



**Project Acronym: Europeana v2**  
**Grant Agreement number: 270902**  
**Project Title: Europeana Version 2**

## **MS28: Data Ingestion Plan**

<b>Revision</b>	[1.0]
<b>Date of submission</b>	1 July 2012
<b>Author(s)</b>	Annelies van Nispen - Europeana Foundation
<b>Dissemination Level</b>	[Public]

## REVISION HISTORY AND STATEMENT OF ORIGINALITY

### Revision History

Revision No.	Date	Author	Organisation	Description
1.0	28 June 2012	Annelies van Nispen	Europeana Foundation	Final version Data Ingestion Plan

### Statement of originality:

This Milestone contains original unpublished work except where clearly indicated otherwise.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.

## Index

<b>INTRODUCTION</b> .....	4
<b>Chapter 1 Data Ingestion Procedure</b> .....	5
<b>Procedure</b> .....	5
<b>Step 1</b> .....	5
<b>Step 2</b> .....	6
<b>Step 3</b> .....	6
<b>Step 4</b> .....	6
<b>Step 5</b> .....	6
<b>Step 6</b> .....	6
<b>Chapter 2 Content via Aggregators and (data) providers -Aggregation Landscape</b> .....	7
<b>Chapter 3 Content via Projects</b> .....	8
<b>Chapter 4 Current status of content in Europeana</b> .....	9
<b>Content report from the 8th of June 2012</b> .....	10

## INTRODUCTION

This document will describe how the workflow to deliver content into Europeana is organised. The main actors in this are:

- Business
- Projects
- Operations-Ingestion

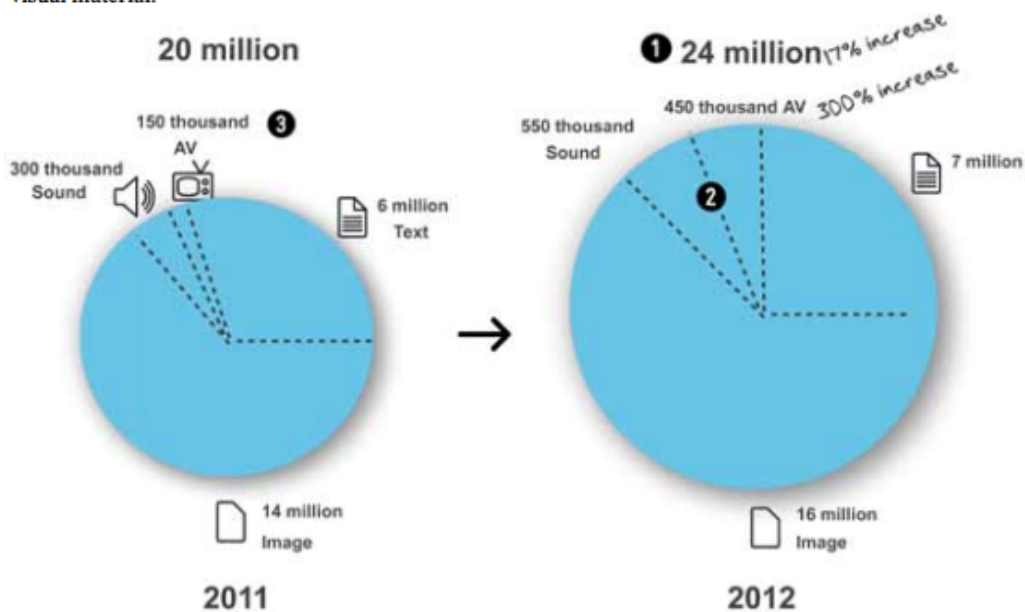
This document will describe the current procedure and planning of data ingestion flow. The content is delivered by:

- Aggregators and (data) providers
- Projects

In the [Annual Plan 2012](#) the following Objective, Key Tasks and KPI's are set:

### 1.2 Content

**Objective:** Improve the overall quality of the content repository with a special emphasis on the audio-visual material.



#### Key tasks:

1. Increase the Europeana repository and make routes for content delivery easy to understand.
2. Audio-visual material is viewed significantly more frequently than other media types on Europeana, yet it is under-represented in the repository. An increase in the amount of AV material from the existing and new partners is required.
3. Perform a GAP analysis of the content and so that the gaps are apparent to both partners and the European Commission.

#### KPIs

- KPI 4: A minimum of 24 million objects is available in the Europeana Repository.
- KPI 5: A minimum of 1 million sound and video material is available in the Europeana repository.

## Chapter 1 Data Ingestion Procedure

This document will describe the how the workflow to deliver content into Europeana is organized. The main actors in this are:

- Business
- Projects
- Operations

### Procedure

The procedure for delivering content to Europeana is available on [Europeana Pro](#) and this is the workflow for content delivery between Business and Operations-Ingestion. The content coming in via Projects follows this procedure. Projects and Business department manage this workflow. When projects are ready for delivery, they are handed over to the Operations department

### Procedure

Aggregators, projects and direct content providers should follow the steps below to contribute data to Europeana. Partners contributing indirectly to Europeana will be provided with the necessary requirements for the next steps by the aggregator or project in which they are involved.

Partner Request to Europeana
1. Data Exchange Information 1.1 Re-routing Information
2. Establish partnership between Europeana and Cultural Heritage Organisation 2.1 Sign Europeana Data Exchange Agreement
3. Cultural Heritage Organisation starts workflow with Operations Team 3.1 Sign Content Contribution Form
4. Operations Team provides feedback on harvesting and metadata mapping
5. Ingestion of provided datasets
6. Ingestion of updated and new datasets

Steps involved when delivering content to Europeana:

### Step 1

A partner institution interested in contributing data to Europeana will need to fill in the *Data Exchange Information Form*. The form identifies types and numbers of objects, data owners etc.

