



Abu-MaTran

Automatic building of Machine Translation

PIAP- GA-2012-324414

D2.1. Workshop on software management

Dissemination level	Public
Delivery date	2013/06/30
Status and version	Final, v1.0
Authors and affiliation	Gema Ramírez-Sánchez (Prompsit), Antonio Toral (DCU)



Project funded by the European Community under the Seventh Framework Programme for Research and Technological Development



Table of Contents

Executive Summary.....	2
1 Workshop Description	3
2 Conclusions.....	4
Annex A. Workshop Materials.....	5

Executive Summary

This deliverable corresponds to task T.2.1. (*Workshop on software management for researchers*) within work package 2 (*Dissemination and outreach*) aiming at the dissemination of the project for different audiences (industrial stakeholders, academia and general public).

This particular task was carried out by the industrial partner, Prompsit Language Engineering, during the secondment of Gema Ramírez-Sánchez in Dublin City University (DCU) in March and April 2013. It was developed under the supervision of Dr. Antonio Toral and the support of DCU's technical and administrative staff.

The workshop took place at DCU's School of Computing premises on 19th April 2013 with the participation of 18 people from the Centre for Next Generation Localisation and DCU. The attendants were mainly project managers and administrators, but also early stage researchers. It was organised as an 8-hour workshop combining theory and hands-on activities with the participation of the invited speaker, Dr. Lluís Padró, from the Universitat Politècnica de Catalunya, who gave a talk about the Freeling¹ project.

As part of the workshop, 8 best practices for software management in research environments were outlined and discussed among the participants. It was agreed that some actions could take place to disseminate those best practices specially among early stage researchers.

All workshop materials (participant's guide, workshop slides and Dr. Lluís Padró video talk) that were produced as part of this task have been made available to the general public through the Abu-MaTran website (written material) or through DCU's School of computing website (video talk²). The participant guide and slides have been published under a free license.

The workshop received positive feedback and, as future work related to this activity, partner Prompsit is preparing a survey to get to know the adhesion to the best practices of all the partners. It has been also planned to prepare a shorter version of the workshop to be carried out during Prompsit's secondments at the remaining partners.

1 <http://nlp.lsi.upc.edu/freeling/>

2 <http://www.computing.dcu.ie/videos/llu%C3%ADs-padr%C3%B3-seminar-talk-19th-april-2013>

1 Workshop Description

The *Workshop on Software Management*, which was finally named *Workshop on Agile Software Management for Research Environments*, aimed at outlining some basic best practices on software management for people working in research, both newcomers and senior researchers or project managers. The topics covered by the workshop were taken from practices that have proven to be very useful in project management in industrial scenarios as well as in Prompsit's experience managing the Apertium³ platform project over the last 8 years.

The main challenge when thinking about software management in a research environment is to find a compromise between the time that the researcher devotes to management and the scientific objectives that are expected to be achieved. Too time-consuming practices will be hardly ever implemented in researcher's daily work while the complete absence of management policies can lead to chaotic situations like inability to recover data or to reproduce an experiment, to get collaborators or even further funding.

Having these parameters in mind, and by addressing some topics that involve the whole research life-cycle, we aimed at defining a set of basic best practices. In all, 8 topics were discussed during the workshop covering:

- coding practices: how to work under version control systems, how to make software stable by implementing regression tests and how useful it is to adhere to standards when coding
- documentation and diffusion practices: how to spread the word about the researcher's work for different purposes and audiences
- dissemination practices: how to make software reusable by packaging and releasing it to make end-user's life easy
- licensing and authorship practices: how to indicate the authors of a software and how to distribute it with a license that suits the objectives of the authors
- administration and management practices: how to further give support and maintain software and projects that comes to a mature status.

During the workshop, 8 best practices were finally outlined:

1. Put your software source code and documents under a version control system.
2. Write regression tests to assess performance of your code (unit tests) and the functionalities of your system (functional tests).
3. Use standards to maximise interoperability, to avoid misunderstandings, to save time and to build upon solid foundations.
4. Document your software, for a given purpose, having in mind who is the reader, taking the pleasure to tell others about your work.
5. Package clean and easy-to-use versions of your software and release them.
6. Choose a license for your software, indicate the authorship.
7. Share your software, let it succeed!
8. Ensure and organise support for your software project, let everybody know how things will

3 www.apertium.org

work.

To support the theory and exercises that were prepared for each topic, we decided to invite Dr. Lluís Padró from the Universitat Politècnica de Catalunya. He is the main researcher and manager of a software project which was created inside a research environment and which is now widely used among the Natural Language Processing community: Freeling .

Dr. Lluís Padró gave a talk exploring all the topics covered by the workshop and some additional information regarding the Freeling project. Attendants were able to exchange questions and answers with him after the talk and during the whole workshop. The talk was recorded by the technical staff of DCU's School of Computing and is available through DCU' School of Computing website at: <http://www.computing.dcu.ie/videos/lluís-padró-seminar-talk-19th-april-2013>

The day was opened with Dr. Lluís Padró talk. After that, each topic was addressed as follows:

- all participants received a workshop guide containing an introduction to every topic with suggested reading and some exercises
- every topic was introduced with the support of ah-doc prepared slides leading to a discussion of the topic
- after discussion, participants were asked to complete the exercises for each topic which were also discussed together
- each topic was concluded with the definition of a best practice

The guide as well as the slides were distributed to all participants and are available under a free/open-source creative commons license through the Abu-MaTran website at:

A copy of both materials is also attached to this deliverable.

2 Conclusions

We can highlight some very positive aspects that arose from the organisation of this workshop:

- Attendants interest: while the expected number of participants was 10 (see Description of Work), we finally got 18 participants
- Useful content: the workshop guide along with the video and the 8 best practices allow for casual off-line users to be introduced to the topic, follow the workshop and adhere to the topics. As an example, at the moment of producing this deliverable, Dr. Lluís Padró talk has been visited more than 250 times.
- Feedback: we received positive feedback about the workshop from the participants who received, in exchange, a sum up of the best practices and a digital copy of the materials.
- Evaluation: a survey is being prepared to evaluate how much adhesion we find nowadays to the best practices at least in the Abu-MaTran research centres.
- Re-usability: A shorter version of the workshop is being also produced to be offered to the rest of the partners as an activity during Prompsit's secondments.

Annex A. Workshop Materials

The written materials used at the workshop (slides and guide) are delivered together with this document in PDF format. The fonts for slides (LibreOffice Open Document format) and guide (LaTeX format) will be delivered on demand and under a Creative Commons Attribution-ShareAlike 3.0 license⁴

⁴ <http://creativecommons.org/licenses/by-sa/3.0/>