THOMAS for the 21st Century:

A Proposal

Submitted by: Michael Plasmeier
                    Massachusetts Institute of Technology
Date: March 9, 2010
Revised: April 16, 2010

Abstract

THOMAS is the official online repository of legislative information for the United States Congress. THOMAS displays both legislation in progress and enacted and is run by the Library of Congress. When it was launched in 1995 it was revolutionary for the time. However, 15 years later, the site has only seen minor updates and has fallen behind the times. The site has many quirks and inconsistencies and requires extensive knowledge of how the system works. Information is split between many different government agencies, some of which are not even linked from THOMAS. Ordinary citizens cannot easily find the bills that Congress is actively considering among the thousands that are submitted without action each year. THOMAS should be rebuilt from scratch to track bills over time. The various sources of information such as the Congressional Record and C-SPAN should be integrated. Among other changes, news feeds should be added to provide at-a-glance access to recent changes. Amendments should be previewed on the actual bill; branch charts should be provided to link similar legislation. The Internet allows citizens to be more engaged in Congress than ever before. The constitutionally mandated "Journal of its Proceedings" for the House and Senate need to be updated to the 21st century to maximize Congress's accountability to the American people.

Table of Contents

Background and Aims

THOMAS is the official online website of legislative information for the United States Congress. THOMAS includes legislation both enacted and in progress. It is run by the Library of Congress. When it was launched in 1995 it was revolutionary for the time. However, 15 years later, the site has only seen minor updates and is no longer state of the art.

THOMAS is not very usable. When it came out, THOMAS made the near real-time information of Congress available online. At the time, this was a large step forward. Before THOMAS, one had to subscribe to the Congressional Record, a biweekly print publication, or use LexisNexis Congressional, which was a paid service. One could also go to a Federal Depository Library to read the Congressional Record. Information was not available in an immediate fashion, nor was it easily accessible. Ordinary citizens had to be fairly motivated to request the information and then wait a few days to receive it. Lawyers and lobbyists whose job was to deal with this information had people on staff to receive the papers in the mail and organize them. Merely having on-demand online information was clearly better than what existed. *good presentation of the problems and limitations of the previous "system"*

The problems with THOMAS are numerous. First, on the homepage, the search box has a toggle switch between keyword and bill number (Figure 1). This is unnecessary. A computer can easily determine when you enter a bill number by using elementary pattern matching. There is no reason that a person should have to select that they are entering a bill number.
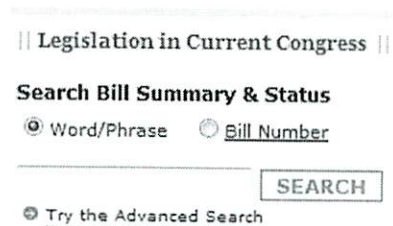
|| Legislation in Current Congress ||

Search Bill Summary & Status
◉ Word/Phrase   ○ Bill Number

SEARCH

◉ Try the Advanced Search

Figure 1 THOMAS Search Box

NEW SEARCH | HOME | HELP

Print Subscribe Share/Save

S.2806
Title: A bill to codify and enhance existing regulations designed to encourage individuals to adopt healthy behaviors through voluntary participation in programs of health promotion and disease prevention.
Sponsor: Sen Ensign, John [NV] (introduced 11/19/2009)   Cosponsors (1)
Related Bills: S.1796
Latest Major Action: 11/19/2009 Referred to Senate committee. Status: Read twice and referred to the Committee on Health, Education, Labor, and Pensions.

All Information (except text)	Text of Legislation	Summary	Major Congressional Actions
Titles	Cosponsors (1)	Committees	All Congressional Actions
Related Bills	Amendments	Related Committee Documents	All Congressional Actions with Amendments With links to *Congressional Record* pages, votes, reports
CBO Cost Estimates	Subjects		

THOMAS Home | Contact | Accessibility | Legal | FirstGov

Figure 2 THOMAS Bill Homepage

Once you do enter a bill number, you are brought to a page with many links, but not much information. The page's browser title bar is "Search Results - THOMAS (Library of Congress)," which provides no information about the page you are currently on. It makes navigation between pages in history impossible and is not easy for search engines to index. Thankfully the URL is unique and permanent which makes sharing and saving of the link possible. However, the URL contains extraneous information and is not easily readable. Clean URLs would be preferable. These would quickly transverse the hierarchy. For example:

do you mean "traverse"?

