

PERICLES - Promoting and Enhancing Reuse of Information
throughout the Content Lifecycle taking account of Evolving
Semantics
[Digital Preservation]

DELIVERABLE 7.3
Report on Training Activities



GRANT AGREEMENT: 601138

SCHEME FP7 ICT 2011.4.3

Start date of project: 1 February 2013

Duration: 48 months

“This project has received funding from the European Union’s Seventh Framework Programme for research, technological development and demonstration under grant agreement no601138”.



Project co-funded by the European Commission within the Seventh Framework Programme (2007-2013)		
Dissemination level		
PU	PUBLIC	X
PP	Restricted to other PROGRAMME PARTICIPANTS (including the Commission Services)	
RE	RESTRICTED to a group specified by the consortium (including the Commission Services)	
CO	CONFIDENTIAL only for members of the consortium (including the Commission Services)	

Revision History

V #	Date	Description / Reason of change	Author
V0.1	19/01/2017	First outline of a content structure	NO (HB)
V0.2	31/01/2017	Approved revised narrative flow	HB, KCL, DOT
V0.3	21/02/2017	First draft including Introduction and rationale + chapter introductions and descriptions	NO (HB)
V0.4	23/02/2017	Revised version	CS (KCL)
V0.5	16/03/2017	Second draft submitted for internal review	CS (KCL)
V0.6	20/03/2017	Review comments and revisions	IK (DOT), NO, KM, KNH (HB)
V0.7	28/03/2017	Revised version after review	KCL
V1.0	04/04/2017	Final version submitted to EU	HB/KCL

Authors and Contributors

Contributors at this WP are the end user partners. WP1 provides feedback to technical partners.

Authors

Partner	Name
KCL	Christine Sauter

Contributors

Partner	Name
HB	Nasrine Olson

Reviewers

Partner	Name
DOT	Ioannis Kotsopoulos
HB	Karen Nowé Hedvall, Kathleen Mahon

Table of Contents

Contents

1	Executive Summary	6
2	Introduction & Rationale	7
2.1	What to expect from this deliverable	7
2.2	Document Structure	7
3	Taming the Diversity	9
3.1	The diverse nature of the developed results	9
3.2	Getting it right: the target audience	9
3.3	Finding the right mode of translation	10
3.4	All in good time	11
4	Design	12
5	Reflections on questions of impact and success measurement	15
6	Reflections on questions of impact and success measurement Fehler! Textmarke nicht definiert.	
6.1	The planning	17
6.2	The survey	18
6.3	Available expertise and resources	18
6.4	Training activities	20
6.4.1	<i>In person interaction</i>	20
6.4.2	<i>Live online formats</i>	21
6.4.3	<i>Screened presentations</i>	22
6.4.4	<i>The PMTP</i>	22
7	Conclusion	24
8	Annex: Overview of activities Fehler! Textmarke nicht definiert.	

Glossary

Abbreviation / Acronym	Meaning
PMTP	PERICLES Modular Training Package
MOOC	Massive Open Online Course
RTD	Research Technology and Development
PET	PERICLES Extraction Tool
PeriCAT	PERICLES Content Aggregation Tool
PROPhet	PERICLES Ontology Population Tool
ERMIR	Entity Registry Model Repository

1 Executive Summary

Although this document was produced as a deliverable within the PERICLES project (2013 to 2017), with the original purpose of reporting on the training events carried out in particular during the last 14 months of the project duration we have widened the scope to give it more value outside the context of the project. The motivation of this document is to provide assistance and recommendations based on our reflections regarding the plans and the actions effectively carried out during the project lifetime. We propose to do this by analysing the experience we made both in terms of success, and lessons learnt. Underlying is the assumption that many RTD projects meet the same types of challenges when it comes to planning their training work package.

The purpose of this document is therefore an exploration of these challenges with the hindsight of our experience, and an attempt at a recommendation of good practices of how to meet these challenges. We will report on our activities, though rather with the aim of giving examples rather than of reporting in full on all our achievements in the training work package. The full account of our work, with all details can be found in the final report of the project PERICLES and on the project website for each activity.

Our main focus areas in this analysis are how to deal with the diversity of the nature, the readiness of mature results and the range of expertise in RTD projects, how to comply with the basic pedagogical principles for designing both training events and pertaining material, and how to measure success given the some of the common constraints in funded projects. On the basis of these reflections, we will then take a closer look at the decisions we made, and the success of our activities and plans or proposals for improvement.

At the start of a project, there are many areas that require a prior discussions to establish a shared understanding of the scope and direction of the training work plan: what target groups to address, what knowledge is needed to carry out training activities or develop training material, what formats are feasible. Wrong assumptions can cost a lot of time, by initiating actions that do not bring about helpful results or by planning activities that require an expertise not necessarily as strongly present in an RTD project as required for these types of action. In addition to that, training requires a subject matter that has reached a certain level of maturity in order to transfer knowledge about it. This implies that training activities come late and one needs to be well prepared, sometimes based on an anticipated expectation of what exactly the outcome will be. And finally, there is the challenge of motivating the RTD team, often not have much pedagogical experience and are under time pressure to finalise their own priorities, to provide adequate input and support to the coordination of training efforts.

2 Introduction & Rationale

2.1 What to expect from this deliverable

The H2020 programme is placing a special emphasis on the notion of “communication”¹ to complement the notion of “dissemination” and “exploitation”. With respect to the idea of a more direct engagement with stakeholders, implying a two-way exchange, training offers a perfect means to reach out to a diversity of audiences.

