

**Progetto Far Conoscere WP3**

**Projet Faire Connaître WP3**

**PITEM PACE**

### **3.2 Definizione degli Strumenti di Sharing**

#### **Regione Piemonte**

##### **Soggetti attuatori**

Istituto di Informatica e Telematica - CNR

CSI Piemonte

## **Sommario**

### [1 Glossario](#)

### [2 Introduzione](#)

### [3 Architettura data provider/aggregatore](#)

### [4 Tracciato di interoperabilità](#)

### [5 Modalità di interscambio](#)

#### [5.1 Harvesting](#)

#### [5.2 Importazione](#)

### [6 Ricerca ed esplorazione](#)

#### [6.1 Ricerca modello Google-like](#)

#### [6.2 Ricerca avanzata](#)

#### [6.3 Navigazione](#)

#### [6.4 Percorsi Tematici](#)

#### [6.5 Presentazione dei risultati a fronte di una ricerca](#)

#### [6.6 Rilevanza dei risultati](#)

#### [6.7 Ordinamento dei risultati](#)

### [7 Multilinguismo](#)

#### [7.1 Ricerca multilingue](#)

#### [7.2 Presentazione dei risultati in più lingue](#)

#### [7.3 Interfaccia Web dell'aggregatore multilingua](#)

### [8 Web User Interface](#)

#### [8.1 Home page](#)

##### [8.1.1 Header](#)

##### [8.1.2 Footer](#)

##### [8.1.3 Contenuto principale](#)

##### [8.1.4 Navigazione](#)

#### [8.2 Minischeda record](#)

#### [8.3 Scheda completa di secondo livello](#)

#### [8.4 Presentazione dei risultati](#)

##### [8.4.1 Risultati della ricerca](#)

##### [8.4.2 Filtri](#)

#### [8.5 Ricerca avanzata](#)

### [9 Analisi Statistiche](#)

### [10 Mantenimento e aggiornamento delle risorse](#)

### [11 Preservazione e protezione dei contenuti digitali](#)

#### [11.1 Preservazione](#)

#### [11.2 Protezione](#)

### [12 Appendice A - Il modello di Europeana](#)

#### [12.1 Modalità di ricerca e presentazione dei risultati](#)

# 1 Glossario

## **Risorsa Originale**

Risorsa culturale così com'è descritta nel sistema nativo, come ad esempio un quadro.

## **Copia Digitale**

Risultato di un processo di digitalizzazione di una risorsa originale. Ad una risorsa originale possono corrispondere più copie digitali, come ad esempio diverse foto aventi come oggetto lo stesso quadro.

## **Metadato**

Attributo che specifica una proprietà della risorsa originale o della copia digitale (ad esempio il titolo di una risorsa originale o il formato della copia digitale).

## **Record**

L'insieme di tutti i metadati relativi a una risorsa originale e alle copie digitali associate.

## **Tracciato di interoperabilità**

Schema delle collezioni di record fornite all'aggregatore.

## **Formato del tracciato di interoperabilità**

Uno tra i seguenti: xml, json, csv, ...

## 2 Introduzione

Secondo la DTD del progetto, questo documento è formato da due parti fondamentali:

1. **Definizione dei protocolli e degli strumenti di sharing.** *Si procederà alla definizione delle modalità di implementazione dell'interoperabilità tra i diversi strumenti disponibili tra i partner ai fini della predisposizione dell'aggregatore che dovrà esporre i dati e le informazioni culturali provenienti dai sistemi dei partner. Condivisione di un protocollo basato sulle regole e i metodi di mantenimento, preservazione e protezione dei contenuti digitali.*
2. **Definizione delle regole e dei metodi di digitalizzazione.** *Tenendo conto dei risultati della fase di analisi, si procederà alla definizione di un documento contenente le regole da seguire obbligatoriamente dai partner per la realizzazione di oggetti digitali al fine di renderli nativamente interoperabili.*

La descrizione dettagliata del punto 2 è affrontata nel documento [3.1.2 Raccomandazioni tecniche](#). Mentre il documento 3.1.1 Analisi Tecnica descrive lo scenario attuale presente presso ciascun partner, questo documento, insieme con il documento 3.1.2 Raccomandazioni Tecniche, descrive la strategia comune da adottare affinché i partner possano condividere le proprie risorse presso un nodo aggregatore. Questa strategia è stata concordata con tutti i partner, a partire da ciascuna situazione specifica. Rispetto al documento 3.1.2 Raccomandazioni tecniche, in cui si descrive come i partner devono conformarsi alla strategia proposta, in questo documento, si descrive come l'aggregatore dovrà implementarla.

### 3 Architettura data provider/aggregatore

L'aggregatore dovrà interfacciarsi con i seguenti sistemi di gestione:

- Biblioteca Digitale Ligure (BDL) (Regione Liguria)
- Giornali del Piemonte (Regione Piemonte)
- Mèmora (Regione Piemonte)
- Catalogo dei Beni Culturali (CBC) (Regione Valle d'Aosta)
- Syracuse (Ville de Nice)

Si stima che, sulle cifre correntemente fornite dai partner, l'aggregatore dovrà gestire circa **2 milioni** di record. La tabella seguente riassume, per ogni partner, una stima di quanti record saranno resi disponibili presso l'aggregatore. Le celle vuote indicano che al momento le informazioni non sono disponibili, ma lo saranno nelle prossime versioni del documento. Si precisa che la tabella seguente è in divenire, nel senso che le risorse già specificate potranno essere soggette a revisione.

Partner	Nome collezione/raccolta/archivio/iniziativa	Formato copia digitale	Numero stimato di RECORD derivanti da digitalizzazioni passate	Numero stimato di RECORD derivanti da nuove digitalizzazioni
Regione Piemonte	Pagine dei giornali locali della provincia di Cuneo	PDF	932.000	
	Pagine dei giornali locali della provincia di Torino	PDF	697.000	95.000
	Schede della provincia di Torino (Memora)	JPG	96.048 (al massimo)	
	Schede della provincia di Cuneo (Memora)	JPG	34.102 (al massimo)	
	Fondi geografici Italia-Francia		13.000	
	Fondo Patrimonio Linguistico			
	Fotografie ex FIF e campagne fotografiche			
	Fondo Libri francesi della Biblioteca Storica di Masino			
Regione Valle d'Aosta	Censimento del patrimonio storico di architettura minore	JPG, TIFF	900	3.500
Ville de Nice	Livres anciens fonds régional		145	1.500
	Livres modernes fonds régional		329	10.000
	Images Fixes fonds régional		288	5.000
	Périodiques fonds régional		2.290	40.000
Regione Liguria	Periodici locali	TIFF, JPG	32.000	circa 15.000
	Cartoline storiche di Sanremo	JPG	1.544	
	Materiale cartografico e iconografico	TIFF, JPG	8	
<b>Totale provvisorio</b>			1.822.254	183.500

*Tabella 1. Risorse messe a disposizione dai partner*

La [Figura 1](#) mostra una possibile architettura logica dell'aggregatore.

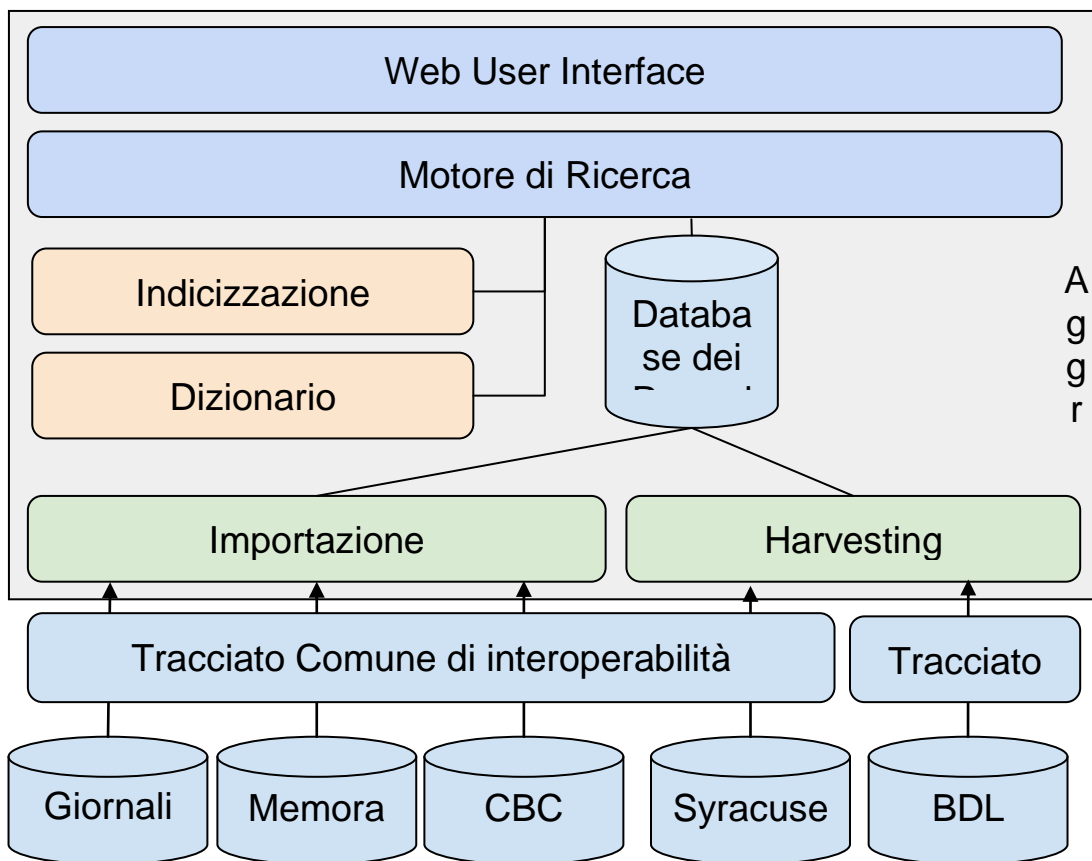


Figura 1. Architettura di massima dell'aggregatore

Nella parte bassa della [Figura 1](#), sono mostrati i sistemi di gestione con cui l'aggregatore dovrà interfacciarsi nel corso del progetto<sup>1</sup>. Ognuno di essi fornirà i propri record nel formato definito dal tracciato comune di interoperabilità, ad eccezione della Regione Liguria, che fornirà i record nel formato proprio, mettendo a disposizione anche il mapping tra il proprio formato e il tracciato comune (descritto nella relativa sezione sul tracciato comune di interoperabilità).

I record dei partner saranno messi a disposizione dell'aggregatore attraverso due modalità:

- importazione (download da un URL o upload, come meglio specificato in seguito)
- harvesting, attraverso il protocollo OAI-PMH.

I Giornali del Piemonte, Memora e VDA utilizzeranno la prima modalità, mentre BDL e VDN la seconda modalità. I dettagli sulle modalità di raccolta dei record sono descritti nella sezione Raccolta dei record.

I record raccolti dall'aggregatore (sia attraverso la modalità di importazione sia attraverso la modalità di harvesting) verranno raccolti nel database dei record, che potrà essere indicizzato, secondo i criteri descritti in seguito. I problemi legati al multilinguismo potrebbero essere gestiti attraverso un dizionario.

L'aggregatore metterà a disposizione un motore di ricerca, che sfrutterà i criteri di indicizzazione definiti per effettuare le ricerche e fornire i risultati finali all'utente, attraverso una Web User Interface (descritti nella sezione Requisiti per la ricerca e la presentazione dei risultati).

<sup>1</sup> Non si esclude che potrebbero essere aggiunti altri sistemi di gestione.

## 4 Tracciato di interoperabilità

L'aggregatore dovrà gestire varie tipologie di risorse fornite dai partner tra cui:

- Testo - Libri, Giornali, Periodici, Documenti Testuali
- Immagini - Opere d'arte, Fotografie, Stampe, Disegni, Mappe
- Audio/Video

I metadati forniti dai partner all'aggregatore saranno conformi ad una struttura che chiamiamo tracciato di interoperabilità, o semplicemente tracciato. Questo schema consente di raggruppare un set di record da inviare all'aggregatore.

Lo schema del tracciato è stato pensato tenendo a mente i seguenti requisiti:

- il numero di elementi deve essere minimo;
- ogni elemento se inserito deve essere motivato (vedi Allegato 1 colonna "motivo");
- i motivi sono legati alla ricerca, che deve rispondere alle domande dove, quando e cosa, e a possibili filtraggi;

Il tracciato di interoperabilità è formato da tre parti fondamentali:

- *header* - contiene la descrizione del pacchetto di record. E' presente una volta sola all'interno del tracciato. Costituito da 5 elementi obbligatori.
- *record/originalResource* - descrizione sintetica della risorsa originale. E' ripetuta per ogni risorsa originale contenuta nel dataset. Costituito da 11 elementi di cui 6 obbligatori
- *record/digitalCopy* - descrizione di una copia digitale della risorsa originale. Una risorsa originale può contenere più copie digitali. Costituito da 5 elementi di cui 4 obbligatori.

Le figure seguenti mostrano la struttura del tracciato. La descrizione completa degli elementi foglia del tracciato è riportata nell'Allegato 1 - Tracciato di Interoperabilità, mentre tutto lo schema del tracciato è riportato nell'Allegato 2 - XML Schema del tracciato di interoperabilità.

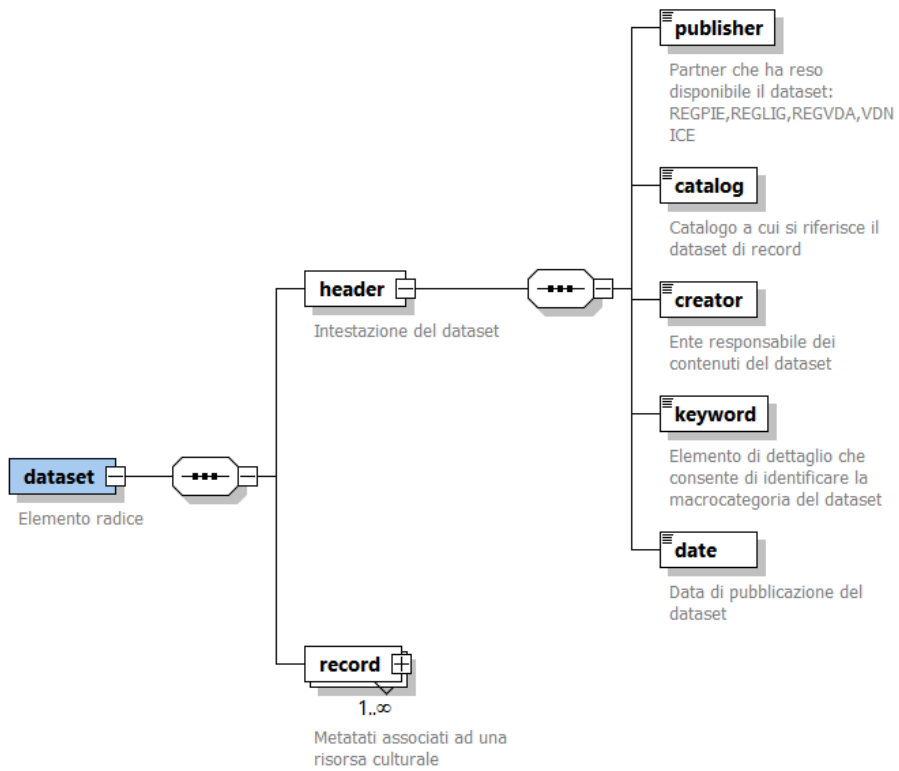


Figura 2. Header del tracciato di interoperabilità



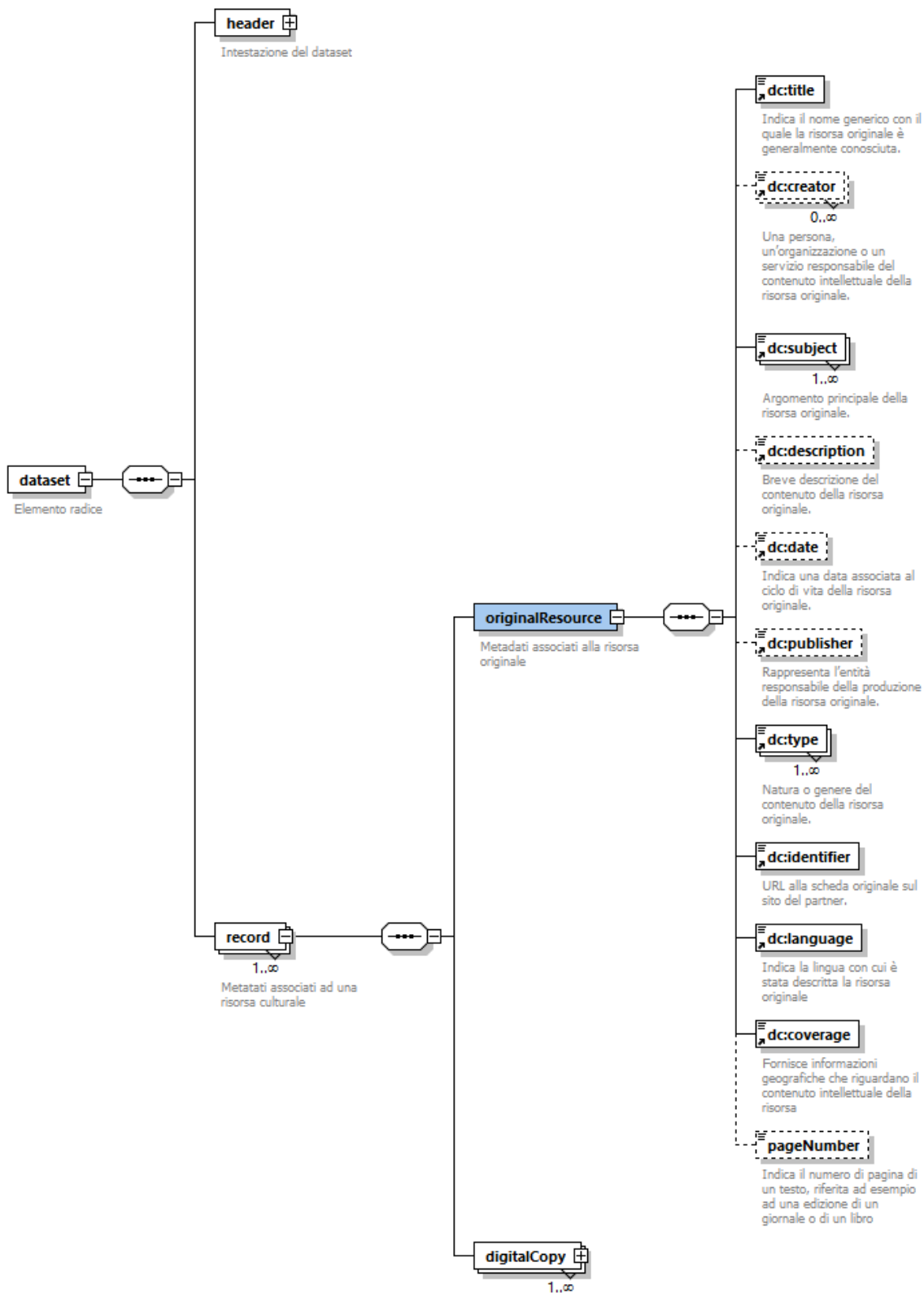


Figura 3. Original Resource

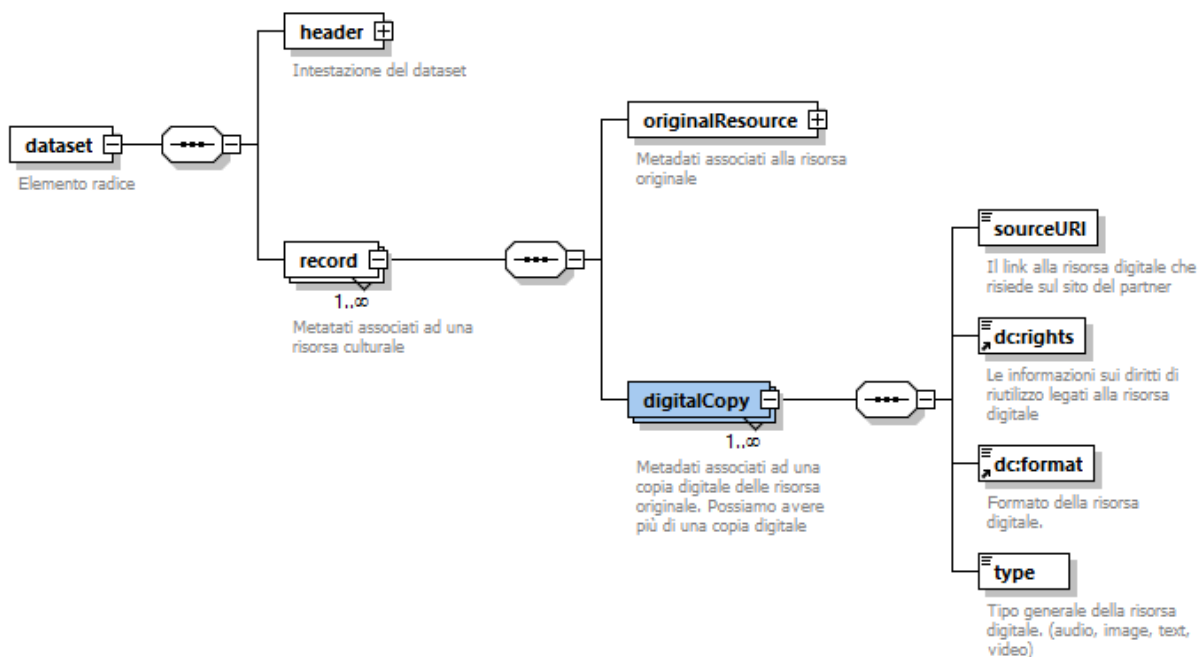


Figura 4. Digital Copy

Questo tracciato di interoperabilità rappresenta lo schema del file statico. I partner che forniranno i dati attraverso OAI-PMH dovranno garantire una corrispondenza tra l'OAI:DC e il tracciato di interoperabilità.

Al momento della stesura del presente documento, tutti i partner, ad eccezione della Regione Liguria forniranno i propri record secondo il tracciato di interoperabilità.

Per quanto riguarda la Regione Liguria, i metadati della Biblioteca Digitale Ligure attualmente sono tutti in formato MAG<sup>2</sup>, che per la descrizione del documento originale utilizza gli elementi del Dublin Core.

Di seguito un esempio <http://bibliotecadigitale.regione.liguria.it/ark:/64639/SGEE001421/134>. Il seguente frammento di codice mostra l'intero elemento <bib> che contiene anche le informazioni sull'esemplare digitalizzato (localizzazione, inventario e collocazione). L'attributo level dell'elemento <bib> indica il livello bibliografico. I codici utilizzati sono quelli dell'unimarc (a=spoglio, m=monografia, s=periodico, c=collezione).

```
<bib level="m">
  <dc:identifier>SGEE001421</dc:identifier>
  <dc:title>Atlas Minor</dc:title>
  <dc:creator>Mercator, Gerardus </dc:creator>
  <dc:publisher>Amsterdam : Adrianus Belthius</dc:publisher>
  <dc:subject>Mondo</dc:subject>
  <dc:date>1610</dc:date>
  <dc:type>cartografia a stampa</dc:type>
  <dc:format>[8], 684 p. ; (180x230 mm)</dc:format>
```

<sup>2</sup> Lo schema e il manuale del formato MAG sono disponibili al seguente indirizzo: <https://www.iccu.sbn.it/it/normative-standard/linee-guida-per-la-digitalizzazione-e-metadati/standard-mag-versione-2.0.1/index.html>

```

<dc:language>IT</dc:language>
<dc:coverage>Mondo</dc:coverage>
<holdings>
  <library>Biblioteca civica Berio, Genova, Italia, IT-GE0036</library>
  <inventory_number>222609</inventory_number>
  <shelfmark>B.S.XVII.A.442</shelfmark>
</holdings>
<piece></piece>
</bib>

```

## Di seguito ecco come dovrebbe essere tradotto il record nel tracciato di interoperabilità

```

<?xml version="1.0" encoding="UTF-8"?>
<dataset xmlns:dc="http://purl.org/dc/elements/1.1/"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:noNamespaceSchemaLocation="./fc.xsd">
  <header>
    <publisher>REGLIG</publisher>
    <catalog>Documenti antichi, manoscritti o a stampa, di carattere geografico
e cartografico</catalog>
    <creator>Biblioteca civica Berio - Genova, IT:GE0036</creator>
    <keyword>Cartografia a stampa</keyword>
    <date>2020-11-11</date>
  </header>
  <record>
    <originalResource>
      <dc:title>Atlas Minor</dc:title>
      <dc:creator>Mercator, Gerardus</dc:creator>
      <dc:subject>Mondo</dc:subject>
      <dc:subject>Cartografia</dc:subject>
      <dc:description>??</dc:description>
      <dc:date>1610</dc:date>
      <dc:publisher>Amsterdam : Adrianus Belthius</dc:publisher>
      <dc:type>image</dc:type><!-- cartografia a stampa -->
      <dc:identifier>http://bibliotecadigitale.regione.liguria.it/
id/SGEE001421</dc:identifier>
      <dc:language>IT</dc:language>
      <dc:coverage>https://www.geonames.org/6295630</dc:coverage><!--Mondo-
->

    </originalResource>
    <digitalCopy>
      <sourceURI>http://bibliotecadigitale.regione.liguria.it/
fedora/objects/mag:134/datastreams/
MMc6601cedd8834f96c1cfc2e04f65cf9a/content</sourceURI>
      <dc:rights>yes</dc:rights>
      <dc:format>image/jpeg</dc:format>
      <type>image</type>
    </digitalCopy>
    <digitalCopy>
      <sourceURI>http://bibliotecadigitale.regione.liguria.it/
fedora/objects/mag:134/datastreams/
MM4c7027861ca89572f3a852c017d8ef7e/content</sourceURI>
      <dc:rights>yes</dc:rights>
      <dc:format>image/jpeg</dc:format>
      <type>image</type>
    </digitalCopy>
    ...
  </record>
</dataset>

```

Di seguito forniamo una tabella di conversione tra i due formati

Tracciato interoperabilità	Tracciato record della Liguria
dataset/header/publisher	
dataset/header/catalog	
dataset/header/creator	metadigit/gen/agency
dataset/header/keyword	
dataset/header/date	
dataset/record/originalResource/dc:title	metadigit/bib/dc:title
dataset/record/originalResource/dc:creator	metadigit/bib/dc:creator
dataset/record/originalResource/dc:subject	metadigit/bib/dc:subject
dataset/record/originalResource/dc:description	
dataset/record/originalResource/dc:publisher	metadigit/bib/dc:publisher
dataset/record/originalResource/dc:date	metadigit/bib/dc:date
dataset/record/originalResource/dc:type	
dataset/record/originalResource/dc:identifier	
dataset/record/originalResource/dc:language	metadigit/bib/dc:language
dataset/record/originalResource/dc:coverage	
dataset/record/originalResource/sequenceNumber	
dataset/record/digitalCopy/sourceURI	metadigit/img/altImg[1]/file@xlink:href
dataset/record/digitalCopy/dc:rights	
dataset/record/digitalCopy/dc:format	metadigit/img/altImg[1]/format/niso:mime
dataset/record/digitalCopy/type	
dataset/record/digitalCopy/sequenceNumber	metadigit/img/sequence_number

*Tabella 2. Tabella di conversione*

## 5 Modalità di interscambio

Questa sezione descrive i metodi con i quali l'aggregatore raccoglierà i record di metadati a partire dai sistemi di gestione disponibili presso i partner.

