

# DELIVERABLE

**Project Acronym:** LoCloud  
**Grant Agreement number:** 325099  
**Project Title:** Local content in a Europeana cloud

---

**D5.1.3: Periodic Report on content ingestion  
Month 24**

**Revision:** [M24 update]

---

**Authors:**

**Kate Fernie (2Culture)  
Gunnar Urtegaard (NRA)  
Silvia Alfreider (NRA)  
Runar Bergheim (AVINET)**

Project co-funded by the European Commission within the ICT Policy Support Programme		
Dissemination Level		
P	Public	X
C	Confidential, only for members of the consortium and the Commission Services	

**Statement of originality:**

This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both

## Table of Contents

<b>1</b>	<b>Introduction</b>	<b>4</b>
<b>2</b>	<b>Content action plans and Event log</b>	<b>5</b>
2.1	The collection workflow	5
2.2	The partners' timetable	6
2.3	Content by partner	8
2.3.1	Collections by partner	9
2.3.2	Content lifecycle status by partner	13
2.3.3	Partner reports	14
<b>3</b>	<b>MINT</b>	<b>19</b>
3.1	MINT reports and metadata statistics	19
3.2	Partner activity on MINT	19
3.2.1	Data imports	20
3.2.2	Data transformations	21
3.2.3	Dataset publications	22
3.2.4	MINT – LoCloud project overall statistics	23
3.3	MINT metadata statistics	24
<b>4</b>	<b>MoRE repository</b>	<b>26</b>
4.1	Metadata ingestion	26
4.1.1	Ingested metadata packages	26
4.1.2	Published metadata packages	27
4.1.3	Package information	29
4.1.4	View items	29
<b>5</b>	<b>Content in Europeana</b>	<b>32</b>
<b>6</b>	<b>Metadata quality</b>	<b>33</b>
<b>7</b>	<b>Conclusions</b>	<b>35</b>
<b>Annex 1:</b>	<b>Event log</b>	<b>36</b>
	Overview of the Event log	36
	Updating the Event log: content providers perspective	36
	User accounts and Log in	36
	Partner information	37
	Information that is stored about collections	38
	Adding events to collections	40
	Events in the life-cycle of a collection	41
	Reports from the event log	42

# 1 Introduction

An important goal for LoCloud is to increase significantly the quantity of high quality digital content accessible through Europeana, provided by small and medium sized cultural institutions. Many small, local institutions have limited IT infrastructure and lack either the requisite staff skills in digitization and digital libraries. LoCloud is exploring the potential of cloud computing technologies for Europeana, both in the direction of an easier to use and a more efficient infrastructure and in the creation of a range of software services that benefit content providers and users.

During the first two years of the project LoCloud has been putting in place the infrastructure and services to support the delivery of content to Europeana. At month twenty-four, LoCloud has completed the implementation of all services. The components of the aggregation infrastructure are in place including MINT, LoCloud Collections, the MORE repository and the series of micro-services made available for metadata enrichment.

The first LoCloud content was delivered to Europeana for test harvesting during month eighteen and was published in the live Europeana portal in early 2015.

The project's support portal and help desk went live in summer 2014, and now provides access to user documentation for each of the LoCloud services. A series of training workshops on LoCloud services were delivered to LoCloud partners in autumn 2014; the training materials used in these workshops are now available online for reference.

This deliverable, D5.1, focuses on monitoring the preparation and ingestion of metadata by LoCloud partners for Europeana.

LoCloud uses an Events Log together with reports available from MINT and MORE to monitor partners' progress with their content.

## 2 Content action plans and Event log

The expected progress in terms of amount and types of content is 4 million items by end of year 3. The project plan anticipated that no content would be provided to Europeana during year 1, with harvesting starting during year 2 and continuing until the end of year 3.

During May and June 2014 the partners' content action plans and timetable for ingestion were reviewed. The results of this monitoring activity were published in an appendix to D1.4 - Consolidated Action Plans.

### 2.1 The collection workflow

The Partners in LoCloud are aware that it is important to follow the workflow of a collection from LoCloud provider to Europeana and provide important data from each phase.

There is no universal definition of the term collection. In LoCloud collection is used about the content listed in the original project application – as well as content from any new data sources associated with the project throughout the execution period. One database/dataset with a uniform structure can hold more than one digital collection in a more traditional use of the term. However, from a technical perspective, a digital collection is all the data which can be extracted from the same set of tables in one database, carrying the same set of attributes – without heed to thematic divisions based on content types, themes etc.

Once a LoCloud partner has reached agreement with a collection “owner” (this may be a local cultural institution within their region or a department in their organisation) to provide the content to Europeana, the following actions need to be completed for each collection:

1. The Europeana data exchange agreement (or and equivalent agreement with an aggregator) needs to be signed
2. Any IPR in the content needs to be cleared
3. Metadata needs to be captured and/or enriched (enrichment is optional)
4. Unique identifiers or URIs need to be provided for the content and metadata
5. The content needs to be published online with its metadata
6. A means of exporting metadata for aggregation needs to be established, this includes:
  - a. Choosing a system for managing the content and metadata – a digital library, repository or other system
  - b. Deciding whether to export metadata in an XML file for upload, or to establish an OAI-PMH repository, or some other means by which the metadata can be harvested
  - c. Deciding which metadata schema to specify as the output format.

LoCloud has established an Event Log (see section 3 below), and each collection being provided to Europeana through LoCloud must be registered in the Event Log. The log is then used to record activities that take place as the collection moves through the LoCloud workflow. The following activities are documented at collection level:

- When the collection is added to the content survey in the Event Log.
- When metadata is extracted about all items in a collection

- When metadata is and mapped to a registered LoCloud intermediary schema using the MINT tool. (i.e. the process of “reading” metadata from local collection management systems, mapping them to a target metadata profile and writing them into a format readable by the chosen repository technology).
- When metadata is normalized. (i.e. the process of transforming metadata from the original notation to EDM or another standard schema format accepted by LoCloud).
- When metadata is published on MINT to the LoCloud MoRE repository.
- When the normalized metadata is harvested and ingested in the LoCloud MoRE repository.
- When metadata is enriched.
- When metadata is validated as being ready for harvesting by Europeana.
- When collection metadata are harvested from the LoCloud aggregator by Europeana.
- When the collection is added to Europeana and is available to end-users.

## 2.2 The partners’ timetable

A timetable for the aggregation of partners’ content was prepared following the review of partners’ content action plans in May and June 2014. Progress with this timetable is monitored in this deliverable. Partners’ timings have varied vary from their initial plans but good progress has been made during year two.

### Key to the table:

TW = LoCloud training workshop

MiT = Test data on MINT

MoT = Test data on MORE

ET = Test data to Europeana for review

MR = Test data to LoCloud for review

H = Harvest planned

W = Wikimedia harvest

C Mi = Completed on MINT

C Mo = Completed on MORE

E\* = published and live in Europeana

EI = sent to Europeana for ingestion

	Sep-14	Oct-14	Nov-14	Dec-14	Jan-15	Feb-15	Mar-15	Apr-15	May-15	Jun-15	Jul-15	Aug-15
ABMR	MiT							H				
ADS	MiT		TW						H			
AIT	MiT		TW					H				H
BGB	MiT			TW					H			
BJC	MiT		TW									H
CAHAI		TW					MR		MiT			
CG33		TW						H	H	H		
CUT		E*				E*						
DP	MiT MoT		TW			C Mo	EI			H		
FMNF							H					
FRS	MiT						C Mo	EI				
Future Li				TW								
GKR:	MiT	TW										
HU	MiT			TW		ET	C Mo					
IPCHS			TW	C Mi	C Mo	E*						
Jara		E*		TW		E*						
KUAS				TW	MiT							
MECD			TW		MiT	MoT	ET					
PL	MiT		TW									
PSRL	MiT			TW		C Mo	E I				H	
PSNC	MiT											
PrifUK KAEG		TW						H		H		
NPU		TW			C Mo	C Mo	E I					
NRA	MiT	TW										
RCE	MiT	TW					H					
VUKF			TW	C Mi	C Mo	E*						
UDE		TW										
Basela nd						MiT	ET					
Paul Maye art							W					

#### Notes

Archaeologie und Museum Baseland has signed a Memorandum of Understanding with LoCloud to provide its content to Europeana.

## 2.3 Content by partner

The report below is provided on the Event log and shows estimated number of digital objects by partner: [http://locloudlog.avinet.no/report\\_content\\_by\\_partner.aspx](http://locloudlog.avinet.no/report_content_by_partner.aspx). Please note that the actual number of objects published on Europeana can increase as a result of collection growth, or decrease as a result of quality assurance measures.

