

Creating Sustainable Documentation Using Hypermedia

A Master's Thesis

Presented to

Department of Information Design and Technology

State University of New York
Institute of Technology

Utica, NY

In Partial Fulfillment
Of the Requirements for the

Master of Science Degree

By

Jeff Savage

January 5, 2012

TABLE OF CONTENTS

INTRODUCTION	4
STATEMENT OF THE PROBLEM	4
SIGNIFICANCE OF THE PROBLEM	5
SOURCES OF DATA	5
PROJECT DESIGN	6
DESIGN OVERVIEW	6
DELIVERY METHODS.....	6
WIKI SOFTWARE SELECTION	7
NAVIGATION DESIGN.....	8
CONTENT SELECTION	8
DOCUMENT SUBDIVISION.....	8
DOCUMENT DESIGN DETAILS	9
DASHBOARD DESIGN	10
PROJECT IMPLEMENTATION	10
IMPLEMENTATION - ORGANIZATION.....	11
IMPLEMENTATION – CREATING CHUNKS OF CONTENT.....	11
LITERATURE REVIEW	13
INDEXING BY ASSOCIATION	13
HYPERTEXT, HYPERMEDIA, AND PROJECT XANADU	15
THE WORLD WIDE WEB	17
EVERYTHING IN THE WORK ENVIRONMENT IS LIVE HYPER-DOCUMENT STUFF	17
CONVERSATIONAL KNOWLEDGE CREATION	19
CONCLUSION	19
PROJECT EVALUATION.....	20
PROJECT OBJECTIVE - CENTRALIZATION AND CONVERGENCE	21
PROJECT OBJECTIVE - HYPERMEDIA CONVERSION AND CREATION	22
PROJECT OBJECTIVE - SUBDIVIDE CONTENT	23
PROJECT OBJECTIVE - ASSEMBLING SUBDIVIDED DOCUMENTS.....	24
PROJECT OBJECTIVE - COMMON VOCABULARY	27
PROJECT OBJECTIVE - ORGANIZE CONTENT LOGICALLY	28
CONCLUSIONS	28
IMPLICATIONS AND FUTURE RESEARCH	31
REFERENCES	33

TABLE OF FIGURES

FIGURE 1: LIST OF CHUNKS THAT MAKE UP THE 26 COMPLETED DOCUMENTS	24
FIGURE 2: LIST OF ASSEMBLED DOCUMENTS IN THE WIKI.....	26
FIGURE 3: CHOSEN VOCABULARY AND SOURCE ALIASES USED IN VARIOUS DOCUMENTATION	28
FIGURE 4: SUB-DOCUMENT USAGE PATTERNS (IGNORE DOCUMENT TITLES ON X-AXIS)	29

Introduction

Statement of the Problem

When I began working in Information Technology in a public college in Western New York, I quickly realized that our approach to documentation and user technology education was lacking in several areas. Among the problems was a scattered approach to distribution. Between network shares, an employee-only intranet, and a public help desk website with a vintage ‘1995 look’ to it, there was confusion as a content creator about where to place documentation. I am sure that sentiment has been shared by users who were looking for information.

There has also been an ongoing problem of creating and distributing content using the Microsoft Word and PDF formats. In a time when people access information from devices of varying abilities, putting information in proprietary containers which require special plugins or applications to view seems like a misguided approach. This problem will be exacerbated over time, as documents evolve to contain more web-based content.

A more significant problem involves the pace of technological change and the failure of existing solutions to address the related challenges. When information contained in a Word document or PDF changes, it can be very difficult to ensure that all parties are referencing the proper version of the document. Considering the increasing rate of change, anything but a living document can easily be outdated in very little time.

Along those same lines, there has been no mechanism to this point to create sustainable documentation. By sustainable documents, I mean those documents which may be linked from many other documents, but if something changes, that change only needs to be made in one

place. Creating a sustainable set of documentation in container formats such as Word and PDF is a near impossibility.

Significance of the Problem

Given that nearly everyone in the College, from students to faculty and staff, are affected by their ability to use the systems supported by Information Technology, providing effective and timely documentation on those systems is fairly important. A lack of documentation can lead to unnecessary calls to the help desk or other supporting staff. A documentation void can also create something I have perceived as a reverse network effect. In lieu of proper, accessible documentation, pockets of bad information can form and spread through unofficial channels, undermining any efforts to provide knowledge-transfer.

From my perspective as a content creator, I need to know that documentation I create is going to be findable by the people who need or want it. Otherwise, the time spent creating that document is probably better saved for the phone calls and emails I will field as a result of the lack of findable documentation.

Sources of Data

I have completed an archetype of a hypermedia site which I have designed to address many, if not all, of the aforementioned problems. Before beginning the project, I selected and outlined 26 documents which would serve as an example of sustainable documentation in a centralized environment.

Before starting on that project, I completed a literature review on the topic of hypermedia, which is embedded later in this document.

I have also prepared interview questions and have interviewed a few people for their perspective on the new hypermedia site, the differences between the old and new approaches, and what they see as opportunities for improvement.

Project Design

Design Overview

The design of a documentation platform entailed several different design decisions, from choosing hypermedia as the medium to the choice of a particular product on which to build the site, and even how to divide and assemble the information contained within each document.

Delivery Methods

This is an archetypal project in which I am trying to demonstrate the benefits of centralizing documentation in a location that is available from almost anywhere and through almost any device. Hypermedia was the natural choice of media to pursue that path, but building a website from scratch was impractical from a time and quality perspective.

In addition, it is somewhat unsustainable to have to write HTML code any time a document needs creating or updating. While the archetype will be strictly my project with no contribution other than existing documentation from anyone else, I am still aiming to recreate real-world usage patterns where multiple people of varying skill levels are contributing and updating content as needed.

Fortunately, this is not an uncommon problem, which is why there are so many choices in content management software products. The first consideration in this case was cost, so even though many organizations may choose to license a commercial product like Confluence for a project like this, I needed to keep costs close to zero.