L'analisi tecnica ha evidenziato che:

- la Biblioteca Digitale Ligure e il Portale della Ville de Nice forniscono i dati tramite server OAI-PMH,
- Mèmora e Giornali del Piemonte della Regione Piemonte e il Catalogo dei Beni Culturali della Regione Valle d'Aosta mettono a disposizione dati statici, in XML (Regione Piemonte) e XML o JSON (Regione Valle d'Aosta).

## 5.1 Harvesting

L'harvester dell'aggregatore dovrà funzionare da client OAI-PMH:

- dovrà interrogare i provider OAI-PMH
- dovrà gestire file XML strutturati secondo lo schema del tracciato comune oppure secondo altri schemi di cui viene fornito un mapping con il tracciato comune.

### **Biblioteca Digitale Ligure**

URL del provider OAI-PMH:

<http://bibliotecadigitale.regione.liguria.it/oaiprovider/>

Esempio di utilizzo:

<http://bibliotecadigitale.regione.liguria.it/oaiprovider/?verb=GetRecord&identifier=XXXX> (sostituire XXXX con l'identificativo del record da scaricare). La barra prima del punto interrogativo è necessaria.

In generale si segue la normale sintassi dei verbi previsti dal protocollo OAI-PMH.

L'unico formato di metadati attualmente supportato è il MAG.

### **Portale della Ville de Nice**

E' in corso la predisposizione del provider OAI-PMH.

## 5.2 Importazione

Per i partner che non hanno a disposizione un server OAI-PMH, l'aggregatore dovrà rendere possibile il caricamento di file statici, contenenti i metadati esportati. In questo caso, l'aggregatore dovrà fornire due modalità di caricamento dei file statici:

3. download - URL (ad esempio su HTTP) contenente il file statico (esempio <http://www.foo.com/records/records.xml>)
4. upload - il partner carica il file statico sull'aggregatore (ad esempio tramite FTP).

Nel caso della modalità download, l'aggregatore dovrà essere in grado di recuperare i metadati, tramite protocolli FTP o HTTP da URL forniti dai partner. Nel caso della modalità upload, l'aggregatore dovrà ospitare un server su cui i partner potranno inviare i file contenenti i record da integrare. In entrambe le modalità, i file forniti dai partner saranno disponibili:

- nei formati XML o JSON
- con lo schema del tracciato comune o altro schema di cui viene fornito un mapping con il tracciato comune.

Al momento della stesura di questo documento i dettagli tecnici relativi alle URL fornite dai singoli partner non sono disponibili. Non appena tali informazioni saranno disponibili, il presente documento sarà aggiornato con le URL relative alle risorse messe a disposizione da Giornali del Piemonte, Memora e Catalogo dei Beni Culturali.



## 6 Ricerca ed esplorazione

Per quanto riguarda le modalità di ricerca, sono stati identificati quattro principali meccanismi, già descritti nell'analisi dei modelli di diffusione (Report 3.1.5 Inventario dei Siti Web), che l'aggregatore dovrebbe essere in grado di implementare:

- *ricerca modello Google-like*, per gli utenti non esperti del dominio, attraverso una casella di testo in cui l'utente può inserire una o più parole chiave che il sistema utilizza per fare la ricerca in alcuni campi testuali relativi alle risorse contenute nel database;
- *ricerca avanzata*, utilizzata da utenti più esperti che hanno conoscenza del dominio, in cui è possibile impostare dei filtri attraverso una maschera di ricerca che mostra alcuni dei campi di metadati disponibili. Sarebbe utile dare all'utente anche la possibilità di utilizzare degli operatori logici;
- *navigazione*, attraverso sezioni del sito che raggruppano le risorse in base ad alcuni criteri (ad esempio temi del progetto, argomenti, tipologie, collezioni, periodi di tempo, persone, fonti).
- *percorsi tematici* creati ad-hoc, che permettono la valorizzazione del patrimonio disponibile perché mostrano quali sono le risorse più interessanti per l'esplorazione.

### 6.1 Ricerca modello Google-like

Questo tipo di ricerca è quella più familiare per l'utente generico e si ispira alla barra di ricerca di Google. Si raccomanda che l'interfaccia utente sia la più semplice possibile. L'inserimento dei termini di ricerca deve essere guidato con il sistema di autocompletamento che parte dopo l'inserimento di almeno 3 caratteri.

La ricerca sarà effettuata sui campi del tracciato comune che rispondono alle seguenti domande:

- Chi
  - autore (originalResource/dc:creator)
  - editore (originalResource/dc:publisher)
- Cosa
  - titolo (originalResource/dc:title)
  - descrizione (originalResource/dc:description)
  - soggetto (originalResource/dc:subject)

Opzionalmente si potrebbe aggiungere una barra di ricerca per data (originalResource/dc:date).

### 6.2 Ricerca avanzata

Nei sistemi di gestione di risorse digitali, questo tipo di ricerca è ormai in disuso, in quanto sostituita quasi del tutto dai filtri. Tuttavia, alcuni sistemi di gestione prevedono ancora la modalità di ricerca avanzata.

La ricerca avanzata sarà effettuata sui campi del tracciato comune che rispondono alle domande presenti nella ricerca Google-like (chi e cosa) con l'aggiunta dei seguenti campi:

- Quando
  - data (originalResource/dc:date)
- Dove
  - località (originalResource/dc:coverage)
  - ente presso cui è conservata la risorsa (header/creator)

La modalità di ricerca avanzata prevederà in modo implicito che i vari campi siano messi in modalità AND. Opzionalmente si potrà prevedere l'uso dell'operatore logico OR applicabile ad ogni singolo campo.

## 6.3 Navigazione

La navigazione verrà effettuata sui campi obbligatori del tracciato record comune in modo da poter navigare tra tutte le risorse del progetto. Si possono prevedere i seguenti tipi di navigazione, implementati come dei filtri:

- Cosa
  - temi del progetto (originalResource/dc:subject)
  - tipologia di risorse (originalResource/dc:type)
- Dove
  - località (originalResource/dc:coverage)
  - ente presso cui è conservata la risorsa (header/creator)
- Quando
  - data (originalResource/dc:date)

Similmente alle tipologie di ricerca, anche la navigazione dovrebbe permettere di effettuare dei filtri sui risultati recuperati.

## 6.4 Percorsi Tematici

Si può prevedere la definizione di alcuni percorsi tematici che valorizzino i collegamenti tra alcune risorse presenti nelle diverse collezioni, come ad esempio *i monumenti lungo la costa italo-francese*. Queste selezioni tematiche richiedono però un lavoro redazionale a parte. L'aggregatore dovrebbe prevedere la possibilità di ospitare questi nuovi percorsi tematici, in una sezione del sito finale.

## 6.5 Presentazione dei risultati a fronte di una ricerca

A seguito di una ricerca, l'utente accede ad un **primo livello di dettaglio** dei risultati, in cui viene presentata una lista di minischede descrittive, corrispondenti ai record estratti. Ciascuna minischeda riporterà un'anteprima (a seconda del tipo di risorsa), e i metadati principali (luogo, data, autore, soggetto, descrizione breve).

Cliccando su ciascuna scheda di primo livello l'utente potrà accedere alla **scheda completa di secondo livello**, che presenta tutti i metadati forniti all'aggregatore, tra i quali ci sarà anche il link alla scheda della risorsa presente sul sito del partner fornitore.

Cliccando sul link della risorsa, l'utente potrà consultare la **scheda originale** sul sito del partner fornitore delle risorse.

Sui risultati del primo livello sarà possibile effettuare dei filtraggi tramite l'uso di facet. I possibili filtri dovrebbero includere i seguenti campi, presenti nel tracciato comune:

- Chi
  - autore (originalResource/dc:creator)
  - editore (originalResource/dc:publisher)
- Cosa
  - soggetto (originalResource/dc:subject)
  - tema del progetto (originalResource/dc:subject)



- tipologia di risorsa (originalResource/dc:type)
- lingua (originalResource/dc:language)
- licenza (digitalCopy/dc:rights)
- formato (digitalCopy/dc:format)
- Dove
  - località (originalResource/dc:coverage)
  - ente presso cui è conservata la risorsa (header/creator)
- Quando
  - data (originalResource/dc:date)

Due filtri particolarmente importanti sono quelli relativi alla copertura spaziale e alla data. Rispetto alla copertura spaziale, il filtro dovrebbe permettere di restringere la località di interesse o selezionare alcune località, gerarchicamente collegate alla località principale, risultato della ricerca principale (ad esempio, se la ricerca principale ha dato come risultato tutte le risorse collegate al luogo Italia, questo filtro permette di restringere la ricerca al luogo Pisa).

Per quanto riguarda il filtro per data, si dovrebbe fornire una modalità per restringere l'intervallo di tempo o selezionare più intervalli di tempo contenuti nel periodo risultato della ricerca principale (ad esempio se la ricerca principale ha dato come risultato tutte le risorse del 1500, questo filtro permette di restringere la ricerca alle sole risorse del 1520).

## 6.6 Rilevanza dei risultati

A fronte di una ricerca, la rilevanza permette di presentare prima i risultati più attinenti ad una ricerca e poi man mano tutti gli altri risultati. Questo meccanismo è possibile grazie all'assegnazione di pesi ai singoli campi su cui è effettuata la ricerca. La rilevanza di un risultato dipende dal tipo di ricerca effettuata. Pertanto i valori dei pesi dei campi dovranno essere determinati di volta in volta dal motore di ricerca dell'aggregatore e non possono essere definiti a priori.

La rilevanza di un risultato dipende dai seguenti criteri:

1. presenza delle parole chiave in uno dei campi utilizzati per la ricerca (che dipendono dal tipo di ricerca).
2. numero di volte che la parola chiave compare nei campi principali

Sulla base di questi criteri, il motore di ricerca dell'aggregatore potrà elaborare un proprio algoritmo per la definizione della rilevanza dei risultati a fronte di una ricerca.

## 6.7 Ordinamento dei risultati

A fronte di una ricerca, l'ordinamento permette all'utente di scegliere l'ordine con cui visualizzare i risultati di una ricerca. Il motore di ricerca dell'aggregatore permetterà di effettuare l'ordinamento sulla base almeno dei seguenti campi:

- autore(originalResource/dc:creator)
- località (originalResource/dc:coverage)
- data (originalResource/dc:date).

## 7 Multilinguismo

I sistemi di gestione disponibili presso i partner mettono a disposizione risorse principalmente in una delle seguenti due lingue: francese o italiano. Inoltre per facilitare l'utilizzo del motore di ricerca alla comunità internazionale si raccomanda anche il supporto della lingua inglese. Pertanto, al fine di migliorare la ricerca e uniformare la presentazione finale dei risultati, è necessario gestire il problema del multilinguismo.

Ci sono tre aspetti riguardanti il multilinguismo:

1. *ricerca multilingue*: indipendentemente dalla lingua, la ricerca recupera sempre le stesse risorse. Ad esempio, se si cerca "ponte" o "pont" o "bridge" come soggetto, la ricerca produce sempre lo stesso risultato;
2. *presentazione dei risultati in più lingue*: i risultati della ricerca saranno presentati nella lingua selezionata;
3. *interfaccia Web dell'aggregatore multilingua*: l'utente può scegliere la lingua di visualizzazione del sito Web, che cambierà gli elementi testuali dell'interfaccia come ad esempio i menu, le etichette, i titoli delle sezioni ecc.

### 7.1 Ricerca multilingue

La ricerca multilingue è riferita a campi di ricerca che possono essere di due tipi:

1. appartenenti ad un vocabolario controllato
2. testo libero

La seguente tabella mostra i campi del tracciato record appartenenti al tipo 1 e i vocabolari associati in italiano, francese e inglese:

	vocabolario		
campo	italiano	francese	inglese
publisher	REGPIE, REGVDA, REGLIG, VDNICE		
dc:subject	preistoria costruito storico resistenze moderne sacro e profano	préhistoire bâti historique résistances modernes sacré et profane	prehistory historical construction modern resistances sacred and profane
dc:type	collezione immagine oggetto fisico testo suono	collection image objet physique texte du son	collection image physical object text sound
dc:rights	si si con limiti richiedere autorizzazioni	oui oui avec des limites demander des autorisations	yes yes with restrictions require authorization
type	audio testo immagine	audio texte image	audio text image

	video	vidéo	video
--	-------	-------	-------

Tabella 3. Campi con valori estratti da dizionari.

Il secondo tipo di ricerca (testo libero) richiede che il testo venga tradotto in tutte le lingue supportate o dal data provider o dall'aggregatore. I testi tradotti nelle varie lingue verranno indicizzati per la ricerca.

Nell'ambito del progetto, la traduzione dei campi a testo libero non è prevista dai data provider, per cui dovrebbe essere demandata totalmente all'aggregatore. I partner dovranno comunque specificare la lingua di ciascun record, come definito nel tracciato di interoperabilità.

Una soluzione alternativa potrebbe essere di non effettuare la traduzione dei campi a testo libero, ma di tradurre il testo di ricerca ed eseguire più ricerche, una per ogni lingua in cui sono salvati i record (italiano, francese). Questa soluzione prevede che per ogni record sia specificata la lingua originaria (italiano o francese). La figura seguente mostra un esempio di questo concetto. Naturalmente deve essere specificata anche la lingua utilizzata per la ricerca.

L'utente cerca ad esempio la parola *bridge*, specificando come lingua l'inglese. L'aggregatore dispone di un traduttore, che traduce la parola *bridge* in francese (*pont*) e in italiano (*ponte*). La ricerca delle due parole è pertanto effettuata separatamente: nei record italiani l'aggregatore esegue la ricerca della parola *ponte*, mentre nei record in francese l'aggregatore esegue la ricerca della parola *pont*. I risultati sono poi restituiti in modo aggregato all'utente finale, specificando comunque la lingua originale di ogni record.

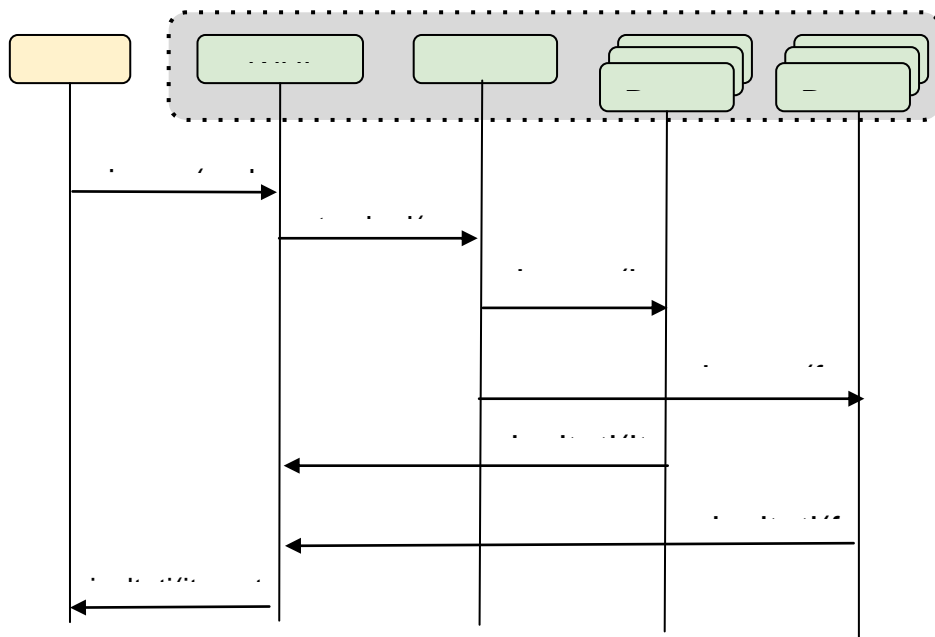


Figura 5. Possibile soluzione per ricerca multilingue.

## 7.2 Presentazione dei risultati in più lingue

Indipendentemente da come sia effettuata la ricerca, i risultati finali dovrebbero essere restituiti all'utente finale in modo omogeneo, cioè nella stessa lingua. Quindi anche se i record originari sono memorizzati in una lingua, se l'utente finale ha scelto un'altra lingua di visualizzazione, l'aggregatore dovrebbe essere in grado di tradurre il testo contenuto nei metadati dei record restituiti nel risultato.

Sebbene nel caso specifico del progetto Pitem Pa.CE. i risultati di una ricerca non possano essere tradotti e modificati, in quanto frutto di un'opera di catalogazione, l'aggregatore potrebbe mettere a disposizione dell'utente finale una funzionalità di traduzione automatica per la comprensione del testo.

Per i campi i cui valori appartengono ad un vocabolario controllato si prevede una traduzione fornita a tantum, mentre per i campi con testo libero (descrizione, luoghi ecc) una soluzione a questo problema potrebbe essere l'uso di librerie o API che permettono la traduzione automatica dei risultati in fase di presentazione finale, senza andare ad inficiare la lingua di memorizzazione dei record. Una soluzione di questo tipo è possibile in quanto per ogni record è specificata in modo obbligatorio la lingua utilizzata. Bisogna però prevedere di marcare alcuni campi come non traducibili (ad esempio titolo e autore).

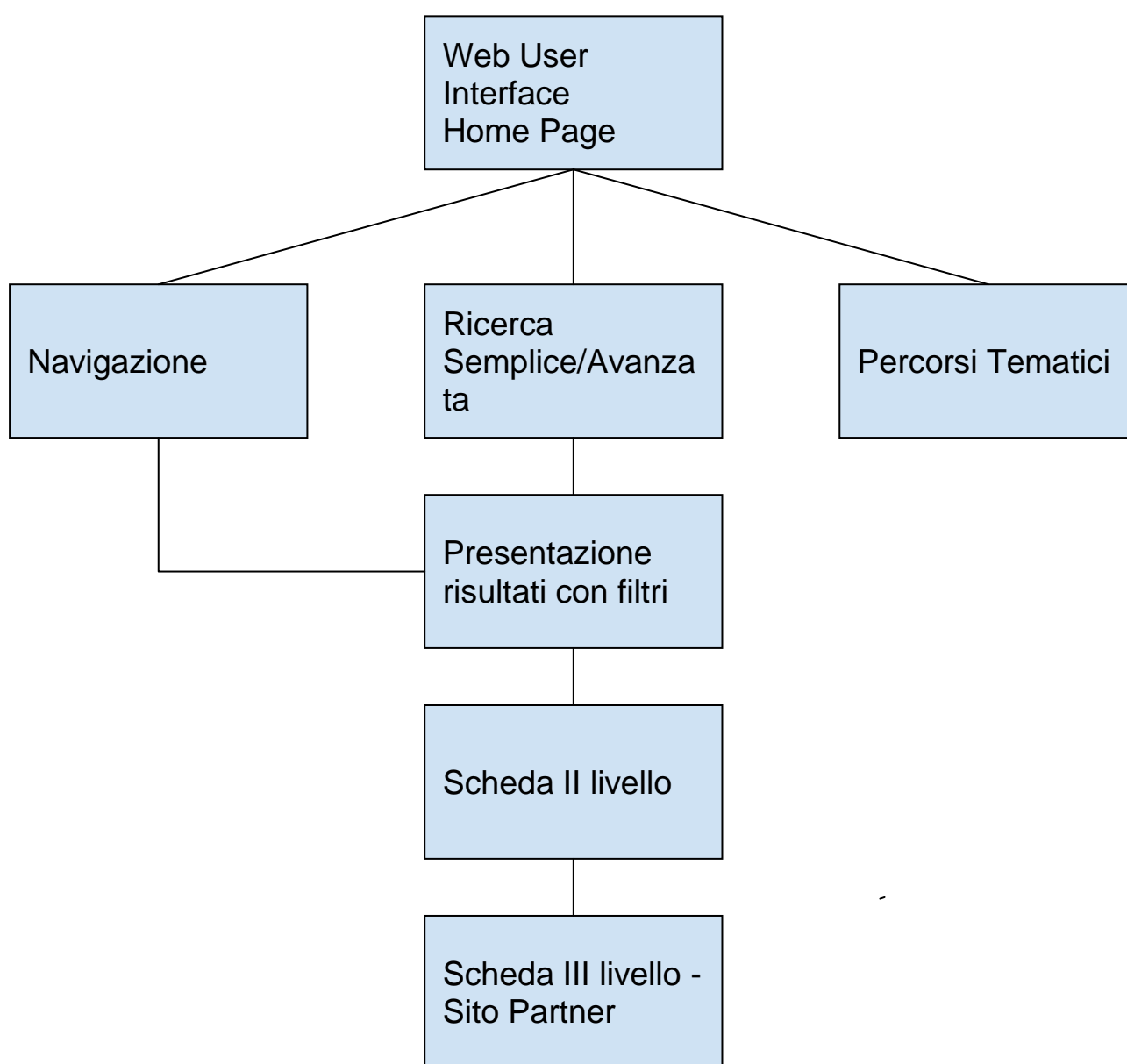
### 7.3 Interfaccia Web dell'aggregatore multilingua

Al fine di poter raggiungere un pubblico più ampio, uno dei requisiti fondamentali per l'aggregatore è che offra all'utente la possibilità di navigare l'interfaccia Web (menù, etichette, titoli) in più lingue. L'aggregatore dovrebbe presentare l'interfaccia Web almeno nelle seguenti lingue: italiano, francese e inglese UK. Il selettore per la scelta della lingua deve trovarsi in un punto immediatamente visibile per l'utente.

## 8 Web User Interface

La Web User Interface dovrà rispondere ai criteri di accessibilità e usabilità, seguendo le normative vigenti. Inoltre, dovrà essere responsive, al fine di permetterne un utilizzo anche su dispositivi mobili come smartphone e tablet.

La figura seguente ([Figura 6](#)) mostra il flusso di navigazione della Web User Interface. A partire dalla Home Page, l'utente potrà scegliere tre modalità di navigazione delle risorse digitali: navigazione, ricerca semplice (Google-like), ricerca avanzata e percorsi tematici. Nel caso di navigazione e ricerca, verrà presentata una pagina con le schede di primo livello associate alle risorse digitali (minischede record), rispondenti ai criteri selezionati, su cui sarà possibile effettuare dei filtri. Cliccando su una singola minischeda, sarà possibile accedere ai dettagli su quella risorsa (scheda completa di secondo livello). Tra i dettagli sulla risorsa sarà disponibile anche il link alla scheda della risorsa originale sul sito del partner.