Partner	Partner country	Partner name	Number of collections	Digital objects
ABMR	SE	Stiftelsen Läns museet Västernorrland	7	8229
ADS	UK	University of York - Archaeology Data Service	5	19062
AIT	AT	Angewandte Informationstechnik Forschungsgesellschaft mbH	6	34100
BGB	RS	Biblioteka grada Beograda (Belgrade City Library)	2	1550
BJC	RO	Biblioteca Judeteana 'O.Goga' Cluj ((Octavian Goga Cluj County Library)	1	1000
CAHAI	IS	Minjastofnun Íslands /The Cultural Heritage Agency of Iceland	2	1200
CG33	FR	Conseil Général de la Gironde - Archives Départementales de la Gironde	3	1897573
CUT	CY	Cyprus University of Technology	8	3905
DP	IE	Discovery Programme	9	23000
FMNF	PT	Fundação Museu Nacional Ferroviário / National Railway Museum	2	2200
FRS	IT	Fondazione Ranieri di Sorbello	7	6515
Future	GR	Future Library	9	
GKR	HR	Gradska knjižnica Rijeka	3	2240
HU	TR	Hacettepe Universitesi	1	1000
IPCHS	SI	Javni Zavod Republike Slovenije za Varstvo Kulturne Dediscine (Institute for Heritage Protection)	2	11133
Jara	SI	Jara. Zavod za razvoj knjižnic	2	6564
KUAS	DK	Kulturarvsstyrelsen (Danish Agency for Culture)	1	200000
MECD	ES	Ministerio de Educacion, Cultura y Deporte (Ministry of Education, Culture and Sport)	32	134471
NPU	CZ	Národní památkový ústav (National Heritage Institute)	2	6486
NRA	NO	National Archive Norway	1	2020000
PL	BE	Provincie Limburg	2	60000
PrifUK	SK	Univerzita Komenského	1	1000



Partner	Partner country	Partner name	Number of collections	Digital objects
KAEG		Prírodovedecká fakulta Katedra aplikovanej a environmentálnej geofyziky		
PSNC	PL	Instytutu Chemii Bioorganicznej PAN - Poznan Supercomputing and Networking Centre	1	18300
PSRL	Bulgaria	Pencho Slaveykov Regional Library	5	1754
RCE	NL	Ministerie van onderwijs, cultuur en wetenschap - Rijksdienst voor het Cultureel Erfgoed (Cultural Heritage Agency)	6	53874
UDE	DE	Universitaet Duisburg-Essen (University of Duisburg-Essen)	8	
VUFK	LT	Vilniaus Universitetas (Vilnius University - Faculty of Communication)	1	17065
			<b>Total</b>	<b>4,535,021</b>

### 2.3.1 Collections by partner

The report below is provided on the Event log and shows estimated number of digital objects in each collection that partners plan to provide, listed by their planned ingestion date: [http://locloudlog.avinet.no/report\\_collections\\_by\\_planned\\_ingestiondate\\_list.aspx](http://locloudlog.avinet.no/report_collections_by_planned_ingestiondate_list.aspx). Please note that the actual number of objects published on Europeana can increase as a result of collection growth, or decrease as a result of quality assurance measures. Note also that the date of harvesting is when the partners plan to harvest their content on MINT or MORE, not date of harvesting or publication by Europeana

Partner	Collection	Harvest date	Digital objects
CUT	Archive of Limassol Municipality	10.2014	500
CUT	POSTAL SERVICES	10.2014	1100
VUFK	Lithuanian hillforts	2.2015	17065
FMNF	Trains of Portugal	3.2015	200
FMNF	Museum collection	3.2015	2000
HU	Vehbi Koç & Ankara Research Centre (VEKAM)	3.2015	1000
IPCHS	Works of art database, restoration	3.2015	4133
NPU	VAL - significant archaeological sites	3.2015	5986
DP	Irish Architectural Archive images	3.2015	1000
IPCHS	Archaeological research database	4.2015	7000
GKR	Local museum items	4.2015	500
AIT	Don Juan Archive Vienna: theatre related texts	4.2015	24000
AIT	Hugo Montfort Digital Edition	4.2015	100
AIT	Numismatic Collection at the University of Graz	4.2015	3500

Partner	Collection	Harvest date	Digital objects
AIT	University of Graz: Archaeological Collections	4.2015	500
AIT	Visual Art of South-Eastern Europe	4.2015	3000
MECD	Biblioteca de Asturias	4.2015	23004
MECD	Biblioteca Pública del Estado en Pontevedra	4.2015	2043
CG33	Archives of Gironde collections: iconography	4.2015	15,886
CG33	Archives of Gironde collections: passports	5.2015	44,600
PSNC	Teatr NN	5.2015	18300
PrifUK KAEG	PriFUK KAEG collection	5.2015	1000
BGB	Belgrade local history	5.2015	1550
GKR	Local history collections, newspapers and magazines	5.2015	1040
GKR	Local heritage Crikvenica Collection Postcards	5.2015	700
ADS	Grey Literature Library	5.2015	12000
ADS	ImageBank	5.2015	450
ADS	PSAS	5.2015	4000
ADS	Star Carr Archives Project	5.2015	2312
ADS	Wessex Archaeology Archive	5.2015	300
ABMR	Birgittamuseet - medicine history	5.2015	470
ABMR	Kubikenborgs skolas intresseförening	5.2015	12
ABMR	Landsarkivet	5.2015	103
ABMR	Samtidsbild	5.2015	2800
ABMR	Ulvöarnas kulturarv	5.2015	1
ABMR	Ånge kommun (Ånges fotosamling )	5.2015	4843
FRS	Pictures	5.2015	1000
FRS	Porcelain	5.2015	661
FRS	Textiles	5.2015	274
CG33	Archives of Gironde collections: civil registry	6.2015	1,473,841
KUAS	Regin tool for 100 museum collections	6.2015	200000
CUT	Church of Cyprus	6.2015	2000
CUT	Church of Cyprus	6.2015	4
CUT	CY Police Forces	6.2015	100
CUT	Cyprus Broadcasting Corporation	6.2015	1
CUT	Ministry of Agriculture	6.2015	100
CUT	Press and Information Office CY Government	6.2015	100
KUAS	Regin tool for 100 museum collections	6.2015	200000
CG33	Archives of Gironde collections: other	7.2015	358,246
PSRL	Calendars	7.2015	57
PSRL	Invitations	7.2015	120
PSRL	Postcards	7.2015	777
PSRL	Seals	7.2015	100
PSRL	Postcards	7.2015	700
NPU	SOVAMM	8.2015	500
AIT	Europa Nostra Archive	8.2015	3000
ABMR	ABF Kramfors	10.2015	

Partner	Collection	Harvest date	Digital objects
MECD	Federación para la Etnografía y el desarrollo de la Artesanía Canaria (FEDAC)	12.2015	85000
MECD	Fundación Cela	12.2015	3224
MECD	Instituto de Educación Secundaria San Isidro	12.2015	1200
FRS	Books	12.2015	1500
FRS	Maps	12.2015	130
FRS	Paintings	12.2015	150
FRS	Prints	12.2015	2800
BJC	Local libraries in Cluj County		1000
BGB	New content providers		
CG33	Partner collections	Late 2015	
Future	Ilioupoli	Late 2015	
Future	Keratsini-Drapetsona	Late 2015	
Future	Korinth	Late 2015	
Future	Kozani	Late 2015	
Future	Levadia	Late 2015	
Future	Naupaktos	Late 2015	
Future	Trikala	Late 2015	
Future	Chania	Late 2015	
Future	Drama	Late 2015	
DP	Archaeological research images		5000
DP	Church Body Library		
DP	DP image collection		
DP	Dublin Institute of Advanced Studies		
DP	Hunt Museum Limerick		2000
DP	Leo Swan Aerial Collection		
DP	Members of AARG, private collections of archaeological sites		
DP	Royal Society of Antiquaries of Ireland		15000
Jara	KAMRA		6164
Jara	User generated content		400
RCE	Archaeological depot, province of Gelderland		
RCE	Archaeological reports		17000
RCE	Controlled vocab of Dutch archaeology (RNA)		
RCE	Historic Cultural Landscapes		36874
RCE	Other archaeological depots and museums (not present in DiMCoN)		
RCE	Regional/local museums		
MECD	Museo Arqueológico de Linares. Monográfico de Cástulo		
MECD	Museo Arqueológico Provincial de Ourense		
MECD	Museo Casa de los Tiros		
MECD	Museo de Albacete		
MECD	Museo de Artes y Costumbres Populares de		

Partner	Collection	Harvest date	Digital objects
	Sevilla		
MECD	Museo de Artes y Costumbres Populares del Alto Guadalquivir		
MECD	Museo de Cádiz		
MECD	Museo de Ciudad Real		
MECD	Museo de Cuenca		
MECD	Museo de Guadalajara		
MECD	Museo de Historia		
MECD	Museo de Huelva		
MECD	Museo de Huesca		
MECD	Museo de Jaén		
MECD	Museo de las Peregrinaciones y de Santiago		
MECD	Museo de los Orígenes		
MECD	Museo de Málaga		
MECD	Museo de Mallorca		
MECD	Museo de Santa Cruz		
MECD	Museo de Teruel		
MECD	Museo de Zaragoza		
MECD	Museo del Parque Cultural de Molinos. Colección Eleuterio Blasco Ferrer		
MECD	Museo Etnológico de Ribadavia		
MECD	Museo Juan Cabré		
MECD	Museo Martín Almagro de Albarracín		
MECD	SG for state museums		20000
MECD	Shadous: memoria fotográfica		
CAHAI	Archaeological sites		1200
CAHAI	Municipality of Skagafjorour, local museum		
NRA	Local history		2020000
PL	Placenames		10000
PL	Various		50000
UDE	Center for International Light art in Unna		
UDE	City Archives of Witten		
UDE	Conserve the Sound Essen		
UDE	Gustav-Lübcke-Museum Hamm		
UDE	Kunstmuseum Mülheim an der Ruhr in der Alten Post		
UDE	Local association Hattingen/Ruhr / Museum in the Iron House		
UDE	Museum of the city Gladbeck in the Water tower Wittringen		
UDE	The Ruhr-University Bochum - Modern Art Gallery		
			4,535,021

### 2.3.2 Content lifecycle status by partner

The Event log is designed to track the activities involved in preparing content for harvesting and aggregation. This report summarises the events by partner. It is intended to be updated by partners as their collections move through the workflow. The report is available at:

[http://locloudlog.avinet.no/report\\_collections\\_by\\_event2.aspx](http://locloudlog.avinet.no/report_collections_by_event2.aspx).

