Avoiding Spoilers On MediaWiki Fan Sites Using Memento

By

Shawn M. Jones

sjone@cs.odu.edu

Committee:
Dr. Michael L. Nelson
Dr. Michele C. Weigle
Dr. Irwin B. Levinstein

OLD DOMINION
UNIVERSITY

Warning: This presentation may contain spoilers

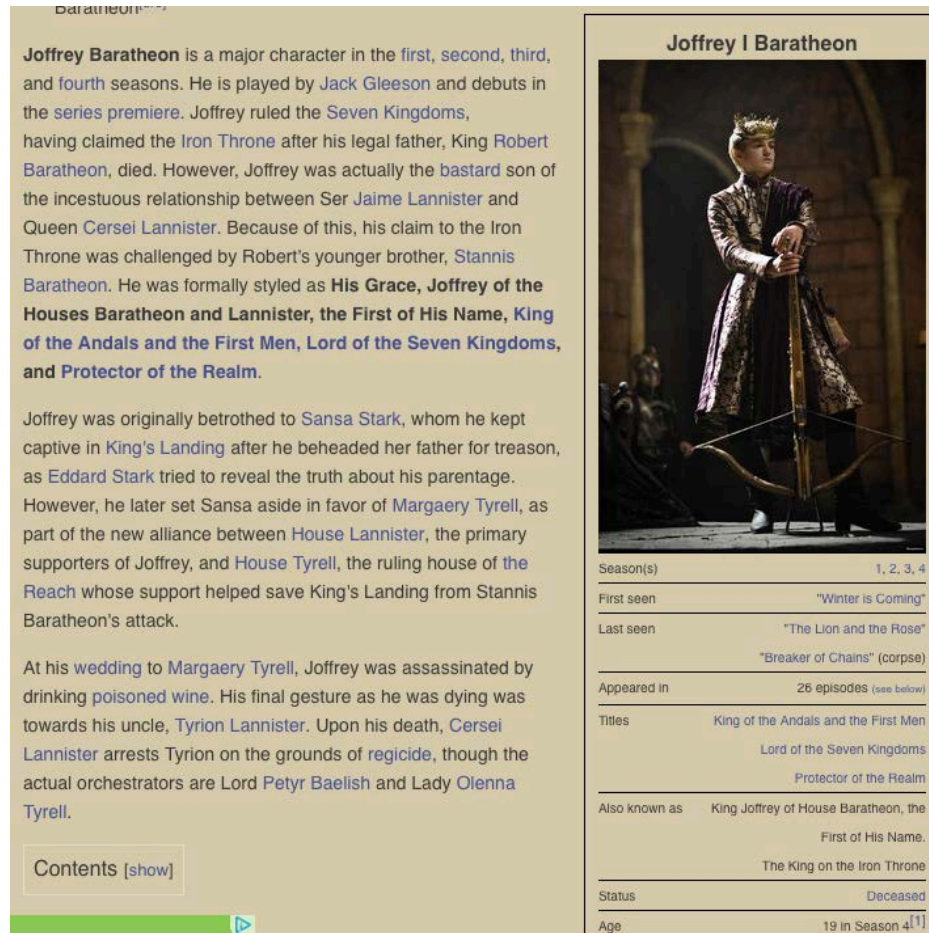# It Started With A Discussion At Work About This Guy From *Game Of Thrones*

So We Use Fan Wikis, Because They Are Useful For Our Discussions

Baratheon

Joffrey Baratheon is a major character in the first, second, third, and fourth seasons. He is played by Jack Gleeson and debuts in the series premiere. Joffrey ruled the Seven Kingdoms, having claimed the Iron Throne after his legal father, King Robert Baratheon, died. However, Joffrey was actually the bastard son of the incestuous relationship between Ser Jaime Lannister and Queen Cersei Lannister. Because of this, his claim to the Iron Throne was challenged by Robert's younger brother, Stannis Baratheon. He was formally styled as **His Grace, Joffrey of the Houses Baratheon and Lannister, the First of His Name, King of the Andals and the First Men, Lord of the Seven Kingdoms, and Protector of the Realm**.

Joffrey was originally betrothed to Sansa Stark, whom he kept captive in King's Landing after he beheaded her father for treason, as Eddard Stark tried to reveal the truth about his parentage. However, he later set Sansa aside in favor of Margaery Tyrell, as part of the new alliance between House Lannister, the primary supporters of Joffrey, and House Tyrell, the ruling house of the Reach whose support helped save King's Landing from Stannis Baratheon's attack.

At his wedding to Margaery Tyrell, Joffrey was assassinated by drinking poisoned wine. His final gesture as he was dying was towards his uncle, Tyrion Lannister. Upon his death, Cersei Lannister arrests Tyrion on the grounds of regicide, though the actual orchestrators are Lord Petyr Baelish and Lady Olenna Tyrell.

Contents [show]

Joffrey I Baratheon

Season(s)	1, 2, 3, 4
First seen	"Winter is Coming"
Last seen	"The Lion and the Rose"
	"Breaker of Chains" (corpse)
Appeared in	26 episodes (see below)
Titles	King of the Andals and the First Men
	Lord of the Seven Kingdoms
	Protector of the Realm
Also known as	King Joffrey of House Baratheon, the First of His Name.
	The King on the Iron Throne
Status	Deceased
Age	19 in Season 4[1]

http://gameofthrones.wikia.com/wiki/Joffrey_Baratheon

But The Current Page For This Character Contains Spoilers

But The Current Page For This Character Contains Spoilers

But The Current Page For This Character Contains Spoilers

Joffrey Baratheon is a major character in the first, second, third, and fourth seasons. He is played by Jack Gleeson and debuts in the series premiere. Joffrey ruled the Seven Kingdoms, having claimed the Iron Throne after his legal father, King Robert Baratheon[...] ...stard son of the incestu[...] ...ter and Queen Ce[...] the Iron Throne wa[...] Stannis Baratheon[...] ...rey of the Houses Baratheon and Lannister, the First of His Name, King of the Andals and the First Men, Lord of the Seven Kingdoms, and Protector of the Realm.

Joffrey was originally betrothed to Sansa Stark, whom he kept captive in King's Landing after he beheaded her father for treason, as Eddard Stark tried to reveal the truth about his parentage. However, he later set Sansa aside in favor of Margaery Tyrell, as part of the new alliance between House Lannister, the primary supporters of Joffrey, and House T[...] Reach whose support helped save[...] Baratheon's attack.

At his wedding to Margaery Tyrell, [...] drinking poisoned wine. His final gesture as he was dying was towards his uncle, Tyrion Lannister. Upon his death, Cersei Lannister arrests Tyrion on the grounds of regicide, though the actual orchestrators are Lord Petyr Baelish and Lady Olenna Tyrell.

Contents [show]

Joffrey I Baratheon

He was

Joffrey was assassinated

Titles	King of the Andals and the First Men
	Lord of the Seven Kingdoms
	Protector of the Realm
Also known as	King Joffrey of House Baratheon, the First of His Name.
	The King on the Iron Throne
Status	
Age	

Deceased

But The Current Page For This Character Contains Spoilers

Joffrey Baratheon is a major character in the first, second, third, and fourth seasons. He is played by Jack Gleeson and debuts in the series premiere. Joffrey ruled the Seven Kingdoms, having claimed the Iron Throne after his legal father, King Robert Baratheon [...] [...] stard son of the incestu [...] [...] ter and Queen Ce [...] the Iron Throne wa [...] Stannis Baratheon [...] rey of the **Houses Baratheon and Lannister, the First of His Name, King of the Andals and the First Men, Lord of the Seven Kingdoms, and Protector of the Realm.**

Joffrey was originally betrothed to Sansa Stark, whom he kept captive in King's Landing after he beheaded her father for treason, as Eddard Stark tried to reveal the truth about his parentage. However, he later set Sansa aside in favor of Margaery Tyrell, as part of the new alliance between House Lannister, the primary supporters of Joffrey, and House T [...] Reach whose support helped save [...] Baratheon's attack.

At his wedding to Margaery Tyrell, [...] [...] nal gesture as he was dying was [...] nister. Upon his death, Cersei [...] e grounds of regicide, though the [...] Petyr Baelish and Lady Olenna

He was

Joffrey was assassinated

poisoned wine.

Contents [show]

Joffrey I Baratheon

Titles	King of the Andals and the First Men
	Lord of the Seven Kingdoms
	Protector of the Realm
Also known as	King Joffrey of House Baratheon, the First of His Name.
	The King on the Iron Throne
Status	
Age	

Deceased

But The Current Page For This Character Contains Spoilers

Joffrey Baratheon is a major character in the first, second, third, and fourth seasons. He is played by Jack Gleeson and debuts in the series premiere. Joffrey ruled the Seven Kingdoms, having claimed the Iron Throne after his legal father, King Robert Baratheon ~~stard son of~~ the incestu ~~ter and~~ Queen Ce ~~the Iron~~ Throne wa ~~Stannis~~ Baratheon ~~rey of the~~ **Houses Baratheon and Lannister, the First of His Name, King of the Andals and the First Men, Lord of the Seven Kingdoms, and Protector of the Realm.**

Joffrey was originally betrothed to Sansa Stark, whom he kept captive in King's Landing after he beheaded her father for treason, as Eddard Stark tried to reveal the truth about his parentage. However, he later set Sansa aside in favor of Margaery Tyrell, as part of the new alliance between House Lannister, the primary supporters of Joffrey, and House T Reach whose support helped save Baratheon's attack.

At his wedding to Margaery Tyrell, ~~nal gesture as he was dying was~~ ~~niste~~ ~~e gro~~ ~~Pety~~

Contents [show]

Joffrey I Baratheon

He was

Joffrey was assassinated

poisoned wine.

Upon his death,

Deceased

Status
Age

First of His Name,
The King on the Iron Throne

We All Enjoy Some Episodic Fiction So Much...

…That Fans Have Created Wikis…

…And The Rest Of Us Read Them

So, What If We Could Avoid The Spoilers By Using Past Wiki Pages?

"We've had vicious kings, and we've had idiot kings...but I don't know if we've ever been cursed with a vicious idiot for a king!"
—Tyrion Lannister, summarizing King Joffrey.[src]

Joffrey Baratheon is a major character in the first, second and third seasons. He is played by Jack Gleeson and debuts in the series premiere. Joffrey is the current ruler of the Seven Kingdoms, having claimed the Iron Throne after his father, King Robert Baratheon, died. However Joffrey is actually the bastard son of the incestuous relationship between Ser Jaime Lannister and Queen Cersei Lannister. Because of this, his claim to the Iron Throne is challenged by Robert's younger brother, Stannis Baratheon. He is formally styled as **His Grace, Joffrey of the Houses Baratheon and Lannister, the First of His Name, King of the Andals and the First Men, Lord of the Seven Kingdoms, and Protector of the Realm.**

Joffrey was originally betrothed to Sansa Stark, whom he kept captive in King's Landing after he beheaded her father for treason, as Eddard Stark tried to reveal the truth about his parentage. However, he later set Sansa aside in favor of Margaery Tyrell, as part of the new alliance between House Lannister, the primary supporters of Joffrey, and House Tyrell, the ruling house of the Reach whose support helped save King's Landing from Stannis

Joffrey I Baratheon

Season(s)	1, 2, 3, 4
First seen	"Winter is Coming"
Last seen	"Two Swords"
Appeared in	25 episodes (see below)
Titles	King of the Andals and the First Men
	Lord of the Seven Kingdoms
	Protector of the Realm
Also known as	King Joffrey of Houses Baratheon and
Age	19 in Season 4[*]
Origin	King's Landing

Status — Alive

http://gameofthrones.wikia.com/wiki/Joffrey_Baratheon?oldid=125053

Order Of Discussion

- Background
- Related Work
- TimeGate Heuristics
- Theory Of Spoiler Probability
- Measurements Of Spoiler Probability
- Spoilers In The Wayback Machine
- The Memento MediaWiki Extension
- Future Work
- Conclusions

Building To The Naïve Spoiler Concept

BACKGROUND

Most Of Us Are Familiar With The Web Browser And HTML

The Web Browser is how we view web pages

HTML is what the browser parses to render the page

Hypertext Transfer Protocol (HTTP)
Is What Delivers Pages To The Browser

HTTP request

(method) URI (HTTP version)
(headers)
(required blank line)
(entity - optional)

User Agent

HTTP response

(HTTP version) (status code)
(headers)
(required blank line)
(entity - optional)

Origin Server
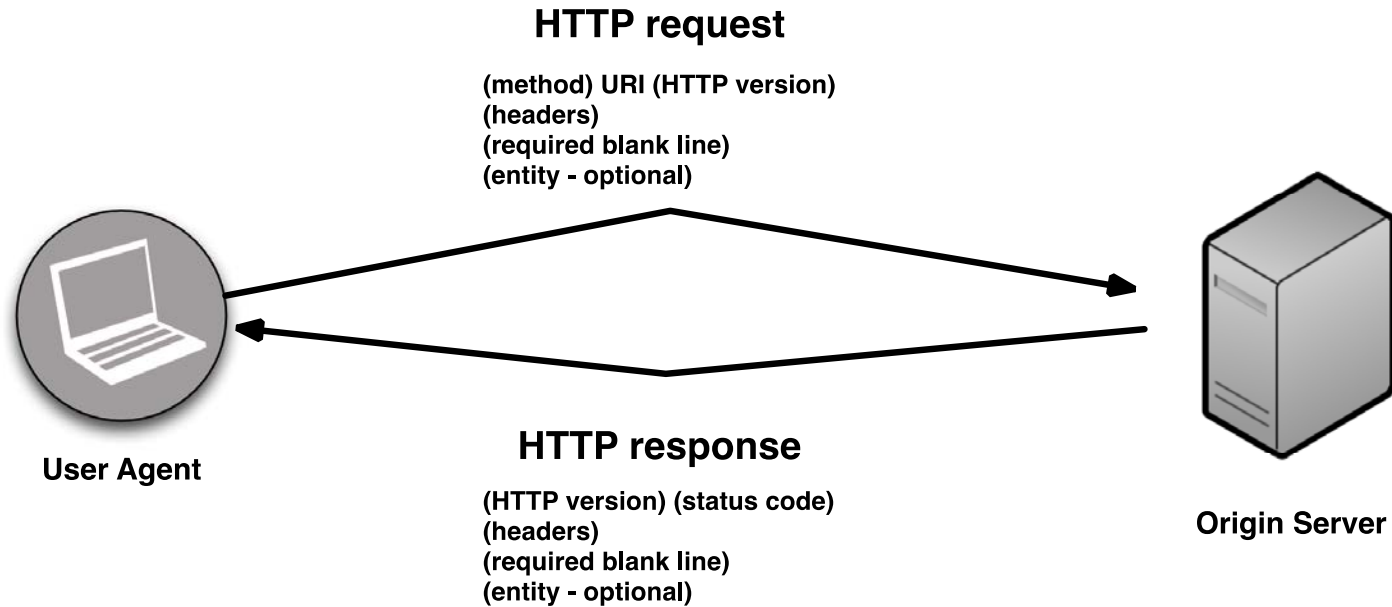
Here Is An Example HTTP Request From Google Chrome

```
1   GET /wiki/The_Hunger_Games HTTP/1.1
2   Host: en.wikipedia.org
3   Accept: image/webp,*/*;q=0.8
4   Accept-Encoding: gzip,deflate,sdch
5   Accept-Language: en-US,en;q=0.8
6   Referer: https://plus.google.com/
7   User-Agent: Mozilla/5.0 (Macintosh; Intel Mac OS X 10_9_4) AppleWebKit/537.36 (KHTML,
        like Gecko) Chrome/36.0.1985.143 Safari/537.36
```

And The Interesting Parts Of Our Request...

```
1  GET /wiki/The_Hunger_Games HTTP/1.1
2  Host: en.wikipedia.org
3  Accept: image/webp,*/*;q=0.8
4  Accept-Encoding: gzip,deflate,sdch
5  Accept-Language: en-US,en;q=0.8
6  Referer: https://plus.google.com/
7  User-Agent: Mozilla/5.0 (Macintosh; Intel Mac OS X 10_9_4) AppleWebKit/537.36 (KHTML,
       like Gecko) Chrome/36.0.1985.143 Safari/537.36
```

Get me /wiki/The_Hunger_Games from en.wikipiedia.org

Here Is An Example HTTP Response From Wikipedia's Server For That Page

```
1   HTTP/1.1 200 OK
2   Accept-Ranges: bytes
3   Age: 352057
4   Cache-Control: private, s-maxage=0, max-age=0, must-revalidate
5   Connection: keep-alive
6   Content-Encoding: gzip
7   Content-language: en
8   Content-Length: 13769
9   Content-Type: text/html; charset=UTF-8
10  Date: Fri, 15 Aug 2014 13:43:03 GMT
11  Last-Modified: Mon, 11 Aug 2014 11:55:18 GMT
12  Server: Apache
13  Set-Cookie: GeoIP=US:Norfolk:36.9312:-76.2397:v4; Path=/; Domain=.wikipedia.org
14  Vary: Accept-Encoding,Cookie
15  Via: 1.1 varnish, 1.1 varnish
16  X-Cache: cp1052 hit (4), cp1065 frontend hit (906)
17  X-Content-Type-Options: nosniff
18  X-UA-Compatible: IE=Edge
19  X-Varnish: 350285837 350035840, 1753471343 1233627904
20
21  ... entity begins here
```

And The Interesting Parts Of Our Response…

```
1    HTTP/1.1 200 OK
2    Accept-Ranges: bytes
3    Age: 352057
4    Cache-Control: private, s-maxage=0, max-age=0, must-revalidate
5    Connection: keep-alive
6    Content-Encoding: gzip
7    Content-language: en
8    Content-Length: 13769
9    Content-Type: text/html; charset=UTF-8
10   Date: Fri, 15 Aug 2014 13:43:03 GMT
11   Last-Modified: Mon, 11 Aug 2014 11:55:18 GMT
12   Server: Apache
13   Set-Cookie: GeoIP=US:Norfolk:36.9312:-76.2397:v4; Path=/; Domain=.wikipedia.org
14   Vary: Accept-Encoding,Cookie
15   Via: 1.1 varnish, 1.1 varnish
16   X-Cache: cp1052 hit (4), cp1065 frontend hit (906)
17   X-Content-Type-Options: nosniff
18   X-UA-Compatible: IE=Edge
19   X-Varnish: 350285837 350035840, 1753471343 1233627904
20
21   ... entity begins here
```

OK, I Have What You Are Looking For

This Is How Big It Is

This Is When It Was Last Changed

Newline Indicating Start Of HTML

HTML Starts Here And You Know How Big It Is, So Stop Reading When You Get 13769 Bytes

Introducing Web Architecture:
Resources Can Have Many Representations

URI http://www.example.com/buffy/

identifies

Resource

Article about the character
Buffy the Vampire Slayer

Metadata:
Content-Language: en
Content-Type:
application/xhtml+xml

<!DOCTYPE html PUBLIC "...
<html xmlns="http://www....
<head>
<title>Character: Buffy</title>
...
</html>

represents

**English XHTML
Representation**

Metadata:
Content-Language: es
Content-Type:
text/html

<!DOCTYPE html PUBLIC "...
<html>
<head>
<title>Carácter: Buffy</title>
...
</html>

represents

**Spanish HTML
Representation**

Metadata:
Content-Language: zh
Content-Type:
application/pdf

%PDF-1.4
%?஄
5 0 obj
<</Length 6 0 R/Filter /
FlateDecode>>
stream
...

represents

**Chinese PDF
Representation**

In 1996 Tim Berners-Lee Discussed Different Dimensions Of Representation

Time	A resource may vary with time. For example, "The Wall Street Journal" varies with time. Each issue is a time-specific resource, which does not change with time. Most home pages on the Web change with time, in a less periodic way.
Language	When a document is translated, it is useful to be able to refer to it either in the generic, or to a particular specific translation.
Content-Type	A given resource may have mny ways in which it can be represented on the wire, using different content-types (in HTTP terms). As an example, an image may be represented in PNG or JFIF format.
Target medium	A given resource may be targetted specifically to a specific medium, such as a printer, being displayed on laptop screen, being displayed on a cellphone, or being projected onto a large screen for an audience. (This is currenltly available for selecting CSS stylesheets, but is not done at the HTTP content negotiation level)

http://www.w3.org/DesignIssues/Generic.html

These Representations Are Requested And Identified By HTTP Headers

Request Header	Response Header	Dimension
Accept	Content-Type	Content type of Representation
Accept-Language	Content-Language	Language of Representation
Accept-Encoding	Content-Encoding	Medium of Representation
Accept-Charset	Content-Type	Character Set of Representation

These Headers Were Specified In RFC 2616

Their Updated Specification Is In RFC 7231

These Representations Are Requested And Identified By HTTP Headers

Request Header	Response Header	Dimension
Accept	Content-Type	Content type of Representation
Accept-Language	Content-Language	Language of Representation
Accept-Encoding	Content-Encoding	Medium of Representation
Accept-Charset	Content-Type	Character Set of Representation

The term for this is **Content Negotiation**

It works so well that many are unaware of its ubiquity

Content Negotiation Does Not Create Representations,
It Only Directs The User To Ones That Already Exist…

Using Our Headers Example, Chrome Tells The Server What You Want

```
1  GET /wiki/The_Hunger_Games HTTP/1.1
2  Host: en.wikipedia.org
3  Accept: image/webp,*/*;q=0.8
4  Accept-Encoding: gzip,deflate,sdch
5  Accept-Language: en-US,en;q=0.8
6  Referer: https://plus.google.com/
7  User-Agent: Mozilla/5.0 (Macintosh; Intel Mac OS X 10_9_4) AppleWebKit/537.36 (KHTML,
       like Gecko) Chrome/36.0.1985.143 Safari/537.36
```

I prefer image/webp, and
I want it compressed with gzip, and
I want it in US English

Using Our Headers Example, Wikipedia Tells Chrome What It Returns

```
1   HTTP/1.1 200 OK
2   Accept-Ranges: bytes
3   Age: 352057
4   Cache-Control: private, s-maxage=0, max-age=0, must-revalidate
5   Connection: keep-alive
6   Content-Encoding: gzip
7   Content-language: en
8   Content-Length: 13769
9   Content-Type: text/html; charset=UTF-8
10  Date: Fri, 15 Aug 2014 13:43:03 GMT
11  Last-Modified: Mon, 11 Aug 2014 11:55:18 GMT
12  Server: Apache
13  Set-Cookie: GeoIP=US:Norfolk:36.9312:-76.2397:v4; Path=/; Domain=.wikipedia.org
14  Vary: Accept-Encoding,Cookie
15  Via: 1.1 varnish, 1.1 varnish
16  X-Cache: cp1052 hit (4), cp1065 frontend hit (906)
17  X-Content-Type-Options: nosniff
18  X-UA-Compatible: IE=Edge
19  X-Varnish: 350285837 350035840, 1753471343 1233627904
20
21  ... entity begins here
```

My response is compressed with gzip, and My response is in US English

This response is in text/html with a character set of UTF-8

Memento Finally Completes The Set By Including Time

Request Header	Response Header	Dimension
Accept	Content-Type	Content type of Representation
Accept-Language	Content-Language	Language of Representation
Accept-Encoding	Content-Encoding	Medium of Representation
Accept-Charset	Content-Type	Character Set of Representation
Accept-Datetime	Memento-Datetime	Time of the Representation

RFC 7089 Defines Memento, Allowing Us To Negotiate In Time

Memento Works In Several Steps

(1) CLIENT:
HEAD <URI-R>
Accept-Datetime: DDDDD

SERVER:
HTTP/1.1 200 OK
Link: <URI-G, rel=timegate>

Original Resource (URI-R) — Current Page

(2) CLIENT:
GET <URI-G>
Accept-Datetime: DDDDD

SERVER:
HTTP/1.1 302 Found
Location: <URI-M>
Vary: Accept-Datetime

TimeGate (URI-G) — Resource That Redirects To Archived Pages

User running Memento Client

(3) CLIENT:
GET <URI-M>

SERVER:
HTTP/1.1 200 OK
Memento-Datetime: DDDDD

Memento (URI-M) — Archived Page From The Past

At First The Client Hopes To Get A TimeGate URI From the Server, But…

①

CLIENT:
HEAD <URI-R>
Accept-Datetime: DDDDD

Most Servers Do Not Yet Return This Data

So Memento Clients Default To Known TimeGates

SERVER:
HTTP/1.1 200 OK
Link: <URI-G, rel=timegate>

Original Resource (URI-R)

Current Page

②

CLIENT:
GET <URI-G>
Accept-Datetime: DDDDD

User running Memento Client

SERVER:
HTTP/1.1 302 Found
Location: <URI-M>
Vary: Accept-Datetime

TimeGate (URI-G)

Resource That Redirects To Archived Pages

③

CLIENT:
GET <URI-M>

SERVER:
HTTP/1.1 200 OK
Memento-Datetime: DDDDD

Memento (URI-M)

Archived Page From The Past

Then, A Browser Asks A TimeGate For A Specific Page From A Specific Datetime

```
GET /web/http://www.cnn.com/ HTTP/1.1
Host: web.archive.org
Accept: image/webp,*/*;q=0.8
Accept-Datetime: Wed, 19 Mar 2003 01:25:35 GMT
Accept-Encoding: gzip,deflate,sdch
Accept-Language: en-US,en;q=0.8
Referer: http://www.cnn.com/
User-Agent: Mozilla/5.0 (Macintosh; Intel Mac OS X 10_9_4) AppleWebKit/537.36 (KHTML,
    like Gecko) Chrome/36.0.1985.143 Safari/537.36
```

I Want This Page On This Date

1
CLIENT:
HEAD <URI-R>
Accept-Datetime: DDDDD

SERVER:
HTTP/1.1 200 OK
Link: <URI-G, rel=timegate>

Original Resource (URI-R)

2
CLIENT:
GET <URI-G>
Accept-Datetime: DDDDD

SERVER:
HTTP/1.1 302 Found
Location: <URI-M>
Vary: Accept-Datetime

TimeGate (URI-G)

User running Memento Client

3
CLIENT:
GET <URI-M>

Memento (URI-M)

SERVER:
HTTP/1.1 200 OK
Memento-Datetime: DDDDD
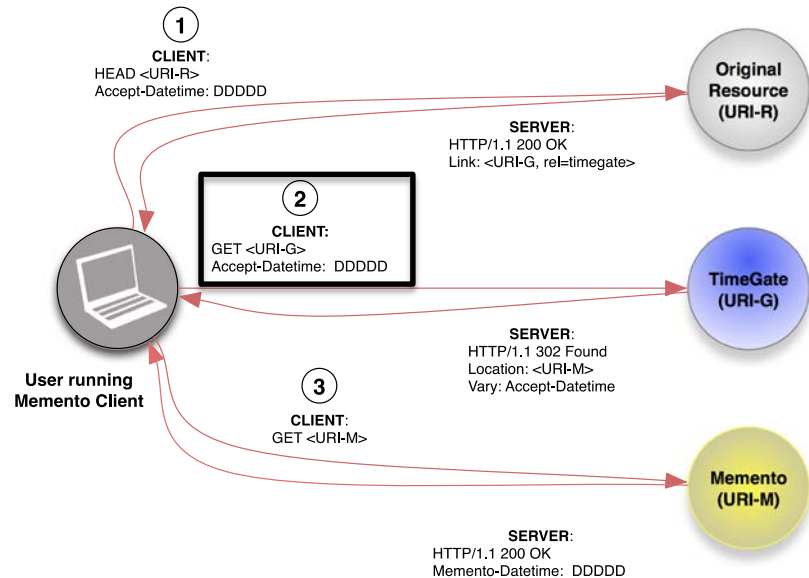
And The TimeGate Tells The Browser Where It Can Go

Do Not Stop Here, I Am Redirecting You Elsewhere
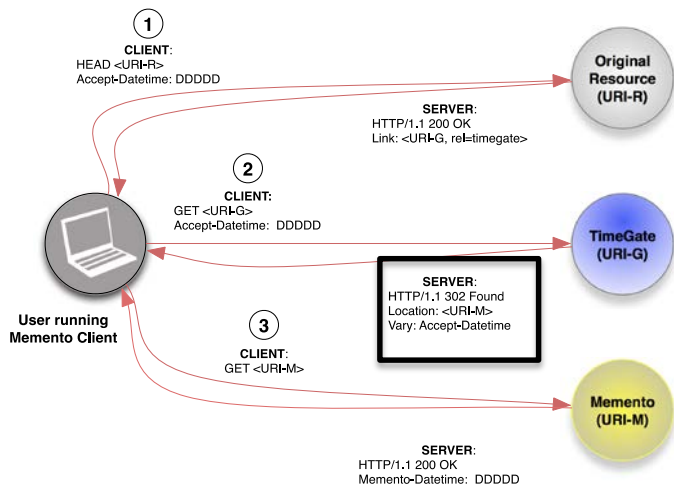

```
HTTP/1.1 302 Moved Temporarily
Connection: keep-alive
Content-Type: text/html
Date: Fri, 22 Aug 2014 17:27:57 GMT
Link: <http://www.cnn.com></http:>; rel="original", <http://web.archive.org/web/
    timemap/link/http://www.cnn.com></http:>; rel="timemap"; type="application/link-
    format", <http://web.archive.org/web/20000620180259/http://www.cnn.com></http:>;
    rel="first memento"; datetime="Tue, 20 Jun 2000 18:02:59 GMT", <http://web.archive
    .org/web/20030215012445/http://www.cnn.com></http:>; rel="prev memento"; datetime
    ="Sat, 15 Feb 2003 01:24:45 GMT", <http://web.archive.org/web/20030320060210/http
    ://www.cnn.com></http:>; rel="memento"; datetime="Thu, 20 Mar 2003 06:02:10 GMT",
    <http://web.archive.org/web/20030321181645/http://www.cnn.com></http:>; rel="next
    memento"; datetime="Fri, 21 Mar 2003 18:16:45 GMT", <http://web.archive.org/web
    /20140822104304/http://www.cnn.com></http:>; rel="last memento"; datetime="Fri, 22
    Aug 2014 10:43:04 GMT"
Location: http://web.archive.org/web/20030320060210/http://www4.cnn.com/
Server: Tengine/2.0.3
Set-Cookie: wb_total_perf=7554; Expires=Fri, 22-Aug-2014 17:28:57 GMT; Path=/web/http
    ://www.cnn.com/
Transfer-Encoding: chunked
Vary: accept-datetime
X-Archive-Playback: 0
X-Archive-Wayback-Perf: [Ind...
    RobotsRedis: 1, RobotsTotal: 1, Total: 7554]
X-Link-JSON: {"closest":{"wb_url":"http://web.archive.org/web/20030320060210/http://
    www.cnn.com/","timestamp":"20030320060210","status":"200"}}
X-Page-Cache: MISS
```

Go To This URI To Get The Best Memento For The Date You Requested

Then The Browser Gets A Memento Like Any Other Web Page

Diagram labels:

① CLIENT: HEAD <URI-R> Accept-Datetime: DDDDD

SERVER: HTTP/1.1 200 OK Link: <URI-G, rel=timegate>

② CLIENT: GET <URI-G> Accept-Datetime: DDDDD

SERVER: HTTP/1.1 302 Found Location: <URI-M> Vary: Accept-Datetime

③ CLIENT: GET <URI-M>

SERVER: HTTP/1.1 200 OK Memento-Datetime: DDDDD

User running Memento Client

Original Resource (URI-R)

TimeGate (URI-G)

Memento (URI-M)
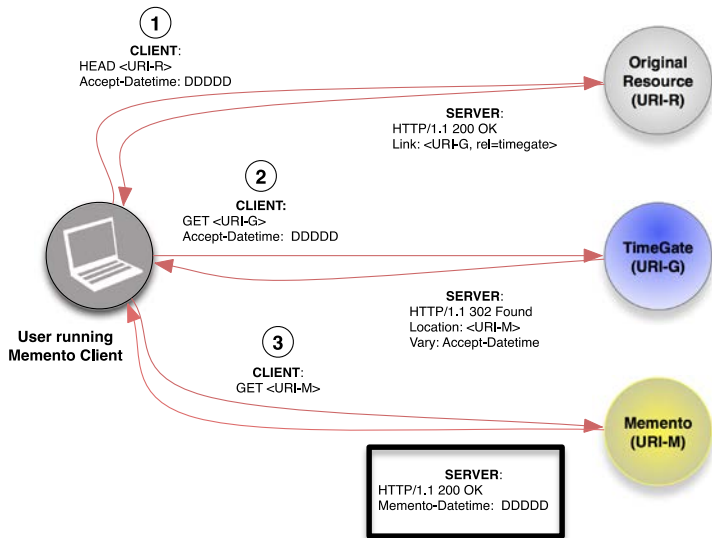
```
HTTP/1.1 200 OK
Server: Tengine/2.0.
Date: Fri, 22 Aug 20
Content-Type: text/h
Content-Length: 80824
Connection: keep-alive
set-cookie: wayback_server=6; Domain=archive.org; Path=/; Expires=Sun, 21-Sep-14
    17:23:10 GMT;
Memento-Datetime: Fri, 01 Jun 2001 04:51:29 GMT
Link: <http://www.cnn.com/>; rel="original", <http://web.archive.org/web/timemap/link/
    http://www.cnn.com/>; rel="timemap"; type="application/link-format", <http://web.
    archive.org/web/http://www.cnn.com/>; rel="timegate", <http://web.archive.org/web
    /20000620180259/http://www.cnn.com/>; rel="first memento"; datetime="Tue, 20 Jun
    2000 18:02:59 GMT", <http://web.archive.org/web/20010601045124/http://www.cnn.com
    />; rel="prev memento"; datetime="Fri, 01 Jun 2001 04:51:24 GMT", <http://web.
    archive.org/web/20010601045129/http://www.cnn.com/>; rel="memento"; datetime="Fri,
     01 Jun 2001 04:51:29 GMT", <http://web.archive.org/web/20010601050038/http://www.
    cnn.com/>; rel="next memento"; datetime="Fri, 01 Jun 2001 05:00:38 GMT", <http://
    web.archive.org/web/20140822104304/http://www.cnn.com/>; rel="last memento";
    datetime="Fri, 22 Aug 2014 10:43:04 GMT"
X-Archive-Guessed-Charset: UTF-8
X-Archive-Orig-server: Netscape-Enterprise/4.1
X-Archive-Orig-expires: Fri, 01 Jun 2001 04:52:29 GMT
X-Archive-Orig-set-cookie: CNNid=cf30
    -2037 16:00:00 GMT; path=/; domai
X-Archive-Orig-date: Fri, 01 Jun 2001
X-Archive-Orig-content-type: text/htm
X-Archive-Orig-last-modified: Fri, 01
X-Archive-Orig-connection: close
X-Archive-Wayback-Perf: [IndexLoad: 677, IndexQueryTotal: 677, RobotsFetchTotal: 3,
    RobotsRedis: 3, RobotsTotal: 3, Total: 974, WArcResource: 139]
Set-Cookie: wb_total_perf=974; Expires=Fri, 22-Aug-2014 17:24:11 GMT; Path=/web
    /20010601045129/http://www.cnn.com/
X-Archive-Playback: 1
X-Page-Cache: MISS
```

This Is The Datetime Of This Memento

Here's Everything I Know About Archives For That Page

Memento Also Provides TimeMaps As A Machine-Readable List Of Mementos

```
<http://lostpedia.wikia.com/wiki/The_Numbers>; rel="original",
<http://web.archive.org/web/timemap/link/http://lostpedia.wikia.com/wiki/The_Numbers>;
     rel="self"; type="application/link-format"; from="Wed, 31 Dec 2008 03:44:05 GMT";
     until="Tue, 01 Jul 2014 04:48:08 GMT",
<http://web.archive.org/web/http://lostpedia.wikia.com/wiki/The_Numbers>; rel="
     timegate",
<http://web.archive.org/web/20081231034405/http://lostpedia.wikia.com/wiki/The_Numbers
     >; rel="first memento"; datetime="Wed, 31 Dec 2008 03:44:05 GMT",
<http://web.archive.org/web/20090119183329/http://lostpedia.wikia.com/wiki/The_Numbers
     >; rel="memento"; datetime="Mon, 19 Jan 2009 18:33:29 GMT",
<http://web.archive.org/web/20090203143759/http://lostpedia.wikia.com/wiki/The_Numbers
     >; rel="memento"; datetime="Tue, 03 Feb 2009 14:37:59 GMT",
<http://web.archive.org/web/20090204193446/http://lostpedia.wikia.com/wiki/The_numbers
     >; rel="memento"; datetime="Wed, 04 Feb 2009 19:34:46 GMT",
....deletia
```

Memento For Chrome
Lets Users "Right-Click Into The Past"

http://ws-dl.blogspot.com/2013/10/2013-10-14-right-click-to-past-memento.html

Mink Lets Users See All Mementos For A Page

http://ws-dl.blogspot.com/2014/10/2014-10-03-integrating-live-and.html

Who Maintains Mementos?

