



UPDATER

Project Acronym: ODS

Project Title: Open Discovery Space: "A socially-powered and multilingual open learning

infrastructure to boost the adoption of eLearning resources"

Open Discovery Space - Updater

Document revision: 1.0

Version: 14.0.5

Authors:

UAH

Pr	Project co-funded by the European Commission within the ICT Policy Support Programme				
Dissemination Level					
Р	Public				
PP	Project Participants	Х			
С	Confidential, only for members of the consortium and the Commission Services				





WP: WP9

Type ¹: O

Project Co-ordinator

Company name: INTRASOFT International S.A.

Name of representative: Antonis Ramfos

Address: Rue Nicolas Bove 2b, L -1253

Luxembourg

Phone number: + 352 44 10 12 2186

Fax number: + 352 44 10 12 2359

E-mail: Antonis.ramfos@intrasoft-intl.com

Website: www.opendiscoveryspace.eu

¹Deliverable Type: P (Prototype), R (Report), O (Other)





Table of Contents

1 Introduction		ductionduction	. 6
	1.1	Definition	. 6
	1.2	Requirements	. 6
		The ODS identifiers	
	1.4	XML files reading	. 7
	1.5	Readable names for Drupal vocabularies	13
		elopment Issues	





List of Figures

Figure 1: XML files location	7
Figure 2: Harvested data to be added to the ODS portal	
Figure 3: Processed data files added to the ODS portal	
Figure 4: Invalid data files discarded by the updater	
Figure 5: Updater log sample	





-	•	4				
	10	t.	nt	1 1	a h	les
		u	UI		uv	





1 Introduction

1.1 Definition

The updater is an application in PHP which is the responsible to create the educational object nodes for the ODS Drupal using LOM²-based XML files generated from harvested content from different repositories. Each XML file represents an educational object node in the Drupal portal.

1.2 Requirements

- Drupal 7.22
- Drush 5.8
- MySQL 5.1.6.3
- Apache 2.2.17
- PHP 5.3.5

1.3 The ODS identifiers

This version of the udpater (14.0.5) rests on the premise that the current educational object nodes in the ODS portal could have the new two ODS identifiers that identify a specific resource: ODS general identifier (machine name: field_ods_general_identifier) and ODS metadata identifier (machine name: field_ods_metadata_identifier). However, it is mandatory that all the new ones must include them (files are discarded is if they don't contain the ODS identifiers). Therefore, in order to determine if a new node is in the ODS portal yet, the updater follows the next process:

- 1. First, the updater checks if drupal nodes have the ODS identifiers filled: ODS general identifier (machine name: field_ods_general_identifier) and ODS metadata identifier (machine name: field_ods_metadata_identifier).
- 2. If not, then the updater checks the next ODS educational objects fields: LO identifier (machine name: field_lo_identifier³) and Data Provider (machine name: field_data_provider), where the Data Provider represents the ID of the repository term included in the *Repository* vocabulary (machine name: repository).

_

² Learning Object Metadata: http://ltsc.ieee.org/wg12/files/LOM_1484_12_1_v1_Final_Draft.pdf

³ Problem: this field is not mandatory according to the ODS Application Profile, however we need to check this field when we don't have the ODS identifiers. Therefore, this field should be mandatory.





If the updater finds the node in the portal, then it adds the new ODS identifiers to the existing Drupal node (if they were empty), and updates the information of this node with the new content.

1.4 XML files reading

The location of the harvested content is defined in the variable 'ods_updater_xml_root_file_path' defined in the ODS Drupal portal (*Administration menu > Development > Variable editor*) as shown in Figure 1.

