defined meaning for HTTP methods other than HEAD and GET.

The "Memento-Datetime" response header is used by a server to indicate that the response contains a representation of a Memento, and its value expresses the datetime of the state of an Original Resource that is encapsulated in that Memento. The URI of that Original Resource is provided in the response, as the Target IRI (see [RFC5988](#) [[RFC5988](#)]) of a link provided in the HTTP "Link" header that has a Relation Type of "original" (see [Section 2.2](#)).

The presence of a "Memento-Datetime" header and associated value for a given resource constitutes a promise that the resource is stable and that its state will no longer change. This means that, in terms of the Ontology for Relating Generic and Specific Information Resources (see W3C.gen-ont-20090420 [[W3C.gen-ont-20090420](#)]), a Memento is a FixedResource.

As a consequence, "Memento-Datetime" headers associated with a Memento MUST be "sticky" in the following ways:



- o The server that originally assigns the "Memento-Datetime" header and value MUST retain that header in all responses to HTTP HEAD/GET requests (with or without "Accept-Datetime" header) that occur against the Memento after the time of the original assignment of the header, and it MUST NOT change its associated value.
- o Applications that mirror Mementos at a different URI MUST NOT change the "Memento-Datetime" header and value of those Mementos unless mirroring involves a meaningful state change. This allows, for example, duplicating a Web archive at a new location while preserving the value of the "Memento-Datetime" header of the archived resources. In this example, the "Last-Modified" header will be updated to reflect the time of mirroring at the new URI, whereas the value for "Memento-Datetime" will be sticky.

#### **2.1.1.1. Values for Accept-Datetime**

Values for the "Accept-Datetime" header consist of a MANDATORY datetime expressed according to the [RFC 1123](#) [RFC1123] format, which is formalized by the [rfc1123-date](#) construction rule of the BNF in Figure 1, and an OPTIONAL interval indicator expressed according to the [iso8601-interval](#) rule of the BNF in Figure 1. The datetime MUST be represented in Greenwich Mean Time (GMT).

Examples of "Accept-Datetime" request headers with and without an interval indicator:

Accept-Datetime: Thu, 31 May 2007 20:35:00 GMT

Accept-Datetime: Thu, 31 May 2007 20:35:00 GMT; -P3DT5H;+P2DT6H

The user agent uses the MANDATORY datetime value to convey its preferred datetime for a Memento; it uses the OPTIONAL interval indicator to convey it is interested in retrieving Mementos that reside within this interval around the preferred datetime, and not interested in Mementos that reside outside of it. Not using an interval indicator is equivalent to expressing an infinite interval around the preferred datetime.

The interval mechanism can be regarded as an implementation of the functionality intended by the q-value approach that is used in regular content negotiation. The q-value approach is not supported for Memento's datetime negotiation because it is well-suited for negotiation over a discrete space of mostly predictable values, not for negotiation over a continuum of unpredictable datetime values.



```

accept-dt-value = rfc1123-date *SP [ iso8601-interval ]
rfc1123-date = wkday "," SP date1 SP time SP "GMT"
date1          = 2DIGIT SP month SP 4DIGIT
                ; day month year (e.g., 20 Mar 1957)
time           = 2DIGIT ":" 2DIGIT ":" 2DIGIT
                ; 00:00:00 - 23:59:59 (e.g., 14:33:22)
wkday          = "Mon" | "Tue" | "Wed" | "Thu" | "Fri" | "Sat" |
                "Sun"
month          = "Jan" | "Feb" | "Mar" | "Apr" | "May" | "Jun" |
                "Jul" | "Aug" | "Sep" | "Oct" | "Nov" | "Dec"
iso8601-interval = ";" *SP "-" duration *SP ";" *SP "+" duration
duration = "P" ( dur-date | dur-week )
dur-date = ( dur-day | dur-month | dur-year ) [ dur-time ]
dur-year = 1*DIGIT "Y" [ dur-month ] [ dur-day ]
dur-month = 1*DIGIT "M" [ dur-day ]
dur-day = 1*DIGIT "D"
dur-time = "T" ( dur-hour | dur-minute | dur-second )
dur-hour = 1*DIGIT "H" [ dur-minute ] [ dur-second ]
dur-minute = 1*DIGIT "M" [ dur-second ]
dur-second = 1*DIGIT "S"
dur-week = 1*DIGIT "W"

```

Figure 1: BNF for the datetime format

#### [2.1.1.2.](#) Values for Memento-Datetime

Values for the "Memento-Datetime" headers MUST be datetimes expressed according to the [rfc1123](#)-date construction rule of the BNF in Figure 1; they MUST be represented in Greenwich Mean Time (GMT).

An example "Memento-Datetime" response header:

```
Memento-Datetime: Wed, 30 May 2007 18:47:52 GMT
```

#### [2.1.2.](#) Vary

The "Vary" response header is used in responses to indicate the dimensions in which content negotiation was successfully applied. This header is used in the Memento framework to indicate both whether datetime negotiation was applied or is supported by the responding server.

For example, this use of the "Vary" header indicates that datetime is the only dimension in which negotiation was applied:

```
Vary: negotiate, accept-datetime
```



The use of the "Vary" header in this example shows that both datetime negotiation, and media type content negotiation were applied:

Vary: negotiate, accept-datetime, accept

### **2.1.3. Location**

The "Location" header is used as defined in [RFC 2616](#) [[RFC2616](#)]. Examples are given in [Section 3](#) below.

### **2.1.4. Link**

The "Link" response header is specified in [RFC5988](#) [[RFC5988](#)]. The Memento framework introduces new Relation Types to convey typed links among Original Resources, TimeGates, Mementos, and TimeMaps. Already existing Relation Types, among others, aimed at supporting navigation among a series of ordered resources may also be used in the Memento framework. This is detailed in Link Header Relation Types ([Section 2.2](#)), below.

## **2.2. Link Header Relation Types**

The "Link" header specified in [RFC5988](#) [[RFC5988](#)] is semantically equivalent to the "<LINK>" element in HTML, as well as the "atom:link" feed-level element in Atom [RFC 4287](#) [[RFC4287](#)]. By default, the origin of a link expressed by an entry in a "Link" header (named Context IRI in [RFC5988](#) [[RFC5988](#)]) is the IRI of the requested resource. This default can be overwritten using the "anchor" attribute in the entry.

### **2.2.1. Memento Framework Relation Types**

The Relation Types used in the Memento framework are listed in the remainder of this section, and their use is summarized in the below table. [Appendix A](#) shows a Memento request/response cycle that uses all the Relation Types that are introduced here.

Relation Type	Original Resource	TimeGate	Memento
original	NA, except see <a href="#">Section 3.1.2.1</a>	REQUIRED, 1	REQUIRED, 1
timegate	RECOMMENDED, 0 or more	REQUIRED, 1 in case of <a href="#">Section 3.2.2.4</a>	RECOMMENDED, 0 or more
timemap	NA	RECOMMENDED, 0 or more	RECOMMENDED, 0 or more



memento	NA, except see	REQUIRED, 1 or more	REQUIRED, 1
	<a href="#">Section 3.1.2.1</a>		or more
+-----+	+-----+	+-----+	+-----+

Table 1: The use of Relation Types

#### **2.2.1.1. Relation Type "original"**

"original" -- A "Link" header entry with a Relation Type of "original" is used to point from a TimeGate or a Memento to their associated Original Resource. In both cases, an entry with the "original" Relation Type MUST occur exactly once in a "Link" header. Details for the entry are as follows:

- o Context IRI: URI-G, URI-M
- o Target IRI: URI-R
- o Relation Type: "original"
- o Use: REQUIRED
- o Cardinality: 1

#### **2.2.1.2. Relation Type "timegate"**

"timegate" -- A "Link" header entry with a Relation Type of "timegate" is used to point both from an Original Resource or a Memento to a TimeGate for the Original Resource. In both cases, the use of an entry with the "timegate" Relation Type is RECOMMENDED. Since more than one TimeGate can exist for any Original Resource, multiple entries with a "timegate" Relation Type MAY occur, each with a distinct Target IRI. Since a TimeGate has no mime type, the "type" attribute MUST NOT be used on Links with a "timegate" Relation Type. Details for the entry are as follows:

- o Context IRI: URI-R or URI-Mj
- o Target IRI: URI-G
- o Relation Type: "timegate"
- o Use: RECOMMENDED
- o Cardinality: 0 or more

In the special case (see [Section 3.2.2.4](#)) where a TimeGate redirects to another TimeGate for the Original Resource, a "Link" header entry



with a Relation Type of "timegate" MUST be used to point from the former to the latter.