How Are Mementos Acquired?
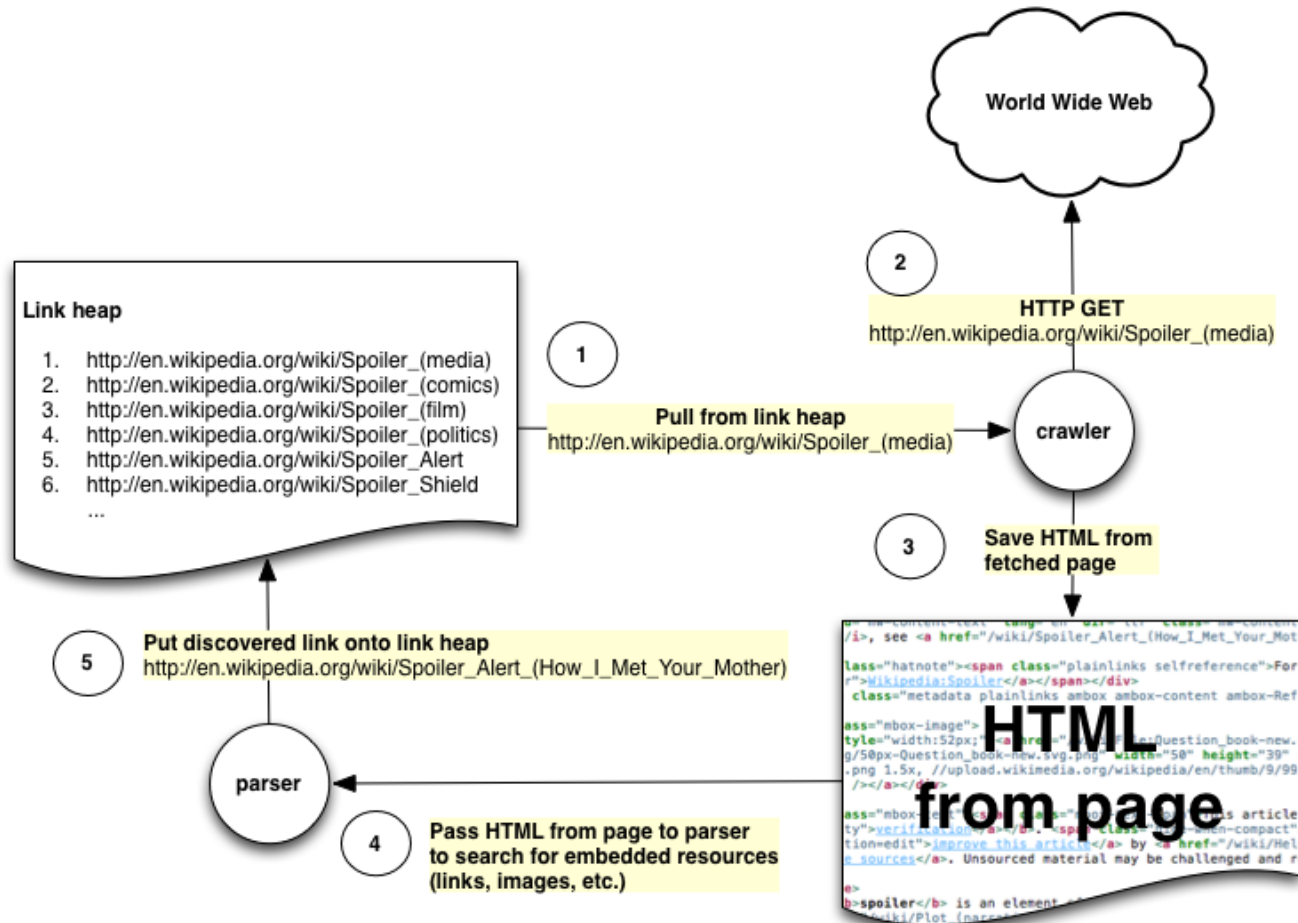
The Internet Archive Provides The Wayback Machine

http://archive.org/web/

Using The Wayback Machine
One Can Select A URI And Datetime...

http://web.archive.org/web/*/http://lostpedia.wikia.com

…And See The Memento From The Internet Archive
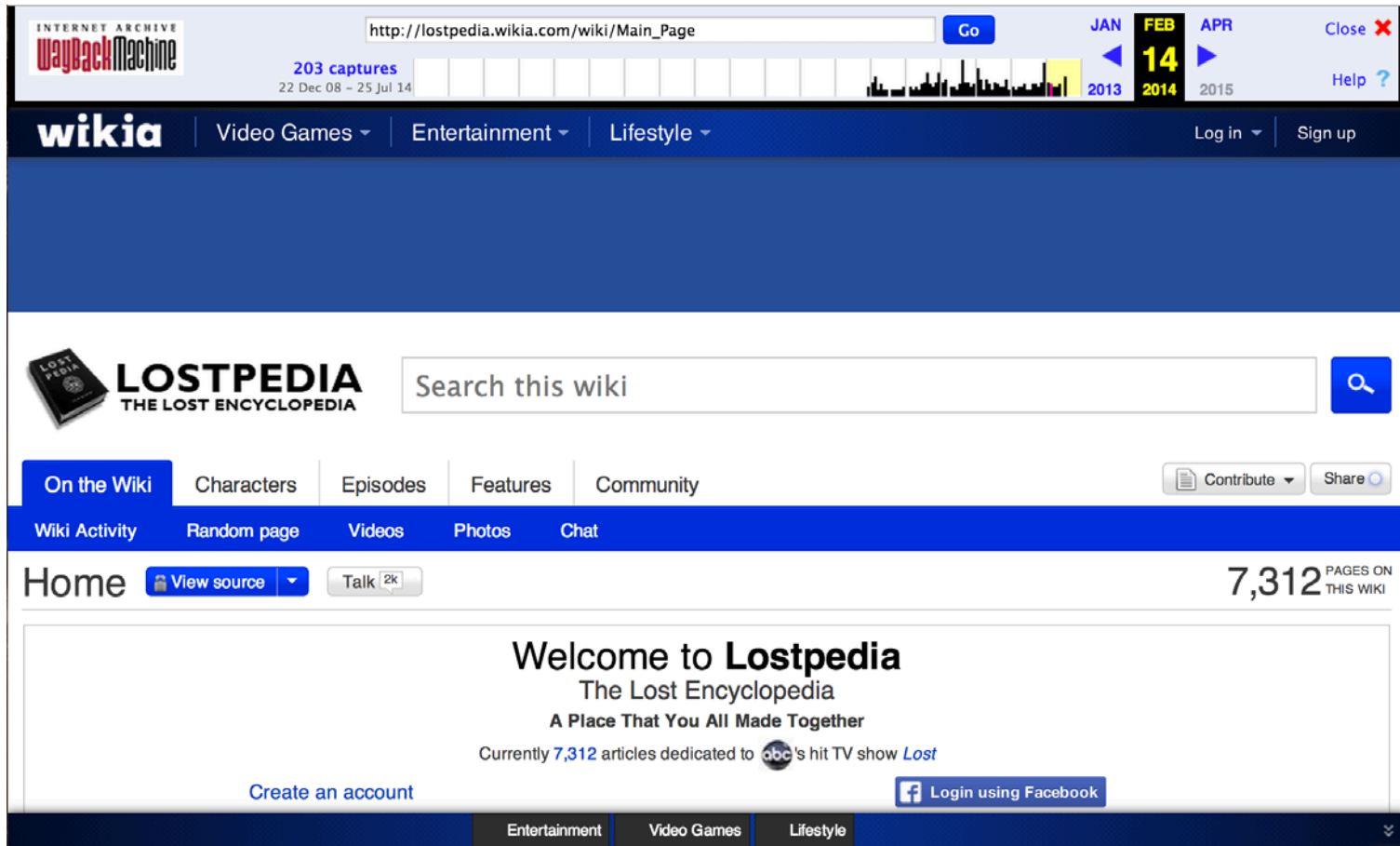
http://web.archive.org/web/20140214065217/http://lostpedia.wikia.com/wiki/Main_Page

The Internet Archive Contains Mementos For Wiki Pages

Here Is An Example Wiki Article…

http://en.wikipedia.org/wiki/Abraham_Lincoln
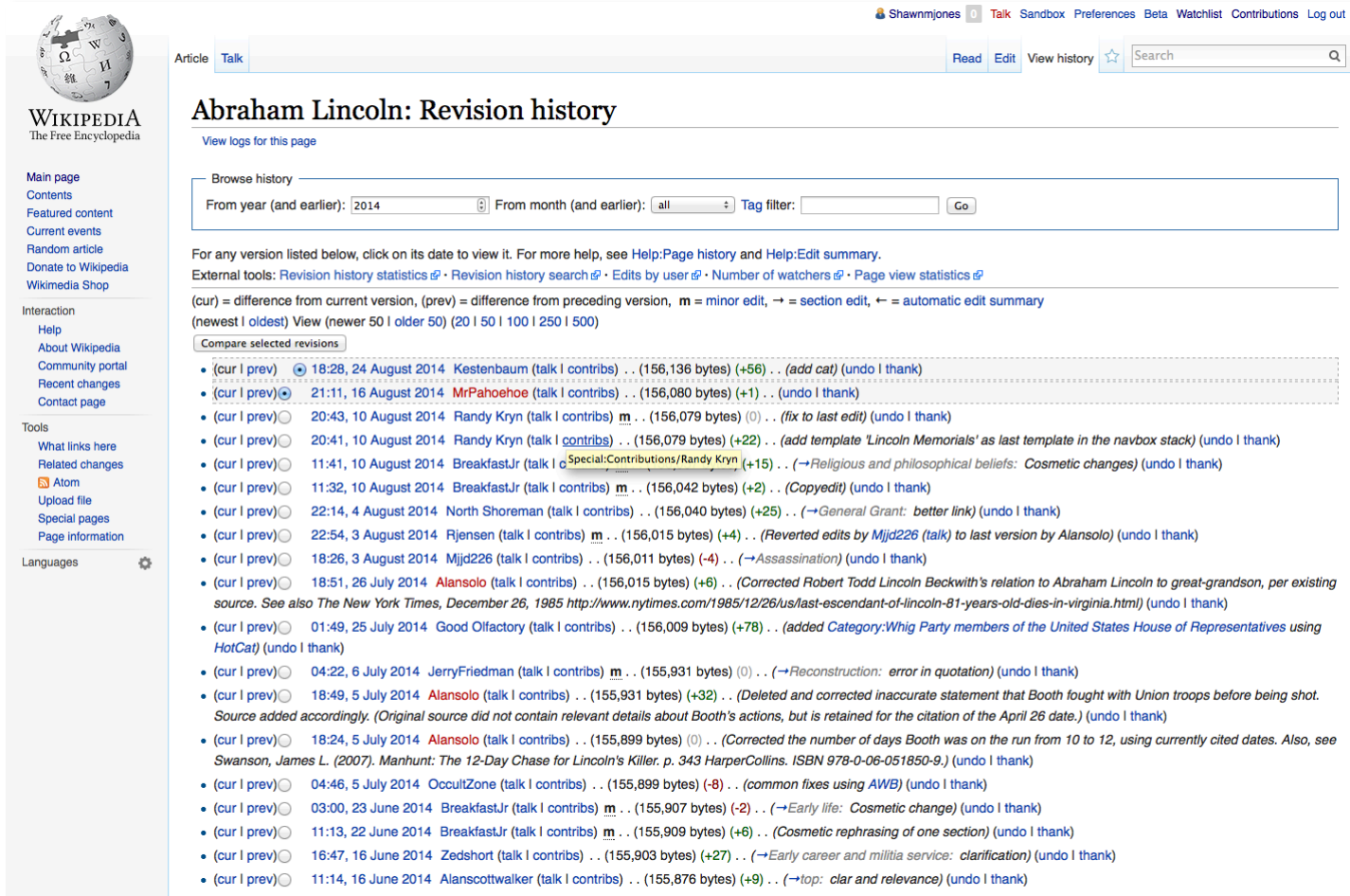
Wikis Are Edited By Multiple Authors, Resulting In Many Revisions…

http://en.wikipedia.org/w/index.php?title=Abraham_Lincoln&action=history

Most Interesting: Wikis Keep
EVERY REVISION OF A PAGE!!!

http://en.wikipedia.org/w/index.php?title=Abraham_Lincoln&action=history

With Every Revision Preserved, One Can Visit Old Revisions Of Pages

This is an **old revision** of this page, as edited by **The Cunctator** (talk | contribs) at 01:01, 4 November 2001 *(*Hooray for Lincoln!)*. It may differ significantly from the **current revision.**

http://en.wikipedia.org/w/index.php?title=Abraham_Lincoln&oldid=345783631

For A Wiki, Every Web Archive Memento Can Be Tied To A Wiki Revision

For A Wiki, Every Web Archive Memento Can Be Tied To A Wiki Revision

We Also Know Which Revisions Were Missed…

Revisions

r_{j-4} r_{j-3} r_{j-2} r_{j-1} r_j

Mementos

m_{k-1} m_k

t_1 t_2 t_3 t_4 t_5 t_6 t_7 t_8 t_9 t_{10} t_{11} t_{12} t_{13} t_{14} t_{15}
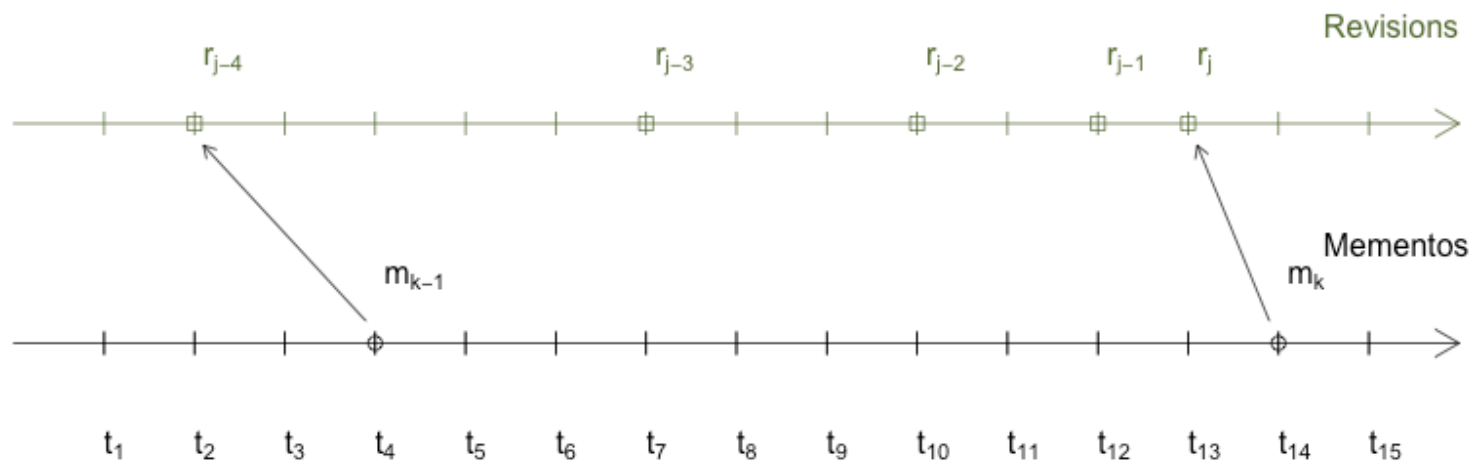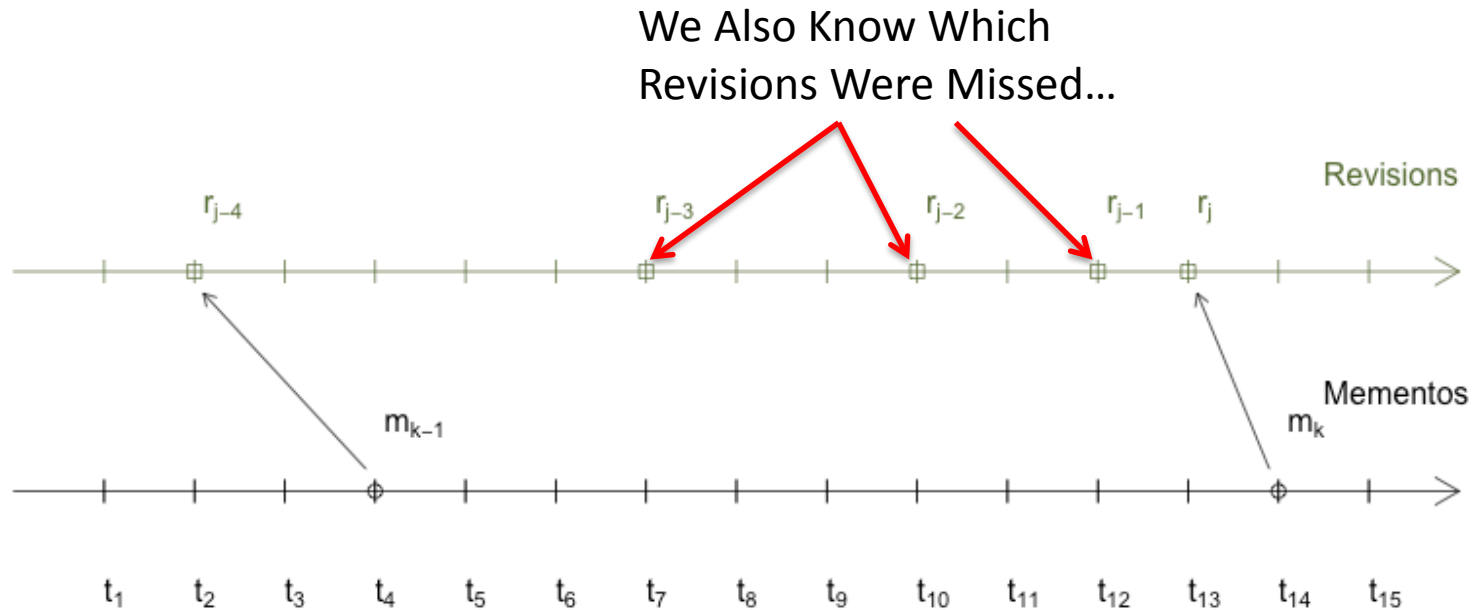
The Web Archives Are Sampling The Wiki Revisions…

Consider The Timeline Of Every Episode In A Series
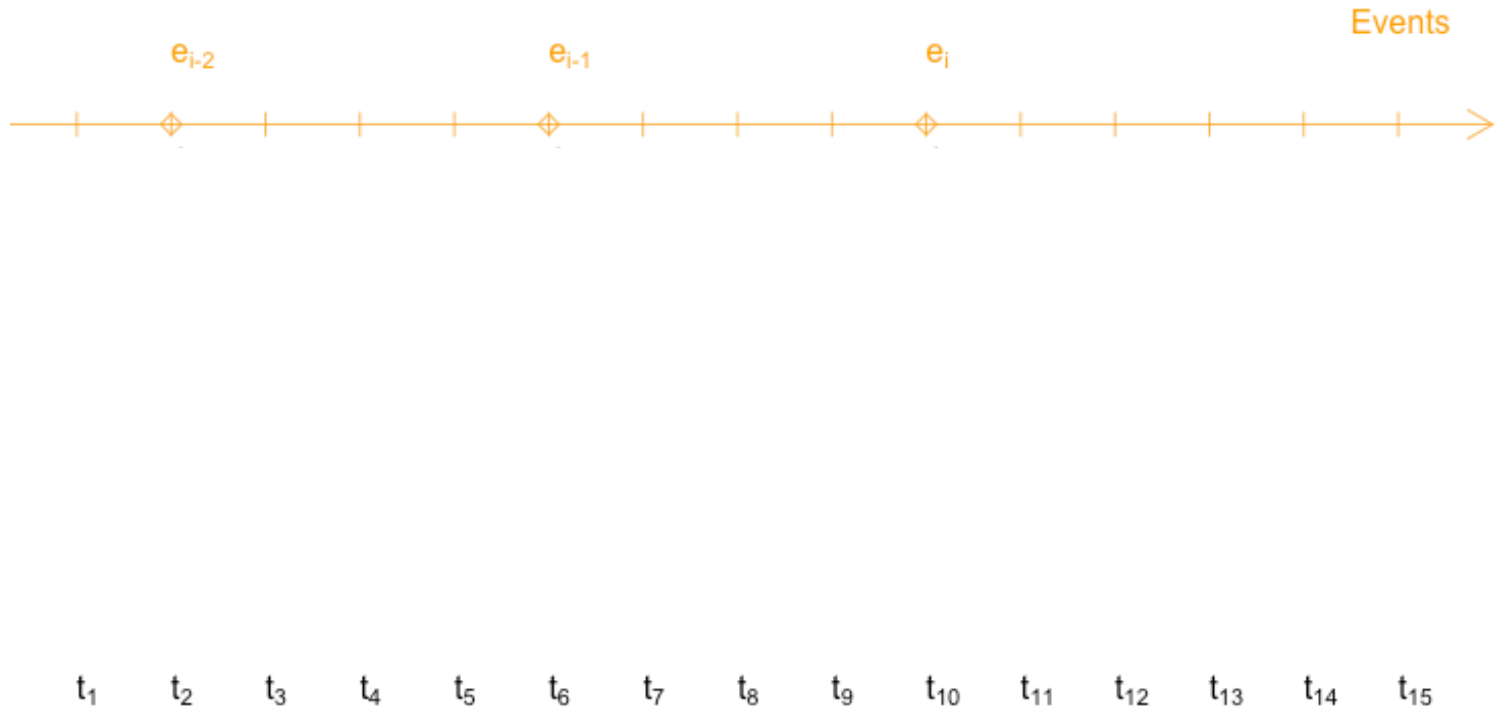
Now Consider A Timeline
Of Wiki Revisions Created By Fans
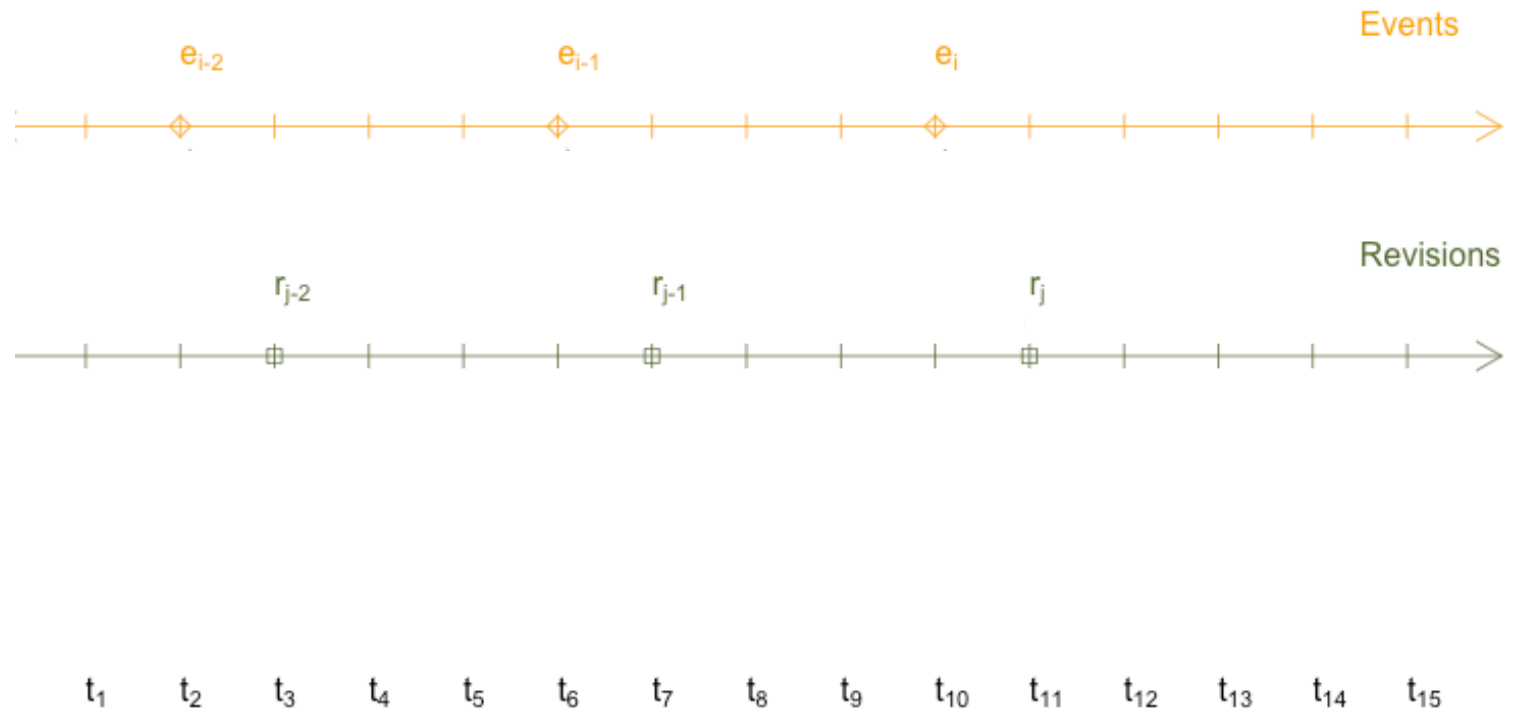
Finally Consider A Third Timeline For Mementos Created From Those Revisions
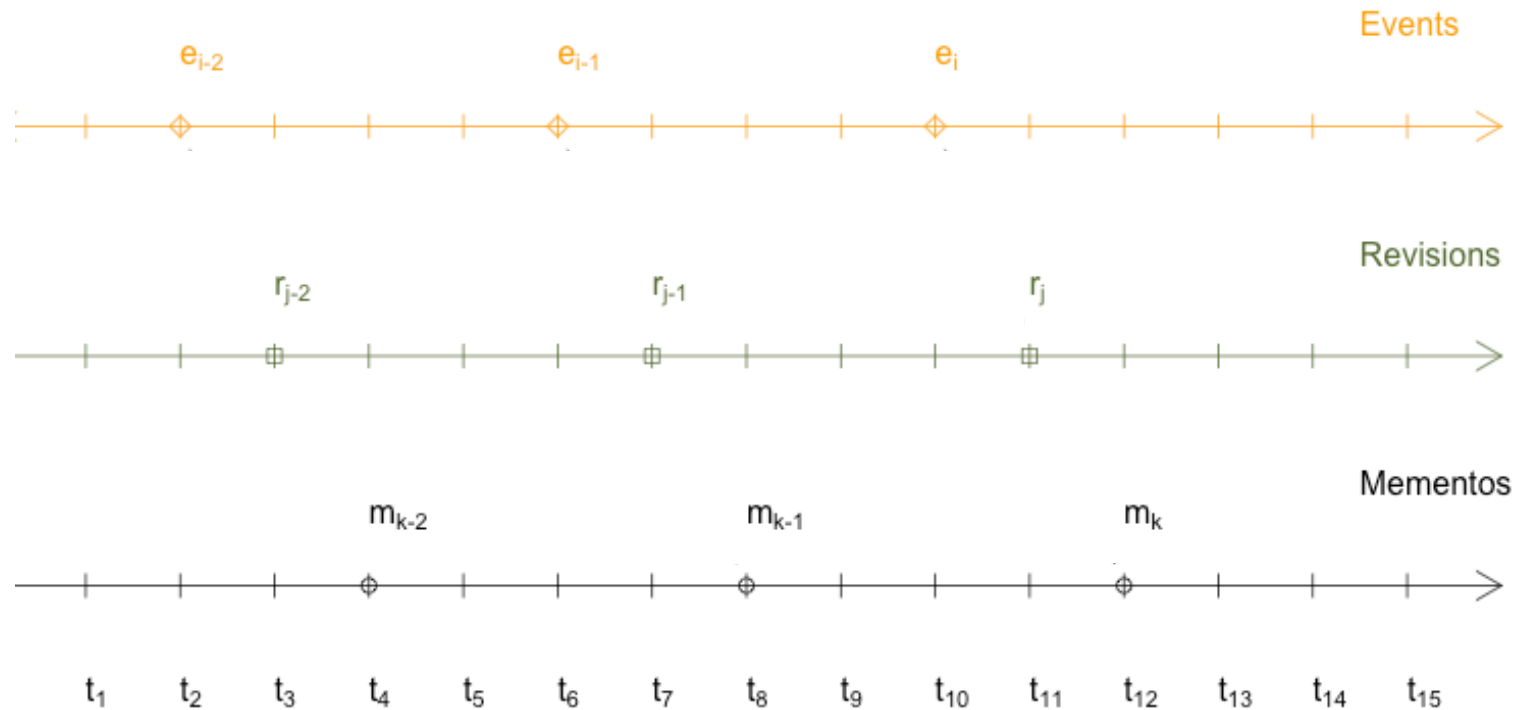
Steiner Noticed That
Events Inspire Wiki Revisions

... or *episodes* inspire wiki edits.

http://arxiv.org/abs/1303.4702

Bringing Us To The Naïve Spoiler Concept: Revisions After An Episode Contain Spoilers

$$t_{r_j} \geq t_{e_i} \implies spoiler$$

$$t_{r_j} < t_{e_i} \implies safe$$

Bringing Us To The Naïve Spoiler Concept: Revisions After An Episode Contain Spoilers

$$t_{r_j} \geq t_{e_i} \implies spoiler$$

$$t_{r_j} < t_{e_i} \implies safe$$

We Are Not The Only Ones Seeking To Avoid Spoilers…

RELATED WORK

So, What Have Others Done About Spoilers?

Hidden Text!!!

TO EVERYONE WHO WILL NOT RECEIVE *SERIES 4* OF DOWNTON ABBEY UNTIL JANUARY 2014, THIS IS A **WARNING**: THIS WIKI WILL BE UPDATED AND EDITED AS THE EPISODES BROADCAST IN THE UNITED KINGDOM (SEPTEMBER - DECEMBER 2013)! PROCEED AT YOUR OWN RISK! **THERE WILL BE SPOILERS ABUNDANT ON MOST OF THE PAGES!** YOU HAVE BEEN WARNED!

Spoiler Notices!!!

- Will They or Won't They?: Jim and Pam. They did, sometime between seasons 3 and 4).
  - Andy and Erin.
  - Michael and Holly also seemed poised to become this until they began going out a few episodes in (which Michael hilariously lampshades):
    Michael: "Why are you helping her? You're not even dating." She's my friend... and... ultimately my strategy is to merge this into a relationship... without her even knowing. They will.
  - Andy and Erin have finally gotten together too.
    - Although Erin dumps Andy in season 9.

Blocking Social Media Posts!
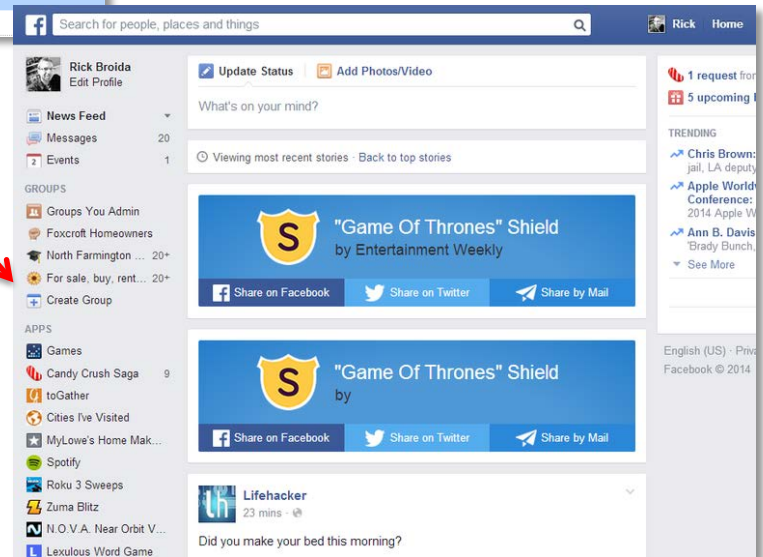
Existing Academic Studies Have Dealt With Social Media

- Separate studies conducted by Johns and the team of Schirra, Sun, and Bently studied two-screen viewing
- Boyd-Graber, Glasgow, and Zajac attempted to use machine learning to find spoilers in social media
- To avoid spoilers fans would avoid, or abandon:
  - Social media
  - Online web pages
  - TV shows

- This results in lost revenue to advertisers!

Our Goals

- Work with wikis, not social media

- Not just warn the user!

- Not hide the data!

- Show the user what existed before the spoiler was revealed, **so the resource is still useful**.

