

2017

Open Geoportal lands to Europe: use cases and improvements from Geodata@Polimi


Marcella Samakovlija

Leonardo Campus Library, Archive and Library Systems, Politecnico di Milano p.zza Leonardo Da Vinci, 32, Milano, Italy

Francesco Bartoli

Geobeyond S.r.l. via Marchesa Augusta 68, 02040, Vacone (RI), Italy

Follow this and additional works at: <https://scholarworks.umass.edu/foss4g>

 Part of the [Cataloging and Metadata Commons](#), and the [Collection Development and Management Commons](#)

Recommended Citation

Samakovlija, Marcella and Bartoli, Francesco (2017) "Open Geoportal lands to Europe: use cases and improvements from Geodata@Polimi," *Free and Open Source Software for Geospatial (FOSS4G) Conference Proceedings*: Vol. 17 , Article 11.

DOI: <https://doi.org/10.7275/R59W0CNC>

Available at: <https://scholarworks.umass.edu/foss4g/vol17/iss1/11>

This Paper is brought to you for free and open access by ScholarWorks@UMass Amherst. It has been accepted for inclusion in Free and Open Source Software for Geospatial (FOSS4G) Conference Proceedings by an authorized editor of ScholarWorks@UMass Amherst. For more information, please contact scholarworks@library.umass.edu.

Open Geoportal lands to Europe: use cases and improvements from Geodata@Polimi

Marcella Samakovlija^{a,*}, Francesco Bartoli^b

^a*Leonardo Campus Library, Archive and Library Systems, Politecnico di Milano p.zza Leonardo Da Vinci, 32, Milano, Italy*

^b*Geobeyond S.r.l. via Marchesa Augusta 68, 02040, Vacone (RI), Italy*

Abstract: TeDOC is part of the University Library System, owns a Great collection of historical and current maps, paper and digital and provides at their distribution to teachers and students, as part of the educational and research activities of the Politecnico di Milano. We were searching for a tool able to find, display and download current and historical cartography and, after a few years of work, we have become the first Europeans to use Open Geoportal (OGP) as our working tool. The component to ingest the data to OGP has evolved considerably, allowing better integration with GeoServer and allow automation of the procedures for adding layer via the importer function with the REST API. In this way, it is not necessary to be an expert operator to load data into the OGP. The procedure of data download has been customized to allow students to download prepackaged packages prepared by the administrators (some complex data need additional documentation to be correctly read, such as legends or lists of codes). One of our needs was to interact with other universities for sharing data, and OGP allows us to do this. For this reason, the Metadata that we have prepared follow the ISO 19139 standard and are in double language, allowing the widest possible dissemination and facilitating the exchange. What we hope is that our experience can become a first European base for the OGP Federation, also a new way to share data between Italian and European Universities.

*Corresponding author

Email address: marcella.samakovlija@polimi.it (Marcella Samakovlija)

1. Introduction

For a long time, we were looking for a tool that would be able to combine geographic research with the search for the most current cartographic typologies (raster, vector, scanned maps), but at the same time not exclude from research the historical cartographies that, for Italian copyright reasons, cannot be transformed into digital elements (Kollen et al. 2013). Usually, the historical cartographies information is contained in the classical University Catalogue (MARC21 and unimarc formats).

Open Geoportal (<https://github.com/OpenGeoportal>) is the tool we have chosen since it allows geographically searching of different typologies of maps in an easy and fast way. The community of Universities and Institutions linked to the Open Geoportal project has created this tool and all the necessary components in a balanced way, always following long debates, comparisons and experiments in a climate of mutual collaboration. We believe that sharing this knowledge from the community is crucial in identifying our closest issues and identifying the most appropriate solutions to overcome them.

Our needs in particular are:

- Allow an inexperienced operator to upload and process the data independently;
- Maintain the original format of the data as they were acquired by the owner Institutions;
- Allow our users to download the necessary data without necessarily going to the library;
- Allow to search and download data that our institution does not have directly and at the same time, allow sharing our cartographies with the federated Institutions (Muratori and Samakovlija 2015).