According to these information Europeana will be able to decide whether the partner can provide the data directly or through other Aggregators such as Projects or National Initiatives.

## **Step 2**

The Cultural Heritage Organisation and the Europeana Office formalise their direct collaboration by signing the *Europeana Data Exchange Agreement*.

## **Step 3**

The Cultural Heritage Organisation submits the *Content Contribution Form* (provided previously by the Europeana Operations Team) containing: type of submission (new/update), licensing and metadata information, transfer mechanism that will be used (preferably OAI-PMH) etc. Within one month the Europeana Operations Team reviews the submission request and gives feedback to the Organisation.

## **Step 4**

Consecutive to step 3, the Europeana Operations Team validates the data transfer infrastructure of the Cultural Heritage Organisation (Europeana Pro Portal, [Technical Requirements](#)).

Subsequently, the Operations Team provides guidance and feedback on harvesting issues and metadata mapping. The submitted datasets must conform to the 'ESEv3.4 or EDM specifications' and the '*Mapping and Normalisation Guidelines*'. These documents, together with the ESE v3.4 XML Schema, *Europeana Rights Guidelines* and *Europeana Portal Image policy* can be found in the [Technical Requirements](#) Section on the Europeana Pro Portal. In addition to the previous resources the Cultural Heritage Organisation can use the Content Checker. The Content Checker is a web tool that validates the mapping against the XML schema, and that simulates the ingestion of the data in a copy of the Europeana Portal.

## **Step 5**

The Europeana office runs the operations through the Europeana production environment and the publication process. When the operations are completed the Cultural Heritage Organisation receives a notification from the Europeana Operations Team that the data is published on the Europeana Portal.

## **Step 6**

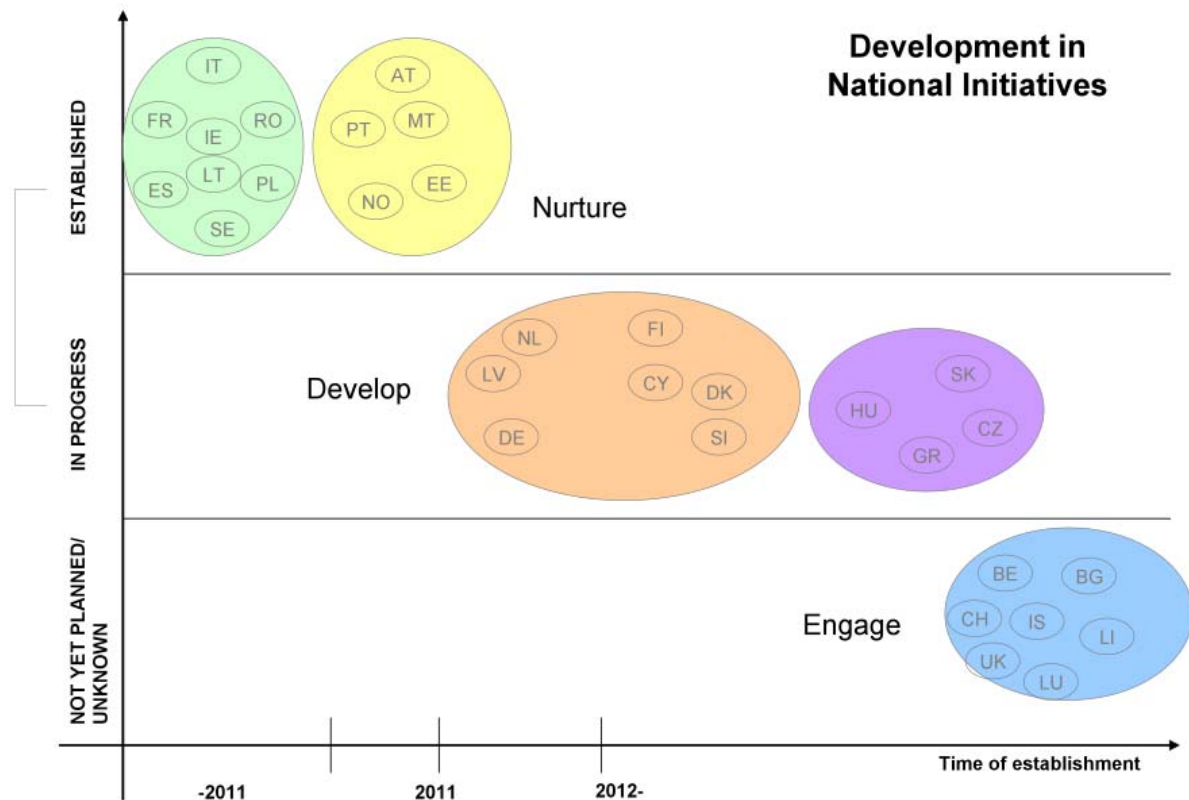
When the Cultural Heritage Organization updates or creates new datasets, the Europeana Operations Team will start the operations again from step 5

## Chapter 2 Content via Aggregators and (data) providers - Aggregation Landscape

The business department is responsible for developing the aggregation landscape and sets the priorities for delivering content to Europeana.

In the Strategic Plan, Annual Plan and the Content Strategy the objectives, tasks and KPI's are set (see Introduction). Business and Operations have weekly meetings to manage the incoming workflow and aggregation landscape.