TimeGates Hold The Key To Avoiding Spoilers

TIMEGATE HEURISTICS

TimeGates Can Be Represeted By A Function

$$M = \mathcal{G}^h(R, t_a)$$

M = memento returned (URI-M)
R = original resource (URI-R)
t_a = desired datetime

h = heuristic being used for TimeGate

mindist Is The Most Widely Used Heuristic

Minimum Distance From t_a With No Bounds

$$\mathcal{G}^{mindist}(R, t_8) = m_2@t_7$$

Mementos

| m_1 | | m_2 t_a | m_3 | m_4 |

t_1 t_2 t_3 t_4 t_5 t_6 t_7 t_8 t_9 t_{10} t_{11} t_{12} t_{13} t_{14} t_{15}

$$\mathcal{G}^{mindist}(R, t_9) = m_3@t_{10}$$

Mementos

| m_1 | | m_2 | t_a m_3 | m_4 |

t_1 t_2 t_3 t_4 t_5 t_6 t_7 t_8 t_9 t_{10} t_{11} t_{12} t_{13} t_{14} t_{15}
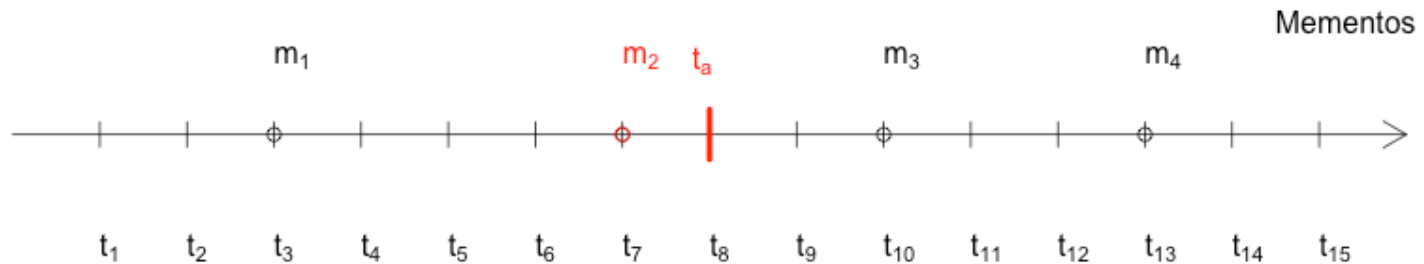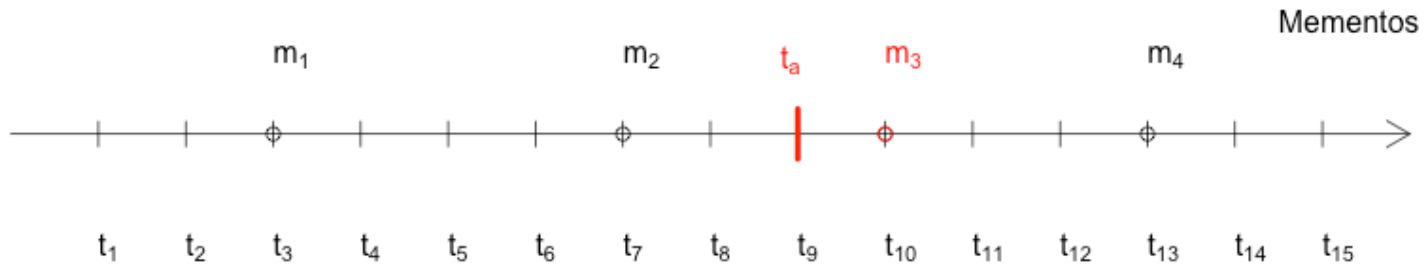
But mindist Can Lead To Spoilers…

Minimum Distance – no bounds

$$\mathcal{G}^{mindist}(R, t_8) = m_2@t_7$$

Mementos

m_1   m_2 t_a   m_3   m_4

t_1 t_2 t_3 t_4 t_5 t_6 t_7 t_8 t_9 t_10 t_11 t_12 t_13 t_14 t_15

$$\mathcal{G}^{mindist}(R, t_9) = m_3@t_{10}$$

spoiler

Mementos

m_1   m_2   t_a   m_3   m_4

t_1 t_2 t_3 t_4 t_5 t_6 t_7 t_8 t_9 t_10 t_11 t_12 t_13 t_14 t_15
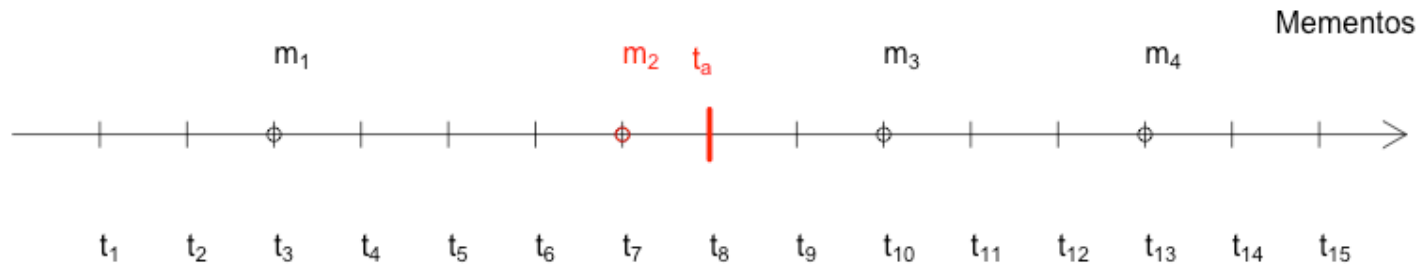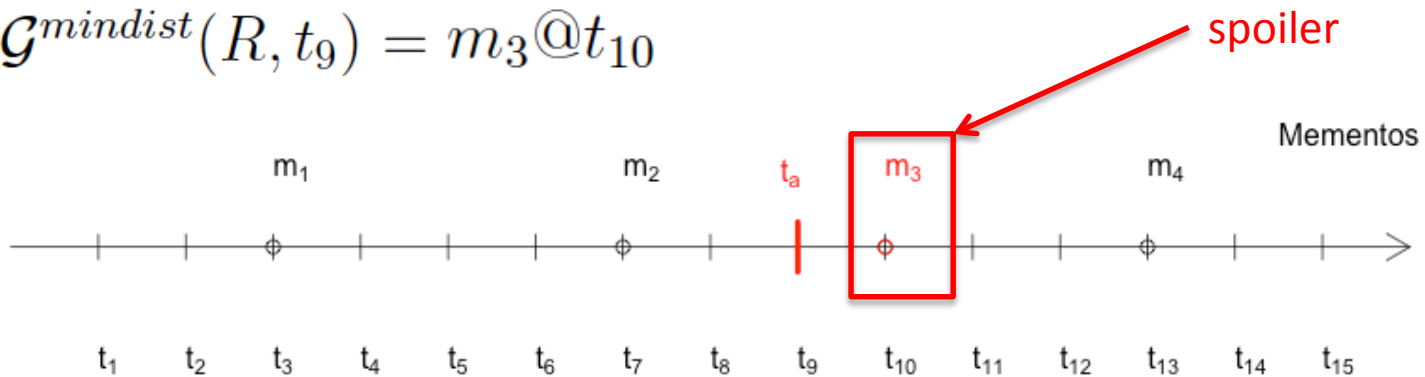
minpast Does Not Lead To Spoilers

Minimum Distance In The Past Where Upper Bound = t_a
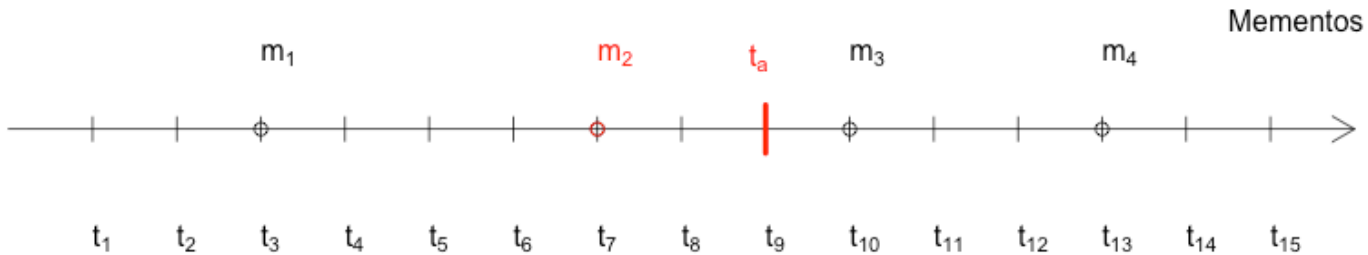
$$\mathcal{G}^{minpast}(R, t_8) = m_2 @ t_7$$

minfutr Is The Opposite Of minpast

Minimum Distance In The Future Where Lower Bound = t_a
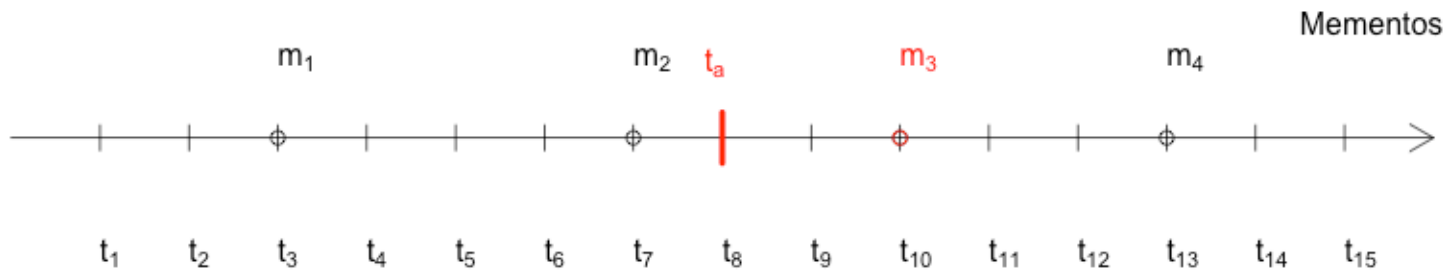
$$\mathcal{G}^{minfutr}(R, t_8) = m_3 @ t_{10}$$

minfutr Always Leads To Spoilers

Minimum Distance In The Future Where Lower Bound = t_a
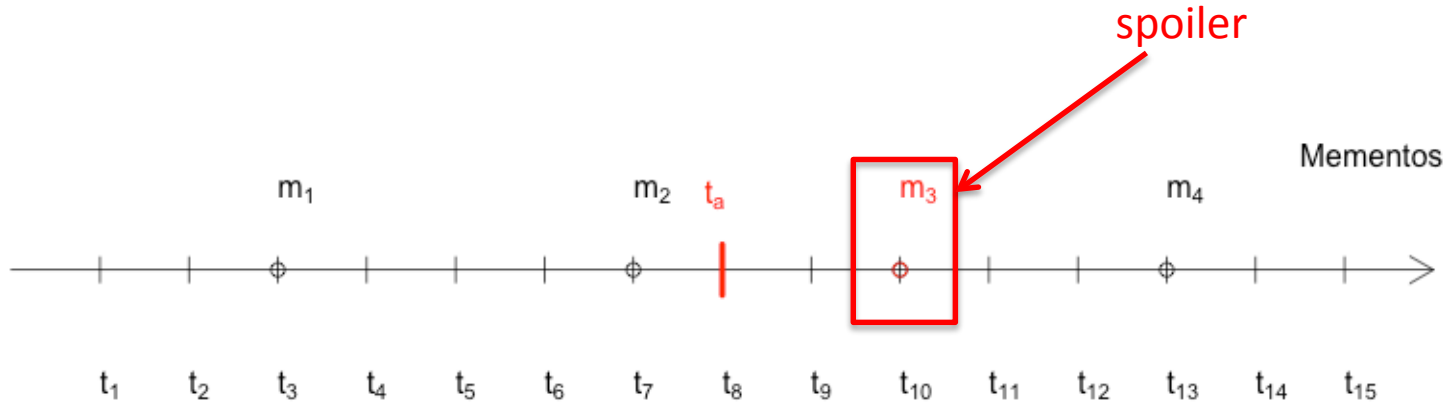
$$\mathcal{G}^{minfutr}(R, t_8) = m_3 @ t_{10}$$

Other Heuristics Can Lead To Spoilers

- **minnear** – bounds specified by user/system
- **eqpast** – compare on both sides of t_a, pick past if equal, mindist if not
- **eqfutr** – compare on both sides of t_a, pick future if equal, mindist if not
- **simpast** – compare on both sides of t_a, pick past if similar, mindist if not
- **simfutr** – compare on both sides of t_a, pick future if similar, mindist if not

We Compared These Heuristics Based On Performance And Spoiler Avoidance

TABLE 7: Summary of TimeGate Heuristics

Heuristic	Uses cost metric	Uses Content of Mementos	Potential running time	Reliably avoids spoilers?
mindist	X		$O(n)$	no
minpast	X		$O(n)$	yes
minfutr	X		$O(n)$	no
minnear	X		$O(n)$	no
eqfutr	X	X	$O(n + c)$	no
eqpast	X	X	$O(n + c)$	no
simfutr	X	X	$O(f + n)$	no
simpast	X	X	$O(f + n)$	no

Even If Modified To Default To minpast, These Heuristics Perform More Poorly

TABLE 7: Summary of TimeGate Heuristics

Heuristic	Uses cost metric	Uses Content of Mementos	Potential running time	Reliably avoids spoilers?
mindist	X		$O(n)$	no
minpast	X		$O(n)$	yes
minfutr	X		$O(n)$	no
minnear	X		$O(n)$	no
eqfutr	X	X	$O(n+c)$	no
eqpast	X	X	$O(n+c)$	no
simfutr	X	X	$O(f+n)$	no
simpast	X	X	$O(f+n)$	no

minpast Is Best For Avoiding Spoilers

TABLE 7: Summary of TimeGate Heuristics

Heuristic	Uses cost metric	Uses Content of Mementos	Potential running time	Reliably avoids spoilers?
mindist	X		$O(n)$	no
minpast	X		$O(n)$	yes
minfutr	X		$O(n)$	no
minnear	X		$O(n)$	no
eqfutr	X	X	$O(n+c)$	no
eqpast	X	X	$O(n+c)$	no
simfutr	X	X	$O(f+n)$	no
simpast	X	X	$O(f+n)$	no

Using mindist Can Be Hazardous For Avoiding Spoilers

THEORY OF SPOILER PROBABILITY

Remember: Consider The Timeline Of Every Episode In A Series

Events

e_{i-2} e_{i-1} e_i

t_1 t_2 t_3 t_4 t_5 t_6 t_7 t_8 t_9 t_{10} t_{11} t_{12} t_{13} t_{14} t_{15}

Remember: Consider A Timeline Of Wiki Revisions Created By Fans

Remember: Consider A Third Timeline For Mementos Created From Those Revisions

Remember Our Naïve Spoiler Concept

$$t_{r_j} \geq t_{e_i} \implies spoiler$$

$$t_{r_j} < t_{e_i} \implies safe$$

Spoiler Areas Exist Where mindist Returns The User A Memento From The Future Of Their Desired Datetime
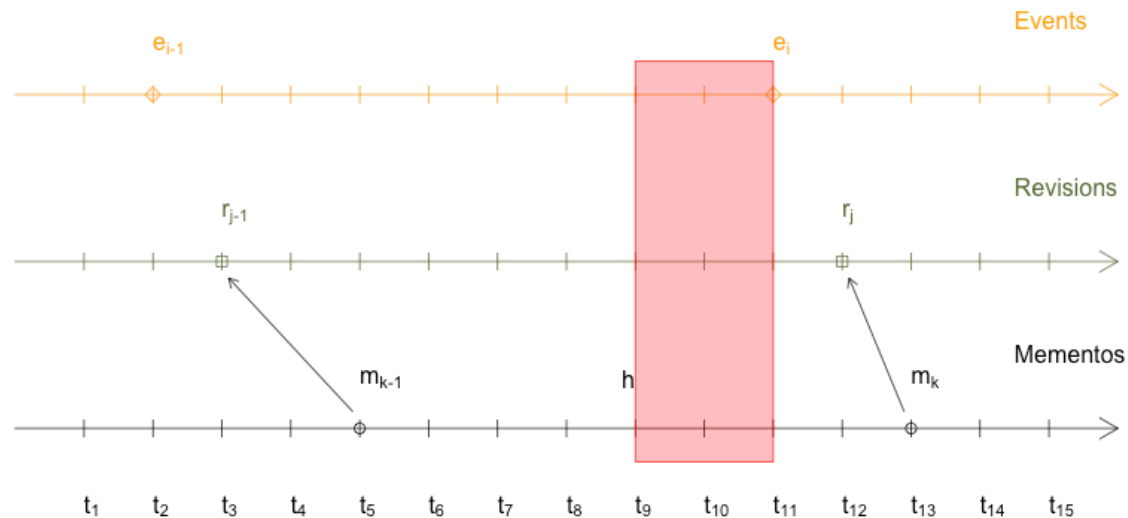
In this example, using mindist, if the user requests a memento with a datetime between t_9 and t_{11}, denoted by the red area, they will get m_k (which is r_j), which exists in the future, even though they chose a datetime **before** m_k!

Conditions Arising From These Timelines Are Defined By When They Occur

- Pre-Archive – occurs prior to first memento

- Archive-Extant – occurs after first memento

- Post-Archive – occurs after last memento

Condition: Pre Archive Spoiler Area

Spoiler area exists for e_3, and one still exists for e_2 because m_1 maps to r_j, which is after e_3
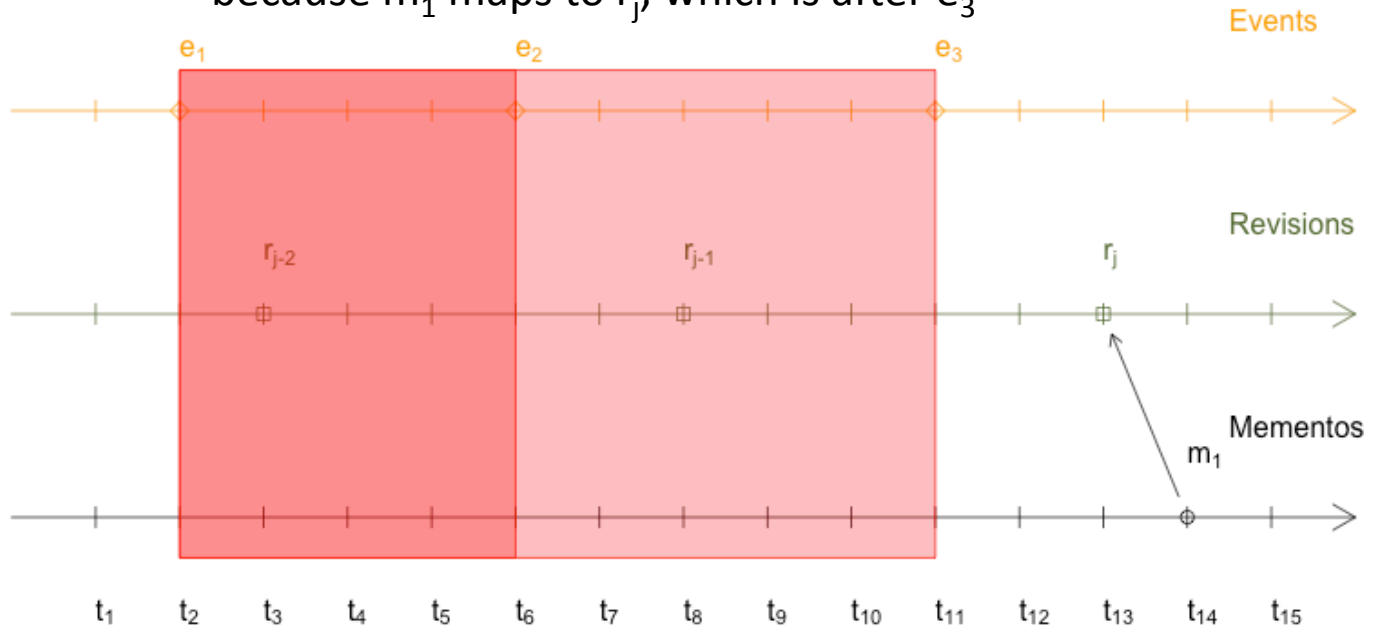
$$[t_s, t_f] = \mathcal{S}_a(e_i) = \begin{cases} (t_{e_1}, t_{e_i}) & \text{if} & t_{e_i} < t_{r_j} \wedge r_j \equiv m_k \\ (0,0) & \text{otherwise} \end{cases}$$

Condition: Pre Archive Safe Area

No spoiler area for e_3, but one still exists for e_2
because m_1 maps to r_j, but r_j is before e_3
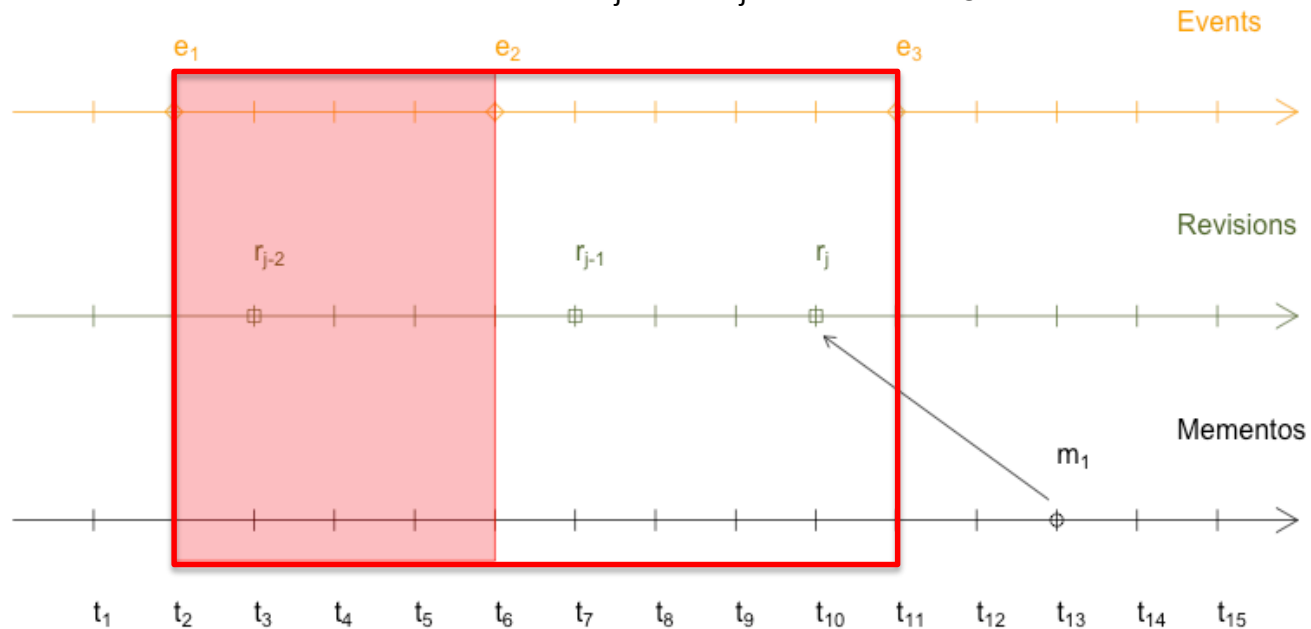
$$[t_s, t_f] = \mathcal{S}_a(e_i) = \begin{cases} (t_{e_1}, t_{e_i}) & \text{if} \quad t_{e_i} < t_{r_j} \wedge r_j \equiv m_k \\ (0, 0) & \text{otherwise} \end{cases}$$

Condition: Archive Extant Spoiler Area
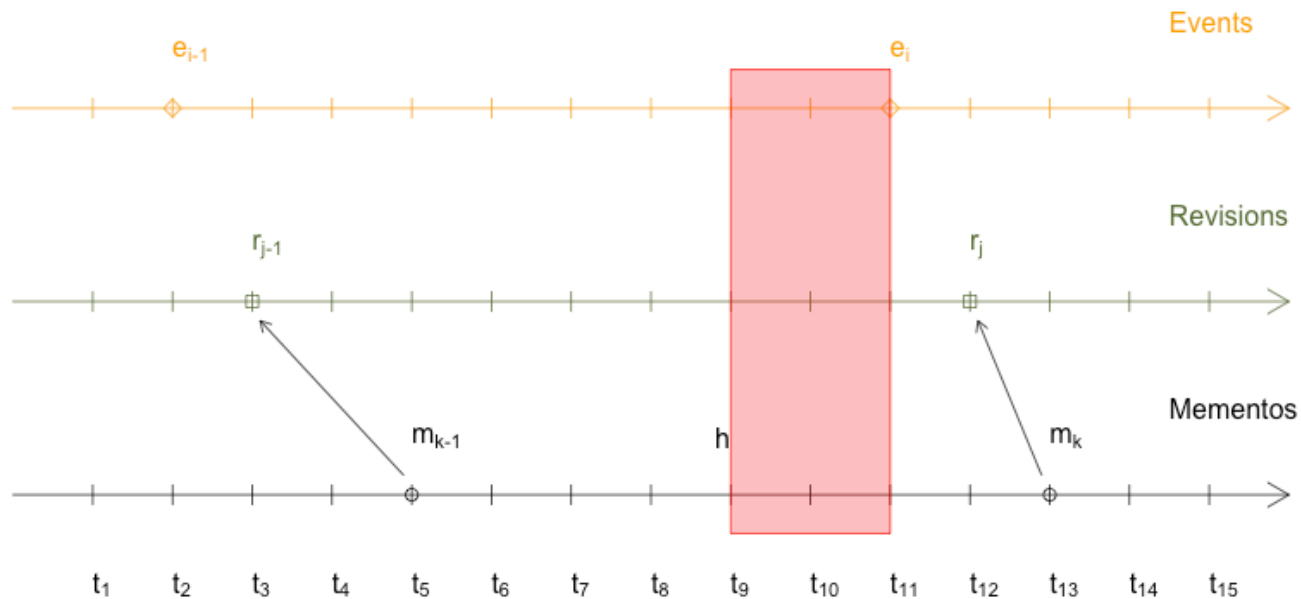
$$[t_s, t_f] = \mathcal{S}_b(e) = \begin{cases} (t_h, t_e) & \text{if} \qquad t_h < t_e < t_{r_i} \ \wedge \ r_j \equiv m_k \ \wedge \\ & \qquad\qquad t_h = \frac{t_{m_{k-1}} + t_{m_k}}{2} \\ (0, 0) & \text{otherwise} \end{cases}$$

Condition: Archive Extant Safe Area EHR
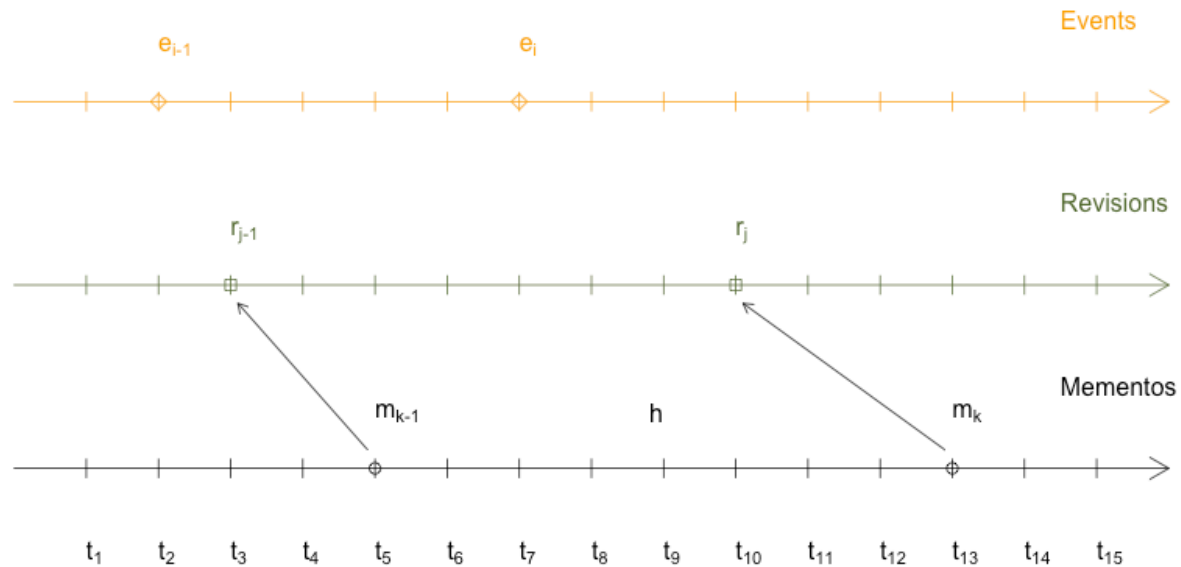
$$[t_s, t_f] = \mathcal{S}_b(e) = \begin{cases} (t_h, t_e) & \text{if} & t_h < t_e < t_{r_i} \ \wedge \ r_j \equiv m_k \ \wedge \\ & & t_h = \frac{t_{m_{k-1}} + t_{m_k}}{2} \\ (0,0) & \text{otherwise} \end{cases}$$

Condition: Archive Extant Safe Area HRE
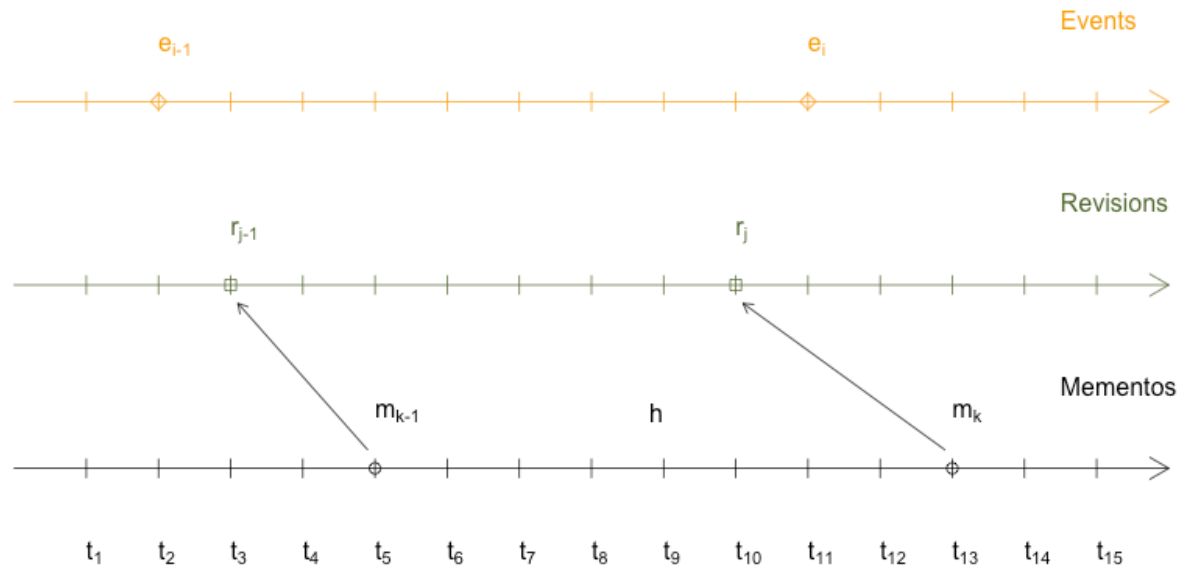
$$[t_s, t_f] = \mathcal{S}_b(e) = \begin{cases} (t_h, t_e) & \text{if} \quad t_h < t_e < t_{r_i} \ \wedge \ r_j \equiv m_k \ \wedge \\ & \quad t_h = \frac{t_{m_{k-1}} + t_{m_k}}{2} \\ (0, 0) & \text{otherwise} \end{cases}$$

The Area Between the First and Last Episodes Is A Potential Spoiler Zone

Using Spoiler Areas And A Potential Spoiler Zone, We Can Calculate The Probability Of Spoiler For A Page

$$Pr(spoiler) = \frac{s}{c}$$

s = # of seconds where we are in a spoiler area
c = # of seconds between e_1 and e_n

Actual Spoiler Areas From Actual Wikis

MEASUREMENTS OF SPOILER PROBABILITY

Sixteen Lucky Fan Wikis Were Selected From wikia.com…

TABLE 9: Fan wikis used in the spoiler areas experiment

Television Show (Network)	Wiki URI .wikia.com	# of Pages	t_{r_1}	t_{e_1}	% of pages in Internet Archive
the Big Bang Theory (CBS)	bigbangtheory	1120	2007-12-14	2007-09-24	68.8%
Boardwalk Empire (HBO)	boardwalkempire	2091	2010-03-18	2010-08-23	80.6%
Breaking Bad (A&E)	breakingbad	998	2009-04-27	2008-01-20	76.0%
Continuum (Showcase)	continuum	258	2012-11-13	2012-05-27	86.8%
Downton Abbey (BBC)	downtonabbey	784	2010-10-04	2010-09-26	53.1%
Game of Thrones (HBO)	gameofthrones	3144	2010-06-24	2011-04-17	75.8%
Grimm (NBC)	grimm	1581	2010-04-14	2011-10-28	57.5%
House of Cards (Netflix)	house-of-cards	251	2013-01-11	2013-02-01	97.2%
How I Met Your Mother (CBS)	how-i-met-your-mother	1709	2008-07-21	2005-09-19	58.7%
Lost (ABC)	lostpedia	18790	2005-09-22	2004-09-22	39.1%
Mad Men (AMC)	madmen	652	2009-07-25	2007-06-03	85.0%
NCIS (CBS)	ncis	5345	2006-09-25	2003-09-23	93.2%
Once Upon A Time (ABC)	onceuponatime	1470	2011-08-09	2011-10-23	79.9%
Scandal (ABC)	scandal	331	2011-06-07	2012-04-05	82.8%
True Blood (HBO)	trueblood	1838	2008-10-06	2008-09-07	74.1%
White Collar (USA)	whitecollar	506	2009-10-30	2009-10-23	79.1%

We See That Not All Pages Are Archived

TABLE 9: Fan wikis used in the spoiler areas experiment

Television Show (Network)	Wiki URI .wikia.com	# of Pages	t_{r_1}	t_{e_1}	% of pages in Internet Archive
the Big Bang Theory (CBS)	bigbangtheory	1120	2007-12-14	2007-09-24	68.8%
Boardwalk Empire (HBO)	boardwalkempire	2091	2010-03-18	2010-08-23	80.6%
Breaking Bad (A&E)	breakingbad	998	2009-04-27	2008-01-20	76.0%
Continuum (Showcase)	continuum	258	2012-11-13	2012-05-27	86.8%
Downton Abbey (BBC)	downtonabbey	784	2010-10-04	2010-09-26	53.1%
Game of Thrones (HBO)	gameofthrones	3144	2010-06-24	2011-04-17	75.8%
Grimm (NBC)	grimm	1581	2010-04-14	2011-10-28	57.5%
House of Cards (Netflix)	house-of-cards	251	2013-01-11	2013-02-01	97.2%
How I Met Your Mother (CBS)	how-i-met-your-mother	1709	2008-07-21	2005-09-19	58.7%
Lost (ABC)	lostpedia	18790	2005-09-22	2004-09-22	39.1%
Mad Men (AMC)	madmen	652	2009-07-25	2007-06-03	85.0%
NCIS (CBS)	ncis	5345	2006-09-25	2003-09-23	93.2%
Once Upon A Time (ABC)	onceuponatime	1470	2011-08-09	2011-10-23	79.9%
Scandal (ABC)	scandal	331	2011-06-07	2012-04-05	82.8%
True Blood (HBO)	trueblood	1838	2008-10-06	2008-09-07	74.1%
White Collar (USA)	whitecollar	506	2009-10-30	2009-10-23	79.1%

None of these are 100%

In fact, 38% of the total is not available in the Internet Archive.

For Pages With Mementos
We Sought To Find Spoiler Areas

Revisions from Wiki export pages

Episode Dates From epguides.com

Mementos from Internet Archive TimeMaps

Pre-Archive Equation

$$[t_s, t_f] = \mathcal{S}_a(e_i) = \begin{cases} (t_{e_1}, t_{e_i}) & \text{if} \quad t_{e_i} < t_{r_j} \wedge r_j \equiv m_k \\ (0, 0) & \text{otherwise} \end{cases}$$

Archive-Extant Equation

$$[t_s, t_f] = \mathcal{S}_b(e) = \begin{cases} (t_h, t_e) & \text{if} \quad t_h < t_e < t_{r_i} \wedge r_j \equiv m_k \wedge \\ & \quad t_h = \frac{t_{m_{k-1}} + t_{m_k}}{2} \\ (0, 0) & \text{otherwise} \end{cases}$$

List of Spoiler Areas $[\ (t_{s1}, t_{f1}), (t_{s2}, t_{f2}), \ldots (t_{sn}, t_{fn})\ ]$

Spoiler Areas Do Exist!