"111.congress.gov/s1234/text" would be for the text of the 111th Senate bill number 1234.

Basic Information

The bill's homepage presents a small summary of the bill and its current status along with many links (Figure 2). The "All information" page pulls together some of the other pages on this bill's site (Figure 3). However, it certainly does not include ALL information but it also duplicates some information from the home screen. The

NEW SEARCH | HOME | HELP

Print Subscribe Share/Save

S.2806
Title: A bill to codify and enhance existing regulations designed to encourage individuals to adopt healthy behaviors through voluntary participation in programs of health promotion and disease prevention.
Sponsor: Sen Ensign, John [NV] (introduced 11/19/2009)   Cosponsors (1)
Related Bills: S.1796
Latest Major Action: 11/19/2009 Referred to Senate committee. Status: Read twice and referred to the Committee on Health, Education, Labor, and Pensions.

Jump to: Summary, Major Actions, All Actions, Titles, Cosponsors, Committees, Related Bill Details, Amendments

SUMMARY:

NONE

MAJOR ACTIONS:

NONE

ALL ACTIONS:

11/19/2009:
    Read twice and referred to the Committee on Health, Education, Labor, and Pensions.

TITLE(S): (italics indicate a title for a portion of a bill)

NONE

COSPONSORS(1), ALPHABETICAL [followed by Cosponsors withdrawn]:   (Sort: by date)

    Sen Carper, Thomas R. [DE] - 11/19/2009

COMMITTEE(S):

    Committee/Subcommittee: Activity:
    Senate Health, Education,   Referral, In Committee
    Labor, and Pensions

RELATED BILL DETAILS: (additional related bills may be indentified in Status)

    Bill: Relationship:
    S.1796   Related bill identified by CRS

AMENDMENT(S):

NONE

THOMAS Home | Contact | Accessibility | Legal | FirstGov

Figure 3 THOMAS Bill All Information

list of titles is both on summary page, as well as its own page. Why must this be in two places at once?

The redesigned site should place this information on the bill's homepage in a well-designed manner to

segment the different types of information. For example, instead of a large banner or link reading

"Committees" the list of committees should simply be provided. Providing the information upfront

makes it more visible and reduces the number of clicks needed to find things.

Help and Process

THOMAS does provide a link to a help page

which explains what the difference between the

titles is. However, the link is not very visible. I have

been using THOMAS for years and only discovered

the link today (Figure 4 red box). Small explanation

texts should be places around the redesigned site

where possible to explain terms and procedures

that which people may not be familiar with. Where needed, small help icons should be placed near to

relevant sections to provide a lengthier explanation of procedures and terms.
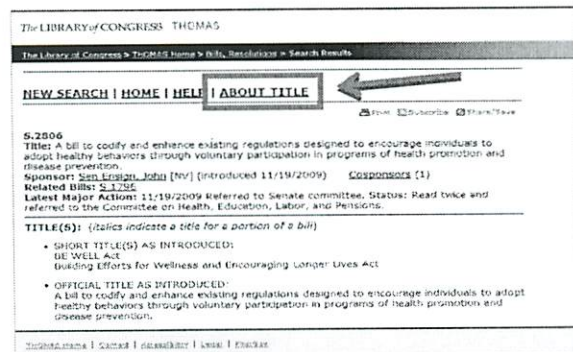
Figure 4 THOMAS Title page for S.2806 with attention drawn to the About Title link

Four Different Views

The page containing the text of the legislation has even more duplicate features and unneeded links.

The links on the top of the text page seem to be useless (Figure 5 Blue Box). Most of them are not even

links; they are just placeholders for links. The other links to the homepage and help are in a different

format than other places on the site. The text itself is displayed in a readable manner, as opposed to

being in a font that which is hard to read on screen. THOMAS also has a link to an XML version of the bill.

This is very well done, as the URL is permanent and predictable, and an XML schema is defined and a

style sheet is available. Providing an XML format with a permanent and predictable URL makes it easy

for developers to reuse the information in innovative applications at no cost to the government.