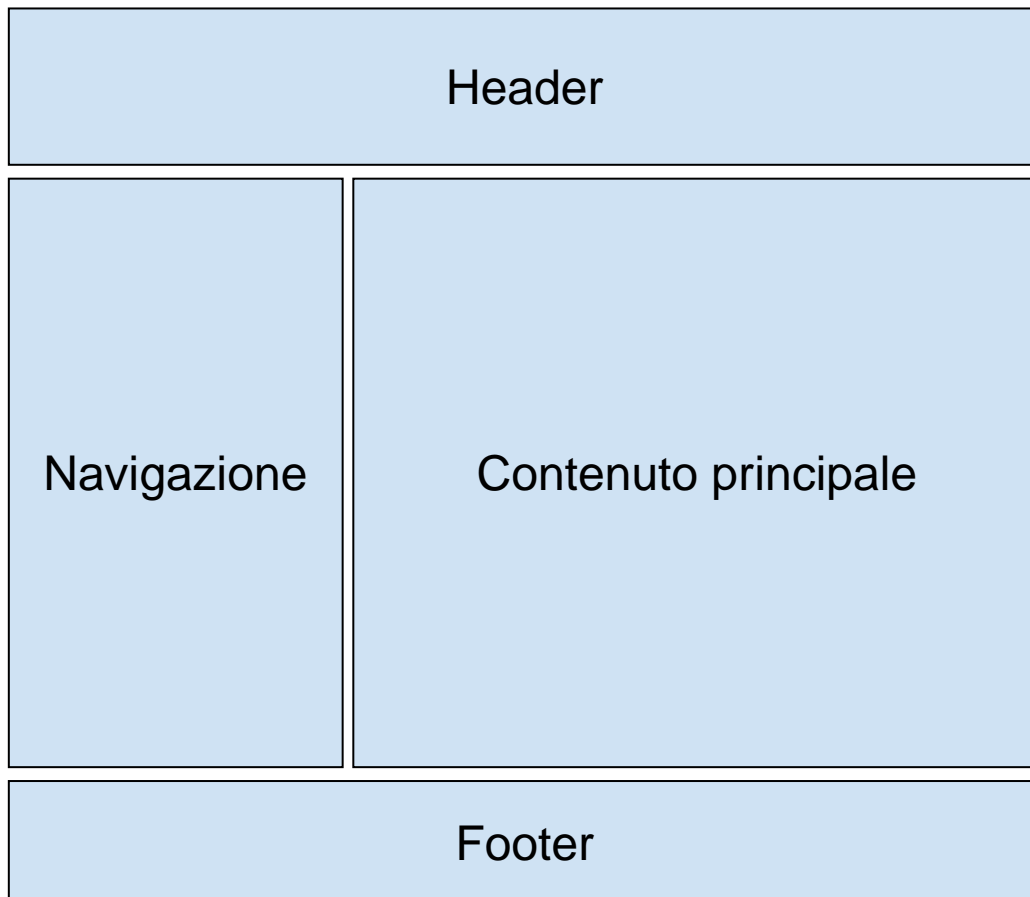
*Figura 6. Diagramma di flusso della user experience.*

Nelle sezioni seguenti vengono presentati dei mock-up delle varie parti dell'interfaccia, con una descrizione (non necessariamente esaustiva) del loro contenuto e delle funzionalità. Le dimensioni

sono solo indicative delle proporzioni relative tra gli elementi. Il colore giallo indica gli elementi interattivi.

## 8.1 Home page

La home page dell'applicazione dovrà essere organizzata in modo semplice, con le classiche sezioni presenti di solito in una pagina web. La [figura 7](#) mostra la suddivisione nelle 4 sezioni previste: header, navigazione, contenuto principale, e footer.



*Figura 7. Home Page.*

### 8.1.1 Header

L'header conterrà il logo del progetto, il titolo, una brevissima descrizione del progetto e i bottoni per la selezione della lingua dell'interfaccia (italiano, francese e inglese).

Conterrà inoltre la casella di ricerca per la ricerca semplice (Google-like), e il bottone per andare alla ricerca avanzata.

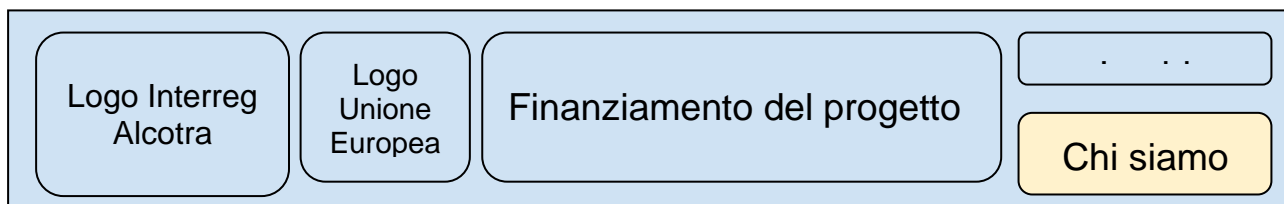


*Figura 8. Header.*

### 8.1.2 Footer

Il footer conterrà tutti i loghi necessari (Interreg Alcotra, Unione Europea, logo di ciascun partner) e l'indicazione dell'origine dei fondi che finanziano il progetto.

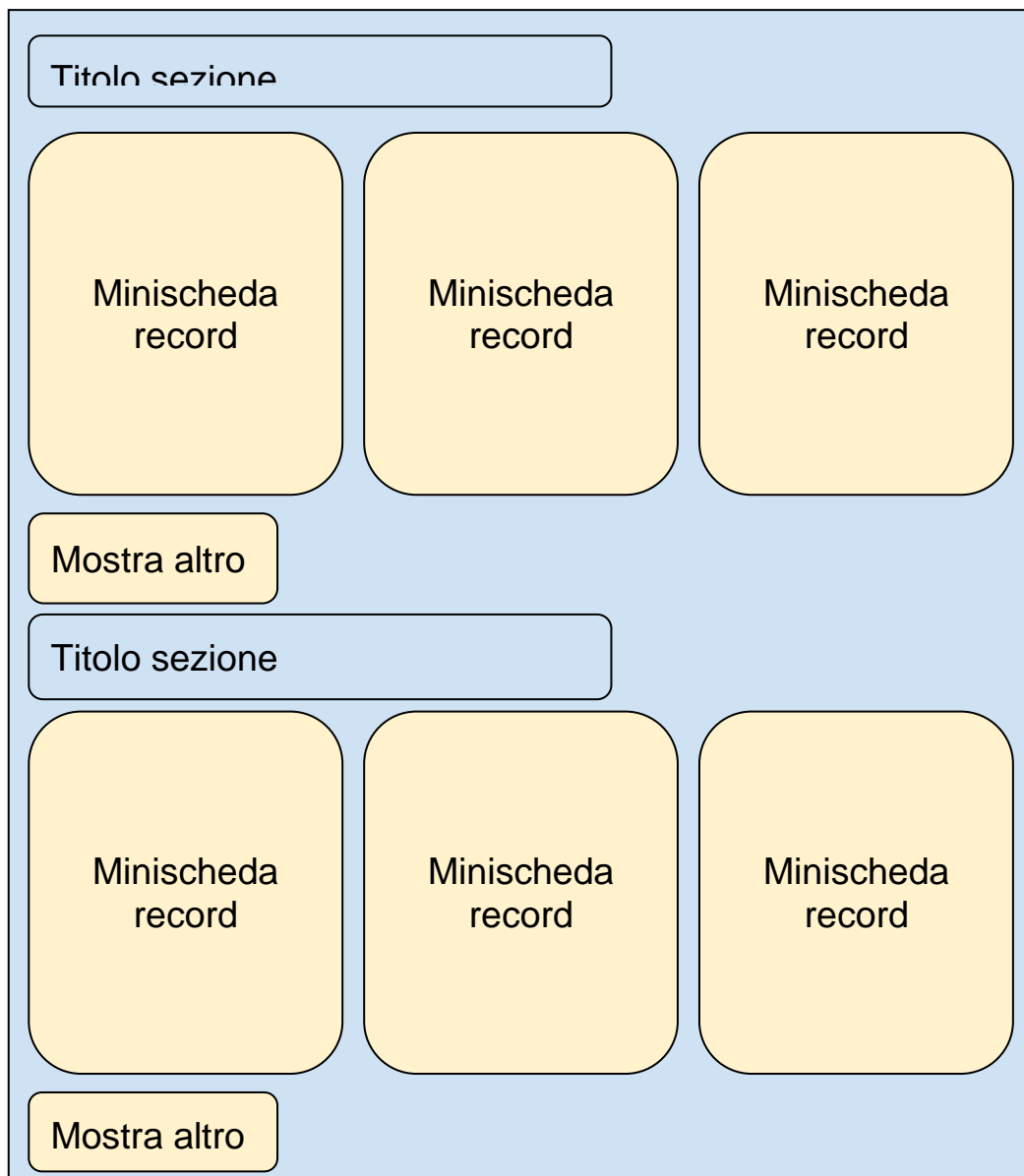
Sarà inoltre presente il link ad una pagina di descrizione dettagliata del progetto e dei partner.



*Figura 9. Footer.*

### 8.1.3 Contenuto principale

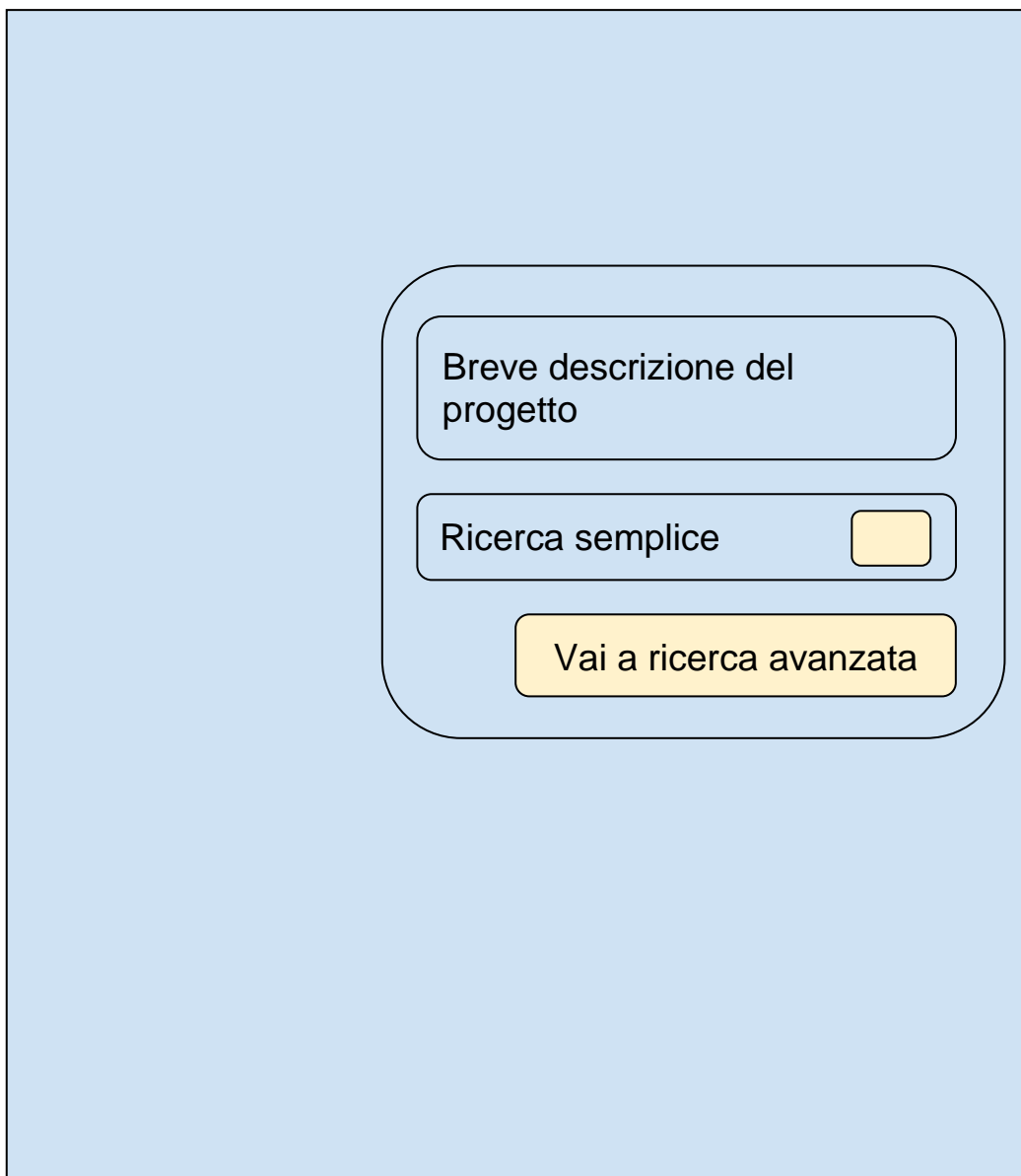
Il contenuto principale nella home page dovrebbe presentare delle schede di esempio, rappresentative del contenuto dell'aggregatore ([Figura 10](#)). Nella prima release dell'aggregatore, queste schede potrebbero essere estratte attraverso degli automatismi da definire, ad esempio partendo dai temi del progetto o dalle singole collezioni. In prospettiva, per le successive release, si potrebbero identificare dei soggetti che, con una redazione apposita, possano predisporre delle schede ad-hoc.



*Figura 10. Contenuto principale con minischede di esempio.*

Se questo non fosse possibile, la sezione del contenuto principale dovrà mostrare un'immagine, e, sovrapposta, una sezione contenente una breve descrizione del progetto, la casella di ricerca e il bottone per la ricerca avanzata, come mostrato in [Figura 11](#), che si troveranno qui invece che nell'header come precedentemente descritto.

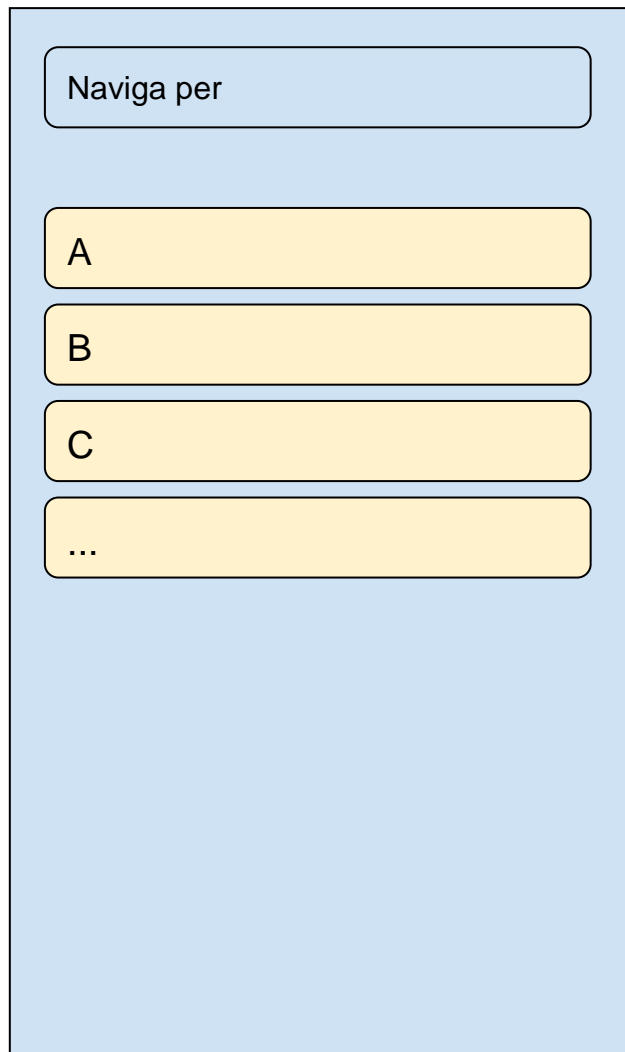




*Figura 11. Contenuto principale con casella di ricerca.*

#### 8.1.4 Navigazione

Questa sezione è dedicata alla navigazione dei contenuti senza aver inserito parole chiave per la ricerca. Cliccando sui pulsanti (i cui valori sono definiti nella [sezione 6.3](#)) vengono filtrati i record contenuti nell'aggregatore e vengono mostrate nella sezione del contenuto principale alcune minischede ordinate in base ai criteri definiti nella [sezione 6.7](#), mentre la barra di navigazione sarà sostituita dalla barra dei filtri.



*Figura 12. Barra di navigazione.*

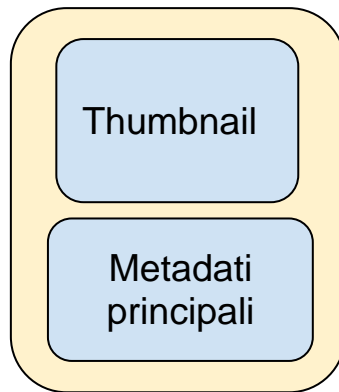
## 8.2 Minischeda record

La minischeda di un record deve presentare in uno spazio ridotto una thumbnail (ricavata a partire dall'immagine originale della risorsa, se presente) e i metadati principali relativi a quel record. E' adatta ad essere mostrata in modalità mosaico insieme ad altre schede nella stessa pagina.

Nel caso in cui non sia presente un'immagine, verrà mostrata un'immagine placeholder dipendente dal tipo di risorsa (`originalResource/dc:type`) a cui il record fa riferimento.

I metadati principali che sono stati identificati per essere riportati nella la minischeda sono indicati nella [sezione 6.5](#).

La minischeda è un elemento cliccabile, e se cliccata conduce alla scheda completa di secondo livello, che viene mostrata riempiendo tutto lo spazio a disposizione nella sezione Contenuto principale.



*Figura 13. Minischeda di primo livello.*

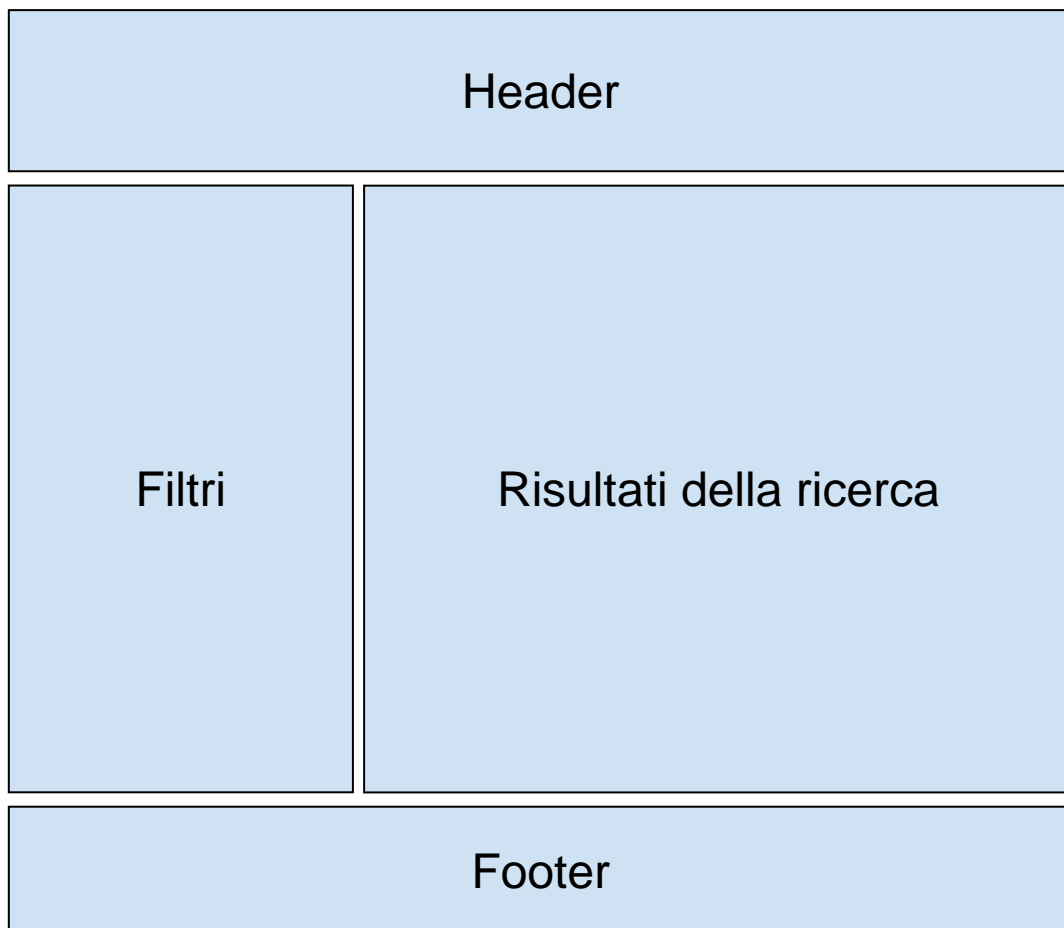
### 8.3 Scheda completa di secondo livello

Questa scheda viene presentata nello spazio del contenuto principale o nello spazio dei risultati della ricerca ([Figura 14](#)).

L'immagine principale è indicata tra i metadati come immagine originale (digitalCopy/sourceURI). Se sono presenti più immagini, l'aggregatore dovrà scegliere la prima e se possibile dovrebbe essere la stessa che era presente nella minischeda. Possono essere presenti altre immagini aggiuntive per la stessa risorsa. Tutte le immagini sono cliccabili per essere mostrate a risoluzione massima.

La scheda presenta anche il titolo della risorsa e la sua descrizione. In basso è presente anche una tabella contenente tutti i metadati contenuti nel tracciato record condiviso. Tra i metadati è presente anche il link alla risorsa originale sul sito del partner, che dovrà ovviamente essere cliccabile.

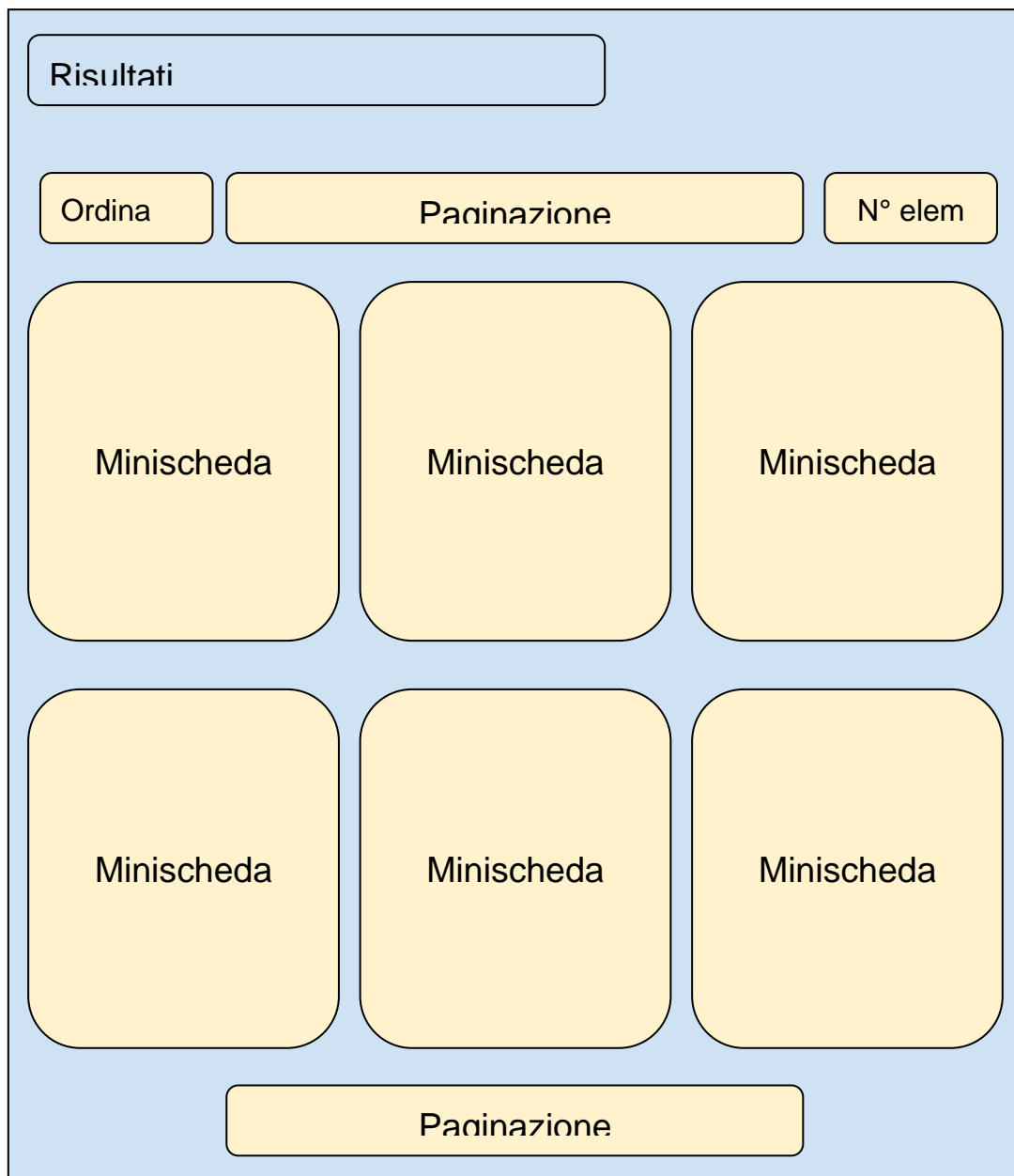




*Figura 15. Presentazione dei risultati a fronte di una ricerca.*

#### 8.4.1 Risultati della ricerca

La sezione che mostra i risultati della ricerca sarà organizzata come in [Figura 16](#): le minischede (non necessariamente tutte della stessa dimensione) saranno posizionate a mosaico. Saranno inoltre presenti i bottoni per la paginazione (uno in cima alla pagina e uno sul fondo), il selettore per il criterio di ordinamento, e il selettore del numero massimo di schede mostrate per pagina.