### Overview of the Aggregation Landscape



## Chapter 3 Content via Projects

The projects deliver content to Europeana. Regularly the delivering starts at a late stage in the project. The Projects team of Europeana is responsible for managing. There are regular meetings between Projects, Business and Operations to manage the incoming workflow.

Overview of Projects providing content and their status

<b>Project</b>	<b>Content or Tech?</b>	<b>Status</b>	<b>Start</b>	<b>Finish</b>
TEL Plus	Providing Content	Complete	01-10-07	31-12-09
Arrow	Providing Technology	Complete	01-09-08	28-02-11
AccessIT	Providing Technology	Complete	01-05-09	29-04-11
Athena	Providing Content	Complete	01-11-08	29-04-11
Europeana Travel	Providing Content	Complete	01-05-09	29-04-11
Europeana Local	Providing Content	Complete	01-06-08	31-05-11
Europeana Version 1	Providing Technology & Content	Complete	01-02-09	29-07-11
EFG	Providing Content	Complete	01-09-08	31-08-11
MIMO	Providing Content	Complete	01-09-09	31-08-11
Europeana Connect	Providing Technology & Content	Complete	02-05-09	31-10-11
Apenet	Providing Content	Complete	15-01-09	30-12-11
Judaica Europeana Assets	Providing Content	Complete	01-01-10	30-12-11
	Providing Technology & Content	Complete	01-04-10	30-03-12
BHL-Europe	Providing Content	Complete	01-05-09	30-04-12
Europeana Regia	Providing Content	In Progress	01-01-10	29-06-12
PrestoPrime	Providing Technology	In Progress	01-01-09	29-06-12
EUScreen	Providing Content	In Progress	01-10-09	28-09-12
Europhoto	Providing Content	In Progress	01-05-10	31-10-12
Europeana Libraries	Providing Content	In Progress	01-01-11	31-12-12
CARARE	Providing Content	In Progress	01-02-10	31-01-13
Europeana Inside	Providing Technology	In Progress	01-04-12	01-04-13
AccessITplus	Providing Content	In Progress	01-05-11	30-04-13
HOPE	Providing Content	In Progress	01-05-10	30-04-13
ThinkMOTION	Providing Content	In Progress	01-06-10	31-05-13
DCA (Digitising Contemporary Art)	Providing Content	In Progress	01-01-11	28-06-13
ECLAP	Providing Content	In Progress	01-07-10	28-06-13
Arrow Plus	Providing Technology	In Progress	01-04-11	30-09-13
Linked Heritage	Providing Content	In Progress	01-04-11	30-09-13
Natural Europe	Providing Content	In Progress	01-10-10	30-09-13
Paths	Providing Technology	In Progress	01-01-11	31-12-13
EFG-1914	Providing Content	In Progress	15-02-12	14-02-14
Europeana Version 2	Providing Technology & Content	In Progress	01-10-11	28-02-14
Openup!	Providing Content	In Progress	01-03-11	28-02-14
Europeana Collections 1914-1918	Providing Content	In Progress	01-05-11	30-04-14
Europeana Awareness	Providing Technology & Content	In Progress	01-01-12	31-12-14
3D Icons	Providing Content	In Progress	01-02-12	30-01-15
DM2E (Digitised Manuscripts to Europeana)	Providing Content	In Progress	01-02-12	30-01-15
Europeana Photography	Providing Content	In Progress	01-02-12	30-01-15
Europeana Fashion	Providing Content	In Progress	01-03-12	27-02-15
Newspapers Online	Providing Content	In Progress	01-02-12	27-02-15
APEX	Providing Content	In Progress	01-03-12	31-03-15

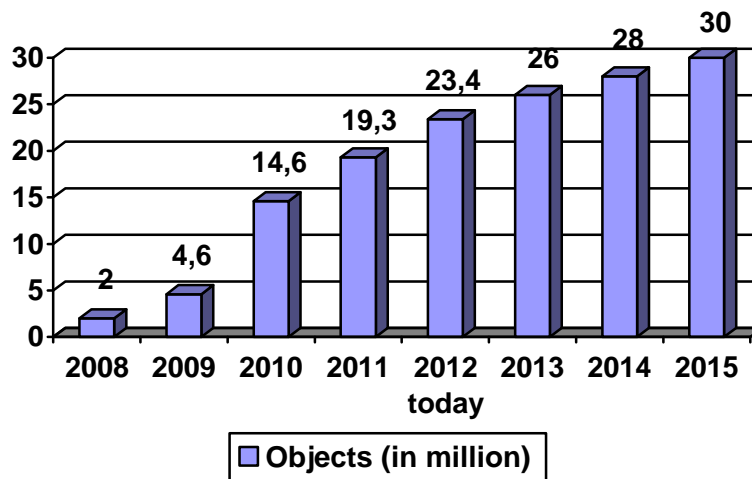


## Chapter 4 Current status of content in Europeana

During the past 12 months, the amount of content as known a raise of 26,48%<sup>1</sup>. In total there are 23.4 million of objects available at Europeana.eu. The KPI for 2012 is 24 million. The current number is 23.4 million objects and the Europeana Operations Team hopes to publish 23.9 million objects for the June publication.

As shown in graph number 1 below, this progress is the continuation of Europeana's exponential content growth (number of objects in million).

Graph 1: Europeana content development



The content report from the 8<sup>th</sup> of June 2012 gives the most up to date numbers. Europeana publishes a monthly content report.

The Operations department follow-ups on the priorities set by the Business and Projects department in the Strategic Plan, the Annual Plan and the Collection Strategy and development Plan.