However, THOMAS provides four ways to print a bill, which all result in a different looking document (Figure 5 Red Box). The link with the printer icon invokes the browser's print function to print the current page. The site does have a print style sheet, but it does not strip out all of the navigation elements which should not be printed (Figure 6a). Printing the navigation elements wastes paper and clutters the printout with useless navigation links. Second, there is a link to a "Printer Friendly" version of the text. This is on a separate HTML webpage and does not include navigational elements (Figure 6b). In a browser which *that* supports print style sheets, this separate "Printer Friendly" page is not needed. In addition, with a browser that supports XML definitions and style sheets, including all the latest versions of major browsers, the XML version can also be printed (Figure 6c). Lastly, the PDF version can also be printed. The PDF version is the cleanest, most official looking, and easiest to read version of the bill (Figure 6d). The four different printing options result in four very different looking pages.
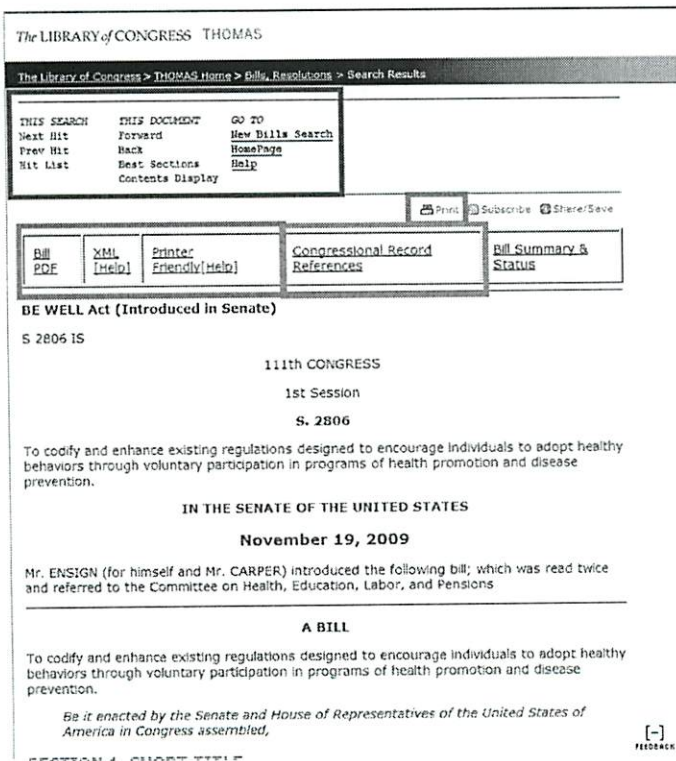
Figure 5 THOMAS Full Text Page

The Library of Congress > THOMAS Home > Bills, Resolutions > Search Results

THIS SEARCH	THIS DOCUMENT	GO TO
Next Hit	Forward	New Bills Search
Prev Hit	Back	HomePage
Hit List	Best Sections	Help
	Contents Display	

Print   Subscribe   Share/Save

Bill PDF	XML [Help]	Printer Friendly[Help]	Congressional Record References	Bill Summary & Status

BE WELL Act (Introduced in Senate)

S 2806 IS

111th CONGRESS

1st Session

S. 2806

To codify and enhance existing regulations designed to encourage individuals to adopt healthy behaviors through voluntary participation in programs of health promotion and disease prevention.

IN THE SENATE OF THE UNITED STATES

November 19, 2009

Mr. ENSIGN (for himself and Mr. CARPER) introduced the following bill; which was read twice and referred to the Committee on Health, Education, Labor, and Pensions

A BILL

To codify and enhance existing regulations designed to encourage individuals to adopt healthy behaviors through voluntary participation in programs of health promotion and disease prevention.

Be it enacted by the Senate and House of Representatives of the United States of America in Congress assembled,

SECTION 1. SHORT TITLE.

This Act may be cited as the `Building Efforts for Wellness and Encouraging Longer Lives Act' or the `BE WELL Act'.

SEC. 2. PROGRAMS OF HEALTH PROMOTION AND DISEASE PREVENTION.

(a) Internal Revenue Code of 1986- Section 9802 of the Internal Revenue Code of 1986 is amended--

(1) by redesignating the second subsection (f) as subsection (g); and

(2) by adding at the end the following:

`(h) Programs of Health Promotion and Disease Prevention-

`(1) APPLICABILITY- The following shall apply with respect to a program of health promotion or disease prevention for purposes of subsection (b)(2)(B). Such programs shall be referred to as `wellness programs'.

111TH CONGRESS 1ST SESSION

S. 2806

To codify and enhance existing regulations designed to encourage individuals to adopt healthy behaviors through voluntary participation in programs of health promotion and disease prevention.