*Figura 16. Sezione dei risultati della ricerca.*

#### 8.4.2 Filtri

La barra dei filtri verrà usata per raffinare i risultati di una ricerca. Dovranno essere presenti i filtri descritti in [sezione 6.5](#), eventualmente eliminando quelli per cui nei risultati di ricerca non siano presenti dei valori. La [Figura 17](#) mostra una possibile organizzazione della barra, con i filtri uno sotto l'altro, collassabili o espandibili tramite un bottone. Per ogni filtro vengono mostrati tutti i valori presenti nei risultati della ricerca, e, cliccabile, il numero di elementi presenti, così da dare all'utente l'informazione su quante schede rimarranno visualizzate.

The diagram shows a vertical sidebar with a light blue background. At the top is a rounded rectangle labeled "Filtri". Below it is a section titled "Temi del progetto" with a toggle switch to its right. Under this section are two rows: the first row has a rounded rectangle labeled "Tema1" and a yellow rounded rectangle labeled "N°"; the second row has a rounded rectangle labeled "Tema2" and a yellow rounded rectangle labeled "N°". Below these rows are four dots "...".

The next section is titled "Autori" with a toggle switch to its right. Under this section are two rows: the first row has a rounded rectangle labeled "Autore1" and a yellow rounded rectangle labeled "N°"; the second row has a rounded rectangle labeled "Autore2" and a yellow rounded rectangle labeled "N°". Below these rows are four dots "...".

At the bottom of the sidebar is a single rounded rectangle containing three dots "..." and a yellow toggle switch to its right.

*Figura 17. Filtri per raffinare i risultati.*

## 8.5 Ricerca avanzata

Il bottone per la ricerca avanzata presente nell'header (o nella sezione del contenuto principale) mostra nella sezione del contenuto principale i campi per la ricerca avanzata ([Figura 18](#)). Sarà presente una casella di input per ciascuno dei campi di ricerca elencati nella [sezione 6.2](#), più una, in cima, che sarà nuovamente la casella di ricerca semplice.

I risultati saranno presentati nella stessa sezione della pagina, mentre nella sezione a sinistra comparirà la barra dei filtri al posto di quella di navigazione, come per la ricerca semplice.

Ricerca avanzata

Tutti i campi

Autore

Titolo

....

....

Cerca

*Figura 18. Ricerca avanzata.*

## 9 Analisi Statistiche

Il sito dell'aggregatore dovrà raccogliere una serie di statistiche tra cui le seguenti:

- numero di accessi
- tipologie di risorse maggiormente consultate
- le parole chiave ricercate
- profondità della ricerca effettuata (se l'utente si ferma solo al primo livello oppure arriva anche alla scheda sul sito del partner)
- download dei dati
- la lingua con cui gli utenti effettuano le ricerche
- paese da cui proviene la ricerca
- tipo di dispositivo utilizzato dall'utente.

L'aggregatore dovrebbe fornire anche l'andamento dei precedenti indicatori nel tempo. Inoltre, dovrebbe mettere a disposizione dei partner una pagina per la consultazione delle statistiche



descritte. In alternativa, periodicamente, potrebbe inviare a ciascun partner un resoconto delle suddette statistiche.

## 10 Mantenimento e aggiornamento delle risorse

L'aggregatore effettuerà l'aggiornamento delle risorse da tutti i data provider con una frequenza trimestrale. L'aggiornamento da parte dei data provider consiste nel trasmettere l'intero insieme dei metadati relativi alle risorse (solo per quelle che hanno subito aggiornamenti), perché la trasmissione dei soli aggiornamenti richiederebbe lo sviluppo di ulteriori sistemi di controllo sulle variazioni.

I partner che adotteranno la modalità di importazione tramite URL preferibilmente dovranno utilizzare dei permalink. Essi si impegneranno a segnalare all'aggregatore eventuali cambiamenti nelle URL, contenenti i file statici di interscambio, qualora non fossero in grado di fornire dei permalink. Il sistema di importazione dei metadati dell'aggregatore dovrà essere quindi configurabile per la gestione di eventuali cambiamenti nelle URL.

Il mantenimento del sistema di importazione dei metadati e dell'aggregatore stesso sarà garantito dalla Ville de Nice per tutta la durata del progetto e successivamente per un tempo indefinito.

Il mantenimento delle risorse sarà garantito da parte dei partner, che si impegneranno anche a diffondere e pubblicizzare l'aggregatore presso altri enti, in modo che questi possano conoscerlo e partecipare, mettendo a disposizione le proprie risorse.

# 11 Preservazione e protezione dei contenuti digitali

L'aggregatore dovrebbe prevedere una interfaccia di amministrazione per la gestione dei contenuti digitali e per il monitoraggio dello stato. L'interfaccia per la gestione dei contenuti digitali dovrebbe permettere l'accesso ai dati forniti dai partner, il loro aggiornamento ed eventuali modifiche. L'interfaccia per il monitoraggio dello stato dovrebbe permettere il monitoraggio dello stato fisico della macchina che ospita il software dell'aggregatore, come ad esempio lo stato dei supporti fisici, come descritto in seguito.

I meccanismi descritti in seguito sono stati elaborati sulla base di quello fornito dalla Library of Congress<sup>3</sup>.

## 11.1 Preservazione

Per preservazione dei documenti digitali si intende la conservazione e il mantenimento dei dati nel tempo. L'aggregatore dovrebbe prevedere un sistema di monitoraggio dello stato dei supporti fisici utilizzati per la memorizzazione dei contenuti e prevedere la migrazione a nuovi supporti nel caso di obsolescenza o degrado dei supporti utilizzati. In caso di funzionamento normale, si potrebbe prevedere un controllo generale dei supporti su base annua. In caso di anomalie, l'aggregatore dovrebbe segnalarle tempestivamente, tramite appositi sistemi di notifica, agli addetti autorizzati.

Analogamente, anche le macchine che ospitano i sistemi di gestione dei partner dovranno prevedere dei meccanismi di monitoraggio dello stato dei supporti fisici.

## 11.2 Protezione

Per protezione dei documenti digitali si intende il processo che impedisce la modifica dei documenti da parte di soggetti non autorizzati o una eventuale perdita accidentale.

Per impedire la modifica non autorizzata dei documenti, l'aggregatore dovrà implementare dei meccanismi di controllo degli accessi alla gestione dei contenuti digitali. Soltanto gli utenti autorizzati dovrebbero poter accedere al sistema di gestione dei contenuti provenienti dai vari partner. A tal proposito, si deve prevedere almeno un utente amministratore del sistema, che, in base alle necessità dettate ad esempio dai vari progetti, può creare o eliminare utenti abilitati all'accesso di una o più parti del sistema di gestione.

Per prevenire la perdita accidentale dei documenti, l'aggregatore dovrà implementare delle politiche di backup periodico dei dati e conseguenti politiche di ritenzione.

---

<sup>3</sup> <https://www.loc.gov/programs/digital-collections-management/inventory-and-custody/responsibility-for-digital-collections-content-and-storage-systems/>

## 12 Appendice A - Il modello di Europeana

Il modello di Europeana si basa sulla pubblicazione di risorse digitali appartenenti a collezioni, che includono (ma non sono limitati a) i seguenti requisiti<sup>4</sup>:

- 1) i metadati su Europeana relativi ad una risorsa devono puntare ad una risorsa digitale liberamente accessibile sul Web
- 2) gli utenti devono poter accedere alle risorse senza autenticarsi
- 3) preferire la qualità delle risorse piuttosto che la quantità
- 4) dare maggiore priorità al miglioramento delle risorse esistenti piuttosto che aggiungerne nuove
- 5) valutare se i nuovi contenuti soddisfano i temi proposti da Europeana.

La politica di Europeana per la prioritizzazione della pubblicazione delle risorse si basa sull'analisi statistica degli accessi al portale (ad esempio, oggetti maggiormente visitati), in modo da definire il tipo di audience. Prima di poter inviare una collezione di risorse ad Europeana, il fornitore di dati deve firmare l'Europeana Data Exchange Agreement (DEA)<sup>5</sup>, che stabilisce i termini in base ai quali Europeana può fare uso delle anteprime e dei metadati descrittivi forniti dai fornitori di dati.

Europeana definisce due tipologie di linee guida per la pubblicazione di risorse: linee guida per il contenuto<sup>6</sup> e linee guida per i metadati<sup>7</sup>. Le linee guida per il contenuto prevedono 4 livelli di pubblicazione (tiers)<sup>8</sup>, che dipendono dall'obiettivo che si pone il fornitore di risorse. I livelli di pubblicazione sono i seguenti:

- 1) Livello 1. Europeana come motore di ricerca: "Voglio che gli utenti trovino le mie collezioni"
- 2) Livello 2. Europeana come vetrina: "Voglio che gli utenti trovino e visualizzino le mie collezioni su Europeana"
- 3) Livello 3. Europeana come piattaforma di distribuzione: "Voglio che gli utenti trovino, visualizzino e possano utilizzare le mie collezioni, ma non a scopo di lucro"
- 4) Livello 4. Europeana come piattaforma per il libero riuso: "Voglio che gli utenti trovino, visualizzino e possano utilizzare le mie collezioni in qualsiasi modo vogliano".

La tabella seguente mostra i requisiti per la digitalizzazione delle risorse appartenenti ai vari tiers.

---

<sup>4</sup> <https://pro.europeana.eu/post/europeana-content-strategy>

<sup>5</sup> <https://pro.europeana.eu/page/the-data-exchange-agreement>

<sup>6</sup> [https://pro.europeana.eu/files/Europeana\\_Professional/Publications/Publishing\\_Framework/Europeana\\_publishing\\_framework\\_content.pdf](https://pro.europeana.eu/files/Europeana_Professional/Publications/Publishing_Framework/Europeana_publishing_framework_content.pdf)

<sup>7</sup> [https://pro.europeana.eu/files/Europeana\\_Professional/Publications/Publishing\\_Framework/Europeana\\_publishing\\_framework\\_metadata\\_v-0-8.pdf](https://pro.europeana.eu/files/Europeana_Professional/Publications/Publishing_Framework/Europeana_publishing_framework_metadata_v-0-8.pdf)

<sup>8</sup> <https://pro.europeana.eu/post/publishing-framework>



TIER	MEDIA				
	IMAGE	SOUND	VIDEO	TEXT	3D
1	- edm:type is IMAGE - Thumbnail is available* - An image is available** with resolution >= 0.1mp	- edm:type is SOUND - working edm:isShownAt	- edm:type is VIDEO - An image is available with resolution >= 0.1mp as edm:object - Thumbnail is available - A video is available** or working edm:isShownAt	- edm:type is TEXT - An image is available** with 0.1mpx (ie. also means that tech metadata is available) OR working edm:isShownAt	- edm:type is 3D - A 3D resource is available (with tech metadata)
2	- edm:type is IMAGE - Thumbnail is available* - An image is available** with resolution >= 0.42mpx (~800x533)	- edm:type is SOUND - A sound is available** OR embeddable media as edm:isShownBy	- edm:type is VIDEO - An image is available with resolution >= 0.1mp as edm:object - Thumbnail is available - A video is available** with vertical resolution equal or higher than 480 pixels OR embeddable media as edm:isShownBy	- edm:type is TEXT - A text resource is available** with mimetype "application/pdf" OR image is available** with resolution >= 0.42mpx	- edm:type is 3D - A 3D resource is available**
3	- edm:type is IMAGE - Thumbnail is available* - An image is available** with resolution >= 0.95mp (1200x800) - associated image has either open or restricted license	- edm:type is SOUND - A sound is available** OR embeddable media as edm:isShownBy - associated sound has either open or restricted license	- edm:type is VIDEO - An image is available with resolution >= 0.1mp as edm:object - Thumbnail is available - A video is available** with vertical resolution equal or higher than 480 pixels OR embeddable media as edm:isShownBy - associated video has either open or restricted license	- edm:type is TEXT - A text resource** is available with mimetype "application/pdf" OR image is available** with resolution >= 0.95mpx (1200x800) - associated image or text resource has either open or restricted license	- edm:type is 3D - A 3D resource is available** - associated 3D resource has either open or restricted license
4	- edm:type is IMAGE - Thumbnail is available* - An image is available** with resolution >= 0.95mp (1200x800) - associated image has only open license	- edm:type is SOUND - A sound is available** OR embeddable media as edm:isShownBy - associated sound has only open license	- edm:type is VIDEO - An image is available with resolution >= 0.1mp as edm:object - Thumbnail is available* - A video is available** with vertical resolution equal or higher than 480 pixels OR embeddable media edm:isShownBy - associated video has only open license	- edm:type is TEXT - A text resource is available with mimetype "application/pdf" OR image with resolution >= 0.95mp (1200x800)** - associated image or text resource has only open license	- edm:type is 3D - A 3D resource is available** - associated 3D resource has only open license

Figura 19.

Relativamente alle linee guida per i metadati, sono definiti tre livelli (tiers), con requisiti differenti, che dipendono dalla percentuale di metadati compilati per la descrizione di risorse. I tiers sono i seguenti:

- 1) A - almeno il 25% dei metadati rilevanti relativi ad una risorsa ha almeno un tag corrispondente alla lingua;
- 2) B - almeno il 50% dei metadati rilevanti relativi ad una risorsa ha almeno un tag corrispondente alla lingua;
- 3) C - almeno il 75% dei metadati rilevanti relativi ad una risorsa ha almeno un tag corrispondente alla lingua;

I metadati forniti per ogni risorsa devono essere rappresentati utilizzando l'Europeana Data Model (EDM)<sup>9</sup>, che prevede alcuni campi obbligatori e altri facoltativi. Ogni metadato deve essere codificato secondo la codifica UTF-8. I campi dell'EDM per la descrizione di una risorsa culturale sono i seguenti:

<sup>9</sup> <https://pro.europeana.eu/page/edm-documentation>

Properties for edm:ProvidedCHO	
+ dc:contributor	dcterms:isReplacedBy
dc:coverage	dcterms:isRequiredBy
+ dc:creator	+ dcterms:issued
+ dc:date	dcterms:isVersionOf
→ dc:description	dcterms:medium
dc:format	dcterms:provenance
+ dc:identifier	dcterms:references
✓ dc:language (if edm:type = TEXT)	dcterms:replaces
+ dc:publisher	dcterms:requires
dc:relation	○ dcterms:spatial
dc:rights	dcterms:tableOfContents
+ dc:source	○ dcterms:temporal
○ dc:subject	edm:currentLocation
→ dc:title	edm:hasMet
○ dc:type	edm:hasType
+ dcterms:alternative	edm:incorporates
dcterms:conformsTo	edm:isDerivativeOf
+ dcterms:created	+ edm:isNextInSequence
dcterms:extent	edm:isRelatedTo
dcterms:hasFormat	edm:isRepresentationOf
dcterms:hasPart	edm:isSimilarTo
dcterms:hasVersion	edm:isSuccessorOf
dcterms:isFormatOf	edm:realizes
+ dcterms:isPartOf	✓ edm:type
dcterms:isReferencedBy	owl:sameAs

Figura 20.

✓ = **Mandatory property**

→ Blue = **at least** one of the blue properties should be present (and can be used alongside each other)

○ Red = **at least** one of the red properties should be present (and can be used alongside each other)

◆ Green = **at least** one of the green properties should be present (and can be used alongside each other)

+ = recommended property

Detailed characteristics of all properties are given in the full tables following this summary.

Figura 21.

Se si vuole includere delle risorse all'interno di una o più collezioni tematiche, occorre definire delle parole chiave specifiche nei campi *dc:type* e *dc:subject*. Per ogni tema, Europeana definisce un vocabolario controllato di parole chiave. Ad esempio, per il tema archeologia, le parole chiave sono le seguenti: *archaeology*, *archaeological sites*, *field archaeology*, *buildings archaeology*, *excavations (sites)*, *archaeological objects*.

Europeana definisce tre criteri per valutare la qualità dei metadati:

- 1) lingua, che permette di definire la lingua di ogni metadato, attraverso il tag *xml:lang*;
- 2) elementi abilitanti, che definiscono dei metadati non obbligatori, ma appunto abilitanti per l'applicazione della risorsa in particolari scenari;
- 3) classi contestuali, che definiscono il contesto della risorsa, come ad esempio una descrizione dei luoghi o delle date relative alla risorsa. Per le classi contestuali, Europeana raccomanda l'uso di riferimenti ai Linked Data.

Dal punto di vista della licenza, mentre i metadati associati alle risorse possono essere utilizzati senza alcuna restrizione, le risorse digitali possono essere rilasciate secondo una delle 14 licenze specificate dalla politica di Europeana<sup>10</sup>.

## 12.1 Modalità di ricerca e presentazione dei risultati

Come descritto nella [Tabella 5](#), del documento 3.1.5 Inventario Siti Web, Europeana permette la navigazione dei contenuti secondo le seguenti modalità: colori, fonti, persone, argomenti, gallerie, periodi di tempo, esposizioni. Le modalità di esplorazione sono mostrate attraverso un menù a tendina (mostrato nella [Figura 22](#)).



Figura 22 Modalità di esplorazione dei contenuti in Europeana.

Selezionando una delle opzioni del menù a tendina si può accedere ad un'ulteriore possibilità di scelta. Ad esempio, selezionando l'opzione I periodi di tempo, è mostrata la seguente lista cliccabile ([Figura 23](#)).



Figura 23. Lista cliccabile dei vari periodi di tempo in Europeana.

<sup>10</sup> <https://pro.europeana.eu/page/available-rights-statements>

Scegliendo poi uno degli oggetti della lista (ad esempio 1700-1799) si apre la lista degli oggetti relativi a quel periodo ([Figura 24](#)).

The screenshot shows the Europeana search interface. On the left, there is a sidebar with filters under 'AFFINA LA TUA RICERCA'. The 'COLLEZIONI' section includes 'All Items' (selected), '1914-1918', 'Archaeology', 'Art', and 'Fashion'. The 'MEDIA' section includes 'Immagine (307,860)', 'Testo (110,472)', '3D (195)', and 'Suono (1)'. Below these are sections for 'POSSO USARLO?', 'PAESE FORNITORE', and 'LINGUA'. The main content area shows search results for '1 - 12 dei 418,528 risultati'. Two items are visible: 'O.L.Vrouw van Loreto (1750)' and 'zonnemonstrans (1728)'. Each item has a description, a link to 'Vedi tutto Museum De Mindere, Sint-Truiden', and an 'Immagine' icon. The first item is a paper with a seal, and the second is a golden monstrance.

Figura 24. Lista degli oggetti relativi al periodo di tempo 1700-1799 in Europeana.

Oltre alle semplici esplorazioni, Europeana mette a disposizione anche le Esposizioni, che sono una sorta di mostre preconfezionate su un dato argomento. Si può accedere alle singole esposizioni attraverso un altro menù a tendina ([Figura 25](#)).

The image shows two side-by-side screenshots of the Europeana 'Esposizioni' menu. The left screenshot is labeled 'a) Italiano' and shows a dropdown menu with the following items: 'The Sound of the Trenches', 'Still a Working Title? Social and Employment Realities', 'Blue Skies, Red Panic', 'The Rise of Literacy in Europe', 'Heritage at Risk', 'Pioneers', and 'Tutte le esposizioni'. Below the menu, it says '1 - 12 dei 418,528 risultati'. The right screenshot is labeled 'b) Francese' and shows a similar dropdown menu with the following items: 'The Sound of the Trenches', 'Still a Working Title? Social and Employment Realities', 'Blue Skies, Red Panic', 'The Rise of Literacy in Europe', 'Heritage at Risk', 'Pioneers', and 'Toutes les expositions...'. Below the menu, it says '3,162,550 oeuvres d'art, objets, livres, v'.

Figura 25 Menù a tendina per la scelta delle esposizioni in Europeana.

Un esempio di esposizione è dato da Heritage at Risk<sup>11</sup>, che analizza il patrimonio culturale a rischio. Questa esposizione è divisa in sottoargomenti, ognuna delle quali è navigabile ([Figura 26](#)).

<sup>11</sup> <https://www.europeana.eu/portal/en/exhibitions/heritage-at-risk>



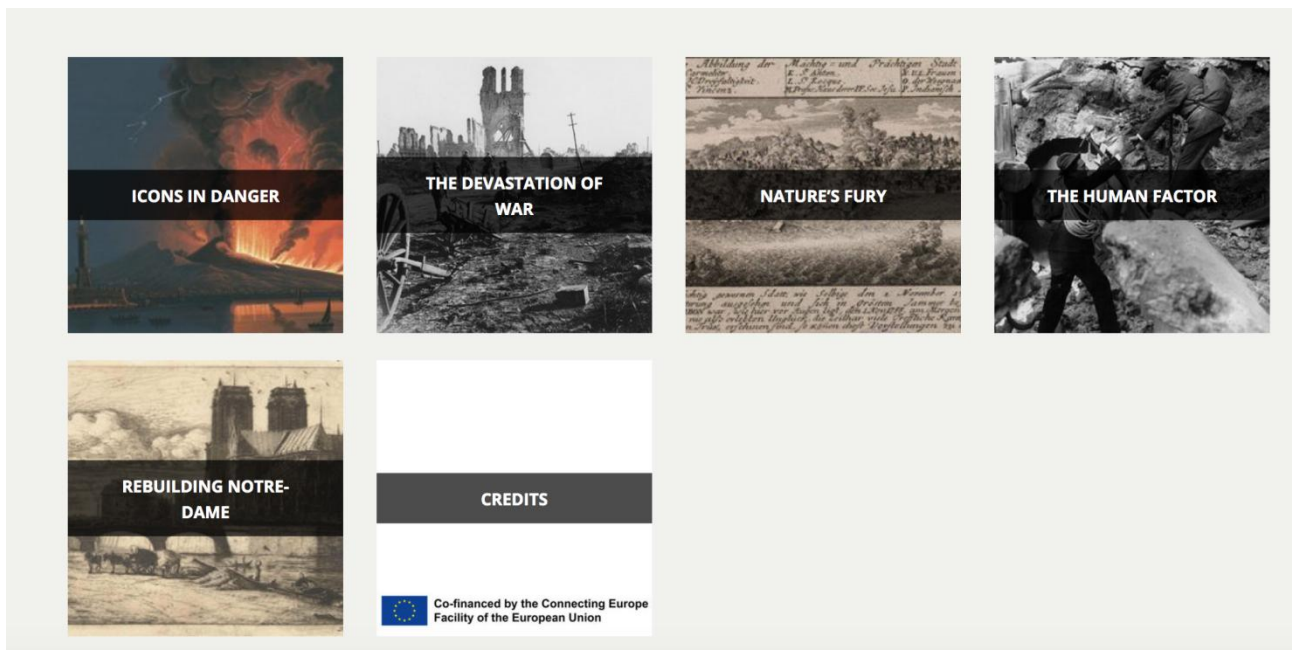


Figura 26. Sottoargomenti dell'esposizione Heritage at Risk in Europeana.

La [Figura 27](#) mostra un esempio di modalità di presentazione dei risultati di primo livello.

### Self-portrait (1887) | Gogh, Vincent van

Vincent moved to Paris in 1886, after hearing from his brother Theo about the new, colourful style of French painting. Wasting no time, he tried it out in several self-portraits. He did this mostly to avoid having to pay for a model. Using rhythmic brushstrokes in striking...

Pittura

[Vedi tutto Rijksmuseum](#)

Immagine



Figura 27. Presentazione dei risultati di primo livello in Europeana.

## 3.2 Définition des outils de partage

### Région Piémont

#### **Délégués**

Istituto di Informatica e Telematica - CNR

CSI Piemonte

# Sommaire

## [1 Glossaire](#)

## [2 Introduction](#)

## [3 Architecture fournisseur de données/agrégateur](#)

## [4 Ensemble de métadonnées communes pour l'interopérabilité](#)

## [5 Modalité d'échange](#)

### [5.1 Collecte](#)

### [5.2 Importation](#)

## [6 Recherche et exploration](#)

### [6.1 Recherche style Google](#)

### [6.2 Recherche avancée](#)

### [6.3 Navigation](#)

### [6.4 Parcours thématiques](#)

### [6.5 Présentation des résultats d'une recherche](#)

### [6.6 Pertinence des résultats](#)

### [6.7 Classement des résultats](#)

## [7 Multilinguisme](#)

### [7.1 Recherche multilingue](#)

### [7.2 Présentation des résultats dans plusieurs langues](#)

### [7.3 Interface graphique de l'agrégateur multilingue](#)

## [8 Interface graphique web](#)

### [8.1 Page d'accueil](#)

#### [8.1.1 En-tête](#)

#### [8.1.2 Footer](#)

#### [8.1.3 Contenu principal](#)

#### [8.1.4 Navigation](#)

### [8.2 Mini-fiche record](#)

### [8.3 Fiche complète de deuxième niveau](#)

### [8.4 Présentation des résultats](#)

#### [8.4.1 Résultats de la recherche](#)

#### [8.4.2 Filtres](#)

### [8.5 Recherche avancée](#)

## [9 Analyses statistiques](#)

## [10 Maintien et mise à jour des ressources](#)

## [11 Préservation et protection du contenu numérique](#)

### [11.1 Préservation](#)

### [11.2 Protection](#)

## [12 Annexe A - Le modèle de Europeana](#)

### [12.1 Modalités de recherche et présentation des résultats](#)

# 1 Glossaire

## **Ressource d'origine**

Ressource culturelle telle que décrite dans le système natif, comme par exemple une peinture.

