

Proposal for piloting an “Open Educational Resource portal service” for the R&E community

Date: 14 February 2014

Author: Peter Szegedi (TERENA)

Contributors: Ilias Hatzakis (GRNET), Eli Shmueli (IUCC), Vicente Goyanes (UVigo), Antonio Vieira de Castro (ISEP)

Contents

1. Introduction and motivations	1
2. Aim of the TERENA small project	2
3. Technical details.....	3
4. Delivering the pilot.....	6
4.1. Participation.....	6
4.2. Timelines and costs.....	7
4.3. Deliverables.....	8
5. Opportunities and Risks	8
References	9

1. Introduction and motivations

The TERENA task force TF-Media (2010-2013) concluded with a project plan to try and implement a European-level OER metadata repository service for the benefit of the Research and Education community gathered under TERENA/GÉANT. The fundamental principles of such a platform/service have been discussed and summarised by the task force. For details, please visit [1]:

<https://confluence.terena.org/pages/viewpage.action?pageId=33751325>

*By definition, **OER (Open Educational Resources)** are teaching, learning, and research objects that reside in the public domain or have been released under an intellectual property license that permits their free use and re-purposing by others. Open educational resources include full courses, course materials, modules, textbooks, streaming videos, tests, software, and any other tools, materials, or techniques used to support access to knowledge.*

There is a large interest around the global education community in establishing and maintaining OER or Learning Object (LO) repositories as exemplified by the number of existing repositories (e.g., MERLOT [2], MAOR [3], OER commons [4], Learning Resource Exchange for Schools from European Schoolnet [5]), organizations building and sustaining them (e.g., MITOpenCourseWare [6]), contributors integrating learning objects in repositories (e.g., OpenContent [7]), and users of these learning objects (e.g., Universities, Libraries). The fundamental reasons are:

- the growing educational demands in all countries,
- the limited capacity of face to face education to fulfil the demand in a timely manner (i.e. emerging MOOCs),
- the effort and cost involved to build multimedia learning materials, and the new possibilities offered by the Internet.

While it is a fact that millions of LO/OER can be found on the Internet using search engines like Google, there is no guarantee that a query will lead to trustable, properly licenced material on which high quality open education can be built. **Well-managed OER repositories that aggregate high quality content offer a solution to this problem.**

Many of the universities operate local content repositories (using community open-source or home-grown software tools) where they store recorded lectures, handbooks, presentations and other teaching-learning materials. In some countries, the NREN provides a centralise repository to universities where all the content can be aggregated to and made available for public or selected user groups. Some of the repositories only aggregates the metadata of the OER (i.e. the information about the learning object) and leaves the content in its originating domain (i.e. in the local content repository of the University). These repositories often called metadata repositories or simply referatories.

The main motivation for developing a metadata repository (European-level aggregation point or referatory) and an OER portal (federated single access web front-end) service would be to support the NRENs and their stakeholders (i.e. the broader TERENA/GÉANT Community) in engaging with open education by providing value-added support services.

The OER service intends to aggregate metadata (not the content) at the European-level and helps Universities and NRENs stepping to the next level (reach the critical mass e.g., in terms of the number of objects) towards exposing their OER to global repositories (such as GLOBE [9], for instance).

2. Aim of the TERENA small project

The primary aim of the TERENA small project is to develop the first working prototype of the OER service (including the metadata aggregation engine and the web portal front-end) and pilot a service for the broader TERENA/GÉANT Community in 2014. The pilot service can then be taken over by the GN4 project for further technical enhancement and service development aiming to roll out in production (Fig.1).

(Note that the New Idea Form to make the OER pilot service part of GÉANT production service portfolio has already been submitted, see attached.)