#### **2.2.1.3. Relation Type "timemap"**

"timemap" -- A "Link" header entry with a Relation Type of "timemap" is used to point from both a TimeGate or a Memento to a TimeMap resource from which a list of Mementos known to the responding server is available. Use of an entry with the "timemap" Relation Type is RECOMMENDED, and, since multiple serializations of a TimeMap are possible, multiple entries with a "timemap" Relation Type MAY occur, each with a distinct Target IRI, and each with a MANDATORY "type" attribute to convey the mime type of the TimeMap serialization. Details for the entry are as follows:

- o Context IRI: URI-G or URI-Mi
- o Target IRI: URI-T
- o Relation Type: "timemap"
- o Target Attribute: "type"
- o Use: RECOMMENDED
- o Cardinality: 0 or more

Further details about TimeMap serializations are provided in [Section 3.4](#).

#### **2.2.1.4. Relation Type "memento"**

"memento" -- A "Link" header entry with a Relation Type of "memento" is used to point from both a TimeGate and a Memento to various Mementos for an Original Resource. This link MUST include a "datetime" attribute with a value that matches the "Memento-Datetime" of the Memento that is the target of the link; that is, the value of the "Memento-Datetime" header that is returned when the URI of the linked Memento is dereferenced. In addition, the link MAY include an "embargo" attribute to convey the datetime until which the Memento will remain inaccessible. The value for both the "datetime" and "embargo" attributes MUST be a datetime expressed according to the [rfc1123](#)-date construction rule of the BNF in Figure 1 and it MUST be represented in Greenwich Mean Time (GMT). This link MAY also include a "license" attribute to associate a license with the Memento; the value for the "license" attribute SHOULD be a URI. The link SHOULD also include a "type" attribute to convey the mime type of the Memento that is the target of the link. Use of entries with the



"memento" Relation Type is REQUIRED and it MUST be as follows:

For all responses to HTTP HEAD/GET requests issued against a TimeGate or a Memento in which a Memento is selected or served by the responding server:

- o One "memento" link MUST be included that has as Target IRI the URI of the Memento that was selected or served;
- o One "memento" link MUST be included that has as Target IRI the URI of the temporally first Memento known to the responding server;
- o One "memento" link MUST be included that has as Target IRI the URI of the temporally most recent Memento known to the responding server.
- o One "memento" link SHOULD be included that has as Target IRI the URI of the Memento that is previous to the selected Memento in the temporal series of all Mementos (sorted by ascending "Memento-Datetime" values) known to the server;
- o One "memento" link SHOULD be included that has as Target IRI the URI the Memento that is next to the selected Memento in the temporal series of all Mementos (sorted by ascending "Memento-Datetime" values) known to the server.
- o Other "memento" links MAY only be included if both the aforementioned previous and next links are provided. Each of these OPTIONAL "memento" links MUST have as Target IRI the URI of a Memento other than the ones listed above.

For all responses to HTTP HEAD/GET requests issued against an existing TimeGate or Memento in which no Memento is selected or served by the responding server:

- o One "memento" link MUST be included that has as Target IRI the URI of the temporally first Memento known to the responding server;
- o One "memento" link MUST be included that has as Target IRI the URI of the temporally most recent Memento known to the responding server.
- o Other "memento" links MAY be included, and each of these OPTIONAL links MUST have as Target IRI the URI of a Memento other than the two listed above.

Note that the Target IRI of some of these links may coincide. For example, if the selected Memento actually is the first Memento known



to the server, only three distinct "memento" links may result. The value for the "datetime" attribute of these links would be the datetimes of the first (equal to selected), next, and most recent Memento known to the responding server.

The summary is as follows:

- o Context IRI: URI-G, URI-Mj
- o Target IRI: URI-M
- o Relation Type: "memento"
- o Target Attributes: "datetime", "embargo", "license"
- o Use: REQUIRED
- o Cardinality: 1 or more

#### **2.2.2. Other Relation Types**

Web Linking [RFC5988](#) [[RFC5988](#)] allows for the inclusion of links with different Relation Types but the same Target IRI, and hence the Relation Types introduced by the Memento framework MAY be combined with others as deemed necessary. As the "memento" Relation Type focuses on conveying the datetime of a linked Memento, Relation Types that allow navigating among the temporally ordered series of Mementos known to a server are of particular importance. With this regard, the Relation Types listed in the below table SHOULD be considered for combination with the "memento" Relation Type. A distinction is made between responding servers that can be categorized as systems that are the focus of [RFC5829](#) [[RFC5829](#)] (such as version control systems) and others that can not (such as Web archives). Note that, in terms of [RFC5829](#) [[RFC5829](#)], the last Memento (URI-Mn) is the version prior to the latest (i.e. current) version.

Memento Type	<a href="#">RFC5988</a> system	non <a href="#">RFC5988</a> system
First Memento (URI-M0)	first	first
Last Memento (URI-Mn)	last	last
Selected Memento (URI-Mj)	NA	NA
Memento prior to selected Memento (URI-Mi)	predecessor-version	prev
Memento next to selected Memento (URI-Mk)	successor-version	next



Table 2: The use of Relation Types

### 3. The Memento Framework, Datetime Negotiation component: HTTP Interactions

This section describes the HTTP interactions of the Memento framework for a variety of scenarios. First, Figure 2 provides a schematic overview of a successful request/response chain that involves datetime negotiation. Dashed lines depict HTTP transactions between user agent and server. [Appendix A](#) shows these HTTP interactions in detail for the case where the Original Resource resides on one server, whereas both the TimeGate and the Mementos reside on another. Scenarios also exist in which all these resources are on the same server (for example, Content Management Systems) or on different servers (for example, an aggregator of TimeGates). Note that, in Step 2 and Step 6, the HTTP status code of the response is shown as "200 OK", but a series of "206 Partial Content" responses could be substituted without loss of generality.

```

1: UA --- HTTP GET/HEAD; Accept-Datetime: Tj -----> URI-R
2: UA <-- HTTP 200; Link: URI-G ----- URI-R
3: UA --- HTTP GET/HEAD; Accept-Datetime: Tj -----> URI-G
4: UA <-- HTTP 302; Location: URI-Mj; Vary; Link:
    URI-R,URI-T,URI-M0,URI-Mn,URI-Mi,URI-Mj,URI-Mk ----- URI-G
5: UA --- HTTP GET URI-Mj; Accept-Datetime: Tj -----> URI-Mj
6: UA <-- HTTP 200; Memento-Datetime: Tj; Link:
    URI-R,URI-T,URI-G,URI-M0,URI-Mn,URI-Mi,URI-Mj,URI-Mk -- URI-Mj

```

Figure 2: Typical Memento request/response chain

- o Step 1: In order to determine what the URI is of a TimeGate for an Original Resource, the user agent issues an HTTP HEAD/GET request against the URI of the Original Resource (URI-R).
- o Step 2: The entity-header of the response from URI-R includes an HTTP "Link" header with a Relation Type of "timegate" pointing at a TimeGate (URI-G) for the Original Resource.
- o Step 3: The user agent starts the datetime negotiation process with the TimeGate by issuing an HTTP GET request against its URI-G thereby including an "Accept-Datetime" HTTP header with a value of the datetime of the desired prior state of the Original Resource.
- o Step 4: The entity-header of the response from URI-G includes a "Location" header pointing at the URI of a Memento (URI-Mj) for the Original Resource. In addition, the entity-header contains an HTTP "Link" header with a Relation Type of "original" pointing at



the Original Resource, and an HTTP "Link" header with a Relation Type of "timemap" pointing at a TimeMap (URI-T). Also HTTP Links pointing at various Mementos are provided using the "memento" Relation Type, as specified in [Section 2.2.1.4](#).

- o Step 5: The user agent issues an HTTP GET request against the URI-Mj of a Memento, obtained in Step 4.
- o Step 6: The entity-header of the response from URI-Mj includes a "Memento-Datetime" HTTP header with a value of the datetime of the Memento. It also contains an HTTP "Link" header with a Relation Type of "original" pointing at the Original Resource, with a Relation Type of "timegate" pointing at a TimeGate associated with the Original Resource, and with a Relation Type of "timemap" pointing at a TimeMap. The state that is expressed by the representation provided in the response is the state the Original Resource had at the datetime expressed in the "Memento-Datetime" header. This response also includes HTTP Links with a "memento" Relation Type pointing at various Mementos, as specified in [Section 2.2.1.4](#).

The following sections detail the specifics of HTTP interactions with Original Resources, TimeGates, Mementos, and TimeMaps under various conditions.

### **[3.1.](#) Interactions with an Original Resource**

This section details HTTP GET/HEAD requests targeted at an Original Resource (URI-R).

#### **[3.1.1.](#) Step 1: User Agent Requests an Original Resource**

In order to try and discover a TimeGate for the Original Resource, the user agent SHOULD issue an HTTP HEAD or GET request against the Original Resource's URI. Use of the "Accept-Datetime" header in the HTTP HEAD/GET request is OPTIONAL.