Spoiler Areas for http://bigbangtheory.wikia.com/wiki/Sheldon_Cooper
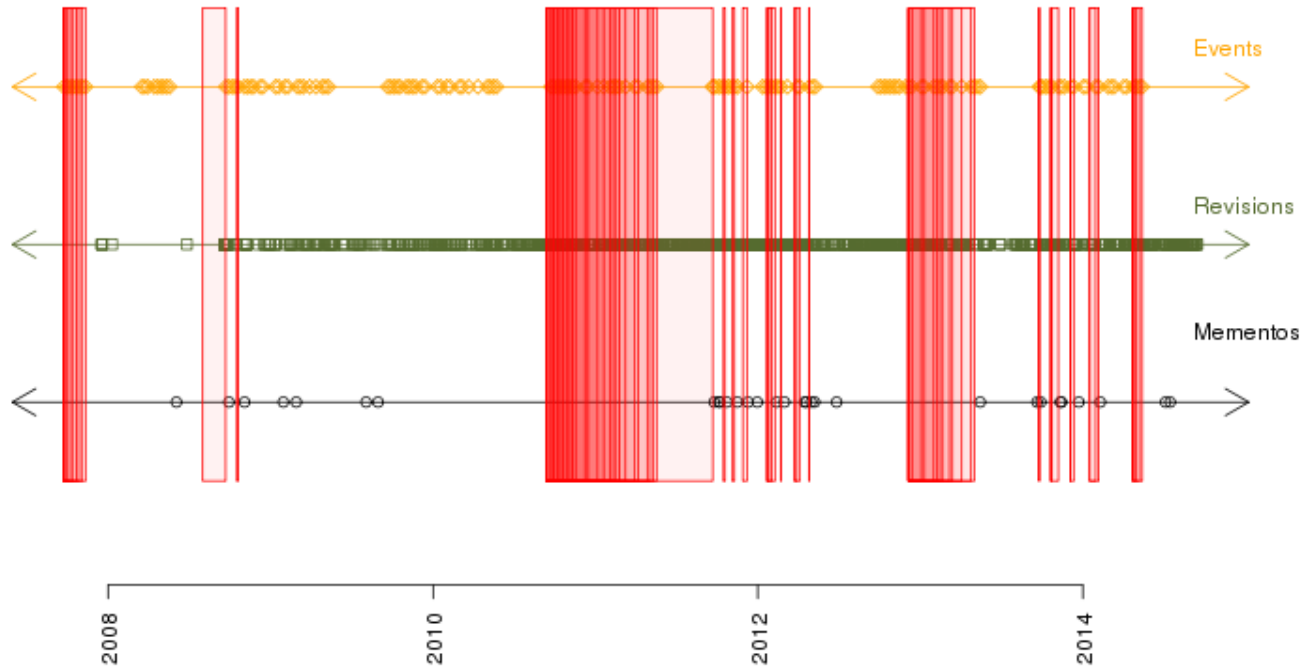
Spoiler Areas Do Exist!

Spoiler Areas for http://bigbangtheory.wikia.com/wiki/Sheldon_Cooper

One Can See Whole Seasons Clumped Together
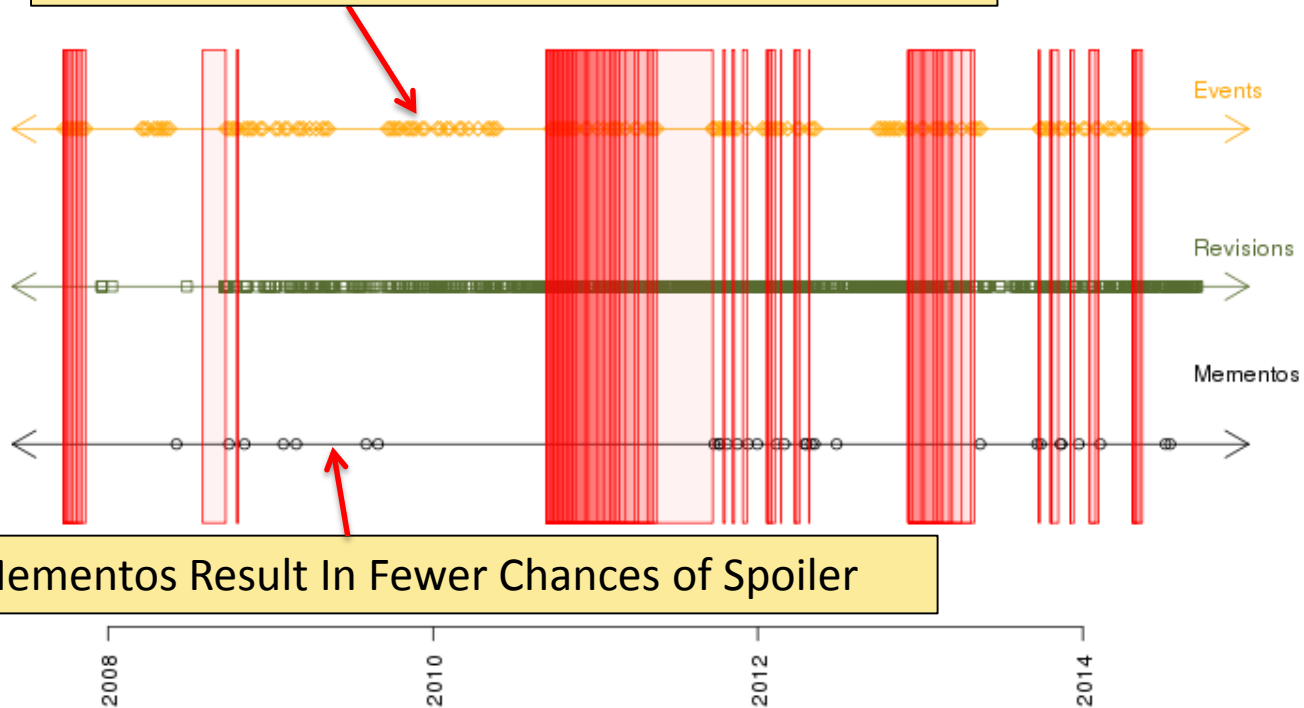
Fewer Mementos Result In Fewer Chances of Spoiler

We Discovered Different Categories of Pages

- Normal
- Wiki-Before-Show
- Season-In-A-Day

Normal Pages
Are Started After The Series Starts

Spoiler Areas for http://lostpedia.wikia.com/wiki/Kate_Austen
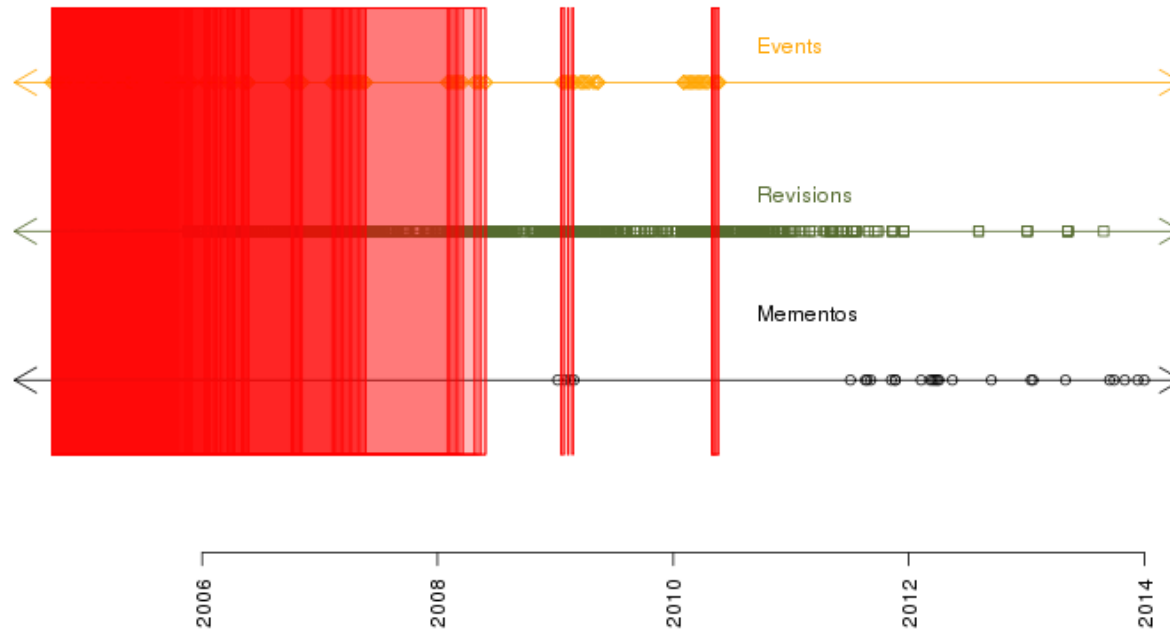
Normal Pages
Are Started After The Series Starts

Spoiler Areas for http://lostpedia.wikia.com/wiki/Kate_Austen

Pre-Archive Spoiler Areas Begin At First Episode

Events

Revisions

Mementos

2006　2008　2010　2012　2014

LOSTPEDIA
THE LOST ENCYCLOPEDIA
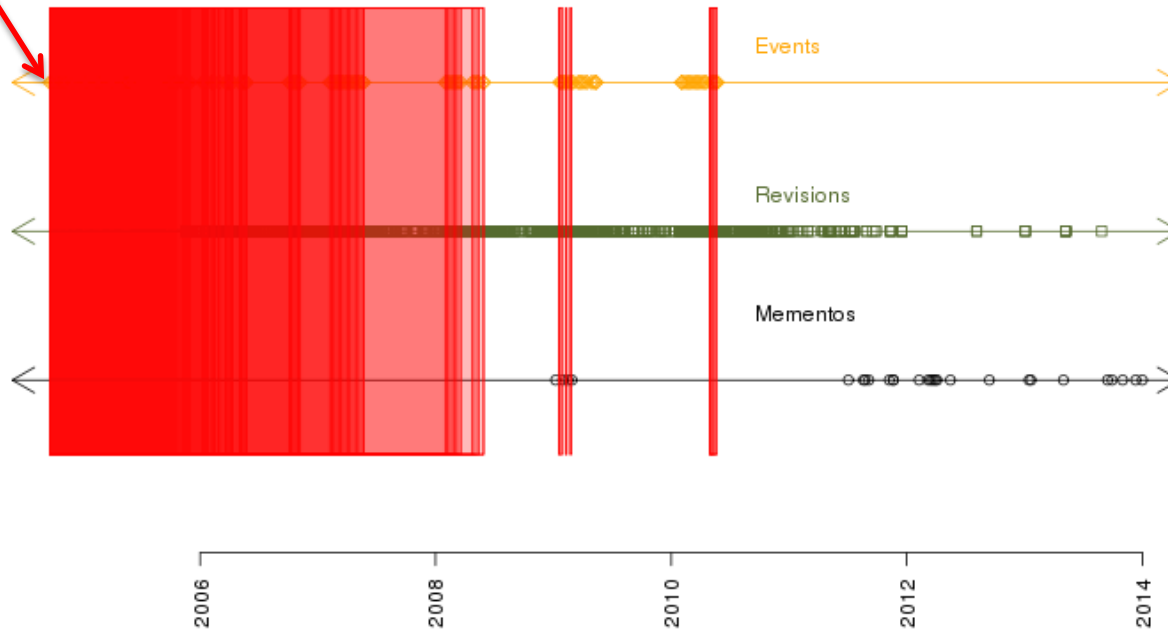
Wiki-Before-Show Pages
Are Started Prior To The First Episode

Spoiler Areas for http://gameofthrones.wikia.com/wiki/Daenerys_Targaryen
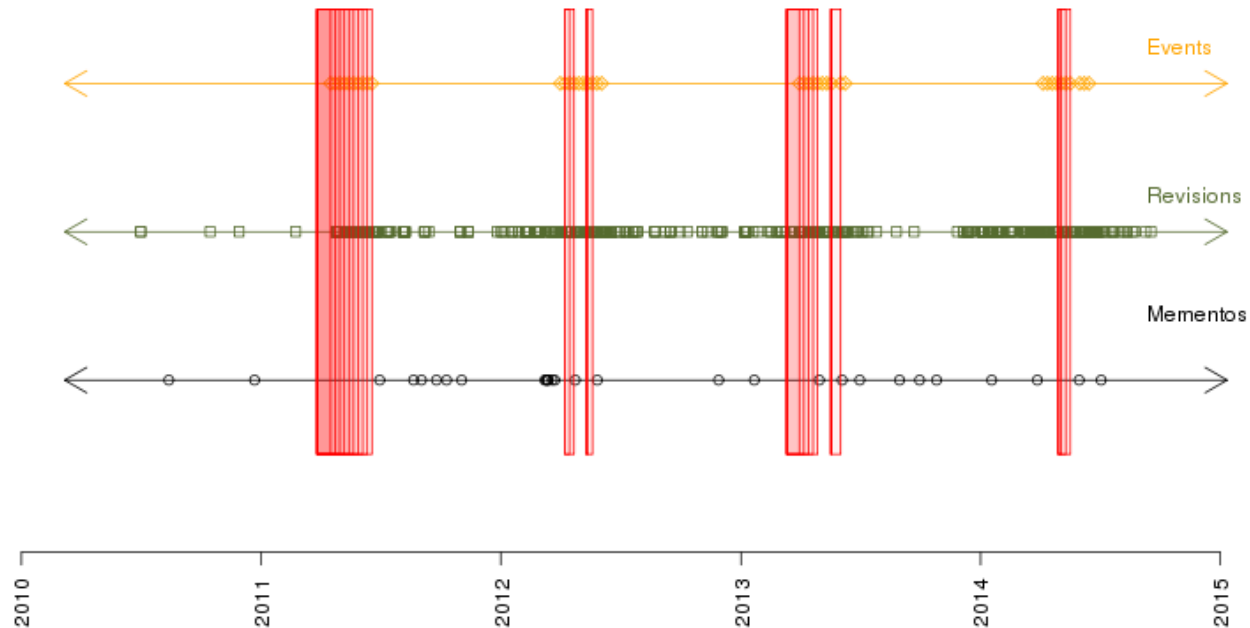
Season-in-a-Day Pages
Have Many Episodes In A Single Day

Spoiler Areas for http://house-of-cards.wikia.com/wiki/Frank_Underwood
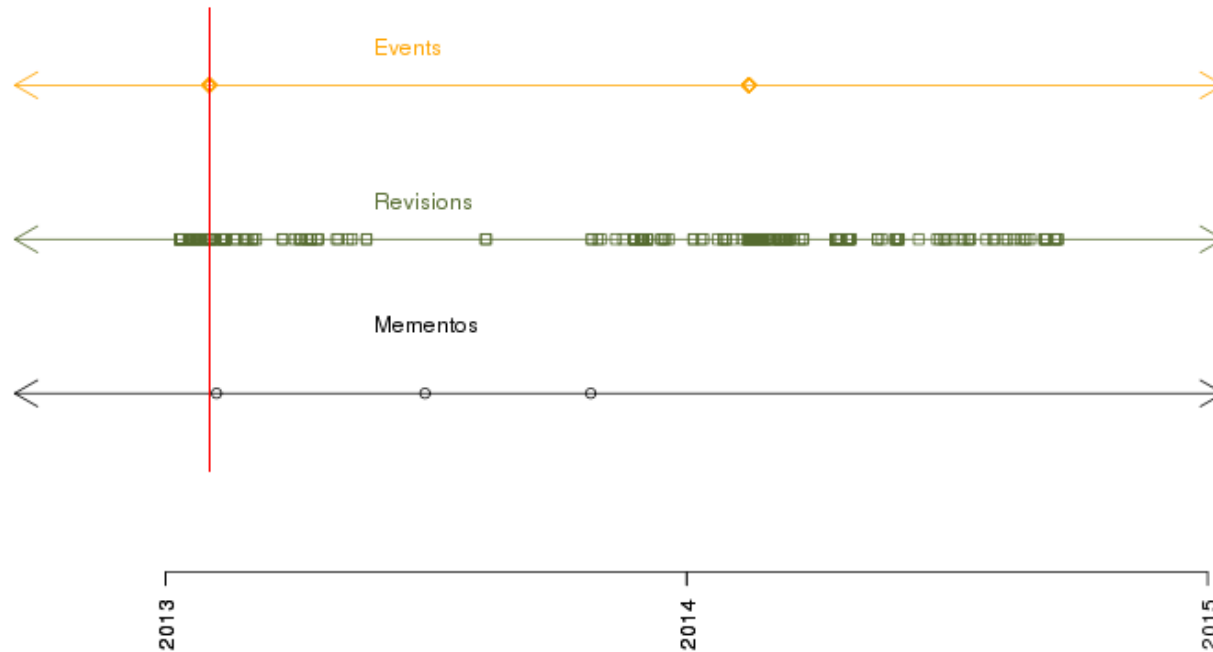
For *Normal* Pages, We See A Variety of Probabilities

TABLE 11: Spoiler probabilities for most popular pages within each fan wiki

Wiki	Page Name	Probability of Spoiler	# of Spoiler Areas	# of Revisions	# of Mementos
bigbangtheory	Sheldon Cooper	0.31	69	1958	30
boardwalkempire	Nucky Thompson	0.15	31	290	15
breakingbad	Walter White	0.43	40	882	20
continuum	Keira Cameron	0.54	21	104	5
downtonabbey	Sybil Branson	0.42	23	580	3
gameofthrones	Daenerys Targaryen	0.16	24	768	29
grimm	Nick Burkhardt	0.39	30	795	5
house-of-cards	Frank Underwood	0.0	13	380	3
how-i-met-your-mother	Barney Stinson	0.55	120	588	13
lostpedia	Kate Austen	0.67	94	3531	27
madmen	Mad Men Wiki	0.22	36	250	85
ncis	Abigail Sciuto	0.67	182	404	11
onceuponatime	Emma Swan	0.36	34	1210	11
scandal	Main Page	0.60	31	250	14
trueblood	Eric Northman	0.28	47	931	14
whitecollar	Neal Caffrey	0.29	38	199	8

For *Wiki-Before-Show* Pages, We Also See A Variety of Probabilities

TABLE 11: Spoiler probabilities for most popular pages within each fan wiki

Wiki	Page Name	Probability of Spoiler	# of Spoiler Areas	# of Revisions	# of Mementos
bigbangtheory	Sheldon Cooper	0.31	69	1958	30
boardwalkempire	Nucky Thompson	0.15	31	290	15
breakingbad	Walter White	0.43	40	882	20
continuum	Keira Cameron	0.54	21	104	5
downtonabbey	Sybil Branson	0.42	23	580	3
gameofthrones	Daenerys Targaryen	0.16	24	768	29
grimm	Nick Burkhardt	0.39	30	795	5
house-of-cards	Frank Underwood	0.0	13	380	3
how-i-met-your-mother	Barney Stinson	0.55	120	588	13
lostpedia	Kate Austen	0.67	94	3531	27
madmen	Mad Men Wiki	0.22	36	250	85
ncis	Abigail Sciuto	0.67	182	404	11
onceuponatime	Emma Swan	0.36	34	1210	11
scandal	Main Page	0.60	31	250	14
trueblood	Eric Northman	0.28	47	931	14
whitecollar	Neal Caffrey	0.29	38	199	8

Our Model Breaks Down For *Season-In-A-Day*

TABLE 11: Spoiler probabilities for most popular pages within each fan wiki

Wiki	Page Name	Probability of Spoiler	# of Spoiler Areas	# of Revisions	# of Mementos
bigbangtheory	Sheldon Cooper	0.31	69	1958	30
boardwalkempire	Nucky Thompson	0.15	31	290	15
breakingbad	Walter White	0.43	40	882	20
continuum	Keira Cameron	0.54	21	104	5
downtonabbey	Sybil Branson	0.42	23	580	3
gameofthrones	Daenerys Targaryen	0.16	24	768	29
grimm	Nick Burkhardt	0.39	30	795	5
house-of-cards	Frank Underwood	0.0	13	380	3
how-i-met-your-mother	Barney Stinson	0.55	120	588	13
lostpedia	Kate Austen	0.67	94	3531	27
madmen	Mad Men Wiki	0.22	36	250	85
ncis	Abigail Sciuto	0.67	182	404	11
onceuponatime	Emma Swan	0.36	34	1210	11
scandal	Main Page	0.60	31	250	14
trueblood	Eric Northman	0.28	47	931	14
whitecollar	Neal Caffrey	0.29	38	199	8

13 spoiler areas exist for this page, but they all have length 0

The Probabilities Do Not Follow A Known Distribution

Histogram of Overall Spoiler Probabilities from 16 Wikia Sites

50% Of The Pages Have A Spoiler Probability < 0.66

Cumulative Distribution Function of Probabilities of Encountering a Spoiler Using Mindist on Mementos From 16 Wikia Sites

Percentage of wiki pages

Probability of Encountering a Spoiler
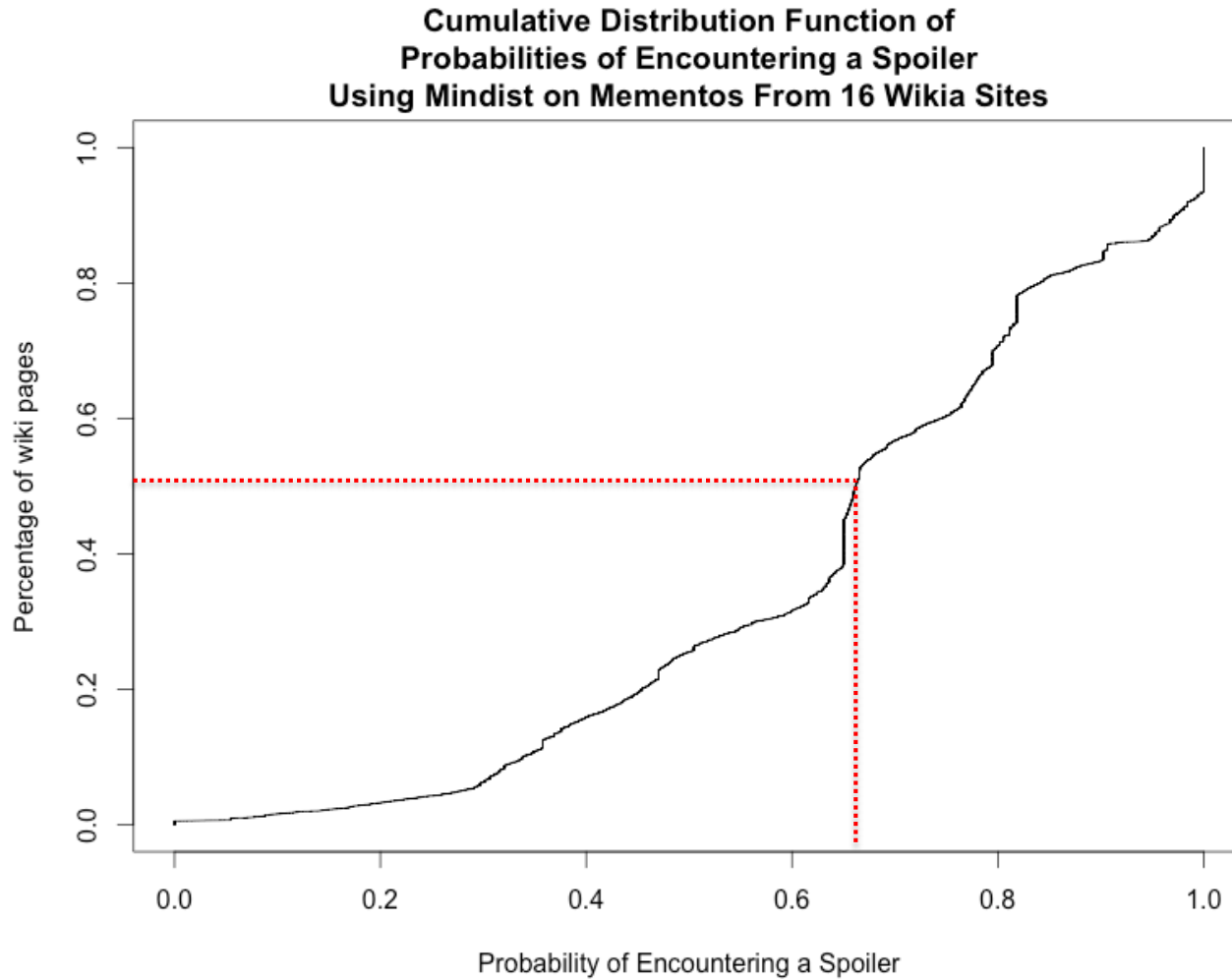
Remember Missed Updates?

These Revisions Were Missed
By The Web Archive...
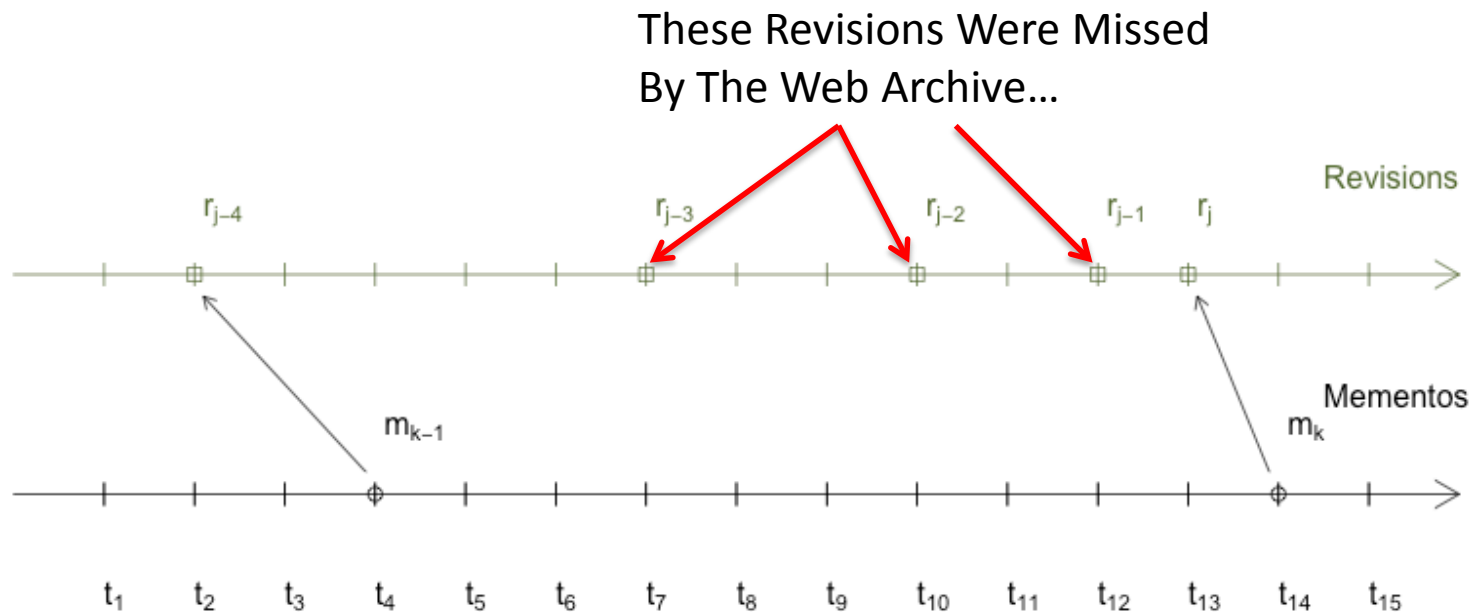
We Had An Opportunity To Measure Missed Updates

History of
Missed Updates Per Day
For 15000+ Wiki Pages
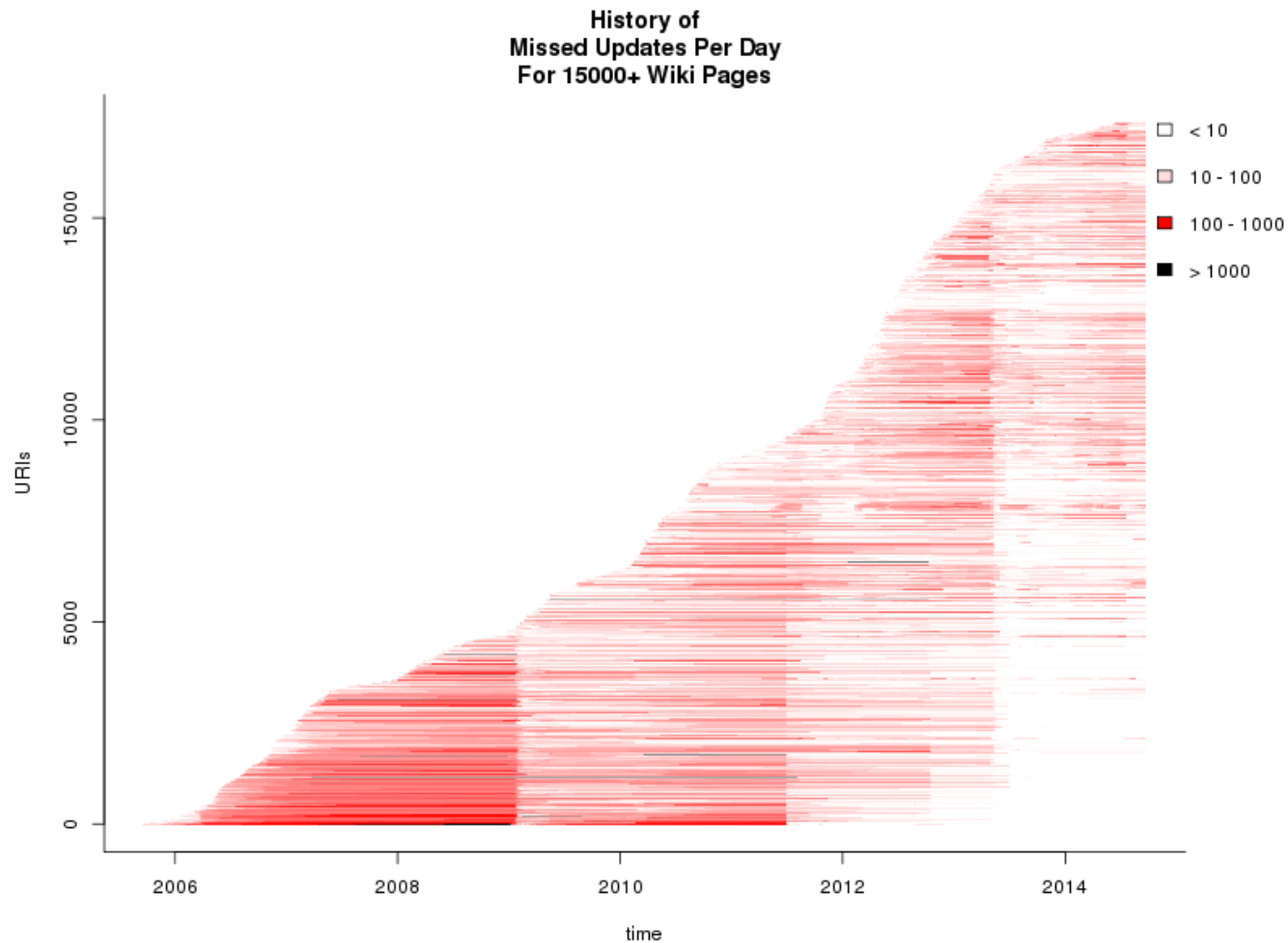
We See Lines Where The Archive Adopts A More Aggressive Policy
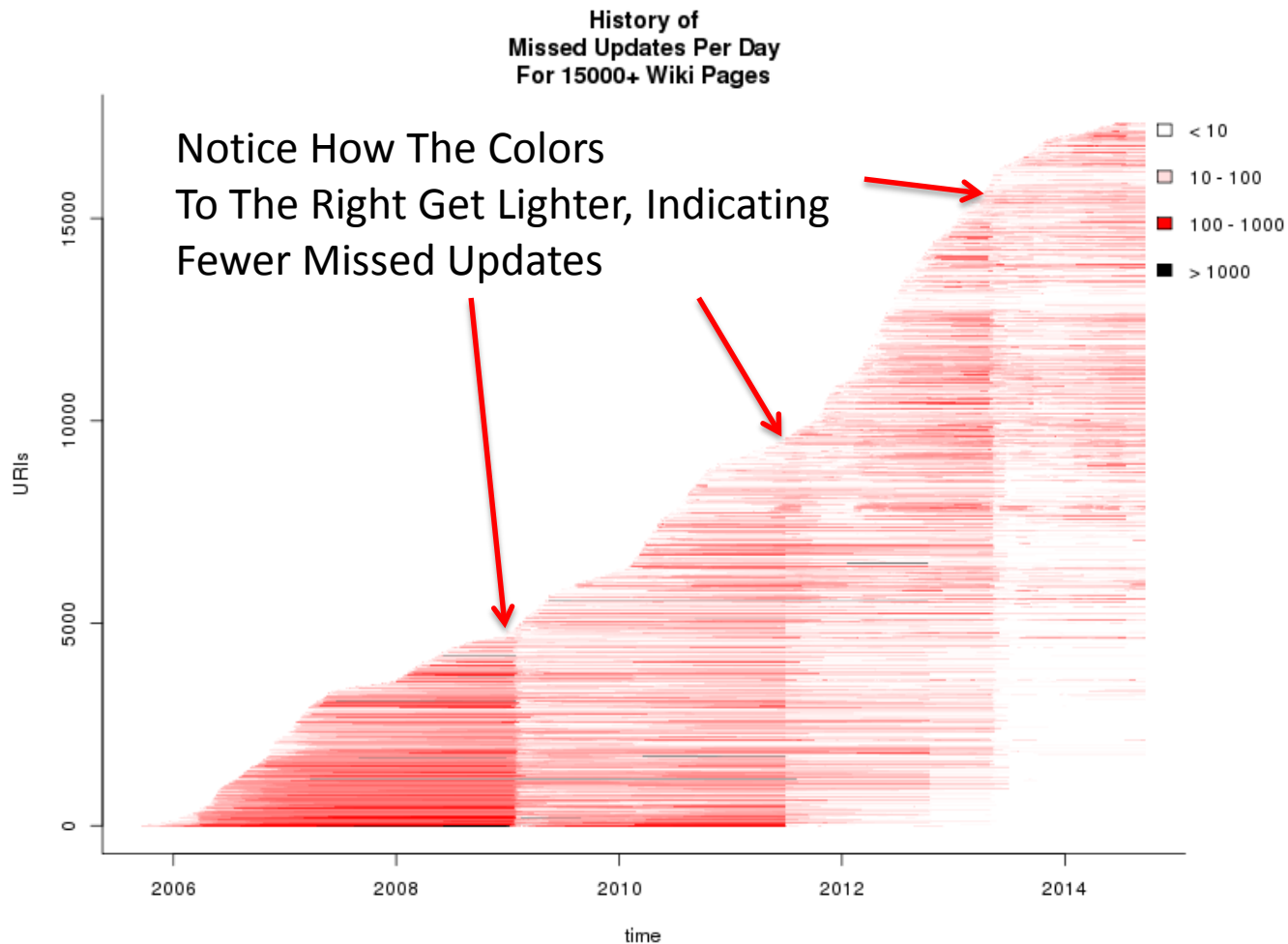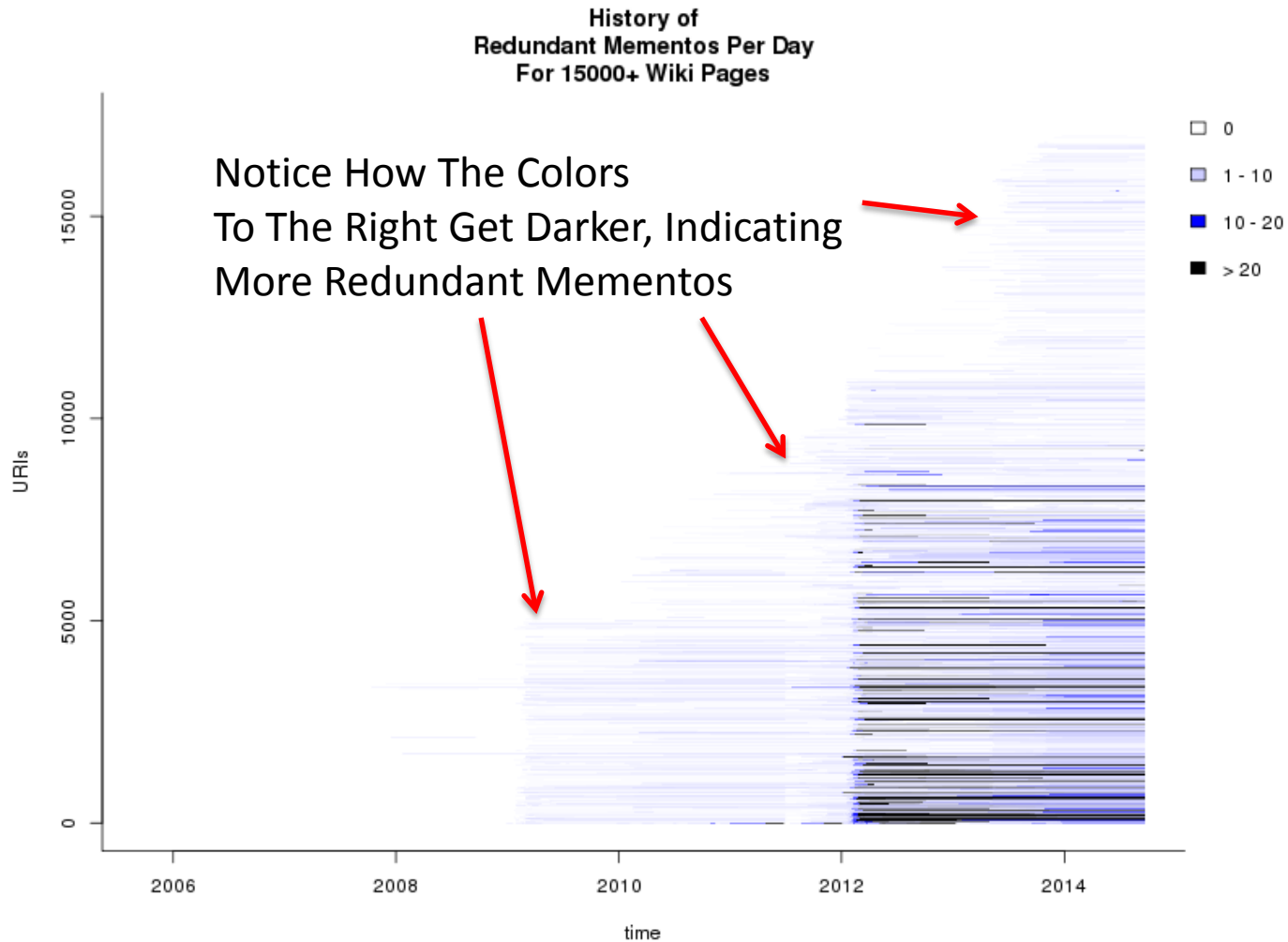
History of Missed Updates Per Day For 15000+ Wiki Pages

