

Beneficiary's Acronym: NMI3-II/WP6

Name of person completing this report: E. Farhi

Date of completion: Jan 30th 2015

1. Publishable summary (stay short and concise max 1 page for whole section 1)

1.1 Short description of your WP objectives and context (inspire yourself from Annex I)

The Data Analysis Standards work-package (WP6) is a networking activity which aims at determining how to develop, deploy and operate a common data analysis software infrastructure to facilitate joint software development for the neutron scattering and muon community in Europe. An analysis of the existing software landscape and coding practices is the starting point of this activity. Then, a limited set of data treatment functionalities will be implemented in an inter-facility common infrastructure, demonstrating the applicability of high-end coding standards. As part of this effort, a set of recommendations for coding software, as well as for project management at the level of an EU common software infrastructure for neutron/muon facilities, will be proposed. This way, we wish to motivate scientist programmers to adopt common standards in order to favour software interoperability and collaboration.

1.2 Work performed since the beginning of the project, a description of the main results achieved so far.

An initial review of the existing software landscape was achieved as the initial deliverable of this project, in conjunction with the edition of an evaluation DVD containing most of this software. Then, we studied the software infrastructure of these evaluated software, in terms of software design, tools used to develop, adopted languages, dependencies and the involvement of the user community. The Mantid project <www.mantidproject.org> was selected as a test platform to implement a prototype software. No other collaborative software project was envisaged in the frame of this work-package. As Mantid was initially designed to handle data acquired at neutron/muon pulsed sources, we chose to evaluate the ability of Mantid to support instruments installed at continuous neutron sources. Code was produced to be able to import and reduce data from a set of instruments installed at the ILL (France), the LLB (France), and the PSI (Switzerland). The efficiency of the produced code was compared with existing legacy programs such as LAMP <<http://www.ill.eu/?id=3463>>. The code was officially contributed to the Mantid project as of release 3.3 (Januray 2015). The last deliverable of this work-package focuses on the evaluation of the produced code, and more generally of the Mantid project. This work-package is now completed.

1.3 Expected final results and their potential impact and use (including the socio-economic impact and the wider societal implications of the project so far).

This work-package aims at showing that a common software solution can be shared at most neutron/muon facilities. The chosen solution, here Mantid, has been used as a prototype infrastructure for this purpose, but other initiatives with similar standards could also be envisaged. Our reports and produced code can be used by any group producing scientific software, and more specifically in the neutron/muon user community. Scientists in EU can benefit from a shared solution by lowering the learning curve when travelling across facilities. It is hoped that a properly designed, unified software aimed at simplicity can participate to attract more scientists to use the neutron/muon facilities.

2. Project objectives for the period

2.1 Project objectives

High resolution detectors and higher data rates, among others, make new demands on software that require professional software solutions. In particular, new instruments constitute a considerable challenge for software provision and failure to address this issue leads to a delay in the scientific impact of new investment.

The importance of data analysis has now become clear and is becoming a focal point of efforts to optimise scientific production.

The aim of this networking activity is to determine how to develop, deploy and operate a common data analysis software infrastructure to facilitate joint software development in the neutron scattering community.

The Data Analysis Standards networking work-package (WP6) will begin with two review tasks: review existing data analysis software and practices of software developers and review existing solutions for a common data analysis infrastructure. Based on the reports from these tasks, the key phase of the project will be to develop a small number of prototype software solutions in areas chosen by the partners. The goal will be to deploy functionality that exists at some facilities in the common framework thereby making it accessible to and optimised for all partners.

Task 1 : Review existing data analysis software and practices of software developers

Task 2 : Review existing solutions for a common data analysis infrastructure

Task 3 : Develop prototype software in chosen solution for representative applications

Task 4 : Evaluate prototype software

2.2 Work progress and achievements

Please provide a concise overview of the progress of the work in line with the structure of Annex I to the Grant Agreement.

For each work package, except project management, which will be reported in the management section, please provide the following information:

- A summary of progress towards objectives and details for each task;
- Highlight clearly significant results;
- If applicable, explain the reasons for deviations from Annex I and their impact on other tasks as well as on available resources and planning;
- If applicable, explain the reasons for failing to achieve critical objectives and/or not being on schedule and explain the impact on other tasks as well as on available resources and planning (the explanations should be coherent with the declaration by the project coordinator) ;
- a statement on the use of resources, in particular highlighting and explaining deviations between actual and planned person-months per work package and per beneficiary in Annex I (Description of Work);
- If applicable, propose corrective actions.

The work package effectively started in June 2012 with the recruit of a trained staff, that is 4 months after the launch of the NMI3-II project.

Task1: Review existing data analysis software and practices of software developers

In order to ease the evaluation step of the project, a set of easily available software for neutrons and muons science was identified and gathered into a Live DVD, that is an operating system holding on a DVD, that can be used without affecting the computer content. This DVD can be used on any current computer to evaluate the pre-installed software, as well as during workshops, tutorials and

training sessions. This preliminary task, which was not initially in the list of deliverables, required about a month.

