**Protocol of the File Formats session 3.3**

**library of standardised actions to describe reduction steps already performed**

**participants**:  
Jochen Stahn (chair), Artur Glavic, Alexandros Koutsioumpas

**initial idea**:

In an **human-readable** file, the reduction steps already performed should be mentioned, probably also the actual state of the data reduction (footprint corrected, backgroung not subtracted, ...)

**elaboration on this**:

* The reduction software can rather easily provide this information
* It seems several programs also have and use this feature.
* One can get more specific, e.g. [footprint correction: with reference sample] or [footprint correction: using sample size and alpha\_i]
* If there are some **standard procedures** used often and by various codes, a defined keyword can be used, where the definition (probably with formulae) is provided by an dictionary on the orso web page (or so).
* A slightly different approach is to include the **echo of the program call** in the header, with a description of what the arguments mean. More general: to provide the full parameter set used for the data reduction and some kind of a short documentation.

**format of an ASCII file**

This was not the scope of this session, nevertheless we ended up there.

* The purpose of this file is to provide the most important information to the user. SHe should find out the state of data reduction and applied corrections easily. An echo of the program call or the like enables a fast repetition, probably with changed parameters.
* The header should have a standardised format (JSON, YAML, ...) AND it should be easily readable by humans.
* The column description must be correct and should have an assignment of the column number to the property. The latter is important in case there are (many) more columns available then the minimum set of 3 (or 4).
* For compatibility with most existing analysis software, the first 4 columns should always be " qz | Rqz | sigma\_Rqz | sigma\_qz ". These can be followed by any number of columns with any content.
* Complex data sets or a huge amount of meta-data (a more complete set of information) should be handled with the nexus format. They wont be *easily readable* any format. And then a hierarchical structure facilitates to find the required information.