

CCEL 2020

2013-08-08

1. Introduction

The CCEL's last strategic plan was entitled *CCEL 2006-2011 and Beyond*. We “used up” that strategic plan entirely, implementing or at least exploring all of the suggested strategies. It is time for a new plan, a vision and roadmap for the future.

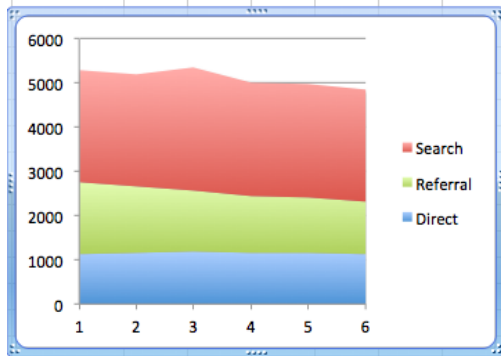
CCEL's efforts for the last five to six years have been largely on Hymnary.org, though we have done some updates to CCEL.org as well, including a new graphical theme and a couple of new software systems. We've also added a few books. However, the website hasn't changed significantly and feels a bit dated.

In the last five years, traffic measured as page views has decreased 35% while the number of visits has declined a small amount. One crucial factor in decreased page views is the decrease in the number of pages per visit from 5.88 to 4.37.

Of course, there are many factors in the decline in the number of page views, including increased competition. However, the number of visits referred by Google and other search engines has actually increased slightly since 2007. Therefore changes in usage patterns of our users must be a key consideration. One theory is that fewer users are doing long-form reading in a Web browser, preferring mobile devices and eBook readers. Then too, new users tend to view fewer pages than repeat users. We could be losing some of our regular, heavy users and seeing more transients.

Mobile computing—including smartphones and tablets such as the iPad—are changing the way the Internet is used. Android and iOS phone users prefer apps to Web browsing, but they do a fair amount of Web browsing on mobile devices as well. Mobile accesses of the CCEL grew from 7.5% in 2011 to 15% in 2012. Clearly we must plan for a large proportion of our users to access the site on mobile devices—perhaps as much as 50% of Web access will be from mobile devices in a few years. We have a mobile website for CCEL, m.ccel.org, which is an improvement over the full website, but it still has some usability issues. In addition, best practices today are for “responsive” websites, which work for a variety of screen sizes, rather than separate mobile sites.

	2007	2008	2009	2010	2011	2012
Search	2550	2517	2807	2594	2573	2548
Referral	1613	1510	1379	1275	1256	1172
Direct	1129	1162	1175	1151	1147	1137
Pages/visit	5.88	5.39	5.25	4.94	4.74	4.37
Pageviews	31122	27963	28147	24792	23586	21217



Another change affecting us is the decline of physical media. CD sales, which were once a key part of our financial model, have declined to close to zero, although users still purchase a few CD images as downloads. In addition, we have deemphasized requests for donations. Users who wanted to download PDF files were previously asked if they would care to make a donation to support the download. They were given the option of a free download. We removed the requests for donations, which were creating some bad feelings among users, and PDF downloads have doubled, but donations and subscriptions have declined, from nearly \$50k per year in 2007 to less than \$10k. (This decline is not all in response to the change in PDF policy—that policy was only instituted a year or two ago—and the trend toward fewer subscriptions has other contributing factors.)

At the same time, sales of eBooks formatted for the Kindle and iBook stores have grown dramatically, more than making up for the decline in CD sales, subscriptions, and donations. In short, the financial model has transitioned from CDs, subscriptions, and advertising to formats and advertising. With this new financial model, our self-generated revenue reached an all-time high last year, \$22,000 more than the previous year. This change to more mobile access and sales of Kindle and iBook editions could be a primary factor in the decrease in the number of CCEL page views as well.

Our groups system could also use rethinking and reworking. Our search engine is lagging behind Google in semantic capabilities. The book user interface could be improved with new capabilities. In short, the time has come for a new strategic planning process for the next five to seven years and renewed emphasis on improving the CCEL. The seven year timeframe would bring us to 2020.

This document will cover the Classic Christian Book-related activities of CCEL. The CCEL, as a project, also runs Hymnary.org, and it is possible that it will have additional websites or projects in the future. These will be treated separately.

2. Process

The strategic planning process we plan to adopt involves reviewing the *mission* of the CCEL and identifying *objectives* for helping us meet that mission for the next five to seven years. When the objectives are set, we will consider various *strategies* for meeting those objectives. Finally, we will prioritize and plan the implementation of those strategies in a *roadmap*.

For each objective, we will perform SWOT analysis, considering with respect to the objective our internal strengths and weaknesses and the external opportunities and threats. We will consider the resources required to implement the strategy and the possible return, financially and in terms of meeting our mission. Then we will give a rough prioritization.

3. Mission

In 2006, the CCEL Board of Directors approved the following brief mission statement:

The Christian Classics Ethereal Library (CCEL) seeks to build up the church by making classic Christian literature widely available and promoting its use for edification and study by interested Christians, seekers and scholars. The CCEL accomplishes this by selecting, collecting, distributing, and promoting valuable literature through the World Wide Web and other media.

The values guiding the collection and activities of the CCEL include:

Selectivity—Writings included in the library are considered valuable in some way to build up the church. In addition to Christian classics, there may be secondary literature on the classics and other works valuable for the mission.

Orthodoxy—Writings in the library for the most part represent orthodox Christianity as understood by mainstream Protestant, Catholic and Orthodox Christians, with greater emphases on Reformed and Protestant writings.

Ecumenicity—The CCEL strives to expose readers to a variety of Christian perspectives.

Community—The CCEL builds community through promotion to and engagement with and among Christians, seekers and scholars worldwide.

Efficacy—The CCEL seeks to provide effective resources and tools for edification, worship, leadership, ministry training, Christian scholarship, and education. The CCEL also seeks to provide educational and scholarly opportunities for interested faculty members and students of Calvin College.

Efficiency—All choices including literature, media, technologies, languages, partners, etc., will be made strategically with the goal of good stewardship of resources and competencies to build up the church worldwide.

Sustainability—The CCEL seeks to create a sustainable model for building up the church through its activities.

The director and the advisory board do not see a need to revise the mission statement.

4. Objectives

At the January, 2013 advisory board meeting, the board considered several objectives proposed by the director. After some discussion and with some modifications, the objectives are to

1. Serve a broad audience of people who want to read classic Christian literature, demographically and by location.
2. Have some modern-language classics and other resources
3. Offer a website that helps users discover interesting books and find answers to questions, encouraging them to read classic Christian books
4. Have a thriving community of people studying these books together and discussing them
5. Enable volunteers to contribute meaningful work
6. Offer opportunities for Calvin students and other faculty
7. Be financially self-sustaining and long-lasting

4.1. Breadth of audience

The audience of the CCEL according to our last user survey is weighted toward middle-aged men. This is perhaps not surprising for a website that caters to theologians and pastors as well as others, but we would like to broaden the audience to include a better age and gender distribution as well as more users who are not professionals. It is perhaps inevitable that we will cater to users who like to read. The “location” part of the objective implies serving users around the world, including in places where Internet access is poor. If we can find a way to address users whose primary language is other than English, so much the better, though that is beyond our current capabilities.

Financially, the net result of investing in the library must be increased page views or subscriptions, in order to be self-sustaining. An increase of page views by 50% after five years would mean about \$30,000 more in ad revenue per year as well as a boost in other revenue—justifying at least that much in additional spending to achieve that increase.

4.2. Modern-language texts

The second objective could perhaps have been considered a strategy for achieving the first objective. Because most of the books on the CCEL are in the public domain, they are largely dated 1922 or earlier. In fact, most of them were written or translated in the nineteenth century. The language is often a barrier, especially for younger readers. Many professors will choose to assign a newer translation of a work rather than use a (free) translation at the CCEL. One way to broaden the

CCEL's appeal would be to have some more modern translations of popular works and other modern resources.

4.3. Website

The third objective is about meeting our mission of promoting Christian classics by offering a website that meets users' needs. A consideration of the CCEL's current design led the board to suggest that it could be improved in three areas: discovery, finding answers to questions, and encouraging users to read.

The CCEL has long excelled at offering good books, presented on the Web and available in various other formats. However, discovery tools and browsing capabilities have been more limited. The search engine works fairly well and got 2.5 million queries in 2012, but it doesn't have spelling correction or semantic search tools, making it a little less useful than state-of-the-art search tools. It does allow users to search by scripture reference, author, and book, three key capabilities that bring users to our site. However, fully 37% of the top 2,500 default searches were for topics such as prayer, compassion, faith, grace, forgiveness. Clearly, users would like to be able to find resources on such topics, and this is the area the CCEL has not served well.

In addition to offering a way of finding information on such topics, we should offer better tools for discovering good books to read and browsing them. One board member commented that more context is needed in various situations. Users should feel as though the CCEL is a good place to explore, with many interesting options offered.

Finally, usability for reading is absolutely fundamental for meeting our mission. If we want to encourage users to read these classics, we must offer a top-notch reading experience. We have put considerable effort into making the experience good, but more can be done, and more systematically, to optimize it.

In the past, many changes were made to the website in an attempt to improve the look and experience. However, little data was gathered analyzing the result. We need to perform more user testing and analysis of usage patterns to assess any changes we make to the website to make sure they are serving their purpose and increasing usability and usage. Assessment must be built into initiatives in the planning stage.

4.4. Community

User discussions can be an excellent way to encourage users to read and engage these classic books. The online discussion groups are doing reasonably well, with 16,742 members enrolled 26,231 times in groups. However, many of these groups are not directly on discussion of the books, but on topics such as prayer, Bible study, theological discussion, etc. The discussion groups and community features of the site do play a role in promoting the use of the site, but the board advised that the

priority of promoting the social aspects of the website should be a little lower than the highest priority activities.

4.5. Enable volunteers

Volunteers have performed an important role in the development of the CCEL, having contributed multiple person-years of labor in helping prepare books. In addition, the volunteer work is good for the volunteer, in that it gives them something useful to do with their time and it gets them into a close reading of one of these classics. Thus, supporting and enabling volunteers has always been one of the ways we meet our mission.

However, while there have been hundreds of people volunteering their time, most volunteers don't contribute any useful work (a few contribute the vast majority), and they all require significant time in developing a relationship, giving jobs, training, etc. A significant amount of time must be invested in working with volunteers to enable them to contribute. We need a way to enable volunteers that is inviting, sustainable, and does not require excessive resources from the CCEL staff.

4.6. Offer opportunities at Calvin

The mission statement and values make it clear that the CCEL is not solely focused externally, but also at providing opportunities for students and faculty at Calvin. The CCEL has met that mission well for students, with scores of interns supported over the years. We have offered opportunities for faculty, so far without any significant takers. We can perhaps make additional opportunities available. Recently-retired faculty might have more time to contribute.

4.7. Be self-sustaining

The last objective is to generate enough revenue to keep the project going and give it longevity. In the last few years, self-generated revenue from advertising, subscriptions, and sales, has grown to the point where it will sustain the project at a basic level—perhaps enough to support the business manager, developer, and some time from the director. Other initiatives, such as developing new capabilities or adding large numbers of resources, have required additional funding through grants or contributions. This model seems to have worked for us so far.

5. Strategies

Strategies are particular initiatives that we may take to reach one or more of the objectives. Typically, implementing strategies will require some of our (limited) resources of time and money, and they will be likely to generate financial return. In addition, some will help more with achieving objectives than others, so we will have to balance many factors in prioritizing. We should also ensure that all of our objectives are addressed by suitable strategies.