The purpose of this document is to share insights about challenges and opportunities relating to training engagements in a project such as PERICLES, which is a four-year multifaceted, multidisciplinary integrated *Research and Innovation* project, producing a large set of results and multiple tools, while addressing a broad range of audiences and involving partners with varying interests and expertise in different areas. The emphasis lies on the nature of the project which entailed a range of challenges.

While the training activities and material will be outlined later in this document, the main body of this document will be dedicated to a reflective discussion of training considerations, the core aim of the efforts, envisaged impact, and the lessons learnt.

At the start of a project, there are many potential pitfalls for the knowledge transfer work plans: assumptions on what “training” means, what type of training is suitable for what type of results, what targets groups are the ones to address, what knowledge is needed to carry out training activities or develop training material. Wrong assumptions can cost a lot of time, by initiating actions that do not bring about helpful results or by planning activities that require an expertise not necessarily as strongly present in an RTD project as required for these types of action. Certainly in addition to that, training needs an objective that has reached a certain level of maturity in order to transfer knowledge about it. This implies that training activities come late and one needs to be well prepared, sometimes based on an anticipated expectation of what exactly the outcome will be.

With the training work package, we have aimed to reach out and share the knowledge produced within the project. The complexity of the findings that result from such a technical and exploratory research project necessitated different ways of approaching diverse audiences, both at individual and community levels, geographically spread and with very different background and experience. We tried to tackle the challenge of aligning training needs, resources, and emergent opportunities, as much as actual efforts with the activity plans that had been drawn long before the start of the project.

In the sections that follow we elaborate on some of these issues.

2.2 Document Structure

Sections 3-5 deal with general reflections on how to approach a training work package and plan within RTD projects.

Section 3 looks at the diversity of results and the target audience.

Section 4 discusses pedagogical considerations in the design of training activities and material

Section 5 reflects on constraints for measuring impact

1 <https://ec.europa.eu/research/participants/portal/desktop/en/support/faqs/faq-933.html>

Section 6 revisits the steps from first planning through the different steps and actions taken, explains the rationale behind them, highlights the successful parts and those that constitute lesson learnt, which we would revisit if we started afresh.

3 Addressing the Diversity

The complexities of the challenges we faced in designing training material and events for an RTD project were multidimensional. Not only by its nature, PERICLES constituted a complex research project combining many disciplines, with partners from varying expertise and backgrounds, but also the potential audiences of the project covered a diverse set of groups, from the academic fields to professionals from the preservation and IT developing domains. In addition to those, the different modes of training available these days are very broad adding to the complexity of decisions to be made. A further complexity was added by the temporal dimension, in that much of the project outcomes reached maturity towards the latter part of the project, or that the timing of emerging training opportunities (such as workshops at relevant conferences) would not always align well with the RTD work plan timelines.

A number of steps were taken to meet this set of challenges.

3.1 The diverse nature of the developed results

Different disciplines were involved in creating individual building blocks in the PERICLES overall approach. Communication between these different fields and work packages was central in reaching a shared understanding of the anticipated overall outcome. Production of a glossary of terms is a good example for the outputs of these discussions, which informed also the training team. Keeping informed of the progress in the different areas, monitoring the opportunities that emerge for training objectives, and relevant contributions from different project members for different training activities required constant attention and “fly on the wall” presence in the monthly WP meetings of the training.

We discovered over the course of the first part of the project that although in principle it was a “preservation” project in that it provided an innovative approach in support of long-term preservation, the “preservation” practitioners would not in every case be the appropriate target group for each of the individual results. We found that the best way to proceed was to start with gathering as much knowledge on the anticipated outputs as early as possible, even if the training as such will only take form much later.

With the analysis of the anticipated results, and in particular their planned manifestation (just a written paper or a functional prototype), the next step would be to reflect on who might be interested in the results, who might benefit from a training, and what type of background or what discipline would be required to be able to understand the result. This leads then to the next question to be considered, that of the target audience.

3.2 Getting it right: the target audience

Looking at the wide range of disciplines and results types, we identified the following potential main target groups²:

- Practitioners (professionals involved in preservation, archiving, and data management – across different domains)
- Policy-makers (decision-makers in organisations with digital repositories)

² We refer to the study done in 2014, the results of which are described in D7.1 Report on Training Needs, see <http://pericles-project.eu/deliverables/46>

- IT developers (from the repository organisations, from business organisations or from technical institutions involved in digital data management and processing, or interested in tool development for the preservation community).
- Researchers (PhD candidates) in library science, computer science, linguistic, mathematics, ontology modelling
- and of course: trainers and educators for above groups.

With such a wide array of audiences, we looked into how we should optimally approach the question of which topic and result we should prepare for which audience, and how to best package the respective subject matters.

Clearly, there are more theoretical aspects that would ultimately impact future practices in repository institutions but at this point would be better understood by those reflecting on methodological frameworks of preservation and data stewardship. These results would best be targeted at the respective PhD candidates and researchers where a cross-over into another discipline might fuel their own research with new insights, such as bringing notions of ontology modelling to the library scientists. On the other hand, tools in support of workflows in the daily practice of dealing with digital data would benefit practitioners and the practitioners-to-be most. Some more theoretical aspects that relate strongly to management of data on a higher management level might be addressed towards researchers from the preservation field who often themselves are decision-makers and influencers. While tools that would need to be developed further to be ready for implementation could be introduced to technical staff from the repository organisations and the businesses catering for these organisations.

We will describe in the next step the conceptual choices we made for the different topics and target groups.

3.3 Finding the right mode of translation

The two areas to look into for training activities would concern creation of material such as documents, slides and screencasts, and direct interaction with the audience in webinars and workshops.