## Content report from the 8th of June 2012

Content report from the 8 <sup>th</sup> of June 2012.....	10
1 European Overview .....	10
2 Content in Europeana by country TOP 15.....	10
3 Content in Europeana by content type .....	11
4 Content in Europeana by Copyright Status .....	13
5 Content in Europeana by content provider TOP 15.....	14
6 Content Type in Europeana by Country TOP15.....	14
6.1 IMAGE per Country TOP 15.....	14
6.2 TEXT per Country TOP 15 .....	15
6.3 SOUND per Country TOP 15.....	16
6.4 VIDEO per Country TOP 15 .....	17
6.4 3D per Country .....	18

### European Overview

Europeana.eu currently features content from 34 countries, in 29 languages. Several providers and projects have offered new content to Europeana in the last month: e.g. Rijksdienst voor het Cultureel Erfgoed, DK-National Aggregation Service, LT-Aggregator Service National Library of Lithuania, Faculdade de Ciências Sociais e Humanas da Universidade Nova de Lisboa, Europeana 1914 – 1918, Instituto dos Museus e da Conservação, Rijksmuseum, Amsterdam, Biblioteca de Catalunya, The European Library.

### Content in Europeana by country TOP 15

In total there are 23,4 million objects available in Europeana.eu, as of April 2010. The contribution by country is now more equally distributed. France, Germany, Sweden, Italy, Spain, Norway and the Netherlands provided most of the content in Europeana.eu. Countries contributing less than one percent of the content are Greece, Hungary, Switzerland, Slovakia, Denmark, Luxembourg, Estonia, Malta, Turkey, Iceland, Russia, Bulgaria, Czech Republic, Romania, Latvia, Portugal, Serbia, Lithuania, Croatia and Cyprus.

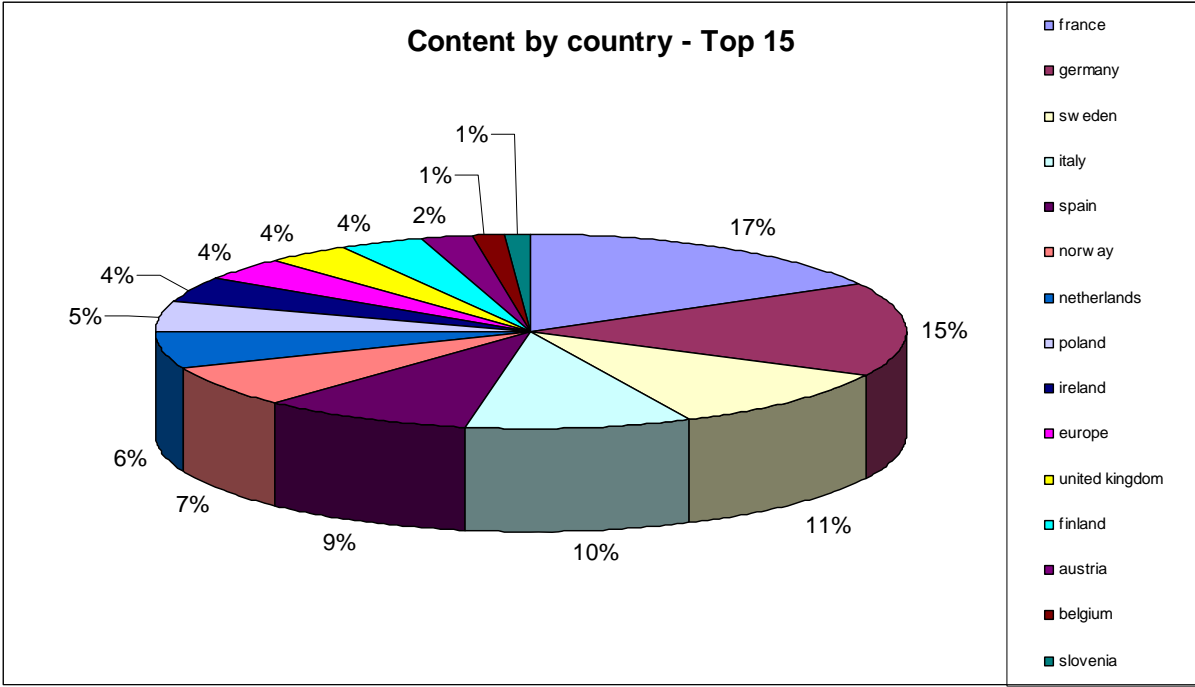


Figure 1 Europeana TOP 15 by country

Country	Total	Percentage
<b>Total of all records</b>	<b>23.431.250</b>	
france	3766370	16,07%
germany	3449297	14,72%
sweden	2360049	10,07%
italy	2199392	9,39%
spain	1980929	8,45%
norway	1557820	6,65%
netherlands	1390681	5,94%
poland	1093705	4,67%
ireland	953995	4,07%
eu rope	872407	3,72%
united kingdom	800797	3,42%
finland	800046	3,41%
austria	449633	1,92%
belgium	330479	1,41%
slovenia	248005	1,06%