For each strategy, we will explain the idea and describe what an implementation might look like. We will describe an assessment procedure. We will analyze our strengths and weaknesses in relation to the proposed strategy and external opportunities and threats. We will consider how the strategy helps us meet objectives, and we will estimate cost and return. Finally, we will estimate a priority. The following possible strategies were prepared with input from the advisory board meeting of January 2013 on an earlier draft:

1. *Financial model*: advertising, sales of eBooks in various formats, etc.
2. *Optimize experience*: optimize the website experience on various devices for our mission, including discovery tools, answers to questions, and a more dynamic front page
3. *Semantic tools*: differentiate the user experience with semantic search and other semantic tools
4. *Treated terms*: create a bible-handbook-like question-answering resource with fully treated entries for common search topics
5. *Dictionary system*: create a new system for accessing digitized or scanned dictionaries, encyclopedias
6. *10,000 titles*: be comprehensive, adding all relevant scanned books, including multiple editions, with automatic/crowd-sourced correction and markup
7. *Enable volunteers*: offer ways for volunteers to help, including an easy-to-use way to make corrections and add markup
8. *New content*: license or create newer content for subscribers
9. *Pithy quotes*: Gather pithy quotes for repurposing
10. *Expand formats*: add capabilities to the formats we offer and offer more books in a variety of formats
11. *Revitalize subscriptions*, perhaps for mobile devices
12. *Improve community*: improve study groups and other community aspects
13. *Distribute eBooks worldwide*: raise funds to distribute eBooks to developing countries
14. *Apply for grants*
15. *Social Media*: promote CCEL activities through Facebook, etc.

16. *20 Languages*

17. *Support Copyrighted Books*

18. *Publisher Partnership*

5.1. Financial Model

The shape our financial model has taken is the result of many years' experience and evolution. Currently, we have less emphasis on asking users to donate or subscribe and more emphasis on selling books in non-Web formats such as Kindle and iBook. Empirically, our current financial model (2011/12 fiscal year, ccel.org only) is that advertising provided 36% of our self-generated income (\$60k) and sales and licensing of our content another 61% (\$100k). The content sales figure includes formats, CDs, iOS apps, subscriptions, and licensing royalties. The financial model under consideration here is for self-generated revenue; grants and gifts have supplemented these figures for each of the last several years.

5.1.1. SWOT Analysis

Strengths

- We have a great deal of content for licensing or sale
- Bible-verse tagging gives us unique capabilities
- Data-mining value of 1000 tagged books
- Four to five million people a year visit our website performing 2.5 million searches
- We have the technical expertise and ability to work at the cutting edge
- Access to student programmers, for inexpensive development
- Access to student writers, for inexpensive creation of content that doesn't have to be of the highest level
- We have access to experts at Calvin (theologians, etc)
- Some of our content is the latest edition and very popular (ECF, MHC, Calcom)
- We have many loyal users who are willing to contribute, help

Weaknesses

- Some of the content could use further proofing and refinement
- Much of our content is nineteenth-century and less readable
- We can't easily undertake large, complicated development projects with students and the inevitable turnover
- Our ability to hire employees is constrained by Calvin rules
- Website usability (including search) could be improved

Opportunities

- eBooks are selling well. Kindles and tablets are popular and growing.
- Traditional publishers are only just starting to make books available as eBooks
- We can add distinctives to our eBooks through custom programming and possibly semantic tools and data mining

- New technologies and network bandwidth make facsimile editions more viable
- Users performed millions of searches per year, many for non-book topics
- Our good Google page rank means we can create new pages that users might be able to find through Google searches
- Data mining, semantic tools, and computational linguistics could give us ways of generating content beyond what most others are able to do
- Google Books and Internet Archive have public domain scans of thousand classic Christian books online
- Our advertising revenue is below industry norms, so there is room for growth

Threats

- Sales of public domain eBooks could decline if publishers refuse to sell them or offer free editions
- Web traffic is quickly going mobile (7.5% in 2011, 15% in 2012). We get less revenue per mobile page view
- Google Books offers almost all of our books, in many editions, free
- CCEL page views have been declining and the trend will likely continue if nothing changes
- It is difficult to hire enough good student programmers

Currently, the majority of our self-generated revenue is from sales of formats and advertising. The subscription program was previously a key component and it could be again. We should look at optimizing these three key streams of revenue. However, we should also evaluate the other, smaller sources of revenue, including affiliate fees, licensing, and sales of CDs/DVDs, and either optimize them or phase them out entirely. We should also identify the effects of the growing mobile use of the CCEL on our financial model.

Sales of eBooks and the subscription program will be addressed in separate strategies, below.

5.1.2. Advertising

Advertising is doing better than ever at CCEL, with a net revenue of \$2.75 per thousand page views for all ads on a page in 2012—a respectable figure when using ad brokers. The ads aren't overly obtrusive, as we don't allow motion in ads. However, higher returns are possible with targeted ads, direct selling, inline placement and motion. One advertising consultant suggests that it should be possible for a startup to get revenue up to \$10 RPM after a few years. However, this would be labor intensive, as a sales person would have to contact advertisers directly and the site would have to be modified. Still, with the potential to increase our advertising revenue from \$60,000 per year to \$225,000, significant effort to optimize is warranted.

A first step would be to optimize our current advertising system. Placement on some pages could be improved. Because we don't allow motion, AdSense will only provide text-based ads, but graphical ads have higher return. We should experiment with different ad types and different placement, noting revenue, tracking of user feedback, and watching traffic indicators such as page views per visit.

If we are willing to modify the site and put effort into securing relevant ads directly, it may be possible to increase ad revenue and make the ads less annoying at the same time. The approach would be to show ads that are relevant to our users and put them into the main flow of content. The most relevant type of ad would presumably be for new Christian books. For example, we could solicit ads from Christian publishers for new books and insert them into our browse lists, labeled as advertisements. Users probably wouldn't mind seeing these ads and they should generate good revenue.

Our problem has been selling ads, which is very time consuming. We may wish to hire an advertising consultant, to help design the advertising program. In addition, since we have much unsold inventory, we should test direct selling by hiring a temporary, part-time advertising sales person and assessing the result.

5.1.3. Affiliate Links

We could also place our own ad-like entries in the browse lists without dealing with publishers. We could select newer classics and create full treatments, as for the other books. They show up in browse lists and search results, just like other books. They would also have book info pages, including the contextual information on our other book info pages, but without a link to read online. Instead, there would be affiliate links for purchasing the book at Amazon.com or Christian Book Distributors (CBD). We could even consider making the books searchable and offering excerpts. This would have the benefit that topical articles pointing to important readings on a topic could point to these pages.

We tried affiliate links like this in the past, and they were never a major source of revenue. They are not likely to be a major source of revenue in the future—they are unlikely to generate \$1,000 in revenue per year even with optimistic assumptions. The decision of whether to use affiliate links should be based primarily on whether adding modern classics to our collection adds value for users.

5.1.4. Effects of Mobile Computing

Mobile traffic reached 7.5% of page views in 2011 and 15% in 2012. It's likely to keep growing for some time, especially since mobile devices such as iPads are more pleasant for browsing and long-format reading. It could reach 50% or more of page views in a few years.

CCEL has a mobile version, m.ccel.org, that is optimized for small devices such as smart phones. It works reasonably well—usability is definitely improved over using

the full site on a smart phone. There are some usability issues—the mobile site book page view currently has a non-critical bug on iPhones. Page turns are not entirely smooth. Some parts of the website are inaccessible. Using the CCEL on an iPad is a bit problematic, too, in that the mobile site is optimized for too small a screen and the desktop site for too large a screen. A responsive site would be better.

Purchasing ePub books is also problematic, and few people are willing to enter name, address, and credit card info on a website on a smart phone. Over all, users view 4.04 pages per visit, compared to 4.43 pages per visit for the desktop site.

Currently, our mobile users see very few ads, though the few ads they see receive a good CPM. In 2012, our mobile ads generated \$1,317 in revenue. Although individual ads received a CPM of over \$4, because there are so few, the effective mobile CPM was \$0.45, compared to \$2.75 for the desktop site.

Mobile users also have little reason to subscribe, unless perhaps to buy 10 eBooks at once. Advertising and subscription revenue together netted \$3 per 1000 mobile visits, compared to \$13 per for the desktop, with virtually all of the difference from advertising revenue. If 50% of our traffic goes mobile, that would result in a loss of nearly \$20k per year in ad revenue at current returns. We should try to raise mobile advertising revenue to desktop levels, although that may be difficult—apparently Facebook also has the problem that mobile revenue is much less than desktop revenue.

Creating a nicer book browse list (discussed below), with inline ads, should improve mobile advertising revenue. We could also put effort into encouraging mobile users to subscribe, although it may be that our time would be better used improving advertising revenue. The subscription program will be discussed in another strategy, below.

We should also promote selling eBook versions and make it very easy to purchase them. Prominent Kindle and iBook links should lead directly to the Kindle or iBook store page for the book. We should also add a link to a store for Android, perhaps Kobo.

If a book being browsed is not available in one of these stores, we can sell our own version directly. However, we shouldn't require users to enter name, address, and credit card information—that is too much of a barrier. Perhaps we can set up an Amazon Web store and take advantage of the fact that people are comfortable ordering from Amazon and probably have a credit card on file already. In any case, the user experience should be carefully optimized to maximize conversions.

5.1.5. Other Revenue

Grants have always been a key component of our revenue. Applying for grants will be discussed in another strategy, below.

Sales of data CDs/DVDs (as physical media and downloads) was at one time the largest component of our revenue, but it has decreased to near zero. In 2012, total revenue was \$4,286. Our CDs are out of date, and they require technical support. Creating a new one would not be cost-effective. It is time to phase them out rather than test and update them for the latest version of Windows. Audio CDs/DVDs don't have compatibility and support issues, so it does no harm to keep them available. They can also be offered as premiums for subscribers, perhaps free as download and with a shipping fee for physical media.

Since licensing revenue has averaged only \$5,000 per year for the last five years, it is probably not worth a significant amount of effort to optimize this source of revenue. However, we could add a contact link for licensing info on our website.

5.1.6. Assessment

In order to assess the success of our financial model, we will need access to certain data, updated regularly and automatically. There should be a way to track this data on the Monitor page and compare to prior time periods. The data to be gathered should show, for selected dates,

- Automatic setup and evaluation of A/B tests for different versions of pages, with the stats below reported separately for A and B pages if appropriate
- Mobile/desktop page views, pages per visit
- Mobile/desktop net CPM for advertising, subscriptions, eBook sales, total
- Mobile/desktop eBook, subscription conversion rates
- Affiliate fee revenue
- Percent of advertising inventory sold
- Click-through rates for inline ads

When we are testing new pages or systems, we should run A/B tests and assess the level of improvement.

5.2. Optimize Experience

5.2.1. Vision

The most important ways we serve our users are by enabling them to find good books and offering a good reading experience and encouragement to read. We also offer aids to research including search, topical resources, and community features. The site ought to feature and highlight these things.

The mission and motivation of the site should be evident in the look and in the way the site works. It ought to draw users into browsing books, finding something good to read, or researching a topic or Bible passage. The site should be up-to-date and engaging, to attract users, and it should be mobile-friendly, with modified pages implemented responsively, so that the m.ccel.org is needed less and less.

We should implement the site so that it will be known as a premiere site for finding good Christian books to read, researching those, and getting access to them or a way of acquiring them. To gain this reputation, it will have to offer top browse and research facilities, and it will have to cover current literature as well as public-domain classics.

5.2.2. Front Page

Front-page visitors are often people who arrive at the site by Web search or by following other links on the Web and then navigating to the home page to learn more about the site. They may also be people showing the site to others. In either case, the primary roles of the front page are to

- Give a sense of the purpose and mission of the site
- Highlight the main tasks that users can perform
- Show how the library is used, including community features
- Serve as a site navigation page, including users-specific elements such as bookmarks and recently-read books

The front page is currently static and graphically oriented, and it has low information density. It looks the same at every visit, affording little incentive to return. It gives a sense of the purpose of the site, as a classic Christian library, and it highlights five common tasks. However, it doesn't give an indication of the activity of other users, and many of the navigation links are buried. Perhaps it can be improved in the following ways:

- Highlight the mission and the missional aspects of the library
- Highlight a few primary tasks for new users, suggesting what to do next
- Show additional navigation options by adding links under the task icons
- Add indications of community activity, such as popular books or topics and editing projects and progress
- Highlight a particular book or study group, changing every day

- Add users-specific navigation options including bookmarks and study groups
- Change from day to day

Search is a fundamental, primary activity at the CCEL, and we should feature it front and center. Perhaps we can add a bigger search box to the right half of the library image, on top of the stacks of books in the image, with a concept similar to that of the search box on the Hymnary.org home pages.

