

Digital Libraries on the Grid to preserve cultural Heritage

A use case:

Federico De Roberto works

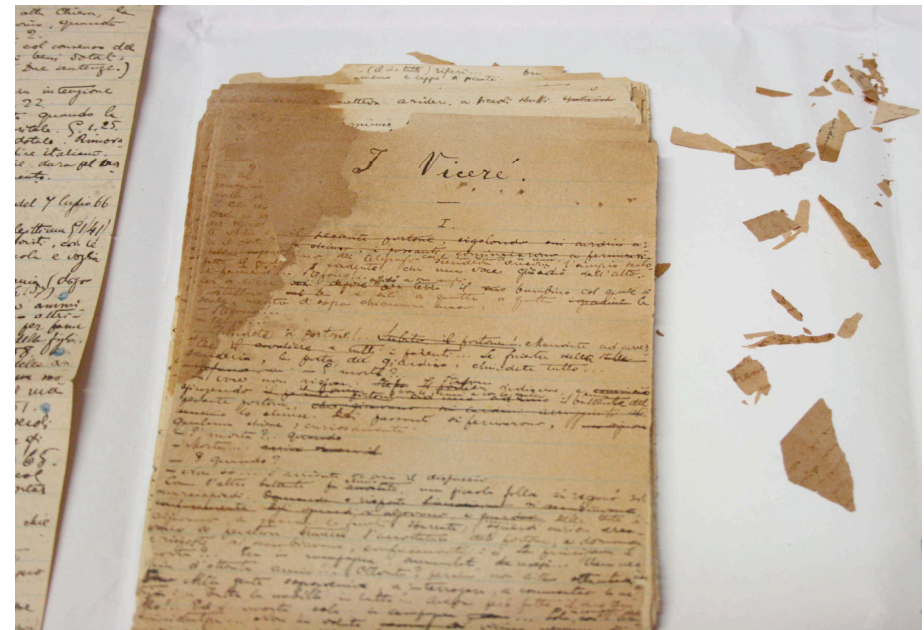
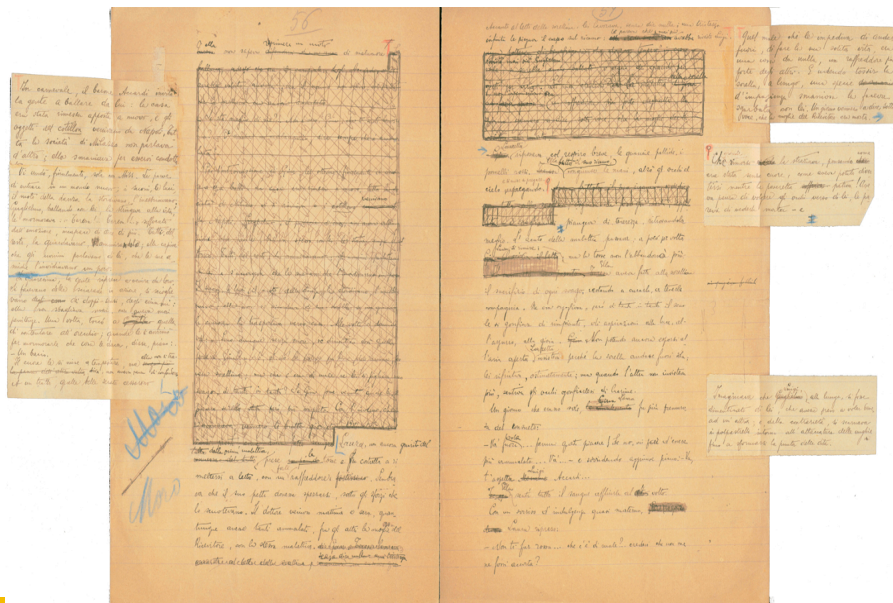
*Antonio Calanducci¹, Jorge Sevilla¹, Monica Saso⁴,
Alessandro De Filippo⁴, Stefania Iannizzotto⁴, Roberto
Barbera^{1,2,3}, Giuseppe Andronico^{1,2}*

¹INFN Catania, ²COMETA Consortium, ³Univ of CT

⁴Faculty of Letters and Philosophy (Univ of Catania)

OGF25/User Forum 4 - Catania, 04/03/09

- De Roberto, an Italian writer of the XIX/XX century, born in Naples, but spending his life in Catania, has left to the humanistic community numerous works
- Those are made up of valuable and hard-to-manage pieces: manuscripts, typescripts, drafts with handwritten corrections, magazines, cuts, sketches, photos, etc.