## **Copie numérique**

Résultat d'un processus de numérisation d'une ressource d'origine. Plusieurs copies numériques peuvent correspondre à une ressource d'origine, par exemple différentes photos prises de la même peinture.

## **Métadonnée**

Attribut qui spécifie une propriété de la ressource d'origine ou de la copie numérique (par exemple le titre d'une ressource d'origine ou le format de la copie numérique).

## **Record**

L'ensemble de toutes les métadonnées d'une ressource d'origine et des copies numériques associées.

## **Ensemble de métadonnées pour l'interopérabilité**

Schéma des collections de record fournies à l'agrégateur.

## **Format de l'ensemble de métadonnées pour l'interopérabilité**

Un des formats suivants : xml, json, csv, ...

## 2 Introduction

Selon la DTD du projet, ce document se compose de deux parties fondamentales :

1. **Définition des protocoles et des outils de partage.** On va présenter les modalités d'implémentation de l'interopérabilité entre différents outils disponibles entre les partenaires pour préparer l'agrégateur qui devra présenter les données et les informations culturelles des systèmes des partenaires. Condivision d'un protocole basé sur les règles et méthodes de maintien, conservation et protection des contenus numériques.
2. **Définition des règles et méthodes pour la numérisation.** En tenant compte des résultats de la phase d'analyse, on va définir un document avec toutes les règles à suivre obligatoirement par les partenaires pour la réalisation d'objets numériques pour les rendre nativement interopérables.

La description détaillée du point 2 est examinée dans le document 3.1.2 Recommandations techniques.

Tandis que le document 3.1.1 Analyse Technique décrit le scénario actuel par chaque partenaire, ce document, ensemble au document 3.1.2 Recommandations techniques, décrit la stratégie commune adoptée par les partenaires pour partager leur ressources avec un nœud agrégateur. Cette stratégie a été agréée avec tous les partenaires, à partir de chaque situation spécifique. Par rapport au document, 3.1.2 Recommandations techniques, dans lequel on décrit comment les partenaires doivent se conformer à la stratégie proposée, dans ce document on décrit comment l'agrégateur devra l'implémenter.

### 3 Architecture fournisseur de données/agrégateur

L'agrégateur devra s'interfacer avec les systèmes de gestion suivants :

- Biblioteca Digitale Ligure (BDL) (Regione Liguria)
- Giornali del Piemonte (Regione Piemonte)
- Mèmora (Regione Piemonte)
- Catalogo dei Beni Culturali (CBC) (Regione Valle d'Aosta)
- Syracuse (Ville de Nice)

On estime que, sur chiffres correctement indiquées par les partenaires, l'agrégateur devra gérer environ **2 millions** de records. Le tableau suivant résume, pour chaque partenaire, une estimation de combien de records seront mis à disposition pour l'agrégateur. Les champs vides indiquent qu'au moment, les informations ne sont pas disponibles, mais il seront indiqués dans les prochaines versions du document. On précise que le tableau suivant n'est pas finalisé, dans le sens que des nouveaux matériaux pourront être ajoutés aux ressources déjà spécifiées, à mesure que les numérisations progressent pendant le projet.

Partenaire	Nom collection/archive/initiative	Format copie numérique	Nombre estimé de RECORD des numérisations passées	Nombre estimé de RECORD des nouvelles numérisations
Regione Piemonte	Pages des journaux locaux de la province de Cuneo	PDF	932.000	
	Pages des journaux locaux de la province de Torino	PDF	697.000	95.000
	Fiches de la province de Torino (Memora)	JPG	96.048 (au maximum)	
	Fiches de la province de Cuneo (Memora)	JPG	34.102 (au maximum)	
	Fonds géographiques Italie-France		13.000	
	Fond Patrimoine Linguistique			
	Photos ex FIF et campagnes photographiques			
	Fondo Libri francesi della Biblioteca Storica di Masino			
Regione Valle d'Aosta	Recensement du patrimoine historique d'architecture mineure	JPG, TIFF	900	3.500
Ville de Nice	Livres anciens fonds régional		145	1.500
	Livres modernes fonds régional		329	10.000
	Images Fixes fonds régional		288	5.000
	Périodiques fonds régional		2.290	40.000
Regione Liguria	Périodiques locaux	TIFF, JPG	32.000	environ 15.000
	Cartes postales historiques de	JPG	1.544	
	Matériel cartographique et iconographique	TIFF, JPG	8	
<b>Totale provisoire</b>			<b>1.822.254</b>	<b>183.500</b>

Tableau 1. Ressources mises à disposition par les partenaires

La [Figure 1](#) présente une possible architecture logique de l'agrégateur.

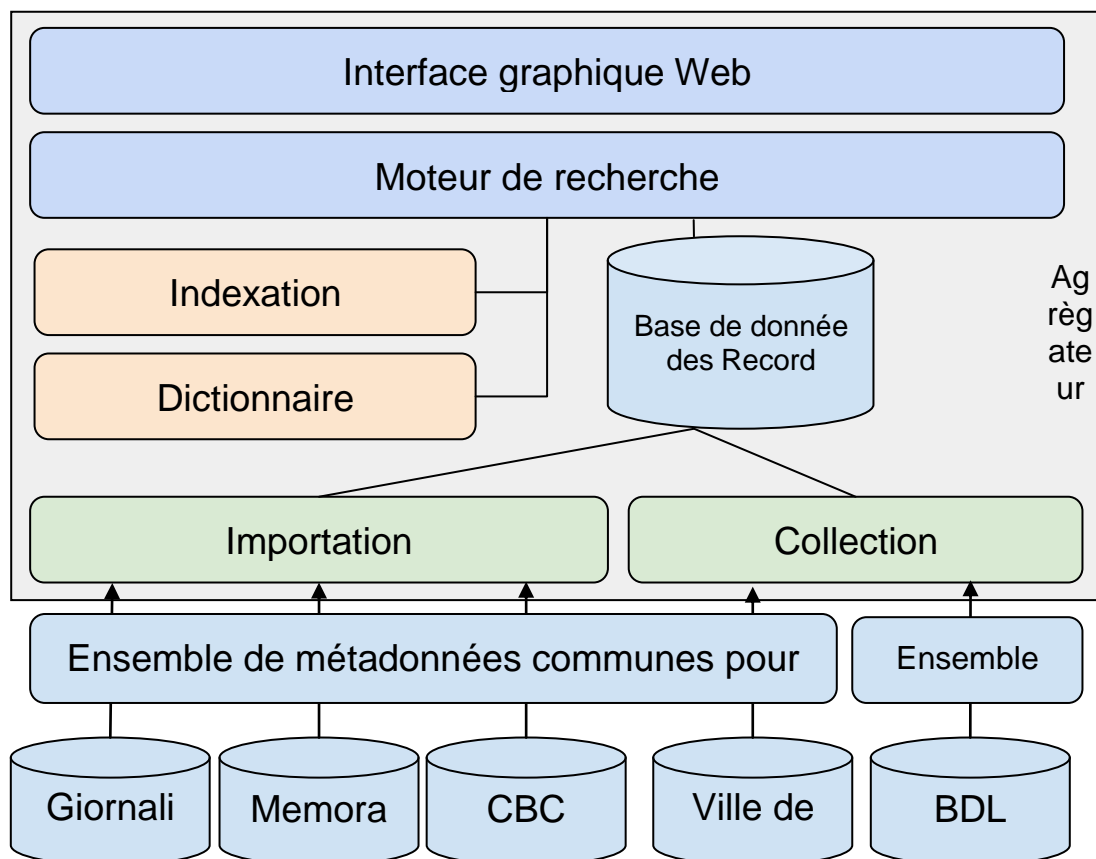


Figure 1. Architecture à grandes lignes de l'agrégateur

Dans la partie en bas de la [Figure 1](#), on trouve les systèmes de gestion avec lesquels l'agrégateur devra s'interfacer dans le cours du projet<sup>12</sup>. Chacun d'eux fournira ses propres records dans le format défini dans l'ensemble de métadonnées communes pour l'interopérabilité, à l'exception de la Regione Liguria, qui fournira les records dans un format propriétaire, avec la correspondance entre son propre format et l'ensemble de métadonnées communes (décrit dans la relative section sur l'ensemble de métadonnées communes pour l'échange).

Les records des partenaires seront mis à disposition de l'agrégateur dans deux modalités :

- importation (téléchargement simple ou par URL, plus de détails en suite)
- collection, par le protocole OAI-PMH.

I Giornali del Piemonte, Memora et VDA utiliseront la première modalité, tandis que BDL et VDN la deuxième. Les détails sur les modalités de collection des records sont décrits dans la section Collection des.

Les records collectés par l'agrégateur (soit par importation ou par collection) seront recueillis dans la base de données des records, qui pourrait être indexée selon les critères décrits ci après. Les problèmes connectés au multilinguisme pourraient être gérées à travers d'un dictionnaire.

L'agrégateur mettra à la disposition un moteur de recherche qui utilisera les critères d'indexation définis pour effectuer les recherches et fournir les résultats finaux aux utilisateurs, à travers d'une interface graphique web (décrits dans la section des exigences pour la recherche et la présentation des résultats).

<sup>12</sup> On n'exclut pas la possibilité que de nouveaux systèmes de gestion pourraient être ajoutés.

## 4 Ensemble de métadonnées communes pour l'interopérabilité

L'agrégateur devra gérer différentes typologies de ressources fournies par les partenaires, notamment :

- Texte - Livres, Journaux, Périodiques, Documents Textuels
- Images - Oeuvres d'art, Photos, Estampes, Dessins, Cartes
- Audio/Vidéo

Les métadonnées fournies par les partenaires à l'agrégateur respecteront une structure qu'on a appelé ensemble de métadonnées communes pour l'interopérabilité. Ce schéma permet de regrouper un ensemble de records à envoyer à l'agrégateur.

Les informations comprises dans l'ensemble de métadonnées communes pour l'interopérabilité respectent les règles suivantes :

- le nombre d'éléments doit être minimale ;
- chaque élément, si ajouté, doit être motivé (voir Annexe 1 colonne "raison") ;
- les raisons sont liées à la recherche qui doit répondre aux questions où, quand et à possibles filtres.

L'ensemble de métadonnées communes pour l'interopérabilité est composé par trois parties fondamentales :

- *header* - contient la description de l'ensemble de records. Il est présent seulement une fois dans l'ensemble de Métadonnées. Il est composé de 5 éléments obligatoires ;
- *record/originalResource* - description sommaire de la ressource d'origine. Il est répété pour chaque ressource d'origine comprise dans l'ensemble de données. Il est composé de 11 éléments, dont 6 obligatoires ;
- *record/digitalCopy* - description d'une copie numérique de la ressource d'origine. Une ressource d'origine peut comprendre plusieurs copies numériques. Il est composé de 5 éléments, dont 4 obligatoires.

Les figures suivantes montrent la structure de l'ensemble de métadonnées. La description complète des éléments de l'ensemble des métadonnées est présentée dans l'Annexe 1 - Ensemble de métadonnées communes pour l'interopérabilité, tandis que le schéma de l'ensemble de métadonnées communes pour l'interopérabilité est présenté dans l'Annexe 2 - XML Schéma de l'ensemble de métadonnées communes pour l'interopérabilité.



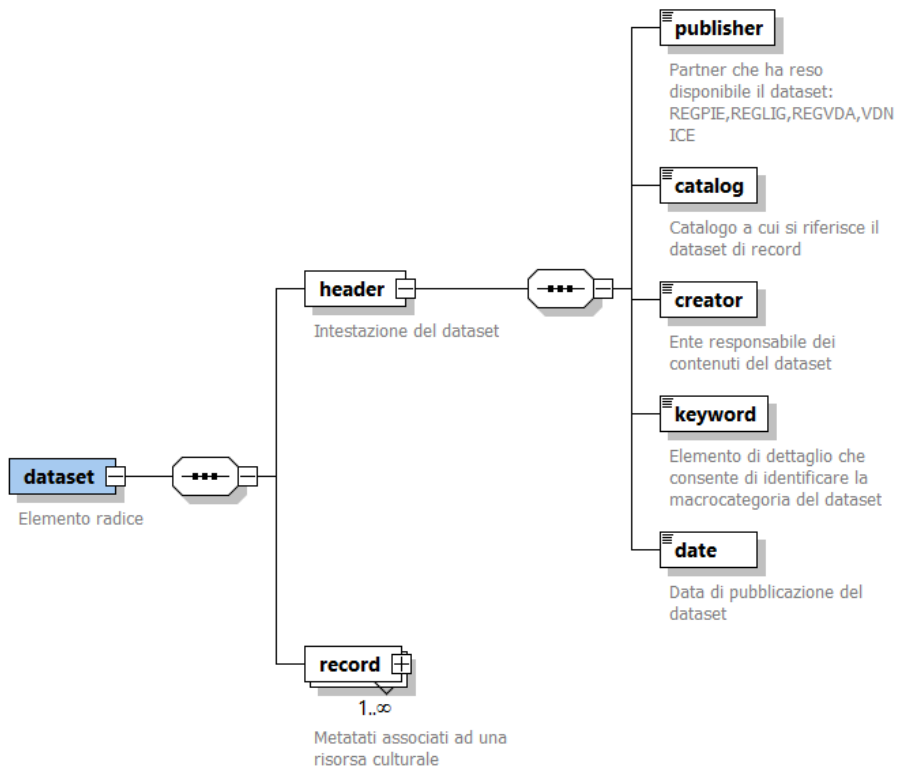


Figure 2. En-tête de l'ensemble de métadonnées communes pour l'interopérabilité (IN ITALIANO)

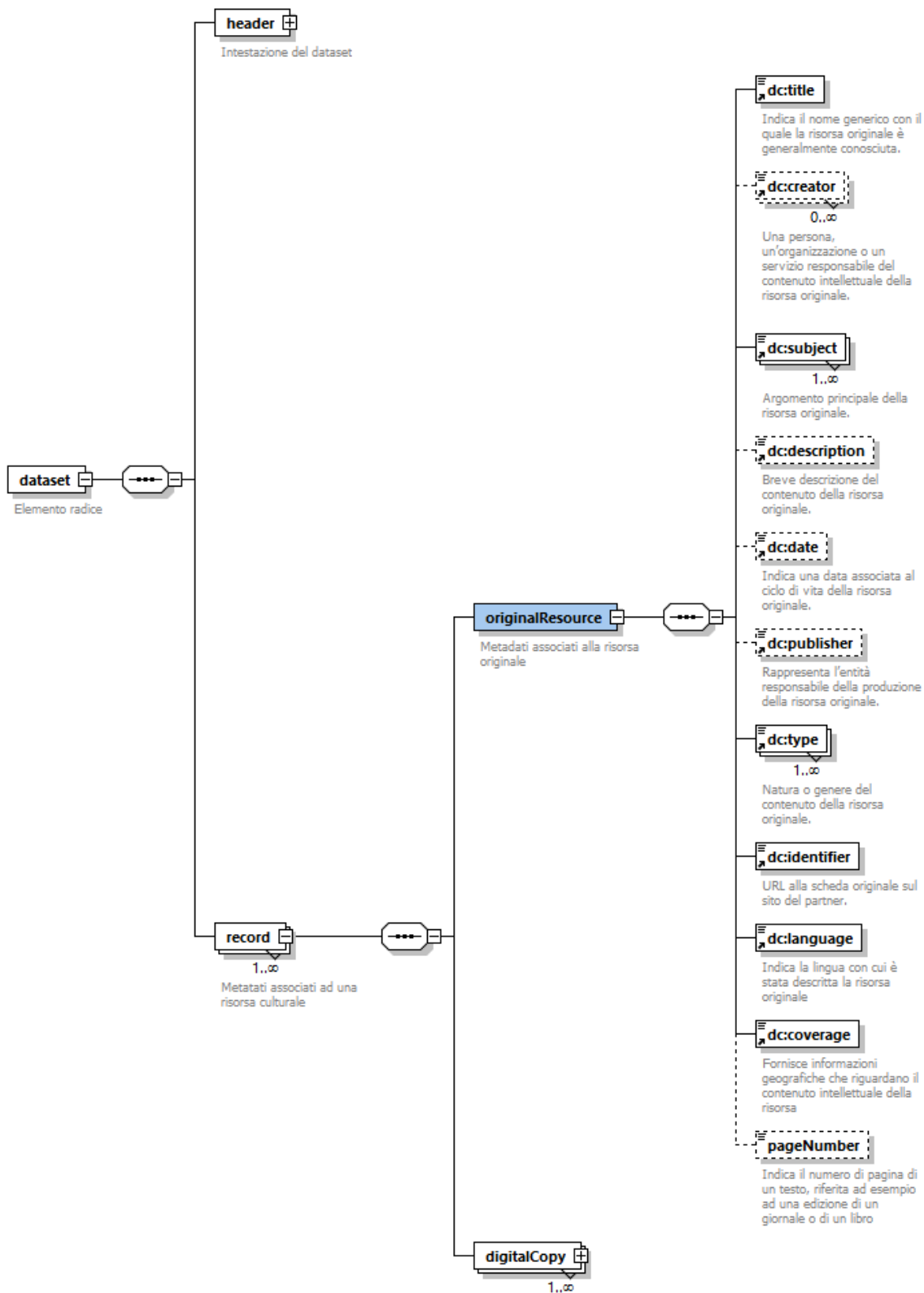


Figure 3. Ressource d'origine

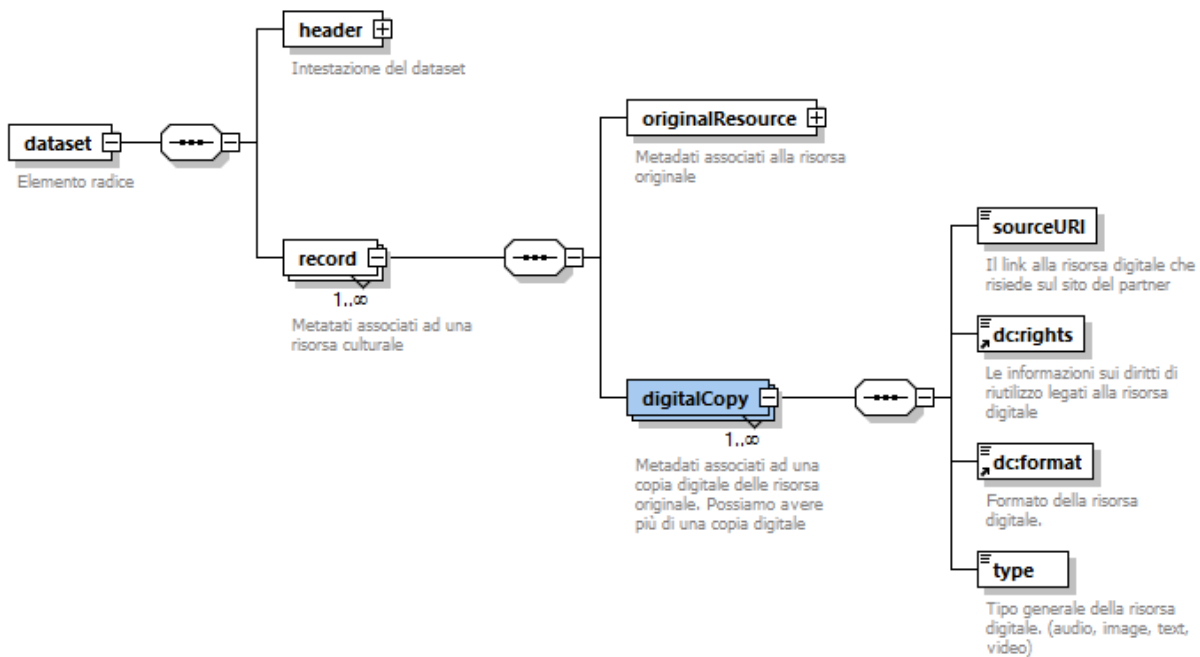


Figure 4. Copie numérique

Cet ensemble de métadonnées communes affecte soit le schéma du fichier statique. Les partenaires qui envoient les données via OAI-PMH devront assurer une correspondance entre l'OAI:DC et les métadonnées communes pour l'interopérabilité.

Au moment de l'élaboration de ce document, tous les partenaires, à l'exception de la Regione Liguria fourniront leurs records selon le ensemble de métadonnées communes pour l'interopérabilité.

Pour ce qui concerne Regione Liguria, les métadonnées de la Biblioteca Digitale Ligure au moment sont tous en format MAG<sup>13</sup>, que pour la description du document d'origine utilise les éléments du Dublin Core.

Ici en suite un exemple <http://bibliotecadigitale.regione.liguria.it/ark:/64639/SGEE001421/134>. Le fragment de code suivant présente l'élément <bib> entier qui contient aussi les informations sur le spécimen numérisé (localisation, inventaire et emplacement). L'attribut level de l'élément <bib> indique le niveau bibliographique. Les codes utilisés sont ceux de l'unimarc (a=dépouillement, m=monographie, s=périodique, c=collection).

```
<bib level="m">
  <dc:identifrier>SGEE001421</dc:identifrier>
  <dc:title>Atlas Minor</dc:title>
  <dc:creator>Mercator, Gerardus </dc:creator>
  <dc:publisher>Amsterdam : Adrianus Belthius</dc:publisher>
  <dc:subject>Mondo</dc:subject>
  <dc:date>1610</dc:date>
  <dc:type>cartografia a stampa</dc:type>
```

<sup>13</sup> Le schéma et le manuel du format MAG sont disponibles au lien suivant : <https://www.iccu.sbn.it/it/normative-standard/linee-guida-per-la-digitalizzazione-e-metadati/standard-mag-versione-2.0.1/index.html>

```

<dc:format>[8], 684 p. ; (180x230 mm)</dc:format>
<dc:language>IT</dc:language>
<dc:coverage>Mondo</dc:coverage>
<holdings>
  <library>Biblioteca civica Berio, Genova, Italia, IT-GE0036</library>
  <inventory_number>222609</inventory_number>
  <shelfmark>B.S.XVII.A.442</shelfmark>
</holdings>
<piece></piece>
</bib>

```

## Ici en suite comme on devrait traduire le record pour l'ensemble de métadonnées communes pour l'interopérabilité

```

<?xml version="1.0" encoding="UTF-8"?>
<dataset xmlns:dc="http://purl.org/dc/elements/1.1/"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:noNamespaceSchemaLocation="./fc.xsd">
  <header>
    <publisher>REGLIG</publisher>
    <catalog>Documenti antichi, manoscritti o a stampa, di carattere geografico
e cartografico</catalog>
    <creator>Biblioteca civica Berio - Genova, IT:GE0036</creator>
    <keyword>Cartografia a stampa</keyword>
    <date>2020-11-11</date>
  </header>
  <record>
    <originalResource>
      <dc:title>Atlas Minor</dc:title>
      <dc:creator>Mercator, Gerardus</dc:creator>
      <dc:subject>Mondo</dc:subject>
      <dc:subject>Cartografia</dc:subject>
      <dc:description>??</dc:description>
      <dc:date>1610</dc:date>
      <dc:publisher>Amsterdam : Adrianus Belthius</dc:publisher>
      <dc:type>image</dc:type><!-- cartografia a stampa -->
      <dc:identifier>http://bibliotecadigitale.regione.liguria.it/
id/SGEE001421</dc:identifier>
      <dc:language>IT</dc:language>
      <dc:coverage>https://www.geonames.org/6295630</dc:coverage><!--Mondo-
->

    </originalResource>
    <digitalCopy>
      <sourceURI>http://bibliotecadigitale.regione.liguria.it/
fedora/objects/mag:134/datastreams/
MMc6601cedd8834f96c1cfc2e04f65cf9a/content</sourceURI>
      <dc:rights>yes</dc:rights>
      <dc:format>image/jpeg</dc:format>
      <type>image</type>
    </digitalCopy>
    <digitalCopy>
      <sourceURI>http://bibliotecadigitale.regione.liguria.it/
fedora/objects/mag:134/datastreams/
MM4c7027861ca89572f3a852c017d8ef7e/content</sourceURI>
      <dc:rights>yes</dc:rights>
      <dc:format>image/jpeg</dc:format>
      <type>image</type>
    </digitalCopy>
    ...
  </record>
</dataset>

```

Le tableau de conversion entre les deux formats est le suivant :

<b>Ensemble de métadonnées communes pour l'interopérabilité</b>	<b>Ensemble de métadonnées de la Ligurie</b>
dataset/header/publisher	
dataset/header/catalog	
dataset/header/creator	metadigit/gen/agency
dataset/header/keyword	
dataset/header/date	
dataset/record/originalResource/dc:title	metadigit/bib/dc:title
dataset/record/originalResource/dc:creator	metadigit/bib/dc:creator
dataset/record/originalResource/dc:subject	metadigit/bib/dc:subject
dataset/record/originalResource/dc:description	
dataset/record/originalResource/dc:publisher	metadigit/bib/dc:publisher
dataset/record/originalResource/dc:date	metadigit/bib/dc:date
dataset/record/originalResource/dc:type	
dataset/record/originalResource/dc:identifiant	
dataset/record/originalResource/dc:language	metadigit/bib/dc:language
dataset/record/originalResource/dc:coverage	
dataset/record/originalResource/sequenceNumber	
dataset/record/digitalCopy/sourceURI	metadigit/img/altImg[1]/file@xlink:href
dataset/record/digitalCopy/dc:rights	
dataset/record/digitalCopy/dc:format	metadigit/img/altImg[1]/format/niso:mime
dataset/record/digitalCopy/type	
dataset/record/digitalCopy/sequenceNumber	metadigit/img/sequence_number

*Tableau 2. Tableau de conversion*

## 5 Modalité d'échange

Cette section décrit les méthodes avec lesquelles l'agrégateur rassemblera les records de métadonnées depuis les systèmes de gestion disponibles chez les partenaires.