Figure 3 shows the use of HTTP HEAD indicating the user agent is not interested in retrieving a representation of the Original Resource, but only in determining a TimeGate for it. It also shows the use of the "Accept-Datetime" header anticipating that the user agent will set it for the entire duration of a Memento request/response cycle.

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



Figure 3: User Agent Requests Original Resource

### **3.1.2. Step 2: Server Responds to a Request for an Original Resource**

The response of the Original Resource's server to the user agent's HTTP HEAD/GET request of Step 1, for the case where the Original Resource exists, is as it would be in a regular HTTP request/response cycle, but in addition MAY include a HTTP "Link" header with a Relation Type of "timegate" that conveys the URI of the Original Resource's TimeGate as the Target IRI of the Link. Multiple HTTP Links with a relation type of "timegate" MAY be provided to accommodate situations in which the server is aware of multiple TimeGates for an Original Resource. The actual Target IRI provided in the "timegate" Link may depend on several factors including the datetime provided in the "Accept-Datetime" header, and the IP address of the user agent. A response for this case is illustrated in Figure 4.

```
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
```

Figure 4: Server of Original Resource Responds

Servers that actively maintain archives of their resources SHOULD include the "timegate" HTTP "Link" header because this link is an important way for a user agent to discover TimeGates for those resources. This includes servers such as Content Management Systems, Control Version Systems, and Web servers with associated transactional archives Fitch [[Fitch](#)]. Servers that do not actively maintain archives of their resources MAY include the "timegate" HTTP "Link" header as a way to convey a preference for TimeGates for their resources exposed by a third party archive. This includes servers that rely on Web archives such as the Internet Archive to archive their resources.

The server of the Original Resource MUST treat requests with and without an "Accept-Datetime" header in the same way:

- o The response MUST either always or never include a HTTP "Link" header with an entry that has a "timegate" Relation Type and the URI of a TimeGate as the Target IRI.



- o The entity-body of the response MUST be the same, for user agent requests with or without a "Accept-Datetime" header.

#### **3.1.2.1. Original Resource is an Appropriate Memento**

The "Memento-Datetime" header MAY be applied to an Original Resource directly to indicate it is a FixedResource (see W3C.gen-ont-20090420 [W3C.gen-ont-20090420]), meaning that the state of the Original Resource has not changed since the datetime conveyed in the "Memento-Datetime" header, and as a promise that it will not change anymore beyond it. This may occur, for example, for certain stable media resources on news sites. In case the user agent's preferred datetime is equal to or more recent than the datetime conveyed as the value of "Memento-Datetime" in the server's response in Step 2, the user agent SHOULD conclude it has located an appropriate Memento, and it SHOULD NOT continue to Step 3.

Figure 5 illustrates such a response to a request for the resource with URI `http://a.example.org/pic` that has been stable since it was created. Note the use of both the "memento" and "original" Relation Types for links that have as Target IRI the URI of the Original Resource.

```
HTTP/1.1 200 OK
Date: Thu, 21 Jan 2010 00:02:12 GMT
Server: Apache
Link:
  <http://a.example.org/pic>
    ; rel="original memento"
    ; datetime="Fri, 20 Mar 2009 11:00:00 GMT"
Memento-Datetime: Fri, 20 Mar 2009 11:00:00 GMT
Content-Length: 255
Connection: close
Content-Type: text/html; charset=iso-8909-1
```

Figure 5: Response to a request for an Original Resource that was created as a FixedResource

Cases may also exist in which a resource becomes stable at a certain point in its existence, but changed previously. In such cases, the Original Resource may know about a TimeGate that is aware of its prior history and hence MAY also include a link with a "timegate" Relation Type. This is illustrated in Figure 6, where the "memento" and "original" Relation Types are used as in Figure 5, and the existence of a TimeGate to negotiate for Mementos with datetimes prior to Fri, 20 Mar 2009 11:00:00 GMT is indicated.



```
HTTP/1.1 200 OK
Date: Thu, 21 Jan 2010 00:02:12 GMT
Server: Apache
Link:
  <http://a.example.org/pic>
  ; rel="original memento"
  ; datetime="Fri, 20 Mar 2009 11:00:00 GMT",
  <http://arxiv.example.net/timegate/http://a.example.org/pic>
  ; rel="timegate"
Memento-Datetime: Fri, 20 Mar 2009 11:00:00 GMT
Content-Length: 255
Connection: close
Content-Type: text/html; charset=iso-8909-1
```

Figure 6: Response to a request for an Original Resource that became a FixedResource

#### **3.1.2.2. Server Exists and Original Resource Used to Exist**

Servers SHOULD also provide a "timegate" HTTP "Link" header in responses to requests for an Original Resource that the server knows used to exist, but no longer does. This allows the use of an Original Resource's URI as an entry point to representations of its prior states even if the resource itself no longer exists. A server's response for this case is illustrated in Figure 7.

```
HTTP/1.1 404 Not Found
Date: Thu, 21 Jan 2010 00:02:12 GMT
Server: Apache
Link:
  <http://arxiv.example.net/timegate/http://a.example.org/gone>
  ; rel="timegate"
Content-Length: 255
Connection: close
Content-Type: text/html; charset=iso-8909-1
```

Figure 7: Response to a request for an Original Resource that not longer exists

In case the server is not aware of the prior existence of the Original Resource, its response SHOULD NOT include a "timegate" HTTP Link. [Section 3.1.2.3](#) details what the user agent's behavior should be in such cases.



### **3.1.2.3. Missing or Inadequate "timegate" Link in Original Server's Response**

A user agent MAY ignore the TimeGate returned in Step 2. However, when engaging in a Memento request/response cycle, a user agent SHOULD NOT proceed immediately to Step 3 by using a TimeGate of its own preference but rather SHOULD always start the cycle by issuing an HTTP GET/HEAD against the Original Resource (Step 1, Figure 3) as it is an important way to learn about dedicated or preferred TimeGates for the Original Resource. Also, cases exist in which the response in Step 2 will not provide a "timegate" link, including:

- o The Original Resource's server does not support the Memento framework;
- o The Original Resource no longer exists and the responding server is not aware of its prior existence;
- o The server that hosted the Original Resource no longer exists;

In all these cases, the user agent SHOULD attempt to determine an appropriate TimeGate for the Original Resource, either automatically or interactively supported by the user. The discovery mechanisms described in [Section 4](#) can support the user agent with this regard.

## **3.2. Interactions with a TimeGate**

This section details HTTP GET/HEAD requests targeted at a TimeGate (URI-G).

### **3.2.1. Step 3: User Agent Negotiates with a TimeGate**

In order to negotiate with a TimeGate, the user agent MUST issue a HTTP HEAD or GET against its URI, its request MUST include the "Accept-Datetime" header to express its datetime preference, and the use of that header MUST be as described in [Section 2.1.1.1](#). The URI of the TimeGate may have been provided as the Target IRI of a "timegate" HTTP "Link" header in the response from the Original Resource (Step 2, Figure 4), or may have resulted from another discovery mechanism (see [Section 4](#)) or user interaction. Such a request is illustrated in Figure 8.

```
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
```

Figure 8: User agent negotiates with TimeGate



### **3.2.2. Step 4: Server Responds to Negotiation with TimeGate**

In order to respond to a datetime negotiation request (Step 3, [Section 3.2.1](#)), the server uses an internal algorithm to select the Memento that best meets the user agent's datetime preference, and redirects to it. The exact nature of the selection algorithm is at the server's discretion but SHOULD be consistent. A variety of approaches can be used including selecting the Memento that is nearest in time (either past or future) or nearest in the past relative to the requested datetime. The common scenario for datetime negotiation with a TimeGate is described in [Section 3.2.2.1](#) but special cases exist, and they are addressed in [Section 3.2.2.2](#) through [Section 3.2.2.9](#).

#### **3.2.2.1. Successful Scenario**

In cases where the TimeGate exists, and the datetime provided in the user agent's "Accept-Datetime" header can be parsed and does not contain an interval indicator, the server selects a Memento based on the user agent's datetime preference. The response MUST have a "302 Found" HTTP status code, and the "Location" header MUST be used to convey the URI of the selected Memento. The "Vary" header MUST be provided and it MUST include the "negotiate" and "accept-datetime" values to indicate that datetime negotiation has taken place. The "Link" header MUST be provided and contain links with Relation Types subject to the considerations described in [Section 2.2](#). The response MUST NOT contain a "Memento-Datetime" header. Such a response is illustrated in Figure 9.



```
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/timemap/http://a.example.org>
    ; rel="timemap"; type="application/link-format",
  <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/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
```

Figure 9: Server of TimeGate responds

Note that if a user agent's "Accept-Datetime" header does not convey an interval indicator, and conveys a datetime that is either earlier than the datetime of the first Memento or later than the datetime of the most recent Memento known to the server, the server's response is as just described yet entails the selection of the first or most recent Memento, respectively. This approach is consistent with interpreting the absence of an interval indicator in the user agent's request as an indication of an infinite interval around its preferred datetime (see [Section 2.1.1.1](#)).