IN THE SENATE OF THE UNITED STATES

NOVEMBER 19, 2009

Mr. ENSIGN (for himself and Mr. CARPER) introduced the following bill; which was read twice and referred to the Committee on Health, Education, Labor, and Pensions

A BILL

To codify and enhance existing regulations designed to encourage individuals to adopt healthy behaviors through voluntary participation in programs of health promotion and disease prevention.

Be it enacted by the Senate and House of Representatives of the United States of America in Congress assembled,

SECTION 1. SHORT TITLE.

This Act may be cited as the "Building Efforts for Wellness and Encouraging Longer Lives Act" or the "BE WELL Act".

SEC. 2. PROGRAMS OF HEALTH PROMOTION AND DISEASE PREVENTION.

(a) INTERNAL REVENUE CODE OF 1986—Section 9802 of the Internal Revenue Code of 1986 is amended—

(1) by redesignating the second subsection (f) as subsection (g); and

(2) by adding at the end the following:

111TH CONGRESS 1ST SESSION

S. 2806

To codify and enhance existing regulations designed to encourage individuals to adopt healthy behaviors through voluntary participation in programs of health promotion and disease prevention.

IN THE SENATE OF THE UNITED STATES

NOVEMBER 19, 2009

Mr. ENSIGN (for himself and Mr. CARPER) introduced the following bill; which was read twice and referred to the Committee on Health, Education, Labor, and Pensions

A BILL

To codify and enhance existing regulations designed to encourage individuals to adopt healthy behaviors through voluntary participation in programs of health promotion and disease prevention.

1   Be it enacted by the Senate and House of Representa-
2   tives of the United States of America in Congress assembled,
3   SECTION 1. SHORT TITLE.
4   This Act may be cited as the "Building Efforts for
5   Wellness and Encouraging Longer Lives Act" or the "BE
6   WELL Act".

Figure 6 The 4 Ways to Print from THOMAS a) Print Stylesheet b) Printer Friendly Page c) XML d) GPO provided PDF

Linking Disclousre

When I first clicked on the PDF link, I was brought to a page explaining that I was leaving THOMAS and being sent to the GPO (Figure 7). This is the result of a governement regulation that requires federal websites to inform you when you are leaving their site. This makes sense if you are leaving a government website to go to a privately controlled site. It makes
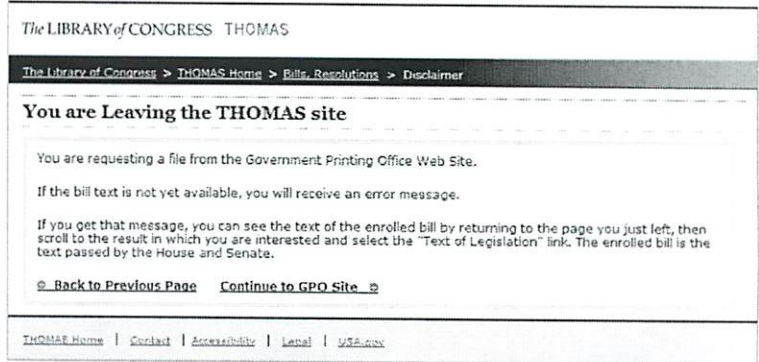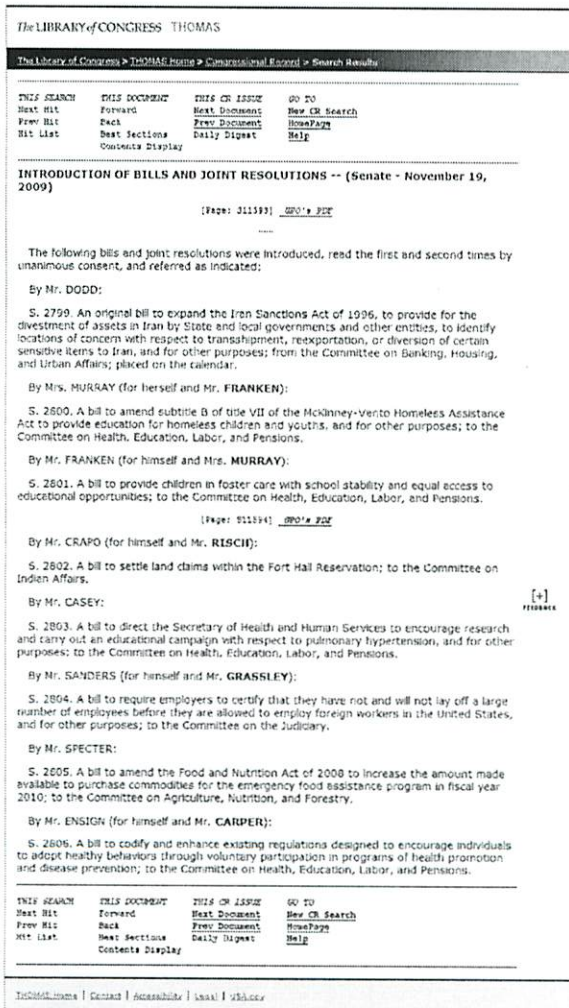