As with the audiences, we examined in the first half of the project the range of potentially suitable formats of knowledge transfer. This implies taking into account not only what topic for which audience, but what type of expression suits best a topic, what means of communication will reach the right audience at the intended level of engagement.

These considerations would take into account opportunistic scheduling of events within the framework of larger events, and common practices within the different communities. Preservation practitioners for example are strongly supported as a community and understand themselves as such. There are platforms catering for them, which we could use. Workshops at preservation related conferences are much more practical than workshops at scientific conferences. Webinars are common practice in preservation communities. They lend themselves as format for discussing ideas and topics or for demonstrating end-user tools. But there are topics such as ontologies that require other formats as webinars would not offer enough space for this type of knowledge transfer.

For the design of training material, our efforts were inspired by the concept of “modularity” which has received much attention in the fields such as management, engineering, innovation where

modularity has been proposed as a powerful means of managing complexity³. Though used very differently in the industry sector, the modular concept has been a source of inspiration in the design of our material, as it allowed us to provide the material in a single structure that offered visibility of the target group addressed and the knowledge expected for using the material.

3.4 All in good time

The challenge of timing was more difficult to “tame”. Organising training activities and producing training material are time-consuming endeavours that involve a process with different steps that often cannot be run in parallel. There has to be a topic or a tool at a mature stage before training instructions can be written or before it can be converted to a course subject with specific related questions and exercises.

Though for project results there is a schedule defined by the work description, one cannot easily infer from them a starting date for the production of the training material, as the topic or the knowledge first needs to be understood before it can be converted to some form of training.

Therefore, it was paramount to try and capture as much knowledge and understanding as possible during the development of the results through early and constant exchange with the producers or researchers in the consortium. The reason is that, in practice, those involved in research are typically not the same persons as those who have a larger role in the training engagements.

As most tools suitable for training can by the very nature of the project dynamics be only available in the last part of a project, it is important to decide early on how to most effectively invest the often limited resources or availability of the RTD teams for the training tasks. The reason is that their aim is to provide as much quality for their RTD results as possible by the final deadline. So it comes down to a question of priority.

In addition to the actual training activities, two more tasks come into play when planning in detail the work package: dissemination in terms of advertising the activities during the project, and sustainability in terms of creating an awareness of where to find training material that would be useful also beyond the project lifetime.

3 Baldwin, C. Y., & Clark, K. B. (1997). Managing in an Age of Modularity. *Harvard Business Review*, 75 (5 (September-October 1997)), pp.84-93. Bask, A., Hsuan, J., Rajahonka, M., & Tinnilä, M. (2016). *Configuring product modularity and service modularity for mass customization strategies*. Paper presented at the World P&OM Conference, Havana, Cuba. Hsuan, J., Frandsen, T., & Raja, J. Z. (2016). *The impact of product and service modularity on business performance – A survey of Danish manufacturers*. Paper presented at the World P&OM conference, Havana, Cuba.

4 Design

Pedagogical considerations are naturally of prime importance in designing training activities and material.

We took a look at the following criteria which are amongst other commonly considered for the design of training resources:

- to be informative
- to be relevant
- to be engaging and interesting
- to be clear and intelligible
- to be combinable with other material

We will expand in the following on these criteria and how we dealt with them in the context of an RTD project.

Information level

The issue here is to decide the level of presupposed knowledge, i.e. information that is already available for the intended user, or if not, can be easily obtained elsewhere. This involves making a decision about how much context could and should be provided in addition to the core information and about accepting that this context has been well provided in other existing training material.

We initiated a survey early on in the project with representatives of the preservation community to help us understand what to provide. The questions we asked referred to the first tool to be delivered: the PERICLES Extraction Tool, short the “PET”, a tool that automatically captures environment information, based on scenarios to be determined by the user. The scenario defines what parts are considered to be significant in the eye of the user.

In PERICLES, we treat environment information as a type of metadata. Then the question for us was to determine what other information we needed to provide beyond the standard demonstration of usage for a tool. In the survey, we had those who knew much about metadata and those who asked for a basic introduction to the concept before moving on to explaining the PET. We needed to decide whether to burden one part of the audience with knowledge they already had, or to risk losing an audience due to lack of prerequisite knowledge necessary to understand the training subject.

Naturally, as an RTD project does not have training as its core remit, constraints as mentioned above in terms of existing expertise and available resources come into play here. Even mere provision of a full list of training material on “metadata” requires a lot of desk research.

The focus therefore was rather on what we could reasonably include. The answer to that question in our case was: provide modular elements so that those with knowledge of metadata could skip introductory elements, while those not too familiar could be made to understand.

Relevance

The question of relevance addresses different aspects. For us, it tied in to the above question of how much context to provide. With the example of the PET, we included explanations of metadata only as far as it paved the way to introducing the notion of “significant environment metadata”. Our decision was based on including context that was relevant to understanding the core knowledge we wanted to provide.

The other aspect of “relevance” is that of pertinence for the audience addressed. With the PET it seemed feasible to address practitioners, i.e. all persons dealing with or creating digital data. An example for a result addressing another type of audience would be back-end and middleware components. An end user will never be directly confronted with these components, therefore the target audience would be IT developers. However, the clear objective of reaching out to IT developers is for them to carry on the work we started, adapt it to their systems, enhance it and provide market maturity. These groups, however, are interested in an introduction to the functionality of the component, so that they can proceed with the installation guidelines and access to the code. This is not what we have classed as “training”.

For this reason, we did not include results of this nature. However, as IT developers will only develop technical solutions in response to an end user requirement, it makes sense to provide some information of the usefulness of the component in ways that the end user can understand. Then this becomes a form of “pedagogical communication” and part of our dissemination efforts in WP8; we did this, for example, with the tool PROPhet,⁴ or the ERMR.⁵ But we did not attempt to create a training module for these components.