The second major consideration is choosing between content management platforms – in this case, the choice seemed to be mostly between blogging platforms and wiki platforms. While blogging software likely provides the easiest interface for content creators, that ease-of-use also comes with a

cost – a reduced ability to re-use content and structure the information, including navigation, dynamically as needed.

Personally, I also have more experience setting up wikis, having used TWiki, MediaWiki, and XWiki for various projects over the years. Since this is an IDT project, rather than an IT project, it seemed ideal to choose a product that I already had some familiarity with or an inordinate amount of time may have been required to even get a platform in place. Even though I do not believe that blog platforms like Wordpress would have presented any problems, focusing my time and energy on the goals of the project, rather than the steps leading up to those goals seemed like a prudent decision.

Wiki Software Selection

As I mentioned, I have set up several wikis before for various class and work projects over the past five years, and I have developed a feeling for the specific advantages and disadvantages of each. XWiki is the most unique and progressive of the three because it is considered an ‘application wiki’, which means that it can encompass and seamlessly integrate other web-based applications. Items like blogs and calendars are add-on apps that can add a great deal of value to the wiki.

Along those same lines, there were some very specific user-centric applications or macros that I thought would bring a great deal to the project beyond the basics typically associated with hypermedia. First is the tag cloud, which will help users identify the most frequently-discussed topics. There is also a built-in activity tracker, which brings the most recent changes in the wiki to the forefront. A floating table-of-contents macro is among the most elegant ways I have come across for providing navigation within a single document. Best of all from a usability perspective is that these aren’t plugins that must be installed on a server, but rather are code snippets that tap into the significant extensibility of this XWiki software platform.

Navigation Design

A primary goal of navigation, beyond the basics, was to provide information opportunistically. I believe the tag cloud and 'activity stream' areas are good examples of allowing a visitor to potentially stumble upon valuable content. For example, maybe I visited this wiki site because I couldn't remember the URI to access email on the web. By clicking on email in the tag cloud and being exposed to a glimpse of every email-related document, I realize that I can also access my email from my phone. These types of opportunistic occasions may also be provided by the activity stream, simply because something unrelated to what I came for what recently updated.

As I discussed in my proposal, high-achievers and lower-achievers fare differently with different navigational tools. By providing several methods of navigation, I hope to appeal to a variety of potential learners.

Content Selection

Content for this project was chosen with little regard for the availability or quality of existing documentation. Rather, content was chosen for the explicit purpose of showing off the virtues of the hypermedia (and by extension, XWiki) platform and how significantly content creators and ultimately consumers benefit from those virtues. I chose content from topics that appeared to overlap in different areas, content that could be subdivided into re-usable chunks, and content that would, at a minimum, include both text and imagery in each document. Regardless of the current state or existence of the document, it will be adapted to utilize the features afforded by the hypermedia platform.

Document Subdivision

Since the primary purpose of content selection was to create a content overlap to the point that I could show the value of reusable chunks, choosing how and when to create those subdivisions was a priority in the document design planning process.

Programmatically, XWiki uses a range of headers (H1 through H5) to demark content. If this discussion had nothing to do with reusing content, these headers would be used to create tables of contents and as anchors within the document. By anchors, I mean that if I want to point to a specific piece of content on another page, the header anchors allow me to point to that specific subsection.

For my purposes, the headers can also be used to embed that chunk in another document. By embedding chunked data whenever possible, in combination with content that has been carefully selected to highlight the advantages of doing so, I believe I will have created the necessary framework to design a more sustainable, living wiki.

Document Design Details

Beyond subdivision of content, there are other standards for each document that I am attempting to reach. First, in order to participate in the tag cloud functionality, I will tag each document with relevant tags to assist consumers in reaching their desired content. By tagging consistently and with a common vocabulary, I believe that the tag cloud will also provide consumers with something of a mental model of the content of the wiki.

Since I believe a consistent approach to content creation and navigation aids the user by posing fewer surprises along the way, I will attempt uniformity in any decision that affects navigation. In this case, that includes a floating table of contents in each document. My subdivisions in each document will use the H2 denotation, or ‘== Content ==’ in XWiki’s native language. Therefore, my uniform table of contents code will also begin with the H2-level header. My preliminary table of contents code follows:

```
{{box cssClass="floatinginfobox" title="**Contents**"}}{{toc start=2 depth=6 numbered=false scope=page /}}{/box}}
```

What this code actually does is define that this box will float to the right side of the page, with a title of ‘Contents’ in bold, and that it will pick up headers beginning at level two (H2) and end at level six (H6). The elements of the table will also not be numbered and will not transcend the page. This bit of

code is also sustainable and reusable because it contains nothing that must be adapted to each individual document.

In piecing together what will be perceived as a single document, the code for that document will look roughly like this:

```
{{box cssClass="floatinginfobox" title="**Contents**"}}{{toc start=2 depth=6 numbered=false scope=page /}} {{/box}}  
== Step 1 ==  
{{include document="SelectedStepOne"/}}  
== Step 2 ==  
{{include document="SelectedStepTwo"/}}  
== Step 3 ==  
{{include document="SelectedStepThree"/}}
```

The end result will look like a coherent, well-designed document, but underneath it all, it will simply be a series of smaller documents pulled together with a table of contents and header titles.

Dashboard Design

Once I have created all of my documents and pages, I will piece them together in a dashboard page, which will include my chosen grouping of content, a tag cloud, and a recently-updated-documents list. The dashboard page, like all other pages, will also include a search function. To facilitate the tag cloud function properly, only completed, assembled documents will receive tags. This step should reduce any chance of a consumer stumbling upon one of my 'chunk documents'.

Project Implementation

The archetype of my project was designed and developed inside a private space of an XWiki instance, which is located here: <https://wiki.flcc.edu/xwiki/bin/view/IDT/WebHome>

In the planning phase, I listed the 26 topics that would serve as the content for my project. Within each of those documents, I determined how each topic would be subdivided for

the purpose of re-using content. I also mapped out which of those ‘chunks’ of information would be shared.