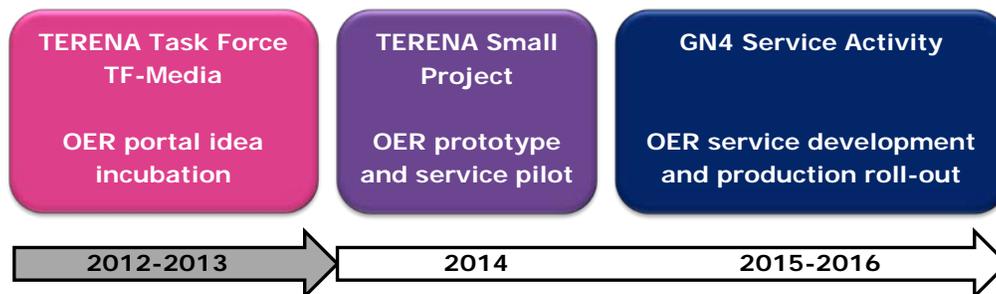


Fig.1 – Timeline of proposals

The TERENA small project is to bridge the gap between the end of TF-Media (now) and the beginning of GN4 (April 2015). The reason why the idea must be tested in a TERENA small project before it's introduced in GÉANT is the fact that the critical mass (in terms of participants, support, interest, etc.) has to be gained before any sustainable service development can be done. OER seems to be a typical “chicken-n-egg” problem at the moment (i.e. without a working prototype it's hard to gain significant interest and without significant interest it's hard to convince the development) therefore, the TERENA small project has to take this initiative. The OER is not the service that the NREN community is desperate to build (e.g., like the Trusted Cloud Drive pilot was in 2012) but it's something that TERENA has to take the lead on (e.g., like the NRENum.net service pilot was in 2008) in order to facilitate the development of future value-added services on top (including MOOCs and others).

3. Technical details

The longer term strategic objectives are as follows:

- Connect the yet scattered institutional/national OERs within the TERENA/GÉANT community (facilitate them to reach the critical mass) and unlock the deep-web by enabling structured searching and reuse of content.
- Create a one-stop-shop broker/aggregator (i.e. the TERENA/GÉANT OER) for institutional/national learning resource organizations, each of them managing and/or federating one or more learning object repositories within the country (such as Nordic OER alliance [14] for instance)
- Finally, TERENA/GÉANT OER joins the Open Course Ware Consortium [8] and/or the GLOBE federation [9] that already has repositories around the globe such as:
 - Education Services Australia
 - LACLO (The Latin American Community of Learning Objects)
 - MERLOT (Multimedia Educational Resources for Learning and Online teaching, USA)
 - OER Africa
 - MAOR- (The Israeli Metadata and Object Repository)
 - The Open University of Japan, Center of ICT and Distance Education (OUJ-CODE)

The architecture of the TERENA OER pilot system must be modular/layered (including the aggregation engine, metadata store, and web services) preferably combining the best of the existing tools such as ARIADNE [10], MAOR, PuMuKit [11], and others. TERENA OER must be a stand-alone OER system that does not use ARIADNE or MAOR services but builds on the (open-source) software components developed and made available by ARIADNE and MAOR projects. TF-Media concluded that in order to start implementing the TERENA OER pilot, a better/deeper understanding on the actual software tools and provided functions is needed. The following OER service-vision has been developed by UVigo:

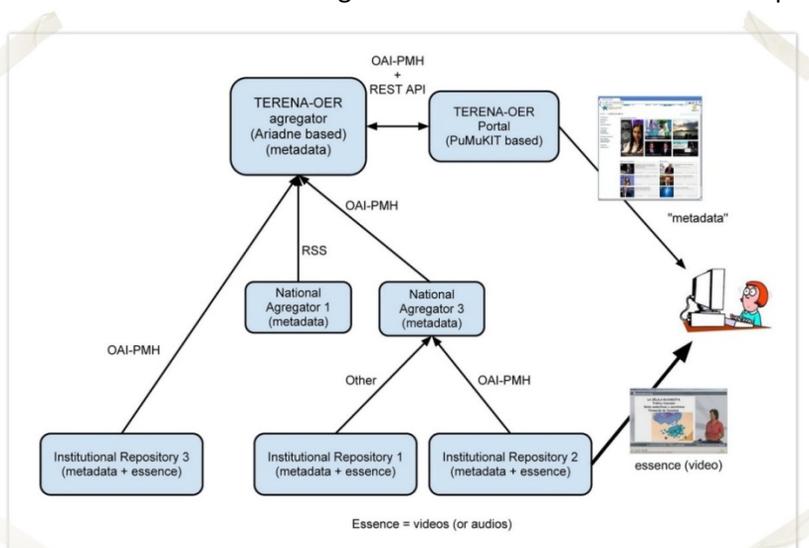


Fig.2 – High-level service architecture proposed by UVigo

The service architecture has further been detailed by GRNET proposing to use ARIADNE tools:

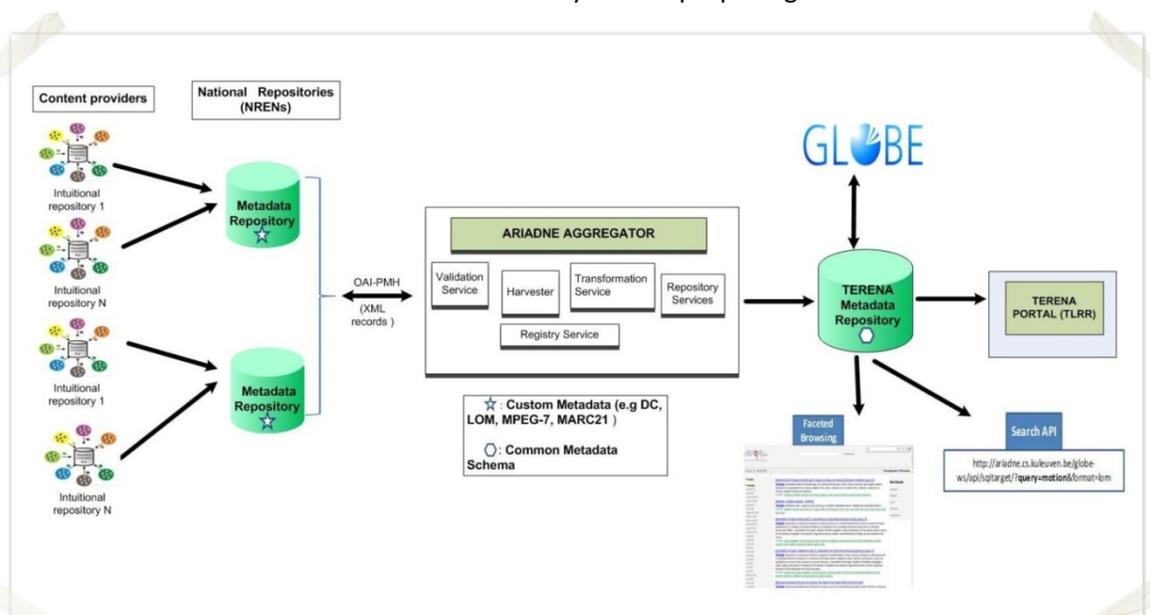


Fig.3 - Proposed infrastructure to collect the NRENs metadata and connect to GLOBE federation

As depicted in the diagram above, by using the ARIADNE infrastructure the metadata records provided (through an OAI-PMH target) by existing repositories could be validated, harvested, transformed (according to the defined common metadata schema) and published into a single repository over which the web portal front-end could be built.

A very-first-look of a potential web portal front-end has been mocked-up by UVigo:

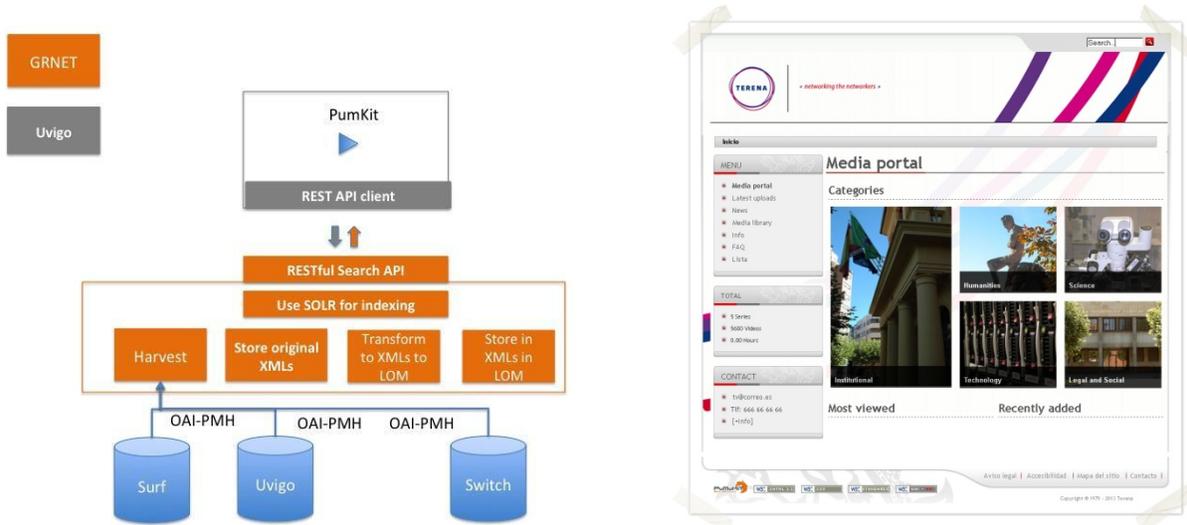


Fig.4 – Service components and web front-end mock-up

The key technical challenges have been identified as follows:

- Gradual approach to content type: The TERENA OER should focus on audio-visual content in the first place (having something catchy to show off). Later on, other type of materials such as data sets, papers, etc. can also be collected.
- Gradual approach to repositories: TERENA OER should focus on (work with) only 2-3 "friendly repositories" in the first place.
 - The following repositories have already been contacted and tested by TF-Media:

	Repository Name	Records Harvested	Metadata Schema
	collection.switch.ch	619	oai_dc
	dspace.iu.lv	1009	oai_dc
	repositorio.portalobaa.org	56	oai_dc
	riunet.upv.es	21902	oai_dc
	scam.kmr.se	7351	oai_lom
	tv.campusdomar.es	978	oai_dc
	wikwijs.samendelen.nl	26054	oai_dc
	www.mtg-maloposka.org.pl	181	oai_dc

Fig.5 – Repositories contacted by TF-Media

- Information has to be collected on the actual categories of the content in order to represent them into the PuMuKit-based web portal.
- It has to be understood how different taxonomies applied by different repositories can match for instance with the UNESCO taxonomy [12] of categories.
- Harvesting “good quality” metadata remains a challenge even in case of these friendly repositories, therefore a small sample/reference content repository (including for instance only TERENA videos) is being proposed that sets the standards for others to follow.
- Gradual approach to automated metadata gathering: Translating key words to categories or mapping different taxonomies are not trivial issues therefore out of the scope of the pilot phase. Categories have to be used that we are getting from “friendly repositories”.

Eventually, the TERENA OER pilot can result in best practices and implementation options for those repositories that have no adequate quality metadata to be taken into account in the production service development and expansion phase hopefully covered by GN4.

4. Delivering the pilot

The TERENA small project can be delivered in four tasks over 9 months (relaxed timeline):

1. Definition of the minimum requirements for a common metadata schema (flexible, scalable, standard-based, etc.) taking into account the information model of the pre-selected “friendly repositories”. The potential piloting of a sample/reference repository with “good quality” metadata will also be considered by this task.
2. Implementation of the ARIADNE-based metadata harvesting engine in the TERENA network.
3. Development and deployment of the PuMuKit-based web portal front-end (web template).
4. Integration of software components and piloting of the metadata harvesting, validating, transforming and publishing service.

4.1. Participation

The following NRENs and institutes have expressed their interest in joining the TERENA small project:

- GRNET, IUCC, SWITCH, FCCN, RedIRIS, NIIF
- University of Vigo (UVigo), Instituto Superior de Engenharia do Porto (ISEP), University of Pierre and Marie Curie (UPMC), Universitat Politècnica de València (UPV), Kaunas University of Technology (KUT), Tel Aviv University (TAU)

The following repositories have been identified as “friendly repositories”:

- SWITCHcollection, Campus do Mar, MAOR, Nordic OER alliance

Although at a first stage of the pilot the “friendly repositories” will be connected to the TERENA OER infrastructure, all the repositories listed in Fig.5 will be further studied in order to identify content that can be used for the TERENA OER pilot. Additional mechanism at the TERENA OER aggregator will be integrated to that end to filter content, transform and enrich metadata. In addition to this, synergies with large OER initiatives at EU level will be explored such as with the Open Discovery Space project [13] in which GRNET is a key partner maintaining and hosting the metadata aggregation infrastructure.