Another example of deciding what training might be relevant for whom, are the theoretical or abstract results, such as for example, models as driving force of management systems. How to use models to capture knowledge, to make complex analysis and provide meaningful answers might not have been at the top list of interests, for example, of library scientists. The fact that models are now being considered as a feasible way to deal with complexities in the digital world is commonplace in many disciplines, but have yet to reach others. This is an area of knowledge that could be of benefit from other disciplines, for example to library science students attending a course on semantic ontologies, as we have successfully proved. Hence the importance here will be how to entice the target audience to acquire this new knowledge.

Engaging and interesting

Making the activities and resources engaging and interesting for learners and participants is naturally a core success criteria. The aspects to be taken into account are:

- a) opportunities for interaction with other learners/participants
- b) hands-on application or exercises to embed the knowledge in practical scenarios
- c) variety of media
- d) involvement of the participants through inviting them to share their knowledge and experience
- e) share responsibility for the knowledge transfer by prompting the participants/learners to reflect on the relevance of the new knowledge for themselves.

Though we were limited in our use of formats, with the range of activities we did carry out were able to cover all these aspects. The workshops all covered a) and b) and c), the surveys provided e) even if not carried as far as we could have in a larger training framework (such as a semester seminar or a training camp). The webinars allowed for d) and with the PMTP training material we covered c).

Clear and intelligible

Producers of training material must have an understanding of the subject and of the expertise of the target audience. In distinct environments, this is a given: take a university where the instructor is an expert on the subject and the participants are aspiring academics with a wish to develop academic

4 <https://youtu.be/9oLBwkr-GOc>

5 <https://youtu.be/N1NYi68qsWI>

thinking and working of specific subjects. Depending on the programme level the instructor will know what standard of knowledge is to be expected.

In contrast, an RTD project with no core remit in instruction will not be able to cover thoroughly the topic and provide appropriate training environment to suit all audiences.

The important issue here is the “translation”. The gap in the communication between IT developers and scientists and professional with no special IT background has been an emerging topic in the past years triggered by the pervasiveness of IT challenges in all parts of social and professional lives.⁶ There are efforts to find ways to translate IT knowledge into words that a person with no technical or scientific (in terms of natural sciences) background would understand.⁷

Often the training team of an RTD project is at the intersection of both groups. As mentioned in the introduction, it is paramount for exactly this reason that the training team closely follows the discussions in the RTD teams and talk over aspects they do not understand with the experts. The training team needs to understand the gist of the tools to be able to translate it into an interesting and intelligible form of communication. It needs to be packaged for a relevant target interested in training as a form of engagement and in gaining new knowledge.

Combinable

Although we did provide training on specific tools as stand-alone modules (be it in the form of workshops, webinars or training modules), it is unlikely that outside the PERICLES project someone will train on e.g. only the PET tool. The training package on the PET could serve as a communication support between the IT developer interested in adapting it and the practitioner needing to understand and decide on its usefulness for his work. But as a training resource, it is more interesting as a part of a wider training context. One could imagine that a trainer or a teacher researches material on metadata, finds our package, likes it and uses it as part of e.g. a seminar on metadata and its scope.

In this spirit, we discovered that we could provide valuable material for trainers, that in itself might not be useful as stand-alone training sessions, but an interesting use case for a specific topic area. We included two use case modules for exactly this purpose: “Space Data Specificities and Lifecycle” and “Recent Developments in the Conservation of Digital Art”.⁸

The modularity of most of the training packages fosters the inclusion into other training activities; the trainer/teacher can decide which part of the modules would fit into their own syllabus.

⁶ This was also a recurrent topic in the PERICLES conference “Acting on Change” <http://pericles-project.eu/page/PERICLESconference2016>

⁷ for example the initiatives in the field of „Public engagement with science“ <https://www.publicengagement.ac.uk/science-engagement>

⁸ The latter is planned to be a video which at the time of writing this deliverable was still in the final production phase and will possibly be uploaded both on the project website and on the TATE website end of March or beginning of April.

5 Reflections on questions of impact and success measurement

Measuring impact, though there is general consensus that it is good practice, is not an easy undertaking, in particular if the activities of which the impact is to be measured are taking place late in the project, not allowing for enough time to receive significant feedback. There are two approaches that can feasibly be considered: one would be using platforms or designs that have proven successful, and through them have reason to expect a long-term impact, or conjecture on the success of the activities carried out during the project, most of them towards the end of the project as a measure for anticipated future success.

The actual impact of training would be to foster knowledge, to enhance expertise and understanding of aspects of preservation planning and management (as discussed and proposed by), and to influence practices towards improved workflows. To measure this, we would need to engage in long-term observation and interaction with the groups that use our modules for training and attended the workshop. If we can hope for that kind of impact beyond the project's lifetime, then it is important that the communities are aware of the training material in the first place.

We therefore strongly recommend (something we ourselves only understood via hindsight), that the training activities and material have a dissemination plan in place early in the project. This can be prepared without waiting for the results to have achieved a maturity level suitable to starting training activities. With the little time left for impact assessment, it is important to ensure as widespread as possible an awareness of training activities and availability of training material (beyond the end of the project).

An example would be to make use of the potential of social media. Suppose that an event could be used to encourage participants to tweet about it, with a specific # leading to the material used or a location of such material, and that some research was done on influential organisations that could be used for re-tweeting and following. Then an analysis of the tweets could show areas where more awareness and interest could be created, lead to a better understanding of the direction of future efforts.

In PERICLES, the PMTP page was mentioned wherever we had an opportunity to do so. We also produced leaflets for distribution at the events.