Screen sizes for desktop PCs are growing, so additional content below the current bottom of the page would be visible to many users. We could highlight additional navigation options below the task icons as follows:

Research

- Search
- Encyc of Christianity
- Early Church Fathers

Study

- Study Bible
- Commentaries
- Topical Bible

Read

- Browse books
- New and popular
- What we're reading

Meditate

- Daily readings
- Forums
- Book discussions

Give

- Proofread books
- Lead study groups
- Donate

Below these navigational links, we could have more dynamic displays to highlight popular books, current discussions, and the like, in a magazine-like layout that changes each day. We could include elements such as the following:

- Top 5 most popular books of the last week, tiny view
- Most active forum discussions of the last week
- User bookmarks, recent annotations, recently viewed books
- Activity in groups that the user is a member of
- Pithy quote of the day

- Excerpt of the most recent newsletter
- Recently edited books
- Percent done thermometer for a current book project
- Top volunteer proofreaders, etc.
- A zoomable map of the world with dots showing recent accesses and what is being read

All of these options make sense for a large desktop display, but for mobile devices and smaller screens, we'll have to reduce the amount of content judiciously. Menus are too fiddly for mobile devices, so it's best to have all information available by scrolling down the page or touching to expand a hidden section.

5.2.3. Discovery

One advisory board member commented that discovery doesn't work well. More context is needed in browse lists. Simply seeing a list of authors or titles tells nothing about why you might want to read a book. There are few good tools for helping find something interesting to read, depending on a user's needs and desires.

By adding a good browsing and discovery facility, we could replace all of the different browse pages and the recommendations page. This facility could replace the "read" homepage.

For each book we could define a full-width, 1.5" high view that contains lots of context—cover art, title, author, date written, first part of blurb, subject, star ratings, importance, popularity last week/month/year, estimated error rate, reading level, first couple lines of blurb, etc.—visually interesting, with enough data to encourage browsing and aid discover.

We could then offer a single browse page with all the books theoretically shown (scrolling up would load in more via Ajax). The page could be ordered by author, title, popularity, importance, date of first publication, etc. There would be widgets for narrowing the list of books shown: a place to type in a keyword, a subject selector via LC call number, language, formats available, whether recommended by ccel, etc. There would also be a widget for controlling the sort order: alphabetically, by popularity, rating, etc. It should be possible to construct a URL to display a particular list of books, such as the most popular of the last week.

In order to be a comprehensive tool for finding good Christian books, we'll have to display copyrighted modern classics as well as PD books included in our library. However, the number of such books in the list should be strictly limited so that most of the books in the list are in the library.

The list could be ever-scrolling, so that when the user nears the bottom of the list, additional entries are loaded via Ajax. Entries could perhaps be stored in a local

HTML5 Web database for performance reasons, especially on mobile devices with finger scrolling.

This arrangement doesn't work well for sidebar ads, but it would be a natural for inline ads the same size as book entries. We could offer ads after every few books.

We could also offer a topic drill-down widget on the research page. This widget would offer high-level subjects from the Library of Congress hierarchy, with users able to click to drill down to more particular subjects, ending in books.

A similar facility could be offered for browsing authors, with enough contextual information about each author that users will be able to determine at a glance which authors to research more carefully. The author browse page would include a low, full-width view of each author giving picture, name, years, importance, popularity, and one-line description.

5.2.4. Browsing Books

After looking through a browse list, users may click on a book for more information. These book information pages also need more context, so that users are better able to decide whether to read the book. They should be more encouraging of users to read more about the book and think about it, and they should offer more information about the book, in a way that encourages browsing. We need a full treatment for books.

The full treatment should be graphically interesting and have many sections, like Amazon's book pages or Hymnary.org hymn pages. Long sections should show just the first two inches or so of content, with a link to expand or bring you to another page with full details.

To help users decide whether to read a book, helpful context could include

- Cover art, title, author, LOC subject, language, date of first publication
- Measures: readability, importance of book, popularity, proofing status, number of citations, length of Wikipedia article, star rating from CCEL users
- Search inside this book
- Blurb about the book
- Other references – what we're reading articles, mentions in topical articles
- A note about the author, with picture and link to author page
- Info about this edition and other editions available (automatically generated)
- People, places, concepts addressed in this book (automatically generated)
- Links to reviews at Amazon or elsewhere
- Link to Wikipedia article
- Formats available
- Study guide (for use in small-group study, etc)
- Amazon, CBD links for buying in print

Therefore, full treatments of books should include the following elements:

- Blurb
- Cover art
- OCLC id number (for MARC record)
- Wikipedia identifier
- Amazon, CBD identifiers
- Optional: notes (expert-written literature guide for terms and books)

In addition, someone should go through all of our dictionaries, encyclopedias, and newsletters to find mentions of our books in them and add them to the full treatment pages for the books.

When term treatments are created with literature surveys, references should be added to the book full treatments if they are mentioned.

Getting experts to write articles about topics and books could be difficult. Perhaps we can find someone who is experienced in the area and make and offer to help them prepare and publish a book with the results of this work—an introduction to Christian literature, say. We could pay them an honorarium and agree to make a book of it all, and make a Kindle edition of the book. The author could seek a print publisher if desired. We could also try to find such a book already in existence and try to license its content.

We could also add a way for users to suggest readings that are important for key topics, as follows: when a users selects a passage, the box that pops up could say “This is an important reading on the following topic:” with a topic drill-down box allowing the user to specify a topic.

5.2.5. Reading Books

Reading books is the most important activity at CCEL, an activity we must encourage in order to meet our mission, so the experience should be as good as we can make it. The experience for our current read pages is pretty good, but there is room for improvement. Currently our read pages have a toolbar along the top with controls for Table of Contents, Previous section, Next section, and Settings. The default font is sans-serif and rather small. It is possible to set a preference to use a serif font and increase the font size, but not enough for my preference, and the serif font is Times—not my favorite. There are two tall AdSense ad blocks in the right sidebar and one at the bottom. These ad blocks aren’t very attractive.

Settings allow users to specify whether the footnotes should be in the margin, at the bottom of the page, or hidden. It is also possible to set a Bible version preference, add a subject tag to the page, and search the book. Arrow keys can be used for navigation. Double-clicking to select a word or phrase enables users to highlight, add a note, and look up a word in a dictionary. In all, the experience is pretty good,

with a couple of areas that could be improved. However, there are quite a few distractions on the page.

We need our book pages to offer an excellent reading experience, with good typography, plenty of whitespace, and minimal distractions. Eliminating advertising would be desirable, but with half of our ads on book pages, we can't afford to eliminate them entirely. Instead, perhaps we can find an advertising strategy that is effective and distracting for only a fraction of the time. If we make the ads relevant and interesting for users, they are less objectionable.

We should also keep in mind that mobile traffic is increasing, and it is likely that within 3-5 years the majority of long-format readers will be doing so from an iPad, smart phone, or similar device. Therefore we should design for the tablet, with the Web as an afterthought, rather than the other way around. One major consideration is that there isn't room for sidebars in a tablet portrait view. Therefore we should use a sidebar-optional design, with a sidebar used for helps for the book being read rather than advertising. All links and controls should be large enough to use on a mobile device, and usability should be tested on tablets and smart phones.

Safari's Reader mode offers highly readable versions of Web pages. Like Reader mode, our book-reading pages should have black text on a white background, with pages 6" wide or so at tablet viewing distances, 7.5" wide at desktop distances. Margins should be $\frac{3}{4}$ inch. It should use a large, readable serif font so that there are 10-12 words on a line in the default setting. We should use full justification and automatic hyphenation since that is the practice in print books and screen resolutions now available support adequate kerning.

Safari's reader mode displays text in what look like print pages, with shadow surround, one for each section it displays. In my testing it sometimes showed the current page and at other times the current one and the next one, with page numbers at the top of each. We could consider a dark semi-transparent screen cover to make everything but the page darken.

Perhaps we can improve on reader mode and load additional sections via AJAX so that users can scroll continuously to the end of the book, or backwards in the book, as well as using the table of contents and arrow keys for navigation. Note that using ever-scrolling pages will reduce page views. Therefore for assessment we will have to measure ad impressions and time on site per visit, rather than counting page views.

The page header should largely function as before. It should show the book short title, next and previous arrows, and icons for ToC and settings. Single-clicking the page in the header area should cause the page header to fade out, and clicking again should cause it to reappear.

Page navigation should happen in several ways. Up and down arrow keys should quickly scroll to the previous or next screen of text with a slide-up/down animation.

The left and right arrow keys and navigation links on the page headers should take users to the previous or next section. If multiple sections are loaded, the current section is defined as the one currently visible, or if several sections are displayed, the bottom-most section whose end is visible. On tablets, swipes replace arrow keys.

Footnotes should be shown in floating divs, visible or hidden, and clicking the footnote marker should toggle the state. There should be settings options to show or hide footnotes, which setting becomes the new user default. Page breaks should continue to have a floating icon, and bookmarks and notes should be shown with inline icons. On desktop computers, words in the text for which we have treatment pages should show underline on hover.

One additional idea to consider is a page image mode with images that can be reflowed and resized, which is possible when you have page image coordinates for each word. Users would be able to select the page image view and the text should fade out and image fade in.

Info Sidebar

Our book reading pages are a key to meeting our mission, and the capabilities and ease of use we offer will be an important part of our differentiation from the likes of Google Books and Internet Archive.

The semantic tools strategy (described below) will enable us to offer study aids such as scripture popups, Bible commentaries, and treatments of theological topics, people, idols, books, etc. These study aids may be helpful or distracting, depending on user needs and preferences, so they should be optional.

For the user interface on tablets, in landscape mode we can default to one main column with an info sidebar as essentially a second column, on the right. It should be resizable and closable. It should be visually distinguished from the main column, but not totally separated. Perhaps this can be accomplished with a subtle buff background color but no separating border. It should have smaller margins and smaller text. Selecting a word or passage brings up helps and other relevant information and options in the info sidebar. In portrait mode, the information by default appears in a popup on top of the text. In either case, there should be an option in the popup to open the info sidebar and show the information there. In a desktop Web browser, the same options should be available, with the info sidebar appearing by default.

The info sidebar should have a common section at the top and bottom with two tabs in the middle: Helps and ToC. The common section above the tabs should show

- A search box to enter search or navigation terms
- A settings icon

- A bookmark icon to add a bookmark and a show bookmarks link. There should be a box in which users can add a bookmark name, defaulting to Bookmark 1, Bookmark 2, etc.

The search box should allow a user to search and navigate the book. The user should be able to enter full-text search terms, navigation terms (Matt 5:8, chapter 1, contents), and treated topic terms. For navigation terms, it jumps immediately to the specified section. For full text and term searches, hits can be displayed on the helps tab, with clicking on one navigating to that section and highlighting the hit.

The show bookmarks option should show the bookmarks for this book by name on the helps tab. Clicking on one should navigate to that spot.

The settings icon should bring up the settings popup. It should function much as it does now, with the following settings options:

- Font size (smaller, larger) and style (serif, sans-serif)
- Bible version preference
- Footnotes: (show or hide)
- Arrow key navigation checkbox
- A Show helps checkbox that causes page breaks to be shown in the margin, treated terms to be underlined, semantic markup indicated, etc
- A Share link (with an “include my annotations” and “selected text only” checkboxes)

The Share link would allow a professor, for example, to select a particular reading that spans multiple sections, add annotations, and make a link to it. This implies that it must be possible to create a URL that shows a selection from a number of sections and shows user annotations (when thus authorized). No URL should be able to show user annotations unless a user explicitly shares them with the above permalink option, of course.