4.2. Timelines and costs

Tasks	M1	M2	M3	M4	M5	M6	M7	M8	M9
1. Metadata schema									
2. Harvesting engine									
3. Web portal front-end									
4. Integration and piloting									

Fig.6 - Pilot project tasks and their dependencies

Tasks	Estimated efforts	Estimated cost	Partners
1. Information model and metadata schema development and recommendations (reference repository)	2PM	12.000 EUR	IUCC/Tel Aviv University (TBC)
2. ARIADNE harvesting engine deployment and testing with “friendly repositories”	3PM	10.000 EUR	GRNET/third-party (TBC)
3. PuMuKit-based web front-end development according to the requirements of a state-of-the-art OER	2PM	6.000 EUR	University of Vigo/third-party (TBC)
4. Software integration and pilot service development	2PM	6.000 EUR	All potential partners (TBC)
5. <i>Project management</i>	<i>1PM</i>	<i>9.000 EUR</i>	<i>TERENA indirect cost</i>
TOTAL	10PM	43.000 EUR	

Fig.7 - Estimated efforts and costs

4.3. Deliverables

Deliverables will be provided in the form of living documentation on the TERENA Wiki pages as well as working software prototypes.

- Deliverable 1 – Information model and metadata schema development and recommendations for the harvesting engine and web front-end implementation (M6)
- Deliverable 2 – Pilot service description and prototyping (M9)

5. Opportunities and Risks

The TERENA OER pilot is going beyond the set-up of a single portal. It aims at delivering an open infrastructure for aggregating OER that can be used by the TERENA community to support educational activities. This will open the door for many new OER related services for the TERENA members such as the development of an OER microsite for a specific topic and/or specific collections. Such OER discovery microsites can be integrated in any portal of the TERENA members so they can build an OER section powered by TERENA. The table below presents a SWOT analysis for the TERENA OER pilot.

Strengths	Weaknesses
<ul style="list-style-type: none"> • European-level OER service governed and operated by NRENs • One-stop-shop for primary NREN stakeholders (academia) • Metadata aggregation only • Key building block for value-added services (e.g., MOOCs) 	<ul style="list-style-type: none"> • Lack of good quality metadata • Consensus on information model and metadata schema • Thematic repository vs. Specialized repository • Lecture recording vs. Learning Object
Opportunities	Threats
<ul style="list-style-type: none"> • Reaching the critical mass (keep the barrier low for small institutional reps.) • Being able to participate in global federation (e.g., GLOBE) • Provide/Transfer knowledge within the academic community 	<ul style="list-style-type: none"> • Low priority service for most of the NRENs, some reluctance • Strong competitors coming from outside the TERENA/GÉANT community

Fig.8 - SWOT analysis

References

- [1] TF-Media OER pilot: <https://confluence.terena.org/pages/viewpage.action?pageId=33751325>
- [2] MERLOT: <http://www.merlot.org/merlot/index.htm>
- [3] MAOR: http://maor.iucc.ac.il/about_en.php#
- [4] OER commons: <http://www.oercommons.org/>
- [5] Learning Resource Exchange for Schools from European Schoolnet
<http://lreforschools.eun.org/web/guest/for-cont-providers>
- [6] MITOpenCourseWare <http://ocw.mit.edu/index.htm>
- [7] OpenContent <http://www.opencontent.org/>
- [8] Open Course Ware Consortium <http://www.ocwconsortium.org/>
- [9] GLOBE <http://globe-info.org/>
- [10] ARIADNE tools <http://www.ariadne-eu.org/content/technologies>
- [11] PuMuKit <https://code.google.com/p/pumukit/>
- [12] UNESCO OER Declaration <http://www.unesco.org/new/en/communication-and-information/access-to-knowledge/open-educational-resources/what-is-the-paris-oer-declaration/>
- [13] Open Discovery Space project <http://www.opendiscoveryspace.eu/>
- [14] Nordic OER <http://nordicoer.org/>