This is illustrated in Figure 10 that shows the response from a TimeGate exposed by a MediaWiki server to a request by a user agent that has an "Accept-Datetime: Mon, 31 May 1999 00:00:00 GMT" header. Note that a link is provided with a "successor-version" Relation Type but not with a "predecessor-version" Relation Type.



```
HTTP/1.1 302 Found
Server: Apache
Content-Length: 709
Content-Type: text/html; charset=utf-8
Date: Thu, 21 Jan 2010 00:09:40 GMT
Location:
  http://a.example.org/w/index.php?title=Clock&oldid=1493688
Vary: negotiate, accept-datetime
Link: <http://a.example.org/w/Clock>; rel="original",
      <http://a.example.org/Special:TimeMap/http://a.example.org/w/Clock>
      ; rel="timemap",
      <http://a.example.org/w/index.php?title=Clock&oldid=1493688>
      ; rel="first memento"; datetime="Sun, 28 Sep 2003 01:42:00 GMT",
      <http://a.example.org/w/index.php?title=Clock&oldid=1493854>
      ; rel="successor-version memento"
      ; datetime="Tue, 30 Sep 2003 14:28:00 GMT",
      <http://a.example.org/w/index.php?title=Clock&oldid=337446696>
      ; rel="last memento"; datetime="Tue, 12 Jan 2010 19:55:00 GMT"
Connection: close
```

Figure 10: A TimeGate's response to a request for a Memento with a datetime earlier than that of the first Memento

### **3.2.2.2. Accept-Datetime with Interval Indicator Provided**

In case, in Step 3, the datetime provided in the user agent's "Accept-Datetime" header can be parsed, and contains an interval indicator, the response depends on whether the server is or is not aware of Mementos with datetimes within the expressed interval. If the server is aware of such Mementos, the server's response MUST be as in [Section 3.2.2.1](#).

However, if the responding server is not aware of any Mementos with "Memento-Datetime" values within the expressed interval, the server's response MUST have a "406 Not Acceptable" HTTP status code. The use of the "Vary" header MUST be as described in [Section 3.2.2.1](#). The use of the "Link" header MUST be as described in [Section 2.2](#). Specifically, the use of links with a "memento" Relation Type MUST follow the rules for the case where no Memento is selected by the responding server ([Section 2.2.1.4](#)) and it is RECOMMENDED that the server provides "memento" links pointing at Mementos that have "Memento-Datetime" values in the temporal vicinity of the interval expressed by the client. The response MUST NOT contain a "Memento-Datetime" header.

As a result, a user agent that allows for the provision of an interval indicator in requests SHOULD anticipate possible "406 Not Acceptable" responses and provide the capability for their



resolution. For example, the client can leverage the "memento" links returned by the responding server, can extend its preferred interval, or can remove it from further requests.

Figure 11 shows a user agent using an "Accept-Datetime" header conveying an interval of interest starting 5 hours before and ending 6 hours after Tue, 11 Sep 2001 20:35:00 GMT. Figure 12 shows the "406 Not Acceptable" response from the TimeGate that has links to the first and last Memento, as well to two Mementos, one on either temporal side of the user agent's preferred interval.

```
GET /timegate/http://a.example.org HTTP/1.1
Host: arxiv.example.net
Accept-Datetime: Tue, 11 Sep 2001 20:35:00 GMT; -PT5H;+PT6H
Connection: close
```

Figure 11: User agent expresses interval of interest in Accept-Datetime header

```
HTTP/1.1 406 Not Acceptable
Date: Thu, 21 Jan 2010 00:06:50 GMT
Server: Apache
Vary: negotiate, accept-datetime
Link: <http://an.example.org>; rel="original",
      <http://arxiv.example.net/timemap/http://a.example.org>
      ; rel="timemap";type="application/link-format",
      <http://arxiv.example.net/web/20000915112826/http://a.example.org>
      ; rel="memento first"; datetime="Tue, 15 Sep 2000 11:28:26 GMT",
      <http://arxiv.example.net/web/20080708093433/http://a.example.org>
      ; rel="memento last"; datetime="Tue, 08 Jul 2008 09:34:33 GMT",
      <http://arxiv.example.net/web/20010910082200/http://a.example.org>
      ; rel="memento"; datetime="Mon, 10 Sep 2001 08:22:00 GMT",
      <http://arxiv.example.net/web/20010912034100/http://a.example.org>
      ; rel="memento"; datetime="Wed, 12 Sep 2001 03:41:00 GMT"
Content-Length: 1732
Connection: close
Content-Type: text/plain; charset=UTF-8
```

Figure 12: A TimeGate's response indicating it has no Mementos within the interval of interest

### **3.2.2.3. Multiple Matching Mementos**

Because the finest datetime granularity expressible using the [RFC 1123](#) [RFC1123] format used in HTTP is seconds level, cases may occur in which a TimeGate server is aware of multiple Mementos that meet the user agent's datetime preference. This may occur in Content



Management Systems with very high update rates. The response in this case MUST be handled as in [Section 3.2.2.1](#), with the selection of one of the matching Mementos.

As an example, Figure 13 shows a hypothetical response from a TimeGate on a MediaWiki server to a request for a Memento for the Original Resource `http://a.example.org/w/Clock` for which two Mementos exist for the user agent's preferred datetime.

```
HTTP/1.1 302 Found
Server: Apache
Content-Length: 705
Content-Type: text/html; charset=utf-8
Date: Thu, 21 Jan 2010 00:09:40 GMT
Vary: negotiate, accept-datetime
Location:
  http://a.example.org/w/index.php?title=Clock&oldid=322586071
Link: <http://a.example.org/w/Clock>; rel="original",
  <http://a.example.org/Special:TimeMap/http://a.example.org/w/Clock>
  ; rel="timemap"; type="application/link-format",
  <http://a.example.org/w/index.php?title=Clock&oldid=1493688>
  ; rel="first memento"; datetime="Sun, 28 Sep 2003 01:42:00 GMT",
  <http://a.example.org/w/index.php?title=Clock&oldid=337446696>
  ; rel="last memento"; datetime="Tue, 12 Jan 2010 19:55:00 GMT",
  <http://a.example.org/w/index.php?title=Clock&oldid=322586071>
  ; rel="memento"; datetime="Sun, 31 May 2009 15:43:00 GMT",
  <http://a.example.org/w/index.php?title=Clock&oldid=326164283>
  ; rel="memento successor-version"
  ; datetime="Sun, 31 May 2009 15:43:00 GMT"
  <http://a.example.org/w/index.php?title=Clock&oldid=326164283>
  ; rel="memento predecessor-version"
  ; datetime="Sun, 31 May 2009 15:41:24 GMT"
Connection: close
```

Figure 13: A TimeGate's response to a request that has multiple Mementos with a matching datetime

#### **[3.2.2.4](#). TimeGate Redirects to another TimeGate**

Cases may exist in which a TimeGate's response entails a redirects to another TimeGate, for example, because the responding TimeGate is aware that the other TimeGate is able to more precisely respond to a client's datetime preference. In such cases, the TimeGate's response MUST have a "302 Found" HTTP status code, and the "Location" header MUST be used to convey the URI of the other TimeGate. The "Vary" header MUST be provided and it MUST include the "negotiate" and "accept-datetime" values to indicate that, although datetime negotiation has not taken place, the responding TimeGate is capable



of it. The "Link" header MUST be provided and contain links with Relation Types subject to the considerations described in [Section 2.2](#). Specifically, the use of links with a "memento" Relation Type MUST follow the rules for the case where no Memento is selected by the responding server ([Section 2.2.1.4](#)). Also, a link with a "timegate" Relation Type MUST be provided that has as Target IRI the URI of the TimeGate to which the current TimeGate is redirecting the client. The response MUST NOT contain a "Memento-Datetime" header.