Extrapolate impact from participants' feedback

We found that the most useful feedback was received in follow-up conversations while written responses remained limited throughout. The development of questionnaires as a feedback tool was valuable. We recommend to always follow-up on training events with a feedback communication.

In addition, another recommended practice would be to have a written report on conversations or questions asked during and after the workshop, often during coffee breaks, as many participants appeared reluctant to speak up in front of the group. Impressions by the organising team, note-takers, moderators and presenters are always helpful to annotate questions and give input for improvement.

It should be noted that feedback received in connection to events, depending on the questions asked, can only provide indications about the immediate usability and envisaged future impact.

Another form of feedback would be the tweets by attendees of an event.

Ensuring impact by proven design and long-term platforms

We purposefully sought and followed proven concepts that had led to broad impact. One such example is a set of modular training material that the Swedish National Agency for Education had produced⁹. These are broadly used by educators throughout Sweden affecting their work and practices. We were inspired by the design ideology of that collection and with their permission adapted their concept to suit our needs in the creation of the PERICLES Modular Training Packages PMTP.

Another way to ensure impact is to collaborate with platforms and key stakeholders that increase visibility beyond the end of the projects, for example working with the Digital Preservation Coalition and Open Preservation Foundation, and offering webinars for their members. The webinar was recorded and is part of the resources offered through DPC on their website.

Impact and the question of legacy

As the given time constraints to properly measure or even monitor impact, the two “sister” activities of training gain in importance. Just as mentioned above, in addition to project results dissemination, a proper dissemination of the training material needs to assure wide awareness, if the material is to make an impact. But the question of this material as project legacy needs to be addressed as part of the “exploitation” planning. Knowing that project websites quickly become obsolete, even if kept alive, the result is that they are rarely used because they seem to be related to a past, without updates. They naturally fall into oblivion, although the content available may still be useful and valuable.

The optimal way would be to include the material into a wider context and an active environment, which can continuously attract users. In this context the question of “past or present” will not be asked, instead, the question will become: “is this content of use and interest to me?”.

The PERICLES partner Högskolan i Borås has agreed to extend the life of PERICLES modular training packages by hosting the collection on their server¹⁰. A local HB website for PERICLES will be created which will also provide links to other PERICLES-related resources such as publications and more. It is intended that the PhD course “Dynamics of Knowledge Organisation” (see below for more information) developed within the project will be offered again in an improved format. The local HB-PERICLES webpage will be a hub that will continue to connect the PhD course, the PMTP, and potentially the work of the new PhD in Knowledge Organisation (to be appointed in 2017).

After the first run of the MOOC¹¹ adapted from the module and the PhD course “Dynamics of Knowledge Organisation”, the feasibility and desirability of a repeat run will be investigated. By all these plans, we are confident that our training work package output will have a life beyond the end of the project. In addition to these efforts by HB, further training material directly linked to project results on the hub (such as the PET), will also be hosted on the PRESERVEWARE portal.¹² This portal is conceived as a digital preservation hub hosting preservation tools (not only from PERICLES but from other providers and projects) and contextual documents such as publications, guidelines and training material.

To ensure that material can be found, it makes sense to have distributed access to the material: the HB site and the project website for the PMTP and PRESERVEWARE for individual training elements or modules linked to the tools.

⁹ as part of the competence and training programme Läslyftet developed by Skolverket (the Swedish National Agency for Education - <https://www.skolverket.se/om-skolverket/andra-sprak/in-english>).

¹⁰ <http://www.hb.se/PERICLES>

¹¹ on Udemy <https://www.udemy.com/dynamics-of-knowledge-organisation/>

¹² <http://www.presereware.com/>

6 Lessons learnt: revisiting decisions and outcome

In the following we will revisit decisions made and assess our activities and their success taking into account the above described considerations and recommendations.

6.1 The planning

Planning of training activities is to be carried out as an iterative action as they strongly depend on other developments and plans in the project on which the training team has little or no influence. The first planning phase happens in the project submission phase, when the application is written, which, in case of acceptance, will become an integral part of the grant agreement. In RTD projects, the key actors in the innovative part of the proposed project typically are the drivers of the application description. At this point the necessary discussions and investigations informing a training plan will not yet have taken place. This bears the risk that communication work packages descriptions rely on assumptions as to target groups and topics. It is strongly recommended that these assumptions are re-visited and verified at the beginning of the project in discussions with the RTD groups.

In the case of PERICLES, it was proposed to carry out a survey for training requirements, followed by a plan that would determine, describe and schedule the training activities. The plan was to provide for event type activities, while in parallel training material would be produced. And all would culminate in a big event; in our case we planned for a summer school. This was motivated by the pedagogical expertise in the project from the fields of library science, and training of academics and practitioners in the framework of “nestor summer schools”.¹³ During the course of the project, it became obvious that only parts of the results were suitable for training practitioners. Some of the results required knowledge of the theoretical discourse and approaches in research on preservation solutions.

The plan was therefore revised as the conditions and the parameters for designing training were better understood. In the second phase, the focus moved to better understanding the results and to finding which target group would benefit from them. In particular the challenge of the diverse disciplines, the different levels of prerequisite knowledge, and the question of which means lent themselves most suitably to transport the content of the training became a guiding principle in this second phase. This led to the modular approach reflected in the PMTP.

The plan for events was kept flexible, to accommodate an opportunistic approach, as events were taking place in existing frameworks, and therefore depended on acceptance of proposals. Of course this implied that a result would provide enough stable information at the time of the submission of a call, which was often months before the actual event and the delivery of the result. Another original plan had been to organise events at the use case partner organisations. This proved more difficult, as the topics were very theoretical, and of more interest to the partners with staff in preservation, rather than the partners where preservation was not part of their defined job profiles. We therefore quickly turned to seeking opportunities for external workshops.