Partner name	No of collections	IPR cleared	Captured	Enriched	Published	PURIs	Exported	Harvested	Validated
ABMR	7	6							
ADS	5	4	4		4	4	4		
AIT	6	5	3		4	3			
BGB	2	1			1				
BJC	1	1							
CAHAI	2	1	1			1			
CG33	3	2	2		1	1	1		
CUT	8	0		8					
DP	9	6	7		2	1			
FMNF	2	1	2						
FRS	7	7	5			7	3	4	3
Future	9	0							
GKR	3	0							
HU	1	1	1	1				1	
IPCHS	1	1	1		1	1	1		
IPCHS	2	0		1	1	2		1	
Jara	2	3	2		2	1	1		
KUAS	1	1	1		1	1	1		
MECD	32	0	2						
NPU	2	1	1		1	1	1	1	1
NRA	1	0							
PL	2	1			1		1	1	1
PrifUK KAEG	1	1		1					
PSRL	5	3	3		3		3		
RCE	6	0							
UDE	8	0							
VUFK	1	1	1		1	1	1	1	

### 2.3.3 Partner reports

ABMR	<p><b>Stiftelsen Länsmuseet Västernorrland</b></p> <p><b>Museums collections</b> – A regional aggregator. ABMR has completed testing of its collections on MINT. Waiting for a local partner to complete uploading its content. Harvesting planned in April.</p>
ADS	<p><b>University of York - Archaeology Data Service</b></p> <p><b>Archive collections</b> – Test mappings have been completed on MINT. Harvesting planned for spring-summer 2015. DEA signed.</p>
AIT	<p><b>Angewandte Informationstechnik Forschungsgesellschaft mbH</b></p> <p><b>Local partners' datasets</b> – mapping for the first dataset (the Don Juan Archive) is now being prepared on MINT. Harvesting is planned in April 2015. DEA signed.</p> <p>The Europa Nostra Archive took time to find cataloguing software and work was delayed, but they would like to enter Europeana via LoCloud and the first data is anticipated later this summer.</p>
BGB	<p><b>Biblioteka grada Beograda (Belgrade City Library)</b></p> <p><b>Belgrade local history</b> plans to use LoCloud Collections, testing has been completed and work is now getting underway to import their digital objects and metadata.</p>
BJC	<p><b>Biblioteca Judeteana 'O.Goga' Cluj ((Octavian Goga Cluj County Library)</b></p> <p><b>Local libraries in Cluj County</b> – working with local partners using Omeka. Metadata is currently in preparation and is planned for harvesting in summer 2015.</p>
CAHAI	<p><b>Minjastofnun Íslands /The Cultural Heritage Agency of Iceland</b></p> <p><b>Sarpur</b> – CMS for 53 Icelandic museums. Tests in progress to provide the collection of <b>Skagafjörður Heritage Museum</b> to Europeana via LoCloud. XML export prepared and a mapping is in preparation on MINT. DEA signed.</p>
CG33	<p><b>Conseil Général de la Gironde - Archives Départementales de la Gironde</b></p> <p>Metadata sample from OAI-PMH repository was sent to LoCloud for review, some mandatory elements were missing and work is now in progress to improve the export format.</p> <p><b>Iconography</b> - harvesting planned in March/April 2015  <b>Passports</b> - harvesting planned in April 2015  <b>Civil registry</b> - harvesting planned in May 2015</p>

	<p><b>Other CG33 collections</b> - harvesting planned in June 2015</p> <p><b>External partners' collections</b> - harvesting planned in autumn 2015</p>
CUT	<p><b>Cyprus University of Technology</b></p> <p><b>Local collections</b> – CUT is in dialogue with a large number of local partners and is acting as a regional aggregator using Omeka. Harvesting has commenced and the first collections have been published on Europeana. Further harvesting is anticipated during 2015. DEA signed.</p>
DP	<p><b>Discovery Programme</b></p> <p><b>DIAS and Leo Swan collections</b> ready on MORE for harvesting by Europeana in March. DEA signed.</p> <p>Further collections in preparation.</p>
FMNF	<p><b>Fundação Museu Nacional Ferroviário / National Railway Museum</b></p> <p><b>FMNF museum</b> object collection is currently being prepared for harvesting. Europeana DEA has been signed.</p> <p><b>Trains of Portugal</b> collection is currently being prepared for harvesting. DEA yet to be signed.</p> <p><b>Municipality of Vila Nova de Famalicão</b> will now provide its content via the Portuguese National Aggregator for Archives.</p> <p><b>Archives</b> management software has not been working since 2013, organisational and finance issues mean that FMNF has not been able to replace the system and is currently unable to progress its archive collections.</p>
FRS	<p><b>Fondazione Ranieri di Sorbello</b></p> <p><b>Samira collection</b> is ready on MORE for harvesting by Europeana. Publication is blocked temporarily while an early test version of the dataset is removed from Cultura Italia and Europeana. The test data was mistakenly exposed for harvesting on the Regione Umbria portal and is now being removed. DEA signed.</p> <p><b>Local institutions</b> – discussions are in progress with local institutions wishing to provide content to Europeana via LoCloud. MOUs and DEAs in the process of being exchanged.</p>
Future	<p><b>Future Library</b></p> <p><b>Library partners</b> – progress has been pending completion of LoCloud Collections, discussions with libraries have now re-started. Two libraries have signed the DEA.</p>
GKR	<p><b>Gradska knjižnica Rijeka</b></p> <p>Currently working on a mapping of metadata from their Omeka instance to EDM, in MINT and MORE.</p>

HU	<p><b>Hacettepe Universitesi</b></p> <p><b>VEKAM collection.</b> Feedback was given by LoCloud and Europeana on sample metadata and following improvements the metadata is now ready on MORE for harvesting by Europeana in March. DEA signed.</p>
IPCHS	<p><b>Javni Zavod Republike Slovenije za Varstvo Kulturne Dediscine (Institute for Heritage Protection)</b></p> <p><b>Works of art database, restoration collection</b> – harvesting completed and the collection is now live in Europeana. DEA signed.</p> <p><b>Cultural monuments reports collection</b> – now in preparation delivery due late 2015. Rights cleared.</p> <p><b>Archaeological reports</b> – delays in clearing the IPR related to these reports led to an alternative collection, Cultural Monuments reports, being prepared.</p>
Jara	<p><b>Jara. Zavod za razvoj knjiznic</b></p> <p><b>Kamra</b> collection has been harvested twice and content is now live in Europeana. Regular updates are planed throughout the project. DEA signed.</p> <p><b>User generated content</b> – planned for delivery late 2015.</p>
KUAS	<p><b>Kulturarvsstyrelsen (Danish Agency for Culture)</b></p> <p><b>Regin</b> - 100 museum collections. LIDO export has been established. Tests have been completed on MINT. Planned for delivery in spring 2015. DEA signed.</p>
MECD	<p><b>Ministerio de Educacion, Cultura y Deporte (Ministry of Education, Culture and Sport)</b></p> <p>Staff changes and an internal reorganisation lead to some changes in the content planned for delivery.</p> <p><b>Museums collection</b> is in preparation and will be available for harvesting shortly.</p> <p><b>Virtual Library of Bibliographical Heritage</b> – sample EDM metadata was sent to LoCloud and Europeana for review, some corrections were requested and harvesting is pending their completion.</p> <p><b>Photo collection</b> is due to be ready at the end of 2015.</p>
NPU	<p><b>Národní památkový ústav (National Heritage Institute)</b></p> <p><b>VAL - significant archaeological sites</b> is ready on MORE for harvesting by Europeana in March. Issue with X and Y coordinates being reversed has been fixed. DEA signed.</p>



NRA	<p><b>National Archive Norway</b></p> <p>Tests have been completed on MINT and MORE. Discussions are in progress with content partners. Due to deliver by end 2015.</p>
PL	<p><b>Provincie Limburg</b></p> <p><b>Erfgoedregister</b> covers more than 200 data provider collections. Is ready for harvesting. Edm:provider issue to be resolved with Europeana. DEA signed.</p>
PrifUK KAEG	<p><b>Univerzita Komenského Přírodovědecká fakulta Katedra aplikovanéj a environmentálnej geofyziky</b></p> <p>Two collections are in preparation, the smaller collection should be ready for harvesting by April/May 2015 a and the bigger one by June 2015. DEA signed.</p> <p>In dialogue with local institutions with private collections.</p>
PSNC	<p><b>Instytutu Chemii Bioorganicznej PAN - Poznan Supercomputing and Networking Centre</b></p> <p><b>Teatr NN</b> – Tests have been completed on MINT, harvesting is planned in spring/summer 2015.</p>
PSRL	<p><b>Pencho Slaveykov Regional Library</b></p> <p><b>Calendars, Invitations, Postcards and Seals collection</b> ready on MORE for harvesting by Europeana in March. DEA signed.</p> <p>Updates anticipated in summer 2015 and January 2016.</p>
RCE	<p><b>Ministerie van onderwijs, cultuur en wetenschap - Rijksdienst voor het Cultureel Erfgoed (Cultural Heritage Agency)</b></p> <p>Various tests are in progress on MINT. Harvesting planned later in 2015.</p>
UDE	<p><b>Universitaet Duisburg-Essen (University of Duisburg-Essen)</b></p> <p>Working with the DDB in Frankfurt. Plans underway for a workshop for local content partners on LoCloud Collections etc.</p>
VUFK	<p><b>Vilniaus Universitetas (Vilnius University - Faculty of Communication)</b></p> <p><b>Lithuanian hillforts collection</b> has been harvested by Europeana and is now published on Europeana. DEA signed.</p>
Mayer	<p><b>Paul Mayeart</b></p> <p>Collection on Wikimedia has been harvested using the LoCloud Wikimedia application and metadata is now on MORE, validation and enrichment work in progress.</p>

Baseland	<b>Archäologie und Museum Baseland</b>  <b>Baseland Museums collection</b> – aggregator for museums in the Basel region. Tests have been completed on MINT and MORE, metadata is now in preparation for harvesting. EDM export format. DEA signed.