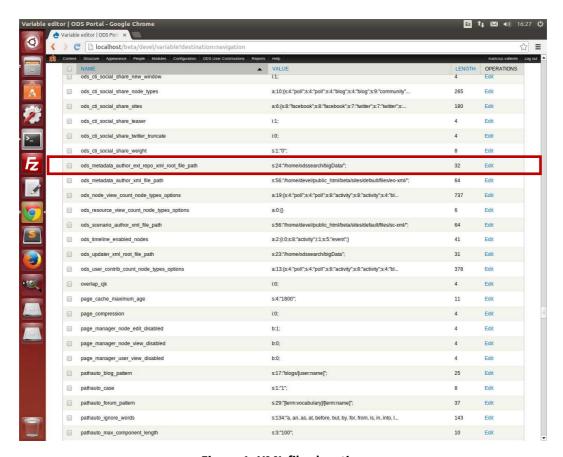


Figure 1: XML files location

The harvested data will be added in the 'new' folder of this location for the updater processing (see Figure 2). Each folder included in the '/home/odssearch/bigData/new' path is a repository and it stores the different XML files that are part of the repository (e.g. if the location of a XML file is /home/odssearch/bigData/new/ORGANIC_EDUNET/ODS_ORGANIC_EDUNET_1030.xml, then the repository name for this file will be 'ORGANIC_EDUNET').



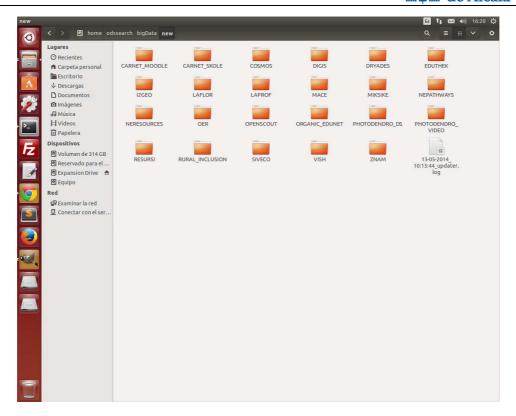


Figure 2: Harvested data to be added to the ODS portal

During the execution of the updater each processed file is moved to the 'old' folder of the ODS updater XML root file path for later data retrieval (see Figure 3), and invalid files are stored in the 'error' folder of this file path (see Figure 4)

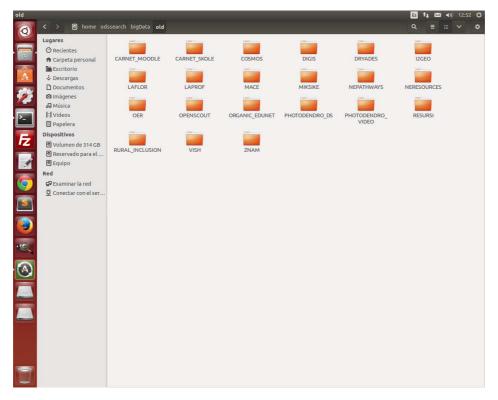


Figure 3: Processed data files added to the ODS portal





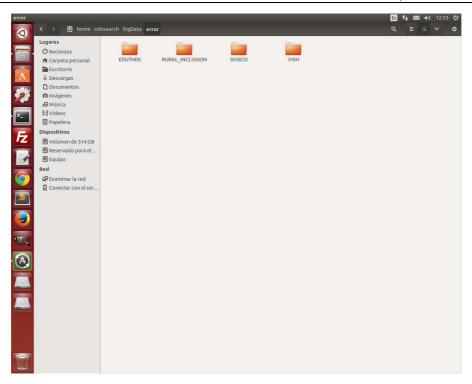


Figure 4: Invalid data files discarded by the updater

Not all the harvested data are used in the updater for the creation of the educational object nodes. Table 1 illustrates the mapping between the LOM fields that the updater processes from the XML files and the ODS Drupal fields from the 'Educational Object' content type (machine name: educational_object). It is worth noting that the updater fills the next ODS Drupal field with the current location of the XML file ('old' folder) in order to be able to retrieve this resource from the server for later processing of the whole content: ODS file location (machine name: field_ods_file_location).