A next iteration of the plan saw the re-consideration of the original idea of a summer school as the culminating event. Though the “nestor school” took place as envisaged (15-17 February, 2016) with

¹³ https://nestor.sub.uni-goettingen.de/school_2016/index.php

PERICLES presenting those tools that had end-user interfaces and maturity level in hands-on exercises for German archive and preservation students, academics and practitioners, it was decided to dedicate a big event to the IT developers and research-practitioners of that community.

All knowledge transfer activities (WP7-9) required close collaboration, to avoid redundancy and gaps, but in particular to align and coordinate activities. Therefore, it was decided that the three teams should join forces in organising the final event.

6.2 The survey

The first element in the training work plan was to carry out a survey. A survey can have a multitude of facets. If it is to be more than a quick anonymous online survey with a simple questionnaire and no real follow-up on the responses, then it might be a useful supporting measure. If done intensively, with interviews for example, then this can be extremely time-intensive. The question remains: what does one expect from the survey, and is it worth the investment?

Often a survey is a means to validate one's assumptions and results. And just as often, a survey serves as a decision-making tool.

In our case, we thought a survey might be helpful in giving us an idea where to start with the production of training material and activities. This endeavour is a good example of how salient it is to first really be clear on the target groups, and therefore have a good idea of the nature of the anticipated training subjects, the scope of expertise available for understanding what input can be expected from the RTD team and needs to be provided on top (such as introductions to a topic field as metadata). Although the planning included a whole list of different potential target groups, the survey focused on what was believed to be the main target group: the practitioners in the field of digital repository management. It was only through engaging with this target audience that we have come to understand that we needed to revisit our initial assumptions of how we should best target to communicate our results.

Though the survey results were ultimately less helpful than expected, it was a valuable way to force us to reflect on the available means of communication/training and on the level of knowledge of the practitioners, as one of our target groups. , It led us to intensive discussions and questioning among the consortium partners.

6.3 Available expertise and resources

After having reflected on questions of target groups, knowledge levels, and options of communicating different types of results, we turned to the RTD teams to assess the existing experience that would help us create training material. Questions on how confident people were to present in public, on experience in teaching or training, (which implied "translating for a non expert audience"), on opportunities to create material for training (e.g. screencasts) were discussed with the teams. In principle, the teams were willing to help and provide input. Yet, there was the question of the "gap to be filled" between input and translation for the target audience. Some could give input in the form of explaining to us what the result or research was about, others could provide presentations or screencasts, while again others were experienced and confident enough for example to be video-filmed for video lessons. The lesson learnt here was that this type of contribution could not be taken for granted, as training is not a mandatory profile required of RTD professionals.

In PERICLES we solved the problem by establishing a matrix that would involve all the criteria explained above, and combine them with available time and resources and expertise. This meant that we had to understand which results would make sense, by seeking answers to questions such as:

- Is training the appropriate format of communication (e.g. IT developers might just need installation instructions and the source code, and would not want to waste time on “training”)?
- Is the result of interest to any of the groups principally open to the anticipated training format (e.g. would middleware be of interest to anyone apart from IT developers)?
- Is the result clear (i.e. is the input from the RTD teams clear enough in itself to be used directly or was the input at least clear enough to the training team to be able to act as translator)?
- Is there enough material available to create training material from?

With this matrix, based on the study of existing documentation produced by the RTD teams (such as papers, posters, and deliverables), and discussions with members of those teams, we established a list of topics for the PMTP and a schedule that we could use in our monitoring call deadlines for interesting conferences.

One factor that we need to mention here, is the aspect of prioritisation which is difficult to answer generally. There are certainly endless means of production and design of training resources. Within a project one might need to prioritise which activities are more valuable. It comes down not only to quantity versus quality, but also to immediate response versus long-time availability. As an example: if time were limited one might need to decide whether to invest time in 1-2 live workshops at conferences or a MOOC (that might be reusable and even if not, might have reached a lot more people than the 1-2 workshops). The difference is that with the workshop, there can be direct feedback for the RTD teams to take back, while the MOOC can provide feedback on the training but analysing the different comments will take time for it to be fed back as input to the RTD teams. Another consideration for how to weigh different options in terms of valuable feedback concerns the investment in workshops or webinars and training documents. In-person workshops with RTD teams would entail a more immediate impact on the RTD work than training documents for self-learning, while written material could provide a more long-lasting impact to a wider circle of people than a workshop or webinar

Sometimes, the DoW is based on high expectations with regard to available expertise in the RTD teams, based on their affiliation. The fact that many of universities are involved in research projects does not imply that individual researchers or software engineers and developers working in centres affiliated to universities or libraries, automatically have pedagogical experience or skills. It is recommendable to take up the question at the beginning of the project, ideally during the submission phase, but at least as part of the first iteration of planning the scope of activities.

In PERICLES we investigated what type of skills and experience was existent. One result was the close collaboration between HB and CERTH on the topic of semantic shift, and this tied in with domain ontologies, also an expertise from CERTH. Ontologies, being at the heart of the model-driven approach, were difficult to translate into “hands-on” workshops or self-learning material for practitioners. But they were relevant for library scientists. Video recordings created as part of this collaboration resulted in an excellent series of seminars conveying these highly theoretical concepts for students in a very lively and attractive video and online PhD course (see also next sections). For the tools developed by IT experts, the team of UGOE had a pedagogics expert involved in the nestor summer school, who helped translate tangible results for training purposes. Video skills were also existent and employed to introduce middleware in an entertaining way, which made the purpose of tools such as PROPhet or ERMR easy to understand for non-technical people. In addition, KCL provided a many “translations” of the knowledge assembled in the diverse documents produced as reports or presentations to a peer audience, to suit the needs of training.