**Content in Europeana by content type**

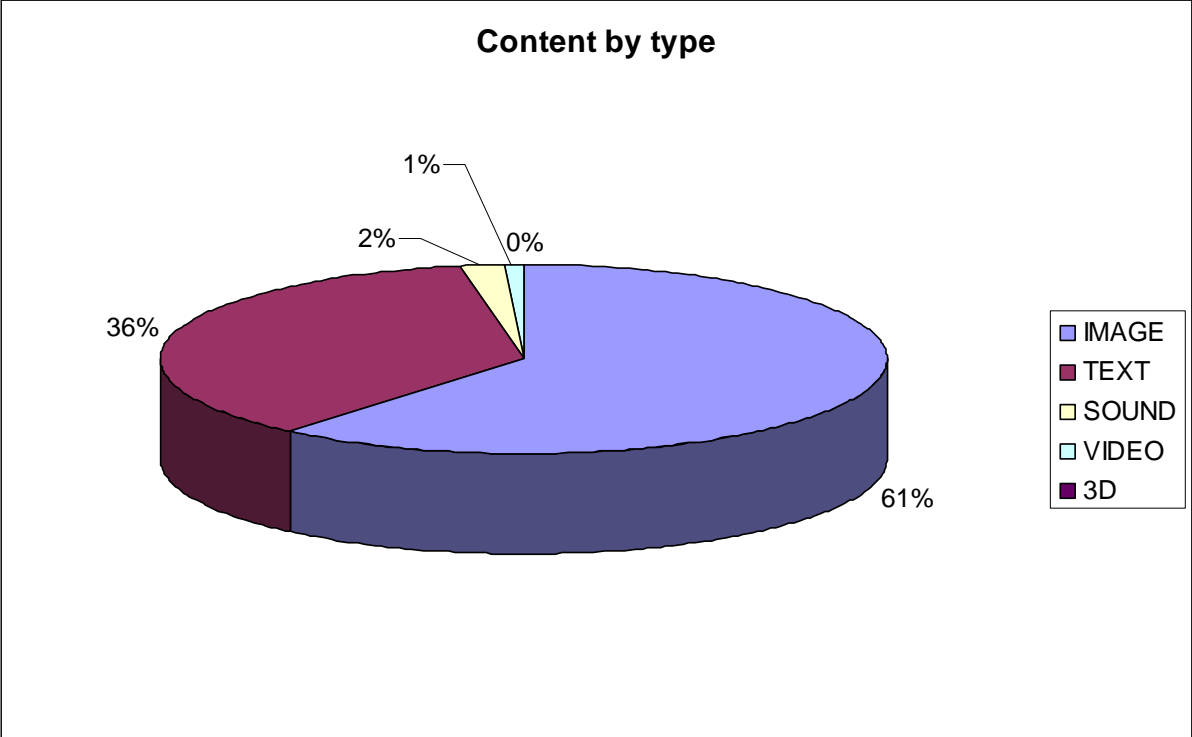


Figure 2 Europeana content by type

Type	Total	Percentage
<b>Total of records</b>	<b>23.431.250</b>	
IMAGE	14339091	61%
TEXT	8467581	36%
SOUND	454070	2%
VIDEO	170506	1%
3D	2	0%

## Content in Europeana by Copyright Status

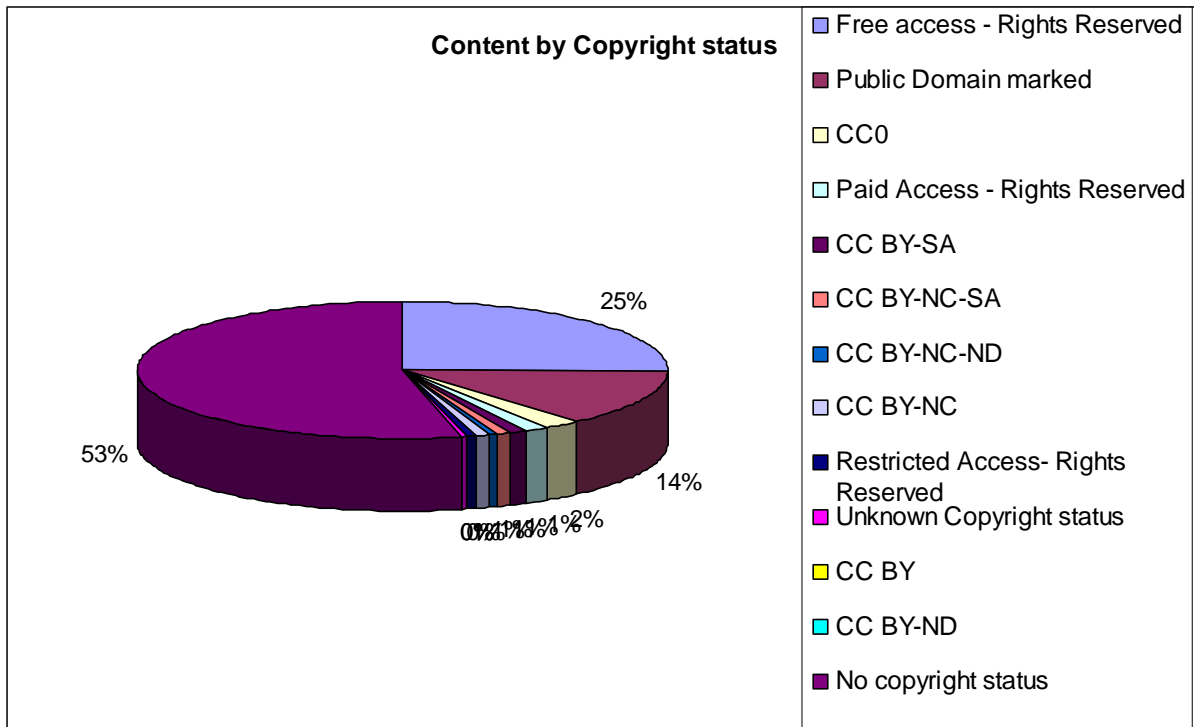


Figure 3 Europeana content by Copyright status

Rights statement	Total	Percentage
<b>Total of records</b>	<b>23.431.250</b>	
Free access - Rights Reserved	5.896.508	25%
Public Domain marked	3.167.252	14%
CC0	503.332	2%
Paid Access - Rights Reserved	317.370	1%
CC BY-SA	278.848	1%
CC BY-NC-SA	176.113	1%
CC BY-NC-ND	158.328	1%
CC BY-NC	156.201	1%
Restricted Access- Rights Reserved	127.169	1%
Unknown Copyright status	70.570	0%
CC BY	41.556	0%
CC BY-ND	3.163	0%
<b>No copyright status</b>	<b>12.534.840</b>	<b>53%</b>