A response in which the client is redirected by TimeGate `http://arxiv.example.net/timegate/http://a.example.org` to TimeGate `http://otherarxiv.example.com/timegate/http://a.example.org` for the Original Resource `http://a.example.org` is illustrated in Figure 14. Note the URI of the latter TimeGate in both the "Location" and "Link" header, in the latter case as the Target IRI of a "timegate" link. Note also that the "memento" and "timemap" links in this response reflect the knowledge of the responding TimeGate, not of the remote TimeGate.

```
HTTP/1.1 302 Found
Date: Thu, 21 Jan 2010 00:06:50 GMT
Server: Apache
Vary: negotiate, accept-datetime
Location:
  http://otherarxiv.example.com/timegate/http://a.example.org
Link: <http://a.example.org>; rel="original",
  <http://arxiv.example.net/timemap/http://a.example.org>
    ; rel="timemap"; type="application/link-format",
  <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://otherarxiv.example.com/timegate/http://a.example.org>
    ; rel="timegate"
Content-Length: 0
Content-Type: text/plain; charset=UTF-8
Connection: close
```

Figure 14: TimeGate redirects to another TimeGate

#### **[3.2.2.5](#). Accept-Datetime and other Accept Headers Provided**

When interacting with a TimeGate, the regular content negotiation dimensions (media type, character encoding, language, and compression) remain available. It is the TimeGate server's responsibility to honor (or not) such content negotiation, and in doing so it MUST always first select a Memento that meets the user



agent's datetime preference, and then consider honoring regular content negotiation for it. As a result of this approach, the returned Memento will not necessarily meet the user agent's regular content negotiation preferences. Therefore, it is RECOMMENDED that the server provides HTTP Links with a "memento" Relation Type pointing at Mementos that do meet the user agent's regular content negotiation requests and that have a value for the "Memento-Datetime" header in the temporal vicinity of the user agent's preferred datetime value.

#### **3.2.2.6. Accept-Datetime Unparseable**

In case, in Step 3, a user agent conveys a value for the "Accept-Datetime" request header that does not conform to the accept-dt-value construction rule of the BNF in Figure 1, the TimeGate server's response MUST have a "400 Bad Request" HTTP status code. With all other respects, responses in this case MUST be handled as described in [Section 3.2.2.2](#).

#### **3.2.2.7. Accept-Datetime Not Provided**

In case, in Step 3, a user agent issues a request to a TimeGate and fails to include an "Accept-Datetime" request header, the response MUST be handled as in [Section 3.2.2.1](#), with a selection of the most recent Memento known to the responding server.

#### **3.2.2.8. TimeGate Does Not Exist**

Cases may occur in which a user agent issues a request against a TimeGate that does not exist. This may, for example, occur when a user agent uses internal knowledge to construct the URI of an assumed, yet non-existent TimeGate. In these cases, the response from the target server MUST have a "404 Not Found" HTTP status code, and SHOULD include a "Vary" header that includes the "negotiate" and "accept-datetime" values as an indication that, generally, the server is capable of datetime negotiation. The response MUST NOT include a "Link" header with any of the Relation Types introduced in [Section 2.2.1](#), and it MUST NOT contain a "Memento-Datetime" header.

#### **3.2.2.9. HTTP Methods other than HEAD/GET**

In the above, the safe HTTP methods GET and HEAD are described for TimeGates. TimeGates MAY support the safe HTTP methods OPTIONS and TRACE in the way described in [RFC 2616](#) [[RFC2616](#)]. Unsafe HTTP methods (i.e. PUT, POST, DELETE) MUST NOT be supported by a TimeGate. Such requests MUST yield a response with a "405 Method Not Allowed" HTTP status code, and MUST include an "Allow" header to convey that only the HEAD and GET (and OPTIONALLY the OPTIONS and



TRACE) methods are supported. In addition, the response MUST have a "Vary" header that includes the "negotiate" and "accept-datetime" values to indicate the TimeGate supports datetime negotiation. Figure 15 shows such a response.

```
HTTP/1.1 405 Method Not Allowed
Date: Thu, 21 Jan 2010 00:02:12 GMT
Server: Apache
Vary: negotiate, accept-datetime
Allow: HEAD, GET
Content-Length: 255
Connection: close
Content-Type: text/html; charset=iso-8909-1
```

Figure 15: Response from a TimeGate accessed with HTTP method other than HEAD/GET

### **3.2.3. Recognizing a TimeGate**

When a user agent issues a HTTP HEAD/GET request against an assumed TimeGate URI (e.g. URI is Target IRI of a link with a "timegate" Relation Type, URI is discovered as described in [Section 4.1](#), etc.), it SHOULD NOT conclude that the targeted resource effectively is a TimeGate and hence will behave as described in [Section 3.2.2](#).

A user agent MUST decide it has reached a TimeGate if the response to a HTTP HEAD/GET request against the resource's URI contains a "Vary" header that includes the "negotiate" and "accept-datetime" values. If the response does not, the user agent MUST decide it has not reached a TimeGate and proceed as follows:

- o If the response contains a redirection, the user agent SHOULD follow it. Note that a chain of redirections is possible, e.g. URI-R -> URI-1 -> URI-2 -> ... -> URI-G
- o If the response does not contain a redirection, or if the redirection (chain) does not lead to a TimeGate, the user agent SHOULD attempt to determine an appropriate TimeGate for the Original Resource, either automatically or interactively supported by the user. The discovery mechanisms described in [Section 4](#) can support the user agent with this regard.

Resources that are not TimeGates (i.e. do not behave as described in [Section 3.2.2](#)) MUST NOT use a "Vary" header that includes the "accept-datetime" value.

In certain cases, it is possible to implement Memento support in such a manner that an Original Resource coincides with its TimeGate, i.e.



URI-R and URI-G are the same. This implementation pattern is NOT RECOMMENDED. It can make determining whether a resource is a TimeGate more challenging, and, more importantly, it may cause problems with caches. Observed caching problems, which implementations must take care to avoid, include:

- o Cache invalidation when switching between a request for the Original Resource and a negotiation with the TimeGate.
- o Delivering a (cached) Original Resource response when a TimeGate response was requested, and vice versa.

### **3.3. Interactions with a Memento**

This section details HTTP GET/HEAD requests targeted at a Memento (URI-M).

#### **3.3.1. Step 5: User Agent Requests a Memento**

In Step 5, the user agent issues a HTTP GET request against the URI of a Memento. The user agent MAY include an "Accept-Datetime" header in this request, but the existence or absence of this header MUST NOT affect the server's response. The URI of the Memento may have resulted from a response in Step 4, or the user agent may simply have happened upon it. Such a request is illustrated in Figure 16.

```
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
```

Figure 16: User agent requests Memento

#### **3.3.2. Step 6: Server Responds to a Request for a Memento**

This section describes possible responses to a request for a Memento. [Section 3.3.2.1](#) discusses the common scenario, whereas [Section 3.3.2.2](#) and [Section 3.3.2.3](#) detail special cases whereby Mementos are archived copies of HTTP responses with 3xx, 4xx and 5xx status codes.

##### **3.3.2.1. Common Scenario**

If the Memento requested by the user agent in Step 5 exists, and is not a special Memento as described in [Section 3.3.2.2](#) and [Section 3.3.2.3](#), the server's response MUST have a "200 OK" HTTP status code or, where appropriate "206 Partial Content", and it MUST include a "Memento-Datetime" header with a value equal to the



archival datetime of the Memento, that is, the datetime of the state of the Original Resource that is encapsulated in the Memento. The "Link" header MUST be provided and contain links subject to the considerations described in [Section 2.2](#). The Target IRI and, when applicable, the datetime values in the "Link" header associated with the "memento" Relation Type SHOULD be the same as conveyed in Step 4, in case the TimeGate and the selected Memento reside on the same server. However, they MAY be different in case the TimeGate and the selected Memento reside on different servers.

Figure 17 illustrates the server's response to the request issued against a Memento in Step 5 (Figure 16).

```
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/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/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/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
```

Figure 17: Server of Memento responds

The server's response MUST include the "Memento-Datetime" header regardless whether the user agent's request contained an "Accept-Datetime" header or not. This is the way by which resources make explicit that they are Mementos. Due to the sparseness of Mementos in most archives, the value of the "Memento-Datetime" header returned by a server may differ (significantly) from the value conveyed by the user agent in "Accept-Datetime".

Although a Memento encapsulates a prior state of an Original Resource, the entity-body returned in response to an HTTP GET request



issued against a Memento may very well not be byte-to-byte the same as an entity-body that was previously returned by that Original Resource. Various reasons exist why there are significant chances these would be different yet do convey substantially the same information. These include format migrations as part of a digital preservation strategy, URI-rewriting as applied by some Web archives, and the addition of banners as a means to brand Web archives.

#### **3.3.2.2. Memento of a 3XX Response**

Cases exist in which HTTP responses with 3XX status codes are archived. For example, crawl-based web archives commonly archive responses with HTTP status codes "301 Moved Permanently" and "302 Found" whereas Linked Data archives hold on to "303 See Other" responses. But also other 3XX responses may be archived.

If the Memento requested by the user agent is an archived version of an HTTP response with a 3XX status code, the server's response MUST have the same 3XX HTTP status code, and it MUST include a "Memento-Datetime" header with a value equal to the archival datetime of the original 3XX response. All other considerations, e.g. pertaining to the use of "Link" header, expressed in [Section 3.3.2.1](#) apply.