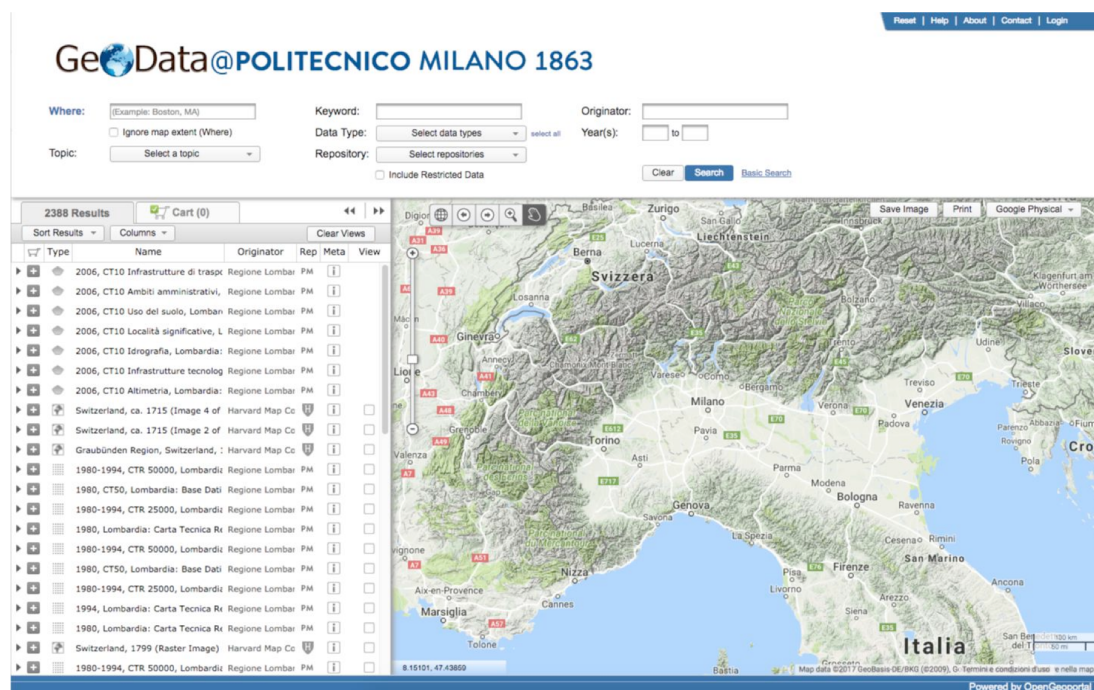


Figure 1: GeoData@Polimi

2. Advanced Integration with Geoserver

2.1. Exploiting the Importer plugin to automate the ingestion of raster and vector data

GeoServer is fundamentally one of the most important component in the architecture of Open Geoportal as it handles all the middleware OGC services for viewing and downloading functionalities. OGPIngest is the primary tool for end users that wants to upload data to Open Geoportal and the old procedure provides a two steps workflow to allow discovering vector and raster records from SOLR and make real data available from GeoServer:

- Load metadata from the OGPIngest GUI to the SOLR collection
- Create manually in GeoServer the associated resource, a layer, and the Web Mapping Service (WMS) or Web Feature Service (WFS) accordingly to what the previous step has declared in the location of its SOLR record.

The new OGPIngest is able to exploit a GeoServer equipped with the Importer plugin (<http://docs.geoserver.org/latest/en/user/extensions/importer/>) and use its REST API interface to automate the previous operations in a single step. Vector datasets such as shapefiles will be loaded as layer stored through a PostGIS datasource connection while raster ones such as JPEG and GeoTIFF will be treated accordingly to WorldImage and GeoTIFF datastore.

Our administrators, with proper privileges to load data in the geoportal, can indeed interact with the old option in the ingest GUI to select this automatic procedure as depicted in the Figure 2. At the end of each successful attempt, a new layer is immediately available and consumable within Open Geoportal for being searched, viewed and downloaded.