Implementation - Organization

Since I had already planned the 26 topics and how I planned to subdivide those topics, I began by creating links to each of the chunks which would be needed to assemble the 26 documents into a consumable format. Creating the links up-front allowed me to gather any existing materials, including existing documents and pictures.

In XWiki, when a link is created to a not-yet-existing page, the link is appended with a question mark icon. As a result, the process of laying out these links in advance of creating content also proved to be a good progress measurement, since the bulk of the work involved was actually creating these content chunks.

I used the same technique when the time came to assemble the final versions of the documents.

Implementation – Creating Chunks of Content

When it came to creating the small chunks of content, there were a few very important considerations in play, one of which caused me to re-think how some information should be subdivided.

The first consideration for these chunks is document independence. In the context of this hypermedia site, document independence means that these small chunks of documents must be able to stand on their own, even though they will never be displayed or read that way. I had to be very careful to avoid referring to previous steps or any particular ‘big picture’ purpose when writing many of these chunks.

Document independence also meant that I strung some would-be chunks together, simply because I realized that they would never be used without each other and in sequence.

This concept of platform independence also caused me to rename a few of these chunks along the way; when I was assembling pieces of different documents, the name of the sub-document I was linking needed to be more neutral than I originally planned for.

On a related note, my design plans ended up calling for a final document structure that looked something like the following.

- Step Title 1
- <include=document1>
- Step Title 2
- <include=document2>

That sequence, which separates the titles of the step from the document itself was something I originally considered unnecessary, aside from added context for the content creator. As it turns out, the separation of title and content added a great deal of flexibility to the process of building documents. For example, in my WebAdvisor documents, the first step is always to go to WebAdvisor. I include a screen shot and a hyperlink. There is at least one case, though, where I use that chunk differently – as a return to WebAdvisor. Because the title of the chunk is in the final document, and not the independent chunk, I am able to provide a little bit of extra context to the user of the document. Ultimately, I think the execution of that design decision will go a long way toward allowing documents designed in this assembled format to retain some cohesiveness.

Literature Review

Indexing by Association

The term ‘hypertext’ was originally coined by Ted Nelson in his 1965 paper, “A File Structure for the Complex, The Changing and the Indeterminate” (Nelson, 1992, errata). While Nelson’s term took hold many years later, it framed a concept put forward in 1945 by Vannevar Bush. In his wide-ranging article for Atlantic Monthly, titled “As We May Think”, Bush touches on a concept that he calls ‘Memex’. Bush’s Memex arises from his frustration with then-current methods of indexing:

When data of any sort are placed in storage, they are filed alphabetically or numerically, and information is found (when it is) by tracing it down from subclass to subclass. It can be in only one place, unless duplicates are used; one has to have rules as to which path will locate it, and the rules are cumbersome. Having found one item, moreover, one has to emerge from the system and re-enter on a new path (Bush, sec. 6, 1945).

Instead, Bush proposes a system that indexes more like the human mind – by association. “With one item in its grasp, it snaps instantly to the next that is suggested by the association of thoughts, in accordance with some intricate web of trails carried by the cells of the brain” (Bush, sec. 6).

Even in 1945, Bush was able to visualize something resembling a modern desktop personal computer, monitor and input devices:

A memex is a device in which an individual stores all his books, records, and communications, and which is mechanized so that it may be consulted with exceeding speed and flexibility. It is an enlarged intimate supplement to his memory. It consists of a desk, and while it can presumably be operated from a distance, it is primarily the piece of furniture at which he works. On the top are

slanting translucent screens, on which material can be projected for convenient reading. There is a keyboard, and sets of buttons and levers (Bush, sec. 6).

With his Memex example, Bush was primarily concerned with accessing reference material, such as library books. He writes of a user wishing to see a certain book and "...taps its code on the keyboard, and the title page of the book promptly appears before him, projected onto one of his viewing positions" (Bush, sec. 6). With these examples, Bush clearly has one of the earliest recorded grasps on hypertext and associative linking, which he summarizes by writing "the process of tying two items together is the important thing" (Bush, sec. 7).

Although Bush often references books and other printed material, his examples are structured in a way that foreshadows the current World Wide Web:

Thereafter, at any time, when one of these items is in view, the other can be instantly recalled merely by tapping a button below the corresponding code space. Moreover, when numerous items have been thus joined together to form a trail, they can be reviewed in turn, rapidly or slowly, by deflecting a lever like that used for turning the pages of a book. It is exactly as though the physical items had been gathered together to form a new book. It is more than this, for any item can be joined into numerous trails (Bush, sec. 7).

Beyond associative linking, Bush discusses more about his vision of the personalization of Memex. In discussing the process of building a trail, Bush writes "occasionally he inserts a comment of his own, either linking it into the main trail or joining it by a side trail to a particular item" (sec. 7). More personalization, according to Bush, could include an individual commentary on the topic: "He inserts a page of longhand analysis of his own. Thus he builds a trail of his interest through the maze of materials available to him" (sec. 7).

Even something bearing resemblance to Wikipedia gets a mention in Bush's article: "Wholly new forms of encyclopedias will appear, ready-made with a mesh of associative trails running through them, ready to be dropped into the memex and there amplified" (sec. 8).

Hypertext, Hypermedia, and Project Xanadu

Building on the ideas of Vannevar Bush, Ted Nelson wrote the first of nine editions of Literary Machines in 1981. Literary Machines is Nelson's version of the story of Project Xanadu, which was an adventure he described as comparable to that of Camelot and the Manhattan Project. Nelson's premise of Xanadu is simple – "the publishing of tomorrow, which will not be on paper" (Nelson, preface). Nelson offers the following basic visions for Xanadu: All published works will be available from your computer screen, including those not yet invented. You, as the user, will be able to make links between anything inside Xanadu, whether by comments, notes, or otherwise. Royalties to each publisher will be delivered instantly upon delivery (Nelson, preface).

Project Xanadu, in addition to sharing some of its vision with Memex, also foreshadows the commercial nature of the current internet in some ways. Nelson even anticipates resistance from big publishers, while noting that "electronic publishing will mean lower-capital entry (and thus smaller publishers)..." (Nelson, preface).

