

SCIENTIFIC WRITING WITH GOOGLE DOCS

Gabriel Wurzer
Technical University of Vienna
Institute for Architectural Sciences
IEMAR: Digital Architecture and Regional Planning
Treitlstrasse 3
1040 Vienna, AUSTRIA

ABSTRACT

Scientific writing is not about getting all words right in the first place; rather, quality comes through repeated revising of a paper. Google Docs is a collaborative text processor that can help collaborating scientists in producing papers faster and without the usual hassle of data exchange and/or different word processors.

INTRODUCTION

If you have a golden hammer, everything looks like a nail.
(Golden Hammer AntiPattern)

There is no single tool that will satisfy all your needs. This is especially true in the Web2.0 arena[\[6\]](#), in which Google Docs has its place. Here, every application is termed a *service* (i.e., by using Google Docs you are using a text processing service). Other services you might be interested in could for example be Wikipedia, the free encyclopedia[\[7\]](#) and google's scholar search[\[8\]](#).

RELATED WORK

Little work has been done to show that Google Docs is capable of handling scientific documents, i.e. documents that contain outputs from research work. A comparison of Google Docs to conventional word processing [\[1\]](#) shows that collaborative writing is direct and instantly viewed by all. Furthermore, peer reviewing and peer revisions are now possible. The literature also mentions that extra care needs to be taken so that the produced paper does not feel like it was written by a committee. That collaborative editing improves work output is yet to be shown. Some work towards this goal was given in [\[2\]](#).

GOOGLE DOCS' BUILT-IN TOOLS

Text formatting

As text editor, Google Docs offers little more than a conventional rich-text editor (see Figure [1](#)): You can format text bold, italic or underlined, change font face, text colour and highlight color. Text may be made super- or subscript. Sadly, there is no symbol table with mathematical notation and Greek symbols. Therefore, formulae are best produced using an external editor and saving the results as image (see "[Inserting objects](#)").

For paragraphs, you can choose between a set of predefined styles (heading, citation, standard text...). Their appearance is fixed, i.e. there is no way in which one could change it. What is painfully missing is a style for numbered headings (although, there is the possibility of numbered lists that can have either roman, arabic or alphabetic enumeration). There is also no caption style (text caption, image caption or table caption). The text spacing is definable (1 lines, 1.5 lines, or more), as are the page breaks. Bullet lists are included as well, however, the bullet symbol is not configurable.

Page-specific styles such as header, footer, footnotes and side-notes are not supported. Page margins are also not specifiable, as are multiple text columns per page. While one could use tables to emulate each of these features, the author strongly discourages this: For layout (as this is), another program should be used that can import rich-text (see).

Normal Paragraph style, with **bold**, *italic* and underlined text. Text in different font face. Text with different color.

1. Numbered List (Arabic)
2. Numbered List (Arabic)
 - i. Numbered List (Roman)
 - ii. Numbered List (Roman)
 - a. Numbered List (Latin)
 - b. Numbered List (Latin)
 - Bullet List
 - Bullet List

Header 1

Header 2

Header 3

Figure 1: Text formatting

Last but not least, spell checking is available for a huge number of languages. Quite similar to other text processors, there is a user dictionary which can take up non-standard vocabulary that scientific documentation often has.

Inserting objects

Inserting images into the document makes them become encapsulated. This means, there is no way to include a graphic by reference (so that changes to the graphic are also reflected in the document). Supported image file types are jpg, gif and png (raster image formats). Therefore, you must convert all images before being able to include them. There is a 2 Mb limit on image file size, which will generally be no problem if the images you use are not page-filling. Images can be scaled proportionally by pressing the shift key while scaling.

Google Docs supports bookmarks that can be used for linking to a specific paragraph in

the document. This feature could be used to create a simple bibliographic index (see Figure 2), however, it is not encouraged. The reason is that changing the bibliography style (e.g. into 'Cohen 1990: "Multivariate Data, Analysis and Impact") would require you to re-type every entry. There are better ways of producing a bibliographic index in Google Docs, given in "[Adding Bibliographic Data](#)". Figure- and table-numbers are only producible manually, which might be a pain (especially if Figures are moved from one part of the document to the other).

As Cohen [1] sees it, multivariate data may not only be...

BIBLIOGRAPHIC INDEX

[1] Cohen, R. 1990: "Multivariate Data, Analysis and Impact"

Figure 2: Using bookmarks to create a simple bibliographic index

Collaboration

There is unique feature that compensates for all shortcomings of Google Docs: The possibility to collaborate while writing a paper, i.e. to work at the same time on the same document. Because work can happen at the same instant in time, there is no need to wait for the latest paper version to be submitted: users can participate in the work when they see fit. Second, because the usual write/review/correct cycle is no longer valid, "peer reviews" can happen at any time, even while the paper is edited. This is similar to software development methods such as pair programming, where writing is usually done *in pairs*: While one developer works on the actual code, the second one focuses on reviewing. As a matter of fact, the quality of the produced software rises.

