

# **THE COLLABORATIVE AUTHORING TOOL: A VISION FOR OPEN TEXTBOOKS; A PLAN TO ACCOMPLISH IT**

ROBERT BEEZER, STEVE BLOOD, DAVID FARMER, AND SALLY KOUTSOLIOTAS

**ABSTRACT.** The current approach to creating, maintaining, and distributing open textbooks has several limitations that hamper the development of open textbooks and prevent a broader adoption of these resources. We propose a web-based collaborative authoring tool which addresses these issues.

The open textbook movement is hampered by a lack of effective tools. Key activities, which should be straightforward, are cumbersome in current practice. Examples include:

## **ISSUE 1. *Enabling community input***

Open textbooks will not thrive without community input. At present there is no easy way for multiple authors to collaborate on an open textbook, nor any simple, low-maintenance mechanism for users to provide feedback or errata. As a result, it tends to be a significant email burden on the author to manage corrections, comments, and feedback. This is due to the fact that it is not possible to directly tie feedback to the location in the book which requires attention.

## **ISSUE 2. *Separating content from presentation***

Currently an open textbook author must handle the entire process of textbook production. Ideally, the author should focus on the scholarly content, leaving the technical, logistical, and design aspects of production to be handled automatically.

## **ISSUE 3. *Supporting true modularity***

Current open textbooks meet the definition of “open” in the sense that the author has made the source available and someone else could take the source and modify it. Typically, the source files have been created without an overall plan, growing as the author writes the book. This makes the files cumbersome to navigate, particularly to the uninitiated. As such, this openness is of limited value, due simply to the difficulty in managing a large collection of raw source files. It can be even more difficult to manage a small number of large files, particularly if the author has not carefully planned for the underlying structure of the document.

## **ISSUE 4. *Providing customized compilations***

Even the best textbooks can benefit from some tweaking to meet the instructor’s needs for a particular group of students in a specific course. Customized editions (or multiple “views”) of a book, which are based on the primary version but which automatically inherit improvements to the primary version, are inherently better than multiple independent versions of the book which must to be maintained separately.

Some of these issues have been successfully addressed for code development of open source software projects. We propose a solution which makes use of ideas that have been used for software development, but which is specifically tailored for open documents.

## 1. THE COLLABORATIVE AUTHORIZING TOOL

### Envisioning the CAT

We seek to develop a general platform, called the *Collaborative Authoring Tool (CAT)*, which supports the creation and maintenance of a wide variety of scholarly documents. This tool embodies two fundamental principles which are missing from current approaches to open textbooks:

#### *The Authority Principle.*

Each document supports a hierarchy of editorial authority. While any reader can suggest additions or changes to the document, only privileged users are granted wider latitude in their ability to modify the document. However, none of these changes appears in the public version without final approval by an editor.

#### *The Structure Principle.*

Each document has a clearly defined and rigid structure, with each component created and edited separately. This formal structure allows the layout and the conversion to multiple formats, such as HTML, print-on-demand, or e-book, to be handled automatically, with little or no input required from the authors.

The Authority Principle addresses issue 1: *Enabling community input*, while the Structure Principle addresses issues 2, 3, and 4, concerning the structural foundation underpinning the document.

### Progress towards the CAT

Currently, these principles are being tested with the *AIM Problem List (AimPL)* tool, available at <http://aimpl.org/>. Initiated at the American Institute of Mathematics (AIM), the AimPL is specifically designed for creating and maintaining the lists of open problems which drive research mathematics. The AimPL represents a special case of the CAT: our goal is to generalize this tool, guided by the experience and knowledge gained through the development of the AimPL. The CAT will function similarly to the AimPL but will be able to handle more general documents.

The success of our initial testing of the AimPL indicates that we have a viable model for implementing both the Authority Principle and the Structure Principle. The shortcoming of AimPL lies in its ability to handle only the problem lists, which have a very specific structure. As planned, the CAT will handle a wide variety of documents, and potentially all scholarly documents.

From the beginning, we regarded the creation of AimPL as the first step in our learning process, one that now enables us to develop a more general framework capable of handling other collaboratively authored documents, including open textbooks. It is this more general program, the CAT, for which we seek funding.

## 2. PLAN FOR CREATING THE CAT

Our plan for developing the CAT is structured around a series of workshops.

Since the AimPL provides an example of the CAT for one particular type of document, the first major hurdle is developing a method for describing the other types of documents that will be served by the CAT. The majority of this initial work will be accomplished through a workshop.

### Part 1: Development of the CAT.

#### Workshop 1: *The Structure of Scientific Documents*

This workshop will examine a variety of documents which will be served by the CAT, with a goal of specifying the schema which will be used to describe the structure of those documents. Once this is established, the design of the CAT software can begin. The workshop will bring together open textbook authors, software developers, people with experience in textbook production (from editors to book designers), and others involved in the open textbook movement.

The 5-day workshop will involve around 30 people. Very few talks will be planned, leaving most of the time for focused work on specific tasks, in large or small groups. AIM has extensive experience hosting such workshops, guiding a large group of people to work together to accomplish a specific goal.

#### PHASE I: *Coding*

Following the workshop, the project leaders will work with programmers to outline the CAT software and coding will begin.

