

Chaining Tools to Create a Knowledge Repo

Michael Rustler¹

Hauke Sonnenberg¹

Christoph Sprenger¹

¹*Kompetenzzentrum Wasser Berlin gGmbH (KWB)*

March 1, 2019

keywords: *knowledge-repo, research-data-management, r, rstudio, blogdown, hugo, hugo-academic*

In small research institutes knowledge is exchanged in an informal way. Staff fluctuation leads to a loss of knowledge as processes and decisions are documented insufficiently. A company-wide knowledge repository helps formalising the documentation process. For example AirBnB established a knowledge repo and also open-sourced its software. However, it is only designed as “knowledge sharing platform between data scientists and other technical professions” and its features are limited to create taggable blog-posts with IPython Notebooks or R markdown, but these are not reproducible yet .

At our research institute also knowledge about code, people, projects, publications, tools and the links between them are important and need to be stored in one place in order to make it meaningful for all, not only technicians.

Thus we developed a toolchain solely based on open-source software, tools and cloud-services. Information from different sources is collected either automatically (using own R packages, e.g. for DataCamp, GitHub, Zenodo, Endnote) or manually (in the form of text file templates). These are exported into markdown files allows us to create a version controlled (Git), static website (Hugo with theme “academic”) which contains all features needed for our small research institute. It is automatically deployed by chaining continuous integration (Gitlab-CI, Travis) with the R package “blogdown” (website generation) and GitLab pages (to host it).

References

- AirBnB (2018). *Knowledge Repo: A next-generation curated knowledge sharing platform for data scientists and other technical professions..* version v0.8.8, <URL: <https://github.com/airbnb/knowledge-repo>>.
- Bion R, Chang R, Goodman J (2018). “How R Helps Airbnb Make the Most of its Data.” *The American Statistician*, 72(1), 46-52. <https://doi.org/10.1080/00031305.2017.1392362>, <URL: <https://doi.org/10.1080/00031305.2017.1392362>>.
- Cushen G (2019). *Academic: The Best Way to Create the Website You Want from Markdown (or RStudio/Jupyter)*. v4.0.0, <URL: <https://sourcethemes.com/academic/>>.
- GitHub (2019). *Pages: Websites for you and your projects*. <URL: <https://pages.github.com/>>.
- GitLab (2019). *Continuous Integration & Delivery*. <URL: <https://about.gitlab.com/product/continuous-integration/>>.
- GitLab (2019). *Pages: Websites for your GitLab projects, user account or group*. <URL: <https://about.gitlab.com/product/pages/>>.
- Hugo (2019). *The fastest framework for building websites*. <URL: <https://gohugo.io>>.
- R Core Team (2018). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. <URL: <https://www.R-project.org/>>.
- Travis-CI (2019). *The simplest way to test and deploy your projects*. <URL: <https://travis-ci.com/>>.
- hugo v0.54.0 (2019). *The world’s fastest framework for building websites*. <URL: <https://github.com/gohugoio/hugo/releases/tag/v0.54.0>>.
- Xie Y (2019). *blogdown: Create Blogs and Websites with R Markdown*. R package version 0.10.4, <URL: <https://github.com/rstudio/blogdown>>.
- Xie Y, Hill AP, Thomas A (2017). *blogdown: Creating Websites with R Markdown*. Chapman and Hall/CRC, Boca Raton, Florida. ISBN 978-0815363729, <URL: <https://github.com/rstudio/blogdown>>.

Source code

The source code for generating this abstract is available here, **Current version:** March 1, 2019; **Corresponding author:** michael.rustler@kompetenz-wasser.de"