The client's handling of a HTTP response with a 3XX status code is not affected by the presence of a "Memento-Datetime" header. The client SHOULD behave in the same manner as it does with HTTP responses with a 3XX status code that do not have a "Memento-Datetime" header. For example:

- o For a response from a Memento that has a 3XX status code and contains a "Location" header, the client SHOULD continue on to the URI specified in that header.
- o For a response from a Memento that has a "300 Multiple Choices" status code, the response body SHOULD be presented to the user to allow selection of a URI.

However, the client MUST be aware that the URI that was selected from the HTTP response with a 3XX status code might not be that of a Memento but rather of an Original Resource. In that case it SHOULD proceed by looking for a Memento of the selected Original Resource.

For example, on April 11 2008 Figure 18 is the response to an HTTP GET request for `http://a.example.org`. This response is archived as a Memento of `http://a.example.org`, and this Memento's URI is `http://arxiv.example.net/web/20080411000650/http://a.example.org`. The response to a HTTP HEAD/GET on this Memento is shown in Figure 19. In essence, it is a replay of the original response with



"Memento-Datetime" and "Link" headers added, to allow a client to understand the response is a Memento. In Figure 19, the value of the "Location" header is the same as in the original response; it identifies an Original Resource. The client proceeds with finding a Memento for this Original Resource. Web archives sometimes overwrite the value that was originally provided in the "Location" header in order to point at a Memento they hold of the resource to which the redirect originally led. This is shown in Figure 20. In this case, the client may decide it found an appropriate Memento.

```
HTTP/1.1 301 Moved Permanently
Date: Fri, 11 Apr 2008 00:06:50 GMT
Server: Apache
Location: http://b.example.org
Content-Length: 0
Content-Type: text/plain; charset=UTF-8
Connection: close
```

Figure 18: Response to the User Agent Request is a Redirect

```
HTTP/1.1 301 Moved Permanently
Date: Thu, 21 Jan 2010 00:09:40 GMT
Server: Apache-Coyote/1.1
Memento-Datetime: Fri, 11 Apr 2008 00:06:50 GMT
Location: http://b.example.org
Link: <http://a.example.org>; rel="original",
    <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/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/web/20080411000650/http://a.example.org>
    ; rel="memento"; datetime="Fri, 11 Apr 2008 00:06:50 GMT",
    <http://arxiv.example.net/web/20080410203051/http://a.example.org>
    ; rel="prev memento"; datetime="Thu, 10 Apr 2008 20:30:51 GMT",
    <http://arxiv.example.net/web/20080412204733/http://a.example.org>
    ; rel="next memento"; datetime="Sat, 12 Apr 2008 20:47:33 GMT"
Content-Length: 0
Content-Type: text/plain; charset=UTF-8
Connection: close
```

Figure 19: Response to a User Agent Request for a Memento of a Redirect; leads to an Original Resource



```
HTTP/1.1 301 Moved Permanently
Date: Thu, 21 Jan 2010 00:09:40 GMT
Server: Apache-Coyote/1.1
Memento-Datetime: Fri, 11 Apr 2008 00:06:50 GMT
Location:
  http://arxiv.example.net/web/20080411000655/http://b.example.org
Link: <http://a.example.org>; rel="original",
  <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/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/web/20080411000650/http://a.example.org>
    ; rel="memento"; datetime="Fri, 11 Apr 2008 00:06:50 GMT",
  <http://arxiv.example.net/web/20080410203051/http://a.example.org>
    ; rel="prev memento"; datetime="Thu, 10 Apr 2008 20:30:51 GMT",
  <http://arxiv.example.net/web/20080412204733/http://a.example.org>
    ; rel="next memento"; datetime="Sat, 12 Apr 2008 20:47:33 GMT"
Content-Length: 0
Content-Type: text/plain; charset=UTF-8
Connection: close
```

Figure 20: Response to a User Agent Request for a Memento of a Redirect; leads to a Memento

### **3.3.2.3. Memento of Responses with Other HTTP Status Codes**

Cases exist in which responses with 4xx and 5xx HTTP status codes are archived. If the Memento requested by the user agent is an archived version of such an HTTP response, the server's response MUST have the same 4xx or 5xx HTTP status code, and it MUST include a "Memento-Datetime" header with a value equal to the archival datetime of the original response. All other considerations, e.g. pertaining to the use of "Link" header, expressed in [Section 3.3.2.1](#) apply.

For example, on April 11 2008, Figure 21 is the 404 response to an HTTP GET request for `http://a.example.org`. This response is archived as a Memento of `http://a.example.org`, and this Memento's URI is `http://arxiv.example.net/web/20080411000650/http://a.example.org`. The response to a HTTP HEAD/GET on this Memento is shown in Figure 22. It is a replay of the original response with "Memento-Datetime" and "Link" headers added, to allow a client to understand the response is a Memento.



```
HTTP/1.1 404 Not Found
Date: Fri, 11 Apr 2008 00:06:50 GMT
Server: Apache
Content-Length: 0
Content-Type: text/plain; charset=UTF-8
Connection: close
```

Figure 21: Response to the User Agent Request is a 404

```
HTTP/1.1 404 Not Found
Date: Thu, 21 Jan 2010 00:09:40 GMT
Server: Apache-Coyote/1.1
Memento-Datetime: Fri, 11 Apr 2008 00:06:50 GMT
Link: <http://a.example.org>; rel="original",
      <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/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/web/20080411000650/http://a.example.org>
      ; rel="memento"; datetime="Fri, 11 Apr 2008 00:06:50 GMT",
      <http://arxiv.example.net/web/20010911203610/http://a.example.org>
      ; rel="prev memento"; datetime="Thu, 10 Apr 2008 20:30:51 GMT",
      <http://arxiv.example.net/web/20010911203610/http://a.example.org>
      ; rel="next memento"; datetime="Sat, 12 Apr 2008 20:47:33 GMT"
Content-Length: 0
Content-Type: text/plain; charset=UTF-8
Connection: close
```

Figure 22: Response to a User Agent Request for a Memento of a 404 Response

#### **3.3.2.4. Mementos Without a TimeGate**

Cases may occur in which a server that hosts Mementos does not expose a TimeGate for those Mementos. This can, for example, be the case if the server's Mementos result from taking a snapshot of the state of a set of Original Resources from another server at the time this other server is being retired. As a result, only a single Memento per Original Resource is hosted, making the introduction of a TimeGate unnecessary. But it may also be the case for servers that hosts multiple Mementos for an Original Resource but consider exposing TimeGates too expensive.



In cases of Mementos without associated TimeGates, responses to a request for a Memento by a user agent MUST be as described in [Section 3.3.2](#) with the exception that it will not contain a HTTP "Link" with a "timegate" Relation Type pointing at a TimeGate exposed by the responding server. It MAY still contain such a Link pointing at a TimeGate exposed elsewhere. Depending on whether one or more Mementos are hosted for an Original Resource, the response may or may not have a HTTP Link with a "timemap" Relation Type. However, the response MUST still contain a "Memento-Datetime" response header with a value that corresponds to archival datetime of the Memento.

Figure 23 illustrates the server's response to the request issued against a Memento in Step 5 (Figure 16) for the case that Memento has no associated TimeGate. In this example, it is also assumed there is only one Memento for the Original Resource, and hence the Links with Relation Types "memento", "first", "last" all point at the same - responding - Memento.

```
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/20010911203610/http://a.example.org>
      ; rel="first last memento"
      ; datetime="Tue, 15 Sep 2000 11:28:26 GMT"
Content-Length: 23364
Content-Type: text/html; charset=utf-8
Connection: close
```

Figure 23: Server of Memento without TimeGate responds

Note that a server issuing a response similar to that of Figure 23 does not imply that there is no server whatsoever that exposes a TimeGate; it merely means that the responding server neither provides nor is aware of the location of a TimeGate.

#### [3.3.2.5](#). Memento Does not Exist

Cases may occur in which a TimeGate's response (Step 4) points at a Memento that actually does not exist, resulting in a user agent's request (Step 5) for a non-existent Memento. In this case, the server's response MUST have the expected "404 Not Found" HTTP Status Code and it MUST NOT contain a "Memento-Datetime" header. Note that the absence of a Memento in an archive is distinct from the case of an archived response with a "404 Not Found" HTTP status code as is described in [Section 3.3.2.3](#)



### **3.3.3. Recognizing a Memento**

When following the redirection provided by a confirmed TimeGate (see [Section 3.2.3](#)), a user agent SHOULD NOT assume that the targeted resource effectively is a Memento and hence will behave as described in [Section 3.3.2](#).