Notice How The Colors To The Right Get Lighter, Indicating Fewer Missed Updates

Legend:
- □ < 10
- □ 10 - 100
- ■ 100 - 1000
- ■ > 1000

Looking At Redundant Mementos, We See The Opposite Effect

History of Redundant Mementos Per Day For 15000+ Wiki Pages

Notice How The Colors To The Right Get Darker, Indicating More Redundant Mementos

Looking At Redundant Mementos, We See The Opposite Effect

History of
Redundant Mementos Per Day
For 15000+ Wiki Pages

Remember That 38% of the Pages Do Not Have Mementos, So This Does Not Reflect All Archived Pages

Legend:
- 0
- 1 - 10
- 10 - 20
- > 20
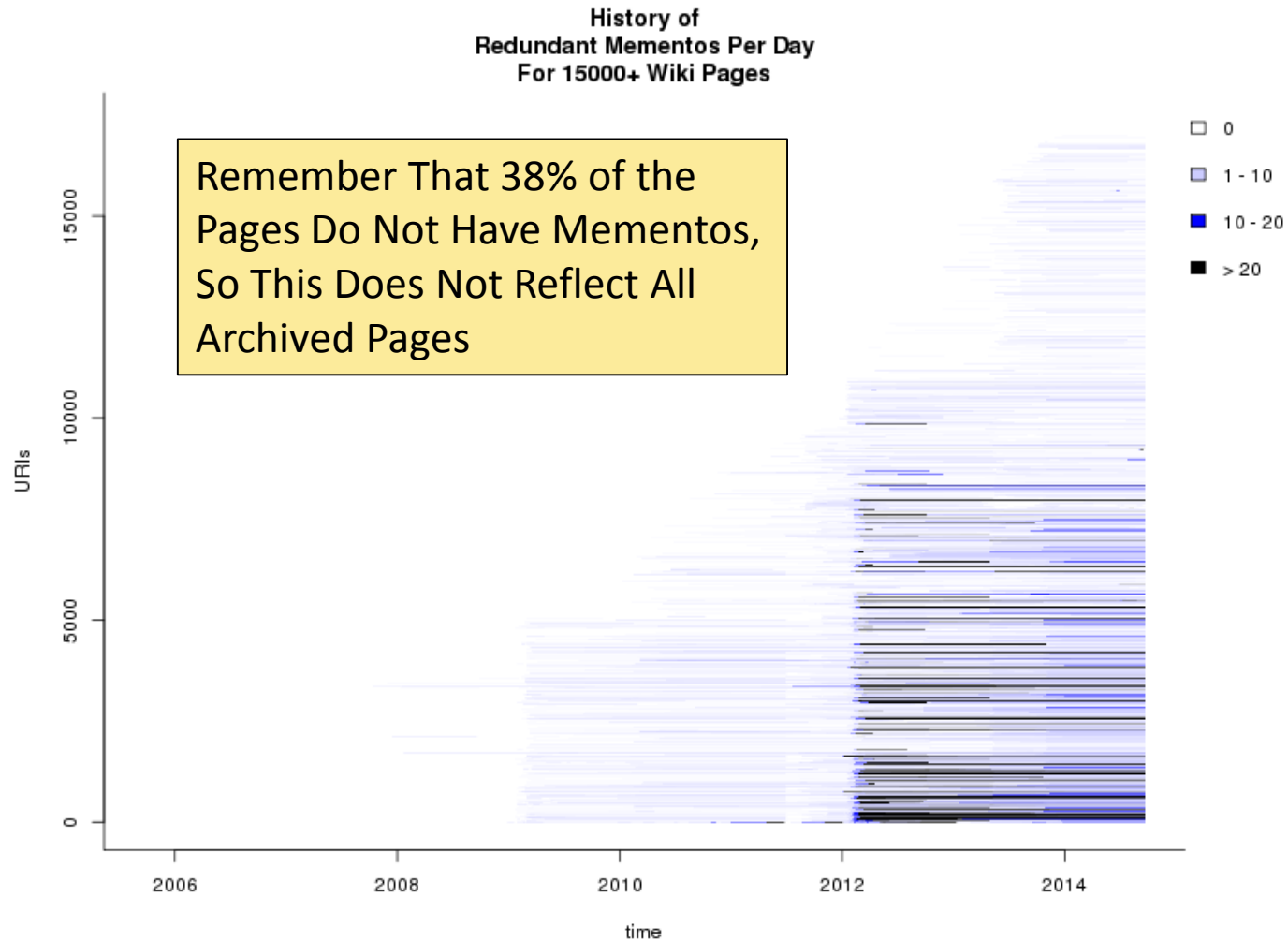
URIs

time

The Wayback Machine Uses mindist, Ergo Users Encounter Spoilers There…

SPOILERS IN THE WAYBACK MACHINE

We Have Access To Some Anonymized Logs From The Wayback Machine

```
1  0.247.222.86 - - [02/Feb/2012:07:03:55 +0000] "GET
      http://web.archive.org/web/20020404020224/http://www.example.com/page1.html HTTP
      /1.1" 200 18875 "
      http://wayback.archive.org/web/*/http://www.example.com/page1.html" "Mozilla/5.0 (
      Macintosh; Intel Mac OS X 10_6_8) AppleWebKit/535.7 (KHTML, like Gecko) Chrome
      /16.0.912.77 Safari/535.7"}
2  0.247.222.86 - - [02/Feb/2012:07:10:02 +0000] "GET
      http://web.archive.org/web/20020405015622/http://www.example.com/page2.html HTTP
      /1.1" 200 18875 "
      http://wayback.archive.org/web/20020404020224/http://www.example.com/page1.html" "
      Mozilla/5.0 (Macintosh; Intel Mac OS X 10_6_8) AppleWebKit/535.7 (KHTML, like
      Gecko) Chrome/16.0.912.77 Safari/535.7"}
```

From these logs, we have:
- which memento the user came from (the referer)
- the memento returned to the user

The Wayback Machine Rewrites URIs, Redirecting Users When They Click On Links To Other Pages That Exist During The Same Time Period.

We Can Infer Desired Datetime And Have The Memento-datetime

http://web.archive.org/web/20020404020224/http://www.example.com

2002	04	04	02	02	24
year	month	date	hour	minute	second
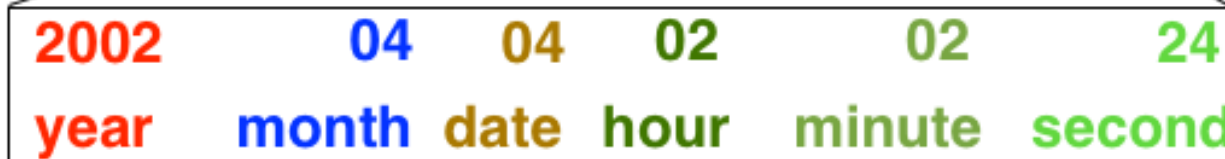
Memento from archive.org for URI-R
http://www.example.com on April 4, 2002 at 2:02:24 AM GMT

From the referrer, we can *infer* their desired datetime

From the memento returned to the user, we can acquire the Memento-Datetime

Using the Logs, We Can See How Many wikia.com Requests End In Spoilers

Wayback Logs

Memento-datetime From Visited URI

Revisions From Wiki Export Pages

Desired Datetime From Referrer

Datetime Of Revision

Did The Wayback Machine Deliver The User To A Revision That Existed In The Future From The Desired Datetime?

No

Yes

Safe
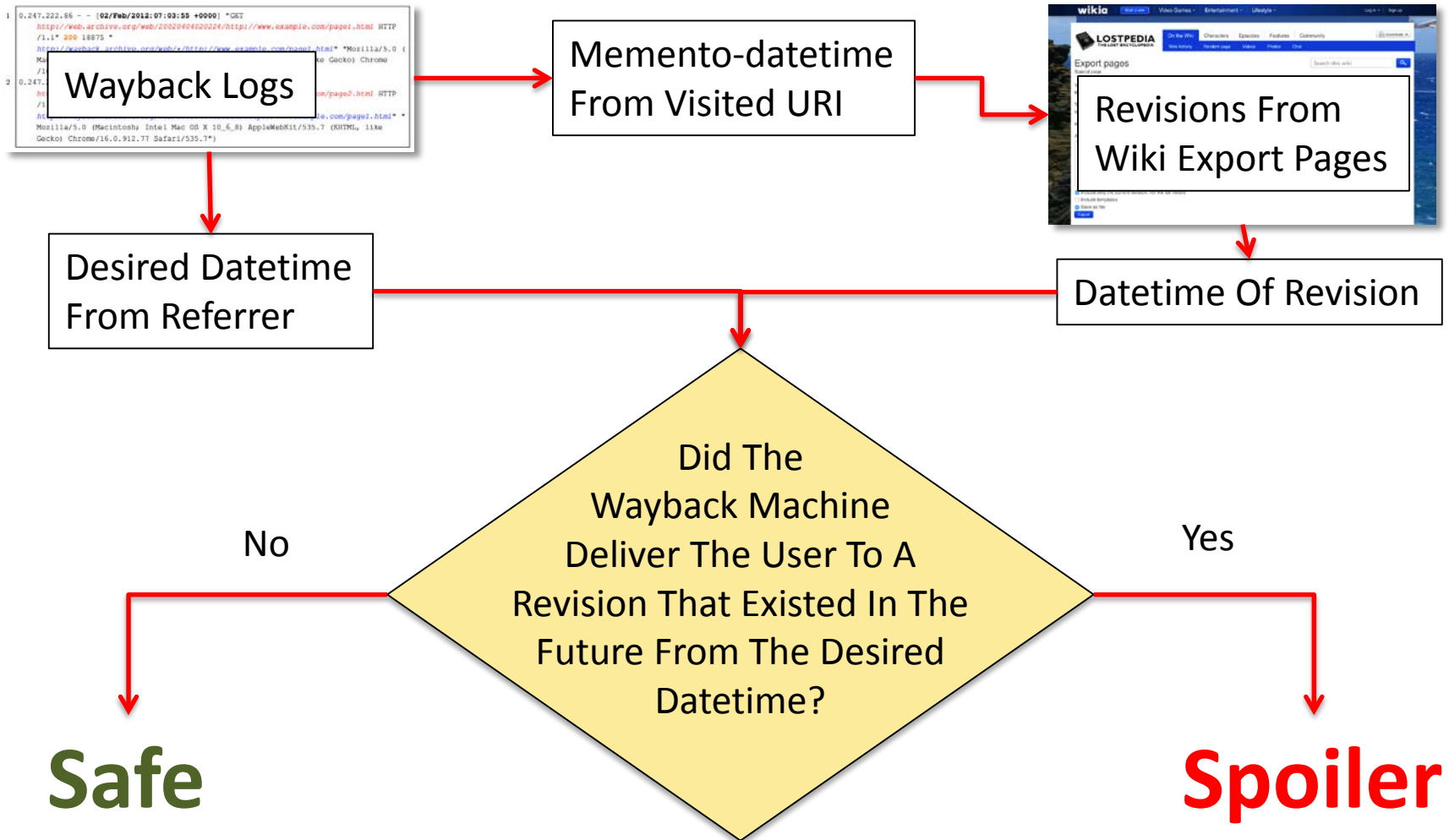
Spoiler

Wayback Machine Log Results
For wikia.com Requests

Wayback Machine Requests For wikia.com:
Spoilers vs Safe

81%

19%

safe
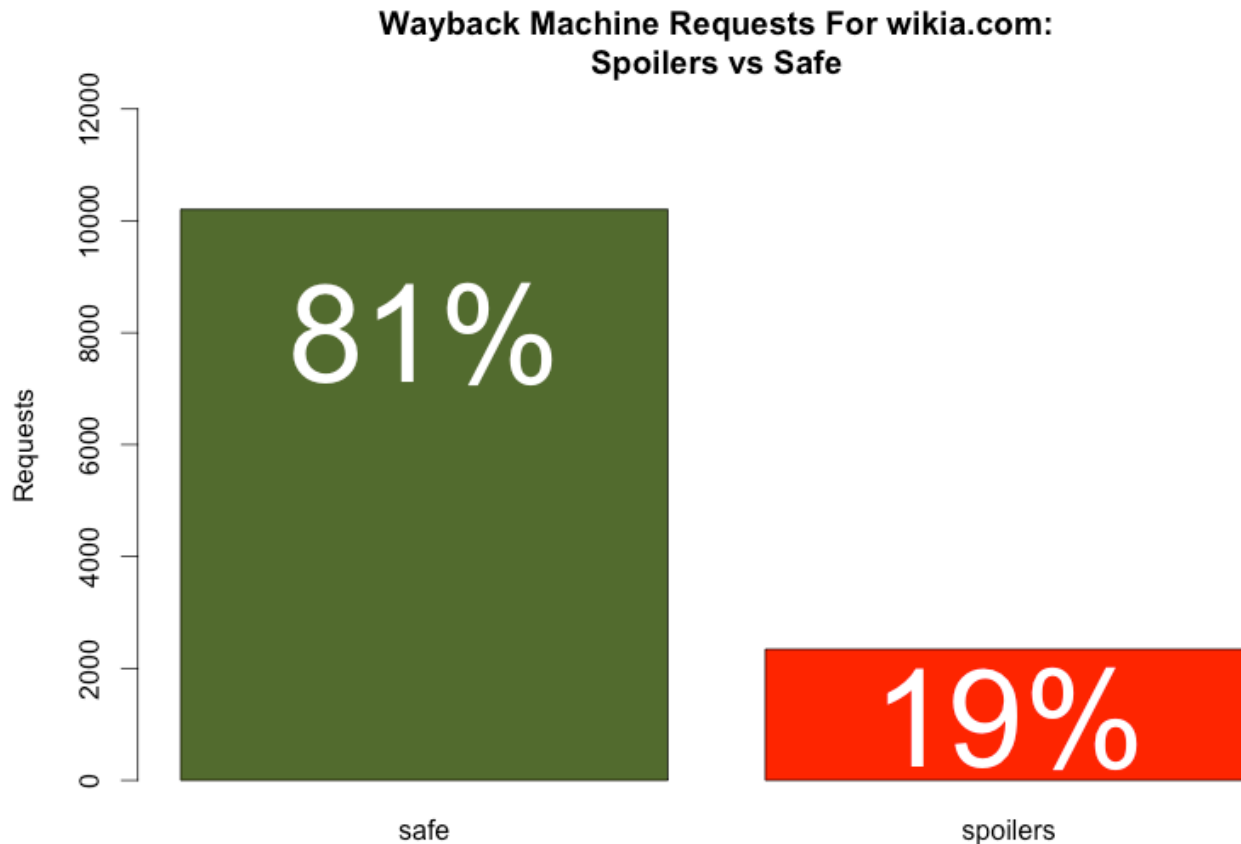
spoilers

We Have No Way To Identify Wikis Other Than By Domain Name,
Data Does Not Include All Wikis...

A Solution For Avoiding Spoilers In Wikis. Wiki Revisions Are Mementos, Too!

THE MEMENTO MEDIAWIKI EXTENSION

We Developed The Memento MediaWiki Extension To Use minpast

MediaWiki Is The Software Used By Wikipedia And Wikia.Com

http://www.mediawiki.org/wiki/Extension:Memento

We Set Up A Demonstration Wiki…

http://ws-dl-05.cs.odu.edu/demo/index.php/Main_Page

Data was exported from: http://awoiaf.westeros.org

For Performance, We Had An Opportunity To Experiment With Alternate Memento Patterns
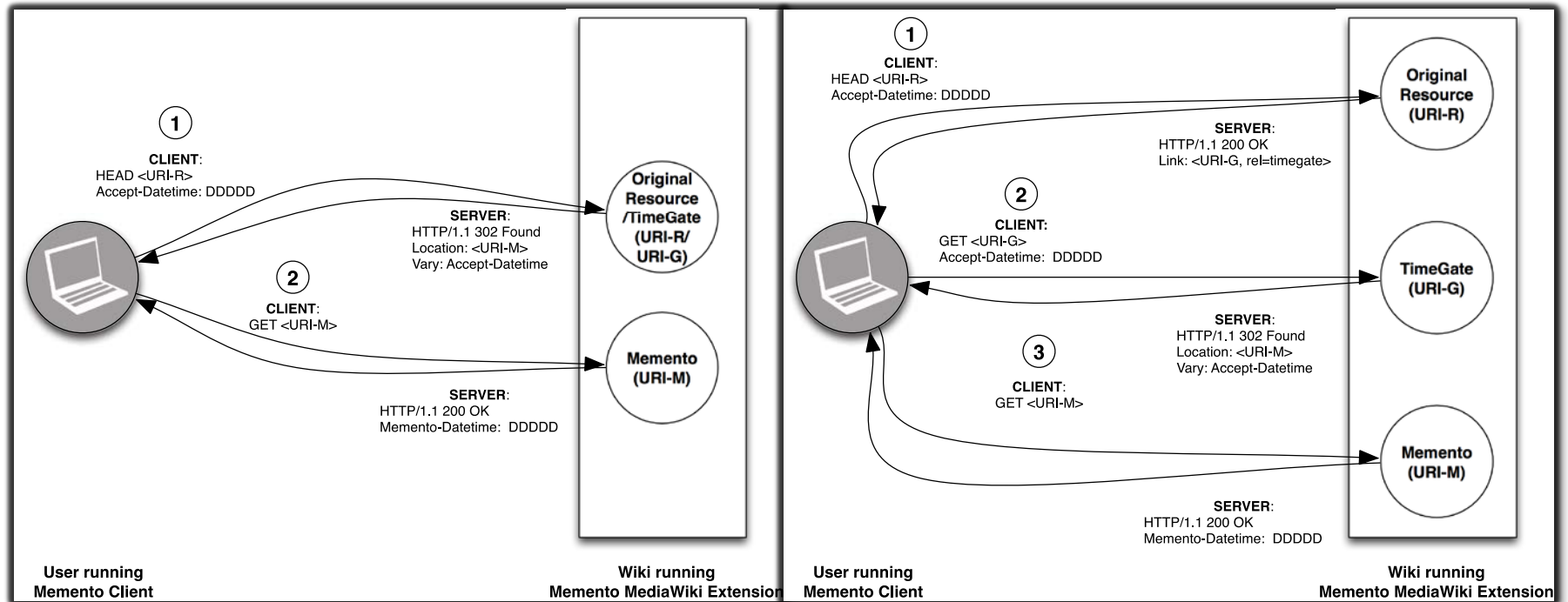
Original Resource Acts As Own TimeGate (Pattern 1.1)

CLIENT:
HEAD <URI-R>
Accept-Datetime: DDDDD

SERVER:
HTTP/1.1 302 Found
Location: <URI-M>
Vary: Accept-Datetime

CLIENT:
GET <URI-M>

SERVER:
HTTP/1.1 200 OK
Memento-Datetime: DDDDD

Original
Resource
/TimeGate
(URI-R/
URI-G)

Memento
(URI-M)

**User running
Memento Client**

**Wiki running
Memento MediaWiki Extension**
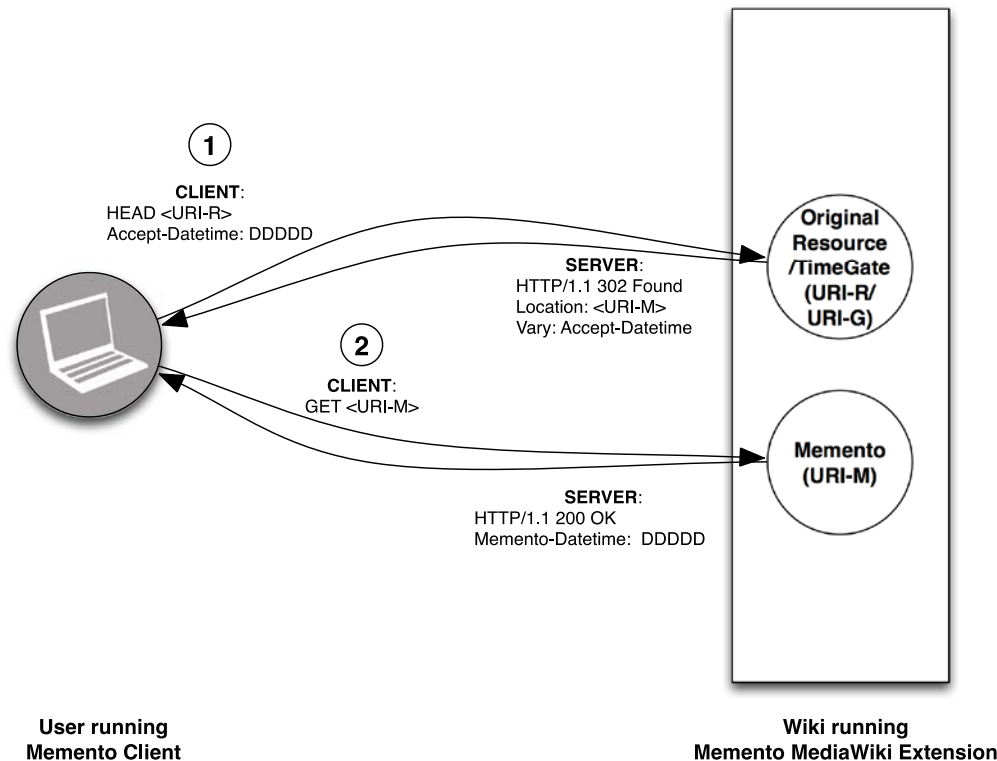
This Pattern Only Requires 2 Requests To Acquire A Memento. Intuitively, It Should Perform Better.

Client Must Ask About TimeGate (Pattern 2.1)

1
CLIENT:
HEAD <URI-R>
Accept-Datetime: DDDDD

SERVER:
HTTP/1.1 200 OK
Link: <URI-G, rel=timegate>

Original Resource (URI-R)

2
CLIENT:
GET <URI-G>
Accept-Datetime: DDDDD

TimeGate (URI-G)

SERVER:
HTTP/1.1 302 Found
Location: <URI-M>
Vary: Accept-Datetime

3
CLIENT:
GET <URI-M>

Memento (URI-M)

SERVER:
HTTP/1.1 200 OK
Memento-Datetime: DDDDD

User running Memento Client

Wiki running Memento MediaWiki Extension
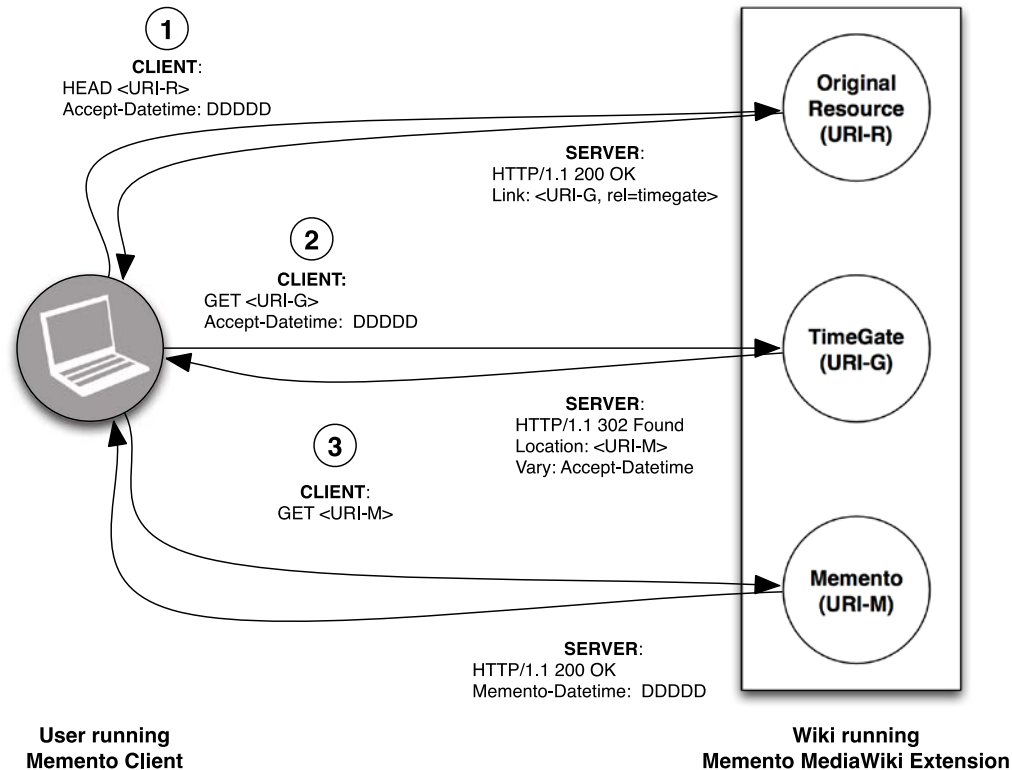
This Pattern Requires 3 Requests To Acquire A Memento.
Intuitively, It Should Perform Worse.

Using Analysis and Experimentation We Found That The Three-Request Pattern Actually Performed Better
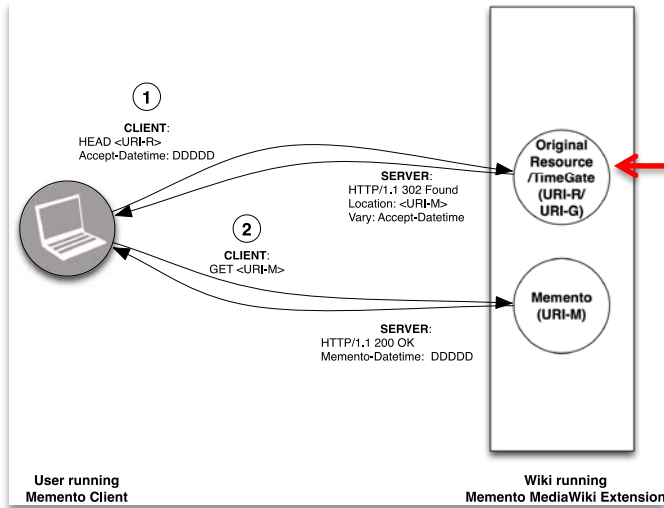
Mediawiki Takes Too Long To Generate A TimeGate Response If The Two Are The Same Page.

The Three-request Pattern Requires An Extra Request, But, Due To Round-trip Time, It Performs Better Unless The User Has A Bandwidth Of 21,926 bps Or Less.
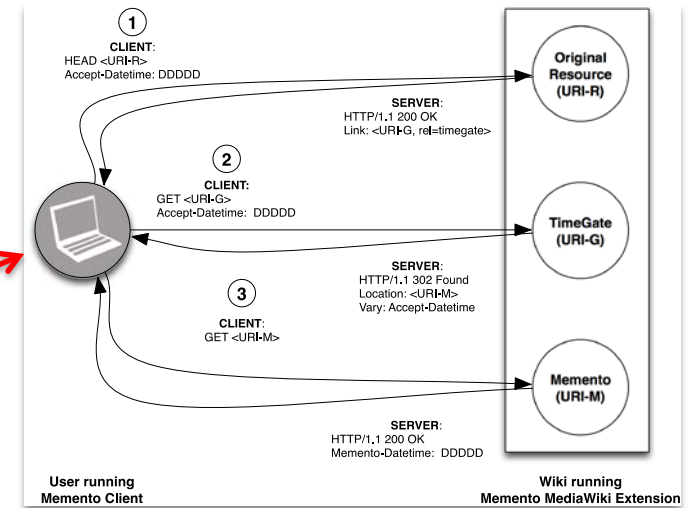
Serving Current Pages Takes The Same Time With Or Without The Extension

**Differences in URI-R Performance
For the Memento Mediawiki Extension
Between No Installation and an All Headers Installation**

Worse Performance

Better Performance

Mean   = -0.0072 s
Std dev =   0.3526 s

We used seige for testing, as detailed in http://arxiv.org/abs/1406.3876
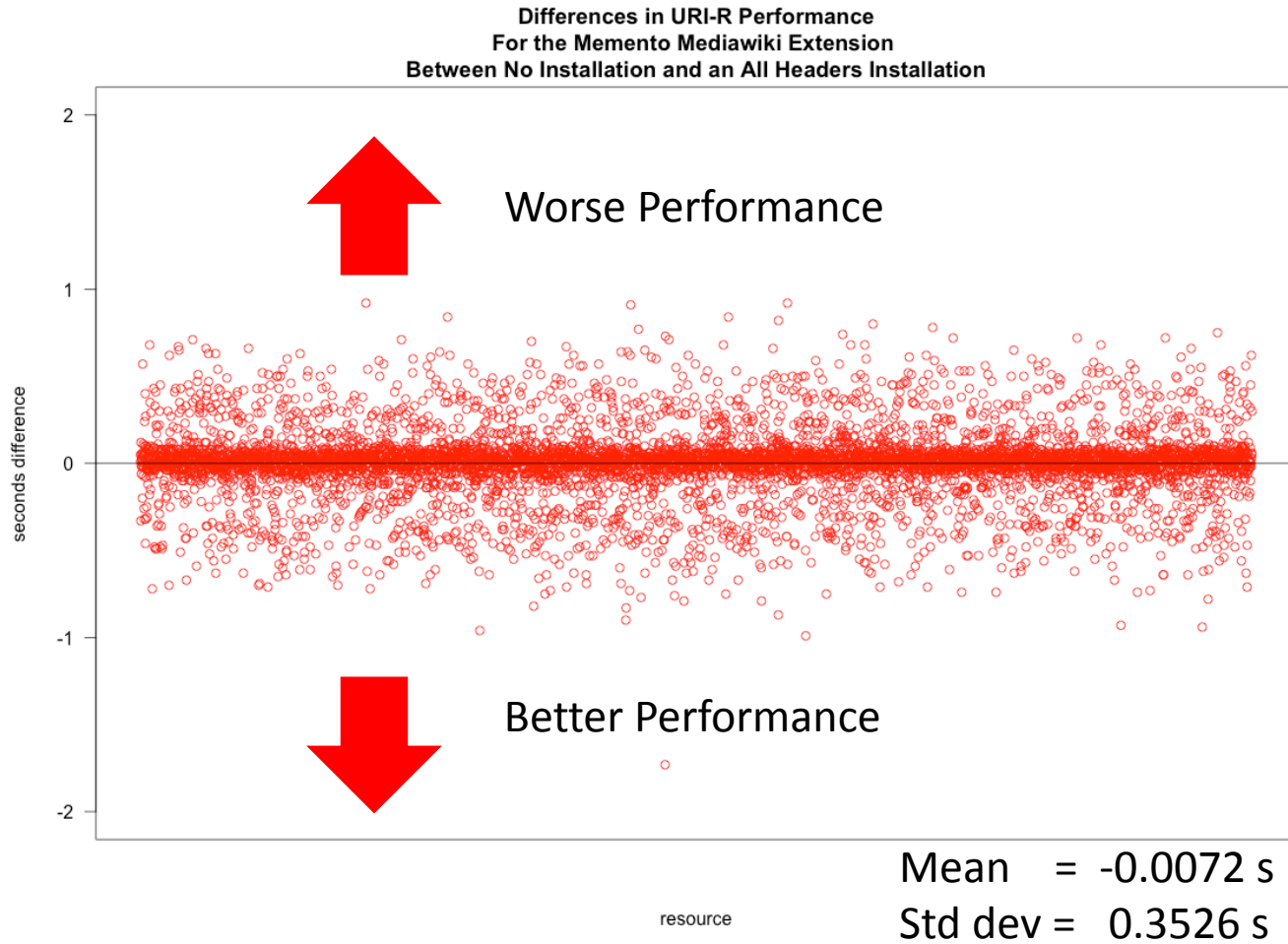
Serving Old Pages (Mementos) Takes The Same Time With Or Without The Extension

Differences in URI-M Performance
For the Memento Mediawiki Extension
Between No Installation and an All Headers Installation