### 3 MINT

The MINT mapping tool (<http://mint-projects.image.ntua.gr/locloud/>) is a web based platform designed to support metadata mapping and ingestion for cultural heritage content and metadata in Europe. The main activities by users of MINT are:

- Registration
- Metadata import
- Creating a metadata mapping
- Metadata transformation
- Publication of data to the LoCloud MORE repository

#### 3.1 MINT reports and metadata statistics

MINT includes reporting functionality that provides information about users' actions and their progress in MINT. These reports allow the following actions to be monitored:

1. Dataset Imports.
2. Dataset Transformations.
3. Dataset Publications.
4. An overview of each Organization's progress by dataset with item counts (imported, transformed, published).
5. A Project overview of all the datasets imported, transformed or published by each partner organization.

#### 3.2 Partner activity on MINT

By March 2015, excluding NTUA, 22 partners had imported datasets to MINT. This activity reflects the hands-on training in MINT offered as part of the three LoCloud training workshops held in autumn 2014. Several partners imported test datasets as part of this training, and several have imported trial exports from their own systems in preparation for harvesting in year 3. As a result of the testing, and repeated exports of the same dataset in different formats, the total number of imported items is not a measure of the content to be provided to Europeana.

### 3.2.1 Data imports

Partner	Name	Latest import	Number of imports	Total imported items
ABMR	<a href="http://rebox.murberget.se">http://rebox.murberget.se</a>	19/03/2014	1	34143
ADS	amarna_for_import-corrected2.xml	28/11/2014	3	1875
AIT	Export KMB DJA_monographs.xml	09/02/2015	1	2983
Basel	<a href="http://gabrielz.myds.me">http://gabrielz.myds.me</a>	12/12/2014	4	200
BGB	BGB xml.zip	12/12/2014	5	8
BJC	<a href="http://greenstone.bjc">http://greenstone.bjc</a> <a href="http://omeka.bjc.ro/omeka/oai">http://omeka.bjc.ro/omeka/oai</a>	14/03/2015	16	1434
DP	DuNoyerTestSetUrls.xml DIASMovieMeta2.xml RSAI_TEST_METADATACSVeditTes testleoswanEDMNS.xml	5/03/2015	11	2840
FRS	output2_xml.zip	11/2/2015	21	8444
GKR	<a href="http://svevid2.codingthrills">http://svevid2.codingthrills</a> .	11/06/2014	16	4608
HU	vekam_son160315.xml	16/03/2015	6	3421
IPCHS	SI_IPCHS_RCG.xml	12/11/2014	3	11629
Jara	Test data	12/12/2014	2	5
KUAS	lido_kid_full_pub.zip	12/12/2014	6	34595
NPU	<a href="http://iispp.npu">http://iispp.npu</a> .	13/01/2015	3	258483
NRA	Various test datasets	5/03/2015	13	7744
PL	export_Modemuseum_001001b.xml	26/01/2015	2	566
PrifUK KAEG	Test data	24/10/2014	1	5
PSRL	Test data	12/12/2014	2	5
RCE	CultGIS.xml Unesco_Werelderfgoed_28992_gml rapporten_will.xml	12/03/2015	9	9
UDE	Test data	24/10/2014	2	5
VUKF	VUFC_piliakalniai_14001-	16/01/2015	6	17118

Notes:

NPU has used the LoCloud MINT tool to import and map an update to its CARARE dataset.

### 3.2.2 Data transformations

Data transformation is the process that takes place once a mapping has been completed between the metadata structure of the imported dataset and a target schema. The report on data transformations shows both the target schema and mapping used, and also the number of items that, once transformed, conform to the schema (are valid according to the schema's rules).

The report shows that 15 partners have transformed datasets using a number of different schema mappings, in two cases experimenting with mappings to different target schemas. The statistics again reflect the hands-on training delivered in the LoCloud workshops, with test mappings and transformations being performed on the sample data provided. The statistics for the number of invalid items, coupled with the evidence for repeat mappings and transformations on the same datasets, demonstrate that the partners are testing the results and making adjustments both to their data output and to their mappings to achieve better quality outcomes.

P	No of datasets transformed	Items	Valid	Invalid	Target schema	Latest date	Mapping
Basel	2	110	110	0	EDM	28/11/2014	Native to EDM
BJC	7	610	596	14	EDM	15/03/2015	Native to EDM
DP	8	351	351	0	EDM	5/03/15	Native to EDM
FRS	22	8445	7654	791	EDM + LIDO	13/02/15	Native to EDM/LIDO
GKR	7	5892	3009	2883	EDM	13/03/2015	Native to EDM
HU	4	2283	2276	7	EDM	9/02/2015	Native to EDM
IPCH S	2	7757	7757	0	EDM + LIDO	25/11/2015	Native to EDM/LIDO
Jara	1	5	0	5	EDM	12/12/2014	Test to EDM
KUAS	1	5	5	0	EDM	12/12/2014	Test to EDM
NRA	5	6784	1067	5717	CARARE + EDM	2/02/2015	Native to EDM/CARARE
NPU	2	13176 3	1317 63	0	CARARE	16/02/2015	Native CARARE to CARARE2.0
PL	2	288	288	0	LIDO	26/01/2015	Native to LIDO
PrifU K KAEG	1	5	5	0	EDM	24/01/2014	Test to EDM

P	No of datasets transformed	Items	Valid	Invalid	Target schema	Latest date	Mapping
RCE	1	2	2	0	CARARE	03/10/2014	Native to CARARE
VU FK	6	17118	17118	0	EDM	21/01/2015	Native to EDM

### 3.2.3 Dataset publications

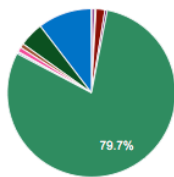
Once partners have completed their mappings and transformations and are satisfied with the results, the final step is publication. Publication is the process by which datasets are packaged with their mappings and sent to the LoCloud MORE repository for ingestion. In the second half of year two, 12 partners published datasets in several cases to test the process and view their results on the MORE repository. During this period 5 partners (DP, HU, IPCHS, NPU and VUKF) published completed datasets which were ready for the final processing steps on MORE before harvesting by Europeana.

Partner	Name	Date	Items
Basel	solr_100.xml	28/11/2014	100
BJC	http://omeka.bjc.ro/omeka/oai-	13/03/2015	95
BJC	http://omeka.bjc.ro/omeka/oai-	14/03/2015	108
BJC	records_v2.xml	16/03/2015	241
DP	DIASOJMeta2.xml	20/02/2015	55
DP	DIASX3DMeta2.xml	20/02/2015	39
DP	DIAS3DPDFMeta2.xml	20/02/2015	81
DP	DIASImageMeta3.xml	20/02/2015	85
DP	LeoSwanEdmElements.xml	06/02/2015	72
FRS	output2_xml_test.zip.tgz(2).tgz	14/01/2015	1470
GKR	http://svevid.locloudhosting.	13/03/2015	1373
HU	vekam_locloud_final.xml	09/02/2015	1138
IPCHS	SI_IPCHS_RCG.xml	25/11/2014	3885
IPCHS	SI_IPCHS_RCG_edm.xml	25/11/2014	3872
KUAS	lido2.xml	19/11/2014	10
KUAS	kunst1.xml	20/11/2014	10
KUAS	lido_mussamfoto_small.xml	11/12/2014	10
NPU	http://iispp.npu.	18/11/2014	125777
NPU	http://iispp.npu.	16/02/2015	5986
NRA	MUSIT photos Oslo 1000 rec.xml	31/10/2014	1000
PL	export_Modemuseum_001001.xml	05/03/2015	283
VUKF	VUFC_piliakalnai_14001-	21/01/2015	3068
VUKF	VUFC_piliakalnai_1-4000_ID7388.	21/01/2015	4000
VUKF	VUFC_piliakalnai_4001-	21/01/2015	4000
VUKF	VUFC_piliakalnai_8001-	21/01/2015	4000
VUKF	VUFC_piliakalnai_12001-	21/01/2015	2000

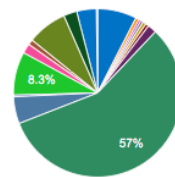
### 3.2.4 MINT – LoCloud project overall statistics.