Figure 7 Leaving THOMAS to see a PDF

Figure 8 THOMAS Congressional Record Search Results

less sense when you are moving from one agency to another. It makes even less sense that the text of a bill can be on the Library of Congress' site in one form, but on the GPO's website in another form. Why are two agencies even involved in this? In addition, what is the difference to any end user, either a regular citizen or a well-informed lobbiest? Why does anyone care that the PDF version is on a different agencies' site? Fortunately, the PDF version is digitally signed. This is very good because it makes sure that the document has not been changed since it was signed. However, the signature is verified by a private company, GeoTrust. Shouldn't the government be the one that is

verifying the private company? This discussion is beyond the scope of this proposal and project.
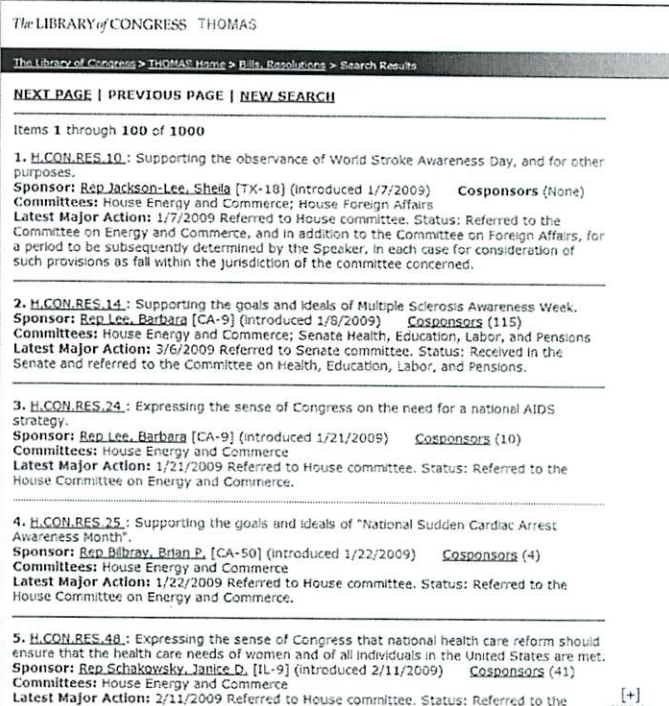
Sparse References

On the bill text page there is a link to the Congressional Record References (Figure 5 Green box). This link is not on the homepage for a bill, only on the text page, which brings up a sparsely formatted page with text extracts from sections of the Congressional Record (Figure 8). There are links to the single page on which that bill was referenced in the printed Congressional Record. The Congressional Record is presented in PDF form, page by page, from the actual printed version of the record. This is not the best way to display this information.

Search

The search engine is hard to use on THOMAS. For example, I searched for "health care" as a keyword. If a citizen stumbled upon THOMAS they might search this term to find the health care bill endlessly discussed throughout 2009. They would be sorely disappointed (Figure 9). The search returns in numerical order all of the bills concerning health. The major health bill is buried in a mountain of results. It is hard to identify at a glance what state the bill is in. Thousands of bills are submitted each term and only a handful get passed. However, they are never marked as abandoned in THOMAS. This may be unavoidable within the current rules of Congress.

Figure 9 THOMAS search for "health care"

Amendments

Amendments are not presented in a clear fashion. In THOMAS, Amendments can only be found by viewing the Congressional Record search results for that bill; the Record only has the list of amendments, not the actual text. Then one will go back to the THOMAS homepage to search for that record using the amendment number. They are then brought to a page where the table of contents of the amendment is displayed – not the entire amendment (Figure 12). Each page of text is linked individually according to how the paper and PDF version of the Congressional Record was put together. This makes no sense. A user wants to read the amendment; the user could care less what page of the Congressional Record the content is on. There does not seem to be an easy way to view the entire text of the amendment on one page on THOMAS. Instead, one must each page of the Congressional record on a separate page! Citizens who know about the GPO's FDsys can use it to navigate to the Congressional Record, navigate to that year, month, and day and input the page number that is mentioned on THOMAS. One can then get a PDF which has all of the amendments submitted separately. This requires too many steps and an advanced knowledge of how each government system operates. There is no reason that the process must be this difficult.