Worse Performance

Better Performance

Mean = -0.0026 s
Std dev = 0.0421 s

We used seige for testing, as detailed in http://arxiv.org/abs/1406.3876
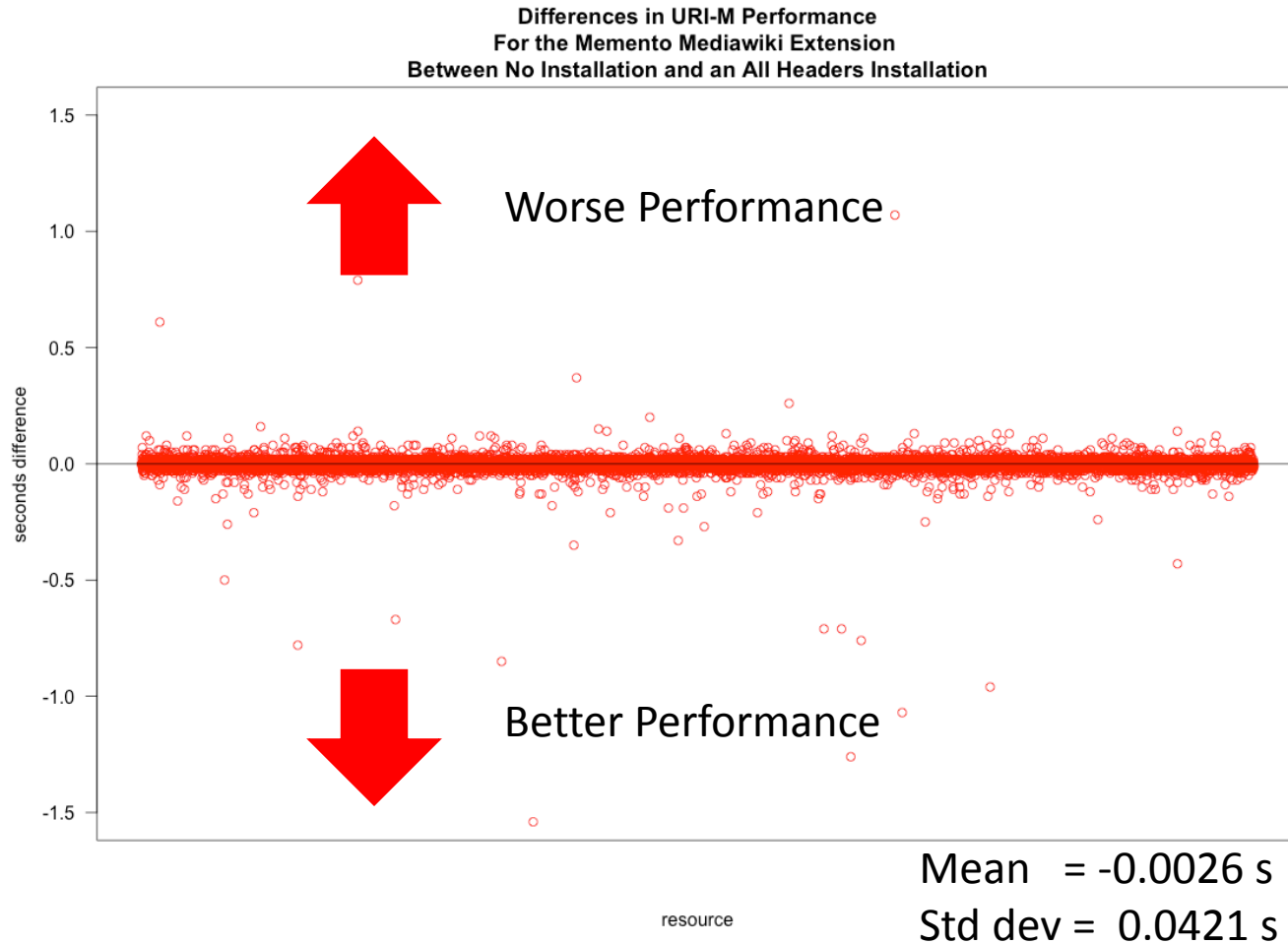
TimeMaps Are Smaller Than Wiki History Pages, So They Perform Better

Differences in size between Memento MediaWiki Extension TimeMaps and MediaWiki History Pages

Better Performance

Mean   = -34.74 kB
Std dev = 31.12 kB

We used seige for testing, as detailed in http://arxiv.org/abs/1406.3876
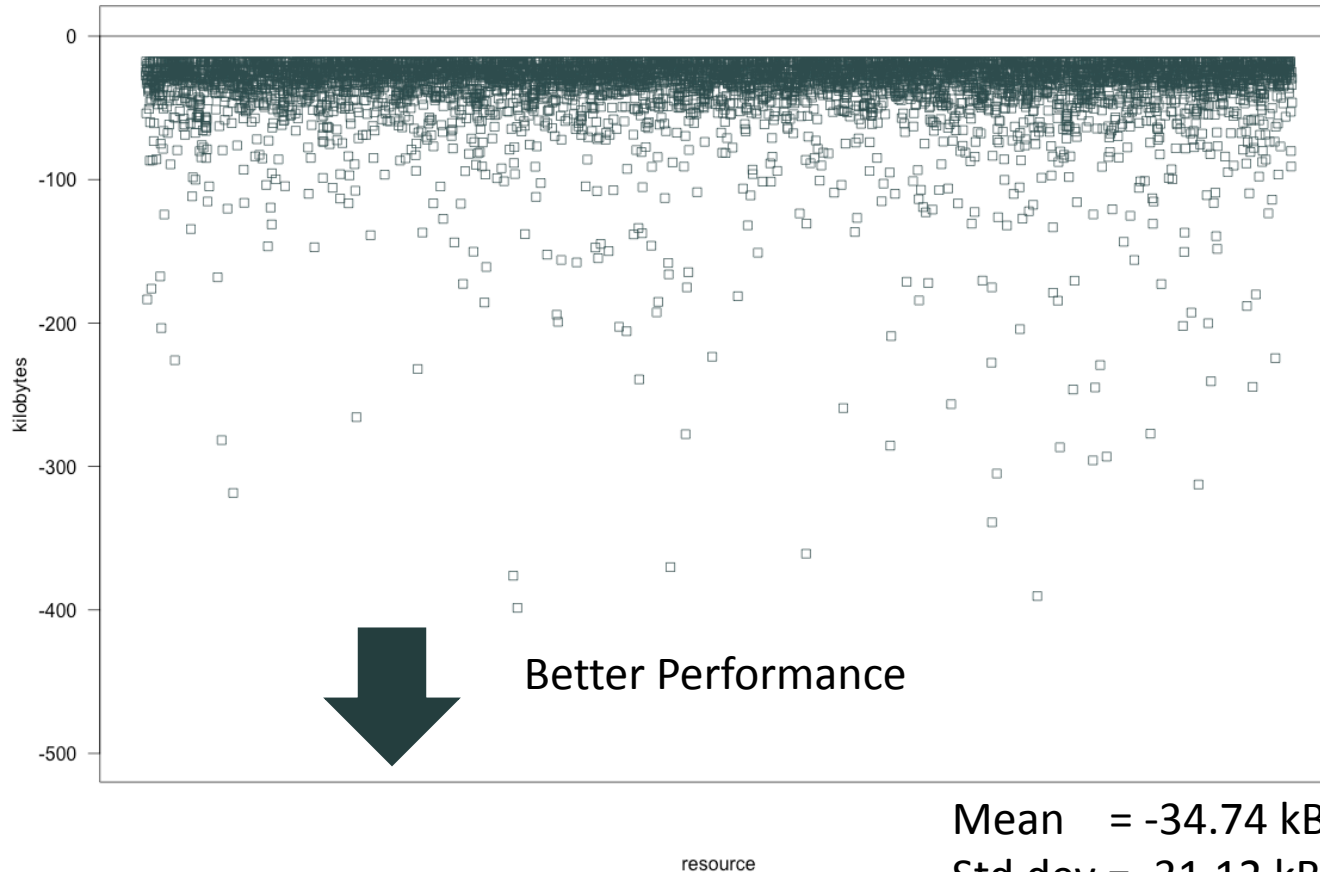
MediaWiki Does Not Just Store Previous Revisions Of Pages...

Images, Stylesheets, and JavaScript have previous revisions stored as well

Then We Wondered, What About Images? They Could Contain Spoilers…

Consider This Map

This Map is important to understanding the content of this article

This image is changed as the article is changed, to reflect its content

http://en.wikipedia.org/wiki/Same-sex_marriage_law_in_the_United_States_by_state

It's The Same Map If Today We Visit The June 5, 2013 Revision

Users can't view this embedded resource as it looked on June 2013 while reading the article from that time period

http://en.wikipedia.org/w/index.php?title=Same-sex_marriage_law_in_the_United_States_by_state&oldid=558400004

What Should Have Happened

This is the current map
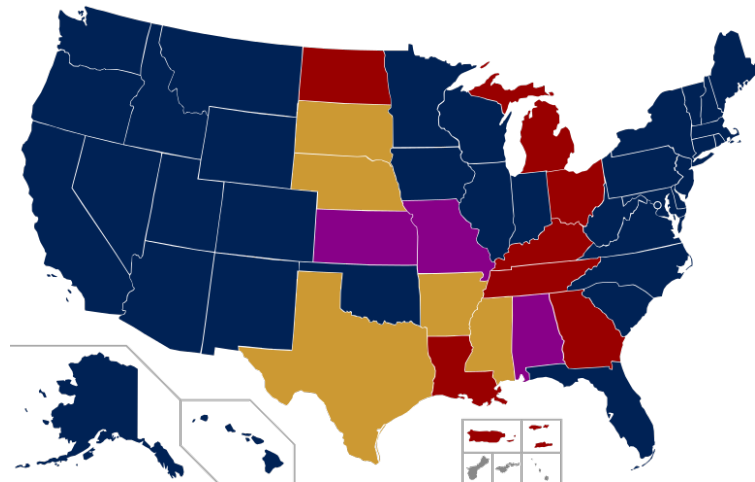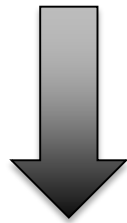
This is the the map from June, 2013 that should have been displayed
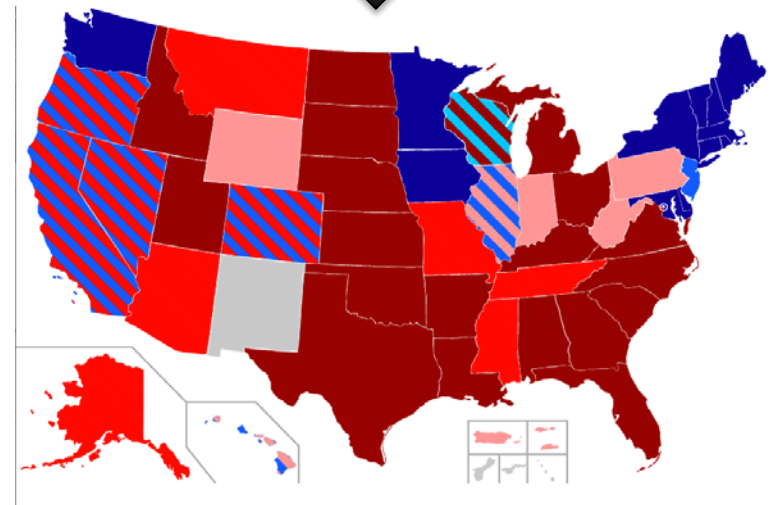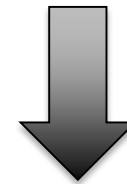
The content of the article won't match the data in this visual aid, possibly confusing a user who wanted historical information on this topic

We Developed A Solution For Images

The $file argument's getHistory() function of the ImageBeforeProduceHTML hook can be used to acquire previous revisions of images

We Could Not Extract Previous Revisions Of CSS And Javascript…

The data is present, but we could not find any way for an extension to access or render it.

We Demonstrated Avoiding Spoilers At WikiConUSA 2014…

http://ws-dl.blogspot.com/2014/06/2014-06-02-wikiconference-usa-2014-trip.html
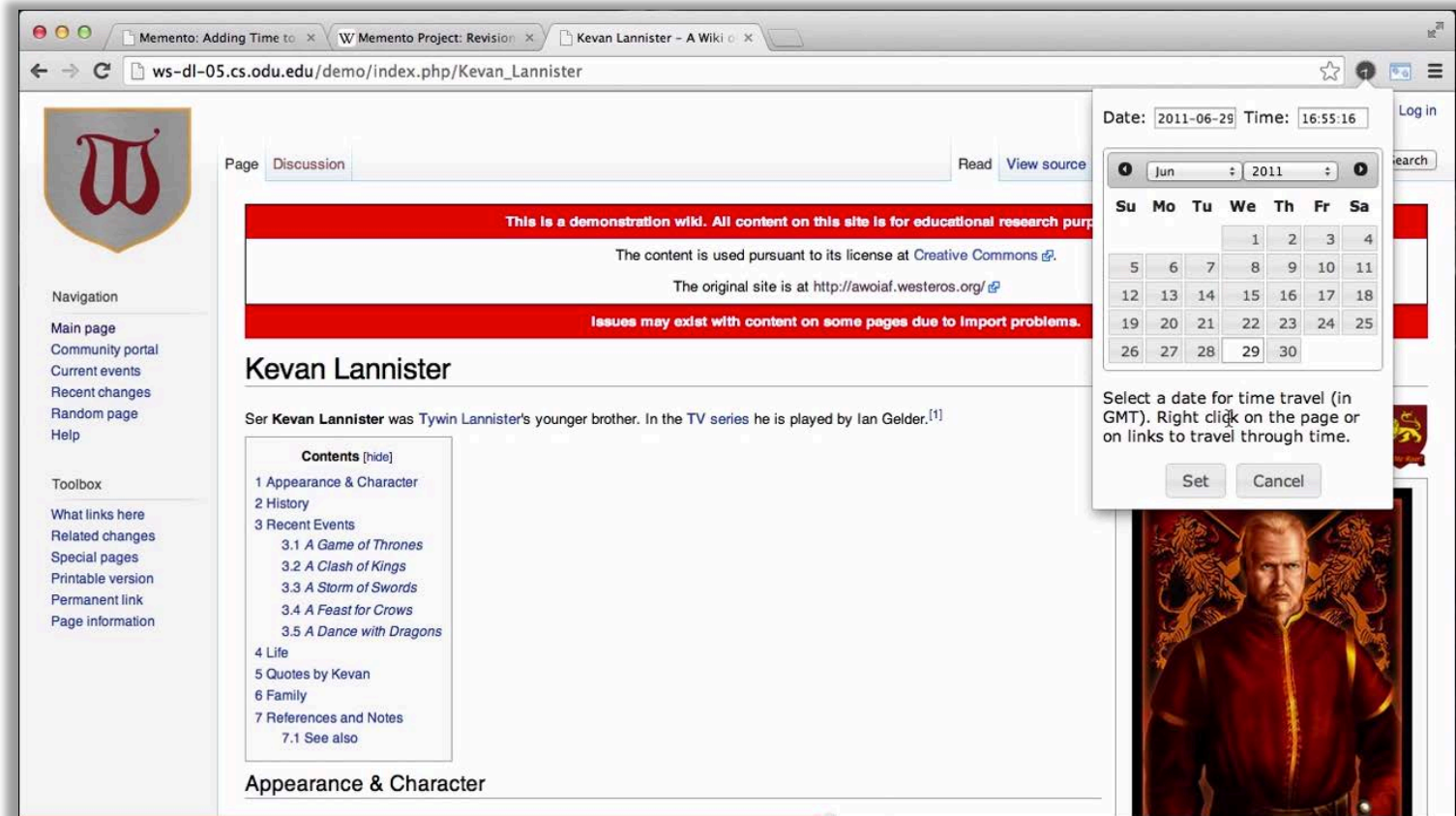
Even The Current Version Of The Page Contains A Spoiler!!!
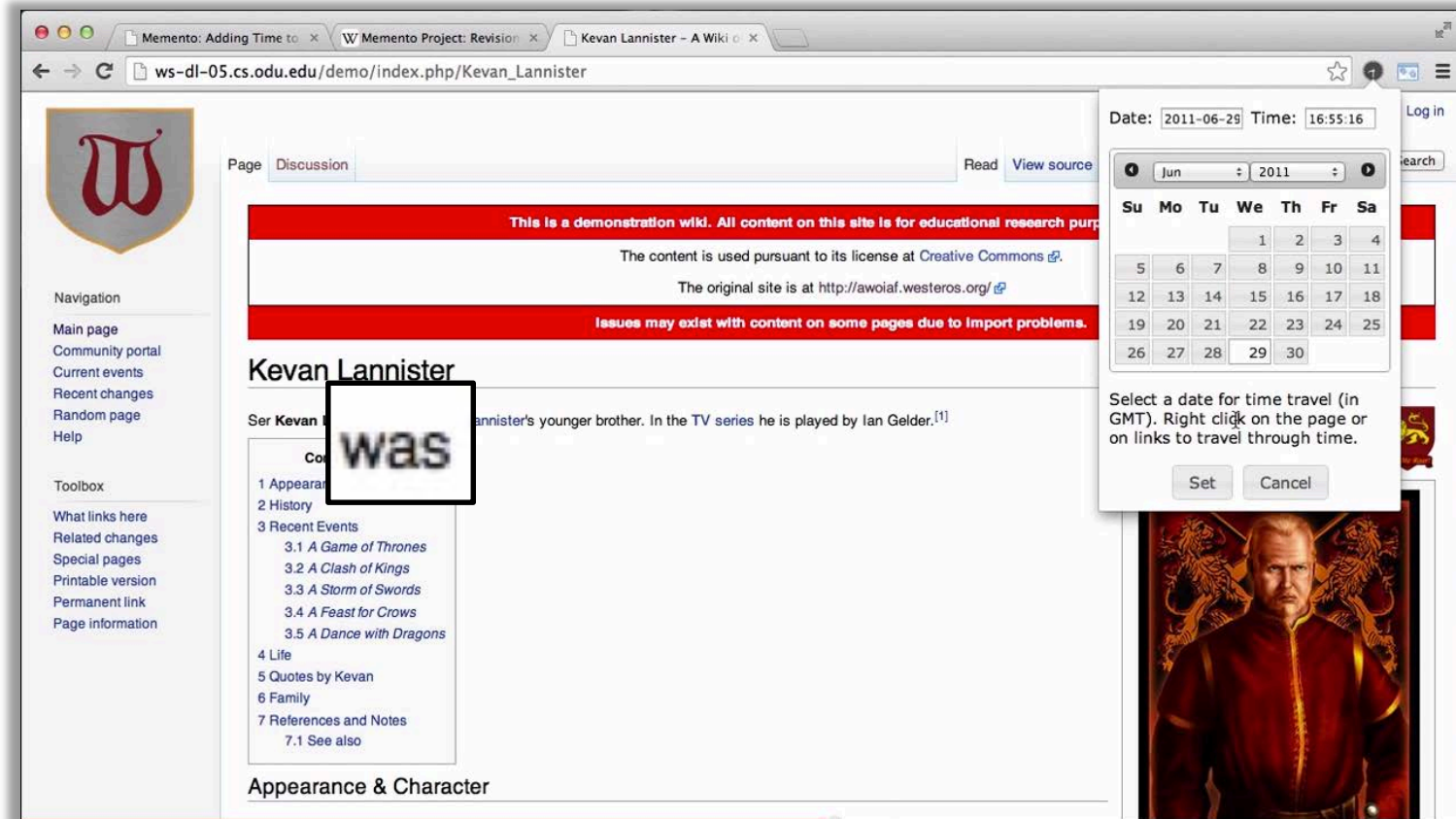
We want to find information about Kevan Lannister, but haven't read the book *A Dance with Dragons* yet. We set the Memento Chrome Extension prior to the release of that book: June 29, 2011.

So We Set Memento For Chrome To The Correct Date…
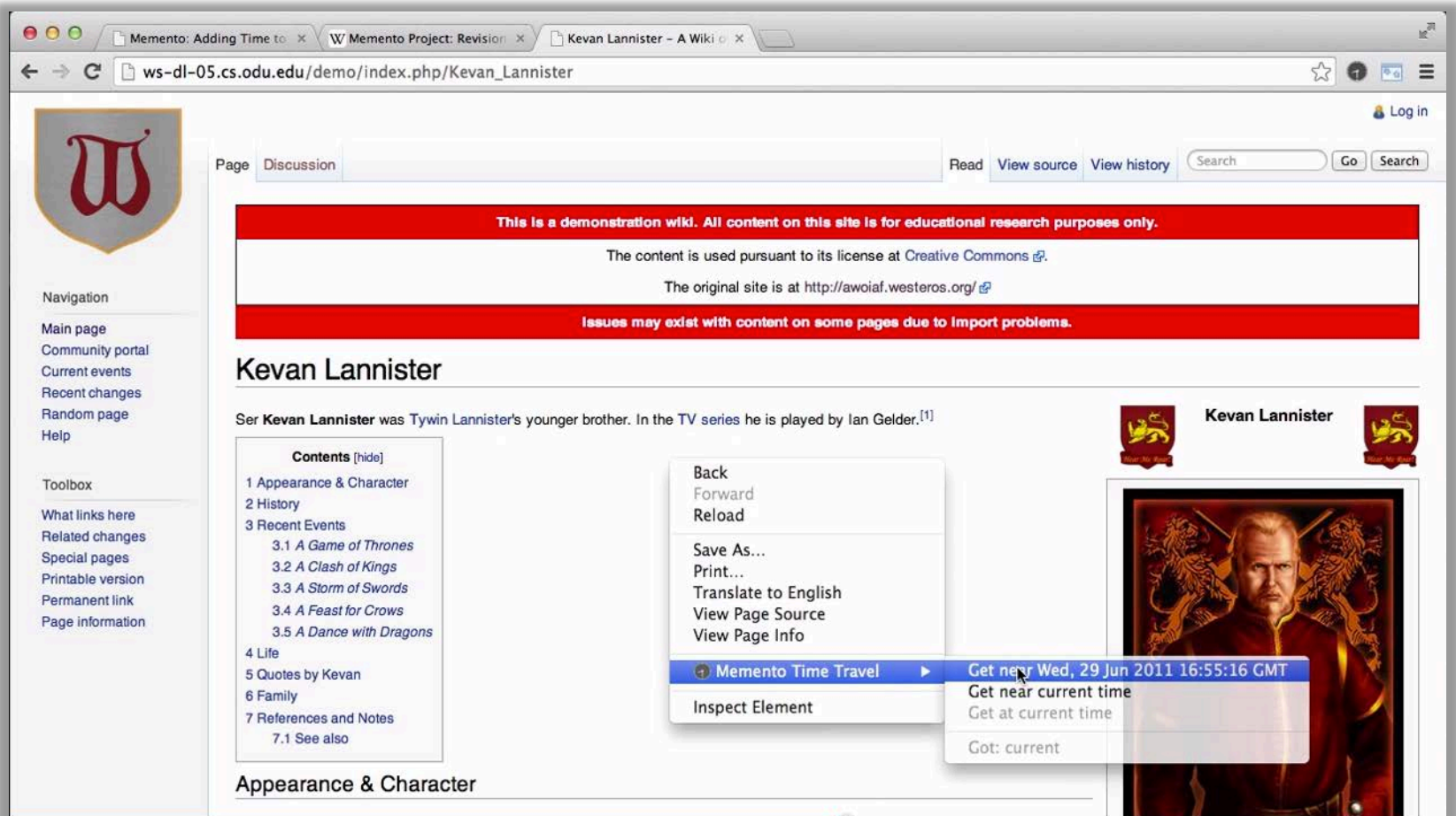
We use the Memento Chrome Extension to request a revision of the page close to, but not over, our requested date.

…And Got A Page Without Spoilers

And We Avoid Spoilers for *A Dance With Dragons*…

Where Do We Go From Here?

FUTURE WORK

How Do We Handle Spoilers For *Season-In-A-Day* Series?

Can We Create A New Heuristic Based On mindist For Spoiler Detection?
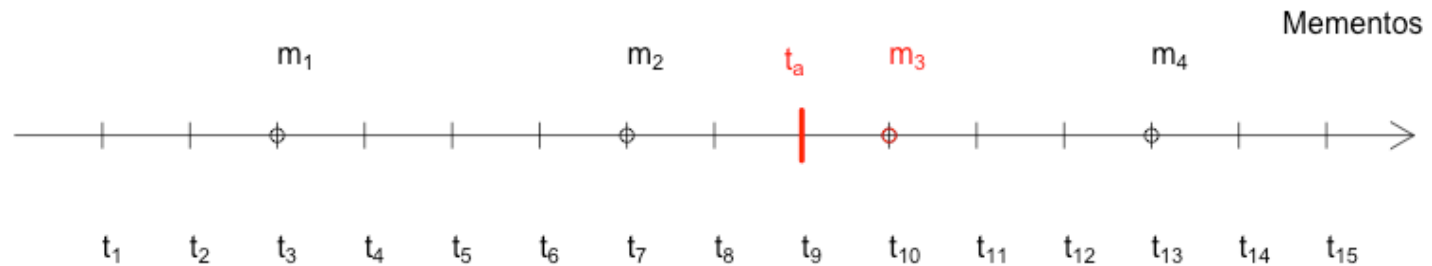
Can we process the content of m_3 and redirect the user to m_2 if we detect spoilers?

Can We Use The Extension To Avoid Spoilers For Sports And The News?

How Do We Use minfutr And minpast To Study Emerging Topics On Wikipedia

We Can Do Further Work On Missed Updates And Redundant Mementos

We Got Here…

CONCLUSIONS

We Introduced The Naïve Spoiler Concept

$$t_{r_j} \geq t_{e_i} \implies \textit{spoiler}$$

$$t_{r_j} < t_{e_i} \implies \textit{safe}$$

We Have Detailed TimeGate Heuristics

TABLE 7: Summary of TimeGate Heuristics

Heuristic	Uses cost metric	Uses Content of Mementos	Potential running time	Reliably avoids spoilers?
mindist	X		$O(n)$	no
minpast	X		$O(n)$	yes
minfutr	X		$O(n)$	no
minnear	X		$O(n)$	no
eqfutr	X	X	$O(n+c)$	no
eqpast	X	X	$O(n+c)$	no
simfutr	X	X	$O(f+n)$	no
simpast	X	X	$O(f+n)$	no

We Showed How To Calculate The Probability Of Encountering A Spoiler
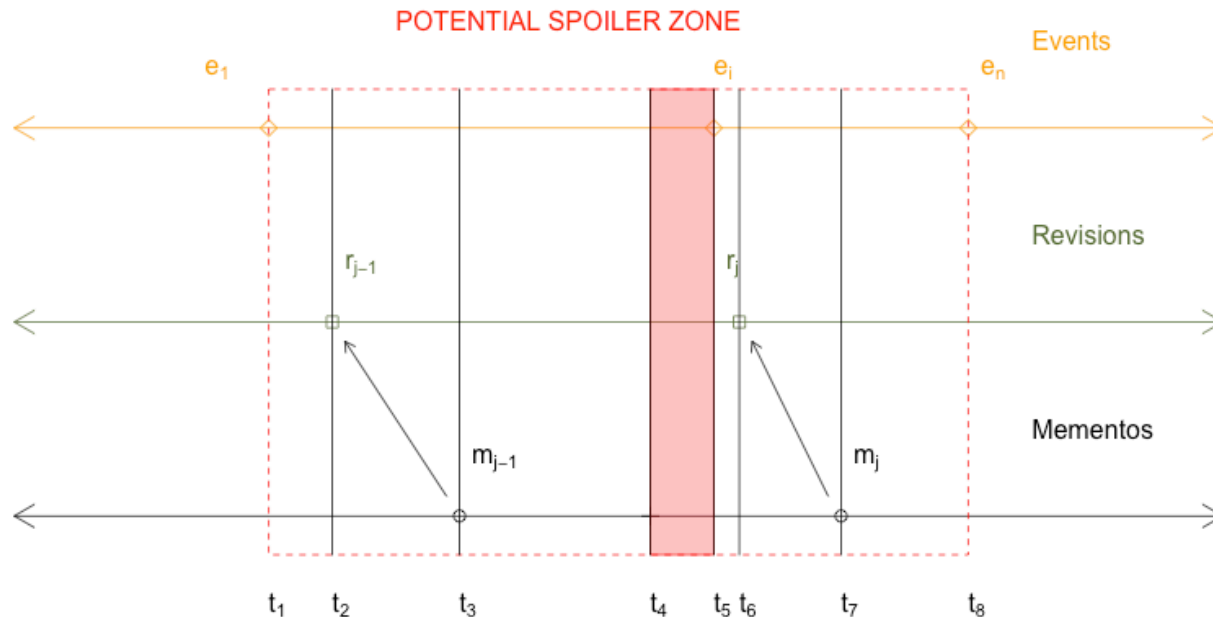
$$Pr(spoiler) = \frac{s}{c}$$

s = # of seconds where we are in a spoiler area
c = # of seconds between e_1 and e_n

We Calculated Real Spoiler Probabilities

TABLE 11: Spoiler probabilities for most popular pages within each fan wiki

Wiki	Page Name	Probability of Spoiler	# of Spoiler Areas	# of Revisions	# of Mementos
bigbangtheory	Sheldon Cooper	0.31	69	1958	30
boardwalkempire	Nucky Thompson	0.15	31	290	15
breakingbad	Walter White	0.43	40	882	20
continuum	Keira Cameron	0.54	21	104	5
downtonabbey	Sybil Branson	0.42	23	580	3
gameofthrones	Daenerys Targaryen	0.16	24	768	29
grimm	Nick Burkhardt	0.39	30	795	5
house-of-cards	Frank Underwood	0.0	13	380	3
how-i-met-your-mother	Barney Stinson	0.55	120	588	13
lostpedia	Kate Austen	0.67	94	3531	27
madmen	Mad Men Wiki	0.22	36	250	85
ncis	Abigail Sciuto	0.67	182	404	11
onceuponatime	Emma Swan	0.36	34	1210	11
scandal	Main Page	0.60	31	250	14
trueblood	Eric Northman	0.28	47	931	14
whitecollar	Neal Caffrey	0.29	38	199	8

We Showed That the Wayback Machine Is Serving Spoilers

Wayback Machine Requests For wikia.com: Spoilers vs Safe

We Developed The Memento MediaWiki Extension To Use minpast

Most of All We Showed That It Is Possible To Avoid Spoilers In MediaWiki Fan Sites Using Memento

Accept-Datetime: Sun, 13 April 2014 00:59:00 GMT

http://gameofthrones.wikia.com/wiki/Joffrey_Baratheon?oldid=125053

Papers/Presentations

- "Avoiding Spoilers in Fan Wikis of Episodic Fiction", with M. L. Nelson (in preparation).
- "Using the Memento MediaWiki Extension To Avoid Spoilers", Presentation, *WikiConference USA 2014,* June 2014.
- "Reconstructing the Past With MediaWiki: Programmatic Issues and Solutions", Presentation, *WikiConference USA 2014,* June 2014.
- "Bringing Web Time Travel to MediaWiki: An Assessment of the Memento MediaWiki Extension", with M. L. Nelson, H. Shankar, and H. Van de Sompel, Tech. Rep. avXiv:1406.3, Old Dominion University, June 2014.

Because What You Have Seen Is Just The Tip Of The Iceberg...

BACKUP SLIDES

Backup Slides

RELATED WORK

Existing Studies: Leavitt and Christenfeld
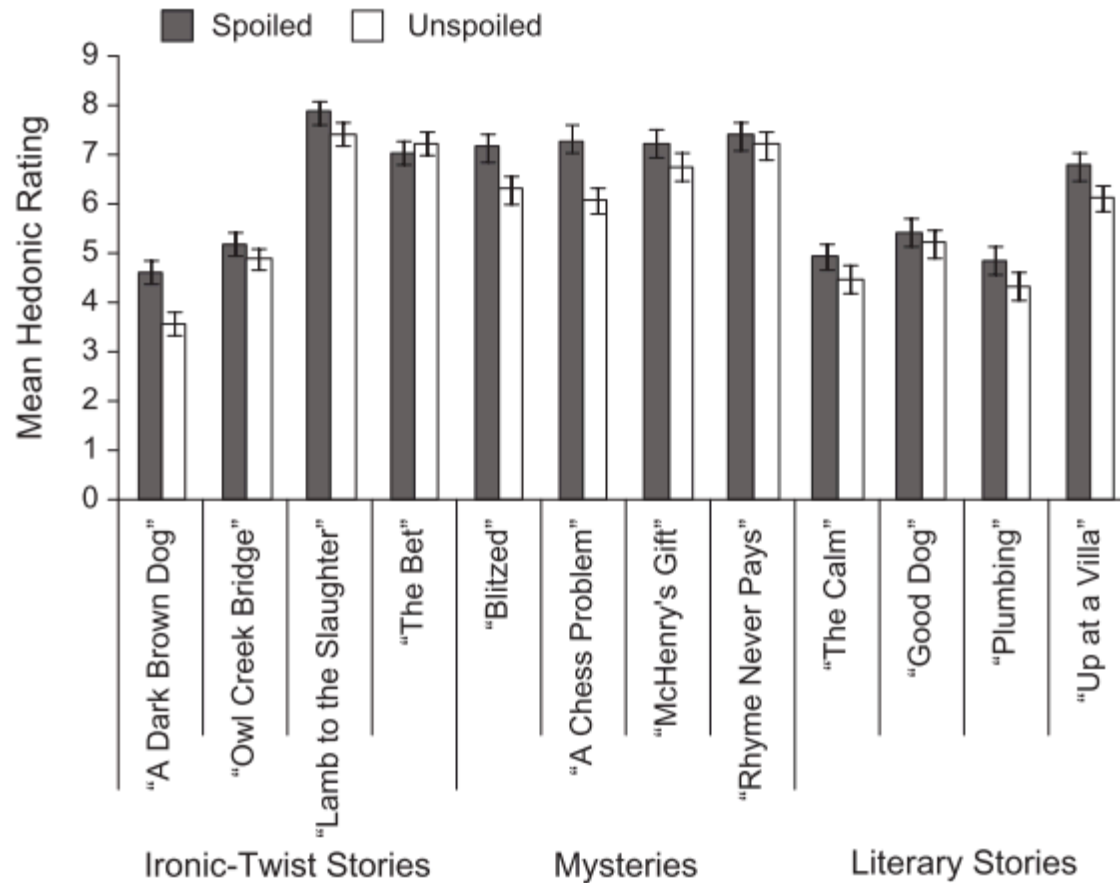
Existing Studies: Release Date Different Per Country

- Schirra, Sun, and Bently's live-tweeting study on *Downton Abbey* unearthed a global problem of avoiding spoilers in shows that air later for fans in a different country.

- This is further echoed by Leaver's essay *The Tyranny of Digital Distance* discussing the release of *Battlestar Galactica* episodes in Australia.

Existing Studies: Johns

- Johns studied two-screen viewing (generic name for live tweeting)

- Fans of particular shows would avoid social media until they had viewed the latest show

- These fans were trying desperately to avoid spoilers

- This contradicts Leavitt's study

Typically Notices Warn Us of Spoilers

The World of Ice and Fire (US, UK) will be released October 28. [edit]
Editors, please refrain from including information from the book prior to November 27.

Synopsis ✏Edit

Note: The following synopsis refers to the "Theatrical cut" version of the film, which features departures from the original novel.

Spoiler warning! Plot and/or ending details follow.

TO EVERYONE WHO WILL NOT RECEIVE SERIES 4 OF DOWNTON ABBEY UNTIL JANUARY 2014, THIS IS A **WARNING**: THIS WIKI WILL BE UPDATED AND EDITED AS THE EPISODES BROADCAST IN THE UNITED KINGDOM (SEPTEMBER - DECEMBER 2013)! PROCEED AT YOUR OWN RISK! **THERE WILL BE SPOILERS ABUNDANT ON MOST OF THE PAGES!** YOU HAVE BEEN WARNED!

MAJOR SPOILER ALERT: This story contains many details of Thursday's season finale episode of ABC's Scandal.

PRETTY LITTLE LIARS #FATALFINALE (Spoiler Alert!)
Seventeen
Subscribe 403,369
25,819
1,474 11

Some Sites Try To Hide The Spoilers With HTML And/Or Javascript

Existing Academic Studies Have Dealt With Social Media

- Separate studies conducted by Johns and the team of Schirra, Sun, and Bently studied two-screen viewing
- Boyd-Graber, Glasgow, and Zajac attempted to use machine learning to find spoilers in social media
- To avoid spoilers fans would avoid, or abandon:
  - Social media
  - Online web pages
  - TV shows

- This results in lost revenue to advertisers!