The common section below the tabs should show stats about the book: accesses, proofing status, progress bar, and the reader’s daily/weekly/yearly percentile rank for pages read compared to other CCEL readers (as described below).

In the tabs section of the info sidebar, the ToC tab should show the table of contents, with the current section highlighted. Clicking an entry should navigate to that section. It should be possible to expand or collapse ToC entries that have children, and initially, the ToC should be collapsed as much as possible while still showing the current section. Long lines should not wrap. The highlight on the current section should automatically update upon navigation or scrolling to a new section, with collapsed ToC entries expanded as needed and the ToC scrolling to keep the current section in view.

The contents of the Helps tab depends on whether text is selected, and if so, whether the selection is a single word or a longer passage. When no text is selected, the box should show

- Grayed-out highlighter icons and an annotate box containing “Select text to add a note or highlight”
- A note saying “Select text to see helps”
- Treated topics and terms in the shown passage, linked to the treatment page

When a single word is selected, it should show

- Highlighter icons
- Annotation box
- The scan of the word, with a box for entering a correction
- A link for correction history
- A dictionary definition, if it is an uncommon English word
- A disambiguation box for named entities, if relevant (“This refers to:” with bulleted options: Jordan [River], Jordan [Country], Michael Jordan, Other _____)
- Any markup for the word
- Options for adding markup
- A CCEL treatment for the word, if available

When a longer passage is selected, the following should be offered:

- Highlighter icons
- Annotation box
- Treated topics and terms in the selection
- Commentary available (selecting one puts you on commentary tab)
- Correction history link
- Options for adding markup

If users who are not logged in highlight or annotate text or change settings, they should be warned that their changes will not last (“log in to maintain changes beyond this session”).

Editing and Differences

The book-reading interface should also support editing with minimal effort for readers to help out. Some editing capabilities have been described above. However, we will also need a “diffs” mode to compare versions of a text. It would be good to be able to see changes made by date, by a particular user, and by processes such as OCR error autocorrection.

We should use the Open Annotation framework for storing corrections, and it should be possible to query the annotation server to find corrections from a particular source and display them as diffs.

Incentives for Reading, Proofing, and Markup

The user interface can also encourage reading, proofing, and markup more explicitly. By offering measures of achievement, goals, and rewards, some (though not all) users may be encouraged to read or proof more. These rewards can serve as a tangible indicator of progress for good activities with fewer inherent tangible results.

For reading, tangible indicators could include a progress bar for books. There could also be daily, monthly, and yearly indicators of reading rank, i.e. number of pages read compared to other CCEL users, as a percentile. These could appear as small vertical progress bars beside the (horizontal) book progress bar on the info page. Clicking them should give a popup with more detail. These would appear for both authenticated and unauthenticated users, though for the latter only reading in that session would be tracked. The detail popup could have a message like “log in to maintain stats.” We could consider having awards, e.g. reader of the day/week/year awards on the detail popup.

More substantial incentives should be offered for proofing and markup. These incentives will be discussed in the 10,000 Titles strategy.

Multiple Columns

For some books, especially the Bible, we may wish to study two editions side by side: a translation and another version such as the original-language version, or a translation and a commentary.

It should be possible to view two aligned editions in separate columns. This would be accomplished by clicking a “see commentary” or “see original language edition” link in the helps sidebar. This would cause a second column to appear in the white section of the page, with the two texts aligned. The short title of each work would appear above it. Whitespace would show for the shorter edition at the end of each section, and it should be possible to adjust column widths. Selecting a passage in one column should also highlight the corresponding passage in the other column.

Advertising

Since we can't use a sidebar for advertising, we will have to place ads inline with text—wide advertisements the width of the page between sections of text and tall advertisements in a floating div inside the text. Every so often we could show an ad with a close box and an “eliminate ads” link. The “eliminate ads” page would talk about helping proof and mark up books or subscribing as ways of eliminating ads. These ads would be inserted dynamically.

Showing an ad at the top of every page or section might be too much or too little, depending on the size of the sections. Instead, we should show an ad on the second screenful of text, and then after every 1200 words or so—about 5 minutes' reading. We can experiment with ad frequency to balance distraction and revenue.

Other platforms

This section proposes a user interface for reading and studying books on the CCEL website, whether on desktop or mobile computing devices. However, as much as possible, this interface could be duplicated for other formats such as Kindle and iBook books, though advertising should not appear in books that users purchased. Adding helps and the ability to proof and mark up in eBook readers would improve and differentiate the experience.

iBook author supports JavaScript widgets that can alter the text being displayed, make AJAX requests, create databases, and the like, so much of this functionality could be implemented on that platform.

Kindle books and ePub 2.0-format books currently do not support JavaScript in a way that would enable us to support this functionality. However, it is possible that a useful subset of it could be implemented by creating two versions of each page, one with the info sidebar showing, and one with it hidden. Since such books cannot use JavaScript, functionality would have to be through links that lead to alternate versions of pages or to a separate browser window.

Resource Implications

The result of implementing this strategy could be a well-differentiated book-reading experience that encourages visitors to return and attracts educational use. It could help us meet our mission of promoting these classic books. It could help encourage users to read and contribute by offering incentives for reading, proofing, and markup. Ideally, it could help build a community of volunteers—the most dedicated sort of users. It could also serve as a differentiator for our books in other formats. In all these ways, it is central to meeting our mission, and it could help to increase traffic and sales of formats.

Implementing all of these capabilities would take several student summers. It would also require that the search modifications and term treatments strategies and possibly the semantic tools and 10,000 titles strategies be implemented and that we put additional effort into securing appropriate ads.

5.2.6. Study Bible

The book-study interface defined above will also work well for Bible study, in general. However, for Bible study there a few special needs:

- Seeing a list of sermons and meditations on a passage

- Seeing hymns related to a passage
- Seeing references in books to passages
- Special book/chapter navigation
- Being able to easily change the translation(s) or commentary(s) shown

The list of sermons, meditations, hymns, and book references related to a passage can be shown in the helps sidebar, perhaps just the first couple of lines if the section would be long. The UI for changing the visible Bible version or commentary could be accomplished with a small down-arrow beside the column title, which could open a list of options. Clicking on the passage could bring up a Bible passage selection widget.

5.2.7. Getting Answers

I have wished for years (decades?) that the CCEL had a way of enabling users to find answer to questions and pointers to key literature on a topic. People come to the website with a research topic in mind, but there is no easy way to find key literature on the topic. The Treated Topics strategy will provide pages with resources for answering these questions. The user interface issue is how users will pose the questions and be directed to the answers.

A two-fold, search and browse strategy seems best. On the research page, we could have a search box with the text “Search for a topic” in it. We had 2.7 million searches at CCEL in 2012, and more than half were of the sort that could be sent directly to a treated topic or book information page.

Beneath the search box we could have a link entitled Topics. This link could lead to a topics browse page, perhaps with several popular topics on one side as a hierarchical, expandable topical taxonomy on the other side. We could add a note or star by each topic that is fully treated.

5.2.8. Search

Search functionality will be treated in the semantic tools strategy below. This section will address the user interface for search.

Currently, there is a small search box on the home page, in the title bar. It should be larger in the title bar on most pages and prominently placed in the center of the home page. In addition, there is a slightly more advanced search form on the Research page, a Bible search box on the study page, and a search-this-book box on each book information page. There should be a search-this-author box on author pages and a book search box on the read page.

The search should have spelling correction, and it should have autocompletion based on common searches people perform at the CCEL and matching entries from the CCEL ontology

```
taj mahal
taj mahal
taj mahal – Monument in Agra, India
taj mahal – Trump Taj Mahal Casino Resort
taj mahal – Musician
taj mahal facts
tai mahal hotel
```

(described in in the Semantics strategy), as in the example from Google in the figure.

The search tabs on results pages are less than ideal. Users want immediate results, and those tabs hide results that may be most appropriate. For example, a search for “Augustine” returns full-text hits; you have to click the “Titles & Authors” tab to see books that Augustine wrote.

Searches should go directly to the best result page when one is (likely) known. Disambiguation can be handled with an autocomplete box, with the option that is currently going to be selected highlighted. Searches that don’t match one of the autocomplete options would default to a full-text search..

Results should be in an ever-scrolling list of mixed content types, with relatively rich context for each hit. Each entry can be about an inch high. The type of the result item should be recognizable by the design, height, background, and image.

Thus, each type of entry will need a small view optimized for search result lists. For books, it could show the cover art on the left, along with title, author, subject, importance, etc. and a highlighted quote hit if relevant. For people, it could show a picture of the person or a person place-holder picture and an indication of the type of person (author, biblical person, etc.). For scripture passages, it could show a Bible icon. For topics, people, etc., show an icon of a dictionary page with the term, pronunciation, and the start of the definition.

5.2.9. Task Pages

The CCEL offers three task landing pages; *Research using our library*, *Study a Bible passage*, *Read a good book*, *Meditate with a daily devotion*, and *Give of your time or money*. These task pages were conceptually to be useful, bookmark-worthy home pages for the mentioned tasks, but the implementations have only been moderately successful.

The *Study* page goes directly to the Study Bible interface, and the *Meditate* page goes straight to a daily devotional resource, both of which seem appropriate, though the resources themselves could be updated.

The *Research* page is intended for users looking to research a topic, book, or author. Possible things to include:

- Advanced search
- Search an author or series widget, set to Early Church Fathers
- Expandable, hierarchical outline of topics for which we have treatments; maybe a graphical representation of CCEL Dictionary of Christianity
- Popular topical articles of the last week
- Article of the Day? (Latest treatment article)

The *Read* page is intended to help users find a good book. As such, it is the logical home for the new faceted-search book browse page. That could be visible at the bottom of the page, with a couple of other blocks at the top of the page, for a top-10 list, a crowd-sourced “what we’re reading” block with a link for past recommendations, and a block pointing to book discussion groups.

The *Give* page should emphasize proofreading and markup, along with a paragraph promoting financial contributions. For graphical interest, it could present graphs and charts tracking top editors and the numbers of edits performed.

5.2.10. Groups, Forums

This topic is covered in Strategy 12.

5.2.11. Assessment

In the past, UI changes have had mixed results. The usability of the site is a key determining factor in its success, so assessing the effect of changes is crucial. We will want to know how well users like the changes and how the changes affect key metrics including the average number of page views per visit and the time spent on site. We should also assess effects on SEO status.

Before going live with a change, we should perform user testing to find any issues and assess how well users like the interface. Before going live, we should gather statistics from Google Analytics on the length of time people spend on a page and the rate at which users exit the site from the page and compare those to the time period after going live. We should also review user feedback after the change goes live.

Assessing the effects on SEO is harder, since there are many hidden variables. However, we can check certain pages before the change and a month after the change, and we can track the number of users reaching the page through Google.

5.2.12. Analysis

Many of the user interface changes proposed here, including the discovery, browse, book-reading, book-editing, question answering, and search interfaces, are dependent on and necessary for other strategies, and they will be assessed elsewhere. A few of the changes—improvements to the home page and landing pages, for example—are relatively low-resource tasks that we should work at continuously.

5.3. Semantic Tools

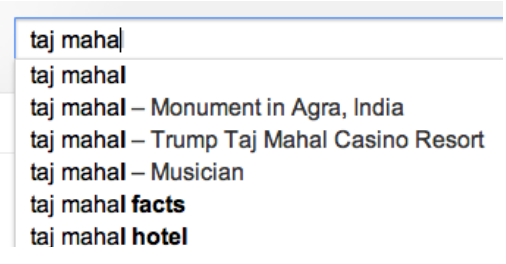
Semantic search attempts to answer user questions directly, as opposed to full-text search, which finds instances of the search string in documents. So, for example, a search for “Augustine’s Confessions” should go straight to that book, not to a page of full-text hits.