Returning to the uniqueness of hypermedia, Nelson promises the simplification of data: "The world of computer data has become tangled and complicated, a mess of disconnected files in which we all wallow today—even privately, now, in our offices and homes" (Nelson, preface). The key to this simplification, according to Ted Nelson, are two tools; the link and the transclusion. Nelson defines transclusion as "part of a document may be in several places – in other places besides the original – without actually being copied there" (preface).

Echoing Vannevar Bush, Nelson takes apart the structure and sequence of language. He notes that we should break from sequential presentation because "...it spoils the unity and structure of interconnection...[and] it forces a single sequence for all readers which may be appropriate for none" (Nelson, p. 1/14). Continuing on the structure of thought, Nelson notes that though is not itself sequential; he calls the shape of thought a structangle. "None of the ideas necessarily comes first; and breaking up these ideas into a presentational sequence is an arbitrary and complex process.

Ted Nelson (1992) defines hypertext as non-sequential writing (p. 1/17). Examples from 1981, according to Nelson, include newspapers and magazines. The advantage of hypertext become apparent when compared to the challenges of traditional, sequential text. "There are basically two difficulties in writing sequential text: deciding on sequence...and deciding what's in and out" (Nelson, p. 1/18).

The relative freedom of hypertext opens the door to other, less sequential styles of reading and writing. Among those styles is what Nelson calls "chunk style hypertext", where "the user, or reader, moves through it by reading one chunk, then choosing the next" (Nelson, p. 1/15). Since Nelson is talking about prototype software, there are no concrete real-world examples of chunked hypertext outside of his Project Xanadu at the time of writing. However, by 2011, chunked hypertext emerged in several distinct forms. Among those forms are blog entries as part of a larger blog and wiki pages as part of the overall space in a wiki. Beyond that, we often see chunked hypertext and hypermedia through RSS, Twitter, and Facebook feeds on various pages. At the same time, each of those technologies also embody the principles of transclusion. When I visit a page that has Facebook content, that content has rarely been copied and pasted in place; instead, it is transcluded from its original location. As a side effect,

transclusion keeps content updated from one central location, no matter how many places it actually appears in.

Project Xanadu is more than a theoretical concept built up around hypermedia. Xanadu source code was eventually released at <http://udanax.com> in 1999 under the X11 license.

The World Wide Web

In 1996, Tim Berners-Lee responded to an invitation to publish in IEEE Computer on the state of the World Wide Web (W3). While acknowledging Ted Nelson and Vannevar Bush as visionaries, Berners-Lee points out the unfortunate truth of 1980:

For all these visions, the real world in which the technologically rich field of High Energy Physics found itself in 1980 was one of incompatible network, disk formats, data formats, and character encoding schemes, which made any attempt to transfer information between dislike systems a daunting and generally impractical task. This was particularly frustrating given that to a greater and greater extent computers were being used directly for most information handling, and so almost anything one might want to know was almost certainly recorded magnetically somewhere” (Berners-Lee, 1996, p. 2).

Berners-Lee points to the flexibility of Universal Resource Identifiers as a key to the growth of the web. “Typically, though, hypertext systems were built around a database of links...guarantee[ing] that links would be consistent, and links to documents would be removed when documents were removed. The removal of this feature was the principle compromise made in the W3 architecture...” (Berners-Lee, p. 3). This compromise allowed a level of scalability that the web would later feed on (p. 3).

Everything in the Work Environment is Live Hyper-document Stuff

Don Engelbart is also a hypertext pioneer; he was part of the team at Stanford Research Institute that built the oN-Line System in the early 1960s (Conklin, 1987, sec. 2.1.2). Engelbart’s vision was roughly: “...placing in computer store all of our specifications, plans, designs,

programs, documentation, reports, memos, bibliography and reference notes, etc., and doing all of our scratch work, planning, designing, debugging, etc., and a good deal of our intercommunication, via the consoles” (Conklin, sec. 2.1.2).

One of the keys to this NLS system was the unique identifier associated with everything inside. “For example, a document might have statements “1”, “1a”, “1a1”, “1a2”, “1b”, etc., though these identifiers did not need to be displayed. Any number of reference links could be established between statements within and between files” (Conklin, sec. 2.1.2). At its peak, NLS and its commercial equivalent, AUGMENT, saw as many as 20 networked mainframe servers across the US (Engelbart, 1995, p. 30).

During the emergence of the web in the mid-1990s, Don Engelbart wrote an article for ACM detailing what he believed should be the next steps in hypermedia. Engelbart’s article focuses on hypertext concepts, including his Concurrent Development, Integration and Application of Knowledge (CoDIAK) paradigm and the Open Hyperdocument System (OHS) concept. These systems followed the oN-Line System (NLS), which was developed and used in the 1960s.

Advancing Ted Nelson’s ambitions for Xanadu, Engelbart sees applications, such as graphical editors and spreadsheets, integrating directly with documents inside his universal knowledge base (Engelbart, 1995, p. 32). He gives the following example of how work revolves around a central knowledge base: “...one could create a Gantt chart within a project management system, and manipulate it as a graph in a charting application or as mail in an email application” (p. 32).

Conversational Knowledge Creation

According to Christian Wagner (2004), “A Wiki is a set of linked web pages, created through the incremental development by a group of collaborating users and the software used to manage the set of web pages (p. 269). Ward Cunningham developed the first Wiki as a tool to communicate specifications for software design (p. 269).

Wikis are a unique implementation of hypertext because they both connect content and serve as markers for new content. In many Wikis, writing in CamelCase creates a link to a not-yet-created page, ready for editing (Wagner, p. 273).

Since Wikis are designed as a multi-user technology, they typically contain several features to avoid conflicts and inconsistencies (Wagner, p. 274). Among those features is an ability to identify orphaned pages, or those pages which aren't linked to (p. 274). Another feature realized because of the nature of a multi-user environment is versioning. Versioning means that the Wiki retains all versions of its pages, allowing for comparison and rollback when necessary (p. 274).