In addition, the actual text of the amendment is very word heavy. It does not provide a visual overview of changes to a document like the diff command does in Linux. This makes it very easy to see what changes have been made at a glance (Figure 13). In addition to benefits for the public, using a computer tool to create these diff files will make the process easier of creating amendments and cause fewer less mistakes. It will also be easier for legislators to scan as the read over the amendment. This may beyond the scope of this proposal however. This system may possibly be able to create diffs by carefully parsing the text of the amendment.

Figure 12 THOMAS Amendment Text

```
27 @synthesize firstTabIndex;                              27 @synthesize firstTabIndex;
                                                        + 28 @synthesize firstReceiptId;
28 @synthesize userLatitude;                               29 @synthesize userLatitude;
29 @synthesize userLongitude;                              30 @synthesize userLongitude;
30                                                         31
31 - (BOOL)application:(UIApplication *)application didFinishLaunchingWi   32 - (BOOL)application:(UIApplication *)application didFinishLaunchir
32 ... application.applicationIconBadgeNumber = 0;         33 .... application.applicationIconBadgeNumber = 0;
33 .... self.firstTabIndex = 0;                            34 .... self.firstTabIndex = 0;
34 ....                                                  + 35 ---->self.firstReceiptId = -1;
                                                        + 36 ---->}
35 .... NSString *msg_type;                                37 .... NSString *msg_type;
36 .... if (msg_type = [[launchOptions objectForKey:@"UIApplicationLaunch   38 .... if (msg_type = [[launchOptions objectForKey:@"UIApplicationLau
                                                        + 39 ---->---->NSLog(@"Get msg_type id: %@", msg_type);
37 ...... if ([msg_type isEqualToString:@"receipt"]) {     40 ...... if ([msg_type isEqualToString:@"receipt"]) {
38 ........ self.firstTabIndex = 0;                         41 .......... self.firstTabIndex = 0;
                                                        + 42 ---->---->}
                                                        + 43 ---->---->NSString *receipt_id;
                                                        + 44 ---->---->if (receipt_id = [[launchOptions objectForKey:@"UIAppl
                                                        + 45 ---->---->---->NSLog(@"Get receipt id: %@", receipt_id);
                                                        + 46 ---->---->---->self.firstReceiptId = [receipt_id intValue];
                                                        + 47 ---->---->}
```

Figure 13 Example of a Diff on a sample of source code

Privately Run Alternatives

There are two popular THOMAS-like sites which have been created by non-profit organizations. Govtrack.us was created by Joshua Tauberer while he was a student. OpenCongress.org was created by the Sunlight Foundation _which_ who advocates for government transparency and runs sites that use government data. Sunlight has done a very good job with OpenCongress.org. As a non-governmental site they can include commenting on the bill, a section to mark up the bill, and a wiki to summarize the bill. A government website would quickly be overrun with flame wars if it accepted comments. This can be seen in the Obama Administration's trial of accepting comments when they first took office. The discussion quickly turned downhill. While it is great that sites have been set up to compete with THOMAS, it is still important to make THOMAS modern. It is still the provider of official government information, which is important to many. It is the first place which citizens often look _for_, because they might know about OpenCongress.org.

Need

Eugene Huang, the director of Director of Government Operations for the FCC's National Broadband Task Force, recently spoke at the Media Lab. He talked about how THOMAS was state of the art when it came out but has fallen behind today. It is not being updated due to lack of attention. Therefore, I am proposing that I lead _(or) to lead_ a team of people to update the site in order to meet the needs of today's audience and allow ordinary citizens to easily interact with government data.

Proposed Project

The US Constitution requires that a journal of be maintained for the House and Senate. The internet provides the opportunity to create a full integrated record of preceding that is easy for Constitutions to understand. A new site should bring information together and make it easy to understand.

