10 Basic Requirements for Text Documents

François Briatte
April 2, 2018

This document is part of the "Computing Tips and Tricks for Students" series.

All requirements below are personal guidelines that apply to producing ***printable text documents*** on a computer. Your mileage may vary, and do not assume that I never fail those guidelines myself.

This document says nothing about the software that you should be using to produce your text documents, but I will assume that you are using either plain text or a word processor such as Google Docs, LibreOffice Writer or Microsoft Word.

1. Your name is on it.

There are special cases, of course, but few: more often than not, the first thing (or second thing, after your title) that should show at the beginning of your document is your full name.

Every year, I pledge never to grade papers sent by students who forgot to indicate their names on their work. I generally relax that pledge on draft papers, but then enforce it on final papers.

Note **– if you convert your document to PDF (Req. #5), your name might end up in the metadata.** Remember this if you need your document to be anonymised, and learn to modify PDF metadata.

2. The document uses headers.

You will almost always need to add headers and footers to your text, in order to include:

- **The production date,** possibly accompanied by a semantic version number if you are writing the kind of document that requires keeping track of several versions. Note that all major word processors also include some form of version control.
- **Page numbers.** Page numbering can be tricky because cover pages – and other things like chapter-specific numbering – might make 'page 4' of a computer file a different thing from 'page 4' of the actual document. Avoid that confusion.
- **A link to the online version of the document** if it is intended to be read mostly on screen. When the link to the document is long and complex (e.g. Google Docs), 'short-link' services like frama.link can produce shorter, memorisable links.

3. The document uses customised stylesheets.

Every single text processor that I know of fails to implement basic readability rules, such as:

- Line spacing should be 115 % or higher
- Default text size should be near 11 pt serif
- Default paragraphs should be followed by a decent nonzero amount of spacing
- Titles are never justified, even when the text is

As a consequence, you will *always* need to use stylesheets to format your text.

Those are my own readability rules, and I actually prefer sans-serif over serif. To me, the most important aspect of formatting is *spacing*, which is very, very poorly handled in many of the text documents that I read:

- For default paragraphs, I typically use 50 % of the text size as the amount of nonzero spacing to be inserted after each paragraph (e.g. 5.5 pt if the line spacing is set to 11 pt).
- For titles, I typically use spacing values equal to 100 % of the title size before the title, and 50 % of the title size after it (e.g. 9 pt before and 14 pt if the title is 18 pt).

4. The margins are realistic.

Page margins are important to keep the average number of characters per line of your document readable by humans. They are also important for printing, since most printers will not be able to handle null margins.

As a rule of thumb, when using A4 paper, *keep using the default margins* set by your word processor.

As a word of advice, do not believe for a second that the extra lines in your cover letter are 'so essential' that they would justify torturing your reader with – single-lined, non-spaced paragraphs that extend to – 1 cm margins. You are fooling yourself.

5. The document is in PDF format.

If your document is meant to be read without any modifications, which is more often the case than not, then you will need to 'print' your paper to a Portable Document Format – which is what the 'PDF' acronym means.

Adobe PDF is a standard format that can be read on all platforms. Microsoft Word – DOC and DOCX – formats are not, and while Rich Text Format (RTF) is cross-platform, it unfortunately suffers from some the same compatibility issues as Word files.

If you are addressing an editable document to an audience that wants to read, not modify, your document, you are making a serious mistake (Req. #9).

If your 21st-century operating system or word processor cannot natively produce PDF files, change your operating system or word processor.

6. The document is backed up somewhere.

Computers and humans are unreliable information storage devices that will periodically destroy valuable information by accident.

Backup your work regularly, or use some 'cloud'-based like Google Docs to continuously save it.

One good reason to use semantic versioning (mentioned in Req. #2) is that it forces you to think of your document going through 'editing rounds' that justify some form of preservation while you are writing. It will also drive you away from naming your files inconsistently, using words like 'new' or 'updated' that your future-you is likely not be able to make sense of.

7. Exactly 0 % of the document is plagiarised.

Plagiarism suppresses the learning benefits that come with writing in your own words.

Expect any learning-oriented institution (such as academia) to expect 0 % plagiarism and to sanction plagiarism harshly, both out of moral considerations for the plagiarised authors and for professional reasons (plagiarising has a quasi-null learning value).

Self-plagiarism and inability to cite and reference sources properly (Req. #8) are variants of plagiarism.

8. Exactly 100 % of our sources are cited.

In academia, we cite our sources using specific styles like the Harvard bibliographic style or some variant of it. If you are writing for an academic audience, learn a bibliographic style and use it.

Applying the same bibliographic style consistently throughout your document is more important than following the exact formatting required by the style guide that you have been asked (or have self-assigned yourself) to follow.

Virtually all 'author-year' bibliographic styles require page numbers for quotes (i.e. quoted citations) and for ideas that show up only at a specific point in the cited text.

9. The document is written for its audience.

Most of what I read is either academic English or academic French. I personally have limited ability to read, and close to zero ability to grade, anything else than academic writing, which emphasises **conceptual clarity and empirical precision**, on top of originality and/or source identification (Req. #8).

Different audiences will require different adjustments to parameters such creativity or originality, but virtually all audiences require concision, i.e. clear ideas expressed in as few words as needed.

Similarly, most audiences will require observing a basic modicum of typographical rules, such as using quotation marks and non-breaking spaces properly, as well as, when your audience is obsessive about typography, thin spaces and em or en dashes.

Last, your audience should guide your choice of document format (Req. #5).

10. The document has reached your reader on time.

Your reader almost certainly operates on a different work timetable than yours. Deadlines reflect this, and should be observed for that single reason.

Deadline extensions are not always grantable, and the probability that they will be granted decreases as time gets closer to the initial deadline.

If you fail to submit on time, as I often do, you will have to suffer the feelings that come with regular failure at satisfying your obligations towards other people, which comes at heavy psychological cost in the medium-to-long term.

In that unfortunate case, the most important things to do are

1. not to blow things out of proportion,
2. analyse the causes of why you submitted late, and
3. apologise, rather than providing excuses.

When documents are to be sent by email, it is always nice to show your reader that you know how to send polite, spell-checked emails with meaningful subject lines, sensible attachment names, and informative body text information. If it helps, treat your email as a 'cover page' for your document(s).