Project overall statistics

Published Items



Imported Items



Overall

Name	Imported	Transformed	Published	Published to EDM	Published to LIDO	Published to CARARE 2.0.4
ABM Resurs	34143	0	0	0	0	0
ANECT	234	117	0	0	0	0
Angewandte Informationstechnik Forschungsgesellschaft mbH	2983	0	0	0	0	0
Archaeologie und Museum Baselland	200	130	100	100	0	0
Archaeology Data Service	1875	0	0	0	0	0
Asplan Viak Internet	13	5	0	0	0	0
Belgrade City Library	8	0	0	0	0	0
Comenius University	5	5	0	0	0	0
Cultural Heritage Agency of The Netherlands	11	2	0	0	0	0
Digital Curation Unit	5	0	0	0	0	0
Hacettepe University, Department of Information Management	2283	2276	1137	1137	0	0
Image Video and Multimedia Laboratory	5	0	0	0	0	0
Ireland	2840	351	332	332	0	0
IVML	5	5	0	0	0	0
Jara	5	0	0	0	0	0
Ministry of Education, Culture and Sport	2792	2792	2792	2792	0	0
National Archives of Norway	7744	1679	743	0	0	0
National Heritage Institute	258483	131763	131763	0	0	131763
NTUA	23512	8622	20	5	5	10
Octavian Goga Cluj County Library	1434	837	437	437	0	0
Pencho Slaveykov Regional Library	5	0	0	0	0	0
Poznań Supercomputing and Networking Center	37739	0	0	0	0	0
Province of Limburg (B)	566	849	283	0	283	0
Ranieri di Sorbello Foundation	8444	7654	1470	1470	0	0
Rijeka City Library	4610	3009	1373	1373	0	0
Testlab	31	13	0	0	0	0
Testlab	0	0	0	0	0	0
The Agency for Culture	34595	5	30	0	30	0
The Cultural Heritage Agency of Iceland	1	0	0	0	0	0
The Institute for the Protection of Cultural Heritage of Slovenia	11629	11629	7757	3872	3885	0
University of Duisburg-Essen	6	0	0	0	0	0
Vilnius University Faculty of Communication	17118	17119	17068	17068	0	0

### 3.3 MINT metadata statistics

In addition to reporting on user's activities, MINT also provides tools to analyse the statistics for individual datasets. The figure below illustrates the statistics for the dataset uploaded by ABMR, showing the element names, the count for each element, the number of distinct values for each element and the average length of the data contained in each element.

**<http://reporx.murberget.se:8080/reporx/OAIHandler> - KAP\_Handel\_lido Statistics**

Xpath	Count	Distinct	Length
lido:lidoWrap	34143		
@xsi:schemaLocation	34143	1	79
lido:lido	34143		
@xsi:schemaLoca...	34143	1	79
lido:lidoRecID	34143	34143	18.7
@lido:source	34143	1	24
@lido:type	34143	1	5
lido:objectPublishedID	34143	34143	55.7
@lido:type	34143	1	3
lido:category	34143		
lido:conceptID	34143	1	42
@lido:type	34143	1	3
lido:term	34143	1	15
@xml:lang	34143	1	2
lido:descriptiveMetad...	34143		
@xml:lang	34143	1	2
lido:objectClassifi...	34143		
lido:objectWork...	34143		
lido:objectW...	34143		
lido:term	68286	2	8
@lid...	68286	1	3
@xm...	68286	2	2
lido:classificati...	34143		
lido:classific...	102429		
@lido:t...	102429	3	12.3
lido:term	102429	3	10.3

By looking at the dataset statistics, it is possible to see if core information is present in the dataset. The core data required for Europeana is:

- Title or Description
- Subject, Type, Coverage or Spatial
- Rights licencing
- A URL or URI pointing either directly to the object or to a landing page where it can be found



This requirement means that to be valid for Europeana, each item in the dataset must contain relevant data in at least 4 elements.

The statistics provided in MINT allow for some additional checks to be made. For example, each record needs a unique identifier – an element that has the same number of distinct values as in the total count for the dataset can provide a unique identifier.

The tool also allows the data that is actually stored in an element to be browsed, which can be very helpful when mapping a dataset but can also provide a way of monitoring the quality. For example users can review the consistency and usefulness of the data actually present.

▼ **lido:displayActor information**

**XPath** `/lidoWrap/lido/descriptiveMetadata/objectRelationWrap/subjectWrap/subjectSet/subject/subjectActor/displayActor`

**Namespace URI** `http://www.lido-schema.org`

**Count** 34757

**Distinct Count** 30021

**Average Length** 18.102425

Value	Count
Centralhotellet	14
Byggnadsfirman	12
Fruktkällaren	12
Ingenjörfirma	12
Nya Damfriseringen	12
Alltjänst	11
J. Näslund	11
J.A. Johansson	11
Nya Modemagasinet	11
Byggnadsfirma	10
Centralaffären	10
E. Eriksson	10
Erik Johansson	10
J. Andersson	10
Nya Modeaffären	10

Prev 1 2 3 4 5 ... 2002 Next

In this example, the element called “DisplayActor” appears to contain a mixture of people’s names and place/organisation names.

## 4 MoRE repository

The MoRE repository (<http://store.locloud.eu>) is a web-based platform to support metadata harvesting, ingestion, validation, enrichment and delivery to Europeana. The main activities on MoRE are:

- Harvesting metadata sources including MINT, LoCloud Collections and also OAI-PMH repositories maintained by LoCloud partners
- Metadata ingestion
- Metadata enrichment
- Metadata transformation to EDM
- Publication of data to Europeana.

### 4.1 Metadata ingestion

The MoRE interface allows administrators to see which packages of content are pending, ingestion and which have been ingested, published, withdrawn or rejected. Currently there are:

- 23 packages pending ingestion
- 29 packages have been ingested but not yet published
- 62 packages have been published
- 56 packages have been withdrawn
- quite a large number of rejected packages

The number of packages that have been withdrawn and rejected reflects testing of MORE by partners as they prepare their datasets for publication, and as a result of the hands-on training during the LoCloud training workshops.

#### 4.1.1 Ingested metadata packages

11 partners have completed the process of ingesting their metadata packages to MORE. The table below shows the metadata format and the status of the metadata. Some metadata packages need to be transformed from the schema provided to EDM, while those provided in EDM are simply ingested. Several packages have been enriched following ingestion. The table below shows that several datasets contain the same number of items. This is an indication of the test activity by partners – with re-ingestion of the same dataset in a slightly different format to evaluate the results.

Package	Partner	Schema	Items	Status
890	ADS	OAI_DC	100	Transformed
884	ADS	OAI_DC	100	Transformed
1275	AIT	ESE	256	Enriched
1683	BJC	EDM	240	Enriched
1674	BJC	EDM	106	Enriched
1306	DP	EDM	72	Ingested
1114	FRS	EDM	1104	Ingested
1074	FRS	EDM	1104	Ingested

Package	Partner	Schema	Items	Status
1056	FRS	EDM	1104	Enriched
1012	FRS	EDM	1104	Ingested
1000	FRS	EDM	1104	Ingested
939	FRS	EDM	1104	Enriched
924	FRS	EDM	1080	Ingested
1571	GKR	OAI_DC	10	Transformed
1570	GKR	OAI_DC	1373	Transformed
1664	GKR	EDM	1373	Ingested
1662	GKR	OAI_DC	1373	Transformed
1660	GKR	EDM	1258	Enriched
1643	GKR	OAI_DC	1373	Transformed
1641	GKR	OMEKA_XML	10	Transformed
922	IPCHS	LIDO	3885	Transformed
919	IPCHS	EDM	3872	Ingested
1024	KUAS	LIDO	10	Transformed
1621	NRA	OMEKA_XML	1	Enriched
1576	PL	LIDO	283	Enriched
1712	PL	EDM	12184	Ingested
1628	PL	EDM	10	Ingested
1269	PSRL	ESE	1054	Ingested
1112	PSRL	ESE	8951	Transformed

#### 4.1.2 Published metadata packages

10 partners have published metadata packages, which is an indication that these are considered ready for harvesting by Europeana. The table below lists both dataset that have already been made available to Europeana for harvesting, and those that are about to be harvested. A small number of test datasets (including the packages from NRA) are included in the table.