- **Digitalization of manuscripts, typescripts, printed works**
 - TIFF Files, one per page, 600 dpi, about 100MB for A3
 - High resolution scans for in-depth examination
 - Multipage PDF, one per work, 300 dpi, varying file sizes 40-400MB
 - Overall examination of works
 - 8000 scans, 3 Terabyte of disk space
 - Different physical formats, A3/A4/custom size
- **Embedded Metadata**
 - TIFF with embedded metadata to provide scan physical features and information about the content
 - ImageWidth, ImageHeight, XResolution, FileSize, CreationDate, ModifyDate
 - Description, Keywords, CaptionWriter, Title, Author, Copyright Status, Copyright Notice
 - Added with Photoshop after the digitalization phase (Adobe XMP format)

- Make those works **accessible** to the humanistic research community
 - **Always on-line**: 24 x 365
 - **Available from everywhere**
 - Simple **easy-to-use interface** for non-expert people
- **Immediately find the desired document**
 - Document organization according the physical and semantic metadata
 - ☐ By type/category
 - ☐ Dynamic filtering of search result sets according the selection of one or more document metadata
- **Long-term preservation (digital preservation)**
 - Multiple copies (**replicas**) spread in different geographical sites
 - Reliability of storage systems and replica redundancy to achieve secure preservation

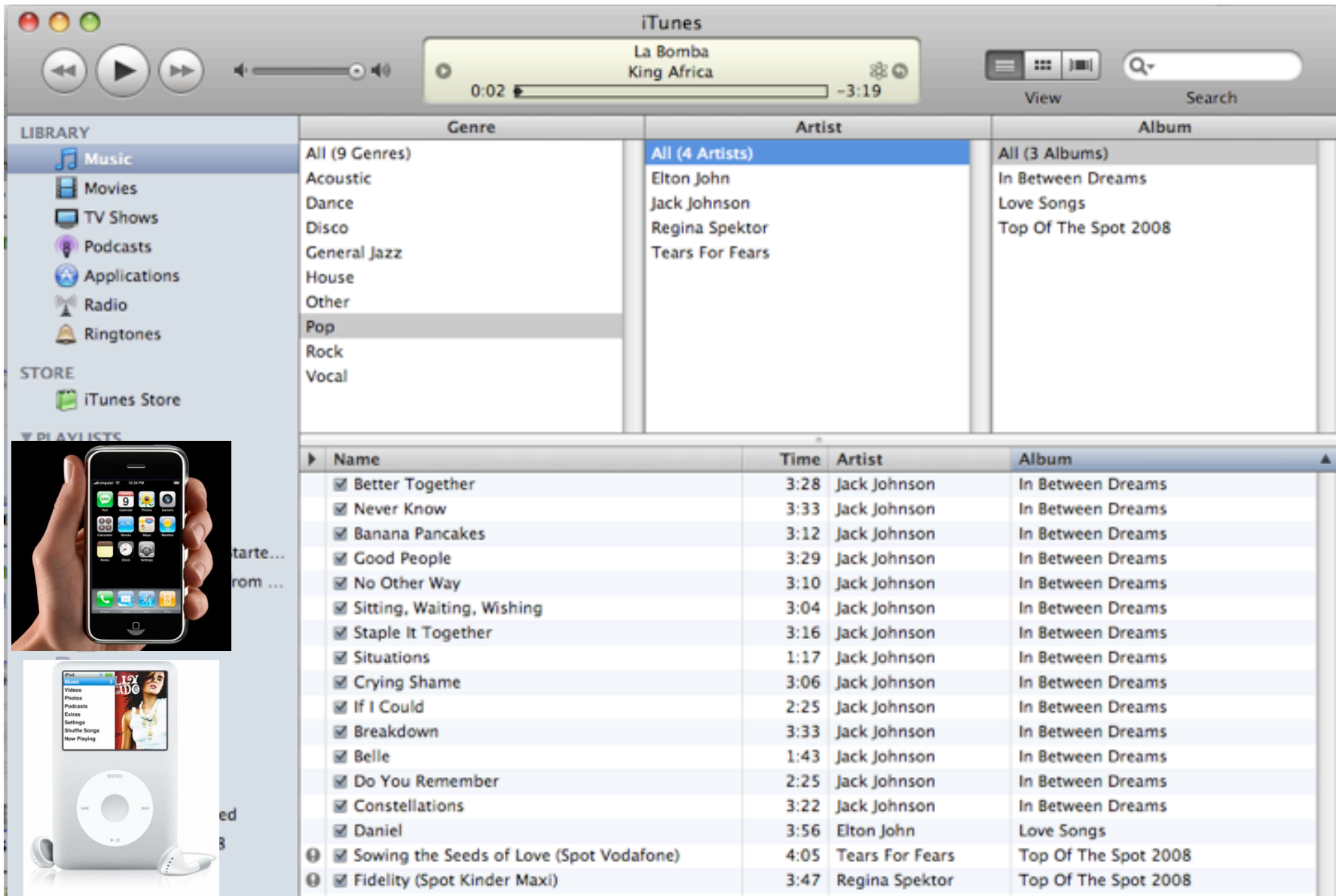
- **Our suggestions:**

- to store the 8000 scans of De Roberto Heritage
----> **Grid Storage Elements**
- to enable an ubiquitous and 24/24h access to scientists
----> **Web Application**
- document organization for a fast search
----> **Metadata Services**
- long-term digital preservation of data
----> redundancy through **Replicas** of files on several Storage Elements
- **easy-to-use** system for searches, organization, upload and download of digitalized documents on the Grid