Figure 10 Senate.gov Pending Legislation

Senate.gov and House.gov

First, there should not even be a visibly separate site or a system externally called THOMAS. When citizens want to know what is going on in the Senate, they go to Senate.gov. Citizens should not have to know or remember that real time legislative information is on a separate page. Senate.gov should display real-time legislative information front and center on the homepage. Senate.gov's main content area is currently dedicated to the history of voting yes or no in the Senate. This is an improvement over the previous features, which have included an online exhibit of the Russell Senate Office building's centennial. While these are worthwhile, visitors going to Senate.gov are looking for real-time information of what is going on in the Senate. Thankful this information is at least present; however it is hidden on the sidebar. The website displays active legislation by theme (Figure 10). This is very useful as it solves the issue I had with searching for "health care." This should be a part of the new site; but the main content area of both the Senate and House websites' should include real time information from THOMAS. When they click on a bill, they are sent to THOMAS. Users would not notice that they were switched from the Senate's site to THOMAS. No disclosure is needed since almost all users simply attribute the source of

the information from the Federal government.

Citizens do not care about inter-agency scope

divisions and do not need to be informed of moving

from a Senate page to a Library of Congress page.

Congressional Record

It is currently very difficult to find out what

went on the last day in Congress. Commercial news

providers do cover Congress, but they only cover

high profile topics and focus more on the politics

than the contents of the bills. Insiders know about

Figure 11 THOMAS Congressional Daily Text View

the Congressional Daily which has updates on everything that went on in the previous day. However,

the Daily is still bound to a paper format and is managed by the GPO on a separate site with only a small

link from THOMAS (Figure 11). The information is online, but it is clear that the paper vision is still

considered primary. The Record should be made part of THOMAS and the integrated online record

should be thought of as the primary record.

Information from the Record should also be linked from where it makes sense. The text of

debates and comments cannot be found at all on THOMAS. The user must know that this information is

provided by the GPO in the Congressional Record on a completely different site. Debates and comments

should be cross referenced wherever possible. For example, debate on a bill should be linked from the

bill's page. In addition, C-SPAN records video of the debates, but these videos are not linked anywhere

from THOMAS or the Congressional Record. This video should be organized and linked from the text of
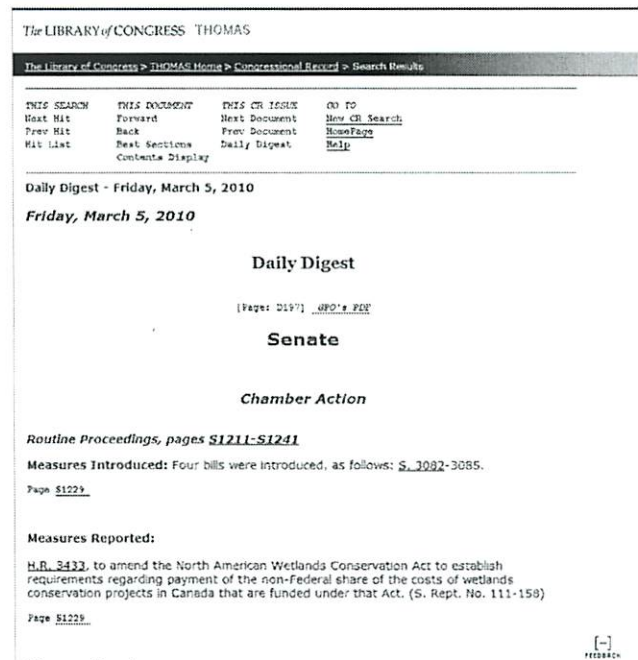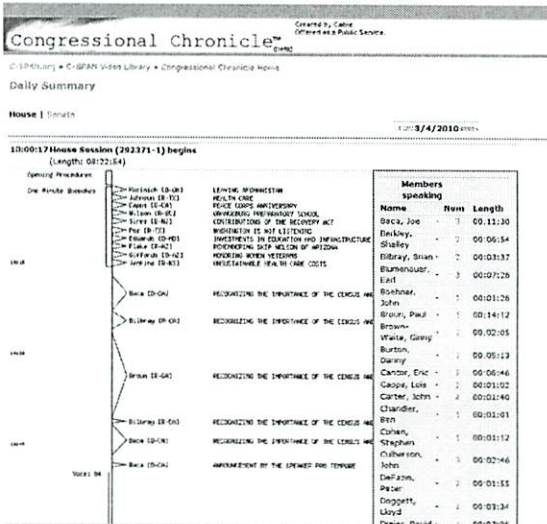
the debate.