#	Partner	Schema	Items	Published schema
1544	CUT	OAI_DC	31	eEDM
1543	CUT	OAI_DC	557	eEDM
1542	CUT	OAI_DC	434	eEDM
1541	CUT	OAI_DC	1	eEDM
1540	CUT	OAI_DC	241	eEDM
1539	CUT	OAI_DC	52	eEDM
1538	CUT	OAI_DC	57	eEDM
1537	CUT	OAI_DC	1	eEDM
1536	CUT	OAI_DC	120	eEDM
1535	CUT	OAI_DC	159	eEDM
1534	CUT	OAI_DC	59	eEDM
1533	CUT	OAI_DC	42	eEDM
1532	CUT	OAI_DC	51	eEDM
1530	CUT	OAI_DC	122	eEDM
1529	CUT	OAI_DC	26	eEDM

1528	CUT	OAI_DC	125	eEDM
1527	CUT	OAI_DC	131	eEDM
1525	CUT	OAI_DC	21	eEDM
1524	CUT	OAI_DC	101	eEDM
1252	CUT	OAI_DC	168	eEDM
1250	CUT	OAI_DC	68	eEDM
1239	CUT	OAI_DC	323	eEDM
1234	CUT	OAI_DC	16	eEDM
1226	CUT	OAI_DC	31	eEDM
1224	CUT	OAI_DC	1032	eEDM
1223	CUT	OAI_DC	29	eEDM
1220	CUT	OAI_DC	1	eEDM
1213	CUT	OAI_DC	87	eEDM
1211	CUT	OAI_DC	2	eEDM
1208	CUT	OAI_DC	9	eEDM
1207	CUT	OAI_DC	37	eEDM
1205	CUT	OAI_DC	63	eEDM
1204	CUT	OAI_DC	21	eEDM
1202	CUT	OAI_DC	337	eEDM
1200	CUT	OAI_DC	11	eEDM
1199	CUT	OAI_DC	12	eEDM
1184	CUT	OAI_DC	209	eEDM
1523	DP	EDM	85	EDM
1521	DP	EDM	81	EDM
1520	DP	EDM	39	EDM
1519	DP	EDM	55	EDM
1127	FRS	EDM	1104	EDM
1705	HU	EDM	1138	EDM
923	IPCCHS	EDM	3872	EDM
1492	Jara	ESE	2395	EDM
1117	Jara	ESE	11906	EDM
1468	NPU	CARARE20	5986	eEDM
1058	NRA	OAI_DC	3	EDM
1032	NRA	OAI_DC	2	EDM
1027	NRA	OAI_DC	2	EDM
1272	PSRL	ESE	1054	EDM
1265	VUFK	EDM	3068	EDM
1264	VUFK	EDM	2000	EDM
1263	VUFK	EDM	4000	EDM
1262	VUKF	EDM	4000	EDM
1261	VUFK	EDM	4000	EDM
1345	Test	EDM	11	EDM
1163	Test	OAI_DC	11	EDM
1146	Test	OAI_DC	11	EDM
888	Test	OAI_DC	11	EDM
875	Test	OAI_DC	11	EDM
873	Test	OAI_DC	29	EDM

The table shows whether the metadata provided to Europeana has been enriched on MORE (eEDM) or was provided to Europeana without enrichment (EDM).

Please note that the series of packages from CUT were harvested directly by MORE from their Omeka repository in a Dublin Core metadata format, enriched and then transformed to EDM

### 4.1.3 Package information

More detailed information is available about each package. The figure below shows the package details including the number of items that were harvested, validated, transformed and enriched. The package statistics show the percentage completeness of the EDM mandatory and strongly recommended elements.

**Package Information** Project LoCloud Home > Projects > LoCloud > 1705

**Package 1705**

Total items: 1138      Metadata source: Hacettepe Universitesi\_MINT

Current Status: **Published**

Harvested   Validated   Ingested   Published

Available metadata schemas: EDM

History   Notifications   Reject Package

View Items   Withdraw

**Statistics**

EDM

Element	Percentage
isShownAt	1.0
aggregatedCHO	1.0
title	1.0
edm:type	1.0
isShownBy	1.0
spatial	1.0
description	0.9
subject	0.75
rights	1.0
provider	1.0
type	1.0
data_provider	1.0

**Package Details**

Type	Results
Harvested items	1138
Validated items (structural)	1138
Validated items (schema)	1138
Items with validation errors	0
Transformed items	1138
Enriched items	0

**Completeness**

EDM

Element	Percentage
total_strongly_recommended	100
total_mandatory	100

These statistics, together with the facility to view items in the package, enable the metadata to be quality assured before it is submitted to Europeana for ingestion.

### 4.1.4 View items

From the package information it is possible to view all of the items contained within.

View Items list of items for selected packages Home > LoCloud > Package 420

Back Project LoCloud :: Package 420 Search

#	Item	Native ID	
1	<a href="#">Γεωργός στο Δάλι</a>	oai:https://apsida.cut.ac.cy:11861	<input type="button" value="OAI_DC"/> <input type="button" value="EDM"/> <input type="button" value="eEDM"/>
2	<a href="#">Καλλιέργειες (ψέκασμα) στην Αθηνώνου</a>	oai:https://apsida.cut.ac.cy:11862	<input type="button" value="OAI_DC"/> <input type="button" value="EDM"/> <input type="button" value="eEDM"/>

For each item it is possible to view the metadata in both its native format, in this case OAI\_DC, and transformed to EDM. This allows for checking of the semantic content of a sample of records within the dataset.

```

- <oai_dc:dc xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/oai_dc/ http://www.openarchives.org/OAI/2.0/oai_dc.xsd">
  <dc:title>Γεωργός στο Δάλι</dc:title>
  - <dc:creator>
    Γραφείο Τύπου και Πληροφοριών (Κυπριακή Δημοκρατία)
  </dc:creator>
  <dc:subject>Αγροτική ζωή--Κύπρος</dc:subject>
  <dc:subject>Agriculture life--Cyprus</dc:subject>
  <dc:subject>Αγρότες--Κύπρος--Ιστορία </dc:subject>
  <dc:description>Ο γεωργός μαζεύει το σιτάρι στο χωριό Δάλι.</dc:description>
  <dc:publisher>Βιβλιοθήκη Τεχνολογικού Πανεπιστημίου Κύπρου</dc:publisher>
  - <dc:publisher>
    Εργαστήριο Ψηφιακής Πολιτιστικής Κληρονομιάς Τεχνολογικού Πανεπιστημίου Κύπρου
  </dc:publisher>
  - <dc:contributor>
    Γραφείο Τύπου και Πληροφοριών (Κυπριακή Δημοκρατία)
  </dc:contributor>
  <dc:date>1950-1960</dc:date>
  <dc:type>Image</dc:type>
  <dc:format>JPG</dc:format>
  <dc:identifier>1A-007-003</dc:identifier>
  <dc:identifier>https://apsida.cut.ac.cy/items/show/11861</dc:identifier>
  - <dc:identifier>
    https://apsida.cut.ac.cy/files/original/5e97b03af8ec01c8332f766eb7b9c8f9.jpg
  </dc:identifier>
  - <dc:source>
    Γραφείο Τύπου και Πληροφοριών (Κυπριακή Δημοκρατία)
  </dc:source>
  <dc:language>EL, EN</dc:language>
  <dc:coverage>35.021093, 33.419368</dc:coverage>
  - <dc:rights>
    Απαγορεύεται η δημοσίευση ή αναπαραγωγή, ηλεκτρονική ή άλλη χωρίς τη γραπτή συγκατάθεση του δημιουργού.
  </dc:rights>
</oai_dc:dc>

```

The figure above shows the native metadata and the figure below the metadata transformed to EDM.

```

- <edm:ProvidedCHO rdf:about="http://more.loccloud.eu/object/CUT/11861">
  <dc:title xml:lang="">Γεωργός στο Δάλι</dc:title>
  <dc:description xml:lang="">Ο γεωργός μαζεύει το σιτάρι στο χωριό Δάλι.</dc:description>
  <dc:subject>Αγροτική ζωή--Κύπρος</dc:subject>
  <dc:subject>Agriculture life--Cyprus</dc:subject>
  <dc:subject>Αγρότες--Κύπρος--Ιστορία </dc:subject>
  <dc:terms:spatial rdf:resource="http://more.loccloud.eu/CUT/11861/SP.2"/>
  <edm:type>IMAGE</edm:type>
- <dc:creator xml:lang="">
  Γραφείο Τύπου και Πληροφοριών (Κυπριακή Δημοκρατία)
  </dc:creator>
- <dc:contributor xml:lang="">
  Γραφείο Τύπου και Πληροφοριών (Κυπριακή Δημοκρατία)
  </dc:contributor>
  <dc:date>1950</dc:date>
  <dc:date>1960</dc:date>
  <dc:publisher xml:lang="">Βιβλιοθήκη Τεχνολογικού Πανεπιστημίου Κύπρου</dc:publisher>
- <dc:publisher xml:lang="">
  Εργαστήριο Ψηφιακής Πολιτιστικής Κληρονομιάς Τεχνολογικού Πανεπιστημίου Κύπρου
  </dc:publisher>
- <dc:source>
  Γραφείο Τύπου και Πληροφοριών (Κυπριακή Δημοκρατία)
  </dc:source>
</edm:ProvidedCHO>
- <edm:WebResource rdf:about="https://apsida.cut.ac.cy/files/original/5e97b03af8ec01c8332f766eb7b9c8f9.jpg">
  <dc:format>JPG</dc:format>
  - <dc:rights xml:lang="">
    Απαγορεύεται η δημοσίευση ή αναπαραγωγή, ηλεκτρονική ή άλλη χωρίς τη γραπτή συγκατάθεση του δημιουργού.
    </dc:rights>
    <edm:rights rdf:resource="http://www.europeana.eu/rights/rr-f"/>
  </edm:WebResource>
- <edm:Place rdf:about="http://more.loccloud.eu/CUT/11861/SP.2">
  <skos:prefLabel>35.021093, 33.419368</skos:prefLabel>
  </edm:Place>
- <ore:aggregation rdf:about="http://more.loccloud.eu/object/CUT/11861#aggregation">
  <edm:aggregatedCHO rdf:resource="http://more.loccloud.eu/object/CUT/11861"/>
  <edm:dataProvider>Βιβλιοθήκη Τεχνολογικού Πανεπιστημίου Κύπρου</edm:dataProvider>
  <edm:provider>LoCloud</edm:provider>
  <edm:rights rdf:resource="http://www.europeana.eu/rights/rr-f"/>
  <edm:isShownAt rdf:resource="https://apsida.cut.ac.cy/items/show/11861"/>
  <edm:isShownBy rdf:resource="https://apsida.cut.ac.cy/files/original/5e97b03af8ec01c8332f766eb7b9c8f9.jpg"/>
  <edm:object rdf:resource="https://apsida.cut.ac.cy/files/original/5e97b03af8ec01c8332f766eb7b9c8f9.jpg"/>
  </ore:aggregation>