Existing Studies: Tsang and Yan

- Fans have abandoned
  - Social media
  - Online web pages
  - TV shows

- Because of the issues in avoiding spoilers

- This results in lost revenue to advertisers!

Hiding Spoilers: Microformats

Artjom Kurapov created a draft HTML microformat for classifying links and images.

Listing 3.1: Examples of xrate microformats for avoiding spoilers, pornography, and violence in links and images

```
<a href="http://www.example.com/who-my-character-fell-in-
    love-with" data-xrate-spoiler="100" data-xrate-sex="20
    ">link on information about this episode</a>

<img src="http://www.example.com/picture-of-character-
    finally-dying" data-xrate-spoiler="100" data-xrate-
    violence="60" />
```

Hiding Spoilers: Tweetdeck

Hiding Spoilers: Tumblr Savior
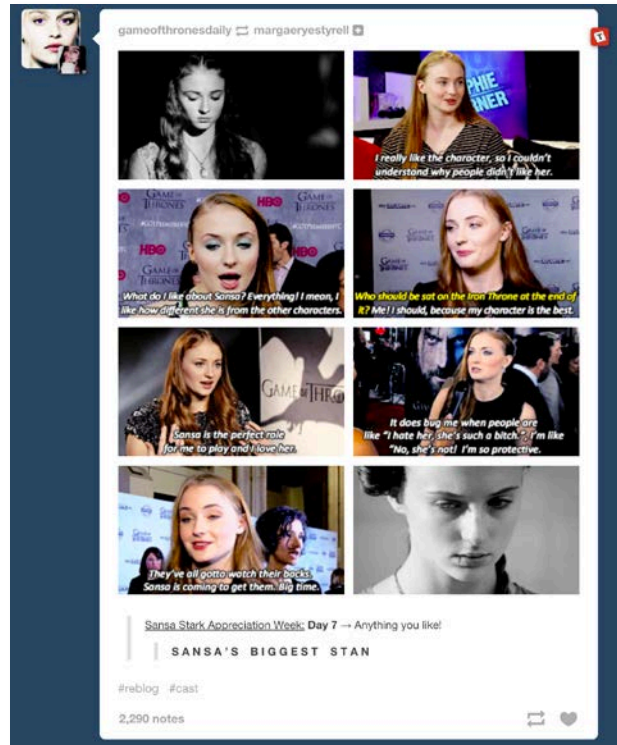
Hiding Spoilers: Netflix Spoiler Foiler
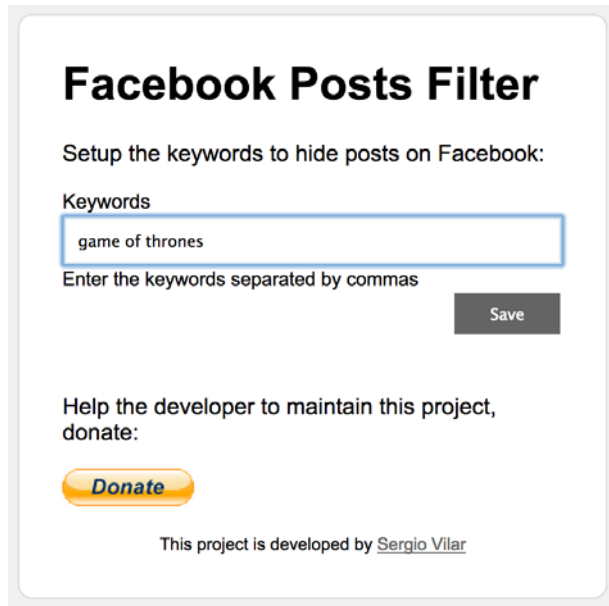
Hiding Spoilers: Netflix Spoiler Foiler
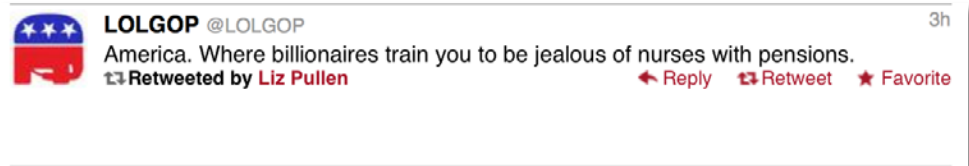
Spoiler Shield Is A Spoiler Blocking Tool For Social Media
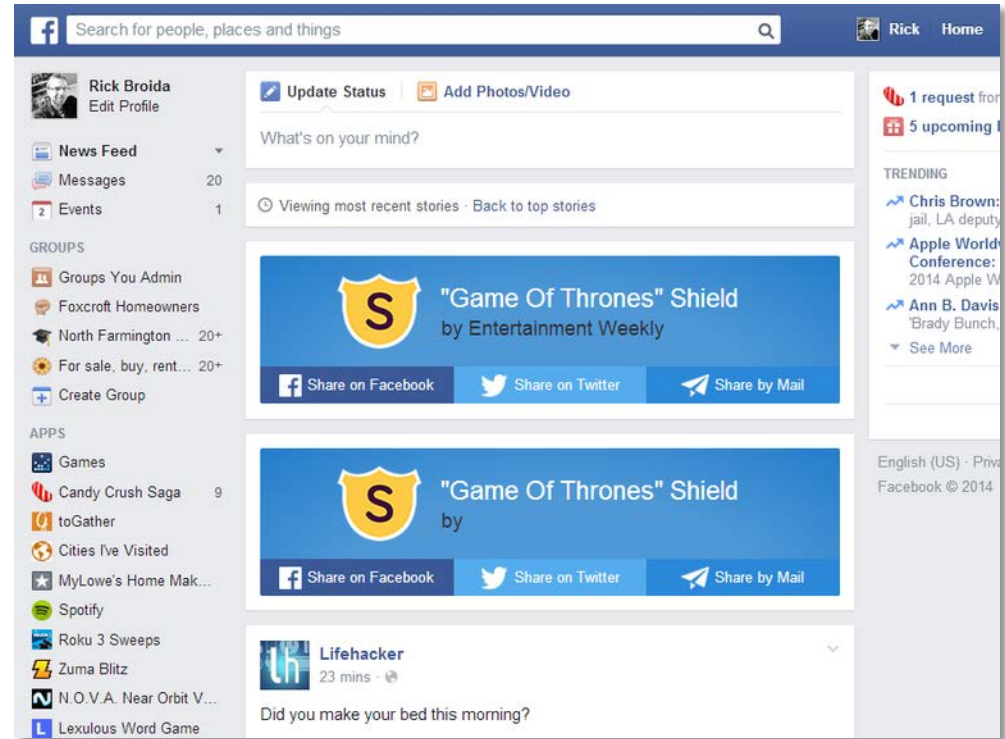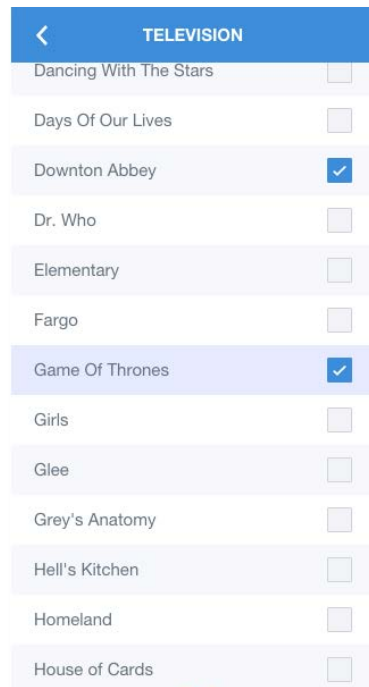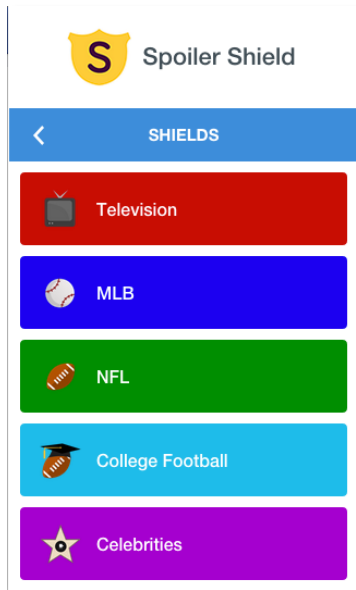
Hiding Spoilers: Facebook Posts Filter And Open Tweet Filter
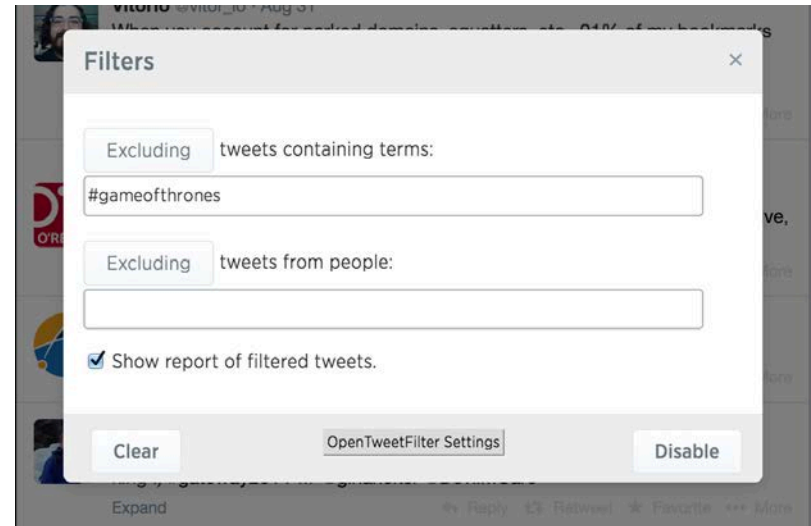
Any Facebook posts with these keywords will not be displayed.

Any tweets with these keywords will not be displayed.

Archiving In MediaWiki: Parsoid
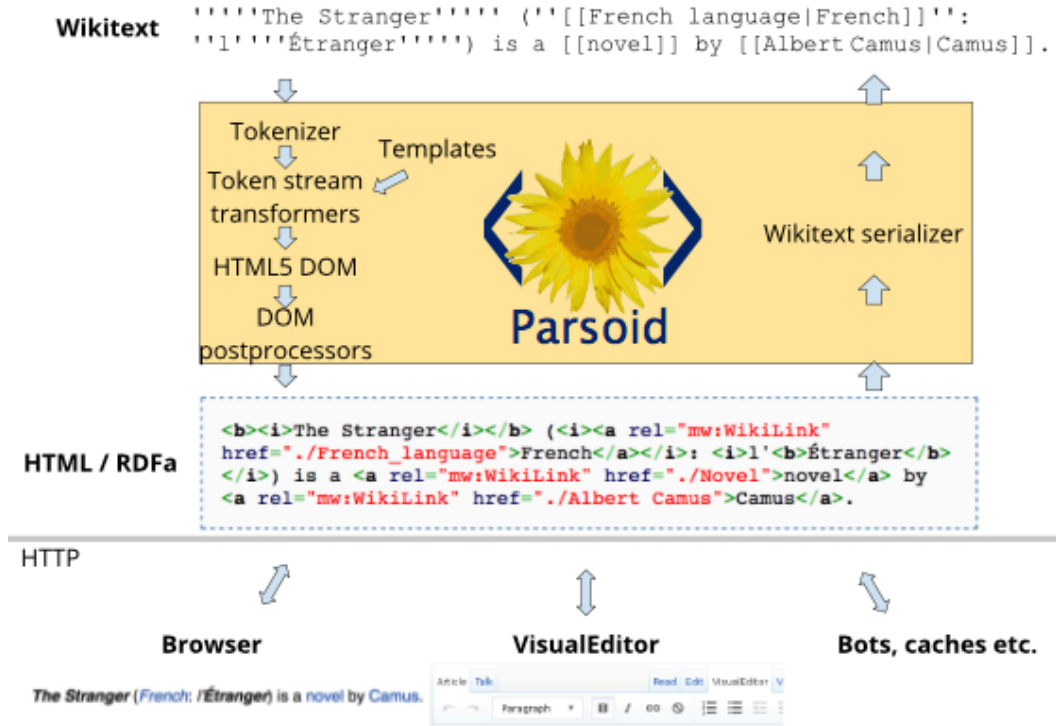
Archiving In MediaWiki: Collection Extension

Collection
From Simple English Wikipedia - the free encyclopedia that anyone can change

You can collect articles, generate and download a PDF file from article collections, order books from a print-on-demand partner and save article collections for later use or to share them.

See Help:Collections for more information about collections.

My Collection

Title: Music Wiki Selection
Subtitle: Articles from Wikipedia

Contents

[Create new chapter] [Sort articles alphabetically] [Clear collection]

- ▼ **Basic topics**   [Rename] [Remove]
  - ▲ ▼ Falsetto   [Remove]
  - ▲ ▼ Singing   [Remove]
  - ▲ ▼ Soprano   [Remove]
  - ▲ ▼ Mezzo-soprano   [Remove]
  - ▲ ▼ Baritone   [Remove]
  - ▲ ▼ Harpsichord   [Remove]
- ▲ ▼ **Composers**   [Rename] [Remove]
  - ▲ ▼ Ludwig van Beethoven   [Remove]
  - ▲ ▼ Johannes Brahms   [Remove]
  - ▲ ▼ Wolfgang Amadeus Mozart   [Remove]
  - ▲ Joseph Haydn   [Remove]

Order Printed Book

You can order a printed book containing your article collection by visiting one of the following print-on-demand partners:

- **Order book from PediaPress** – About PediaPress

Download Collection as PDF

To download an automatically generated PDF file of your article collection, click the button.

Download PDF

Save Collection

If you want to save collections for later use, please log in or create an account.

Seamlessly Viewing Past Versions Of MediaWiki Pages
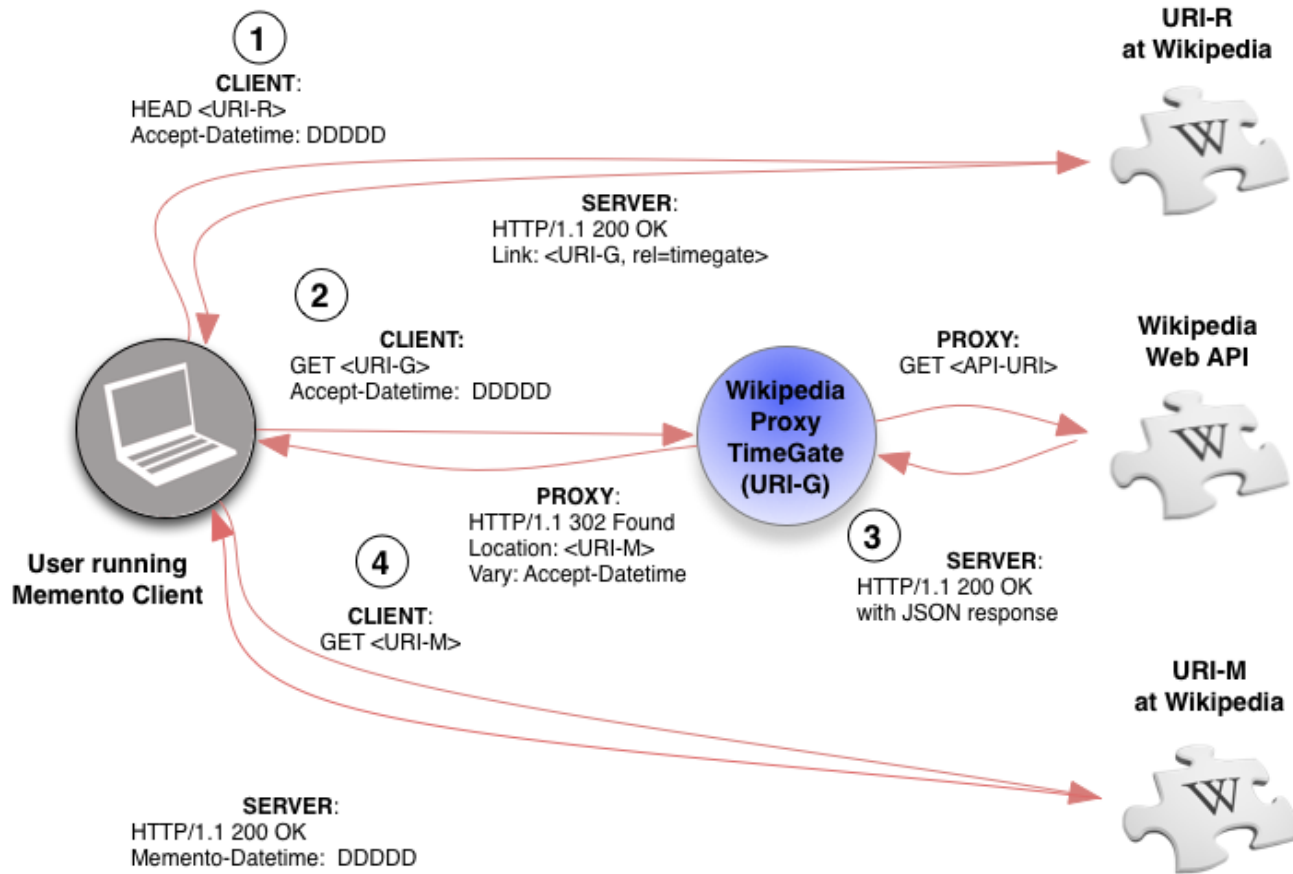
Seamlessly Viewing Past Versions Of MediaWiki Pages
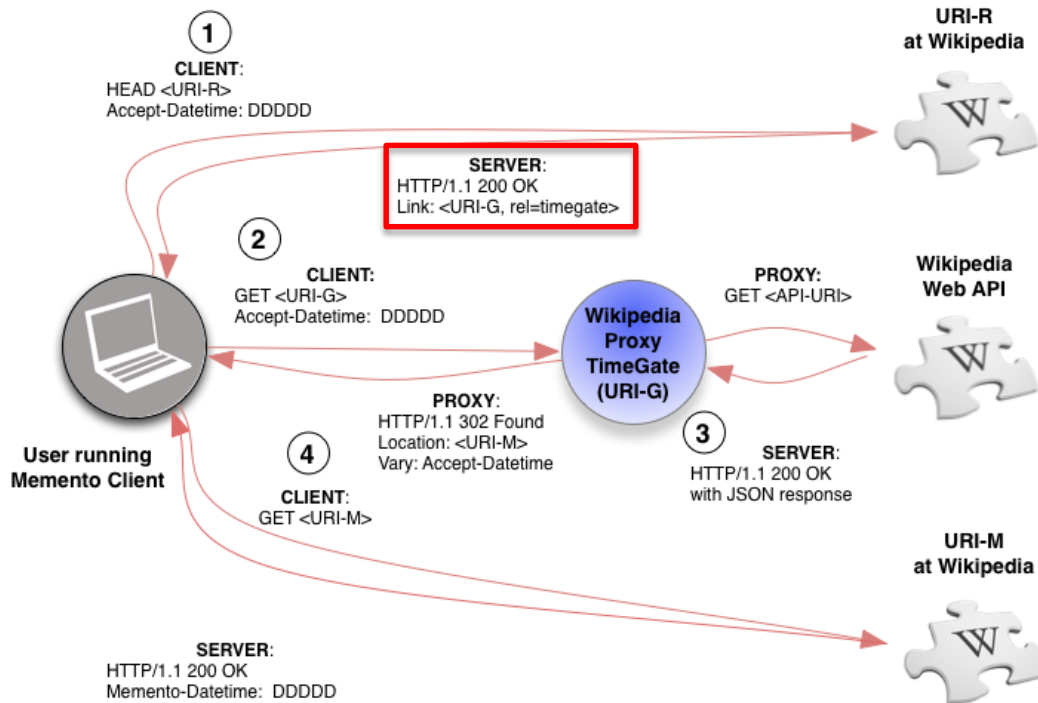
Wikipedia Proxy Does Not Address The Problem For All Wikis.

To Be Generic For Any Wiki, The Wiki Server Itself Would Need To Send A Link Header Back Indicating The Uri Of The TimeGate To Use.

Backup Slides

ALTERNATIVES TO THE MEMENTO MEDIAWIKI EXTENSION (AND WHY IT IS BETTER!)

Memento Extension vs. Manually Getting Page Revision
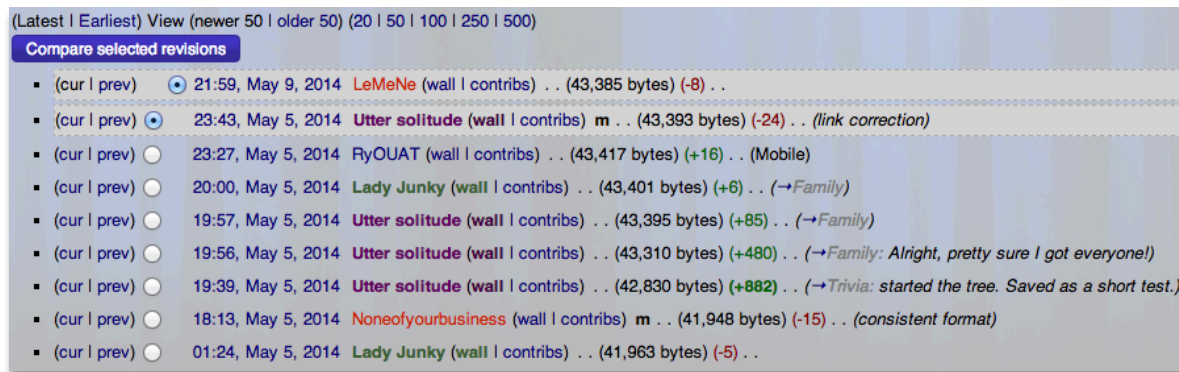
This Is Very Time Consuming.

Memento Let's You *Browse* Through The Whole Web With A Given Date!
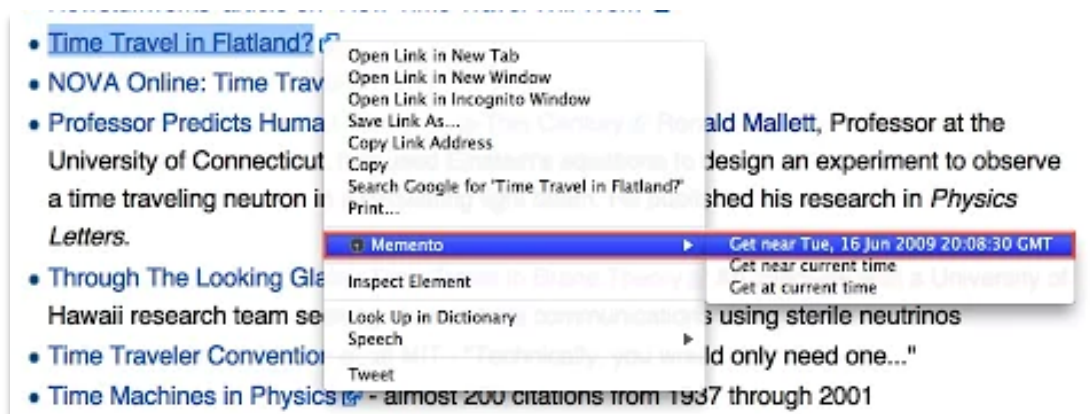
Why Do It When Memento Will Do It For You?

Memento Extension vs. MediaWiki API

JSON:

{"revid":607345961,"parentid":607210719,"timestamp":"2014-05-06T16:07:52Z"}

XML:

<rev revid="607519915" parentid="607345961" user="Marklemagne"
timestamp="2014-05-07T19:00:26Z"/>

Only A Custom MediaWiki Client Can Turn These Oldid Entries Into Uris.

Memento Is A Web Standard Way Of Accessing Past Web Resources And Is Already Implemented For Many Different Applications (Web Archives, Etc.)

Memento Extension vs. MediaWiki API

Link: <http://ws-dl-05.cs.odu.edu/demo-302-recommended-relations/index.php/Daenerys_Targaryen>; rel="original latest-version",

 <http://ws-dl-05.cs.odu.edu/demo-302-recommended-relations/index.php/Special:TimeGate/Daenerys_Targaryen>; rel="timegate",

 <http://ws-dl-05.cs.odu.edu/demo-302-recommended-relations/index.php/Special:TimeMap/Daenerys_Targaryen>; rel="timemap"; type="application/link-format"; from="Sun, 22 Apr 2007 15:01:20 GMT"; until="Fri, 27 Sep 2013 20:48:24 GMT",

 <http://ws-dl-05.cs.odu.edu/demo-302-recommended-relations/index.php?title=Daenerys_Targaryen&oldid=1499>; rel="first memento"; datetime="Sun, 22 Apr 2007 15:01:20 GMT",

 <http://ws-dl-05.cs.odu.edu/demo-302-recommended-relations/index.php?title=Daenerys_Targaryen&oldid=107643>; rel="last memento"; datetime="Fri, 27 Sep 2013 20:48:24 GMT"

Memento Also Follows The RESTful Principle Of "Follow Your Nose", Indicating Additional Resources To Access From Here.

Memento Extension vs. Internet Archive
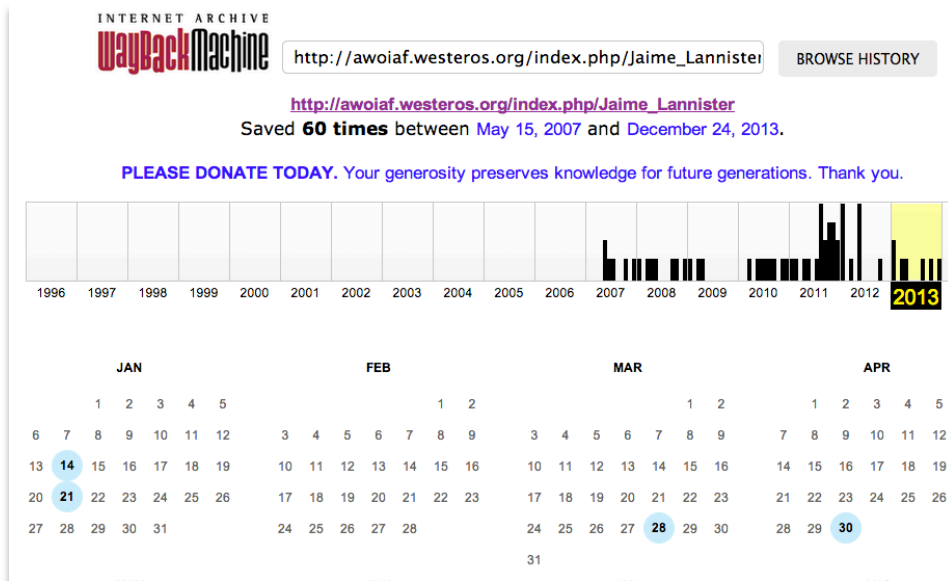
The Internet Archive Only Gets **Some** Of The Revisions Of A Given Page.

MediaWiki Has **All** Of The Revisions Of A Given Page.

Memento Extension vs. Other MediaWiki Time Travel Extensions

While These Extensions Just Work For MediaWiki, ***Memento Works For The Entire Web.***

With The Memento Extensions, One Can Browse The Entire Web Spoiler Free By Seamlessly Accessing Web Archives And Other Resources Through Memento.

Backup Slides

MEMENTO PATTERN COMPARISON IN THE MEMENTO MEDIAWIKI EXTENSION

Assuming An Original Resource Is a TimeGate (Pattern 1.1)
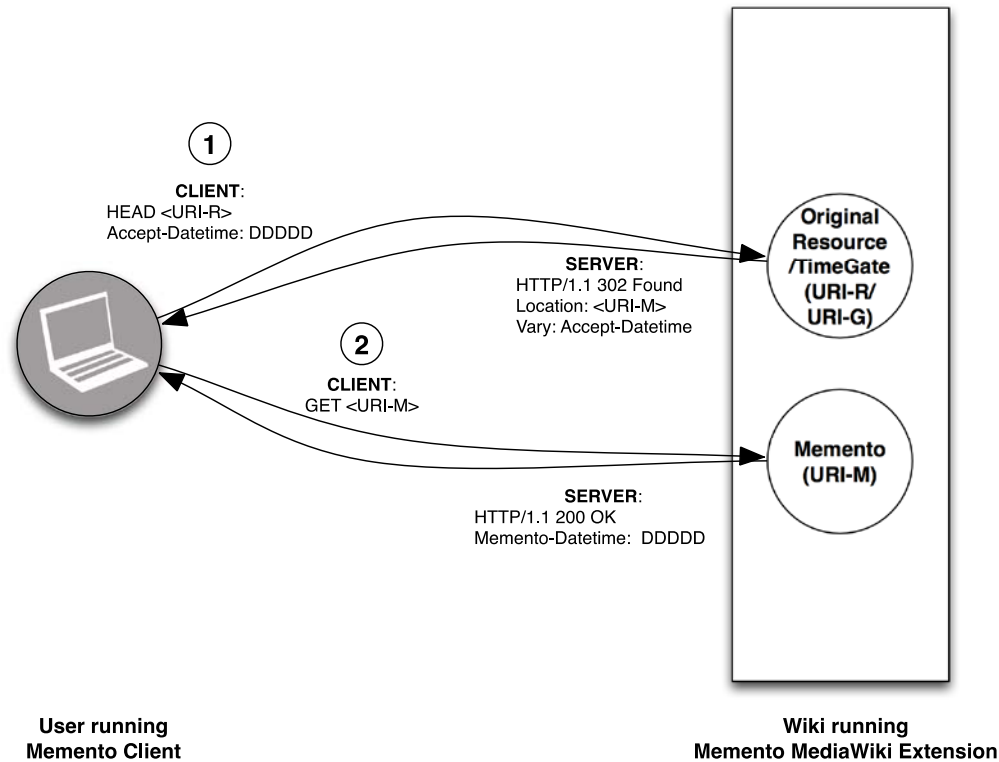
$$d_{p1.1} = B + RTT_B + M + RTT_M$$

Looking For A TimeGate (Pattern 2.1)

1

CLIENT:
HEAD <URI-R>
Accept-Datetime: DDDDD

SERVER:
HTTP/1.1 200 OK
Link: <URI-G, rel=timegate>

2

CLIENT:
GET <URI-G>
Accept-Datetime: DDDDD

SERVER:
HTTP/1.1 302 Found
Location: <URI-M>
Vary: Accept-Datetime

3

CLIENT:
GET <URI-M>

SERVER:
HTTP/1.1 200 OK
Memento-Datetime: DDDDD

Original Resource (URI-R)

TimeGate (URI-G)

Memento (URI-M)

**User running
Memento Client**

**Wiki running
Memento MediaWiki Extension**
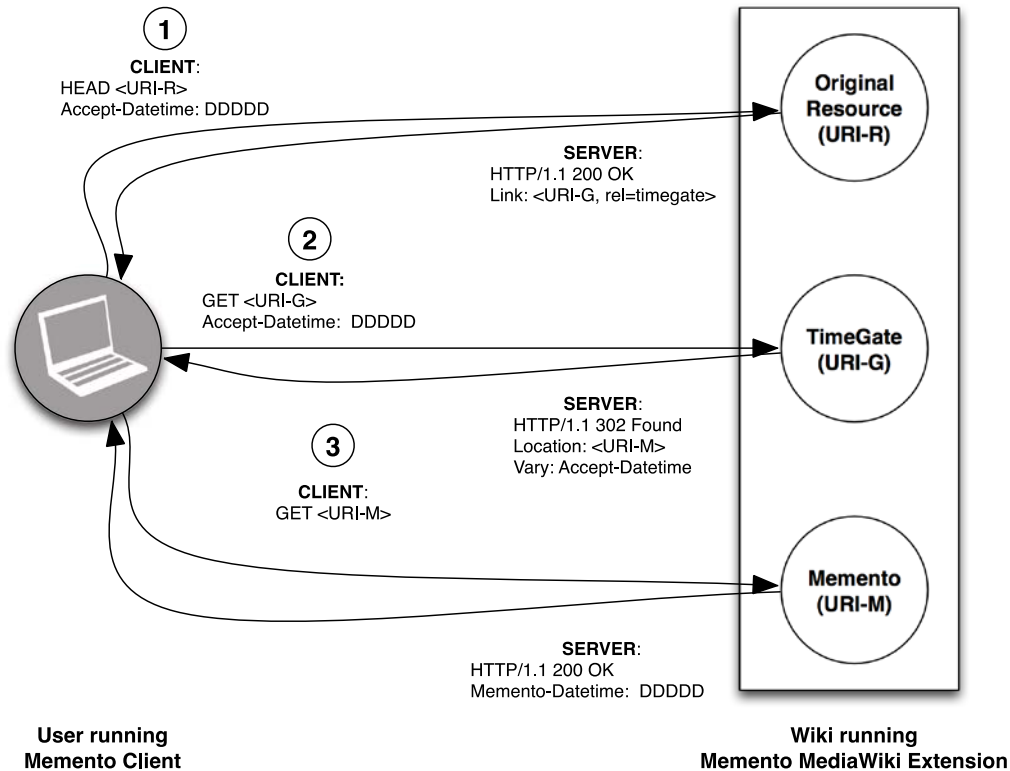
$$d_{p2.1} = a + RTT_a + b + RTT_b + M + RTT_M$$

Comparing These Two Involves Evaluating Their Performance

$$d_{p1.1} < d_{p1.2}$$

$$B + RTT_B + M + RTT_M < a + RTT_a +$$

$$b + RTT_b + M + RTT_M$$

$$B + RTT_B < a + RTT_a + b + RTT_b$$

$$B + RTT_B < a + RTT_a + b + RTT_b$$

$$B < a + RTT_a + b$$

$$B < a + b + RTT_a$$

a = time to generate just a normal wiki page
b = time to perform datetime negotiation when the TimeGate is the same
B = time to perform datetime negotiation when the TimeGate is different
M = time to generate memento
RTT_a, RTT_b, RTT_B, RTT_M = round trip times for *a*, *b*, *B*, and *M*

Using curl
We Obtained the Value of *a*

$$B < a + b + RTT_a$$

With caching, *a* turns out to be about 0.1 seconds on average

Using seige
We Obtained Values for *b* and *B*

Histogram of datetime negotiation times for Special:TimeGate
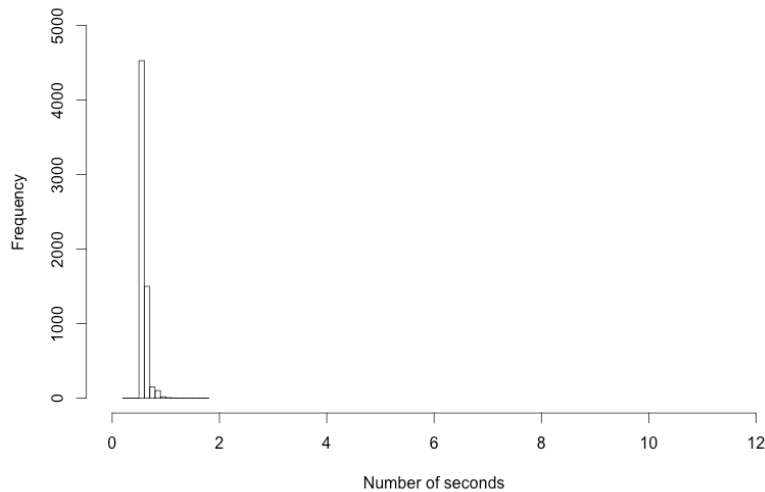
Histogram of datetime negotiation times for URI-R=URI-G
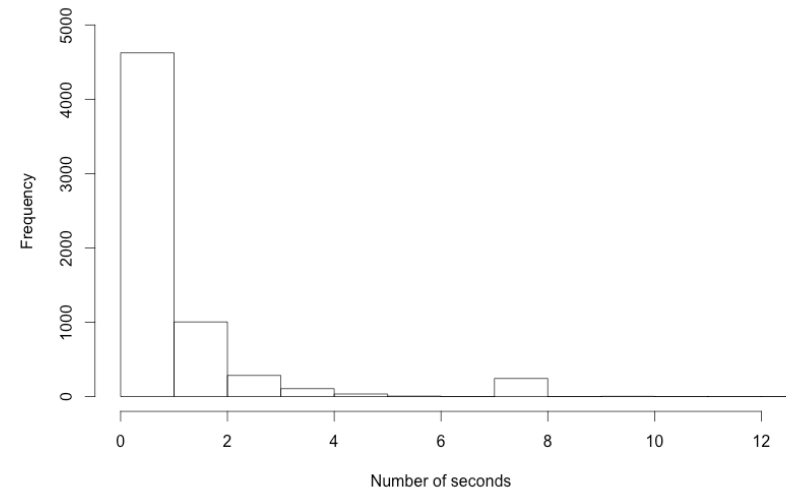
$$B < a + b + RTT_a$$