News Feeds

More should also be done to provide at a glance access to what went on in Congress in a particular day. Since THOMAS was created, Facebook has popularized the news feed. C-SPAN provides a nice prototype of this in the timeline of each day in Congress on their Congressional Chronicle (Figure 14). News feeds should be front and center on the homepages' of THOMAS, the Senate, and the House. Each bill should have a news feed on its homepage. Each committee and member should also have a news feed of everything they said and every vote they cast which is currently contained in the record.

Figure 14 CSAPN Congressional Chronicle

Dependency Chart

It is almost impossible to track a bill as it goes through the process. For example, H.R. 4872 was the revision to H.R. 3590, or the big health care bill. This is not obvious in the current system. A dependency chart should be created which shows the relationship of these bills.

Materials and Budget

This project will require 5 developers and 1 project manager to work for 9 months in order to create the

new version of THOMAS. Programmer time is estimated at $80,000 a year and project manager time is

estimated at $110,000. An additional $20,000 of benefits per person will be provided. Total salary costs

are estimated at $472,500. Space will be provided by the Library of Congress in the Washington D.C.

area and will be covered under their existing budget. $10,000 of equipment is being requested for

development hardware consisting of 6 desktops ($1,000 each) and a development server ($4,000).

Deployment hardware is estimated at $90,000 and will be decided according to the project

requirements during the project.

→ you still haven't labeled the table correctly. This was covered in class.

Expenses	Approximate Cost
Programmer Wages and Benefits	$472,500
Desktop Computer * 6	$6,000
Development Server	$4,000
Deployment hardware	$90,000
Total	$572,500

Michael Plasmeier

plaz@theplaz.com
(610) 513 0390

plaz

theplaz.com

Massachusetts Institute of Technology

Freshman; expected graduation 2013
Interested in Computer Science (6), Public Planning (11), Management (15), Political Science (17)

Research and Work

SocialMenu 2009- Developer, Viral Communications Group, MIT Media Lab
iPhone web programming, Facebook app development, design for human interaction

Residential Computing Consultant 2009- Computer/network support consultant for MIT's Information
Services and Technology group, Client Support division

Websites

GridView 2007- http://apps.facebook.com/gridview Facebook application which lets users put all of their
friends on their profile and makes a photo mosaic of themselves with their friend's photos.

Primary Developer (PHP, MySQL, Facebook Platform)
Created successful concept from idea to implementation
1.1 million installs; >300,000 monthly active users
Managed developers domestically and overseas
Maintained Linux servers

gridview

Tecker 911 2006-2009 http://911.tecker.net Video podcast covering technology for everyday people
100 episodes of high quality programming available online
Designed show concept and branding
Produce show by arraigning shoots, editing, and distribution
Negotiating show onto cable TV

Tecker 8911

Geoov 2009 http://geoov.com AJAX mapper to display all geotagged flickr photos.
Created website from architecture to implementation
Used Google Maps API with client side JavaScript

Geoov!

SeniorQuoter 2007- http://seniorquoter.org Installable, open-source PHP application which lets high school
yearbooks streamline their senior quote collection process.
Designed application from the ground up, including UI
Implemented in PHP, MySQL
Lead team of 4 people to develop version 2

SeniorQuoter

Usability and logo design
Redesigned the skin for my MediaWiki site to increase usability http://wiki.theplaz.com/Nuvola_Skin
Created logos and layout for my projects

Work

Bespoke Media Group 2009- Implement promotions on Facebook; advise companies on social media
strategy
Work closely with clients to explain their options and implement their vision

Haverford High School

4.7 GPA
CTO of Publications Department 2007-2009 Maintained computers and network for newspaper and
yearbook

http://wiki.theplaz.com/Resume

Michael Plasmeier
Assignment #2
Proposal
Revision

The paper is well structured. The problems with the current system are clearly explained in the Background section. The Proposed Project section outlines in general terms how you propose to redo this online repository of legislative information; however, it should mention the involvement of other programmers and a general timeline for the work. These details shouldn't appear for the first time in the Budget section.

There are still some mechanical errors, including punctuation, misspellings, and typos. The table in the Budget section still isn't properly labeled: recall that every figure and every table in a document must be numbered and have a descriptive title, *even if there's only one.* Your table should be labeled at the top: Table 1. Budget for Proposed Project.

Final Grade: B