Wagner also describes how hyperlinks create context for users:

Hyperlinks connect concepts to other concepts, thereby creating context. Aside from the obvious advantage of allowing readers to make connections and to drill down into detail knowledge, hyperlinks are also a potential quality assurance mechanism and relevance indicator. Pages with many links to them indicate a highly useful page. Furthermore, the context identified by a page's hyperlinks (and hyperlinks pointing to it) help define the meaning of a page to a search engine. Modern search engines such as Google are able to interpret link information accordingly (Wagner, p. 277).

Conclusion

For me, the public Internet and World Wide Web was my first exposure to hypertext technology. Despite my own experience, hypertext and hypermedia have been

around in concept form since the mid-1940s. Twenty years later, the terms were coined by Ted Nelson as he and his team began work on an impressive hypertext-based concept known as Project Xanadu.

Around that same time, Don Engelbart and his team at the Stanford Research Institute were working on a project that not only provided the first demonstrable glimpse at hypertext, but also involved the first known implementation of the mouse and windowing display technique (Metz, 2005, p. 131).

Likewise, although public awareness of Wiki technology likely coincided with the launch of Wikipedia in 2001, Ward Cunningham actually developed the first known Wiki in 1995 (Wagner, 2004, p. 283).

Project Evaluation

The purpose of this section is to evaluate the execution of my project relative to objectives set before the project began. Included in the evaluation is how the outcome of the project relates to my research question – *“How can creators of technical documentation leverage a hypermedia environment to organize and display content?”*

As spelled out in my project proposal, this project also had several objectives aside from the main research question. Those objectives are as follows.

- Centralize and converge documentation into the project site
- Convert all proposed documentation into native hypermedia format
- Subdivide, or “chunk”, the content into re-usable pieces

- Demonstrate “chunk-linking”, or the assembly process of building coherent documents
- Develop a common vocabulary
- Organize content in a logical manner

Project Objective - Centralization and Convergence

As I was designing objectives for this project, I tried to focus on objectives that would enhance the experience for either the content creator or the content consumer. This objective happens to suit each of those potential constituent groups. My proposal notes a general sense of disarray when, as a creator of content, I have to decide where to put it. Previous landing spots have included an employee-only intranet site, network shares, and a very outdated public website for the help desk. The mere presence of so many potential locations has discouraged me from creating documentation in the past.

For content consumers, if you have so many possible locations that may contain the information you want, it can be simpler to ask the help desk or someone on staff about a topic for which information may already exist. Ultimately, that outcome defeats the purpose of creating the documentation in the first place.

The site I created for this project is clearly capable of handling any and all of the content already created. With its ability to embed Google documents and YouTube videos, among other externally-hosted content, it should be a very flexible site moving forward. That flexibility would also facilitate the second piece of this objective – convergence. Once all relevant content has been moved into the hypermedia site, the Word documents and PDF files could be removed from other locations. At that point, the current repositories could be removed so as to confuse consumers and all relevant links could be replaced.

Project Objective - Hypermedia Conversion and Creation

As part of this project, I converted existing documents that were created and distributed in Microsoft Word and the PDF format into the site's native hypermedia format. I also created many documents from scratch specifically for this project.

I should point out that I learned native XWiki syntax for this project, which is the most hands-on and likely the most difficult way to get content into the site. The XWiki software offers several easier options for importing and creating content. It is possible, with a bit more server configuration, to allow the software to directly import Microsoft Word documents. I passed on this option because I don't believe that importing entire Word documents intact was in the spirit of this project. There are significant formatting changes that happen when using hypermedia versus a word processor and I wanted these documents to look and feel "native". In addition to those shortcomings, another objective of this objective is subdividing and assembling documents, which would be cumbersome using the document import method. Another conversion tool is the XWiki plugin for Microsoft Word that lets you save your document in XWiki code.

Instead, I saved all of the images from existing documents and pasted in the text before going back to format each piece. That process is not nearly as cumbersome as it sounds. Creating documents from scratch was very similar – I uploaded pictures, edited them if necessary, and created the textual description to go alongside the pictures.

XWiki also features a robust JavaScript-based text editor, which will feel very familiar to regular users of popular word processing software. Because the editor has a more limited feature-set than Microsoft Word, for example, there are fewer opportunities to stray from a typical design standard, even if multiple people are designing documents.

Overall, content creation was very easy and straight-forward. Content creators who are used to Microsoft Word should find the common functions quickly and by now, many people have used JavaScript-based editors in other web-based software like email and even learning management software.

Project Objective - Subdivide Content

The purpose of subdividing content from documents down to steps of a document comes from the idea of creating content once and linking to it many times. The benefit of this approach is that if something changes in, say, a piece of software, you may have many documents referencing a particular piece of information, but the content itself only needs to be updated in one place.

Subdividing content is a very complex process and is very content-sensitive. Many of the subdivisions that I planned for leading up to my proposal were done blindly because there was no document to start from. Once I got into these documents, I realized that I needed to make sure I had a logical reason to subdivide. The best example comes from when I was converting wireless networking documents. In each of these documents, I reached a point where I realized there was no reason to subdivide anymore because there is simply nowhere else to go with the document once you reach certain steps. In other words, once you get into a place where you are looking at wireless networks, there are very few or no other potential forks that the content can branch into, other than the document you are subdividing. Subdivision for subdivision's sake was not a productive outcome of this project, so I adapted several of the subdivision plans to account for some of the more rigid content.