6.4 Training activities

In the following we present the actual activities carried out, where possible grouped by the nature of the activity. We do not elaborate on each event. For more information, we refer to our project website: <http://pericles-project.eu>

6.4.1 *In person interaction*

In person workshops are a very effective form of engaging with the target audience, allowing for rich interaction, group work, hands-on demonstrations, exchanging knowledge and sharing experiences, and receiving direct feedback. They also allow for presentations of interim results.

Amongst the issues to consider was early timing, as calls for workshop submission in the framework of conferences are published often many months before the actual event takes place. Submission might have to “take a gamble” by anticipating that by the time a detailed programme is published, for which the project needs to have provided enough material to satisfy the created expectations. The proposal would have to be as precise as possible, in anticipation of the outcome of the research work. Unless one anticipated a tool being mature enough for a hands-on exercise, aiming for the discussion of an interim result was the basis on which to submit a proposal under controlled risk.

The calls for events needed to be closely watched. In PERICLES, the dissemination team kept a file on the project’s Google drive listing all published conferences and deadlines for calls that included topics we had been working on.

Not all workshops we provided were training workshops. Amongst those that we considered as such were the three IDCC workshops¹⁴: in 2015 on the PET tool, in 2016 on policy modelling and risk appraisal, and in 2017 on the Appraisal, EcoBuilder and Process compiler tools. Each of these included interactive group work. The ones in 2015 and in 2017 required that tools were ready for demonstration and use for hands-on exercises. In 2016, we used paper exercises as a basis for discussion of work in progress (on our appraisal tool and policy modelling), and exchange of respective experience among the attendees.

In a similar fashion, the PET and PeriCAT tools formed the topic of a seminar at the “nestor” winter school which was attended by students, researchers, representatives from industry, technical staff from archives, museums and libraries.

Another type of workshop was demonstration. These were events with not much emphasis on hands-on work, but which invited more discussion and interaction via feedback after intensive demonstration sessions. This we did for example within the framework of the Swedish Book Fair, where professionals dealing with repositories, not only from the memory institutions, but also from domains such as health, attended and engaged with vividly¹⁵.

Besides taking up the opportunities offered by the organisers of conferences, symposia and other events, there was, of course, the option to organise a stand-alone workshop. Here the pros and cons needed to be carefully weighed. While conferences usually attract an audience that can be a pool to draw upon for an integrated workshop, stand-alone workshops need to offer something attractive enough for professionals to invest their often very limited time. Apart from the time investment, there is also the question of justifying the financial investment of travelling to such a seminar or workshop.

14 http://pericles-project.eu/uploads/files/PERICLES_IDCC2015_workshop_programme.pdf, http://pericles-project.eu/uploads/files/PERICLES_Workshop_IDCC16_programme.pdf, http://pericles-project.eu/uploads/files/PERICLES_Workshop_IDCC17_programme.pdf

15 <http://pericles-project.eu/events/68>

We solved this issue in three different ways:

- Though we invited globally, we expected attendance mainly from local people. This was the case with a workshop we did at HB on “From Semantics of Change to Change of Semantics”¹⁶. The widespread dissemination was done within the university itself and the near-by University of Gothenburg. For those who were not local we invited specifically people that we knew had an interest in long-term preservation. The workshop was a set of presentations in a theatre setting, with an ensuing discussion with the audience in a seminar room. It attracted academics in the fields of library and information science and informatics, as well as data service agencies and data provision organisations.
- In two cases, we invited targeted experts to present the more complex aspects of our research: the modelling and the question of re-use of data. We needed feedback from the memory institutions and the scientific community. But this involved bringing them quickly up to speed to where we stood with our research. A mix of presentations, group work and ensuing discussions, hence a typical training type of event, seemed the best way to convey the core aspects in a condensed manner. These activities were also our test grounds in finding ways to communicate our research findings to the broader target audiences represented by such experts. It was a valuable exercise; it showed us how difficult it was to translate some of the more theoretical concepts behind the PERICLES approach, and helped us refine our training and presentation material so as to better reach these audiences.
- The third means was to create a big event that would attract people by promising enough information to justify their time and travel expense: this was the PERICLES conference. One part of the conference was dedicated to training-type workshops that would combine live demonstrations of tools with immediate feedback and discussion.

Our experience showed that the two approaches, the one making use of conferences and similar events and the other specifically inviting key actors in the field for a stand-alone workshop were the most successful in implementing knowledge about what PERICLES is about. In the former, one had a limited influence of who would attend other than the broader community addressed by the conference itself, i.e. students, practitioners, technical or non-technical persons, etc. In the latter approach we could target exactly the persons we thought would be most beneficial to us in terms of feedback, advice, and multiplicative effects on information.

6.4.2 Live online formats

We experimented also with live online formats: webinars and online (distant learning/education) seminars.

Webinars were carried out with the Digital Preservation Coalition and the Open Preservation Foundation. The difference between these webinars and the workshops described above was certainly the time constraints to present and discuss our results. Most public webinars (as opposed to academic online seminars) have a time span of one hour, which means with all the paraphernalia that it leaves only 30 minutes presentation time. The notable difference is the lack of group work or “exercise” with a specific task. While the anonymity of the online room would suggest less inhibition to talk, there is a tendency to use the chat room rather than the microphone, and there is a risk of cross-communication. Discussion tends to need some time before it gets going, but one cannot plan an open-ended webinar and play it by ear.