There were 2.7 million searches at CCEL in 2012. An analysis of the most common sorts shows that 12% were book searches, 11% scripture, 6% definition, and 5% “advanced”, for a total of one third of searches. Fully 2/3 were full text (default) searches. Among the most common 699 default searches, 48% were a book title or author, 31% a topic or person, 16% a Bible passage, and 5% navigation, full text, or other.

Thus, in all, 44% of searches were for a book title or author name, 26% for named topics and named entities such as people or places, and 22% were searches for Bible passages, though the scripture figure would likely be higher if we were considering searches, not just the top 699. The remainder—about 8% in our sample—would be for advanced searches, navigation, and the like. The default search mode—full text—is a kind of search users are rarely interested in. Thus, we can improve search behavior for about 95% of searches by returning immediate results for searches on book authors/titles, topics and people, scripture passages, and navigation terms. Remaining searches can default to full-text.

We can organize search as a number of modules looking for particular kinds of searches and returning likely results with probabilities for disambiguation.

Disambiguation could take place in an autocomplete list, as Google does in the figure, with the most likely interpretation on top, as default. When users select disambiguation options in the list, we should record the options selected and prioritize the disambiguation list for future searches accordingly.



A screenshot of a search autocomplete dropdown menu. The search input field contains the text "taj mahal". Below the input field, a list of suggestions is displayed. The suggestions are: "taj mahal", "taj mahal – Monument in Agra, India", "taj mahal – Trump Taj Mahal Casino Resort", "taj mahal – Musician", "taj mahal facts", and "taj mahal hotel". The word "facts" in the fifth suggestion is bolded.

An author/title search module should look for searches that match an author or title name. We would need an estimate of how likely a search is to be targeting a given author or book. The module could return likely hits with estimated probabilities.

A scripture search module should look for searches that are Bible passages. When a scripture passage is recognized, the search should bring the user immediately to the study bible page for that passage, with commentary showing and with options for other helps or other kinds of scripture search. Scripture searches can also be ambiguous, though. For example, a search for “John” could be a search for a book of the bible, the apostle, a part of the name of an author, etc.

For the purposes of this discussion, we will refer to topics, people, places, and the like all as named entities. To handle searches for named entities, we currently have a definitions search tab. It shows available entries in dictionaries, encyclopedias, and resources like *Nave's Topical Bible*. In the future, we should have a page for each named entity, and searches for such things (after disambiguation) should go directly to the entity treatment page. If there are additional reasons users might search for named entities—for example, a user might want to find mentions of the apostle John in all of the volumes of the *Early Church Fathers*—it should be possible to do such searches from the entity treatment page.

We also need a way of handling navigation and resource type searches. Users typed in search terms such as commentaries, catechism, mp3, maps, dictionary, audio, music, sermons, concordance, kindle, pdf, download, free, etc. We need a way to handle each type of search appropriately, with a response URL. Such searches could lead directly to resources or to other searches. For example, a search for “forums” could go directly to the forums page. A search for “mp3” could lead to the book browse page, showing books that have mp3 audio narrations available.

Users have asked for a way to search the forums. We could implement this as a forum search box on the forums page.

These semantic search tools could also be used to build tools for helping the study of the Bible and other books. In particular, when a section of text is studied that is on a topic, or contains a named entity, our helps on those named entity pages could be made available in the study sidebar or in a popup as well as on the treatment page.

5.3.1. Markup

We will need various kinds of markup and representation of semantic information for a full implementation of semantic tools and the associated strategies, such as treated terms. We will need to represent:

1. Named entities such as people, places, and theological topics, along with properties of the entities, search terms, and disambiguation information
2. Properties connecting these named entities with each other and with external entity collections such as DBpedia
3. Connections between named entities and books, e.g. when a book mentions “John”, which named entity is referred to; what book passages are relevant to the topic “trinity”

We should build an ontology to represent the first two types of semantic information. The third type could be represented as markup in ThML, but it will be added as users read the books, and it could come from other sources such as algorithmic agents or external sources. It may be best to store this information as stand-off markup using the Open Annotation Collaboration (OAC) model. Markup details will require further specification, possibly during the director's sabbatical in the spring of 2014.

5.3.2. Architecture

We will have to build an ontology of named entities with the search terms they might be identified with. We can start with dictionary and encyclopedia terms, search terms that don't fall into another category, and identified named entities in CCEL books. For each entity we will have to add some information such as a name to be used for disambiguation, the type of entity, relationships with other entities, and the like. Building this ontology will be a significant task requiring continued attention. One student worked on the bible names ontology for a summer.

The information added to a book through crowd-sourced corrections and semantic markup would be stored in the form of standoff markup on the server.

For each named entity, we will have to keep track of the number of times it was searched the number of times viewed, etc. These will help us with probability estimation for searching and prioritization for creating full treatments.

In order to support semantic aids for book study, the named entity treatment module will have to offer results as pages, sidebar content, or popups. For good-looking popup ideas, see the book *Stunning CSS3*. For SEO, the full name of the named entity should be in the URL.

We should build a JavaScript library that can be linked to a page to automatically enable the named entity study aids (with popups) for Web pages. We could let other people link this library on their Web pages to offer semantic helps for their Web pages. This JavaScript library would have to be able to identify named entities when text is selected, presumably by sending the text to a server for analysis.

5.3.3. Third Party Library

We should build a JavaScript library that adds these semantic tools to Web pages automatically, simply by linking and invoking the library. The code would find named entities and link them appropriately. In addition, it could link scripture references and citations to CCEL books (see the 10,000 titles strategy, below). If this turns out to be impractical, we could consider sending chunks of text to the server for analysis or using a preprocessor. Once the book has been processed, stand-off markup can be retrieved from the server and the page modified with links appropriately.

We would want to add CCEL branding and links and a small advertisement to each popup. Then, each popup would generate a small amount of revenue and could help increase traffic and Google page rank. We could even consider offering to pay users who link this library according to the number of popup views they generate. For example, it might be plausible to pay 25 cents for each thousand popup page views.

5.3.4. Assessment

Since searching is a critical component of use at the CCEL, we should have a way of displaying dynamic stats on how well search is performing. The goal should be to learn what kinds of searches users are performing, how well the search engine is giving the results users want, whether the pages thus found meet users' needs, and in general how we could serve them better.

Basic search stats for a given timeframe that should be available include number and type of searches, common search terms, number of follow-on searches, amount of time spent with target page, a proxy for whether searches were successful, and the most common searches that were not successful. We could analyze search logs for much of this data, since it probably doesn't have to be real time.

Analysis

Implementation of this strategy would greatly improve the search engine, and it is inherently tied to other strategies including the Named Entity Treatments, UI optimizations, and the semantic markup component of the 10,000 titles strategy, which have major implications for usability, differentiation, and traffic. It will require at least a couple of student-summer for basic implementation of the search modifications. However, many aspects of semantic tools are intimately tied with creating an ontology, treatments for named entities, and book reading tools, so those strategies will have to be analyzed as a group.

5.4. Treated Terms

For years I have wanted the CCEL to have a way to locate resources by topic. What are the most important readings on a theological topic such as *prayer*? What is there in the library about Mary? I would love to be able to type in such a term and see a survey of classic Christian literature on the subject, prepared by an expert. Ideally, in addition to a literature overview, it would have interesting facts, figures, and graphics, reminiscent of a Hymnary full-treatment page. Such a resource would meet a considerable need: an analysis of search logs shows that a quarter of CCEL searches may be for named entities such as theological topics, people, places, etc. In this strategy we would build a system that offers treatments for named entities. One idea for a name is CDOC (*CCEL Dictionary of Christianity*).

Readers of the CCEL would likely be visiting the site to research topics with a view toward what classic Christian literature has to say about them. Notionally, a user might visit the site, type in a search term such as *prayer*, and see an article naming schools of thought and giving pointers to key literature. Users might also want to be able to easily find treatments of the topic in the Early Church Fathers, or in a particular author, or in the whole library. Finally, it may be helpful to have some content or excerpts from additional sources such as the dictionaries and encyclopedias at CCEL or Wikipedia.

The resource should be dynamic in the sense that articles are continuously added and edited. It should be possible to browse and cite older versions of the resource. Two fundamental questions should be addressed concerning the organization of the resource: should it be academic? If not, should it be crowd-sourced and moderated?

The academic option would ensure that all articles are authored by respected academics in theology who are approved by an academic editorial board. This would ensure academic respectability and make the resource usable for academic research, but it would also be difficult and expensive. This might be a plausible approach with major grant funding.

If articles are written in-house, they should attempt to be neutral, giving an overview of schools of thought and key literature, and an author attribution should be given. At that point, we may as well allow third-party article submissions, with attribution and moderation. We should support multiple attributed mini-articles per treatment.

The literature surveys should cover all classic Christian books, not just those in the public domain. For books that the CCEL doesn't have, we could try to offer excerpts. We could try to include the full text in our search engine, with it able to return excerpts. We could also offer affiliate links for purchasing the books CCEL doesn't have.

We could offer similar full treatments for CCEL book and authors, giving background information and context. For authors, we already have our own brief bio in many cases. We could add other pertinent information, such as a one-phrase summary, years, denomination, a complete bibliography, measures of importance (number of citations, length of Wikipedia article, etc.), and links to additional resources. We should offer the capability of searching all of the author's works. Full treatments for CCEL books might include the year originally written, numbers of citations, links to reviews (scrounged from the Web), popularity over time in citations databases, and the like.

There were also a fair number of searches for people, biblical and otherwise. The top five searches were for Mary, Paul, Jesus, Abraham, and Constantine. For people who are not CCEL authors, a full treatment might include name, disambiguation name, years, occupation/description, links to external resources, and a short biographical article with pointers into mentions in classic Christian literature. It should also be possible to research the person from this page, finding references in the library, for example.

These pages would serve as identifiers for named entities and as places to gather the resources we have for those entities—links to Wikipedia, dictionary and encyclopedia entries, a place to launch a search with disambiguation already completed, etc. In order to use standard identifiers for entities as much as possible, we could consider using Wikipedia keys when available. We may also want to make our identifiers and information available on the semantic web.

We may want to use a notion of expert notes in some cases—small writings attributed to a person or source. So, for example, in a treatment of the term *prayer* there might be a note “John Calvin chapter on prayer is an all-time classic. – Prof. John Smith.” Then we could allow users to contribute such notes. It should be possible to comment on and rate these expert notes as well, with new and highly-rated notes rising to the top.

5.4.1. Implementation

We will probably need to have several types of named entities—theological topics, bible people, other people, places, books, authors—each with its own template and full treatment standards. We would want a well-designed schema for each type of entity, perhaps drawing on the schemas for entities in DBPedia. For entities that are not fully treated, we would at least want a disambiguation name, the DBPedia key, and data that can be gleaned from usage, including search terms leading to the entity, search and view counts, and relationships to other entities. We would also want to connect all of our dictionary definitions and encyclopedia entries to appropriate entities. We will want to develop ontologies for these entities, with properties and relationships between them.

For fully treated entities, typically there will be newly-written components centering on treatment of the topic in CCEL and other classic Christian literature, as well as a list of basic properties that should filled out.

It will take time and resources to fully treat entities, and it is likely that only the most important entities will be worth treating fully. Therefore we should consider allowing users to add information. This could take the form of a Wikipedia-style article or an annotated bibliography entry for useful literature on a topic, attributed to the name of the contributor. We could also allow users to fill in properties for entities, e.g. locations or pictures for places. There could also be a way for users to add additional entities to the ontology.

To create full treatments, we could find books and online reference material outlining the literature on key topics. From resources like these our editors could create literature survey articles. To jump-start the process, we could find a public-domain book with lots of literature survey content and use it to create starter articles. We could also consider licensing a modern resource, such as a Zondervan bible handbook.

Another idea would be to find a (retired?) seminary professor who would be interested in writing a book surveying classic Christian literature for key theological topics. We could partner with that professor, adding the articles to CCEL, and publishing a book in Kindle format afterwards.