A user agent MUST decide it has reached a Memento if the response to a HTTP HEAD/GET request against the resource's URI contains a "Memento-Datetime" header with a legitimate value. If the response does not, the following applies:

- o If the response contains a redirection, the user agent SHOULD follow it. Even a chain of redirections is possible, e.g. URI-G -> URI-X -> URI-Y -> ... -> URI-M.
- o If the response by a confirmed TimeGate does not contain a redirection, or if the redirection (chain) that started at a confirmed TimeGate does not lead to a resource that provides a "Memento-Datetime" header, the user agent MAY still conclude that it has likely arrived at a Memento. That is because cases exist in which Web archives and CMS are made compliant with the Memento framework "by proxy". In these cases TimeGates will redirect to Mementos in such systems, but the responses from these Mementos will not (yet) include a "Memento-Datetime" header.

### **3.4. Interactions with a TimeMap**

A TimeMap is introduced to support retrieving a comprehensive list of all Mementos for a specific Original Resource, known to a responding server. The entity-body of a response to an HTTP GET request issued against a TimeMap's URI:

- o MUST list the URI of the Original Resource that the response lists Mementos for;
- o MUST list the URI and datetime of each Memento for the Original Resource known to the responding server;
- o MUST list the URI of one or more TimeGates for the Original Resource except when no TimeGate exists (see [Section 3.3.2.4](#));
- o SHOULD, for self-containment, list the URI of the TimeMap itself;
- o MUST unambiguously type listed resources as being Original Resource, TimeGate, Memento, or TimeMap.

The entity-body of a response from a TimeMap MAY be serialized in



various ways, but the link-value format serialization MUST be supported. In this serialization, the entity-body MUST be formatted in the same way as the value of a HTTP "Link" header, and hence MUST comply to the "link-value" construction rule of "[Section 5](#)". The Link Header Field" of [RFC5988](#) [[RFC5988](#)], and the media type of the entity-body MUST be "application/link-format" as introduced in I-D.ietf-core-link-format [[I-D.ietf-core-link-format](#)]. All links conveyed in this serialization MUST be interpreted as having the URI of the Original Resource as their Context IRI. The URI of the Original Resource is provided in the entity-body as the Target IRI of the link with an "original" Relation Type.

#### **[3.4.1](#). User Agent Requests a TimeMap**

In order to retrieve the link-value serialization of a TimeMap, a user agent SHOULD use an "Accept" request header with a value set to "application/link-format". This is shown in Figure 24.

```
GET /timemap/http://a.example.org HTTP/1.1
Host: arxiv.example.net
Accept: application/link-format;q=1.0
Connection: close
```

Figure 24: Request for a TimeMap

#### **[3.4.2](#). Server Responds to a Request for a TimeMap**

If the TimeMap requested by the user agent exists, the server's response MUST have a "200 OK" HTTP status code (or "206 Partial Content", where appropriate). Note that a TimeMap is itself an Original Resource for which Mementos may exist. For example, a response from a TimeMap could provide a "timegate" Link to a TimeGate via which prior TimeMap versions are available. In this case, the use of the "Link" header is subject to all considerations described in [Section 2.2](#), with the TimeMap acting as the Original Resource.

However, in case a TimeMap wants to explicitly indicate in its response headers for which Original Resource it is a TimeMap, it MUST do so by including a HTTP "Link" header with the following characteristics:

- o The Context IRI for the HTTP Link is the URI of the Original Resource;
- o The Relation Type is "timemap";
- o The Target IRI for the HTTP Link is the URI of the TimeMap.



Because the Context IRI of this HTTP Link is not the URI of the TimeMap, as per [RFC5988](#) [[RFC5988](#)], the default Context IRI must be overwritten by using the "anchor" attribute with a value of the URI of the Original Resource.

The response from the TimeMap to the request of Figure 24 is shown in Figure 25. The response header shows the TimeMap explicitly conveying the URI of the Original Resource for which it is a TimeMap; for practical reasons the entity-body in the example has been abbreviated. Notice also the use of the "license" and "embargo" attributes introduced in [Section 2.2.1.4](#) on the "memento" links in the TimeMap.

```
HTTP/1.1 200 OK
Date: Thu, 21 Jan 2010 00:06:50 GMT
Server: Apache
Link: <http://arxiv.example.net/timemap/http://a.example.org>
    ; anchor="http://a.example.org"; rel="timemap"
    ; type="application/link-format"
Content-Length: 4883
Content-Type: application/link-format
Connection: close

<http://a.example.org>;rel="original",
<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/20000620180259/http://a.example.org>
    ; rel="first memento";datetime="Tue, 20 Jun 2000 18:02:59 GMT"
    ; license="http://creativecommons.org/publicdomain/zero/1.0/",
<http://arxiv.example.net/web/20091027204954/http://a.example.org>
    ; rel="last memento";datetime="Tue, 27 Oct 2009 20:49:54 GMT"
    ; license="http://creativecommons.org/publicdomain/zero/1.0/"
    ; embargo="Tue, 19 Apr 2011 00:00:00 GMT",
<http://arxiv.example.net/web/20000621011731/http://a.example.org>
    ; rel="memento";datetime="Wed, 21 Jun 2000 01:17:31 GMT"
    ; license="http://creativecommons.org/publicdomain/zero/1.0/",
<http://arxiv.example.net/web/20000621044156/http://a.example.org>
    ; rel="memento";datetime="Wed, 21 Jun 2000 04:41:56 GMT"
    ; license="http://creativecommons.org/publicdomain/zero/1.0/",
...
```

Figure 25: Response from a TimeMap



#### **4. The Memento Framework, Discovery Component**

[Section 3](#) describes how TimeGates, Mementos, Original Resources, and TimeMaps can be discovered by following HTTP Links with Relation Types "timegate", "memento", "original", and "timemap", respectively.

Naturally, some of these links can also be embedded into representations of resources that have a media type that allows embedding of typed links. For example, an Original Resource that has an HTML representation can include a "timegate" link by using HTML's LINK element, e.g. `<link href="http://arxiv.example.net/timegate/http://a.example.org" rel="timegate">`. The use of such embedded links is also subject to the considerations of [Section 2.2](#).

In this section additional approaches are introduced that support batch discovery of TimeGates and Mementos. The approaches leverage the Robots Exclusion Protocol.

##### **4.1. Discovering TimeGates Via Robots Exclusion Protocol**

The Robots Exclusion Protocol's robots.txt file [[robotstxt.org](http://robotstxt.org)] is commonly used by Web site owners to give instructions about their site to Web robots. It is used both to protect resources hosted by a server from crawling and to facilitate discovering them. This document introduces the "TimeGate" and "Archived" directives for robots.txt to provide a server-wide mechanism to support TimeGate discovery that SHOULD be used by:

- o Servers of Original Resources;
- o Servers that provide access to Mementos of Original Resources by exposing TimeGates.

A robots.txt file MAY contain zero or more occurrences of the "TimeGate" directive, and each occurrence MUST be followed by one or more associated "Archived" directives. The meaning of the directives is as follows:

- o TimeGate: Conveys the base URL (that is URI scheme, host and path component) that is shared by all URIs of TimeGates of a set of Original Resources.
- o Archived: Indicates - by means of mandatory host and optional path parts of a URI - for which set of Original Resources actual TimeGates are available that have the base URL conveyed in the associated TimeGate directive.



For example, consider a wiki at `http://a.example.org/w/` that supports the Memento framework and exposes TimeGates to access the wiki's history pages at base URL

`http://a.example.org/w/index.php/Special:TimeGate/`. An actual TimeGate for the wiki's `http://a.example.org/w/My_Title` page would then be at `http://a.example.org/w/index.php/Special:TimeGate/http://a.example.org/w/My_Title`. This wiki SHOULD make its TimeGates discoverable by using the directives shown in Figure 26 in its robots.txt file.

```
TimeGate: http://a.example.org/w/index.php/Special:TimeGate/  
Archived: a.example.org/w/
```

Figure 26: robots.txt for a wiki, host of Original Resources, TimeGates, and Mementos

As another example, consider a server of Original Resources at `http://a.example.org/` and <http://www.a.example.org/> that is aware that its resources are regularly crawled by a Web archive that generally exposes TimeGates at base URL

`http://arxiv.example.net/timegate/` and hence has TimeGate `http://arxiv.example.net/timegate/http://a.example.org/` to access Mementos for `http://a.example.org/`. This server SHOULD make the remote TimeGates discoverable by including the directives shown in Figure 27 in its robots.txt file:

```
TimeGate: http://arxiv.example.net/timegate/  
Archived: a.example.org/  
Archived: www.a.example.org/
```

Figure 27: robots.txt for a server of Original Resources aware of remote TimeGates

And, consider a Web archive that crawls a wide range of Original Resources, and exposes TimeGates to access the resulting Mementos at base URL `http://arxiv.example.net/timegate/`. In order to make its TimeGates discoverable, this Web archive SHOULD include the directives shown in Figure 28 in its robots.txt file:

```
TimeGate: http://arxiv.example.net/timegate/  
Archived: *
```

Figure 28: robots.txt for a Web Archive that hosts Mementos for a wide range of Original Resources



#### **4.2. Discovering Mementos via Robots Exclusion Protocol**

Servers can support discovery of their Mementos by crawlers through the use of the Robots Exclusion Protocol, but SHOULD do so in a manner that conveys to crawlers and mirroring applications that the sticky Memento-Datetime behavior (see [Section 2.1.1](#)) MUST be respected. To that end, servers SHOULD use the "User-agent" and "Allow" directives of the Robots Exclusion Protocol in the following manner:

- o User-agent: Has "memento" as its value;
- o Allow: Lists the path that contains Mementos that can be crawled, and for which content can be mirrored subject to the sticky Memento-Datetime behavior.