## Content in Europeana by content provider TOP 15

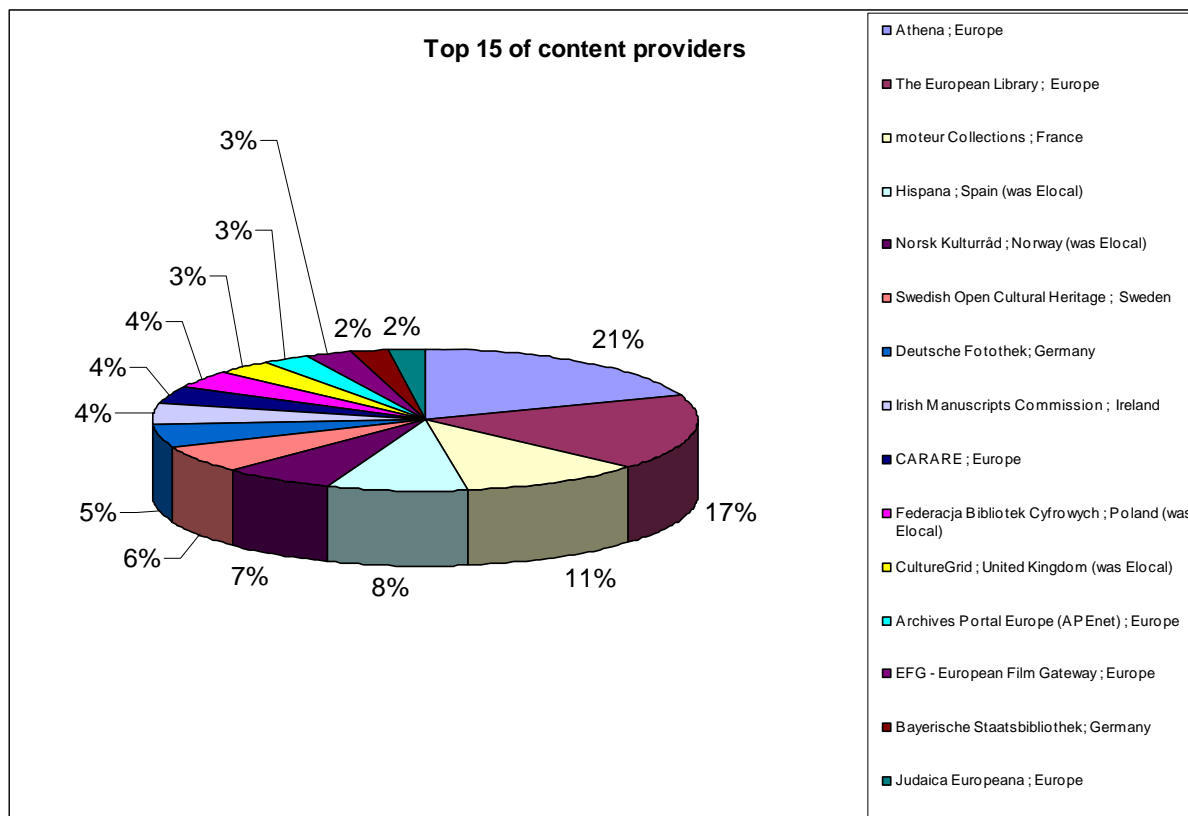


Figure 3 Europeana Content by content provider TOP 15

Provider	Total/provider	Percentage
<b>Total of records</b>	<b>23.431.250</b>	
<b>Athena ; Europe</b>	3.976.984	17%
<b>The European Library ; Europe</b>	3.478.818	15%
<b>moteur Collections ; France</b>	2.205.084	9%
<b>Hispana ; Spain (was Elocal)</b>	1.699.229	7%
<b>Norsk Kulturråd ; Norway (was Elocal)</b>	1.352.518	6%
<b>Swedish Open Cultural Heritage ; Sweden</b>	1.262.773	5%
<b>Deutsche Fotothek; Germany</b>	1.104.123	5%
<b>Irish Manuscripts Commission ; Ireland</b>	907.837	4%
<b>CARARE ; Europe</b>	812.972	3%
<b>Federacja Bibliotek Cyfrowych ; Poland (was Elocal)</b>	783.275	3%
<b>CultureGrid ; United Kingdom (was Elocal)</b>	696.036	3%
<b>Archives Portal Europe (APEnet) ; Europe</b>	555.370	2%
<b>EFG - European Film Gateway ; Europe</b>	546.352	2%
<b>Bayerische Staatsbibliothek; Germany</b>	465.795	2%
<b>Judaica Europeana ; Europe</b>	460.883	2%

