The Journal of Open Source Education

memoiR: R Markdown and Bookdown Templates to Publish Documents

Eric Marcon [1]

1 AgroParisTech, UMR AMAP, CIRAD, CNRS, INRAE, IRD, Univ Montpellier, Montpellier, France

DOI: 10.xxxxxx/draft

Software

- Review
- Repository
- Archive

Submitted: 22 August 2023
Published: unpublished

Summary

`memoiR` is a package for R that provides templates to publish well-formatted documents both in HTML and PDF formats. Documents can be produced locally or hosted on GitHub, where GitHub actions can update the published documents continuously.

Long documents, namely theses, are the main purpose of this package. Along with HTML output to be read online, their PDF version based on the LaTeXclass `memoir` can be highly customized.

Functions are provided to make the publication of the documents on GitHub very easy, including their continuous integration.

Templates include:

- "Memoir": a book template based on memoir. The HTML version is a GitBook or a Bootstap 4 book and the PDF version is formatted by the `memoir` LaTeXtemplate.
- "Stylish Article": an article for self archiving. The HTML versions are optimized to be read online and the PDF version is a two-column, well formatted article to be printed.
- "Simple Article": simpler than the stylish article. The HTML versions are the same and the PDF version follows the classical `article` LaTeXtemplate.
- "Beamer Slideshow": the HTML version is either IOSlide or Slidy and the PDF version is a Beamer slideshow.

A gallery is provided on the GitHub pages of the project.

Statement of need

Reproducibility is a major concern of science (Ioannidis, 2005), that induced propositions to improve it (Munafò et al., 2017).

Experimental design and data curation are out of the scope of this paper but a lot can be done in the last steps of the production of scientific documents, namely data processing, inclusion of the results (including tables and figures) into a written document, and the final formatting. In an academic context, reproducible student work is easier to update and its content more clearly assessable.

Literate programming (Knuth, 1984) is the framework: the code used to process the data is included into the text of the document, and run each time the document is produced. In the R (R Core Team, 2023) environment, its first implementation was `Sweave` (Leisch, 2002), that included code chunks into LaTeXdocuments. The `knitR` package (Xie, 2015) made its processing easy.

JOSE
The Journal of Open Source Education

The emergence of Markdown (Gruber, 2004) and its inclusion in Pandoc (MacFarlane, 2006) motivated the development of R Markdown (Xie et al., 2018) so that R users could switch from LaTeXto Markdown, enlarging the audience with less technical skills needed, and produce HTML outputs along with PDF documents. R Markdown allows producing simple documents (the so-called R Markdown notebooks). The `bookdown` package (Xie, 2016) extends its features to the required standards for scientific publishing, namely bibliographic citations, numbered figures and equations, and cross-references in multi-chapter documents. The powerful features of `bookdown` allow the simultaneous production of HTML and PDF documents, along with other formats such as Microsoft Word. The full process is summarized in figure Figure 1.

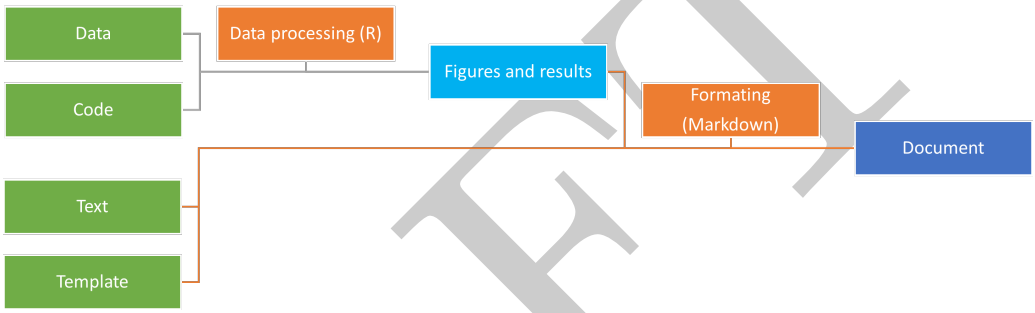

Figure 1: Processing documents with R and Markdown. Data and code are processed by R to obtain figures and results, which are included in the text and formated by R Markdown (actually by `knitr`, `bookdown`, Pandoc and a LaTeXcompiler) with respect to a template.

