

# A semantic approach to scientific logbook used in physics experiments

Gianluca Moro<sup>1</sup>, Manola Carraro<sup>2</sup>

Abstract

A meticulous record of the activities and observations has always been a valuable companion to every scientist, more specifically a “logbook” usually keeps a chronological trace of the evolution of the experiment. A scientific experiment is by definition keen to provide unexpected outcomes, so a strict logbook structure is not very suitable to the task because meaningful features or new kinds of events may emerge during the ongoing of the experimental campaign. On the other hand, the utility of a logbook is empowered by its capability to be searched and analyzed with respect to many different parameters, not always completely defined from the beginning of the experiment.

Semantic wiki represents an interesting compromise between the internal markup language flexibility, typical of a wiki, and the possibility to classify and tag any kind of information for a subsequent analysis. Leveraging the innate features of the Mediawiki framework, we propose a new structure for a scientific logbook to be exploited in the SPIDER and MITICA experiments hosted at Consorzio RFX. Such experiments consist in a series of “shots”, each one characterized by a large amount of information related to different phases:

1. proposal phase: who is requesting this type of shot, the intended effect to explore, the needed diagnostics
2. planning phase: when the shot is planned, in which session, which resources are assigned
3. realtime acquisition of the data during the shot execution
4. post analysis

For phase 3, the realtime data is acquired and stored using MDSPPlus (a set of software tools for data acquisition and storage and a methodology for management of complex scientific data - <https://www.mdsplus.org/>), while the other phases have not a predefined tool. The proposed logbook is related to phase 4; we define a basic data structure tagging using Semantic wiki, proposing some structured table to be filled for each shot, giving the possibility to precompile part of it with fields related to “session values”, which are usually common to a set of consecutive shots.

This approach provides the basis to build a structured set of data that can be easily analyzed and accessed by automatic analysis tools, machine learning algorithms. In addition the proposed solution can be also applied in phase 1 and 2, providing researchers with a comprehensive solution that covers all the experimental phases.

---

<sup>1</sup>Gianluca.Moro@unipd.it - Centro Ricerche Fusione, Università di Padova at Consorzio RFX, Corso Stati Uniti 4, 35127 Padova, Italy

<sup>2</sup>Manola.Carraro@igi.cnr.it - CNR at Consorzio RFX, Corso Stati Uniti 4, 35127 Padova, Italy