L'analyse technique a souligné que :

- la Biblioteca Digitale Ligure et le Portail de la Ville de Nice envoient données à travers serveur OAI-PMH,
- Mèmora et Giornali del Piemonte de la Regione Piemonte et Catalogo dei Beni Culturali de la Regione Valle d'Aosta mettent à disposition données statiques, en XML (Regione Piemonte) et XML ou JSON (Regione Valle d'Aosta).

### 5.1 Collecte

L'extracteur de l'agrégateur devra fonctionner comme un client OAI-PMH :

- il devra interroger les fournisseur OAI-PMH
- il devra gérer fichiers XML structurés selon le schéma de l'ensemble de métadonnées communes ou selon d'autres schémas dont on a une correspondance avec l'ensemble de métadonnées communes.

#### **Biblioteca Digitale Ligure**

URL du fournisseur OAI-PMH :

<http://bibliotecadigitale.regione.liguria.it/oaiprovider/>

Exemple d'utilisation :

<http://bibliotecadigitale.regione.liguria.it/oaiprovider/?verb=GetRecord&identifiser=XXXX> (remplacer XXXX avec l'identifiant du record à télécharger). La barre avant le d'interrogation est nécessaire.

En général, on suit la syntaxe normale des verbes prévue par le protocole OAI-PMH.

Le seul format au présent pris en charge est le MAG.

#### **Portail de la Ville de Nice**

Le fournisseur OAI-PMH est en cours de réalisation.

### 5.2 Importation

Pour les partenaires qui n'ont pas un serveur OAI-PMH, l'agrégateur devra permettre le téléchargement de fichiers statiques, qui comprennent les métadonnées exportées. Dans ce cas, l'agrégateur devra fournir deux modalités de téléchargement des fichiers statiques :

3. réception - URL (par exemple sur HTTP) qui comprend le fichier statique (exemple <http://www.foo.com/records/records.xml>)
4. émission - le partenaire télécharge le fichier statique sur l'agrégateur (par exemple à travers FTP).

Dans le cas de la modalité en réception, l'agrégateur devra avoir la possibilité de récupérer les métadonnées, à travers les protocoles FTP ou HTTP depuis URL envoyés par les partenaires. Dans le cas de la modalité en émission, l'agrégateur devra héberger un serveur sur lequel les partenaires pourront envoyer les fichiers avec les records à intégrer. Dans les deux modalités, les fichiers fournis par les partenaires seront disponibles :

- dans le formats XML ou JSON

- avec le schéma de l'ensemble de métadonnées communes ou autre schéma dont on a la correspondance avec l'ensemble de métadonnées communes.

Au moment de la préparation de ce document, les détails techniques des URL de chaque partenaire ne sont pas disponibles. Dès qu'elles seront disponibles, ce document sera mis à jour avec les URL des ressources mises à disposition par Giornali del Piemonte, Memora et Catalogo dei Beni Culturali.

## 6 Recherche et exploration

Pour ce qui concerne les modalités de recherche, on a identifié quatre mécanismes principaux, déjà décrits dans l'analyse des modèles de diffusion (Rapport 3.1.5 Inventaire des sites web), que l'agrégateur devrait être capable d'implémenter :

- *recherche style Google*, pour les utilisateurs qui ne sont pas des experts du secteur, à travers un champ de texte dans lequel l'utilisateur peut écrire une ou plusieurs mots clé que le système utilise pour faire la recherche dans les champs de texte des ressources dans la base de données ;
- *recherche avancée*, utilisée par des utilisateurs plus experts qui ont une connaissance du secteur, dans laquelle il est possible de configurer des filtres à travers d'un masque de recherche qui montre quelques-uns des champs de métadonnées disponibles. Il serait utile de donner à l'utilisateur la possibilité d'utiliser des opérateurs logiques ;
- *navigation*, à travers sections du site qui regroupent les ressources sur la base de quelques critères (par exemple thèmes du projet, sujets, typologies, collections, périodes de temps, personnes, sources) ;
- *parcours thématiques* créés ad-hoc, qui permettent la valorisation du patrimoine disponible parce qu'ils montrent quelles sont les ressources les plus intéressantes pour l'exploration.

### 6.1 Recherche style Google

Ce type de recherche est le plus familier pour le grand public et s'inspire de la barre de recherche de Google. On recommande que l'interface graphique soit la plus simple possible. La saisie de mots doit être guidée avec le système 'auto-remplir' depuis le quatrième caractère.

La recherche sera effectuée sur les champs de l'ensemble de métadonnées communes qui répondent aux questions suivantes :

- Qui
  - auteur (originalResource/dc:creator)
  - editeur (originalResource/dc:publisher)
- Quoi
  - titre (originalResource/dc:title)
  - description (originalResource/dc:description)
  - sujet (originalResource/dc:subject)

Éventuellement on pourrait ajouter une barre de recherche par date (originalResource/dc:date).

### 6.2 Recherche avancée

Dans les systèmes de gestion des ressources numériques ce type de recherche n'est plus utilisé, car il a été presque complètement remplacé par les filtres. Toutefois, quelques systèmes de gestion prévoient encore la modalité de recherche avancée.

La recherche avancée sera effectuée sur les champs de l'ensemble de métadonnées communes qui répondent aux questions présentes dans la recherche style Google (qui et quoi) en ajoutant les champs suivants :

- Quand
  - date (originalResource/dc:date)
- Où
  - localité (originalResource/dc:coverage)



- institution chez laquelle la ressource est préservée (header/creator)

La modalité de recherche avancée prévoit par défaut que les champs sont associés en modalité AND. Éventuellement on peut prévoir l'utilisation de l'opérateur logique OR applicable à chaque champ.

## 6.3 Navigation

La navigation sera effectuée sur les champs obligatoires de l'ensemble de métadonnées communes pour naviguer entre toutes les ressources du projet. On peut prévoir les types de navigation suivants, implémentés comme des filtres :

- Quoi
  - thèmes du projet (originalResource/dc:subject)
  - typologie de ressources (originalResource/dc:type)
- Où
  - localité (originalResource/dc:coverage)
  - institution chez laquelle la ressource est préservée (header/creator)
- Quand
  - date (originalResource/dc:date)

Semblablement aux typologies de recherche, la navigation aussi devrait permettre de créer des filtres sur les résultats récupérés.

## 6.4 Parcours thématiques

On peut prévoir la définition de quelques parcours thématiques qui valorisent les connexions entre les ressources présents dans différentes collections, tels que, par exemple, *les monuments long de la côte italo-française*. Ces sections nécessitent néanmoins un travail de rédaction additionnel. L'agrégateur devrait prévoir la possibilité d'héberger ces nouveaux parcours thématiques dans une section du site officiel.

## 6.5 Présentation des résultats d'une recherche

A la suite d'une recherche, l'utilisateur arrive à un **premier niveau de détail** des résultats, dans lequel on présente une liste de mini-fiches descriptives, correspondantes aux records extraits. Chaque mini-fiche présentera une prévisualisation (selon le type de ressource), et les métadonnées principales (lieu, date, auteur, sujet, description brève).

En cliquant sur chaque fiche de premier niveau, l'utilisateur pourra accéder à la **fiche complète de deuxième niveau**, qui présente toutes les métadonnées envoyées à l'agrégateur, entre eux aussi le lien à la fiche de la ressource présente sur le site du partenaire fournisseur.

En cliquant sur le lien de la ressource, l'utilisateur pourra consulter la **fiche originale** sur le site du partenaire fournisseur.

Sur les résultats de premier niveau il sera possible d'effectuer des filtre en utilisant des facet. Les filtres possibles devraient comprendre les champs suivants, présents dans l'ensemble de métadonnées communes :

- Qui
  - auteur (originalResource/dc:creator)
  - editeur (originalResource/dc:publisher)

- Quoi
  - sujet (originalResource/dc:subject)
  - thème du projet (originalResource/dc:subject)
  - typologie (originalResource/dc:type)
  - langue (originalResource/dc:language)
  - licence (digitalCopy/dc:rights)
  - format (digitalCopy/dc:format)
- Où
  - localité (originalResource/dc:coverage)
  - institution chez laquelle la ressource est préservée (header/creator)
- Quand
  - date (originalResource/dc:date)

Deux filtres particulièrement importants sont ceux relatifs à la couverture spatiale et à la date. Par rapport à la couverture spatiale, le filtre devrait permettre de limiter la localité d'intérêt ou de sélectionner quelques localités, hiérarchiquement connectées à la localité principale, résultat de la recherche principale (par exemple, si la recherche principale a donné comme résultat toutes les ressources connectées au lieu 'Italie', ce filtre permet de limiter la recherche au lieu 'Pise').

Pour ce qui concerne le filtre par date, on devrait fournir la possibilité de limiter l'intervalle de temps ou sélectionner plusieurs intervalles de temps dans la période résultat de la recherche principale (par exemple si la recherche principale a donnée comme résultat toutes les ressources du 1500, ce filtre permet de limiter la recherche aux seules ressources du 1520).

## 6.6 Pertinence des résultats

A front d'une recherche, la pertinence permet de présenter en évidence les résultats les plus pertinents à une recherche et au fur et à mesure tous les autres résultats. Ce mécanisme est possible grâce à l'attribution de critères sur chaque champ sur lequel la recherche est effectuée.

La pertinence d'un résultat dépend du type de recherche effectuée. Pourtant les valeurs pour les champs seront définies de temps en temps par le moteur de recherche de l'agrégateur et elles ne peuvent pas être définies a priori.

La pertinence d'un résultat dépend des critères suivants :

1. présence des mots clé dans un des champs utilisés pour la recherche (qui dépendent du type de recherche).
2. nombre de fois que le mot clé apparaît dans les champs principaux

Sur la base de ces critères, le moteur de recherche de l'agrégateur pourra élaborer son propre algorithme pour la définition de la pertinence des résultats à front d'une recherche.

## 6.7 Classement des résultats

A front d'une recherche, le classement des résultats permet à l'utilisateur de choisir l'ordre de visualisation. Le moteur de recherche de l'agrégateur permettra d'effectuer le classement sur la base des champs suivants, au minimum :

- auteur (originalResource/dc:creator)
- localité (originalResource/dc:coverage)
- date (originalResource/dc:date).

## 7 Multilinguisme

Les systèmes de gestion disponibles chez les partenaires mettent à disposition les ressources principalement dans une des deux langues : français ou italien. En plus, pour faciliter l'utilisation du moteur de recherche pour la communauté internationale, on recommande aussi le support de la langue anglaise. Pourtant, pour améliorer la recherche et harmoniser la présentation finale des résultats, il est nécessaire de gérer le problème du multilinguisme.

Il y a trois aspects qui concernent le multilinguisme :

1. *recherche multilingue*: indépendamment de la langue, la recherche récupère toujours les mêmes ressources. Par exemple, si on cherche le mot "ponte" ou "pont" ou "bridge" comme sujet, la recherche produit toujours le même résultat ;
2. *présentation des résultats dans plusieurs langues* : le résultats de la recherche seront présentés dans la langue sélectionnée ;
3. *interface graphique de l'agrégateur multilingue* : l'utilisateur peut choisir la langue de visualisation du site web, qui changera les éléments textuels de l'interface tels que, par exemple les menus, les étiquettes, les titres des sections, etc.

### 7.1 Recherche multilingue

La recherche multilingue est référée aux champs de recherche qui peuvent être de deux types :

1. limités à un vocabulaire contrôlé ;
2. texte libre.

Le tableau suivant présente les champs de l'ensemble de métadonnées du type 1 et les vocabulaires associés en italien, français et anglais :

champ	vocabulaire		
	Italien	français	anglais
publisher	REGPIE, REGVDA, REGLIG, VDNICE		
dc:subject	preistoria costruito storico resistenze moderne sacro e profano	préhistoire bâti historique résistances modernes sacré et profane	prehistory historical construction modern resistances sacred and profane
dc:type	collezione immagine oggetto fisico testo suono	collection image objet physique texte du son	collection image physical object text sound
dc:rights	si si con limiti richiedere autorizzazioni	oui oui avec des limites demander des autorisations	yes yes with restrictions require authorization
type	audio testo immagine	audio texte image	audio text image

	video	vidéo	video
--	-------	-------	-------

Tableau 3. Champs avec valeur extrait de dictionnaires.

Le deuxième type de recherche (texte libre) demande que le texte soit traduit dans toutes les langues supportées ou par le fournisseur de données ou par l'agrégateur. Les textes traduits dans les différentes langues seront indexés pour la recherche.

Dans le cadre du projet, la traduction des champs en texte libre n'est pas prévue par les fournisseurs de données, et pourtant elle devrait être confiée totalement à l'agrégateur. Les partenaires devront de toute façon spécifier la langue de chaque record, tel que défini dans l'ensemble de métadonnées communes pour l'interopérabilité.

Une solution alternative pourrait être de ne pas traduire les champs en texte libre, mais de traduire le texte de recherche et effectuer plusieurs recherches, une pour chaque langue dans laquelle les records sont préservés (italien, français). Cette solution prévoit que la langue d'origine (italien ou français) soit spécifié pour chaque record. La figure suivante montre un exemple de ce concept. Naturellement, la langue utilisée pour la recherche doit être spécifiée. L'utilisateur cherche par exemple le mot *bridge*, spécifiant comme langue l'anglais. L'agrégateur dispose d'un traducteur, qui traduit le mot *bridge* en français (*pont*) et en italien (*ponte*). La recherche des deux mots est pour le tant effectuée séparément : dans le records italiens, l'agrégateur recherche le mot *ponte*, tandis que dans les records en français il cherche le mot *pont*. Les résultats sont ensuite présentés d'une façon agrégée à l'utilisateur final, en spécifiant de toute façon la langue d'origine de chaque record.

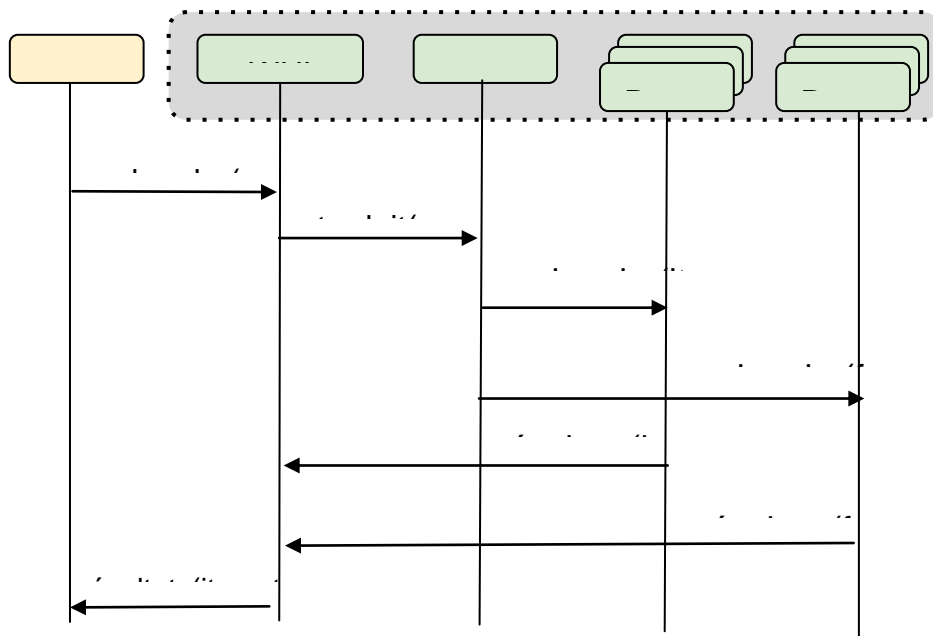


Figure 5. Possible solution pour recherche multilingue.

## 7.2 Présentation des résultats dans plusieurs langues

Indépendamment de comment la recherche a été effectuée, les résultats finals devraient être présentés à l'utilisateur final d'une façon homogène, c'est-à-dire dans la même langue. Donc, même si les record d'origine sont mémorisés dans une langue, si l'utilisateur final a choisi une

autre langue de visualisation, l'agrégateur devrait être capable de traduire le texte contenu dans les métadonnées des records présentées dans le résultat.

Bien que dans le cas spécifique du projet Faire Connaître, les résultats d'une recherche ne peuvent pas être traduits et modifiés, car ils sont le résultat d'un catalogage, l'agrégateur pourrait mettre à disposition de l'utilisateur final une fonctionnalité de traduction automatique pour la compréhension du texte.

Pour les champs dont les valeurs appartiennent à un vocabulaire contrôlé, on prévoit une traduction *na tantum*, tandis que pour les champs en texte libre (description, lieux, etc.) une solution pourrait être l'utilisation de librairies ou API qui permettent la traduction automatique des résultats en phase de présentation finale, sans affecter la langue de mémorisation des records. Une solution de ce type est possible en tant que la langue utilisée est spécifiée obligatoirement pour chaque record. Il faut, de toute façon, prévoir d'identifier des champs spécifiques comme non traduisibles (par exemple titre et auteur).

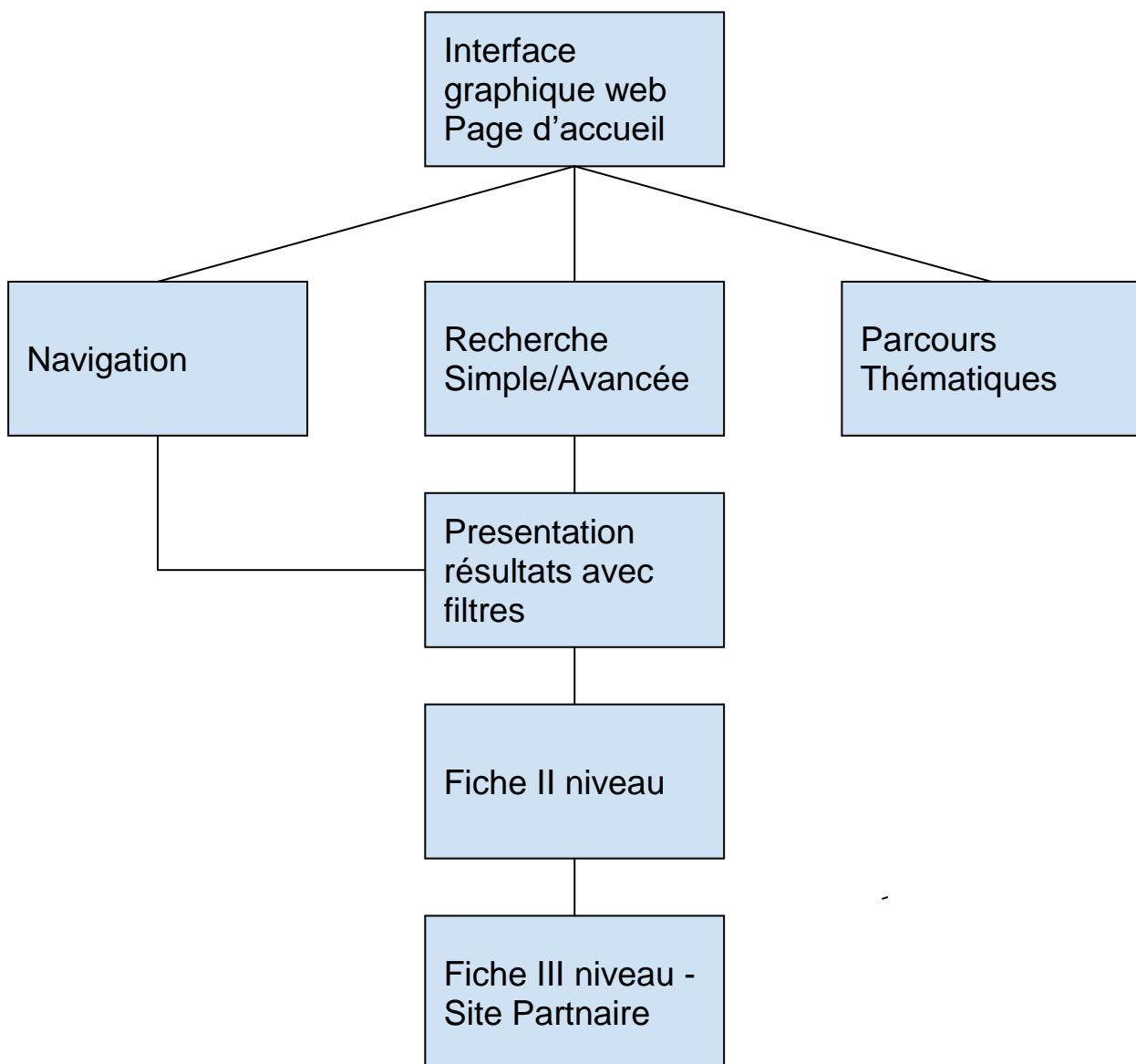
### 7.3 Interface graphique de l'agrégateur multilingue

À la fin d'atteindre le plus grand public, une des nécessités fondamentales pour l'agrégateur est qu'il offre à l'utilisateur la possibilité de naviguer l'interface graphique web (menu, étiquettes, titres) dans plusieurs langues. L'agrégateur devrait présenter l'interface au minimum dans les langues suivantes : italien, français et anglais (UK). Le sélecteur pour choisir la langue doit se trouver dans un point immédiatement visible pour l'utilisateur.

## 8 Interface graphique web

L'interface graphique web devra répondre aux critères d'accessibilité et d'usabilité, en suivant la réglementation en vigueur. En outre, elle devra être responsive, pour en permettre l'utilisation aussi sur appareils mobiles tels que smartphone et tablette.

La figure suivante ([Figure 6](#)) présente le flux de navigation de l'interface graphique web. Depuis la page d'accueil, l'utilisateur pourra choisir trois modalités de navigation des ressources numériques : navigation, recherche simple (style Google)/recherche avancée et parcours thématiques. Dans le cas de la navigation et recherche, l'interface présentera une page avec les fiches de premier niveau associées aux ressources numériques (mini-fiches record) qui répondent aux critères sélectionnés, sur lesquelles il sera possible d'effectuer des filtres. Cliquant sur une mini-fiche, il sera possible d'accéder aux détails sur cette ressource (fiche complète de deuxième niveau). Entre les détails sur la ressource, il y aura aussi le lien à la fiche de la ressource d'origine sur le site du partenaire.



*Figure 6. Diagramme de flux de l'expérience utilisateur.*

Dans les sections suivantes on présente de prototypes des différentes sections de l'interface, avec une description (non nécessairement exhaustive) de leur contenu et fonctionnalités. Les

dimensions sont indicatives des proportions relatives entre les éléments. La couleur jaune indique les éléments interactifs.

## 8.1 Page d'accueil

La page d'accueil devra être organisée d'une façon simple, avec les sections classiques normalement présentes dans une page web. La [figure 7](#) montre la répartition dans les 4 sections prévues : en-tête, navigation, contenu principal et contenu principale, e footer.