The `memoiR` package builds on `bookdown` with two main goals: provide a full set of well-formated templates and simplify the integration with GitHub pages to efficiently publish the documents.

The `memoir` LaTeXpackage (Wilson, 2006) allows publishing high-quality long documents such as books and theses. The `memoiR` package for R makes it available for R users. A wide set of options allows extensive customization of the documents. Other available formats are a simple and a stylish article, and a Beamer slideshow.

To ensure reproducibility, the continuous integration supported by GitHub Actions and publication on GitHub pages are powerful tools. `memoiR` provides the functions to simplify their use, including the production of the necessary scripts.

Other packages provide templates for R Markdown, e.g. `rticles` (Allaire et al., 2023), which gathers LaTeXtemplates for many journals, `rmdformats` (Barnier, 2022) which comes with several HTML output formats or `thesisdown` (Ismay & Solomon, 2023) for theses. In contrast, `memoiR` is a general purpose package, that always allows both HTML and PDF outputs, and was developed for easy continuous integration and publishing.

Acknowledgements

This work benefited from the support of “Investissements d’avenir” of the French National Agency for Research (Labex CEBA, ref. ANR-10-LABX-25-01). The author thanks Mathias Legrand for the original LaTeXclass of the stylish article.

JOSE
The Journal of Open Source Education

References

Allaire, J., Xie, Y., Dervieux, C., R Foundation, Wickham, H., Journal of Statistical Software, Vaidyanathan, R., Association for Computing Machinery, Boettiger, C., Elsevier, Broman, K., Mueller, K., Quast, B., Pruim, R., Marwick, B., Wickham, C., Keyes, O., Yu, M., Emaasit, D., … Hyndman, R. (2023). *Rticles: Article formats for R markdown* [R Package]. https://CRAN.R-project.org/package=rticles.

Barnier, J. (2022). *Rmdformats: HTML output formats and templates for 'rmarkdown' documents* [R Package]. https://CRAN.R-project.org/package=rmdformats.

Gruber, J. (2004). *Markdown*. https://daringfireball.net/projects/markdown/.

Ioannidis, J. P. A. (2005). Why Most Published Research Findings Are False. *PLoS Medicine*, *2*(8), e124. https://doi.org/10.1371/journal.pmed.0020124

Ismay, C., & Solomon, N. (2023). *Thesisdown: An updated R Markdown thesis template using the bookdown package* [R Package].

Knuth, D. E. (1984). Literate Programming. *The Computer Journal*, *27*(2), 97–111. https://doi.org/10.1093/comjnl/27.2.97

Leisch, F. (2002). Sweave: Dynamic Generation of Statistical Reports Using Literate Data Analysis. In W. Härdle & B. Rönz (Eds.), *Compstat* (pp. 575–580). Physica-Verlag HD. https://doi.org/10.1007/978-3-642-57489-4_89

MacFarlane, J. (2006). *Pandoc*. https://pandoc.org/.

Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Percie Du Sert, N., Simonsohn, U., Wagenmakers, E.-J., Ware, J. J., & Ioannidis, J. P. A. (2017). A manifesto for reproducible science. *Nature Human Behaviour*, *1*(1), 0021. https://doi.org/10.1038/s41562-016-0021

R Core Team. (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Wilson, P. (2006). The memoir class. *The PracTeX Journal*, *2006*(3).

Xie, Y. (2015). *Dynamic documents with R and knitr* (2nd ed.). Chapman and Hall/CRC. ISBN: 978-1-4987-1696-3

Xie, Y. (2016). *Bookdown: Authoring books and technical documents with R Markdown*. Chapman and Hall/CRC. ISBN: 978-1-138-70010-9

Xie, Y., Allaire, J. J., & Grolemund, G. (2018). *R markdown: The definitive guide*. Chapman and Hall/CRC. ISBN: 978-1-138-35933-8