Collaborators can be invited as editors or read-only viewers, given that they own a gmail account. This account can only be granted by someone who already owns one, i.e. it is a system of invitation. There is an RSS feed which informs all collaborators of the most recent changes to a document. This way, everyone can stay up to date without actually having to log into google docs.

When collaborating, users can type text directly into the document or use the comment tool to add review comments (see Figure 3). These comments will not be output when publishing the document. As added gimmick, Google Docs keeps track of all revisions made by each collaborator. This way, changes can be rolled back in case of mistakes. Also, different revisions of the document can be compared (e.g., to support questions like: "Which version sounded better?").

A data item which is time-dependent may be written as

$$d = f(t)$$

indent this for better readability -John 1/4/07, 9:08am

Figure 3: Using the comment tool for collaboration

The better alternative is to use a citation software that allows to insert placeholders for citations and produces the list of citations automatically. See "Using a Citation Program".

EXTERNAL TOOLS

Adding Bibliographic Data

Storing Bibliographic data externally (i.e. using a bibliography software) enables the researcher to enter your list of references once, and re-use this data among many publications. The output style with which the citations are written into the Google Docs document must be choosable freely (e.g. [4] or (Duncan and Mulholland, 1995)) without requiring further work. Citavi[3] is a free reference management software that is proven to be a good solution. Sadly, there is only a german version at the moment. EndNote[4] has similar (or supposedly even greater) functionality, but it is not free.

The process for adding bibliographic data is

1. (Enter the references into the reference management of your choice.)
2. For every title you want to cite, insert the corresponding placeholder string into the Google Docs document. An example for such a placeholder would be {Baeza-Yates 1997 #3}. The reference management software will usually allow to copy the placeholder string for a chosen title to clipboard.
3. With all the placeholders entered into the document, export it as .rtf file.
4. Feed the .rtf file into the reference management software. Out will come a .rtf file with all the placeholders replaced by citations (in the style chosen) and a trailing list of citations.

Using a Layout Software for the Layout

A common misconception about text processing software is that it should cover two aspects at once: *text processing* and *layout*. The bare truth is that those two aspects are, if not completely seperatable, then at least hard to bring together under one hood.

- text processing covers all semantic aspects of text editing, i.e. producing a text that is annotated with a *meaning* (e.g. heading, bullet item, ...)
- layout refers to the visual representation of these semantic aspects.

By using a layout software to do the visual representation, much of the shortcomings mentioned earlier can be compensated. The layout software used needs to be able to place .rtf files. InDesign[5], for example, has exactly this capability. The following procedure is capable of producing photo-ready papers:

1. (Make a new layout document for the conference style. Page borders, paragraph style etc. should be adapted to the sepcific visual needs of the conference.)
2. Place the Google Docs .rtf file which already contains the references and list of citations.

3. Check that each paragraph format of the layout document was correctly applied to the placed text.
4. Output the finished layouted document as .pdf

CONCLUSIONS

Google Docs is good for collaborative text processing, but not for layout and citation management. For these tasks, as well as any other functionality that is not readily available in Google Docs, a workaround must be made.

- Formulae and Graphs can be imported as images
- Citations can be inserted in a post-step, using a citation reference management software
- Layout should be handled in a similar fashion - using a layout program to specify the layout of paragraphs, bullet lists etc.

In this pragmatic fashion, Google Docs can become a valuable tool for the scientific community - with no freedom having to be sacrificed for the love of a tool.

BIBLIOGRAPHIC INDEX

- [1] "Comparison of Using Google Docs to Conventional Wordprocessing", http://docs.google.com/View?docID=ah7pzmd9d7gs_bacpff3h9vrjg&revision=_latest, last viewed 4.1.2007
- [2] Baeza-Yates, Ricardo; Pino, Jose A.: "A First Step to Formally Evaluate Collaborative Work". In Proceedings of the international ACM SIGGROUP conference on Supporting group work 1997
- [3] Citavi Reference Management Software. Academic Software Zurich. <http://www.citavi.com/de/allgemeines/english.html>
- [4] EndNote. ResearchSoft. <http://www.endnote.com>
- [5] InDesign. Adobe. <http://www.adobe.com/de/products/indesign>
- [6] O'Reilly, T.: "What is Web 2.0", <http://www.oreilynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html>
- [7] Wikipedia Encyclopedia. <http://www.wikipedia.org>
- [8] Google Scholar. <http://scholar.google.com>