- INFN-developed tool totally gLite based
- It allows to **store, organize, search** and **retrieve** digital assets on a Grid environment with an intuitive front-end
- What we mean by **Digital Assets**:





The screenshot shows the iTunes application window. The top bar displays the current song "La Bomba" by "King Africa" with a progress bar at 0:02 / 3:19. The left sidebar shows the "LIBRARY" section with categories like Music, Movies, TV Shows, Podcasts, Applications, Radio, and Ringtones. The main area is divided into three columns: Genre, Artist, and Album. The "Genre" column shows "All (9 Genres)" and "Pop". The "Artist" column shows "All (4 Artists)" and lists Elton John, Jack Johnson, Regina Spektor, and Tears For Fears. The "Album" column shows "All (3 Albums)" and lists "In Between Dreams", "Love Songs", and "Top Of The Spot 2008". Below these columns is a table of songs with columns for Name, Time, Artist, and Album.

| Name | Time | Artist | Album |
|--|------|-----------------|----------------------|
| <input checked="" type="checkbox"/> Better Together | 3:28 | Jack Johnson | In Between Dreams |
| <input checked="" type="checkbox"/> Never Know | 3:33 | Jack Johnson | In Between Dreams |
| <input checked="" type="checkbox"/> Banana Pancakes | 3:12 | Jack Johnson | In Between Dreams |
| <input checked="" type="checkbox"/> Good People | 3:29 | Jack Johnson | In Between Dreams |
| <input checked="" type="checkbox"/> No Other Way | 3:10 | Jack Johnson | In Between Dreams |
| <input checked="" type="checkbox"/> Sitting, Waiting, Wishing | 3:04 | Jack Johnson | In Between Dreams |
| <input checked="" type="checkbox"/> Staple It Together | 3:16 | Jack Johnson | In Between Dreams |
| <input checked="" type="checkbox"/> Situations | 1:17 | Jack Johnson | In Between Dreams |
| <input checked="" type="checkbox"/> Crying Shame | 3:06 | Jack Johnson | In Between Dreams |
| <input checked="" type="checkbox"/> If I Could | 2:25 | Jack Johnson | In Between Dreams |
| <input checked="" type="checkbox"/> Breakdown | 3:33 | Jack Johnson | In Between Dreams |
| <input checked="" type="checkbox"/> Belle | 1:43 | Jack Johnson | In Between Dreams |
| <input checked="" type="checkbox"/> Do You Remember | 2:25 | Jack Johnson | In Between Dreams |
| <input checked="" type="checkbox"/> Constellations | 3:22 | Jack Johnson | In Between Dreams |
| <input checked="" type="checkbox"/> Daniel | 3:56 | Elton John | Love Songs |
| <input checked="" type="checkbox"/> Sowing the Seeds of Love (Spot Vodafone) | 4:05 | Tears For Fears | Top Of The Spot 2008 |
| <input checked="" type="checkbox"/> Fidelity (Spot Kinder Maxi) | 3:47 | Regina Spektor | Top Of The Spot 2008 |

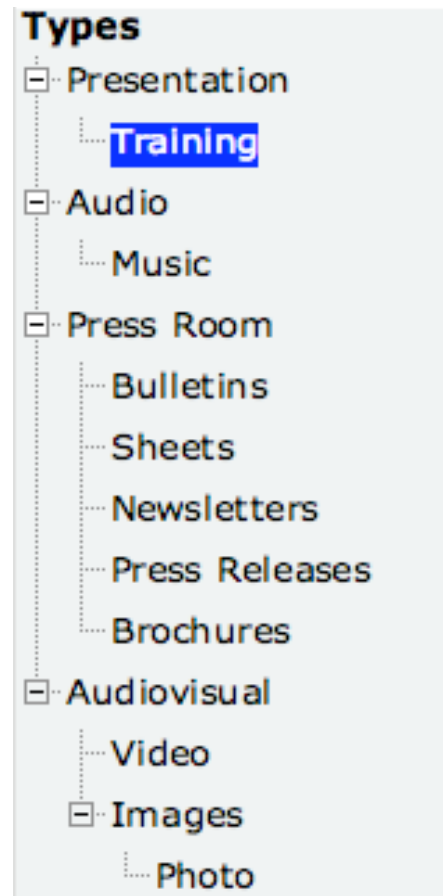
- “Types” and “Categories” definition by repository providers/admins:

- Assets are organized by **type**:

- a list of specific attributes to describe each kind of asset to be managed by the system
- hierarchical (a child type shares and extend parent’s attributes)
- queried during searches

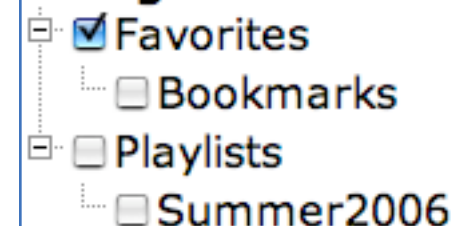
EXAMPLE OF TYPES AND ATTRIBUTES’ LIST

| Type | Attributes’ list |
|--------------|--|
| Audio | Format, Bitrate, Samplerate, Time |
| Music | (Format, Bitrate, Samplerate, Time), Name, Artist, Album, Genre, Tracknumber, Year, Artwork, Lyric, Rating |
| Presentation | Format, NumOfPages |
| Training | (Format, NumOfPages), Title, Runtime, Speaker, Author, Subject, Event, Date, Type |
| (Root) | FileName, SubmissionDate, Description, Keywords, LastModificationDate, Size |





- and/or organized by **category/collection**:
- Group together related assets of different type;
- Useful also to define subsets of assets belonging to the same type
- Multiple category assignment per asset (tagging)


Categories



- Assets can be browsed selecting a type (or category) and selecting one or more **filters**:
 - attributes of the selected types, chosen from a defined list, used to narrow the result set
- Filter application is cascading and context-sensitive: the selection of a filter value dynamically influences subsequent filter values (*“à la iTunes”* browsing)
 - Classical search by description and keywords available too

Speaker 

Event 

Subject 

ALL

Annamaria Muoio

Emidio Giorgio

Gianni Ricciardi

Giuseppe La Rocca

Giuseppe Platania

Tony Calanducci

Valeria Ardizzone

ALL

Cyclops first workshop

Tutorial per gli Insegnanti degli Istituti Tecnici Industriali

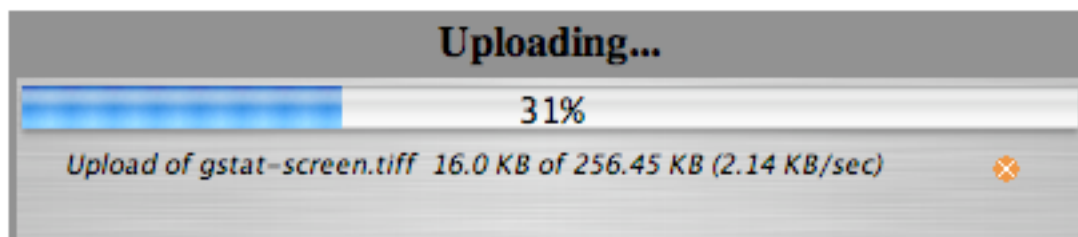
ALL

general presentation

| TITLE | RUNTIME | SPEAKER | SUBJECT | EVENT | AUTHOR | DATE | TYPE | SIZE | FORMAT | NUM |
|----------------------------|---------|-----------------|----------------------|------------------------|-----------------|------|------|---------|--------|-----|
| the GILDA t-infrastructure | 30 | Tony Calanducci | general presentation | Cyclops first workshop | Tony Calanducci | | | 7406292 | pdf | 45 |

- Users can upload their local assets on one or more (creating replicas) Storage Elements of the Grid

- Uploads managed through Java Applets



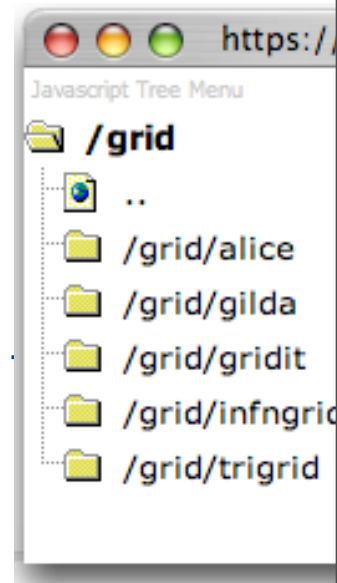
- Files already on SE can be included in a digital library by File Catalogue browser

- Download from SEs to the users' laptop/desktop:

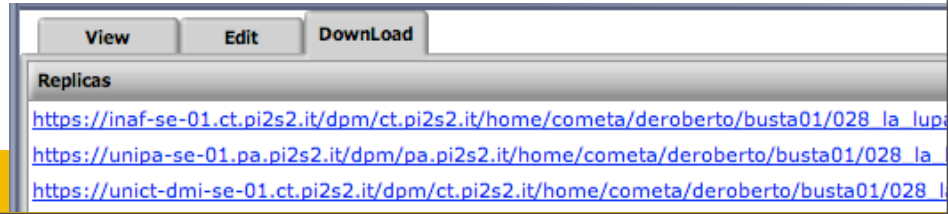
- selection of a replica link from a list
- download java applet

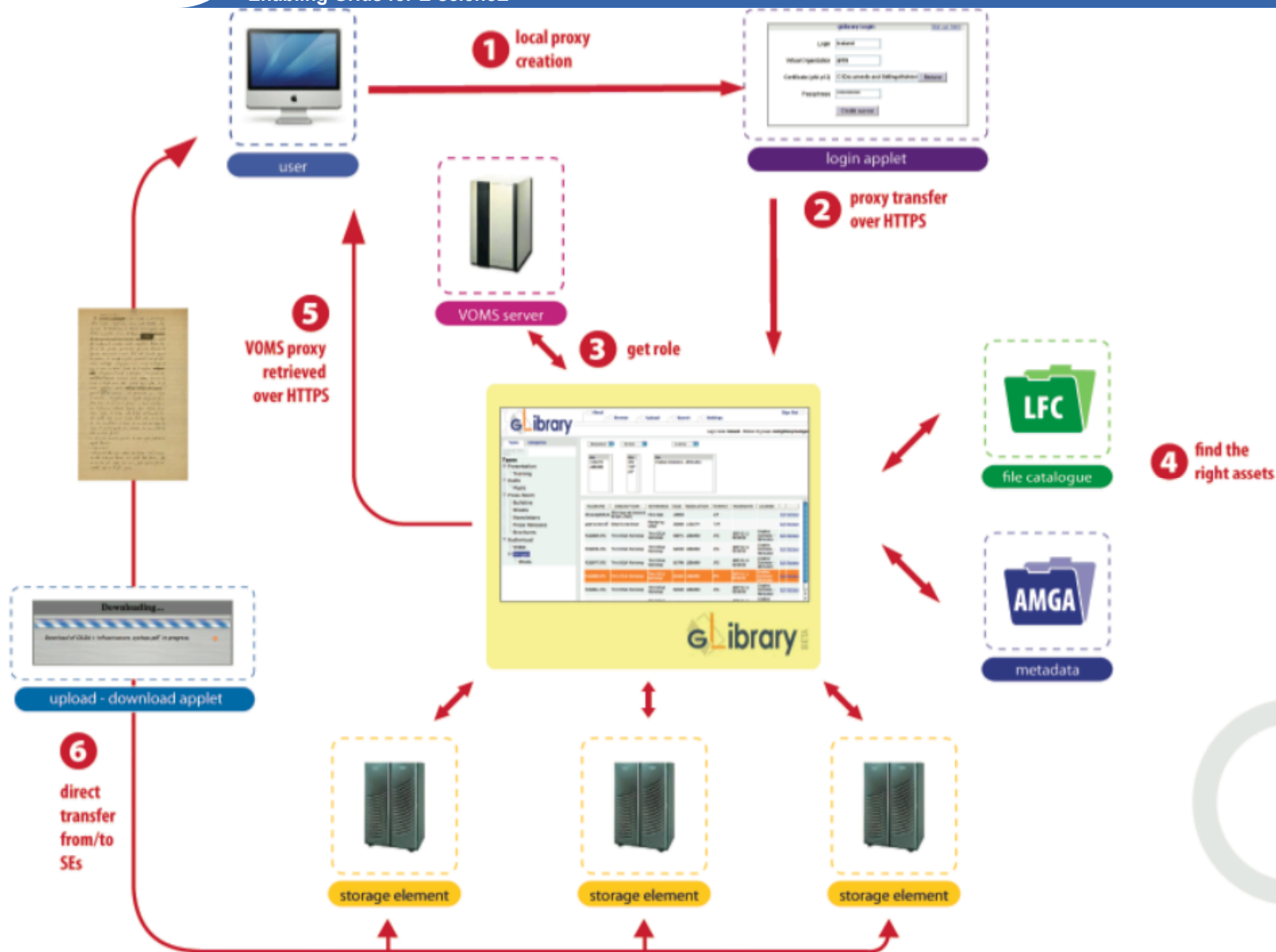
List of replicas:

<srm://infn-se-01.ct.pi2s2.it/dpm/ct.pi2s2.it/home/cometa/generated/2008-06-14/filea203fe0e-08ed-495b-8d16-73a6fa275e8e>
<srm://unime-se-01.me.pi2s2.it/dpm/me.pi2s2.it/home/cometa/generated/2008-06-14/file59c6ebad-6833-4e6d-bc49-f7241e7f45b8>
<srm://unipa-se-01.pa.pi2s2.it/dpm/pa.pi2s2.it/home/cometa/generated/2008-06-14/file680f1e17-530e-4a9e-a9d0-1df0c9cf6d86>



- **Latest NEWS: transfers can be handled over HTTPS with X.509 Grid Cert AuthN/AuthZ**

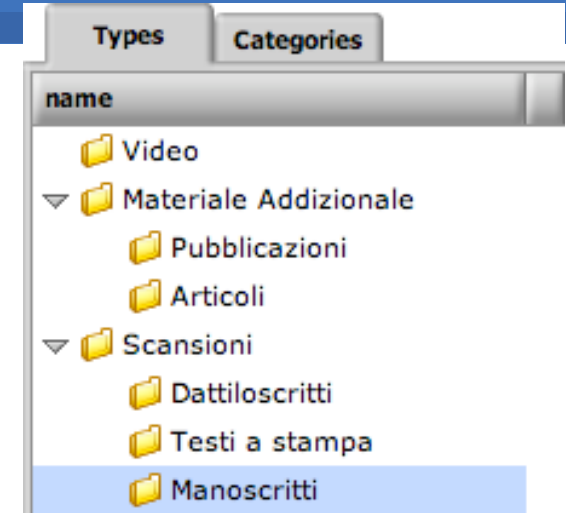




- **Web standards:**
 - AJAX/PHP5/JSON on the client side
- **Grid standards:**
 - Storage Element **SRM** interface to get the TURLs (Transfer URLs)
 - Transfers handled with GridFTP and X.509 cert validated HTTPS
 - X.509 based Globus Security Infrastructure with the VOMS extensions to handle authentication and authorization (ACL based) on Metadata and Storage Elements
 - All grid services implemented with the EGEE gLite middleware (DPM Storage Element, AMGA Metadata Catalogue, LFC File Catalogue, VOMS Services)
- **Other standards:**
 - Subset of XMP Metadata
 - Dublin Core Metadata set will be included in the next release

- **Types** definition for the assets of the DR library
- **Attributes** definition per type. Es:

| Attributo | Valore |
|-----------------|--|
| Title | la lupa |
| Author | federico de roberto, giovanni verga |
| Description | manoscritto della tragedia lirica ... |
| Keywords | verismo, federico de roberto, la lupa, ... |
| CaptionWriter | stefania iannizzotto, alessandro ... |
| CopyrightStatus | copyrighted |
| PageNum | 5 |
| TotalPages | 34 |
| DocumentGenre | tragedia lirica |
| PublicationYear | 1916 |
| Publisher | officine tipo-litografiche barravecchia e balestrini |
| FileType | PDF |
| Resolution | 300 |
| ScanQuality | good |



- **Filter** definition per type. Es:
 - DocumentGenre
 - Title
 - FileType
 - ScanQuality
 - DocumentType
 - PublicationYear
 - PublicationStatus
 - Publisher
 - Location

- There are some libraries that allow automatic metadata extraction from given file types:
 - exiftool
 - Imagero
- We have been used exiftool to extract XMP metadata from TIFF images, e.g.:
 - `$ exiftool -E -XMP:Subject -XMP:Description -XMP:Rights -XMP:Title -XMP:Author -FileName -FileSize 001\ gli illustri\ amanti.tif`
 - Subject : federico de roberto, manoscritti letterari, verismo, gli illustri amanti, la.mu.s.a., facoltà di lettere e filosofia catania, società di storia patria per la sicilia orientale
 - Description : manoscritto de gli illustri amanti, conservato presso la biblioteca della società di storia patria per la sicilia orientale
 - Rights : società di storia patria per la sicilia orientale catania.la.mu.s.a., facoltà di lettere e filosofia, università degli studi di catania
 - Title : gli illustri amanti
 - File Name : 001 gli illustri amanti.tif
 - File Size : 106 MB

https://glibrary.ct.infn.it/glibrary_new/ربول.html

Google

Invio SMS Alice Facebook Home De Roberto DR Google Maps Wikipedia Macity Voda SMS Usingmac.com gLibrary/DRI Rapidshare SA Tutorial app@ICTP ICTP Phone Book

https://glibrary.ct.infn.i... gLibrary Login

Repositories

Current: deroberto2

Trees

Types Categories

name

- Video
- ▼ Materiale Aggiunzionale
 - Pubblicazioni
 - Articoli
- ▼ Scansioni
 - Dattiloscritti
 - Testi a stampa
 - Manoscritti