To develop the list of terms for which we need full treatments, we could start with CCEL search terms, our existing ontologies, our existing books, authors, and series (e.g. Early Church Fathers), and common named entities from CCEL books. We would also want to develop an ontology of theological terms.

5.4.2. Analytics

We will want to gather data and use it to refine our list of entities, improve search results, and assist in disambiguation, and assist named entity markup in books. For these purposes, we will need a dynamic ontology with the ability to add and edit entities dynamically and merge entities. We will need to gather statistics on search terms and the entities users eventually ended up at. We will want to gather search and visit frequency data. We should also attempt to find associations and relationships between entities through usage data, if only of the sort “users who searched for x also searched for y.”

5.4.3. Analysis

Having information on named entities is essential for a number of other strategies, including semantic search and semantic tools. It will improve the usability and utility of the site and offer differentiation from other libraries. It would also add newly-written content which is hopefully easier to read. We will definitely want to

pursue this at least to the extent of automatically building named entities lists, inferring disambiguation information, and automatically building treatment pages.

Writing new content for a large number of full treatment pages would be more expensive. Ideally, with CCEL's good page rank, we would get first-page search rankings for full-treatment pages, and the traffic driven to the CCEL would pay for the cost of preparing full treatments. However, assuming 4 page views per visit, \$3 CPM advertising revenue, and a \$40 cost of fully treating a topic, we might estimate that treatment pages would need to attract 100-300 visits a month to pay for themselves. Currently, only our top 25 search terms attract 300 visits per month through Google searches, and the top 80 or so attract 100 visits per month. So it is unlikely that more than a few dozen full treatments will pay for themselves directly through immediate advertising revenue. On the other hand, topics such as "prayer," "compassion," "faith," or "grace" could bring in many more visits than names of authors and books (the search terms that currently bring in the most visits), if we can get good Google search rankings. Perhaps we should start out by creating 20-50 nice full treatments and determining whether they add value to the site for users and increase traffic.

To develop a full-fledged Encyclopedia of Christianity, especially with scholarly authoring and moderating, would be a much larger undertaking. It would probably require a large grant, collaboration with a well-known theologian or two, and buy-in from the academic community.

5.5. Dictionary System

Currently, our dictionary and encyclopedia system handles books that are specially marked in ThML as dictionaries. When dictionaries are installed, the terms it contains and links to definitions are stored in the database. If you do a definition search, you get a list of definitions with matching titles. Clicking a definition brings you to the appropriate article—usually. There is special ThML handling code to pretend that terms and definitions constitute separate sections in books and to navigate between them. There are a few bugs in the system, especially for book sections that contain multiple definitions.

The system does not handle dictionaries that have not been digitized. We have a number of high-value dictionaries, such as Schaff-Herzog and Hastings that have only been partially digitized, or not digitized at all. It is unlikely that these will be digitized in the near future. We would like to be able to make these usable as facsimile editions without complete digitization. In addition, we need to be able to use these definitions with the new Semantic Tools strategy. This means that it should be tied into our named entity system and the definitions should be retrievable from the database.

In order to support facsimile dictionaries, we will need markup that can reference page images in place of text—perhaps `` links with CSS setting the image offset and handling magnification. We would also need to mark up the dictionaries appropriately. For this, we would need a computerized tool that allows users to drag to select articles and type in the term being defined. It should output appropriate ThML.

Another option would be to use OCR software that outputs word coordinates. We could use OCR output to locate terms being defined, perhaps with manual assistance (click on the terms being defined in each image...). We could then infer page scan coordinates of the definitions from the OCR output.

I would imagine that building the tools could be done in a couple months of student programmer time, and marking up our facsimile encyclopedias could be done in another month or two. However, we already have a partially working system. There may also be synergies with the 10,000 Titles strategy—in locating terms being defined and inferring definition locations. The 10,000 titles strategy might also provide a suitable UI for viewing and improving markup of definitions. Therefore we may wish to delay implementation of this strategy for a couple of years to see how the need shakes out.

5.6. 10,000 Titles

Along with the Semantic Tools strategy, this is another proposed major strategic direction for the CCEL that would likely require grant funding to implement fully. The idea is to be comprehensive in the books we offer—to attempt to offer all public domain books by important Christian authors. Books would be automatically added from scans at Internet Archive, with automatic markup and links to page images. Users would be able to improve the markup and correct errors with minimal barriers. We could also align different editions of books, so that users reading a book could ask to see a passage in another translation or in the original language. The number 10,000 is only suggestive, not an actual goal or estimate of the number of books that would be involved, although expanding our coverage by a factor of 10 seems an appropriately ambitious goal.

We could locate titles by searching for Christian authors at Wikipedia and finding all their books. We also have some bibliographies of important books for theological study, and we could look at reading lists for seminary classes. We could also search a citation database to find theological books with many citations. (If a project such as citeseer doesn't handle theological books, we could consider extending it as a part of this strategy.)

5.6.1. Processing texts

We would estimate the importance of Christian authors by the length of their Wikipedia article, and we could find all scans of books by important authors at Internet Archive. For each book, we could look for other editions on the Web in order to improve our ability to correct OCR errors and infer structure.

We would correct errors in OCR by looking for likely errors using Google Ngrams. We would then look for the word's context in other editions, finding likely corrections, especially for replacements of likely OCR errors. Once corrected, we could compare versions to find all the editions of particular translations.

For automatic markup we could use location and size of text and comparisons with other editions of the same translation to locate title page, table of contents, front matter, body (chapters, sections, paragraphs), and back matter. Alternatively, we could build a Web app that allows an editor to enter some key information about the book by clicking on the title page, ToC, first page, and first page of back matter; page i, page 1, and the first of any other page segments; page headers and footers; etc. If we add 10,000 titles in the next few years, allocating 5 extra minutes of manual labor for each would be worthwhile if it saved about a student-summer of development. In any case, books should have accurate page numbers and table of contents before going online.

We would automatically mark page breaks, image coordinates for words, lines, and paragraphs, scripture references, named entities, and citations. If possible, we should identify and link footnotes. We could use computational linguistics

techniques to automatically find a topic for each section. For dictionaries or encyclopedias, we could have an automatic or manual method for identifying the terms at definitions. Dictionaries should have defined terms identified before going live.

The goal of aligning texts would be to create a list of standoff pointers linking chunks of a text to an exemplar. We could work hierarchically, aligning books/chapters/sections, then paragraphs, then sentences, making the alignments as detailed as possible. It would be good to be able to align an edition with other translations, and with versions in other languages or the original language, if possible. This may have to be done by hand for important works. Finally, we could find named entities from our ontology, marking them appropriately, or marking possibilities if they are ambiguous. With aligned texts, we would be able to offer a standard capability of viewing a passage in another version or language or viewing two versions in parallel columns.

5.6.2. User Interface

This strategy implies that the majority of the books at the CCEL will be automatically created from scans and OCR output. They will have residual OCR errors and errors in markup. Therefore a key aspect of this strategy would be making a nice interface for using such books and for correcting the errors and improving the markup online. We could offer game-like incentives for helping improve the books. An interface is suggested in the Optimize Experience strategy, but that section didn't offer a detailed proposal for the interface for error correction and markup.

We can easily offer a means of correcting OCR errors in a single word by showing a popup when the word is selected. The scan of the word could be shown next to an edit box containing the word. Users could make a correction without logging in, though we should show a message that you have to be logged in to get credit for your correction. Corrections that are unlikely should be automatically rejected.

If a term matches multiple entries in our ontology, we could allow users to disambiguate by showing the disambiguation names of matching named entities and allowing users to click on one. Users could select a section of text and, if it is an appropriate length, click the "this is a pithy quote" button. We could construct similar low-barrier, syntax-free interfaces for other kinds of markup.

This system serves a twofold function. It helps improve the books, and it also gives users a reason for reading and a sense of accomplishment. Therefore it should be possible to make corrections and add markup on mobile devices as well—at least, those with a big enough screen. These capabilities should also be added to Kindle or iBook book if possible. If not possible, perhaps that is a reason for selling book apps, though that would also interfere with sales of formats.

We should reward users for correcting and marking up texts with game-like incentives. This would give our many volunteers an easy way to help, and it would

encourage users to carefully read our books. We could offer incentives such as leveling up, icons by names, listings in top editors of the day, week, or month. We could also offer to eliminate ads or send free CDs for major contributors. We could institute a point system, offering a point for each correction or markup, proofing a page, forum post, etc. We could offer achievement levels, perhaps such as new convert, member, acolyte, deacon, elder, clergy, saint. Points would be deducted if edits are undone. More experienced editors might be better trusted.

5.6.3. Other Capabilities

Citations

One new capability that this strategy could enable is the ability to find and link citations. We can already find scripture references; perhaps we can write a similar program to find references to Augustine's *Confessions*, or to any of the works of Augustine. The open source CiteSeerX project on Sourceforge finds arbitrary citations; we should test that code to see whether it can be applied.

If finding and linking citations were a standard part of the book install process, we would have statistics about citations to books in important Christian literature. This would be an important new semantic capability for the library, making possible new forms of research such as the analysis of citations to books over time, among different groups of people.

Author Search

We should offer a prominent way of searching all of an author's works. This implies having all editions of an author's works and perhaps having quotes attributed to the author as well.

Offsite Books

We could also have book entries for books that are not on our library. We could include them in literature reviews and offer treatments of them, including reviews, etc. We could offer affiliate links for purchasing such books. We could also consider scanning a few of these books and including the text in the search engine. Searches and links in treatments could show brief excerpts.

MARC Records

We use the Library of Congress form of authors' names, and we will also want to use the LoC subject heads for browsing by subject. Perhaps the easiest way to enable this is to add an OCLC record number for a MARC record for each edition. Internet Archive already provides these. Then to get the subject heads or the author name we would retrieve the MARC record from OCLC. It would make sense to require a MARC record pointer for all books and drop everything in ThML that duplicates what can be found in the MARC record.

Markup

It will likely prove impractical or unwieldy to store all of the alignments and image coordinates in the ThML files. In addition, the semantic markup will be added and edited online. Therefore we will need to handle markup in multiple places. We may want to store some markup, such as coordinates and alignments, in separate stand-off markup files.

We will also want to support markup in the database but not in the ThML file. This would also be standoff markup. Therefore the authoritative version of a book would be the ThML file as modified by standoff markup files and entries in the database. Thus, before editing a file, we may have to make a new ThML file incorporating all the sources of standoff markup.

We will need to modify ThML or create new conventions to support some new types of markup. We will want a standard way of handling standoff markup. We will also need to be able to mark up citations and theological topics.

5.6.4. Assessment

If this strategy is implemented, we should assess the accuracy of OCR corrections and automatic markup. We should also attempt to assess whether the rewards are helpful in encouraging volunteers to correct and mark up text. User interfaces should be assessed with user testing and A/B tests.

5.6.5. Analytics

Having 10,000 titles could make digital-humanities style research possible with a suitable search engine. Knowing what users are doing could give insight into the importance and popularity of books and the topics users are researching.

5.6.6. Analysis

If we end up adding 10,000 titles, each would likely have a handful of errors on every page that need correction. That might amount to 50 million errors. To get them corrected in 20 years, users would have to make 7,000 corrections per day, which may be overly ambitious since we currently get 12,000 to 16,000 visits per day and get only a few corrections. However, low barriers and rewards could improve the number of corrections to hundreds per day, and we could attempt to focus the efforts of key volunteers on key books.

This is a large strategy. Including user interface aspects for book correction, automatic markup and tools, standoff markup support, and adding 10,000 titles, it might take 2 student programmers and one content person per summer for five years as well as significant time from the director, etc. It might be reasonable to hope for the strategy to add 50% to our traffic by the end of the five years and for years afterwards. This is a key strategy but it is unclear that it would pay for itself or

that we have the resources to implement it fully, in a reasonable timeframe, without a grant. In the absence of an external grant, it might make sense to work on this project at a reduced level, especially as it relates to other strategies we are implementing such as changes to the book user interface.