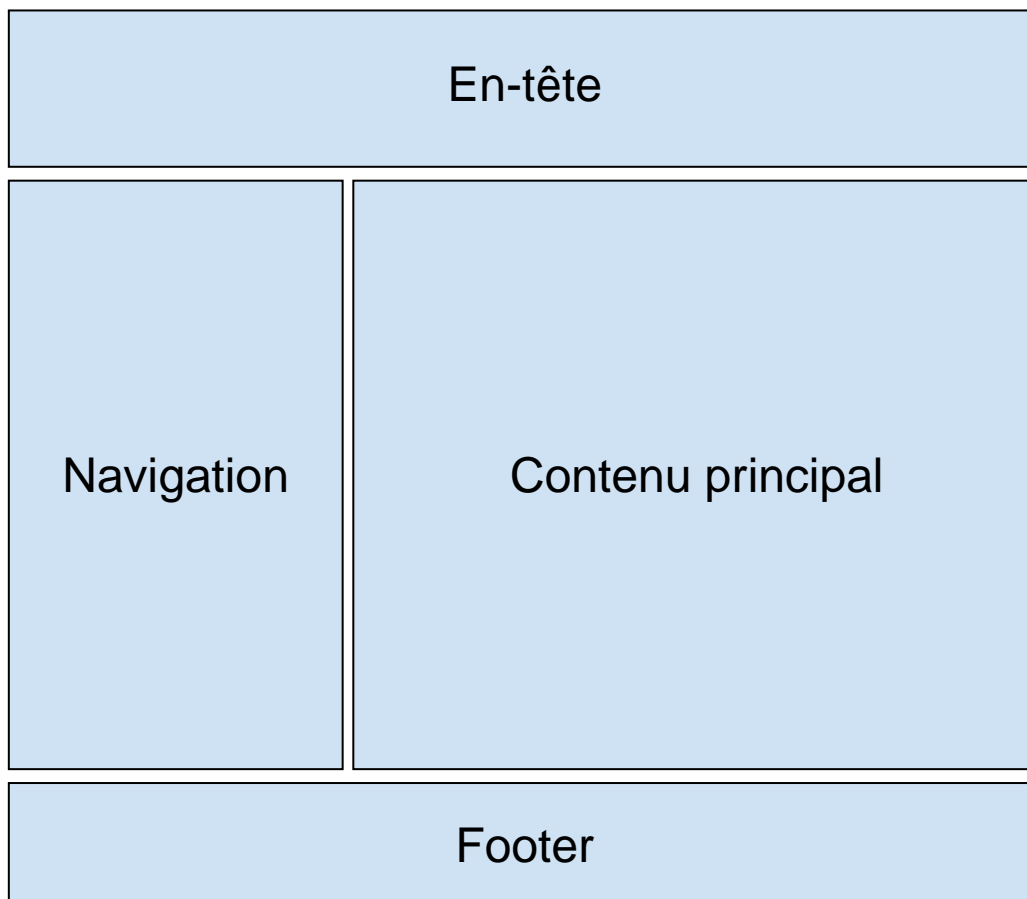


Figure 7. Page d'accueil

### 8.1.1 En-tête

L'en-tête comprendra le logo du projet, une description de synthèse du projet et les boutons pour la sélection de la langue de l'interface (italien, français et anglais).

Il comprendra aussi la boîte pour la recherche simple (style Google), et le bouton pour passer à la recherche avancée.

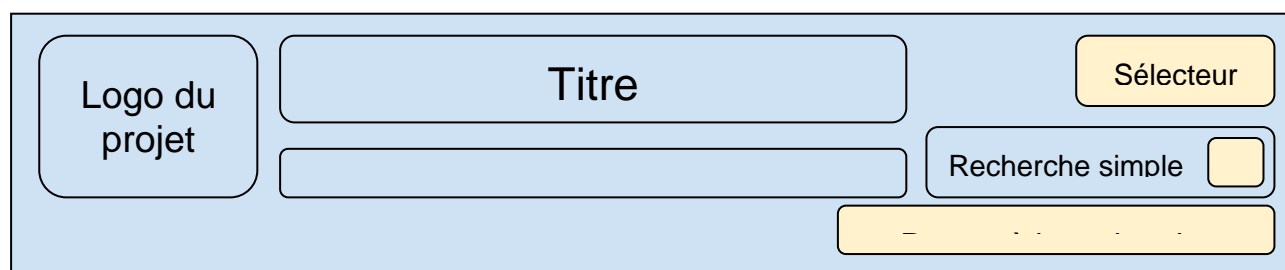
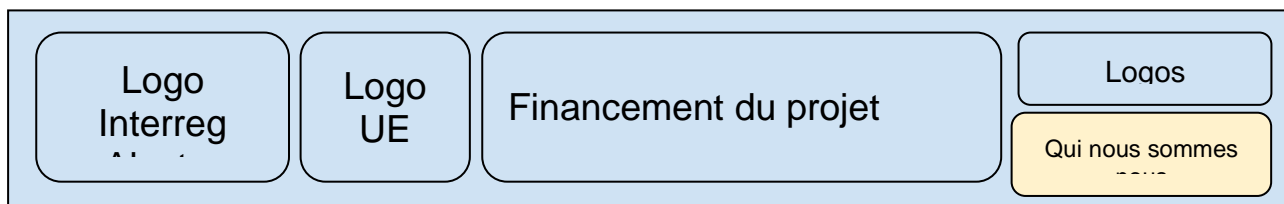


Figure 8. En-tête.

### 8.1.2 Footer

Le footer comprendra tous les logo nécessaires (Interreg Alcotra, Union Européenne, logo de chaque partenaire) et l'indication de l'origine des fonds qui financent le projet.

Il y aura aussi le lien à une page de description de détail du projet et des partenaires.



*Figure 9. Footer.*

### 8.1.3 Contenu principal

Le contenu principal dans la page d'accueil devrait présenter des fiches d'exemple, représentatives du contenu de l'agrégateur ([Figure 10](#)). Dans la première sortie de l'agrégateur, ces fiches pourraient être extraites à travers des automatismes à définir, par exemple depuis les thèmes du projet ou des collections. A terme, pour les sorties suivantes, on pourrait identifier des sujets que, avec une rédaction appropriée, puissent préparer des fiches ad-hoc.



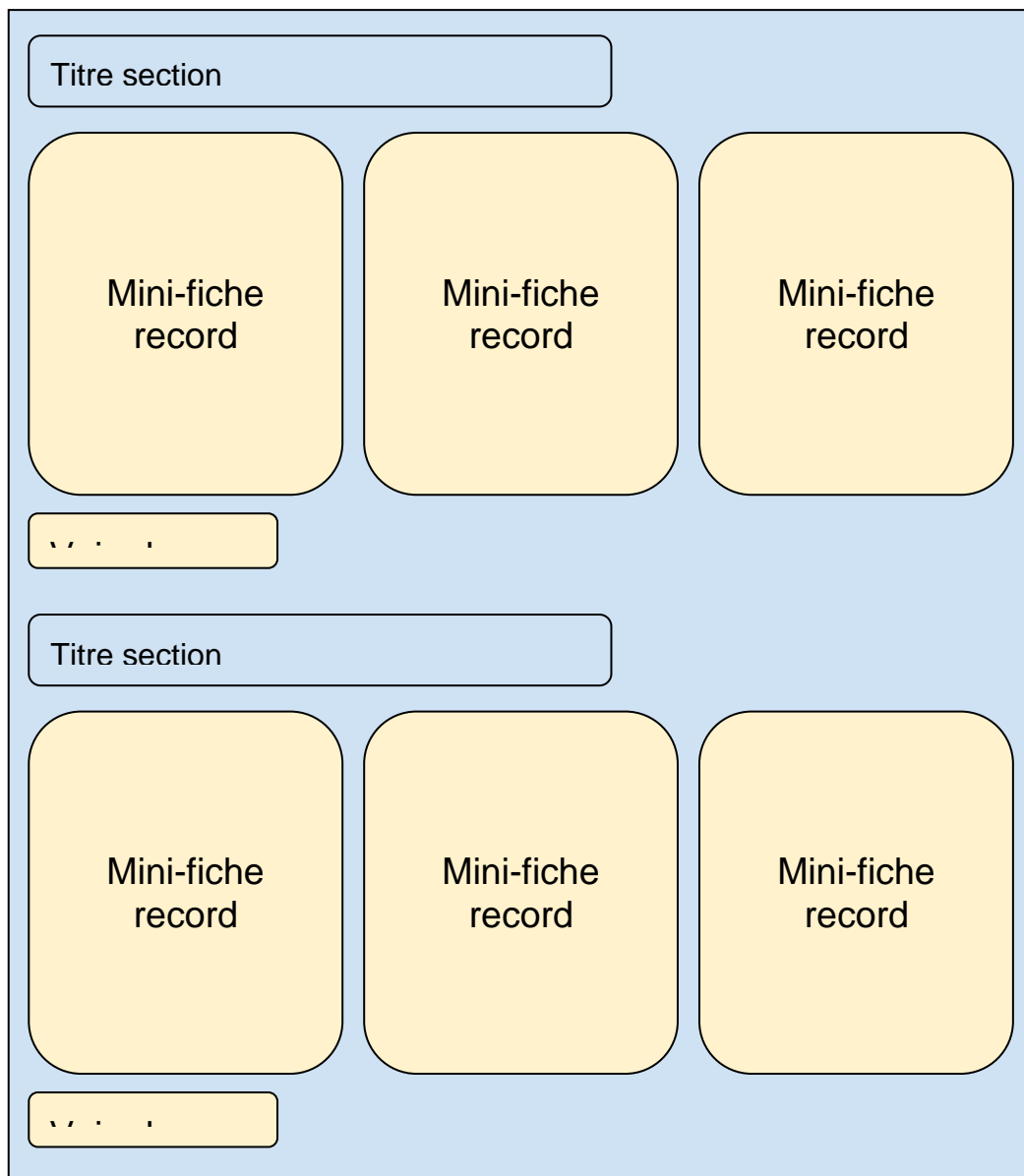


Figure 10. Contenu principal avec mini-fiches d'exemple.

S'il ne serait pas possible, la section du contenu principal devrait présenter une image et, superposée, une section contenant une brève description du projet, la boîte de recherche et le bouton pour la recherche avancée, Tel que présenté dans la [Figure 11](#), qui se trouveront ici, et non pas dans l'en-tête comme décrit avant.

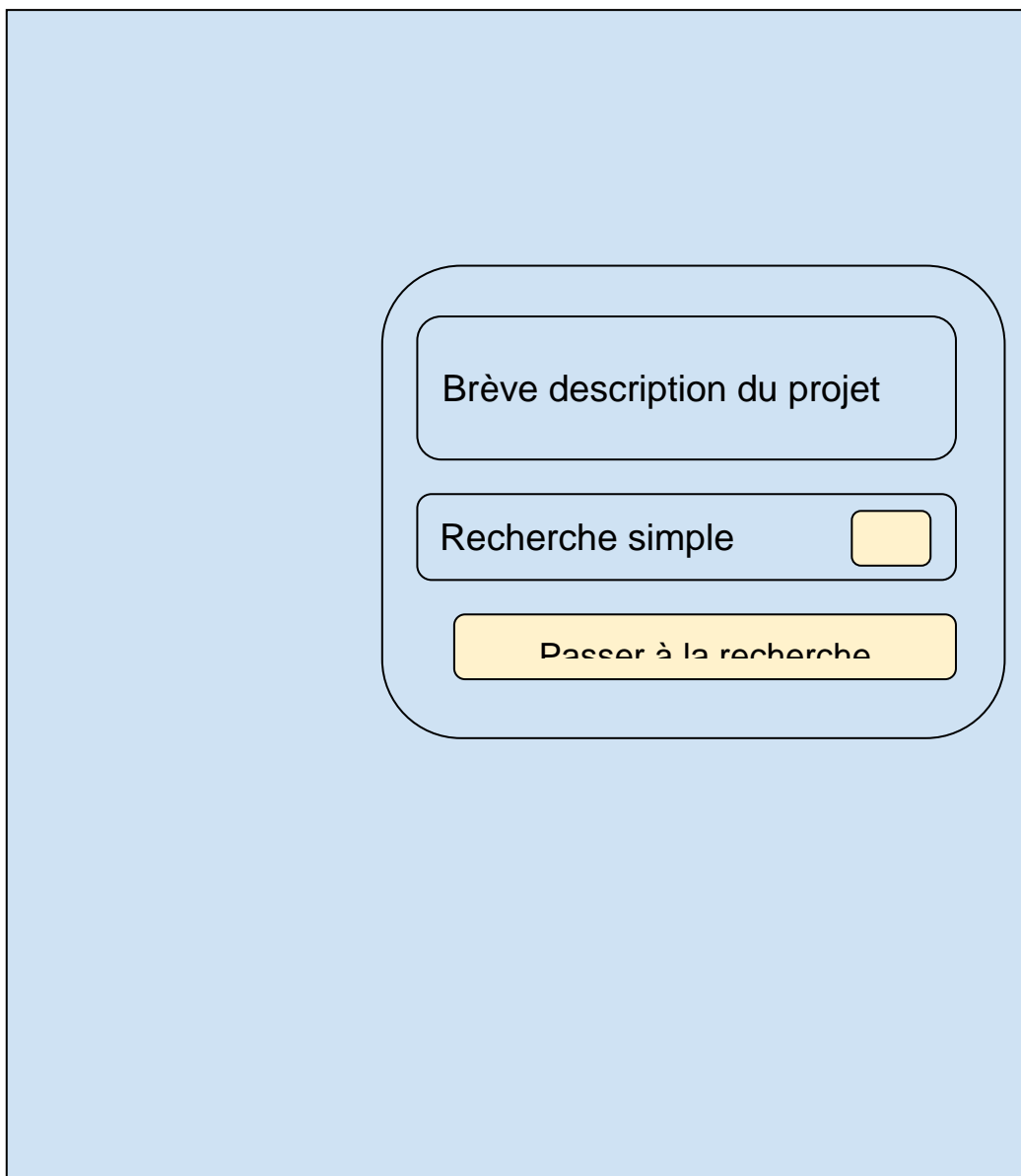


Figure 11. Contenu principal avec boîte de recherche.

#### 8.1.4 Navigation

Cette section est dédiée à la navigation des contenus sans entrer des mots-clés pour la recherche. Cliquant sur les boutons (dont les valeurs sont définies dans la [section 6.3](#)) il est possible de filtrer les records présentes dans l'agrégateur et de voir dans la section du contenu principal quelques mini-fiches ordonnées sur la base des critères définis dans la [section 6.7](#), tandis que la barre de navigation sera remplacée par la barre des filtres.

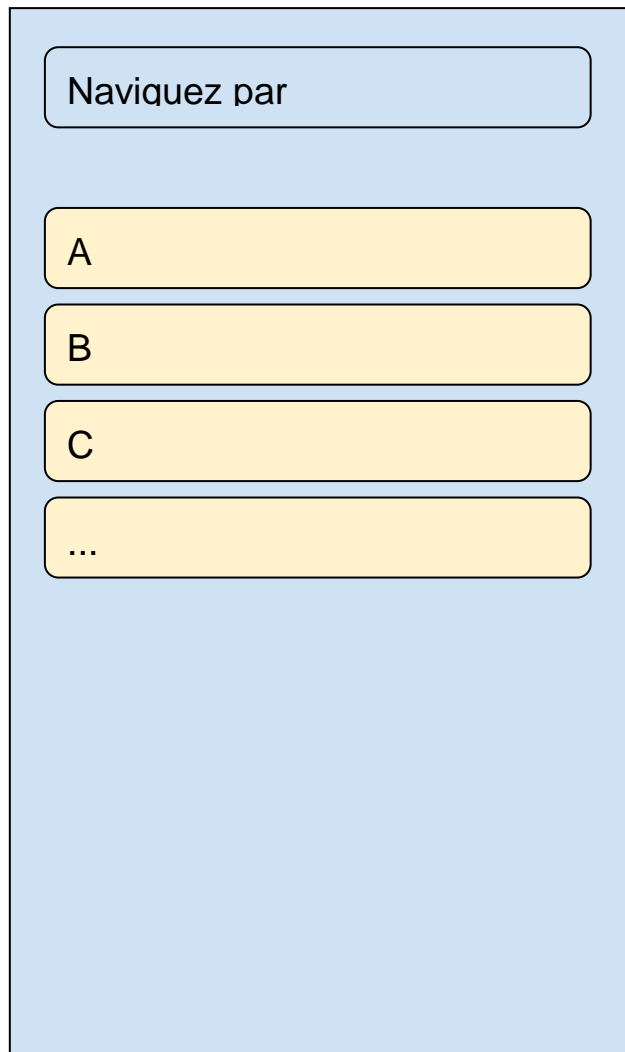


Figure 12. Barre de navigation.

## 8.2 Mini-fiche record

La mini-fiche d'un record doit présenter dans un espace réduit une ongle de pouce (extraite de l'image d'origine de la ressource, si présente) et les métadonnées de ce record. Il est approprié pour être présentée en modalité mosaïque ensemble à d'autres mini-fiches dans la même page.

Au cas où il n'y a pas une image, la mini-fiche présentera une image de substitution qui dépend du type de ressource (`originalResource/dc:type`) auquel le record se réfère.

Les métadonnées principales qui ont été identifiées pour être présentées dans la mini-fiche sont détaillées dans la [section 6.5](#).

La mini-fiche est un élément cliquable, et si l'utilisateur clique il va passer à la fiche complète de deuxième niveau, qui est présentée en remplissant tout l'espace à disposition dans la section contenu principal.

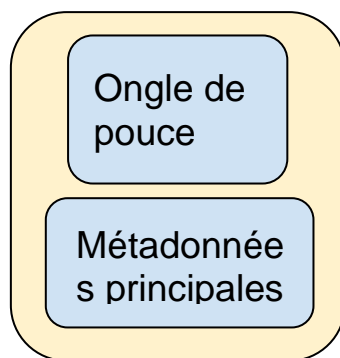


Figure 13. Mini-fiche de premier niveau.

### 8.3 Fiche complète de deuxième niveau

Cette fiche est présentée dans l'espace du contenu principal ou dans l'espace des résultats de recherche ([Figure 14](#)).

L'image principale est indiquée entre les métadonnées comme image d'origine (digitalCopy/sourceURI). S'il y a plusieurs d'images, l'agrégateur devra choisir la première et, si possible, il devrait être la même image présente sur la mini-fiche. D'autres images supplémentaires peuvent être présentes pour la même ressource. Toutes les images sont cliquables pour être visualisées en pleine résolution.

La fiche présente aussi le titre de la ressource et sa description. En bas on trouve un tableau avec toutes les métadonnées contenues dans l'ensemble de métadonnées communes partagée. Entre les métadonnées il y a aussi un lien à la ressource d'origine sur le site du partenaire, qui devra évidemment être cliquable.

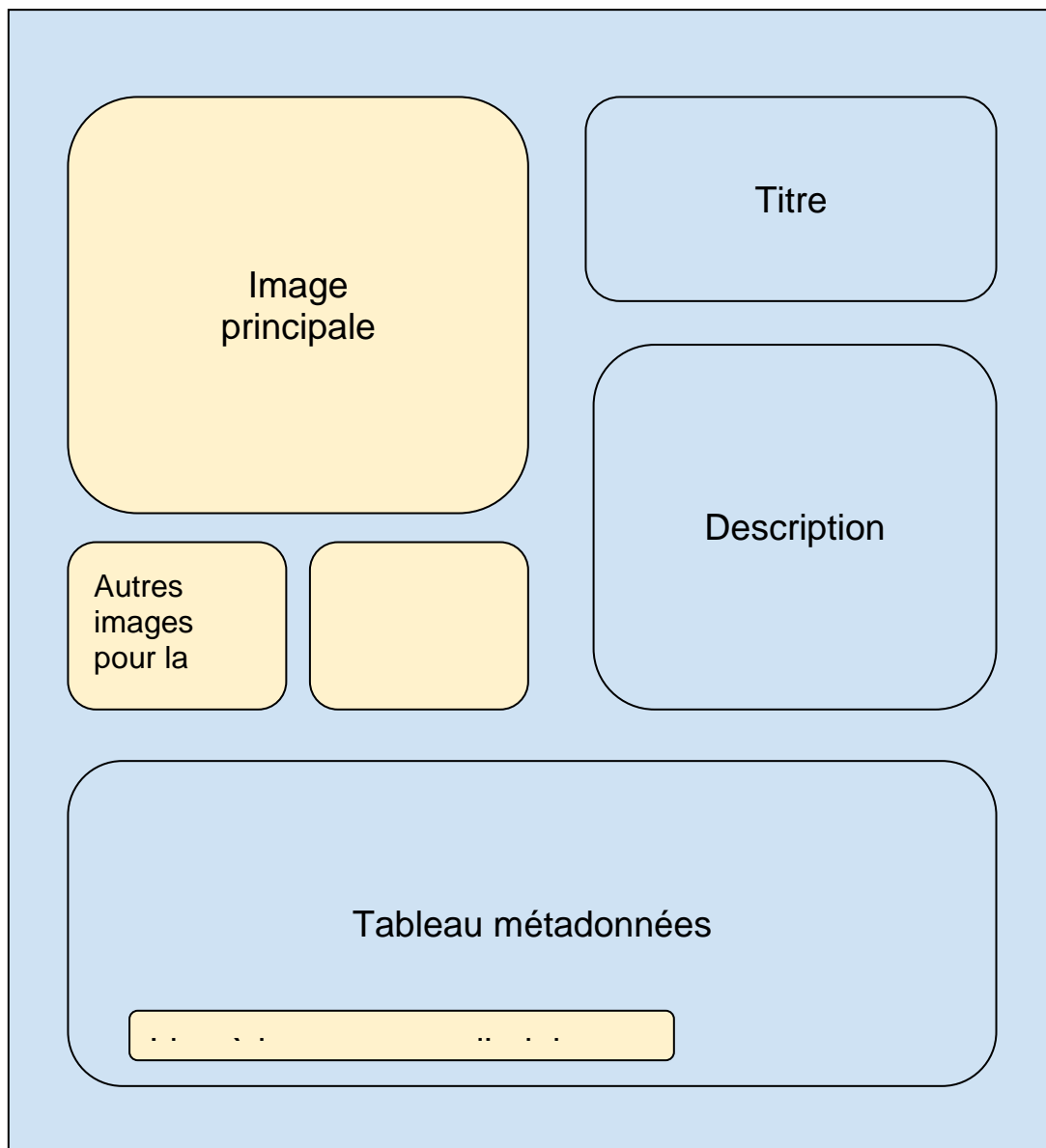
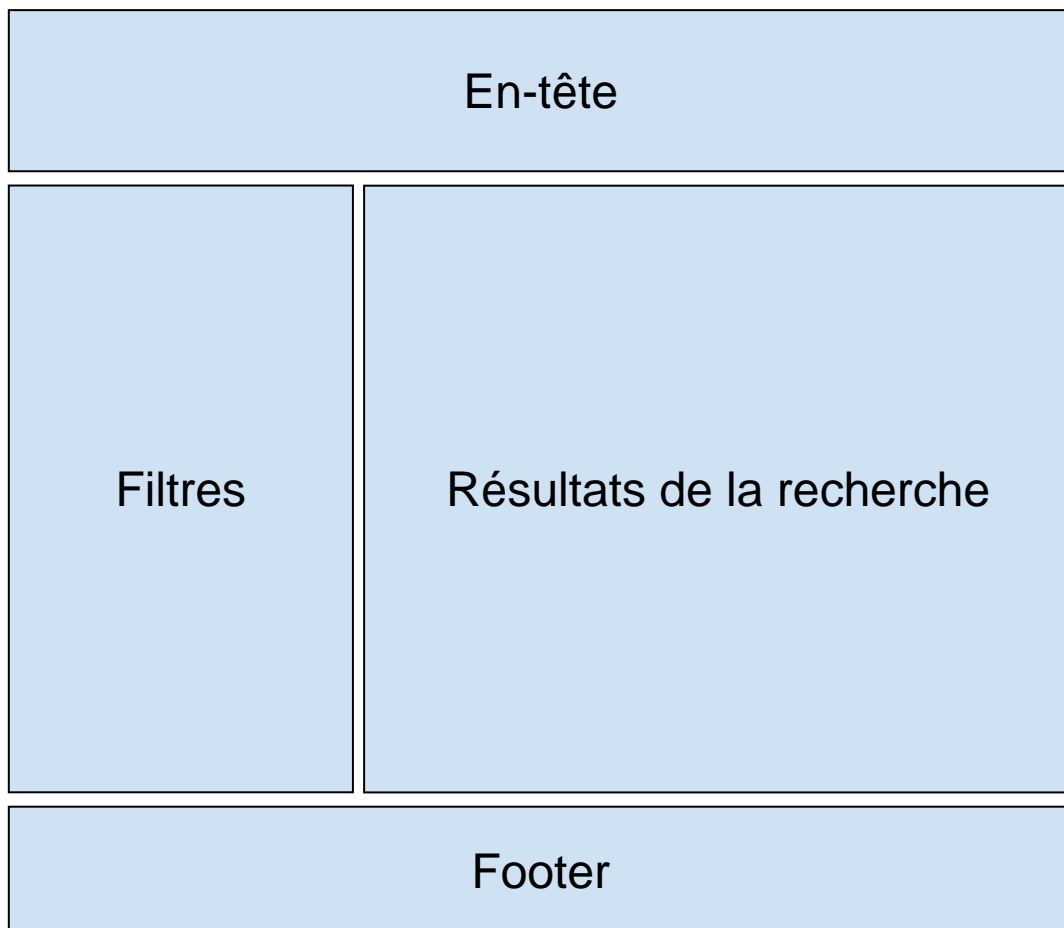


Figure 14. Fiche complète de deuxième niveau.