Filter attributs

FileType DocumentType Author Location

Filter's values

| | | | |
|------|----------------|----------------------------------|----------|
| PDF | dattiloscritto | federico de roberto | busta 01 |
| TIFF | manoscritto | federico de roberto, giovanni v. | busta 02 |
| | -No available- | | busta 11 |

+ Add Filter - Remove Filter

jsevilla

| FILE | Title | Description | PagNum | TotalPages | CaptionWriter | Size | FileType |
|------|---------|--|--------|------------|---|-----------|----------|
| 100 | la lupa | manoscritto della tragedia lirica in due atti, musicata da pierantonio tasca, pubblicata da officine tipo-litografiche barravecchia e balestrini, palermo 1919; altre edizioni: noto, tipografia rosario caruso, ottobre 1932; conservato presso la biblioteca di storia patria per la sicilia orientale | 28 | 34 | stefania iannizzotto, alessandro de filippo | 15736028 | PDF |
| 101 | la lupa | manoscritto della tragedia lirica in due atti, musicata da pierantonio tasca, pubblicata da officine tipo-litografiche barravecchia e balestrini, palermo 1919; altre edizioni: noto, tipografia rosario caruso, ottobre 1932; conservato presso la biblioteca di storia patria per la sicilia orientale | 29 | 34 | stefania iannizzotto, alessandro de filippo | 107433936 | TIFF |
| 102 | la lupa | manoscritto della tragedia lirica in due atti, musicata da pierantonio tasca, pubblicata da officine tipo-litografiche barravecchia e balestrini, palermo 1919; altre edizioni: noto, tipografia rosario caruso, ottobre 1932; conservato presso la biblioteca di storia patria per la sicilia orientale | 29 | 34 | stefania iannizzotto, alessandro de filippo | 15542633 | PDF |
| 103 | la lupa | manoscritto della tragedia lirica in due atti, musicata da pierantonio tasca, pubblicata da officine tipo-litografiche barravecchia e balestrini, palermo 1919; altre edizioni: noto, tipografia rosario caruso, ottobre 1932; conservato presso la biblioteca di storia patria per la sicilia orientale | 30 | 34 | stefania iannizzotto, alessandro de filippo | 107433772 | TIFF |
| 104 | la lupa | manoscritto della tragedia lirica in due atti, musicata da pierantonio tasca, pubblicata da officine tipo-litografiche barravecchia e balestrini, palermo 1919; altre edizioni: noto, tipografia rosario caruso, ottobre 1932; conservato presso la biblioteca di storia patria per la sicilia orientale | 30 | 34 | stefania iannizzotto, alessandro de filippo | 14852580 | PDF |
| | | manoscritto della tragedia lirica in due atti, | | | | | |



| View | Edit | Download |
|-----------------------|---|----------|
| FileName: | 029_la_lupa.tif | |
| TypeID: | 2 | |
| CategoryIDs: | | |
| SubmissionDate: | 2009-03-02 19:54:00 | |
| Description: | manoscritto della tragedia lirica in due atti, musicata da pierantonio tasca, pubblicata da officine tipografiche barravecchia e balestrini, palermo 1919; altre edizioni: noto, tipografia rosario caruso, ottobre 1932; conservato presso la biblioteca di storia patria per la sicilia orientale | |
| Keywords: | verismo, federico de roberto, la lupa, giovanni verga, officine tipo-litografiche barravecchia e balestrini palermo, tipografia rosario caruso noto, società di storia patria per la sicilia orientale catania, manoscritti letterari moderni, la.mu.s.a., facoltà di lettere e filosofia, università degli studi di catania, società di storia patria per la sicilia orientale | |
| LastModificationDate: | 2007-08-11 20:59:00 | |
| Size: | 107433936 | |
| Thumb: | | |
| FileType: | TIFF | |
| ImageWidth: | 5100 | |
| ImageHeight: | 7020 | |
| CreatorTool: | Adobe Photoshop CS2 Windows | |
| XResolution: | 600 | |
| YResolution: | 600 | |
| Title: | la lupa | |
| Author: | federico de roberto | |
| CaptionWriter: | stefania iannizzotto, alessandro de filippo | |
| CopyrightStatus: | Copyrighted | |

Add Filter Remove Filter

Modal Window

View Edit Download

1

1

1

1

103

la lupa

30

34

103 la lupa pierantonio tasca, pubblicata da officine tipo-litografiche barravecchia e balestrini, palermo 1919; altre edizioni: 30 34 rosario caruso noto, società di storia patria per la sicilia orientale 107433772 TIFF

