A R Setup, renv, and Docker

R. Heller · Date: 2026-05-09 · Last changed: 2026-11-10

A.1 R installation and current first-use path

Install R from CRAN. The current candidate was locally validated with R 4.6.1. An IDE such as RStudio is optional. Rendering also requires Pandoc; PDF output additionally needs XeLaTeX/TinyTeX and R with Cairo support. Compiled packages may require a C/C++ toolchain and system libraries; the Linux dependencies are listed in `.github/workflows/render.yml`.

From the repository root, install and check the book's R dependencies:

```
RENV_CONFIG_AUTOLOADER_ENABLED=false Rscript --vanilla scripts/install-book-deps.R
RENV_CONFIG_AUTOLOADER_ENABLED=false Rscript --vanilla scripts/install-book-deps.R --
check
```

The installer installs missing packages from current CRAN and checks that their namespaces load. It does not restore historical versions or certify an immutable software environment. `--check` performs no installation or network access; `--list` prints the package inventory. Use `--library=/path/to/library` to select an R library, or `--include-optional` for packages used only by unevaluated examples and exercises. Optional Python, SQL, OCR and deployment prerequisites must be configured separately.

Internet access is required for installation and uncached provider requests. Configure an activated OpenAlex key locally as `OPENALEX_API_KEY`, outside Git, for requests that need authenticated paging. `SCIENTOMETRICS_OFFLINE=1` requires existing exact-query OpenAlex caches in `_freeze/openalex_cache/`; it does not make other providers offline and cannot supply missing cache entries. The repository `README.md` describes the helper functions in `R/` and the sample data in `data/`.

A.2 Package management with renv

renv can record and restore a project's R package versions. This repository retains a historical `renv.lock` and automatic activation in `.Rprofile`, but that lockfile has **not been qualified for the current book**. The first-use commands above bypass project startup and use the dependency installer. Do not assume that restoring the historical lockfile will reproduce the current validated candidate.

The following is an optional restoration example for a project whose lockfile you have independently validated, rather than a required setup step for this book:

```
install.packages("renv")
renv::restore()
```

Restoration attempts to install the recorded versions; success also depends on compatible R, system libraries and available package sources. In a development environment, add a package and deliberately update the lockfile only after validating the resulting environment:

```
renv::install("newpackage") # Replace with the intended package name.
renv::snapshot()
```

A software lockfile does not freeze live API results. Exact reproduction also requires the same licensed input snapshots and analysis configuration.

A.3 Core packages

The table below lists the primary packages used throughout the book.

Package	Purpose	Chapter(s)
openalexR	OpenAlex API access	5–37
tidyverse	Data wrangling and visualisation	All
igraph	Network analysis	15–20
tidygraph	Tidy network manipulation	15–20
ggraph	Network visualisation	15–20
quanteda	Text mining and corpus management	21–25
topicmodels	LDA topic modelling	22, CS-5
stm	Structural topic models	22
gt	Publication-quality tables	Throughout
targets	Pipeline orchestration	34
plotly	Interactive charts	35
DT	Interactive data tables	35
shiny	Web applications	37
word2vec	Word embeddings	24
uwot	UMAP dimensionality reduction	24

A.4 Docker environment

The repository's `Dockerfile` is an **optional historical recipe**, based on `rocker/verse:4.4.1` and `renv::restore()`. Its restore path has not been qualified for the current book. It also downloads an amd64 Quarto package, so other architectures need a compatible recipe. Docker does not remove the need to validate the lockfile, system dependencies, API access and resulting outputs.

For a deliberate evaluation of that recipe, build it and inspect its R version:

```
docker build -t scientometrics-in-r .
docker run --rm scientometrics-in-r R --version
```

The image's default command calls `bookdown::render_book()` for all formats; it does not launch RStudio Server. Successful construction and execution of this historical image are not claimed here. Use the current first-use path above for the locally validated candidate.

A.5 Render the book

This repository is configured for **bookdown**, with `_bookdown.yml` and `_output.yml`. From the repository root, render HTML and refresh its download manifest:

```
SCIENTOMETRICS_REBUILD_CHUNKS=1 RENV_CONFIG_AUTOLOADER_ENABLED=false Rscript --vanilla
-e 'bookdown::render_book("index.Rmd", output_format = "bookdown::bs4_book")'
Rscript --vanilla scripts/build-download-manifest.R
```

`SCIENTOMETRICS_REBUILD_CHUNKS=1` re-executes analysis chunks while retaining exact-query provider caches. Missing provider caches still require API access. The configured output directory is `docs/`. Serve it locally, for example:

```
python3 -m http.server 8000 --directory docs
```

Then visit `http://localhost:8000`. PDF additionally needs XeLaTeX/TinyTeX; `output_format = "all"` requests HTML, PDF and EPUB. Before distribution, run `Rscript --vanilla scripts/repair-epub.R docs/scientometrics-in-r.epub` to validate the ebook's internal links and bundle its citation and licence files, then run `Rscript --vanilla scripts/build-download-manifest.R`. Local outputs do not update the published site. The README lists the PDF font prerequisites.

Quarto appears in optional examples elsewhere in the book, but `quarto render` is not the build command for this repository. A Quarto CLI is needed only when working through those separate examples.