## Content Type in Europeana by Country TOP15

IMAGE per Country TOP 15

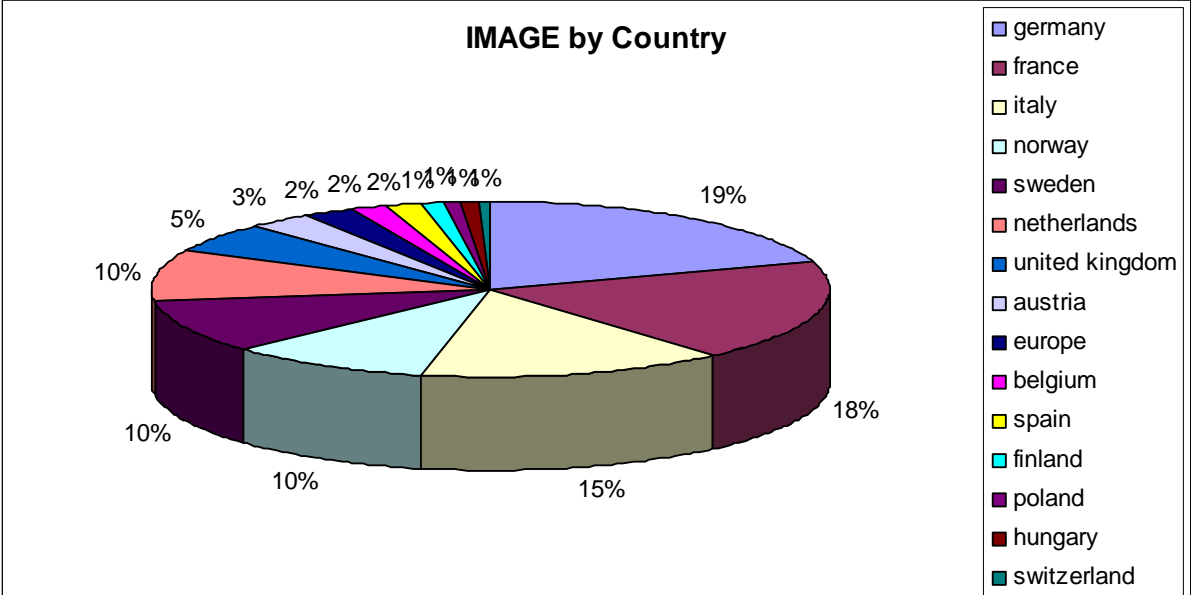


Figure 4 IMAGE by Country TOP 15

Country	Image	Percentage
<b>Total of Image type</b>	<b>14.339.091</b>	
germany	2763223	19%
france	2526844	18%
italy	2034646	14%
norway	1335247	9%
sweden	1329592	9%
netherlands	1315301	9%
united kingdom	751823	5%
austria	415164	3%
europe	337427	2%
belgium	235094	2%
spain	234099	2%
finland	145092	1%
poland	117075	1%
hungary	107603	1%
switzerland	95127	1%

**TEXT per Country TOP 15**

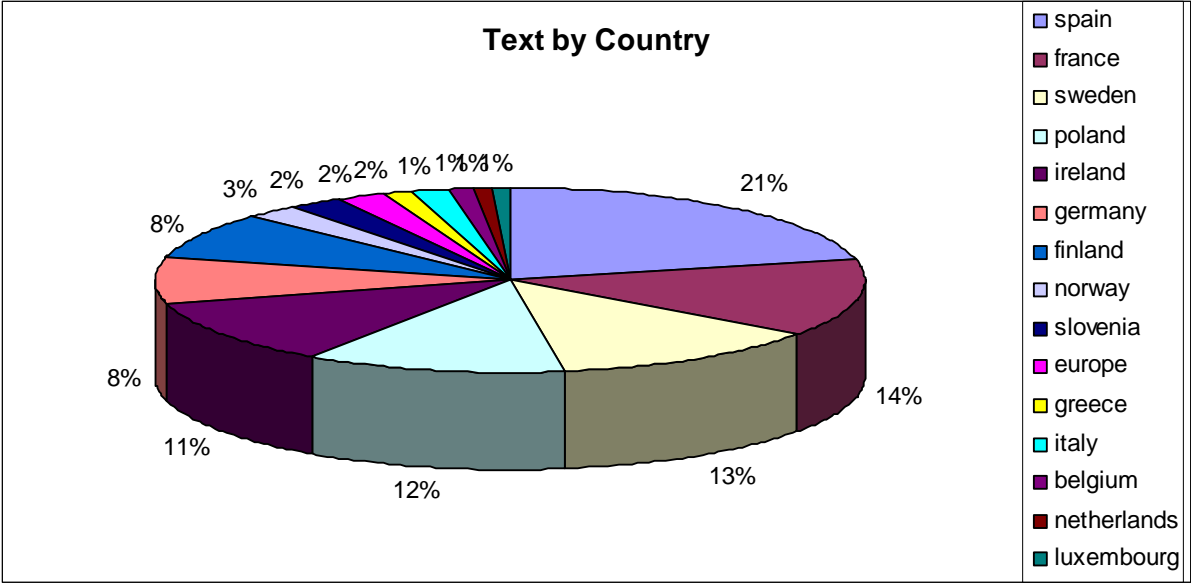


Figure 5 TEXT by Country TOP 15

Country	Text	Percentage
<b>Total of Text type</b>	<b>8.467.581</b>	
spain	1737363	21%
france	1102372	13%
sweden	1029834	12%
poland	975818	12%
ireland	912654	11%
germany	657498	8%
finland	653540	8%
norway	207442	2%
slovenia	196336	2%
europa	186731	2%
greece	128044	2%
italy	121558	1%
belgium	93545	1%
netherlands	70682	1%
luxembourg	65543	1%

