

## Reproducibility Template Draft

Note: This draft is revised version of SIGMOD Template 2016, most of things are the same, but sessions are reorganized and some summarized information is from SIGMOD Guidelines: <http://db-reproducibility.seas.harvard.edu/>

---

The end goal is to successfully reproduce the raw data and relevant plots that the authors used to draw their conclusions. Authors should provide a complete set of scripts to install the system, produce the data, run experiments and produce the resulting graphs along with a detailed Readme file that describes the process step by step so it can be easily reproduced by a reviewer or other researchers.

### A) Hardware Specification

[Here you should include any details and comments about the used hardware in order to be able to accommodate the reproducibility effort. Any information about non-standard hardware should also be included. You should also include at least the following info:]

- 1) Processor (architecture, type, and number of processors/sockets)
- 2) Caches (number of levels, and size of each level)
- 3) Memory (size and speed)
- 4) Secondary Storage (type: SSD/HDD/other, size, performance: random/sequential read or write)
- 5) Network (if applicable: type and bandwidth)

### B) System and Environment Setup

System setup is one of the most challenging aspects when repeating experiments. System setup will be easier to conduct if it is automatic rather than manual. You should test that the system they distribute can actually be installed in a new environment. The documentation should detail every step in system setup]

- 1) Operation System (e.g., the required compiler must be run with a specific version of the OS)
- 2) Configuration for the environment if needed (e.g., environment variables, paths)
- 3) Programming Language: [C/C++/java/...]
- 4) Additional Programming Language info: [optional, e.g., java version]
- 5) Packages/Libraries Needed: [an as thorough as possible list of software packages needed]
- 6) Compiler Info: [full details of compiler and version]
- 7) Procedures to test if the system is configured correctly

[Ideally, there is a script called: ./prepareSoftware.sh]

### C) Datasets Info

- 1) Repository: [url]
- 2) Data generators: [url]

### D) Experimentation and Measurements

- 1) Scripts and how-tos to generate all necessary data or locate datasets

[Ideally, there is a script called: `./prepareData.sh`]

- 2) Scripts and how-tos for all experiments executed and measurements are taken

[Ideally, there is a script called: `./runExperiments.sh`]

3) Scripts for a clean-up phase where the system is prepared to avoid interference with the next round of experiments.

### E) Data Representation and Visualization

[For each graph in the paper, you should describe how the graph is obtained from the experimental measurements.]

- 1) Tools that are used to generate the graphs (e.g., Gnuplot or Matplotlib)
- 2) Scripts (or spreadsheets) how to generate the graphs.

=====