Content Chunks

AccessWA	TimeEntryLinkWA	NewAppointmentOWA	EnterServerIOS	AddWiFiAndroid
LoginWA	SelectPositionWA	AppointmentDetailsOWA	SetSyncIOS	NetworkDetailsAndroid
WhatsUserID	EnterHoursWA	ContactsOWA	AccessPersonalGmail	WiFiIOS

EnterDetailsWA	TimeApprovalWA	CreateListOWA	SelectSettingsGmail	WiFiCredentialsIOS
RetrieveUserID	TimeApprovalSuperWA	MeetingRequestOWA	AccountsAndImportGmail	AirPortOSX
GuestAccessWA	SelectPositionSuperWA	SchedulingAssistantOWA	EnterAddressGmail	CleanUpProfilesOSX
WhatsPasswordWA	ReviewHoursWA	AppListAndroid	POPSettingsGmail	WiFiSettingsOSX
InitialPasswordWA	ApproveTimeSuperWA	SettingsAndroid	AccessEmailAndroid	ConnectWiFiOSX
ClickOKWA	LeavePlanSummaryWA	AccountsAndroid	AddManualAccount	CycleAirPortOSX
ResetPasswordWA	PayAdvicesWA	AddAccountAndroid	ManualSetupAndroid	NetworkAndSharingW7
SelectEmailWA	EnterRequisitionWA	CorporateAndroid	IMAPSettingsAndroid	ManageWirelessW7
TempPasswordWA	SearchForClassesWA	SyncOptionsAndroid	OtherAccountIOS	AddOrManageW7
ChoosePasswordWA	AccessOWA	NameAccountAndroid	OtherCredentialsIOS	EnterCredentialsW7
UpdateEmailWA	NewMessageOWA	AccessSettingsIOS	AXPAddressGmail	ManageWirelessXP
SignUpWA	AddressBookOWA	MailContactsCalsIOS	AXPSettingsGmail	NetworkSettingsXP
RequestFormWA	ComposeOWA	AddAccountIOS	WirelessNetworksAndroid	EnterCredentialsXP
EmployeesLinkWA	CalendarOWA	EnterCredentialsIOS	WiFiSettingsAndroid	NetworkAndSharingVista

Figure 1: List of chunks that make up the 26 completed documents

I don't recall having any concrete plans for when to subdivide coming into this project, beyond simply subdividing at each logical step. That may have been a mistake, but on a smaller-scale project, it was mostly unnoticeable until I got to the later wireless documents. Overall, the plan and execution of subdividing documents was successful, but additional planning could have saved me some revising time.

Project Objective - Assembling Subdivided Documents

The process of assembling documents from chunks of information is easy. Take my EnteringRequisitions document as an average-sized example, since it features seven steps and the assembly is made up of four chunks of information. The code which creates that page follows.

```

{{box cssClass="floatinginfolbox" title="***Contents***"}}{{toc start=2 depth=6 numbered=false scope=page /}}

```

```

== Access FLCC WebAdvisor ==

```

```

{{include document="AccessWA"/}}

```

```

== Log Into WebAdvisor ==

```

```

{{include document="LoginWA"/}}

```


== Access the Employees Menu ==

```
{{include document="EmployeesLinkWA"/}}
```

== Access the Requisitions Menu ==

```
{{include document="EnterRequisitionWA"/}}
```

The key to this entire process is staying organized. I created 85 pieces of re-usable chunks of information, toward an end result of 26 finished, presentable documents. Fortunately for me, I had some variety in content to make it easier to identify the chunks that were needed for any particular document. The other thing that worked in my favor was that I worked sequentially through the topics. In other words, I created the chunks needed for a particular documents and then assembled the document. If all 26 documents had been WebAdvisor documents, I probably would have needed a different approach to organization. I am including a table of final documents below.

Assembled Documents

WebAdvisor

Learn your User ID

Set your password

Reset your password

Update your email address

Enter your time

Approve time entry

Check your leave time

Retrieve your pay advices

Enter a purchase requisition

Find a class

Student Email

Access from your Android device

Access from your iPad, iPhone, or iPod

Access from your personal Gmail account

Employee Email

Send a message

Add an event to your calendar

Create a contact list

Coordinate a meeting

Access from your Android device

Access from your iPad, iPhone, or iPod

Access from your personal Gmail account

Wireless

Configure Android devices

Configure iPads, iPhones, or iPods

Configure Macbooks

Configure Windows 7 laptops

Configure Windows Vista laptops

Configure Windows XP laptops

Figure 2: List of Assembled documents in the wiki

This concern would hold true as the project scales, not only to include more topics and documents, but content creators. I believe that organization is the best way to prevent the

benefits of having re-usable content from being canceled out by confusion over which pieces are necessary for assembly.

If I were doing this on a larger scale, both in people and content involved, I would propose an entirely separate 'space' on the site for the subdivided content. The separate space should be organized topically and chunk names should be uniform and as descriptive as possible. Having a dedicated space would serve at least two purposes – enhancing organizational opportunities and preventing chunks from appearing by themselves in anything I would want users interacting with, including activity streams.

Project Objective - Common Vocabulary

The establishment of a common vocabulary was chosen as an object to ensure that content is both understandable by users in words they can understand and consistent throughout the environment. As a side benefit, it helped to reduce any tendency to bring documents into the hypermedia environment in the voice of whoever first designed the document.

As the only participant in this project, wanting a common vocabulary and voice was relatively easy to accomplish, mainly because phrasing that isn't mine stands out to me. However, there were several words in existing documents that I was uncomfortable with, such as 'uncheck' and 'hit'. Rather than simply transcribing the documents as-is, I added the word to my vocabulary list and chose a more comfortable alias. I believe my final table is much smaller than it would be in a group environment where many people have contributed several documents each. My table is below.

Common Vocabulary Index

Usage	Aliases
Email	Email, E-mail, electronic mail, message
WebAdvisor	WebAdvisor, Student information system, SIS, ERP
iPad, iPod, iPhone	iOS
Deselect	Uncheck, not chosen
Prefix	Insert, begin with

Figure 3: Chosen vocabulary and source aliases used in various documentation

Project Objective - Organize Content Logically

There were at least two solid choices that could have served as a logical organizational structure – topically and by role. Organizing by role at the college – students, faculty, and staff – made some sense originally. In theory, you could limit exposure for each group to the content that applies to them. On the other hand, there are almost no hard lines in this business; faculty and staff tend to use many of the same systems for many of the same reasons and even students occasionally get access to more systems through work-study programs or as student aides.