**SOUND per Country TOP 15**



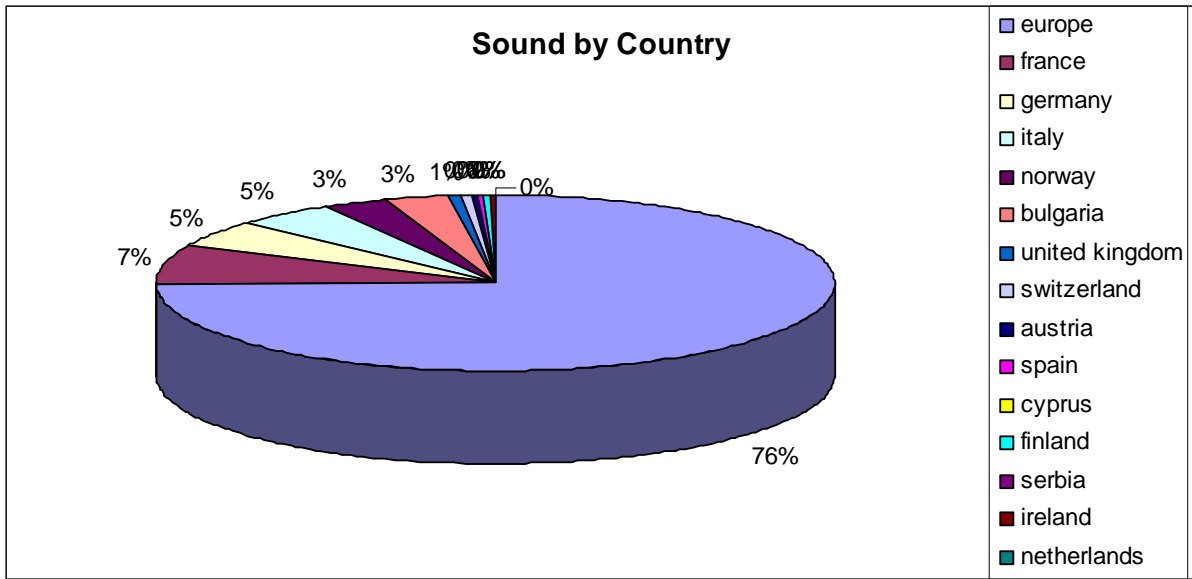


Figure 5 SOUND by Country TOP 15

Country	Sound	Percentage
<b>Total of Sound type</b>	<b>454.070</b>	
europa	338849	75%
france	32809	7%
germany	22147	5%
italy	21056	5%
norway	14576	3%
bulgaria	13738	3%
united kingdom	2944	1%
switzerland	1529	0%
austria	1433	0%
spain	1036	0%
cyprus	881	0%
finland	715	0%
serbia	599	0%
ireland	414	0%
netherlands	337	0%

**VIDEO per Country TOP 15**

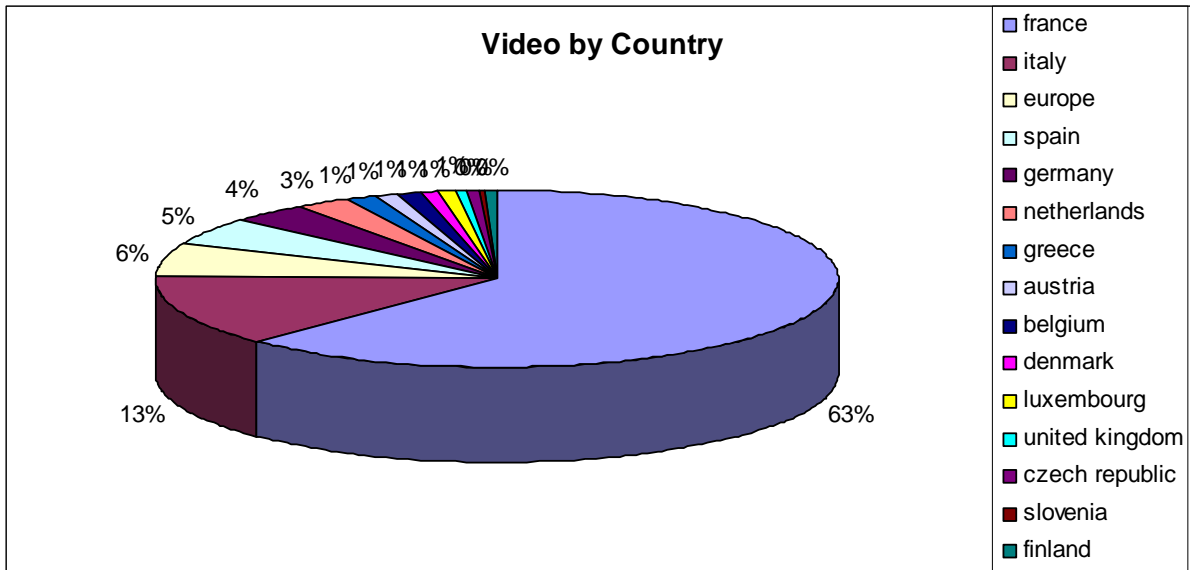


Figure 5 VIDEO by Country TOP 15

Country	Video	Percentage
<b>Total of Video type</b>	<b>170.506</b>	
france	104345	61%
italy	22132	13%
europa	9400	6%
spain	8431	5%
germany	6429	4%
netherlands	4361	3%
greece	2456	1%
austria	1880	1%
belgium	1818	1%
denmark	1395	1%
luxembourg	1323	1%
united kingdom	997	1%
czech republic	743	0%
slovenia	721	0%
finland	699	0%

**3D per Country**

Country	3D	Percentage
<b>Total of Video type</b>	<b>2</b>	
cyprus	2	100%