```

---

## 5 Content in Europeana

The first LoCloud collections have been published in Europeana at the time of writing this report. The content can be browsed by using the following links to search results on [www.europeana.eu](http://www.europeana.eu):

- [LoCloud \(26983\)](#)
  - [Vilnius University Faculty of Communication \(17068\)](#)
  - [Library of Cyprus University of Technology \(4214\)](#)
  - [eheritage.si, Zavod za varstvo kulturne dediščine Slovenije \(3872\)](#)
  - [KAMRA \(1829\)](#)

The content includes:

- 26882 images
- 60 text objects
- 21 video files
- 20 sound files

Whose copyright is as follows:

- 25470 – Rights reserved: Free access
- 1470 CC BY-NC
- 46 Unknown copyright status

Additional content is already in the Europeana publication pipeline and will appear in the portal in due course. Please note that new content is uploaded to the portal by Europeana once a month.



## 6 Metadata quality

In January 2014 Europeana convened a task force to work on metadata quality to discuss how the quality of metadata supplied to Europeana could be improved and how this would improve end user experience. The expected outcomes of the task force were to:

- Document what Europeana means by metadata quality.
- Create best practice guidelines for creating good metadata quality records and good metadata quality fields.
- Make recommendations on metadata fields per media type.
- Provide documents to clarify the discrepancies between metadata displayed in the Europeana portal and provided data.

The timeframe of the task force has been extended and at the time of writing work is expected to be completed by September 2015.

EDM specifies a series of mandatory metadata elements, these are as follows:

- edm:dataProvider
- edm:isShownAt **or** edm:isShownBy
- edm:provider
- edm:rights
- dc:title **or** dc:description
- dc:language for text objects
- dc:subject **or** dc:type **or** dc:coverage **or** dc:spatial
- edm:type

The recommended metadata elements are:

- edm:object
- dcterms:alternative
- dc:creator
- dc:contributor
- dc:date
- dcterms:created
- dcterms:issued
- dcterms:temporal
- dc:publisher
- dc:source
- dcterms:isPartOf
- edm:isNextInSequence (for hierarchical objects)

Thus one aspect of metadata quality is **completion** of the mandatory and recommended metadata elements. Records that do not fulfill the minimum mandatory requirements are automatically rejected by Europeana and will not be published in the portal.

There are other aspects of metadata quality which are routinely checked at the time of publication, these include:

- Uniqueness of identifiers
- Currency of links, i.e. check the links are not broken
- Coherence of the rights statement. This considers whether the edm:rights statement (which is chosen from a standard list published by Europeana at: <http://pro.europeana.eu/available-rights-statements>) consistent with any rights statement included in the dc:rights element.

Other aspects of metadata quality, which can be checked for include:

- Use of international controlled vocabularies such as Geonames
- The length of titles and descriptive text
- Accuracy of geographical coordinates (for example, points lie within the expected region when mapped)

Length of text may not appear to be an aspect of metadata quality, however very short titles and descriptions offer limited information to end-users. They also limit the possibility of enriching the indexing of metadata records by using parsers to identify place names, personal names or other keywords.

The accuracy of geographic coordinates can be checked visually by plotting the coordinates onto a map. If a collection of heritage material from a northern European country is plotted somewhere in the South Pacific, there is a problem with the coordinates. Smaller discrepancies are less obvious, but can be checked and corrected using the LoCloud Geo microservice.

Apart from the availability of a reasonable description of an object, other aspects that Europeana has identified as being important for content re-use and having an impact on the quality of user experience include:

- The availability of a thumbnail or preview
- Provision of a direct link from the metadata record to an object (rather than to a landing page)
- Provision of higher resolution images
- Openly licencing which allows for re-use of the content that is made available

## 7 Conclusions

During the period under review LoCloud partners have made good progress with their content workflows.

Given the diversity of the content being provided by LoCloud partners it is difficult to identify standard metrics. However it is possible to monitor how the metadata compares to the recommendations of the Europeana quality task force, and to focus on improving those aspects that will have most impact on improving the representation of the content in Europeana, its potential for re-use and improving user experiences.

Statistical monitoring tools have been put in place to monitor the conformance of the metadata being provided to mandatory and recommended metadata elements. LoCloud micro-services now implemented on the LoCloud aggregator can be used to address aspects of metadata quality that improve its retrieval and usability. For example, the language of the metadata can be identified and language attributes added to the metadata.

Human monitoring of the metadata is also important. LoCloud partners are monitoring the metadata provided by their partners and working with them to improve its quality. In addition a sample of each collection is monitored by the project coordination, the staff of the LoCloud aggregation service and the content ingestion team at Europeana, and feedback is provided to relevant partner or content creator.

## Annex 1: Event log

LoCloud has put in place a modification of the Events Log system used in EuropeanLocal; the system was set up by Avinet. Data about partners' collections captured during the content survey carried out in April-May 2014 has been ingested into the system, which was launched as a fully functional service in September 2014, in time for monitoring at month 18.

In the context of this tool, the term event relates to when an activity in the LoCloud workflow takes place on an individual collection and is intended to allow progress through the workflow to be monitored.

The event log is online at: <http://locloudlog.avinet.no>

### Overview of the Event log

The Event Log is a simple database with online data entry forms, reporting and analyzing functions.

When a content provider or aggregator has done work on one or several collections, he/she will connect to the Event Log, select the collection involved and add information from the predefined list of events. At a minimum this can be just some very simple core facts that will take a few minutes to add, sufficient to document the action and enter amount of items and objects and date the work was done. This will be enough to maintain data about amount and progress.

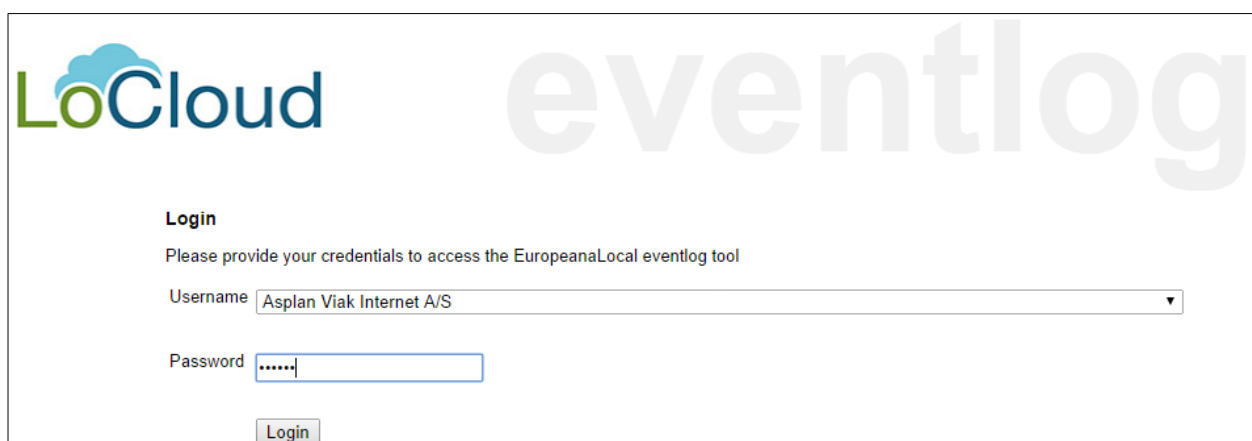
If the content provider or the aggregator encounters any problems or wishes to share lessons learned and ideas about improving the processes, this can be added as a manual report and thus being made available to the rest of the partners in the project.

### Updating the Event log: content providers perspective

#### User accounts and Log in

User accounts have been set up for each partner to enable them to report on progress as the collection moves through the workflow. These accounts are shared with all of the LoCloud support services.

The log-in screen provides a drop-down list of the organisation names of registered users to simplify the process.



**LoCloud** eventlog

**Login**  
Please provide your credentials to access the EuropeanLocal eventlog tool

Username

Password

## Partner information

Once partners are logged into the system they will see a list of options, which allow them to view their collections, a series of reports and their partner details.

The “Add & Edit” section includes the option to select “Partner”.

This displays a list of all partners that the user is authorized to edit. This may be limited to their own organization, but can include other organisations in their region. The administrator view (figure below) shows all registered organisations.