The first benchmark for this phase is producing a new version of the AimPL code which handles the user interface as well as the input and output from the database, but which does not refer to the structure of the document. This code will be written from scratch, but will be informed by the AimPL software.

The second benchmark for this phase is a beta version of the parser which takes the schema describing the structure of a document and sets up the database for storing the document, sets up internal representation of the structure of the document, and encodes how the user can interact with the various components.

At this point the main capabilities of the CAT will be in place and testing may begin.

#### PHASE II: *Initial testing and further coding*

The project leaders will test the software by creating schemas for various documents, ranging from the simple (a glossary, or an annotated bibliography), to the more elaborate (a problem list, or a survey article), to the highly complex (a textbook). Ongoing feedback to the programmers will improve and refine the specification of the schema and the ability of the code to implement the schema.

Once the software is able to handle a variety of documents, attention can turn to the user interface. The project leaders will begin importing existing documents into the CAT. Feedback to the programmers will be used to improve the usability of the program.

The result of this phase will be a version of the software which is ready for beta testing.

**Workshop 2:** *Author Interface Beta Testing*

The focus of this workshop will be on the process of creating documents using the CAT. The workshop will bring together the project leaders, authors of open textbooks, and programmers to test and refine the CAT.

Workshop participants will import existing open documents into the CAT. This will be the first step in giving a new life to those documents, and also to ensuring that the CAT software is easy to use and meets the needs and expectations of a variety of users. The combined efforts of the project leaders, authors, programmers, and other workshop participants will ensure that the key issues are identified and prioritized.

Revising and maintaining documents will be the focus of testing after the workshop.

PHASE III: *Global Beta Testing and Further Coding*

Continued development of the CAT software will bring the code to the point where open documents can be released on the web, available for use by the wider community. The documents incorporated into the CAT at Workshop 2 will serve as the material to beta test this code. Those will become the first documents which are truly open and which support the Authority Principle and the Structure Principle, beyond the problem lists served by the AimPL.

Finally, outstanding issues relating to the authoring interface (arising during and after Workshop 2) will be addressed during this phase.

At this point the CAT will be ready for its official release. This will enable a multitude of open textbooks and other documents to be created, used, and maintained by a broad spectrum of the academic community.

**Part 2: An Explosion of Open Documents.**

Eventually, sufficiently many templates of open textbooks will be developed and in place that a new author can easily mix-and-match existing features to create an appropriate layout for their planned book. Before that occurs, we envision a series of workshops to encourage the development of a critical mass of open textbooks and also serving to expand the CAT to other scholarly areas.

Using a model that parallels the development of the CAT, a plan of continued code development will be implemented during and between the following series of workshops.

**Workshop 3:** *Initiating New Books*

This workshop will bring together approximately 15 groups who have started (and possibly have already completed) a textbook that they would like to release as an open text. At this workshop, the focus will be on topics in mathematics and related disciplines.

**Workshop 4:** *Initiating More New Books*

The CAT has the potential to transform open textbooks in all areas. This workshop will have a similar goal to Workshop 3, but will bring in authors from a broader range of academic disciplines. It may be useful to repeat this workshop several times, to cover a variety of areas.

**Workshop 5: *Improving Existing Material***

There are hundreds of documents on the web which purport to be open textbooks, but which are not sufficiently well developed to be selected as a textbook by most instructors. The goal of this workshop is to identify authors who are willing to transfer their texts into the CAT environment. Once incorporated into the CAT, the process of improving the texts will be greatly simplified.

This is a potential source of a large number of open texts which could be made available quickly. It may be useful to repeat this workshop.

**Workshop 6: *Customized Versions of Books***

A customized “view” of a textbook is a version which consists almost entirely of material from the master copy book, but repackaged for a slightly different purpose or audience. An example is a book in which the material is presented in a different order, or which contains a new or alternate sub-topic. Another example is a “teacher’s edition” of a book.

The goal of this workshop is to implement software that will enable both authors and users to create custom views. The workshop will be structured in two phases: the first will bring together programmers and 2-3 authors who have attempted to create customized views in the past. Tools will be created during this time, allowing a larger set of authors to test the software and provide immediate feedback during the second phase of the workshop.

Implementation of this feature will allow changes to the master copy of the book to be automatically transferred to all the “views” of the book. Views will allow for customized versions of a book, with little maintenance required to keep the customized versions up-to-date. A 10-day workshop of around 15-20 people should be sufficient to design and implement views in the CAT.

At the end of Part 2, we anticipate that the CAT will be self-sustaining as an open-source project, with continuing code development, and will house a rich repository of book templates supporting the ever-increasing availability of open textbooks.

### 3. TIMELINE AND BUDGET

We estimate 18 months to reach the official release at the end of Part 1. Part 2 will require another 9-18 months, with workshops planned every 2 or 3 months.

For Part 1, we estimate the following costs:

Design and management:	\$250k
Programmers:	\$500k
Workshops:	\$100k (2 x \$50k each)
Total (Part 1):	<u>\$850k</u>

For Part 2, we estimate:

Design and management:	\$150k
Programmers:	\$250k
Workshops:	\$200k
Total (Part 2):	<u>\$600k</u>