We should apply for a grant to support this project. Initially, we could consider a Digital Humanities Startup grant to fund the OCR improvements and markup piece, though the amount of such grants is small. Alternatively, we could seek funding from a Christian foundation for a larger chunk of the project. This project may also be of interest to CCEL users. We could promote it on the site and ask for volunteers and donations.

5.7. Enable Volunteers

The support of volunteers is a part of the mission of the CCEL. The CCEL has had several hundred people volunteer to help, and while most aren't very serious, a few have contributed person-years of labor. However, the support of volunteers has never been scalable, so most volunteers get driven away through lack of support. There is precedent for crowd-sourced projects that grow very large, e.g. Gutenberg or Wikipedia, though these have a wide base of potential supporters. Some open-source projects also flourish and require very specialized volunteers. So it should be possible to set up a successful and scalable volunteer support system.

The biggest areas of support volunteers have needed in the past are setting forth the vision, help in selecting projects, instruction on how to mark up texts, and correcting and verifying their work--fixing syntax errors, checking corrections, etc. These must be unnecessary or done by other volunteers. We will still need to present a compelling vision for the work and prioritize projects, but these can be done once and made available on the website.

In order to make the process as foolproof as possible, it should be possible for volunteers to make corrections and add markup through the new book user interface, in such a way that syntax errors are prevented. Of course, it will be possible to make semantic errors, so we need a way for volunteers to review recent edits, prioritized by the likelihood that they are incorrect, and accept or reject or correct them.

We should make the proofing and markup of books very easy and without barrier. So, in the normal book interface, selecting a word by double-clicking or clicking and dragging should bring up a popup or sidebar tool that shows the scan of the word and a box in which a user can make a correction. This should not require a login if we can prevent spam. Corrections should be checked automatically for reasonableness and go into a priority queue for review, but the change should take place immediately.

Rewards might include points for corrections and daily, weekly, and monthly leaderboards. Volunteers should also have levels and be able to level up. Getting enough points could reduce the ads a user sees or earn incentives such as ePub or audio book downloads.

The current Wikibooks system uses a Markdown-style syntax, similar in spirit to Wikipedia syntax, but with CCEL additions. The proposal here is for select-and-click style editing. We should run tests to determine which is easier and faster to use and whether we can prevent syntax errors while allowing free markup. Initially, we can perhaps offer both styles of editing, with a "reveal formatting" button that shows markup and enables free-form editing. The current "correct an error on this page" system will be replaced by popup edits.

5.7.1. Analytics

As to data gathered, prioritization of our efforts would be better informed if we knew how many volunteers were contributing and the number of corrections made. There should be daily edit rates to enable us to track changes over time. We should ultimately gather stats on how many corrections are needed to complete a book. We should have a list of top editors available so we can make an effort to encourage and support them.

5.7.2. Analysis

The support for volunteers is an essential part of the 10,000 titles strategy. We already have a couple of ways volunteers can help, through on-line proofing and the Wikibooks system. We can add pop-style editing as we implement the new UI and the 10,000 titles strategy.

5.8. New Content

One of the major limitations on usage of the CCEL is that the books are generally dated 1922 or earlier—91 years old, as of this writing. In fact, most are nineteenth-century. These books often use old-fashioned or archaic language that is more difficult for modern readers, especially young readers, and they can't deal with topics of current interest to society. Even for old classics, such as Augustine's *Confessions*, professors assigning readings will often choose a newer, copyrighted translation, rather than a public domain translation. There is value in the CCEL's mission, but the target audience could expand significantly with modern writings and best translations.

Of course, the work of authors and editors and the publishing industry are crucially important, and they deserve a reward for the work that they do. So what is the role of a free library of classic Christian books on the Web? Does it simply appeal to users who don't want to pay and are willing to settle for inferior versions of books?

The CCEL adds value to the books. In addition to making them easily accessible, the CCEL adds discovery, searching, study tools, and pointers to important books. Users can browse books, read sections, and discuss them. They can search the works of an author or look for mentions of Bible passages in the *Early Church Fathers*. Instructors can link passages for students to read. The proposed term treatment strategy will allow users to find classic writings on various topics. Study tools enable readers to find helpful resources related to passages.

Nevertheless, offering more modern-language literature would enable us to reach more people. In some cases it may be possible to get permission to include up-to-date writings, and we have some examples. In other cases, we can create our own content. Currently we do this for book blurbs, and in the Treated Terms strategy we will be creating study aids for named entities and theological concepts.

We should also treat modern books as first-class items, indexing and promoting them in the same way we do public domain classics, so that users can discover, browse, and ultimately purchase them. It should be possible to do this in a way that adds value for users.

5.8.1. Updating Existing Literature

We could also experiment with using automatic translation techniques to modernize older works, effectively translating from older English to modern English. This may have a low probability of complete success by itself, but it could make some books more accessible.

We could then ask for volunteers to proof and correct the result of language auto-update. Their task would simply be to make sure it is all grammatical, though they could make further corrections. Someone else (perhaps a CCEL editor) could review their work. The editing user interface would have to be different from the general

editing interface—that one is for corrections and markup only; we don't want people changing books in the standard editing interface.

5.8.2. New Literature

We should keep looking for quality unpublished or Web-published works that can be added to the CCEL. One prime example is SIL's Easy English bible commentaries. These would serve our users well, especially overseas users.

Another option for getting up-to-date content is to license it. Licensing for a one-time fee is likely to be too expensive. If a publisher asked \$10k to license a work, we would have to sell about 30k copies at \$10 or generate 10 million extra page views, neither of which is likely for books that are not best-sellers. And even if we could find a licensing arrangement that would work, the effort to develop the arrangement and the software and the diversion of our attention from other areas suggests that this is not a likely avenue to pursue.

5.8.3. Open Publishing

We could also consider publishing previously unpublished works in a more official, open manner. We would list CCEL as the publisher. The author would retain copyright, though the CCEL would have the right to display and sell the work. We could promote the book in a "new books" section, in the newsletter, and/or with internal ads.

An open publication process might work like this: users could submit manuscripts, which go into a queue and are displayed on a manuscript webpage. With each, we would show topic, author, author bio, anonymous (but authenticated) reviews, rating. Manuscripts that reach a certain level would be selected for publication. These would go into the main CCEL library. This approach has some promise at large scale, but it would require a large and committed user base. It would have to be promoted heavily at academic conferences, etc.

Perhaps it would be better to start small, with notes rather than monograph volumes. Users could contribute sentences or paragraphs on various topics—named entities, scripture passages, and the like. They could submit literature reviews and recommendations on topics or books. These would be displayed in the semantic tools section with new and popular ones at the top of the list, attributed to the author by name. Users would click "like" on notes they like. We could give feedback on a user's profile page, detailing the number and types of contributions and the number of times they were read.

5.8.4. Modern Classics

Users searching or browsing at the CCEL are usually looking for books or authors or searching by topic or scripture passage. It would be helpful to them if information were also available about copyrighted important Christian books. Ideally, we would

digitize such books so that we could search and show excerpts a la Google Books. The ability to link citations to such books would also be helpful. In term treatments, references to such books should be linked.

For key authors and key books, we would create full treatments like those we create for books we do have onsite. The book page should offer look inside and search this book options, and the author page should offer “search the works of this author.” In both cases there should be affiliate links for purchasing books.

To handle copyright issues, we would need to set a safe limit for fair use: for example, at most 10 pages or 10%, whichever is less, can be shown to each IP address in a given six-month period. It may be wise to consult a lawyer to set limits.

Offering a full treatment for all the books of a key author such as CS Lewis would be expensive. We wouldn’t be able to use crowd-source proofing, so it is likely that these books would not be proofread. However, we could try to scan three different editions to minimize OCR errors.

If the 10,000 titles strategy is implemented, it should be easy to add books from page scans. The cost of acquiring, scanning, and adding three editions of each of CS Lewis’ works and creating full treatments and literature surveys might be \$100 per book, for about 40 books—perhaps \$4,000. This collection would need to generate 1,000 page views per day for 4 years to pay back through advertising alone. That seems very plausible. And there would be other revenue opportunities as well, including affiliate fees for sales or advertising fees for listing books.

To assess this strategy we would need to track the cost of adding and treating an author’s books and the advertising and affiliate revenue generated.

5.8.5. Full Treatments for Named Entities

The most likely way of getting modern content online would be to implement the Treated Terms strategy. In that strategy, we create “full-treatment” pages for hundreds or thousands of terms that users search for, along with pointers to relevant literature for those terms.

We may write some of the full treatments, but we can also get a great deal of information from the resources on the semantic web such as DBPedia. These resources, and the treatments we write, will be in modern English, and they will be an important component of the user experience.

5.8.6. Analysis

Adding new content is an important strategy for broadening the appeal of the CCEL, but it is a difficult problem. The automatic/human language update approach seems plausible, especially if much of the development cost is covered by a grant. The small-scale open publishing, in which users contribute small tidbits, may result in

some new content, but it is unlikely that it would result in a significant amount for each of the thousands of terms, scripture passages, and books we hope to cover.

However, the adding select modern resources such as the SIL's Basic English Commentaries will help, and the full treatment pages for named entities will go a fair way toward changing the feel of the site. Allowing user contributions for reading suggestions, etc. will also help.

5.9. Pithy Quotes

The idea in this strategy is to build a system to gather “pithy quotes”—quotes of a couple of sentences or a short paragraph, suitable for devotional reading. We could have a pithy quote widget on the home page, or somewhere else, with a quote for the day. Users could vote it up or down, with new or popular quotes more likely to appear. There should be a button for “another quote”.

Since we want users to see the best quotes, we can’t simply show one quote for everyone. Instead, we should show random quotes biased by quality/newness. For logged in users we should show the highest-rated quote that they haven’t yet seen.

We could enable people to submit a pithy quote by selecting it. The popup tools window would have a “submit as pithy quote” link. We could show the name of the person submitting the quote, if the appropriate box is checked in their user preferences.

This project would not require a large amount of extra programming if we are implementing editing with markup capabilities in the book interface anyway, and it would add a nice dynamic feature to the front page. It would also give us some quotes with devotional value that could be used for newsletters, daily devotionals, and the like.

5.10. Expand Formats

Amazon.com reports that they now sell more eBooks than print books, and our sales of books in Kindle and iBook format are doing very well, providing nearly half of our self-generated income. Some titles do better than others, depending largely on the competition and the reviews they receive.

One significant risk for our sales of eBooks is that they are public domain. This means that fly-by-night publishers can create low-quality editions quickly and sell them very cheaply or give them away free. The result is that the Kindle and iBook stores are hesitant to accept public domain books. The Kindle store accepts them only if they are differentiated somehow, and Lulu has informed us that they will no longer accept new ones. The upside here is that competition won't increase.

One of our strengths is that we have programming capabilities. We could differentiate our books by adding capabilities of the online library to the extent that it is possible in Kindle or iBook format. It would be nice to be able to read a book on an iPad and access semantic aids, make corrections, add markup, and the like. Currently, much of this is possible in iBook format, but Kindle format does not support programmability. This is an area we should continue to explore and develop. This would also offer a significant additional differentiation factor.

Our differentiation factors would then include cover art, an introduction, an index of scripture references, and possibly semantic tools or a column of named entities with links to helps. If we can improve our selection of cover images, that would also be helpful. Some of these added features could also have links back to the CCEL.

We want to develop a reputation for offering quality editions. Therefore we should carefully check them before release, to make sure they are functional, accurate, and attractive.

Our pricing scheme is usually uniform--\$2.99 per volume. However, a couple of years ago we made an effort to optimize prices on the Kindle store based on competition. With \$75,000 in eBook revenue, pricing adjustments that increase revenue by 20% would be worth \$15,000. Optimization may be worth some effort. We could perhaps write a program that optimizes prices using an algorithm such as simulated annealing. It would make semi-random variations, converging on an optimum. Then too, occasionally offering sales on books or even offering them for free could boost their popularity and increase sales when the sale ends.