Figure 29 shows the robots.txt for a server that generally disallows crawling, yet allows agents that respect the sticky Memento-Datetime behavior to crawl Mementos in the /web/ path.

```
User-agent: *  
Disallow: /  
User-agent: memento  
Allow: /web/
```

Figure 29: Restricting crawling to agents that respect sticky Memento-Datetime behavior

#### **5. IANA Considerations**

This memo requires IANA to register the Accept-Datetime and Memento-Datetime HTTP headers defined in [Section 2.1.1](#) in the appropriate IANA registry.

This memo requires IANA to register the "Link" header Relation Types "original", "timegate", "timemap", and "memento" defined in [Section 2.2.1](#) in the appropriate IANA registry.

This memo requires IANA to register the "datetime", "license", and "embargo" attributes for "Link" headers with a "memento" Relation Type, as defined in [Section 2.2.1.4](#) in the appropriate IANA registry.

#### **6. Security Considerations**

Provision of a "timegate" HTTP "Link" header in responses to requests for an Original Resource that is protected (e.g., 401 or 403 HTTP



response codes) is OPTIONAL. The inclusion of this Link when requesting authentication is at the server's discretion; cases may exist in which a server protects the current state of a resource, but supports open access to prior states and thus chooses to supply a "timegate" HTTP "Link" header. Conversely, the server may choose to not advertise the TimeGate URIs (e.g., they exist in an intranet archive) for unauthenticated requests.

Authentication, encryption and other security related issues are otherwise orthogonal to Memento.

## **7. Changelog**

v04 2011-12-20 HVDS MLN RS [draft-vandesompel-memento-03](#)

- o Added description of Mementos of HTTP responses with 3XX, 4XX and 5XX status code.
- o Clarified that a TimeGate must not use the "Memento-Datetime" header.
- o Added wording to warn for possible cache problems with Memento implementations that choose to have an Original Resource and and its TimeGate coincide.

v03 2011-05-11 HVDS MLN RS [draft-vandesompel-memento-02](#)

- o Added scenario in which a TimeGate redirects to another TimeGate.
- o Reorganized TimeGate section to better reflect the difference between requests with and without interval indicator.
- o Added recommendation to provide "memento" links to Mementos in the vicinity of the preferred interval provided by the client, in case of a 406 response.
- o Removed TimeMap Feed material from the Discovery section as a result of discussions regarding (lack of) scalability of the approach with representatives of the International Internet Preservation Consortium. An alternative approach to support batch discovery of Mementos will be specified.

v02 2011-04-28 HVDS MLN RS [draft-vandesompel-memento-01](#)

- o Introduced wording and reference to indicate a Memento is a FixedResource.



- o Introduced "Sticky Memento-Datetime" notion and clarified wording about retaining "Memento-Datetime" headers and values when a Memento is mirrored at different URI.
- o Introduced section about handling both datetime and regular negotiation.
- o Introduced section about Mementos Without TimeGate.
- o Made various changes in the section Relation Type "memento", including addition of "license" and "embargo" attributes, and clarification of rules regarding the use of "memento" links.
- o Moved section about TimeMaps inside the Datetime Negotiation section, and updated it.
- o Restarted the Discovery section from scratch.

v01 2010-11-11 HVDS MLN RS First public version  
[draft-vandesompe1-memento-00](#)

v00 2010-10-19 HVDS MLN RS Limited circulation version

2010-07-22 HVDS MLN First internal version

## **8. Acknowledgements**

The Memento effort is funded by the Library of Congress. Many thanks to Kris Carpenter Negulescu, Michael Hausenblas, Erik Hetzner, Larry Masinter, Gordon Mohr, Mark Nottingham, David Rosenthal, Ed Summers for early feedback. Many thanks to Samuel Adams, Scott Ainsworth, Lyudmilla Balakireva, Frank McCown, Harihar Shankar, Brad Tofel for early implementations.

## **9. References**

### **9.1. Normative References**

- [I-D.ietf-core-link-format]  
Shelby, Z., "CoRE Link Format",  
[draft-ietf-core-link-format-09](#) (work in progress),  
November 2011.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), March 1997.



- [RFC2616] Fielding, R., Gettys, J., Mogul, J., Frystyk, H., Masinter, L., Leach, P., and T. Berners-Lee, "Hypertext Transfer Protocol -- HTTP/1.1", [RFC 2616](#), June 1999.
- [RFC4151] Kindberg, T. and S. Hawke, "The 'tag' URI Scheme", [RFC 4151](#), October 2005.
- [RFC4287] Nottingham, M., Ed. and R. Sayre, Ed., "The Atom Syndication Format", [RFC 4287](#), December 2005.
- [RFC5829] Brown, A., Clemm, G., and J. Reschke, "Link Relation Types for Simple Version Navigation between Web Resources", [RFC 5829](#), April 2010.
- [RFC5988] Nottingham, M., "Web Linking", [RFC 5988](#), October 2010.

## 9.2. Informative References

- [Fitch] Fitch, "Web site archiving - an approach to recording every materially different response produced by a website", July 2003, <<http://ausweb.scu.edu.au/aw03/papers/fitch/paper.html>>.
- [I-D.masinter-dated-uri] Masinter, L., "The 'tdb' and 'duri' URI schemes, based on dated URIs", [draft-masinter-dated-uri-08](#) (work in progress), January 2011.
- [RFC1123] Braden, R., "Requirements for Internet Hosts - Application and Support", STD 3, [RFC 1123](#), October 1989.
- [W3C.REC-aww-20041215] Jacobs and Walsh, "Architecture of the World Wide Web", December 2004, <<http://www.w3.org/TR/webarch/>>.
- [W3C.gen-ont-20090420] Berners-Lee, "Architecture of the World Wide Web", April 2009, <<http://www.w3.org/2006/gen/ont>>.
- [robotstxt.org] "Robots Exclusion Protocol", August 2010, <<http://www.robotstxt.org/robotstxt.html>>.

## [Appendix A](#). [Appendix B](#): A Sample, Successful Memento Request/Response cycle

Step 1 : UA --- HTTP GET/HEAD; Accept-Datetime: Tj -----> URI-R



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

Step 2 : UA <-- HTTP 200; Link: URI-G ----- URI-R

```
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
```

Step 3 : UA --- HTTP GET/HEAD; Accept-Datetime: Tj -----> URI-G

```
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
```

Step 4 : UA <-- HTTP 302; Location: URI-Mj; Vary; Link:  
URI-R, URI-T, URI-M0, URI-Mn, URI-Mi, URI-Mj, URI-Mk ---- URI-G

```
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

Step 5 : UA --- HTTP GET URI-Mj; Accept-Datetime: Tj -----> URI-Mj

```
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
```

Step 6 : UA <-- HTTP 200; Memento-Datetime: Tj; Link: URI-R,  
URI-T, URI-G, URI-M0, URI-Mn, URI-Mi, URI-Mj, URI-Mk ---- URI-Mj

```
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
```

A successful flow with TimeGate and Mementos on the same server



## Authors' Addresses

Herbert VandeSompel  
Los Alamos National Laboratory  
PO Box 1663  
Los Alamos, New Mexico 87545  
USA

Phone: +1 505 667 1267  
Email: [hvdsomp@gmail.com](mailto:hvdsomp@gmail.com)  
URI: <http://public.lanl.gov/herbertv/>

Michael Nelson  
Old Dominion University  
Norfolk, Virginia 23529  
USA

Phone: +1 757 683 6393  
Email: [mln@cs.odu.edu](mailto:mln@cs.odu.edu)  
URI: <http://www.cs.odu.edu/~mln/>

Robert Sanderson  
Los Alamos National Laboratory  
PO Box 1663  
Los Alamos, New Mexico 87545  
USA

Phone: +1 505 665 5804  
Email: [azaro42@gmail.com](mailto:azaro42@gmail.com)