### Add or edit partner information

On this page you can add new partners or edit information about regist

[Add new](#)

Select	Country	Partn
	AT	Angewandte Informationstechnik Forschungsgesellsch
	BE	Provincie Limburg
	BG	Pencho Slaveykov Regional Library
	CY	Cyprus University of Technology
	CZ	Národní památkový ústav (National Heritage Institute)
	DE	Universitaet Duisburg-Essen (University of Duisburg-E
	DK	Kulturarvsstyrelsen (Danish Agency for Culture)
	ES	Ministerio de Educacion, Cultura y Deporte (Ministry o
	FR	Conseil Général de la Gironde - Archives Départemen
	GR	Future Library
	HR	Gradska knjiznica Rijeka
	IE	Discovery Programme
	IS	Minjastofnun Íslands /The Cultural Heritage Agency of
	IT	Fondazione Ranieri di Sorbello

Updating the details is a simple process of selecting the partner to update by clicking on the green button in the first column of the table. The details about the organization are then displayed at the top of the page. Clicking the “edit” link allows the form to become editable.

Country	Norway
Region	N/A
Partner name	Asplan Viak Internet A/S
Password	*****
Plan to use MINT	<input checked="" type="checkbox"/>
Plan to use MORE	<input checked="" type="checkbox"/>
Plan to use LoCloud Collections (LDL)	<input type="checkbox"/>
Plan to test microservices	geo,enrich
<a href="#">Edit</a>	
<a href="#">Add new</a>	

**LoCloud**

**You are currently logged in as: Asplan Viak Internet A/S**

**View planned...**

- [My collections](#)
- [Collections \(by partner\)](#)
- [Content \(by partner\)](#)
- [Content \(by partner\) - diagram](#)
- [Content ingestion plan](#)
- [Content ingestion summary](#)
- [Partner map](#)

**View progress...**

- [Partners' status](#)

**Add & Edit**

- [Partner](#)
- [Collections](#)
- [Event](#)

**Admin...**

- [Logout](#)

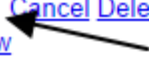
The details include the organisation name, country and region and an indication of which of the LoCloud services the partner plans to use.

Users can update their username and also their password - but we do not encourage this, as we would like partners to maintain the same login across all LoCloud systems.

Country	Norway
Region	N/A
Partner	Asplan Viak Internet A/S
Password	.....
Plan to use MINT	<input checked="" type="checkbox"/>
Plan to use MORE	<input checked="" type="checkbox"/>
Plan to use LoCloud Collections (LDL)	<input type="checkbox"/>
Plan to test microservices	geo.enrich
Longitude	
Latitude	

The longitude and latitude of each partner’s home office is stored to enable a “partner map” to be created.

[Update](#) [Cancel](#) [Delete](#)  
[Add new](#)



### Information that is stored about collections

Clicking on “Collection” under the “Add & Edit” heading displays a list of all collections belonging to the user.


**Add or Edit Collections**

This form allows you to add new collections or edit collections belonging to the currently logged in partner.

**Selected collection details**

Please select a collection below to edit it or [click here to add a new collection](#)

**Collections**

Select	Collection name
	LoCloud Test collection

Each collection can be selected by clicking on the green button in the first column of the table. This displays the collection information at the top of the page.

**Selected collection details**

Content provider	AVINET
Collection name	LoCloud Test collection
Planned harvest date	20.09.2014
Europeana DEA signed	completed
Metadata export format	EDM
Europeana rights statement	1 The Public Domain Mark (PDM)
Thumbnails available	Yes, for all digital objects
Example URL (isShownBy/At)	http://www.avinet.no
Long term consistency of URLs	Yes, for all URLs
Description	This is a test collection for LoCloud
Number of digital objects	1 text objects 1 image objects 1 sound objects 1 video objects 1 3D objects
<a href="#">Edit</a> <a href="#">Delete</a> <a href="#">New</a>	
<a href="#">Add new collection</a>	

Clicking the “edit” link below the detailed information about the collection allows the form to become editable and the details to be updated.

**Selected collection details**

Content provider	<input type="text" value="AVINET"/>
Collection name	<input type="text" value="LoCloud Test collection"/>
Planned harvest date	20.09.2014 <input type="text" value="20.09.2014"/>
Europeana DEA signed	completed <input type="text" value="completed"/>
Metadata export format	<input type="text" value="EDM"/>
Europeana rights statement	1 The Public Domain Mark (PDM) <input type="text" value="1 The Public Domain Mark (PDM)"/>
Thumbnails available	Yes, for all digital objects <input type="text" value="Yes, for all digital objects"/>
Example URL (isShownBy/At)	<input type="text" value="http://www.avinet.no"/>
Long term consistency of URLs	Yes, for all URLs <input type="text" value="Yes, for all URLs"/>
Description	<input type="text" value="This is a test collection for LoCloud"/>
Number of digital objects	<input type="text" value="1"/> text objects <input type="text" value="1"/> image objects <input type="text" value="1"/> sound objects <input type="text" value="1"/> video objects <input type="text" value="1"/> 3D objects
<input type="text" value="Update"/> <input type="text" value="Cancel"/>	
<a href="#">Add new collection</a>	

The information stored about each collection is designed to complete the “Content Contribution Form” requested by Europeana prior to ingestion. It includes:

- The name of the content provider – in some cases this is different from the organization participating in LoCloud.
- Collection name – for example, the name of the collection given in the content survey.

- Planned harvest date – the date when the partner aims to have the collection ready for harvesting by Europeana
- Whether the Europeana Data Exchange Agreement (DEA) has been signed for the collection
- What format the exported metadata from your collection is available in, i.e. EDM, ESE or similar
- The Europeana Rights Statement - a declaration of the usage rights for the content of the collection. Contributors select between 13 predefined licenses specified by Europeana: <http://pro.europeana.eu/available-rights-statements>
- Whether thumbnails are available for the objects in the collection when published in Europeana
- An example URL – an example of the URL to be provided in the EDM isShownBy or isShownAt element (i.e. a direct link to the object or a landing page for the object)
- Description – for notes about the collection or simply a description of its content.
- The number of digital objects categorized by texts, images, sounds, videos and 3d objects

## Adding events to collections

Clicking on “Event” under the “Add & Edit” allows events to be added to each collection. The page opens displaying one of the collections belonging to the currently logged in user. A dropdown list allows users to choose which collection to work on.

**Select the collection you wish to add events for**

LoCloud Test collection ▾ ← Select collection here

---

**Selected collection details**

Content provider	AVINET
Collection name	LoCloud Test collection
Planned harvest date	20.09.2014
Europeana DEA signed	in progress
Metadata export format	EDM
Europeana rights statement	1 The Public Domain Mark (PDM)
Thumbnails available	Yes, for all digital objects
Example URL (isShownBy/At)	<a href="http://www.avinet.no">http://www.avinet.no</a>
Long-term consistency of URLs	Yes, for all URLs
Description	This is a test collection for LoCloud
Number of digital objects	5 text objects 1 image objects 1 sound objects 1 video objects 1 3D objects

← Collection details here

---

**Events**

← List of events here

[Add new event](#)

Date	Event	Status	Information about the event	Delete
09.09.2014	IPR cleared	completed		<a href="#">Delete</a>
09.09.2014	Metadata captured	completed		<a href="#">Delete</a>

The page displays details about the collection and, at the bottom of the page, a list of events for the collection.

Existing events can be deleted by choosing the “Delete” link.



## Events

[Add new event](#)

Date	Event	Status	Information about the event	Delete
09.09.2014	IPR cleared	completed		<a href="#">Delete</a>
09.09.2014	Metadata captured	completed		<a href="#">Delete</a>

New event can be added to the collection by clicking the “Add new events” link below the header “Events”. This displays the editing form for a new event.

### Events

[Add new event](#)

Input date of event	09.09.2014
Select an event type	IPR cleared
Select status for the event	in progress
Comments	
Attachments	<input type="button" value="Choose File"/> No file chosen

[Insert](#) [Cancel](#)

Date	Event	Status	Information about the event	Delete
09.09.2014	IPR cleared	completed		<a href="#">Delete</a>
09.09.2014	Metadata captured	completed		<a href="#">Delete</a>

By default, the form will be filled in with today’s date – this can be changed using the date picker or by entering a date in the format dd.mm.yyyy.

## Events in the life-cycle of a collection

There are eight pre-defined events in the life-cycle of a collection; these are:

1. IPR cleared
2. Metadata captured
3. Metadata enriched
4. Content published
5. Persistent URIs established
6. Metadata exported
7. Metadata harvested
8. Validation completed

For each collection, we would like you to fill in the information as the various stages are reached. Each event has a date and a status that can be either “in progress” or “completed”. Any other value will be assumed to be not done or incomplete.

Users can write a comment related to the event and upload an attachment if relevant, but this is not mandatory.

## Reports from the event log

The Event Log is an important tool for collecting and sharing information about the amount of content and progress in LoCloud. Reports can be selected by country, region, type of content, time span, amount etc.

The Event Log documents what happened to each collection. It will enable many interesting analyses of the data to be performed, for example:

- The number of items and objects harvested from local providers at any given time. A report organized by country will be updated on a daily basis by the system.
- The number of collections, items and objects are entered in the Event Log at any given time? (arranged by country, domain etc).
- How many digital objects from a chosen number of collections are extracted and mapped by a given date.
- How many items were affected by metadata normalization and enrichment from a country, a region, certain types of collections etc in a given period of time.
- How many items and objects are harvested into LoCloud repository by any date, country, region, provider type (museum, archive, library).
- Items and objects harvested by Europeana by end of year 2, end of year 3.
- The number of providers and collections involved in LoCloud at any given time.
- The number of manual reports added, from which country, provider, etc.