Conclusions

How can creators of technical documentation leverage a hypermedia environment to organize and display content?

The most critical aspect of this project was producing and demonstrating sustainable content. Decisions about content to include and how to subdivide it were all primarily based on whether or not portions of each document could be used in more than one document. Naturally, the easiest way to do this would be to focus on documents centered on a single system, such as WebAdvisor. I did opt for more topical variety for the sake of a realistic demonstration.

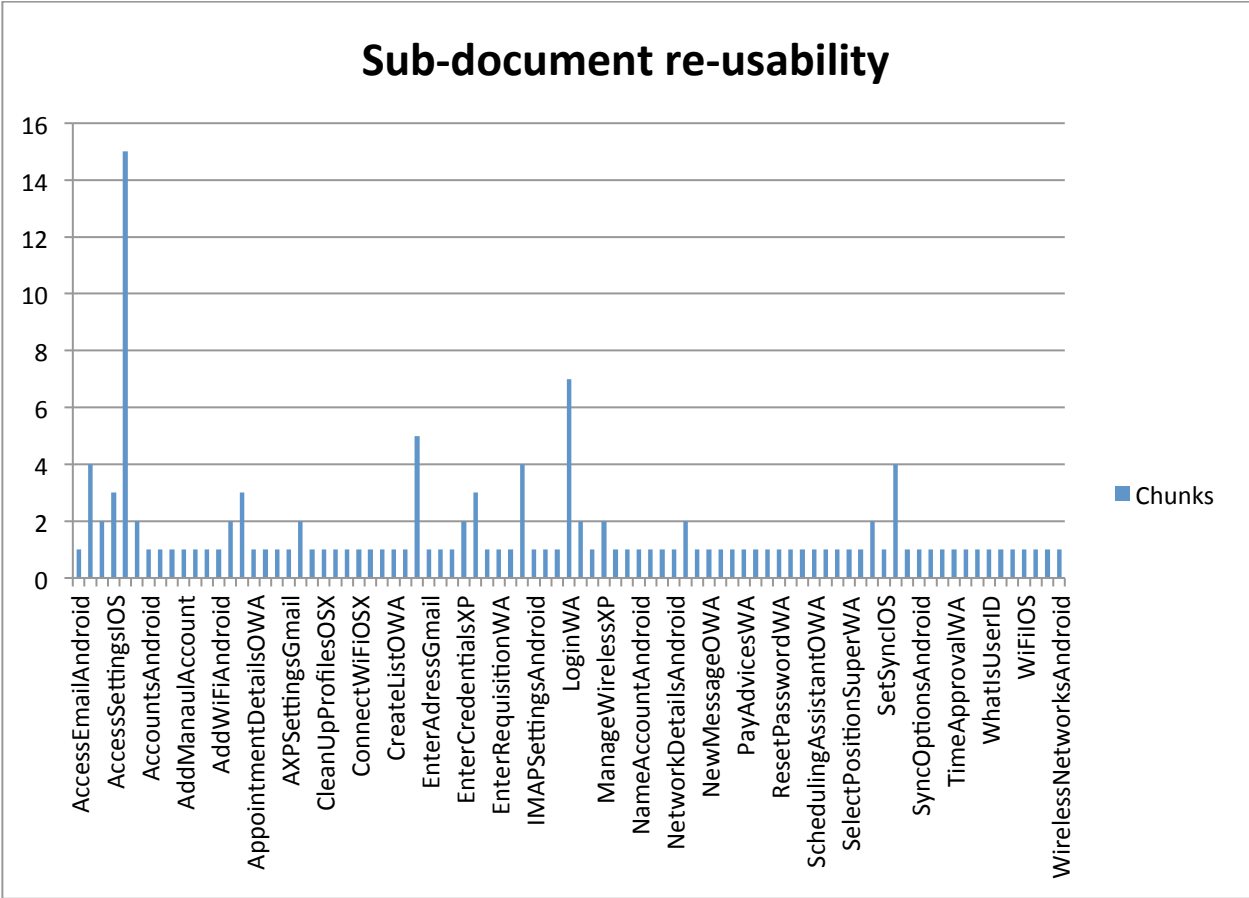


Figure 4: Sub-document usage patterns (ignore document titles on x-axis)

Given the 85 sub-documents comprising 26 finished documents, 18 of those sub-documents were used more than once. By that measure, roughly 21% of the sub-documents found more than one use. I laid out a usage table and averaged the numbers, which gave me the average rate of re-use of all of the sub-documents. That number came in at over 56%. I consider that a fairly impressive number given the topical variety and small sample size.

In my literature review, it was noted that the hypermedia environment itself can introduce a sort of cognitive overhead on the learner. Tools such as simple lists are one suggestion offered to minimize that cognitive impact. However, it was also noted that people of different ages and experience levels tend to pursue information in different ways. In book terms, that may mean tables of contents, indexes, and simple browsing. In hypermedia terms, I

have created a similar variety of options. Those include the simple navigational list, the search function, and the tag cloud. Even though the activity stream function was broken due to the process of isolating my project workspace, it represented another navigational option for high-achievers.

Above, I noted several concerns and adjustments I would make in response to how this project went for me. Organization and planning, in particular, are major issues in a highly-structured environment such as this. While acknowledging my own concerns as valid, I also believe that this project would scale very well from a simple archetype into a multi-user, multi-discipline, multi-topic environment. Despite my own organizational challenges, the software I used is very flexible and well-suited to the type of organization I have proposed.

Along those lines, I think it is vital to look at opportunities that benefit both content creators and content consumers. By isolating sub-documents from documents and creating intuitive, coherent organizational schemes in each environment, both creators and consumers can benefit from rich, versatile organizational schemes that are designed explicitly for them. Without the sub-document and linking approach, that flexibility would be lost; the content would be in the same structure no matter where you looked at it.

There is no doubt that with a re-use rate of over 56%, my time spent creating content was significantly reduced and that the rate would continue to grow as more content is introduced. I almost felt like I was creating a network effect by myself and that effect would only be enhanced by having more people and more content involved.