Table 1: Mapping between LOM fields and ODS Drupal fields

LOM field	Educational Object field	Machine Name	Vocabulary
general/identifier/entry	LO Identifier ODS general identifier (catalog = 'ODS')	field_lo_identifier field_ods_general_identifier	-
general/title/string (language attribute)	Title	title_field	-
general/language	General Language Language	field_general_language language	ODS AP Languages -
general/description/string (language attribute)	Educational Object Description	field_eo_description	-
general/keyword/string (language attribute)	Edu Tags	field_edu_tags	Edu Tags
general/aggregationLevel/value	Aggregation Level	field_aggregation_level	ODS AP Aggregation Level
lifecycle/contribute/entity	Author Fullname	field_author_fullname	-
lifecycle/contribute/date/dateTime	EO Update Date	field_eo_update_date	-
metametadata/identifier/entry	ODS metadata identifier (catalog = 'ODS')	field_ods_metadata_identifier	-
technical/format	Media Type	field_technical_format	ODS AP Technical.Format
technical/location	Resource Link	field_resource_link	-

Open Discovery - Updater Page **10** of **17**





LOM field	Educational Object field	Machine Name	Vocabulary
educational/learningResourceType/value	Learning Resource Type	field_learning_resource_type	ODS AP Educational.LearningResourceType
educational/context/value	Educational Context	field_educational_context	ODS AP Educational.Context
educational/typicalAgeRange/string (language attribute)	Educational TypicalAgeRange	field_educational_typicalagerang	-
rights/cost/value	Rights Cost	field_rights_cost	ODS AP Rights.Cost
rights/copyrightAndOtherRestrictions/value	Rights Copyright	field_rights_copyright	ODS AP Rights.Copyright
classification/taxonpath/taxon/entry/string (language attribute)	Classification TaxonPath Classification Discipline	field_classification_taxonpath field_classification_discipline	- ODS AP Classification.Discipline

As we have already mentioned, it is important to notice that the repository name of each XML file is the name of the folder in which the file is part of, and the educational object field Data Provider (machine name: field_data_provider) will store the ID for the corresponding term in the *Repository* vocabulary.





Finally, once the updater has finished the processing of the XML files it generates a log file which is located in the same path of the harvested data (i.e. '/home/odssearch/bigData/new') with information about the processed XML files. Figure 5 illustrates the content of a log file generated by the updater script.

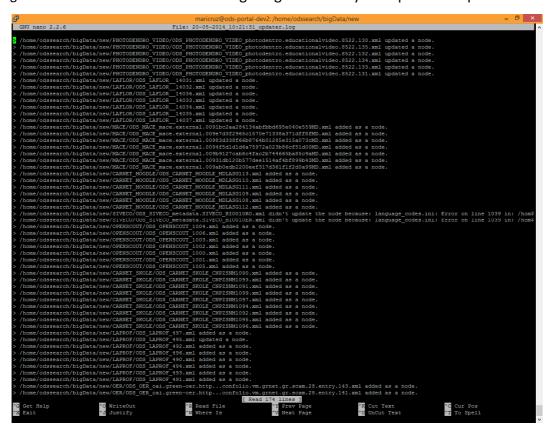






Figure 5: Updater log sample

1.5 Readable names for Drupal vocabularies

Aggregation levels, language codes, languages and repository names from the harvested data must be replaced by an existing taxonomy term in the corresponding ODS Drupal vocabulary (Administration menu > Structure > Taxonomy). To do this, the updater uses different '.ini' files that contain the valid terms to use in the ODS Drupal vocabularies: aggregation_level.ini, languages.ini, language_codes.ini and repositories.ini. The '.ini' files must be located in the next path: <code>/path_to_drupal_installation/odsUpdater</code>. For each .ini file, it is important to check that the corresponding Drupal vocabulary has the same terms than the ini file. Besides, for the languages, if new languages are added the Drupal portal languages configuration has to be updated through the option Administration menu > Configuration > Regional and language > Languages.

aggregation_level.ini:

```
; Drupal Taxonomy: ODS AP Aggregration Level
; Machine name: ods_ap_aggregation_level
1 = "Educational Object"
2 = "Lesson Plan"
3 = "Educational Scenario"
```





language_codes.ini:

= "bg" = "cs" CS da = "da" de = "de" = "de" de-DE = "de" Deutsch el = "el" en-US = "en" eng = "en" english = "en" = "en" = "es" es = "et" et fi = "fi" = "fr" fr ga = "qa" hr = "hr" = "hu" hu it = "it" = "lt" = "lv" lv = "nl" nl = "pl" pl pt ro ru sr = "sv" sv = "en" und

languages.ini:

```
; Drupal Taxonomy: ODS AP Languages
; Machine name: ods_ap_languages
   = "Bulgarian"
bg
            = "Czech"
            = "Danish"
            = "German"
de
de-DE
            = "German"
Deutsch
           = "German"
            = "Greek"
el
            = "English"
eng
english
           = "English"
            = "English"
en-US
            = "English"
             = "Spanish"
es
             = "Estonian"
fi
             = "Finnish"
fr
             = "French"
ga
             = "Irish"
```





hr = "Croatian" hu = "Hungarian" = "Italian" it = "Lithuanian" = "Latvian" ٦v = "Dutch" nl = "Polish" pl = "Portuguese" pt = "Romanian" ro ru = "Russian" = "Serbian" sr = "Slovenian" sv = "Other" und

repositories.ini:

; Drupal Taxonomy: Repository
; Machine name: repository

CARNET_MOODLE = "Carnet" CARNET_SKOLE = "Carnet" COSMOS = "Cosmos" COSMOS HEALTHY = "Cosmos" DIGIS = "DIGIS" = "DIGIS" DIGIS_HEALTHY DRYADES = "Dryades" DRYADES_HEALTHY = "Dryades" EDUTHEK = "Eduthek"

GREEN_OER = "Green OER Comments"

GREEN_OER_HEALTHY = "Green OER Comments"

I2GEO = "i2geo" I2GEO_HEALTHY = "i2geo" LAFLOR = "La Flor" LAFLOR_HEALTHY = "La Flor" LAPROF = "LaProf" LAPROF_HEALTHY = "LaProf" MACE = "MACE" MTKSTKE = "Miksike" MIKSIKE_HEALTHY = "Miksike"

MOODLE = "Moodle Carnet"

MOODLE_HEALTHY = "Moodle Carnet"

NEPATHWAYS = "Nepathways"

NERESOURCES = "Neresources"

OER = "OER"

OPENSCOUT = "OpenScout"

ORGANIC_EDUNET = "Organic.Edunet"

ORGANIC_EDUNET_HEALTHY = "Organic.Edunet"

PHOTODENDRO_DS = "Photodendro"

PHOTODENDRO_VIDEO = "Photodendro"

PHOTODENDRO_VIDEO_HEALTHY = "Photodendro"

RESURSI = "Resursi"

RURAL_INCLUSION = "Rural Observatory"





RURAL_INCLUSION_HEALTHY = "Rural Observatory"

SIVECO = "Siveco LRE"
SIVECO_HEALTHY = "Siveco LRE"

SKOLE = "Skole"

SKOLE_HEALTHY = "Skole"

VISH = "Vish"

ZNAM = "Znam.bg"

ZNAM_HEALTHY = "Znam.bg"





2 DEVELOPMENT ISSUES

This chapter provides a set of remarks about some implementation issues that should be improved:

- The updater considers the first general language label of the XML file to be the language of the Educational Object node. If the XML file doesn't contain a general language label, then the updater assigns the "en" (English) language code to the language field of the Educational Object node.
- The updater considers the first string title with the same language of the Educational Object Node to be the title of the Educational Object Node. If the XML file doesn't contain a title with the same language of the language node, ir order to avoid empty titles, the updater takes the first title as the title of the Educational Object node and it changes the language of the Educational Object node with the language of this first title (when a title doesn't contain the language attribute, the updater assigns the value "und" to the language of this title).
- The typical age range field is processed as a string and perhaps it should be better categorized into a real age range for further analysis in the ODS portal (especially for the search functionality: age filter).
- The updater considers the entry label of the classification/taxonpath/taxon element to be classifications of the Educational Object node. However, sometimes the value of the taxon entry is in the form "term::term::term..." or "term>term>term..." and this is how the classification is displayed in the summary page of the ODS portal.
- The updater considers the last term after the separators "::" in the entry label of the classification/taxonpath/taxon element to be disciplines of the Educational Object node. If the taxon entry doesn't include the separators, a discipline is not assigned.