The evaluation of current software revealed a sparse landscape which is substantially far from the definition of standards. Consequently, handling this diversity required longer than expected. In parallel to this task, getting acquainted with the Mantid project, which is to be evaluated as part of this work-package, required about 6 months. However, learning about Mantid was considered essential before the Task 1 of the project could be completed. For these reasons, the “Report on current software and practices” (Task 1 deliverable D6.1) was produced after its planned date. The Task 1 report, which was published in March 2013 after approval and input from the work-package members, has been published on our project page. It consists in an analysis of 24 existing and easily available software packages. The programming methodologies, user experience, data formats, as well as current practices are detailed. A list of recommendations is then produced.

Task 2: Review existing solutions for a common data analysis infrastructure

The Task 2, which aims at reviewing current practices for a common data analysis infrastructure, mainly deals with technical aspects needed to organise a large software project. These aspects were already partly discussed in the Task 1 report. We have collected and analysed software development items that all contribute to a successful project, including code development, distribution of the software, how to foster a user community, and considerations about the user interface appearance. The report was published in March 2014.

Task 3: Develop prototype software in chosen solution for representative applications

The 3rd Task, which constitutes to core of the work-package, required a strong implication into the Mantid project. Considering the complexity of the Mantid infrastructure, we decided to initiate the Task 3 (prototype coding) before the completion of the Task 1 and Task 2. Actually, importers for Mantid have been produced to support nine continuous source based instrument data sets (IN4, IN5, IN6, D33, D2b, D17, IN16b at the ILL, MiBemol at the LLB, and Focus at the PSI). The corresponding detector geometry for these instruments have been described. In addition, specific algorithms for reduction steps after loading the data were coded. The existing Mantid 'algorithms' for data reduction have then been applied and compared with equivalent 'macros' in LAMP. We find that results from Mantid and LAMP compare well, but small differences arise, due to the way data sets are stored and handled. The maintenance of the early coded Mantid algorithms added-up to the will to support as many instruments as possible within the work-package duration. For this reason, we chose to extend the prototype coding task D6.3 up to the end of the work-package funding. The report, which also contains a print-out of the source code, was published in September 2014. Source code is also available on the work-package web page, and was pushed as a contribution into the Mantid project.

Task 4: Evaluate prototype software

The Task 4 required Task 3 to be completed, and could only be started in Autumn 2014. In addition, the delivery date was postponed due to illness for 2 months. The final evaluation did not consider the accuracy of the produced code compared with existing software. Such a comparison depends on

the technical choices made during the coding phase, as well as Mantid infrastructure internals which are beyond our scope. We hope that our contribution to Mantid will be properly maintained, and improved in the coming years, to satisfy the user needs and expectations. We thus focused the Task 4 on the evaluation of the work-force needed to produce and maintain scientific code when using the Mantid infrastructure. We hope that these estimates can be used to properly scale the work-force in order to serve the neutron/muon community with modern, unified software.

Work-package achievements

Actually, the work-package provides a clearer view about the existing software landscape for neutron/muon scattering, and the methodologies that have been used to produce these tools. Our contribution to the Mantid project is part of an evaluation, but has however brought a set of algorithms dedicated to continuous neutron source instruments. Due to the limited duration and funding of the work-package, we had to restrict most of our study and contribution to the sole Mantid project. We hope that this contribution will be used by the community, either from Mantid itself, or as components in other software. Finally, we have produced a list of recommendations for coding standards and methodologies, that could ease the future development of inter-facility software. The detailed content of these standards is left open for further initiatives.

All deliverables have been published on the NMI3-II web site <<http://nmi3.eu/about-nmi3/networking/data-analysis-standards.html>>, and are available to the public, including the produced source code. The whole work-package deliverables have been completed within the NMI3-II project duration.

3. Deliverables

DELIVERABLES									
Del. no.	Deliverable name	WP no.	Lead beneficiary	Nature	Dissemination level	Delivery date from Annex I (proj month)	Actual / Forecast delivery date	Delivered Yes/ No/ Ongoing	Comments
D6.0	Catalogue of existing software	6	ILL	other	public	-	1	Yes	New deliverable as a DVD on <nmi3.eu> website
D6.1	Report on current software and practices	6	ILL	report	public	2	10	Yes	Late staff recruit, longer than expected training period.
D6.2	Report on solutions for developing a common software infrastructure.	6	ILL	report	public	4	15	Yes	Late staff recruit, longer than expected training period. Extrapolated from D6.1
D6.3	Prototype software in chosen solution	6	ILL	prototype	public	18	27	Yes	Code was developed until the end of the project (month 27). Now contributed to <mantidproject.org>
D6.4	Report on evaluation of prototype software.	6	ILL	report	public	24	32	Yes	Could only be achieved after D6.3 Report delayed due to illness.