## 8.4 Présentation des résultats

Après que l'utilisateur envoie la requête pour une recherche, les résultats sont présentés dans la section du contenu principal et la barre de navigation est remplacée par la barre de filtres, tel qu'on voit dans la [Figure 15](#).



*Figure 15. Présentation des résultats à front d'une recherche.*

#### 8.4.1 Résultats de la recherche

La section qui montre les résultats sera organisée comme on peut voir dans la [Figure 16](#) : les mini-fiches (pas nécessairement toutes de la même taille) seront placées en mosaïque. Il y aura aussi des boutons pour la pagination (un au sommet de la page et un en bas), le sélecteur pour choisir le critère d'ordre et le sélecteur du nombre maximal de fiches présentées par page.

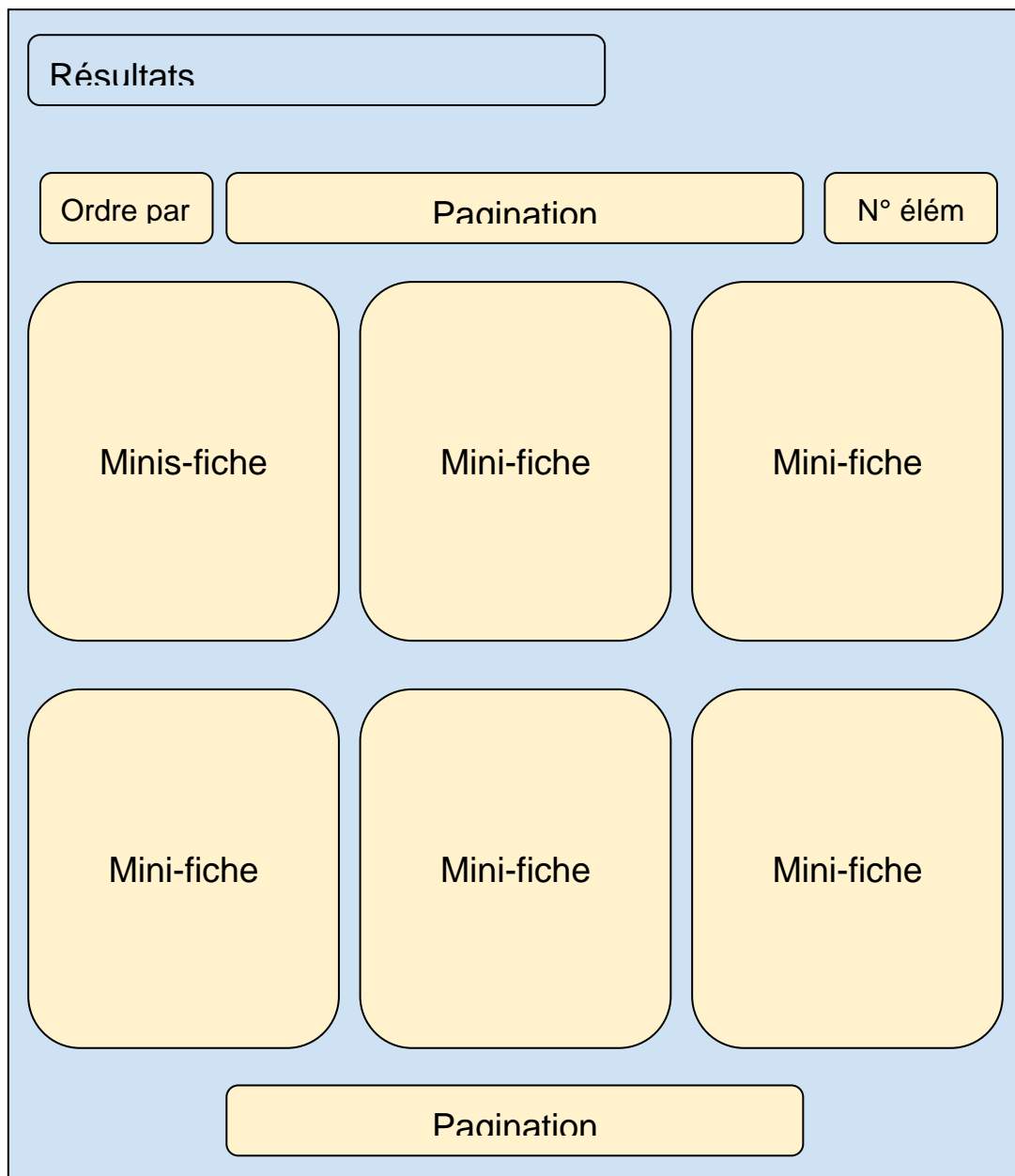


Figure 16. Section des résultats de la recherche.

#### 8.4.2 Filtres

La barre des filtres sera utilisée pour affiner les résultats d'une recherche. Il est nécessaire d'avoir tous les filtres décrits dans la [section 6.5](#), éventuellement en éliminant ceux pour lesquels il n'y a pas de valeurs dans les résultats de la recherche. La [figure 17](#) présente une possible organisation de la barre, avec les filtres l'un sous l'autre, rétractables ou dilatables avec un bouton. Pour chaque filtre, l'interface présente toutes les valeurs présentes dans les résultats de la recherche et le nombre d'éléments présents est cliquable, pour donner à l'utilisateur l'information sur combien de fiches seront visualisées.

The image shows a vertical filter panel with a light blue background. At the top is a rounded rectangle labeled 'Filtres'. Below it are two larger rounded rectangles. The first is titled 'Thèmes du projet' and contains a toggle switch, two rows of buttons labeled 'Thème1' and 'Thème2' next to yellow buttons labeled 'N°', and an ellipsis. The second is titled 'Auteurs' and contains a toggle switch, two rows of buttons labeled 'Auteur1' and 'Auteur2' next to yellow buttons labeled 'N°', and an ellipsis. At the bottom is a single row with an ellipsis and a toggle switch.

Figure 17. Filtres pour affiner les résultats.

## 8.5 Recherche avancée

Le bouton pour la recherche avancée présent dans l'en-tête (ou dans la section du contenu principal) montre dans la section du contenu principal les barres pour la recherche avancée ([Figure 18](#)). Il y aura une zone de saisie pour chacun des champs de recherche énumérés dans la [section 6.2](#), et une en plus, en haut, qui sera de nouveau la barre de recherche simple.

Les résultats seront présentés dans la même section de la page, tandis que dans la section à la gauche apparaîtra la barre des filtres en remplacement de celle de navigation, tel que pour la recherche simple.



Recherche avancée

Tous les champs

Auteur

Titre

....

....

Cherche

Figure 18. Recherche avancée.

## 9 Analyses statistiques

Le site de l'agrégateur devra collecter une série de statistiques, notamment les suivantes :

- nombre d'accès
- typologie de ressources les plus consultées
- mots-clés recherchés
- approfondissement de la recherche effectuée (si l'utilisateur s'arrête seulement au premier niveau ou s'il clique sur le lien de la fiche sur le site du partenaire)
- téléchargement des données
- langue dans laquelle les utilisateurs effectuent les recherches
- pays d'où vient la recherche
- type d'appareil utilisé par l'utilisateur.

L'agrégateur devrait aussi fournir la tendance de ces indicateurs dans le temps. En outre, il devrait mettre à disposition des partenaires une page de consultation de ces statistiques. Alternativement, il pourrait envoyer, périodiquement, à chaque partenaire un rapport avec ces statistiques.

## 10 Maintien et mise à jour des ressources

L'agrégateur effectuera la mise à jour des ressources depuis tous les fournisseurs de données tous les trois mois. La mise à jour pour les fournisseurs de données implique la transmission de l'entier ensemble de métadonnées des ressources (seulement pour celles qui ont été mises à jour), parce que la transmission des seuls éléments mises à jour nécessiterait le développement d'autres systèmes de contrôle sur les variations.

Les partenaires qui adopteront la modalité d'importation à travers de URL préférablement devront utiliser des permalink. Ils s'engageront à notifier à l'agrégateur éventuels modifications dans les URL contenant les fichiers statiques d'échange s'ils ne peuvent pas fournir des permalink. Le système d'importation des métadonnées de l'agrégateur devra donc être configurable pour la gestion d'éventuels changes dans les URL.

Le maintien du système d'importation des métadonnées et de l'agrégateur sera assuré par la Ville de Nice pour la durée du projet et ensuite pour un temps indéfini.

Le maintien des ressources sera assuré de la part des partenaires, qui s'engageront à diffuser et promouvoir l'agrégateur à autres institutions, à fin de qu'elles aussi puissent le connaître et participer, en mettant à disposition leurs ressources.

# 11 Préservation et protection du contenu numérique

L'agrégateur devrait prévoir une interface d'administration pour la gestion du contenu numérique et pour le suivi de l'état. L'interface pour la gestion du contenu numérique devrait permettre l'accès aux données des partenaires, leur mise à jour et éventuelles modifications. L'interface pour le suivi de l'état devrait permettre le suivi de l'état physique de l'appareil qui héberge le logiciel de l'agrégateur, et aussi l'état des supports physiques, comme décrit ici ensuite.

Les mécanismes décrites ici ensuite ont été élaborés sur la base de celui de la Library of Congress<sup>14</sup>.

## 11.1 Préservation

Pour préservation des documents numériques on entend la conservation et maintien des données dans le temps. L'agrégateur devrait prévoir un système de suivi de l'état des supports physiques utilisés pour la mémorisation des contenu et prévoir la migration à nouveau supports dans le cas d'obsolescence ou dégradation des supports utilisés. En cas de fonctionnement normal, on pourrait prévoir un contrôle générale annuel des supports. En cas d'anomalies, l'agrégateur devrait les notifier rapidement, à travers appropriés systèmes de notification, aux services autorisés.

Similairement, aussi les dispositifs qui hébergent les systèmes de gestion des partenaires devront prévoir des mécanismes de suivi de l'état des supports physiques.

## 11.2 Protection

Pour protection des documents numériques on entend le processus qui empêche la modification des documents par sujets non autorisés ou une éventuelle perte accidentelle.

Pour empêcher la modification non autorisée des documents, l'agrégateur devra implémenter des mécanismes de contrôle des accès à la gestion des contenus numériques. Seulement les utilisateurs autorisés devraient pouvoir accéder au système de gestion des contenus des autres partenaires. A ce propos, on doit prévoir au minimum un utilisateur administrateur du système qui, sur la base des nécessités des différents projets, peut créer ou éliminer utilisateurs autorisés à l'accès d'une ou plusieurs parties du système de gestion.

Pour prévenir la perte accidentelle des documents, l'agrégateur devra implémenter des politiques de sauvegarde périodique des données et politiques de retenue.

---

<sup>14</sup> <https://www.loc.gov/programs/digital-collections-management/inventory-and-custody/responsibility-for-digital-collections-content-and-storage-systems/>

## 12 Annexe A - Le modèle de Europeana

Le modèle d'Europeana est basé sur la publication de ressources numériques qui appartiennent à collections qui incluent (mais ne sont pas limités à) les exigences suivantes<sup>15</sup> :

- 1) les métadonnées sur Europeana, relatives à une ressources doivent envoyer à une ressource numérique librement accessible sur le Web
- 2) les utilisateurs doivent pouvoir accéder à la ressource sans authentification
- 3) préférer la qualité des ressources plutôt que la quantité
- 4) considérer plus prioritaire l'amélioration des ressources existantes plutôt qu'en ajouter des nouvelles
- 5) évaluer si les nouveaux contenus respectent les thèmes proposés par Europeana.

La politique d'Europeana pour la priorisation des ressources se base sur l'analyse statistique des accès au portail (par exemple, objets plus visités), pour définir le type d'audience. Avant de pouvoir envoyer une collection de ressources à Europeana, le fournisseur de données doit signer l'Europeana Data Exchange Agreement (DEA)<sup>16</sup>, qui établit les termes sur la base desquels Europeana peut utiliser les prévisualisations et les métadonnées descriptives des fournisseurs de données.

Europeana définit deux typologies de lignes directrices pour la publication de ressources : lignes directrices pour le contenu<sup>17</sup> et lignes directrices pour les métadonnées<sup>18</sup>. Les lignes directrices pour le contenu prévoient 4 niveaux de publication (tiers)<sup>19</sup>, qui dépendent de l'objectif du fournisseur de ressources. Les niveaux de publication sont :

- 1) Niveau 1. Europeana comme moteur de recherche : "Je veux que les utilisateurs trouvent mes collections"
- 2) Niveau 2. Europeana comme vitrine : "Je veux que les utilisateurs trouvent et visualisent mes collectio sur Europeana"
- 3) Niveau 3. Europeana comme plateforme de distribution : "Je veux que les utilisateurs trouvent, visualisent et peuvent utiliser mes collections, mais sans but lucratif"
- 4) Niveau 4. Europeana comme plateforme pour une réutilisation libre : "Je veux que les utilisateurs trouvent, visualisent et peuvent utiliser mes collections de n'importe quelle façon".

Le tableau suivant présente les exigences pour la numérisation des ressources dans les niveaux différents.

---

<sup>15</sup> <https://pro.europeana.eu/post/europeana-content-strategy>

<sup>16</sup> <https://pro.europeana.eu/page/the-data-exchange-agreement>

<sup>17</sup> [https://pro.europeana.eu/files/Europeana\\_Professional/Publications/Publishing\\_Framework/Europeana\\_publishing\\_framework\\_content.pdf](https://pro.europeana.eu/files/Europeana_Professional/Publications/Publishing_Framework/Europeana_publishing_framework_content.pdf)

<sup>18</sup> [https://pro.europeana.eu/files/Europeana\\_Professional/Publications/Publishing\\_Framework/Europeana\\_publishing\\_framework\\_metadata\\_v-0-8.pdf](https://pro.europeana.eu/files/Europeana_Professional/Publications/Publishing_Framework/Europeana_publishing_framework_metadata_v-0-8.pdf)

<sup>19</sup> <https://pro.europeana.eu/post/publishing-framework>

TIER	MEDIA				
	IMAGE	SOUND	VIDEO	TEXT	3D
1	- edm:type is IMAGE - Thumbnail is available* - An image is available** with resolution >= 0.1mp	- edm:type is SOUND - working edm:isShownAt	- edm:type is VIDEO - An image is available with resolution >= 0.1mp as edm:object - Thumbnail is available - A video is available** or working edm:isShownAt	- edm:type is TEXT - An image is available** with 0.1mpx (ie. also means that tech metadata is available) OR working edm:isShownAt	- edm:type is 3D - A 3D resource is available (with tech metadata)
2	- edm:type is IMAGE - Thumbnail is available* - An image is available** with resolution >= 0.42mpx (~800x533)	- edm:type is SOUND - A sound is available** OR embeddable media as edm:isShownBy	- edm:type is VIDEO - An image is available with resolution >= 0.1mp as edm:object - Thumbnail is available - A video is available** with vertical resolution equal or higher than 480 pixels OR embeddable media as edm:isShownBy	- edm:type is TEXT - A text resource is available** with mimetype "application/pdf" OR image is available** with resolution >= 0.42mpx	- edm:type is 3D - A 3D resource is available**
3	- edm:type is IMAGE - Thumbnail is available* - An image is available** with resolution >= 0.95mp (1200x800) - associated image has either open or restricted license	- edm:type is SOUND - A sound is available** OR embeddable media as edm:isShownBy - associated sound has either open or restricted license	- edm:type is VIDEO - An image is available with resolution >= 0.1mp as edm:object - Thumbnail is available - A video is available** with vertical resolution equal or higher than 480 pixels OR embeddable media as edm:isShownBy - associated video has either open or restricted license	- edm:type is TEXT - A text resource** is available with mimetype "application/pdf" OR image is available** with resolution >= 0.95mpx (1200x800) - associated image or text resource has either open or restricted license	- edm:type is 3D - A 3D resource is available** - associated 3D resource has either open or restricted license
4	- edm:type is IMAGE - Thumbnail is available* - An image is available** with resolution >= 0.95mp (1200x800) - associated image has only open license	- edm:type is SOUND - A sound is available** OR embeddable media as edm:isShownBy - associated sound has only open license	- edm:type is VIDEO - An image is available with resolution >= 0.1mp as edm:object - Thumbnail is available* - A video is available** with vertical resolution equal or higher than 480 pixels OR embeddable media edm:isShownBy - associated video has only open license	- edm:type is TEXT - A text resource is available with mimetype "application/pdf" OR image with resolution >= 0.95mp (1200x800)** - associated image or text resource has only open license	- edm:type is 3D - A 3D resource is available** - associated 3D resource has only open license

Figure 19.

En ce qui concerne les lignes directrices pour les métadonnées, trois niveaux (tiers) sont définis, avec des exigences différentes qui dépendent du pourcentage de métadonnées remplis pour la description des ressources. Les niveaux sont :

- 1) A - au moins 25% des métadonnées pertinentes pour une ressource a au au moins une balise correspondante à la langue ;
- 2) B - au moins 50% des métadonnées pertinentes pour une ressource a au au moins une balise correspondante à la langue ;
- 3) C - au moins 75% des métadonnées pertinentes pour une ressource a au au moins une balise correspondante à la langue ;

Les métadonnées fournies pour chaque ressource doivent être représentées en utilisant l'Europeana Data Model (EDM)<sup>20</sup>, qui prévoit quelques champs obligatoires et d'autres facultatifs. Chaque métadonnée doit être codifiée selon le codage UTF-8. Les champs de l'EDM pour la description d'une ressource culturelle sont :

<sup>20</sup> <https://pro.europeana.eu/page/edm-documentation>

Properties for edm:ProvidedCHO	
+ dc:contributor	dcterms:isReplacedBy
dc:coverage	dcterms:isRequiredBy
+ dc:creator	+ dcterms:issued
+ dc:date	dcterms:isVersionOf
→ dc:description	dcterms:medium
dc:format	dcterms:provenance
+ dc:identifier	dcterms:references
✓ dc:language (if edm:type = TEXT)	dcterms:replaces
+ dc:publisher	dcterms:requires
dc:relation	○ dcterms:spatial
dc:rights	dcterms:tableOfContents
+ dc:source	○ dcterms:temporal
○ dc:subject	edm:currentLocation
→ dc:title	edm:hasMet
○ dc:type	edm:hasType
+ dcterms:alternative	edm:incorporates
dcterms:conformsTo	edm:isDerivativeOf
+ dcterms:created	+ edm:isNextInSequence
dcterms:extent	edm:isRelatedTo
dcterms:hasFormat	edm:isRepresentationOf
dcterms:hasPart	edm:isSimilarTo
dcterms:hasVersion	edm:isSuccessorOf
dcterms:isFormatOf	edm:realizes
+ dcterms:isPartOf	✓ edm:type
dcterms:isReferencedBy	owl:sameAs

Figure 20.

✓ = Mandatory property

→ Blue = at least one of the blue properties should be present (and can be used alongside each other)

○ Red = at least one of the red properties should be present (and can be used alongside each other)

◆ Green = at least one of the green properties should be present (and can be used alongside each other)

+ = recommended property

Detailed characteristics of all properties are given in the full tables following this summary.

Figure 21.

Si on veut inclure des ressources dans une ou plusieurs collections thématiques, il est nécessaire de définir des mots-clés spécifiques dans les champs *dc:type* et *dc:subject*. Pour chaque thème, Europeana définit un vocabulaire contrôlé de mots-clés. Par exemple, pour le thème archéologie, les mots-clés sont les suivants : *archaeology*, *archaeological sites*, *field archaeology*, *buildings archaeology*, *excavations (sites)*, *archaeological objects*.

Europeana définit trois critères pour évaluer la qualité des métadonnées :

- 1) langue, qui permet de définir la langue de chaque métadonnée, à travers la balise *xml:lang* ;
- 2) éléments génériques, qui définissent les métadonnées non obligatoires, mais qui permettent d'appliquer la ressource dans scénarios spécifiques ;

- 3) classes contextuelles, qui définissent le contexte de la ressource, tel que par exemple une description des lieux ou des dates relatives à la ressource. Pour les classes contextuelles, Europeana recommande l'utilisation des références aux Linked Data.

Du point de vue de la licence, tandis que les métadonnées associées aux ressources peuvent être utilisées sans aucune limitation, les ressources numériques peuvent être mises à disposition selon une des 14 licences spécifiées par la politique de Europeana<sup>21</sup>.

## 12.1 Modalités de recherche et présentation des résultats

Comme décrit dans le [Tableau 5](#), du document 3.1.5 Inventaire des Sites Web, Europeana permet la navigation des contenus selon les modalités suivantes : couleurs, sources, personnes, sujets, galeries, périodes de temps, expositions. Les modalités d'exploration sont présentées à travers un menu déroulant ([Figure 22](#)).



Figure 22 Modalité d'exploration des contenus dans Europeana.

En sélectionnant une des options du menu déroulant on peut accéder à une autre possibilité de choix. Par exemple, en sélectionnant l'option Périodes de temps, la suivante liste cliquable est visualisée ([Figure 23](#)).



<sup>21</sup> <https://pro.europeana.eu/page/available-rights-statements>



Figure 23. Liste cliquable des différentes périodes de temps dans Europeana.

En choisissant un des objets de la liste (par exemple 1700-1799), la liste des objets relatifs à cette période s'ouvre (Figure 24).

Figure 24. Liste des objets relatifs à la période de temps 1700-1799 dans Europeana.

En plus des simples explorations, Europeana met à disposition aussi les Expositions, qui sont des expositions pré-sélectionnées sur un sujet spécifique. On peut accéder aux différentes expositions à travers un autre menu déroulant (Figure 25).

Figure 25 Menu déroulant pour choisir les expositions dans Europeana.

Un exemple d'exposition est Heritage at Risk<sup>22</sup>, qui analyse le patrimoine culturel en danger. Cette exposition est divisée en différentes sous-thèmes, et chacun est navigable (Figure 26).

<sup>22</sup> <https://www.europeana.eu/portal/en/exhibitions/heritage-at-risk>

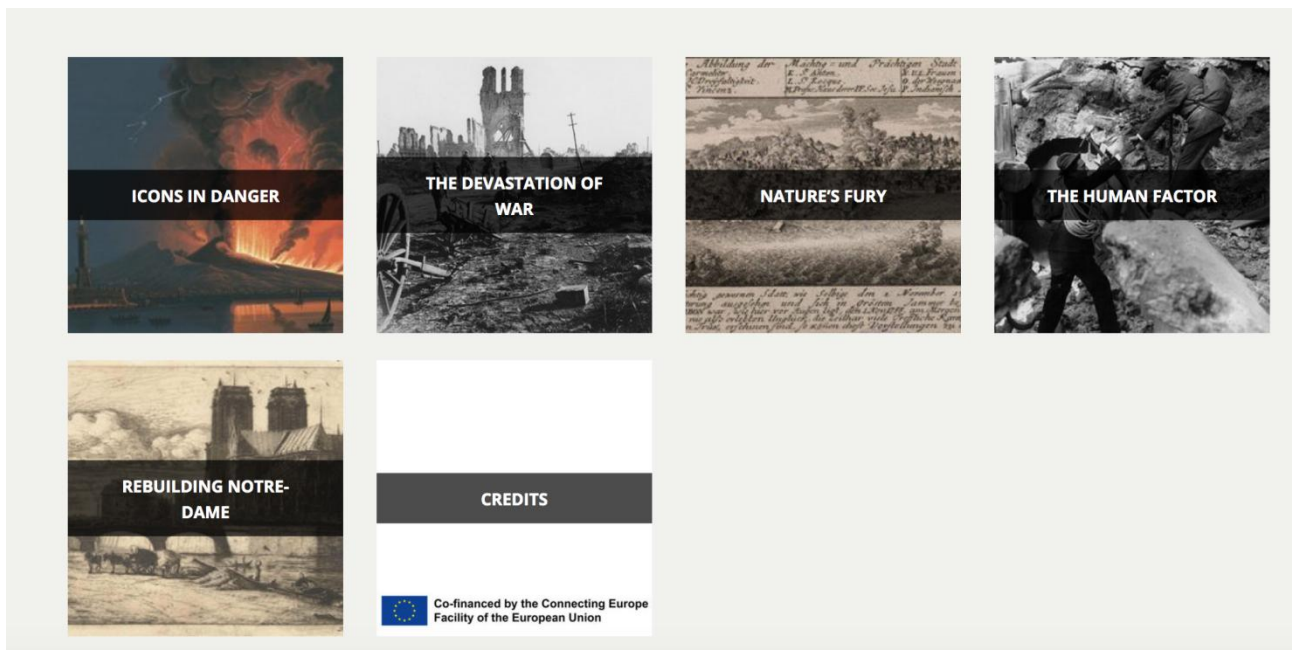


Figure 26. Sous-thèmes de l'exposition *Heritage at Risk* dans *Europeana*.

La [Figure 27](#) présente un exemple de modalité de présentation des résultats de premier niveau.

### Self-portrait (1887) | Gogh, Vincent van

Vincent moved to Paris in 1886, after hearing from his brother Theo about the new, colourful style of French painting. Wasting no time, he tried it out in several self-portraits. He did this mostly to avoid having to pay for a model. Using rhythmic brushstrokes in striking...

Pittura

**Vedi tutto Rijksmuseum**

Immagine



Figure 27. Présentation des résultats de premier niveau dans *Europeana*.