5.11. Revitalize Subscriptions

Subscriptions at one point generated fairly substantial income, but the financial model has changed, so that last year they generated only \$7,000 in income. They also have a cost, in maintenance, management, tech support, redevelopment for new versions of Drupal, and the like. And revitalizing the subscription system is likely to require major design changes to the site and significant time from CCEL leadership.

Unless we make major changes to the site to add significant value that could be offered by subscription, it is unlikely we will want to maintain our subscription option. We may want to drop the subscription system the next time the site is redesigned. The one possible reason for keeping subscriptions is to give users the option of turning off ads. If we offer more prominent ads, regular users might want to turn them off for \$15 a year or some such.

5.12. Improve Community

Community is a fundamental part of the Christian life. Community makes possible the love and hate, the learning and growth, that give value to life. Nevertheless, not all of the Christian life is community. We worship in community but we pray in the closet. We minister in public but we prepare in the desert.

Reading books seems to be an inherently solitary activity, if for no other reason than the single-user nature of the technology. Reading also seems commonly to be associated with an inward focus. However, that is not exclusively the case. Classrooms and book study groups are places where we come together before or after solitary reading. There we learn from, encourage, and inspire each other.

CCEL's attempts to provide a community component to this point have been limited to discussion forums and on-line study groups, which function much like discussion forums. It may be worth thinking about the purposes of community in the context of book reading and the ways in which the CCEL could support these purposes.

In classrooms and book study groups, we encourage each other to read an assignment and provide some level of accountability. We are introduced to readings we wouldn't otherwise have discovered. We answer questions and help each other to understand and apply the reading. We point out flaws or hidden assumptions in readings. We learn to be appropriately critical. And the pleasure of coming together, discussing, and being a part of a community encourages further reading.

Many of these benefits of community could also help the CCEL address its mission. However, we should think about embedding the community aspects at appropriate places in the structure of the CCEL, not add it on in the form of discussion forums.

5.12.1. Shared Annotation

Discussion of books and learning from others should happen while we are reading, within the book, not in an external discussion group later on. It should be possible to make comments and annotations on a passage for personal use or to be shared. It should be possible to add comments to those annotations. It should be possible for a user to email a link to a passage with his or her annotations showing. Trust is an issue, so it should be possible to read about the comment author and select the comments that will be displayed. Of course, it should also be possible to turn off annotations entirely or see only your own.

As a user interface, when annotations are turned on, they should be visible in some way. It may be possible to use hover events on the desktop, but that option is not available on mobile devices, so it may be better to use another approach. One possibility is small icons embedded in the text. Clicking or hovering should bring up a nicely-designed popup containing the picture of the person making the note and the note itself, with a like button and a link for comments on the note. The annotation interface should have a share button to make the comment public.

5.12.2. Discovery

We can support the communal aspect of discovery by allowing users to add expert notes to books and named entity treatments, recommending books for reading for particular topics. Users could submit links to sections of books with annotations. We could also solicit users to contribute small “what we’re reading” notes and share them on the front page or the read page.

5.12.3. Encouragement and Accountability

Our study groups system tried to provide a structure for users to encourage each other to read and to hold organized discussions at particular times. How could social capabilities addressing the same issues be implemented at a more fundamental level in the library?

The shared annotation system would be part of the solution. Annotations would be shared among the members of a defined study group. For the assignment, accountability, and time-based aspect, we could enable users to create reading plans that could be shared with the group. Reports of who has completed the reading could be made available at scheduled times. We would still need some means of holding the time-based discussion, such as a chat room or forum.

This would be a fairly large development project for an area in which we have not had much success in the past, an area in which I am not aware of good examples on the Web. This may not be an area we should implement.

5.12.4. Community and Discussion Groups

The discussion groups seem to fulfill a need, so we should keep them around without investing a large amount of effort. Perhaps we can use better forum software such as PhpBB and improve the appearance with CSS3, using the speech bubbles of the Stunning CSS3 book. If we change the forum and group software, we should transfer the old posts.

Discussion forums need oversight, encouragement, and moderation. It would probably be wise to assign some student hours to supporting the forums. Discussion forums and groups seem to be generating about 50k page views per week, so it may be plausible to spend on the order of \$100 per week to support and promote the forums.

Currently we are able to have discussion groups with closed membership. We should research whether that is a feature we need to maintain and whether it is supported by PhpBB or other modern forum software.

In order to promote a sense of community, it may be good to having a single community page, replacing the Give page. It would have information about the forums, the online book editing project, and recent public annotations, as well as

popular books, hot forum topics, trending searches, and the like. We could also have leader charts for contributors here. We could also add some statistics about CCEL usage. Some of this information could go on the front page as well to give it more of a dynamic, community-oriented feel. Perhaps we could add a kind of feed or ticker within the community icon.

5.12.5. Analytics

In order to track discussion forums, we should know the number of posts and replies written and read each day, by forum. Perhaps these could be displayed on the forum pages themselves as well as in a monitor page. We should also have statistics on shared/private annotations.

5.12.6. Assessment

5.12.7. Analysis

Community features of the site play an important role, but it is a secondary role to discovery and reading and should be prioritized correspondingly. The shared annotation is important for other reasons—supporting educational use, for example—and it is a fundamental part of the annotation/markup strategy.

The community page replacing the give page is a good idea, and it addresses one of the board member's suggestions that the home page have more dynamic features. The discussion groups are not very attractive, and at least some effort should be put into improving the looks. We should use new bulletin board software if it can be done without too much difficulty.

Adding a facility for online book studies is probably the lowest priority of all the community options.

5.13. Distribute eBooks worldwide

One of the early motivations for the CCEL was to make classic Christian books available, and people in developing countries seemed to have the least access to such literature and the most need. The CCEL distributed over 7,000 CDs along with permission to make additional copies.

The demand for CDs has dropped precipitously in the US because of the rise of high-bandwidth Internet access. Now it is easier to search for and view books on the Web. That option is less readily available in developing countries, but educational institutions usually have Internet access of some sort.

If we were to continue seeking to distribute classic Christian books in the developing world, perhaps doing so on a Kindle would be best—they can store hundreds of books and can run for weeks on a single charge. They do not require the use of a computer, and at \$80 or so for the cheapest ones, they are not exorbitantly priced. This might be an area for a foundation grant proposal. If we do prepare a proposal, we could allocate some to licensing relatively current content to fill out the gaps in our library.

However, the fact that we have no contact with theology students, ministers, etc. in developing countries and we really don't know what they need works against this idea. We have no way to assess the effectiveness of such a program. This may be better done by a group that has contact with potential recipients.

5.14. Grants

We may need grant funding to implement all of these strategies in a timely fashion. Projects that may be suitable for grant applications:

- 10,000 titles with OCR correction and crowd-sourced markup: Digging into Data? NEH digital humanities? Mellon?
- Semantic tools: as part of Lilly Preaching Portal project proposal?
- Classics on the big questions: Templeton? A private foundation?
- ABS, theological libraries on Kindles for the world
- Luce foundation, classics in 20 languages

5.15 Social Media

Some sites are starting to spend a significant portion of their promotion budget (e.g. 15%) on social media sites such as Facebook. We don't have a promotion budget *per se*, but some of our activities are essentially targeted at promotion. We need to determine whether Facebook can be an effective way of promoting the CCEL. We could

- Put a Facebook Like button on the home page and elsewhere
- Put a Like button on pithy quotes or other features, to have them included in a newsfeed
- Make occasional posts with paid promotion for new term treatments, books, and the like

5.15 Twenty Languages

In this strategy we would use Google Translate to translate a few key books into many languages. We might start with five books every Christian should read. At the bottom of the book we could add a note to the effect that the book was automatically translated and needs cleanup, asking for volunteers.

Google Translate supports more than thirty languages. We might choose languages where there are lots of readers but few classic Christian books already in the language—Chinese, Korean, Vietnamese, Russian, etc. Google offers a Translator Toolkit page to support translation; we could offer volunteers permission to work on the translation in question.

The results of Google Translate are very rough—enough to get the sense of a passage, with some effort, but not pleasant to read. We would then solicit volunteers to clean up these translations. Perhaps we could find immigrant churches to get volunteers.

Eventually, it would be nice to select a larger number of books, including books that are useful in seminary education. That part of the idea might be fundable. The Theology program at the Luce foundation seems like a candidate for funding. However, the major work would be in soliciting, training, and managing volunteer translators, and we have little track record in doing so. We'd need some proof of success before getting funding.

5.16.1 Evaluation

Setting up CCEL pages in several languages would be a major new direction for the CCEL, aligning well with the mission. The strategy would require setting up interfaces in various languages, but it would not require much programming. However, it would require considerable personal effort to solicit, train, and manage volunteers. There is little potential for added revenue except through grants.

It may be worth finding volunteers to translate some books into a language to test the approach. Then we could consider writing a grant proposal, perhaps to the Theology program at the Luce foundation.

5.17. Copyrighted Books

In this strategy we would add support for copyrighted books to the CCEL. Two recent court decisions have (Author's Guild vs. Hathitrust, Author's Guild vs. Google Books) have supported the finding that scanning copyrighted books and offering snippets on the web is fair use. We could scan some more modern classics and have pages for the books at the CCEL. Users could search the books and view snippets. Requests for additional pages would state the copyright status and offer to allow users to purchase an ePub version of the book. Once published, that user would be able to view the whole book online and download the ePub.

This strategy would probably require the CCEL to become an ePub reseller. We already sell our own ePubs, but we would have to make arrangements with various publishers, which might be time-consuming. We would also have to pay royalties. It would also require adding support to the CCEL for copyrighted books, allowing searching and viewing of snippets.

This strategy addresses the problem of not having books using modern language. We would have such books, at least for snippets and possibly the first chapter. It might be possible to have links to sections that would show the relevant part of the book. For example, if term treatments list key literature, and some of that literature is copyrighted, we could put a version of the book online and link to it.

Another benefit is that this strategy might partially address the issue of not having any modern commentaries. We could put a modern commentary online in this form, and when a user seeks commentary on a passage, we could offer this copyrighted

commentary, at least for small passages. We could do the same thing with Bible Dictionaries or Encyclopedias, perhaps offering a single article, or a part of an article, to users.

5.17.1 Evaluation

This is one of the few strategies we have that addresses the problem of adding recent literature and literature in modern English. If we were to make it clear in book lists that such books are copyrighted and only available in snippets, their presence should be an advantage for CCEL users, as long as it's still easy to find public domain books.

Adding support wouldn't be a huge job—perhaps a student summer. There would be an additional cost for each book we add. If we limit ourselves to books available in ePub format, adding the book might be easy. However, adding markup for commentaries could be time consuming.

The recent court cases make a nice precedent, but we might still want to have an attorney review our implementation and give an opinion on copyright risk.

Becoming an ePub reseller could add an additional revenue stream, but I find it unlikely that it would be very significant. However, it might cover the costs of adding books and make the CCEL more valuable for users and more comprehensive.

This strategy would require time and investment, and it would help us meet our mission, but it seems unlikely compete favorably with other high priorities unless it results in another revenue source such as a partnership with a publisher, as described in the strategy below.

5.18. Publisher Partnership

In this strategy, we would form a partnership with a publisher of Christian books, perhaps Baker Bookhouse. We would add some or all of the publisher's books to the CCEL, as copyrighted books, with support as described in the previous strategy. We would enable users to search these books, view snippets online, and purchase them in ePub format. The publisher would get more visibility for the books and hopefully more sales.

In return, the publisher would purchase advertising at the CCEL at market rates. Perhaps we could replace the Calvin seal with the publisher's logo and a message that the CCEL is sponsored by the publisher, which would appear on every page. Additionally, the publisher could agree to purchase some number of impressions for new books they wish to promote. With 20-30 million ad impressions, the advertising revenue could be \$40-80k. If the advertising is at market rates, the

publisher could fund the cost by shifting some of their advertising budget. The CCEL is an ideal place for book publishers to advertise.

This strategy seems viable and beneficial for both parties, and the potential revenue warrants considerable effort.