In addition to self-evaluation, I also asked a couple of colleagues to give me their impressions on the end result of the product as well as the methodology involved in building the site. While discussing what makes for good documentation, it was agreed that the intentional step-by-step organization and

sticking to layman's terms were important to the quality of the site and to good documentation in general. Distinct, meaningful organization of the site displayed to the user was also noted among the initial impressions (personal communication, October 24, 2011).

On the topic of participating in an environment like the one I created, but in a more collaborative sense, advantages such as writing and updating content once were noted. Conceptually, I gather that the site's structure and organization can be difficult to grasp. While the discussion focused on appealing to participants about the advantages involved, I believe that some way of creating a mental model for users on how the site fits together may work just as well as a sales pitch (personal communication, October 24, 2011).

Implications and Future Research

Before beginning this project, I researched the subject of hypermedia through its history, which yielded a look at several innovative thought processes and projects from decades ago. One of the unique qualities of hypermedia in all its forms over the years is its near-infinite linkability. Despite this intrinsic quality of linkability, some of its uniqueness remains untapped, as noted by Ted Nelson in Literary Machines (1993, preface). Despite the potential, Nelson discusses the increasingly complicated and tangled nature of computer-related data and promises simplification through two simple tools – links and transclusion. Transclusion is the process by which part or all of a document can be included in another by reference, rather than by copying.

My primary objective for this project became leveraging opportunities for transclusion, and by doing so, creating more sustainable documentation. In the world of information technology, you often have two choices when it comes to documentation – do it up-front, or ignore it and prepare to spend a considerable amount of time answering questions later.

On the other hand, creating documentation is not often a fun or easy process. Prior to using the techniques leveraged in this project, any version upgrades or new systems often required complete rewrites of the accompanying documentation. I experienced this first-hand recently on a piece of software we run separate but similar instances of for employees and students, and the prospect of rewriting a basic document and its related documents was my responsibility. For me, this scenario was the crystallization of why Word and PDF documents are ill-suited for this type of documentation. The two documents in question were roughly 95% similar, yet I had to create two distinct documents and store them in two different places. This is exactly this disconnected, tangled mess that Nelson was referring to and transclusion and linking are a well-suited solution to the problem.

I believe that the techniques used in this project will scale well and will create opportunity for further research. As one example – the dynamic will change considerably with the addition of more people to a project like this. Although I used wiki software as the platform for this project, I didn't use it as a wiki in a collaborative sense – I used it simply for its advanced hypermedia features. The network effect of collaborating in an environment like this – and transclusion definitely lends itself to collaboration – was beyond the scope of this project, but was something that I thought of frequently while advancing through the project.

Organization is another aspect of this project that merits further study. Even with a relatively small sample of documents and me as the only participant, organization was a challenge and one that would have to be reconsidered before adding more documents or people to the project. Earlier in this document, I proposed opportunities for changes in the organizational structure of the wiki.

Moving forward in a project such as this while expanding the scale of the project will require adherence to established standards, whatever they are. As someone who has worked in several collaborative systems over the years, I know how difficult it is to get everyone on a project to commit to

simple standards like naming conventions. In my experience, that lack of adherence to standards can be a nuisance, but not terribly disruptive. However, in this environment, even straying from naming conventions could be a destructive action. Educating contributors should go a long way toward creating sustainable standards and should be a point of emphasis in any advancement beyond this point.

This project has focused exclusively on the experience of the content-creator. There is further opportunity to focus on the benefits of the sustainable approach to end-users of the documentation that comes as a result of the project. Does the fact that changes, if necessary, only need to be done in one place lead to more timely updates? Does the flexibility that comes from working in hypermedia lead to better organization for end-users? Is the documentation ultimately more findable and usable? I would expect to see several correlative benefits for end-users if this project were to continue in a meaningful way.

References

- Berners-Lee, T. (1996). WWW: Past, present, and future. *Computer*, 29(10), 69. Retrieved from <http://www.w3.org/People/Berners-Lee/1996/ppf.html>.
- Bush, V. (1945). AS WE MAY THINK. *Atlantic Monthly* (10727825), 298(2), 55-58. Retrieved from <http://www.theatlantic.com/magazine/archive/1969/12/as-we-may-think/3881/1/>.
- Conklin, J., & Microelectronics and Computer Technology Corporation. (1987). A survey of hypertext. Austin, Tex: Microelectronics and Computer Technology Corp. Retrieved from: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.41.9368&rep=rep1&type=pdf>
- Engelbart, D. C. (1995). Toward Augmenting the Human Intellect and Boosting our Collective IQ. *Communications of the ACM*, 38(8), 30-33. Retrieved from EBSCOhost.
- Engelbart, D. C. (1992). Toward High-Performance Organizations: A Strategic Role for Groupware. *Proceedings of the GroupWare '92 Conference*, San Jose, CA, August 3-5, 1992, Morgan Kaufmann Publishers. Retrieved from: <http://www.dougenelbart.org/pubs/augment-132811.html>

Metz, C. (2005). Whose Idea Was It?. PC Magazine, 24(19/20), 130-131. Retrieved from EBSCOhost.

Nelson, Theodor H. (1992). Literary Machines 93.1. Sausalito CA: Mindful Press.

Wagner, Christian (2004) "Wiki: A Technology for Conversational Knowledge Management and Group Collaboration," *Communications of the Association for Information Systems*: Vol. 13, Article 19. Retrieved from: http://wikitrust.peacocktech.com/images/9/96/Wiki-_A_Technology_for_Conversational_Knowledge_Management_and_Group_Collaboration.pdf

XWiki Documentation (n.d.). Retrieved January 5, 2012 from the XWiki.org Documentation Wiki: <http://www.xwiki.org/xwiki/bin/view/Main/Documentation>

XWiki.org (2011). XWiki Enterprise (Version 3.0) [Software]. Available from <http://enterprise.xwiki.org/xwiki/bin/view/Main/Download>