In person workshops, due to their length and the physical presence, allow much more flexibility in guiding the discussion and agenda and in understanding the audience and their requirements or

¹⁶ <http://pericles-project.eu/events/27>

interests. In PERICLES, the real advantage was that the webinars were performed using well-accepted platforms and thus reached the community through trusted channels. Also in the case of the DPC, the recording of the webinar became part of the archived resources that can be accessed by their members. Hence the immediate impact of the webinars seemed less strong than in the workshops, though the long-term availability on the initiatives' websites provides the potential of sustainability.

An **online PhD course** on “Dynamics of Knowledge Organisation” was created in the form of a 7.5 ECTS credit doctoral course that allowed students from several European countries to attend without the need to travel. This format overcame the above described constraints of webinars on intensive discussion found in online formats, as its duration was much longer.

The PhD course had a plan that defined the area of the course, the prerequisites for attendance, the number of credits awarded, the expected learning outcomes, the content of the course and modes of teaching and examination.

The format offered opportunities for participants to ask questions addressed to fellow participants, facilitators, and, most importantly, researchers who developed the PERICLES results around which the course revolved.

By time these courses were delivered, some sections of the project had already become institutionalised as parts of an established educational programme. It would also be safe to say that the integration of some of the project's findings in a number of master's courses and the PhD course are expected to have an impact on students' academic and research lives and their future research. This is also a manifestation of the impact and relevance of PERICLES in shaping future research directions in the area.

6.4.3 Screened presentations

The format closest to live formats is screened material: videos, screencasts, MOOC, filmed lectures and live streams. The advantage is that one can target the audience independently of a specific date and therefore repeatedly over a longer period of time. The audience can flexibly choose the time or frequency of viewing, as no physical attendance is necessary. For those reasons, significantly broader audiences can be targeted.

We managed to produce screened presentations in all variants.

Screencasts. For end-user tools such as PET, PeriCAT, EcoBuilder, Policy Editor and Appraisal we provided screencasts as part of the PMTP modules. The screencasts of the open source tools were posted also on YouTube. They have a demonstration-type of orientation.

Video as a format was used for results that demanded a presentation-type or lecture-type of training. This we applied for the more theoretical results presented in the modules “Dynamics of Knowledge Organisation” and “Ontology engineering and Linked Data”.

MOOC: We also experimented with the format of a MOOC that would allow intuitive and somewhat interactive learning for a very large number of students from all parts of the world, ranging from a few thousands to a few hundreds, depending on the specificity of the topic. In a MOOC course, for instance, participants will be able to post discussion comments that draw their own views and experiences of digital preservation into their conversations with other learners.

6.4.4 The PMTP

The PERICLES Modular Training Package is the repository for the main activities that are not dependent on a live intervention from the PERICLES team. It combines many of the above listed formats with textual formats such as blogs, articles and slides.

The modular structure includes clear indications regarding the target audience, the expected learning outcome, the required time for completion, the level of advancement and moreover also allows to assemble different result types in one repository. It provides flexibility and openness to allow participants to make choices based on their own interests and needs.

In addition to the results already mentioned above, we have more theoretical modules such as “Contextualising Semantics” which explains and contextualises the middleware tools PeriCode and PROPhet rather than providing a “hands-on” demo (which is not possible or wouldn’t make sense).

As OAIS and the continuum theory are both important factors in the PERICLES approach, we were happy to include lectures from our partner University of Edinburgh and the video-recorded introduction to the continuum model by Barbara Reed.

Overall, the PMTP is the most important legacy of the training material. The modularity also allows for distributed dissemination, which is a great advantage beyond the project.

7 Conclusion

The immediate recommendation that all knowledge transfer work packages would give after a project is “more time”. But unless this is arranged for in the programme schemes, there are several recommendations towards capitalising on the available assets and means.

The first would be to carefully consider what to formulate in the description of work, which becomes the legal framework to which then the work performance is tied. In simple terms: don’t make concrete sounding promises just because they look good in an application.

The next recommendation would be: there is a lot of work one can do while waiting for the content of training (or any other knowledge transfer for that matter) to be developed, such as researching institutions and people to contact, platforms to post on, formats of dissemination, discovering opportunities for workshops, webinars etc., training programmes to contact, designs that would economise on the time for producing material and optimise the input (e.g. adopt same input for different activities, reuse input in different formats etc.), and reflect on different ways to quickly capture feedback and perhaps measure impact.

In parallel to that, a recommendation is to stay in close contact with the RTD teams to be able to understand the gist of the upcoming results and identify the adequate target audience. This then allows for preparation such as researching how to be reach and advertise workshops in the right community channels, (specific listservs, platforms announcing events, event calendars, who to talk to for word by mouth dissemination). Presence in meetings also allows to understand what types of format might be appropriate for which result, and discuss potential input from researchers early on so they can include it in their scheduled efforts.

Ultimately, the best recipe seems to be to have as much personal interaction as possible (combined with similar dissemination activities) that will stabilise the awareness, then assure quality products such as video lectures, demos, texts, slides, MOOCs, for use without the personal interaction by the project members, and ways for these material to be easily found during and after the project.

To repeat a statement made at the beginning: within the three knowledge-transfer work packages (dissemination-training-exploitation), training is the closest form of engagement with the stakeholder. The term “training” is perhaps not the best choice to denote this kind of engagement, and the broader use of the term will always seem to require a justification, while the more traditional narrow use will forfeit opportunities of the direct engagement. H2020 with their notion of “communication” have taken up on an ambiguity amongst the knowledge transfer packages, which indicates that there is a great potential for making things clearer, but also for strengthening the remit of these work packages within the project planning.