FileName : 029_la_lupa.tif
 CategoryIDs :
 Description : manoscritto della trage
 LastModificationDate : 2007-08-11 20:59:00
 Thumb :
 ImageWidth : 5100
 CreatorTool : Adobe Photoshop CS2 \\\n
 YResolution : 600
 Author : federico de roberto
 CopyrightStatus : Copyrighted
 PagNum : 29
 PhysicalFormat :
 ScanQuality : good
 DocumentType : manoscritto
 TypeID : 2
 SubmissionDate : 2009-03-02 19:54:00
 Keywords : verismo, federico de ro
 Size : 107433936
 FileType : TIFF
 ImageHeight : 7020
 XResolution : 600
 Title : la lupa
 CaptionWriter : stefania iannizzotto, ale
 CopyrightNotice : la.mu.s.a. - facoltà di le
 TotalPages : 34
 ScanDate :
 DocumentGenre : tragedia lirica
 PublicationYear :

| | |
|--|-----------|
| officine tipo- | |
| schia e | |
| ipografia | |
| società di | |
| icilia orientale | 107433936 |
| TIFF | |
| letterari | |
| facoltà di | |
| iversità degli | |
| età di storia | |
| rientale | |
| roberto, la | |
| , officine tipo- | |
| schia e | |
| ipografia | |
| società di | |
| icilia orientale | 15542633 |
| PDF | |
| letterari | |
| facoltà di | |
| iversità degli | |
| età di storia | |
| rientale | |
| roberto, la | |
| , officine tipo- | |
| schia e | |
| ipografia | |
| rosario caruso noto, società di | |
| storia patria per la sicilia orientale | 107433772 |
| TIFF | |

http://unict-dmi-se-01.ct.pi2s2.it:777/gpfs/cometa/2009-03-02/029_la_lupa.pdf.1...TlcSOu+LUM15HVZf0IYBqnCJs7KAlloQnt

Google

http://unict-dmi-se-01.ct.pi2s2.it:777/gpfs/cometa/2009-03-02/029_la_lupa.pdf.1069362.0?httpstoken=/dp

Google Maps

Facebook | Home | De Roberto DR | Google Maps

Wikipedia | Macity | Voda SMS | Usingmac.com | gLibra

A T T O P R I M O

Nell'ala
sull'imbrunire.

A destra la capanna dei mietitori, a sinistra una gran bica; macchie ed attrezzi rurali sparsi qua e là. In fondo, l'ampia distesa della carica di messe, già velata dalla sera, e il corso sinuoso del fiume e le cune palustri. S'odono passare in lontananza delle canzoni strache, il tintinnio dei campanacci delle mandre che se abbayerare e di tanto in tanto l'uggiolare dei cani sparsi per la sulla quale scorrono delle folate di vento, con un fruscio largamente mature. Negli intervalli di silenzio sembra sorgere e diffondersi delle acque e il trillare dei grilli, incessante. La luna incomincia accesa, sbiancandosi man mano, in un alone afoso.

SCENA I.

Bruno, Malerba, Neli, Grazia, contadini e contadine
duti sulla paglia e su barili vuoti e cestoni rotti
intorno a Filomena che sta narrando una favola

Save PDF to download folder

Basta coi pianti!

Pi... poi vedrete quando un'orfanello
al mondo lascerà'!

Maledetto sia Giuda! E' proprio bella
questa festa ch'io fo!

Ed io? ed io?

Negate! negate!

Non e' vero! Che intendi?

E negate! negate!

Non e' vero!

Che cosa hai visto, di?

- **Contact:** antonio.calanducci@ct.infn.it,
lamusa@unict.it
- **Prototype of the De Roberto Digital Repository:**
 - <https://glibrary.ct.infn.it/deroberto/>
- **Papers:**
 - A. Calanducci, R. Barbera, A. De Filippo, F. De Mattia, S. Iannizzotto, M. Saso, J. Sevilla, D. Vicinanza "*Digital Libraries on the Grid to preserve cultural heritage*" Proceedings of the First EELA-2 Conference - B. Marechal et al. (Eds.) - CIEMAT 2008
 - <https://glibrary.ct.infn.it/t/EELA-3-paper-final.doc>
 - A. Calanducci, C. Cherubino, L. N. Ciuffo, D. Scardaci, "*A Digital Library Management System for the Grid*", Fourth International Workshop on Emerging Technologies for Next-generation GRID (ETNGRID 2007) at 16th IEEE International Workshops on Enabling Technologies: Infrastructures for Collaborative Enterprises (WETICE-2007), GET/INT Paris, France, June 18-20, 2007 (<http://etngrid.diit.unict.it/2007/index.html>).
 - https://glibrary.ct.infn.it/glibrary/downloads/gLibrary_paper_v2.pdf

A large, dark blue question mark is centered on the slide. It is composed of a thick, curved top and a thick, vertical stem. The question mark is superimposed over a faint background image of a globe and various electronic devices like laptops and servers.

Thank you for the attention