

Open-Source formats

What are Open-Source formats?

Much of science relies on computers and software. Many of the traditionally used software is commercial and thus requires a license. The cost for such licenses is a barrier to opening science, since not all institutes or labs will be able to afford the same licenses. This means some researchers might not be able to verify the conclusions of others, or benefit from their shared materials. The use of Open-Source formats aims to prevent this and contribute to data being interoperable (see [Open Data](#)). This involves ensuring data is shared in formats accessible by Open-source software, which can of course be achieved by using Open-source software as much possible throughout the research process.

Benefits of open source formats

- Maximizing reproducibility
- Maximizing usefulness of materials contributing to science and its progress

How to engage?

This is largely based on personal initiative. Search for Open Source variants of software you are using and decide whether and how these could replace your current proprietary software. Sometimes, a full switch to using a new software is the best solution; the School is currently starting the process of switching from using SPSS to using R-based software for statistical analyses. In some cases, sharing data or materials from proprietary software in formats that can be read by or executed by (e.g., computer code) by Open Source variants is the easiest solution. Of course, in line with [FAIR principles](#) you ideally confirm that the Open Source software indeed works for your materials. In this light, it is appropriate to report the software version for which you confirmed this.

Indeed, citing the software and packages you used is important, both from a reproducibility perspective but also as a way to acknowledge the work and time that people spent creating

tools for others. Ideally, you would also include version numbers here. Please see below for an example:

All statistical analyses were run using R 3.3.2 (<https://www.r-project.org>; R Core Team, 2016). In addition to the base version of R, we used the packages dplyr 0.5.0 (Wickham, 2011), effsize 0.7.0 (Torchiano, 2016), ggplot2 2.2.0 (Wickham, 2009), Hmisc 4.0-1 (Harrell, 2016), lm.beta 1.5-1 (Behrendt, 2014), multicon 1.6 (Sherman, 2015), psych 1.6.9 (Revelle, 2016), and yarr 0.1.2 (Phillips, 2016) in our analyses.