From these results:

$$0.22 \leq b \leq 1.75$$
$$0.56 \leq B \leq 12.06$$

Comparing TimeGate Implementations

Differences in URI-G Performance
For the Memento Mediawiki Extension
Between URI-R=URI-G and SpecialPage TimeGate Implementations

Worse Performance
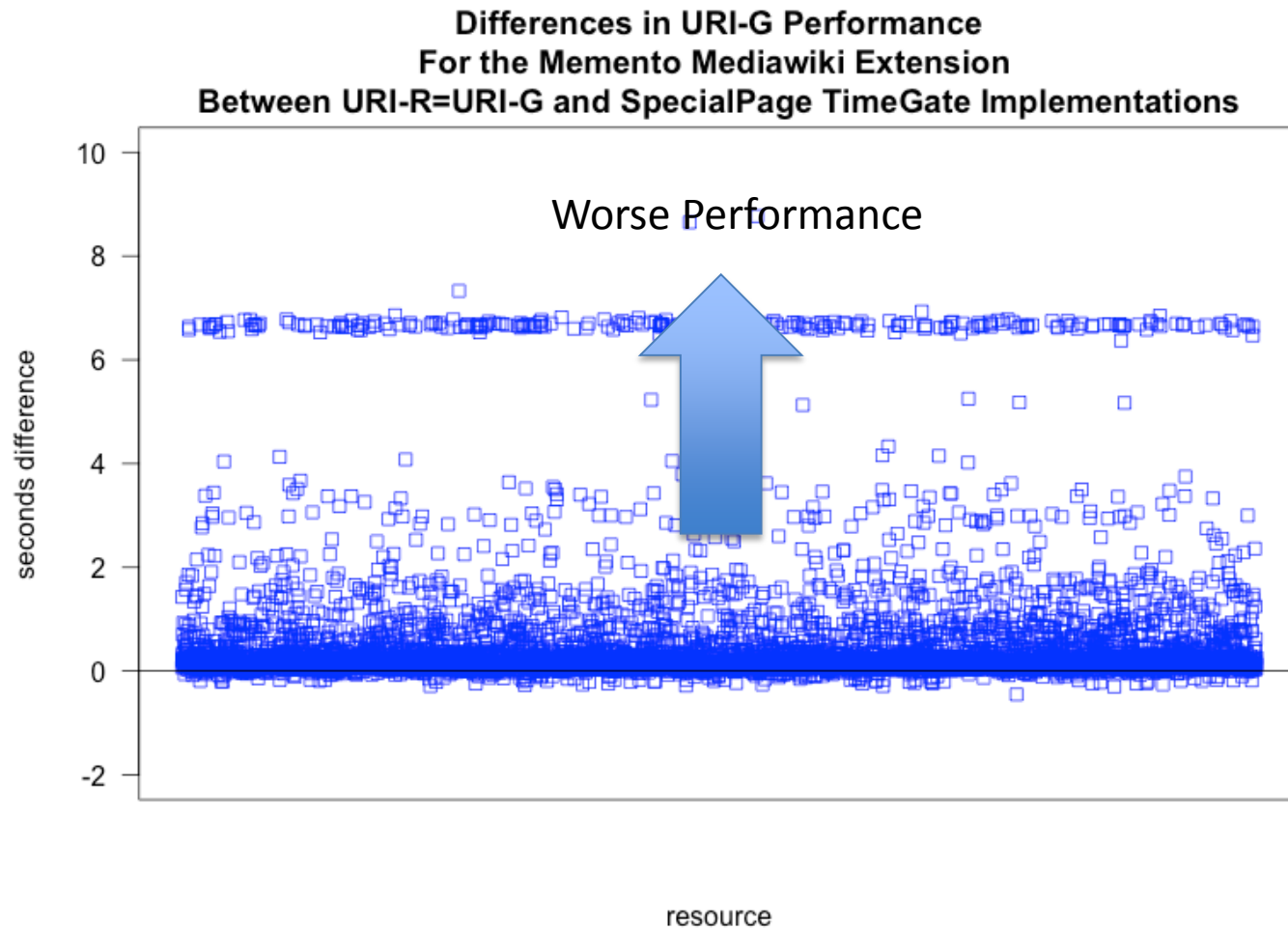
Using Analysis We Worked To Obtain The Value Of RTT_a

Round Trip Time is the sum of transmission delay and propagation delay

$$RTT = d_t + d_p$$

Transmission delay is # of bits divided by the rate of transmission

$$d_t = \frac{N}{R}$$

Requests and responses are typically about 11,840 bits

Assuming a worst case of 1G telephony (28,000 bps), d_t = 0.41 s

Using our previous values for *a, b*, and *B*, we see that such a 1G user would need to experience a d_p of 0.13s for the two-request pattern to perform better.

Continuing Our Analysis We Obtained The Value Of RTT_a

0.13s sounds small, but at the speed of light, this would require the user to be 24,216.7 miles from the server

This is almost the circumference of the Earth!!!

So, At What Value Of d_t Does The Two-Request Pattern Win Out?

$$1.24\ s < d_t + 0.1\ s + 0.6\ s$$

$$1.24\ s < d_t + 0.7\ s$$

$$0.54\ s < d_t$$

$$d_t = \frac{N}{R} \quad \text{From (20)}$$

$$0.54\ s = \frac{11840\ b}{R}$$

$$(0.54\ s)(R) = 11840\ b$$

$$R = \frac{11840\ b}{0.54\ s} = 21926\ bps$$

At bandwidth less than 21,926 bps, the two-request pattern wins out

That's slightly better than 1G telephony!

For most users, the three-request pattern performs better!

ADDITIONAL ARCHIVE EXTANT SAFE AREAS

These Are Identified By The Order Of Occurrence Of Halfway, Revision, and Event

- For example
  - HRE – means:
    1. Halfway mark between two mementos
    2. Revision is created
    3. Event occurs
  - ERH – means:
    1. Event occurs
    2. Revision is created
    3. Halfway mark between two mementos

Condition: Archive Extant Safe Area REH
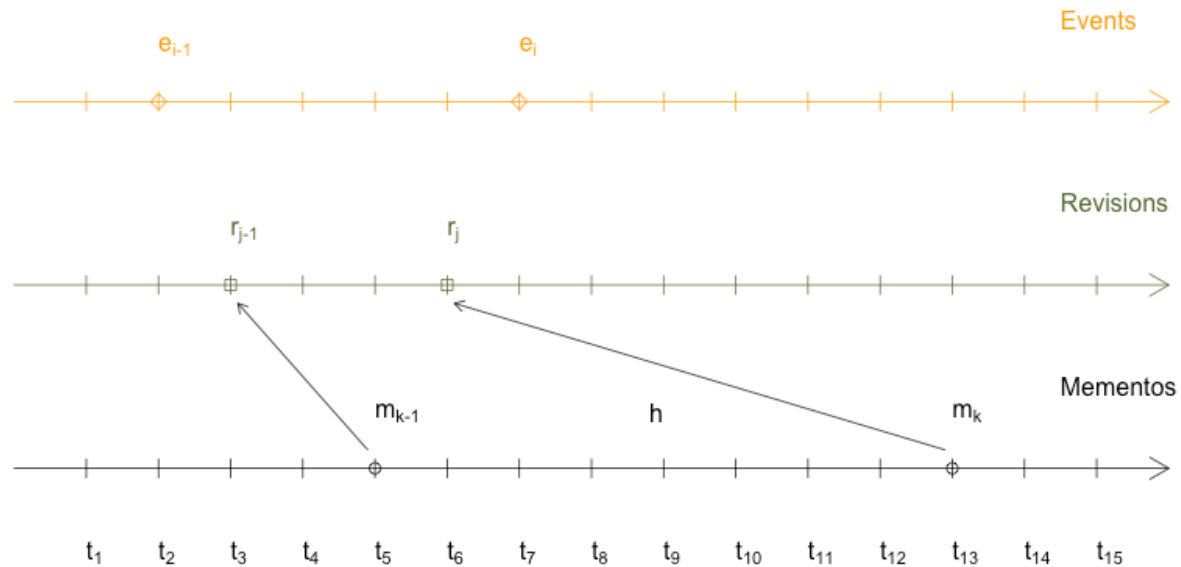
$$[t_s, t_f] = \mathcal{S}_b(e) = \begin{cases} (t_h, t_e) & \text{if} & t_h < t_e < t_{r_i} \ \wedge \ r_j \equiv m_k \ \wedge \\ & & t_h = \frac{t_{m_{k-1}} + t_{m_k}}{2} \\ (0, 0) & \text{otherwise} \end{cases}$$

Condition: Archive Extant Safe Area RHE
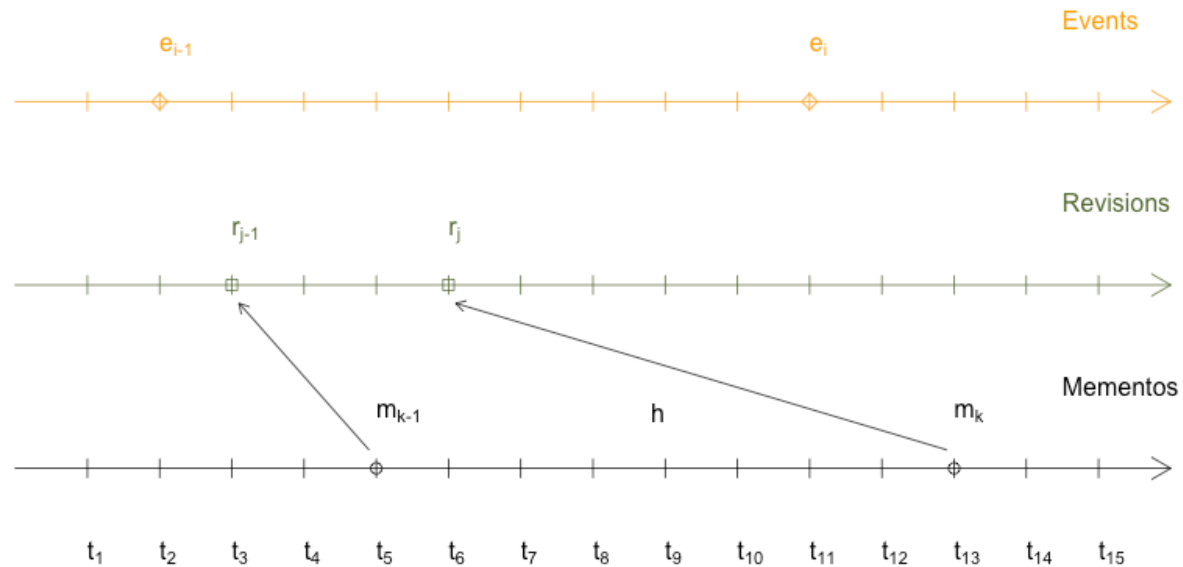
$$[t_s, t_f] = \mathcal{S}_b(e) = \begin{cases} (t_h, t_e) & \text{if} & t_h < t_e < t_{r_i} \;\wedge\; r_j \equiv m_k \;\wedge\; \\ & & t_h = \frac{t_{m_{k-1}} + t_{m_k}}{2} \\ (0,0) & \text{otherwise} \end{cases}$$

Condition: Archive Extant Safe Area ERH
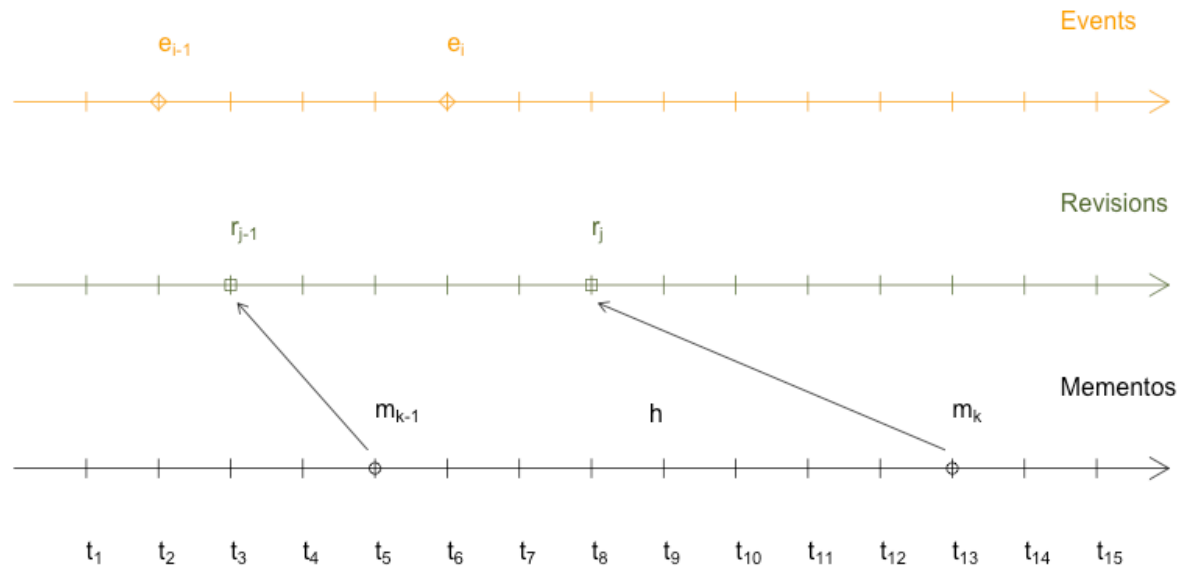
$$[t_s, t_f] = \mathcal{S}_b(e) = \begin{cases} (t_h, t_e) & \text{if} & t_h < t_e < t_{r_i} \ \wedge \ r_j \equiv m_k \ \wedge \\ & & t_h = \frac{t_{m_{k-1}} + t_{m_k}}{2} \\ (0,0) & \text{otherwise} \end{cases}$$

Condition: Post Archive Safe Area RE
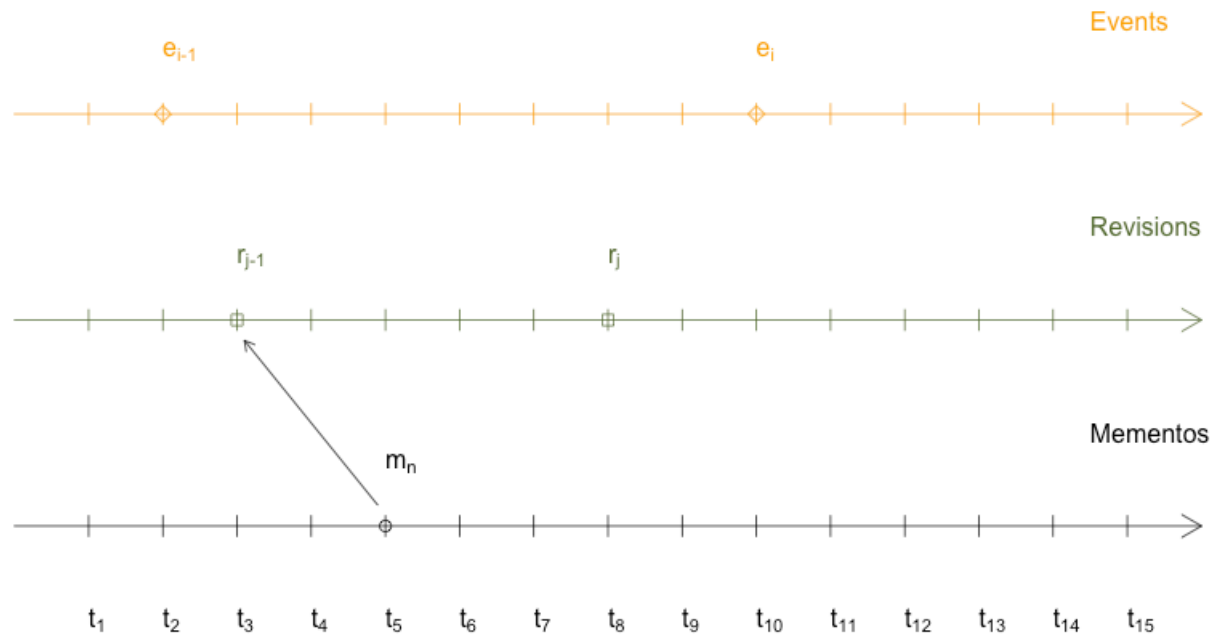
No spoilers after last memento m_n

Condition: Post Archive Safe Area ER
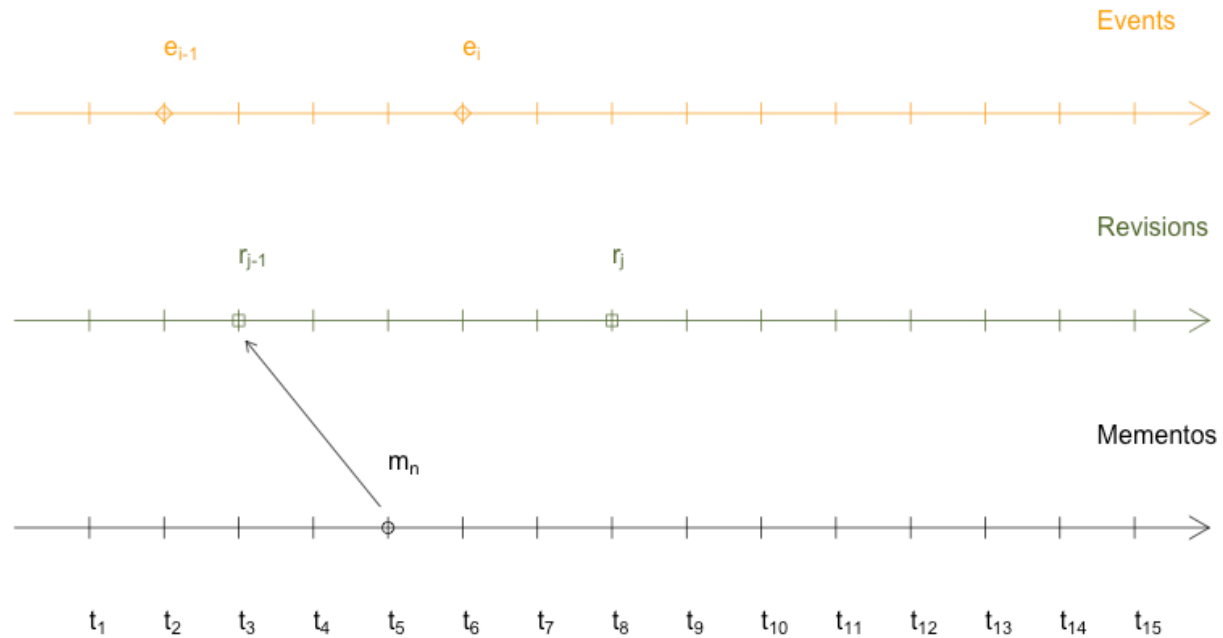
No spoilers after last memento m_n

Backup Slides

ARCHITECTURE OF THE MEMENTO MEDIAWIKI EXTENSION

We Used Class Inheritance For the Different Memento Resource Types

http://www.mediawiki.org/wiki/Extension:Memento
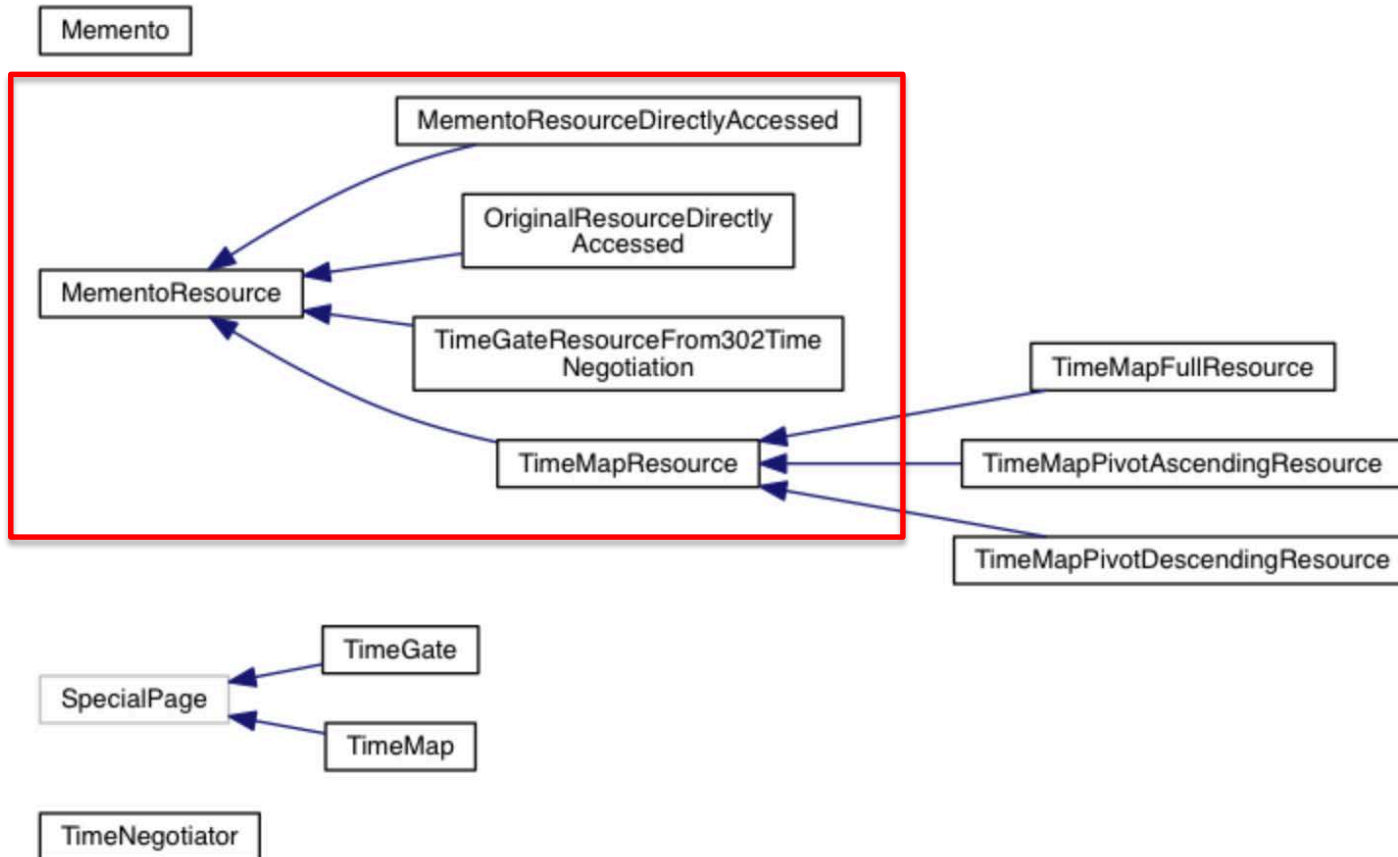
MediaWiki SpecialPages Invoke TimeGate And TimeMap Functionality

http://www.mediawiki.org/wiki/Extension:Memento
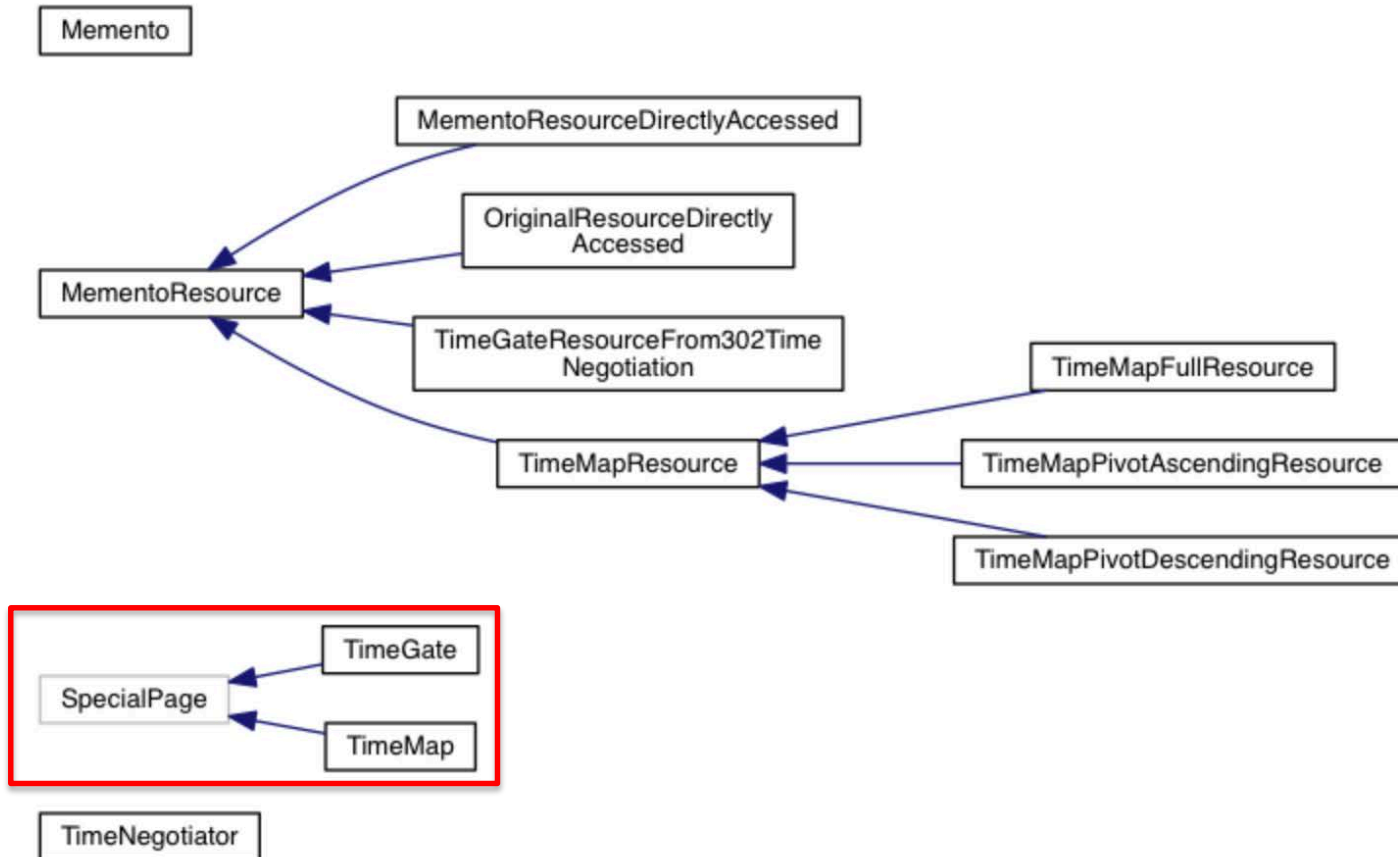
All Datetime Negotiation Is Centralized In The TimeNegotiator Class

http://www.mediawiki.org/wiki/Extension:Memento
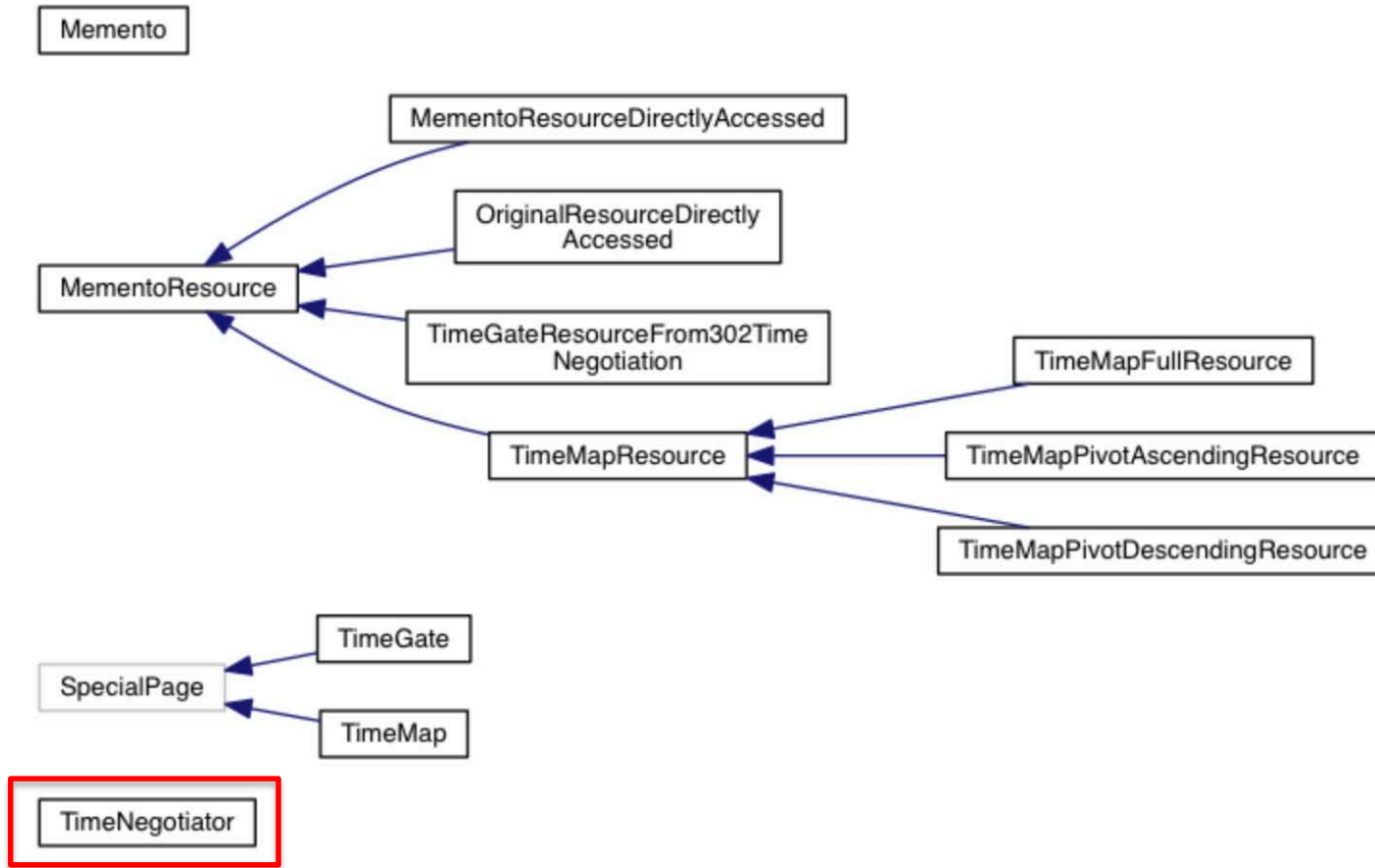
The Memento Class Is The Entry Point For The Extension
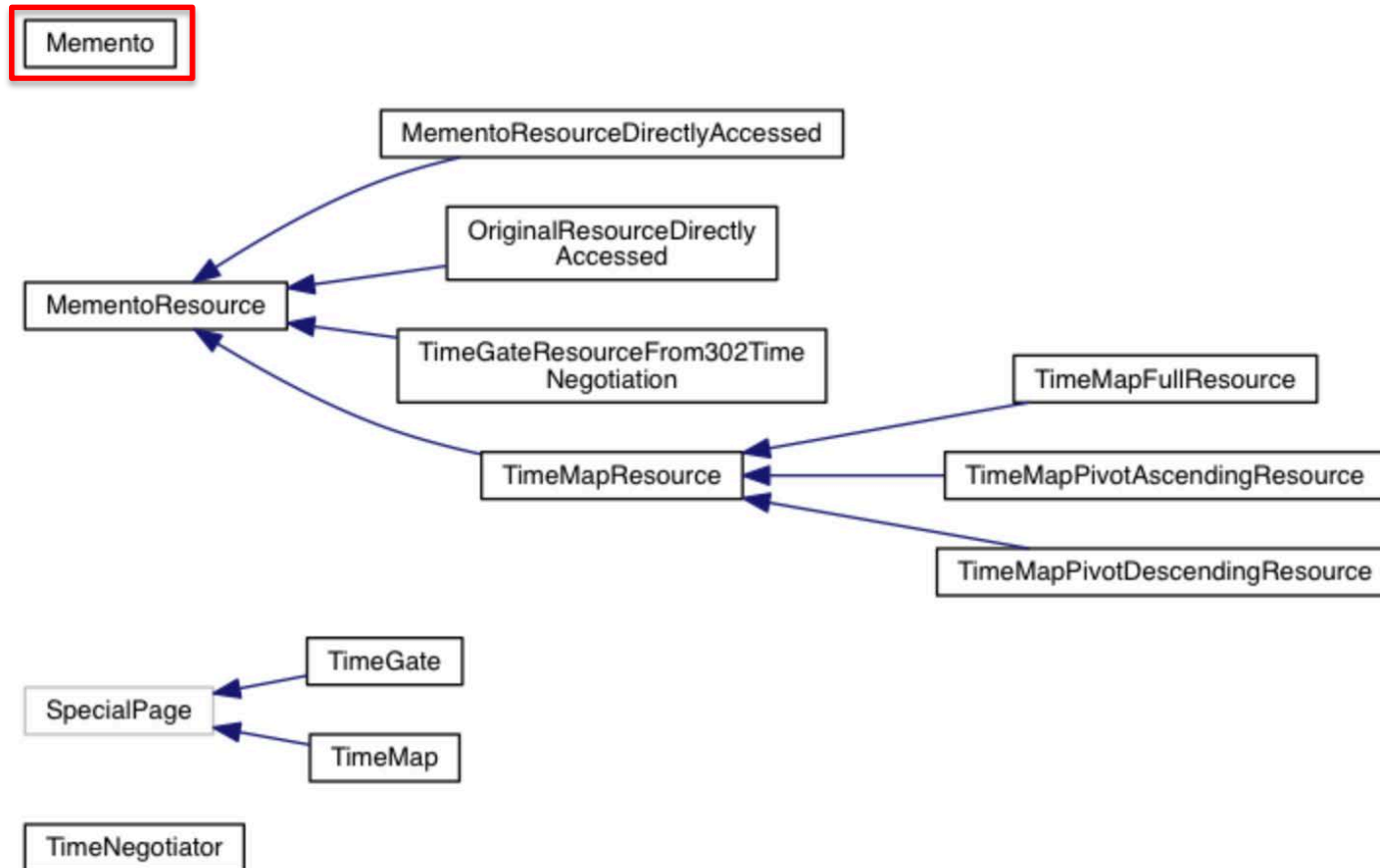
Backup Slides

MISCELLANEOUS

MediaWiki Still Has CSS Issues

Other Uses For The Memento MediaWiki Extension

Evolving laws and legal discourse

Changing relationship between organizations (ICANN vs. Verisign)

Past software contributions (Folding@Home)

Memento Headers Extension