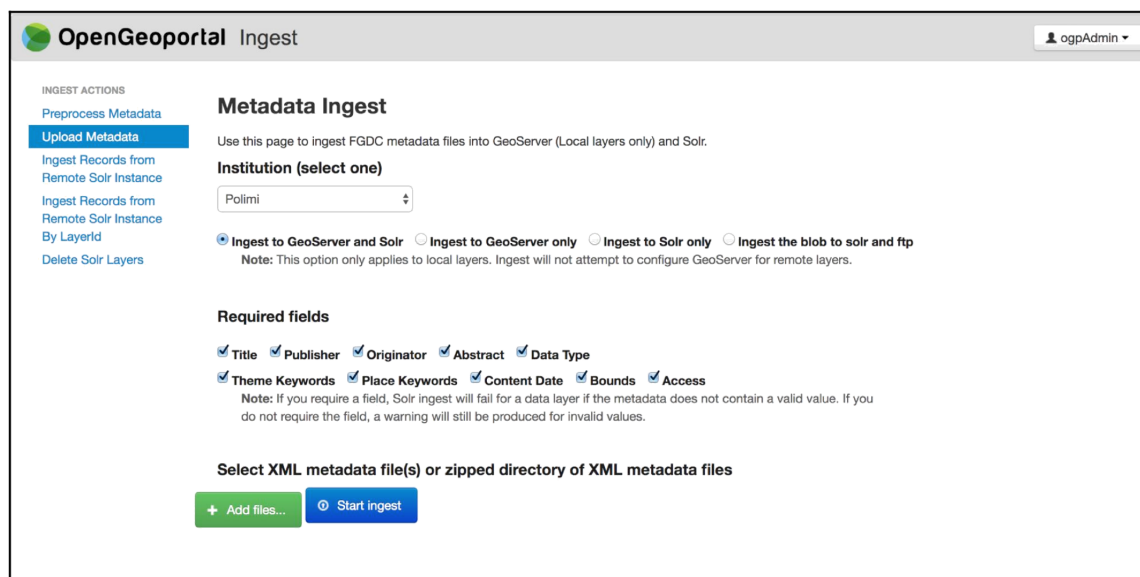


Figure 2: Upload menu of OGPIngest.

Behind the scene, OGPIngest performs a client HTTP call to the REST endpoint of the importer in order to accomplish the creation of the layer and the association with a specific datastore in GeoServer. These calls, for example in case of uploading a shapefile to PostGIS, are actually the same of the following commands:


```
curl -u admin:geoserver -XPOST -H "Content-type: application/json" -d @import.json
"http://localhost:8080/geoserver/rest/imports"
```

Figure 3: HTTP call example to trigger the importer.

```
{
  "import": {
    "targetStore": {
      "dataStore": {
        "name": "postgis"
      }
    },
    "targetWorkspace": {
      "workspace": {
        "name": "public"
      }
    }
  }
}
```

Figure 4: JSON file example to import a Shapefile, import.json.

The response will have an empty import task identification, which can be then used to trigger the real task:

```
curl -u admin:geoserver -F name=myshapefile -F filedata=@myshapefile.zip
"http://localhost:8080/geoserver/rest/imports/14/tasks"
```

Figure 5: HTTP call example to trigger a specific import task.

Then set the task to have a PostGIS target for the import action:

```
curl -u admin:geoserver -XPUT -H "Content-type: application/json" -d
@target.json "http://localhost:8080/geoserver/rest/imports/14/tasks/0/target"
```

Figure 6: HTTP call example to set the target for an import task.

```
{
  "dataStore": {
    "name": "postgis"
  }
}
```

Figure 7: JSON file example to set a target datastore, target.json.

Finally execute the complete action:

```
curl -u admin:geoserver -XPOST "http://localhost:8080/geoserver/rest/imports/14"
```

Figure 8: HTTP call example to execute the complete import job.

3. Management and Download of Blob Data

An improvement from the base Open Geoportal was planned to allow the end users to get assembled collection of generic files for the same datasets with few clicks. The new procedure has been required for handling any kind of blob data, usually a zip file with single or multiple vector/raster datasets correlated by any additional file for styles, data specification and standard metadata compliant to ISO 19139.

The solution has been designed toward a distributed architecture rather than a storage in a local file system. A new component - FTP/SFTP server - has been introduced to host the zip files and create the resource for the specific metadata field location in the Solr records. During the ingestion of blob the operator has a new option Ingest the blob to solr and ftp as depicted above in Figure 9.

The choice of FTP has been taken in order to meet the requirement of a reliable and robust way to exchange and host files with the most secure and consolidated protocol within a safe network zone. That doesn't mean it was selected also for the download procedure whose links will still continue to be served through the HTTP protocol and being protected by the identity and access management system. This also has the advantage to not require any additional integration of the FTP server within the perimeter of Single Sign-On performed by Shibboleth.

The final architecture can be described as per the diagram below:

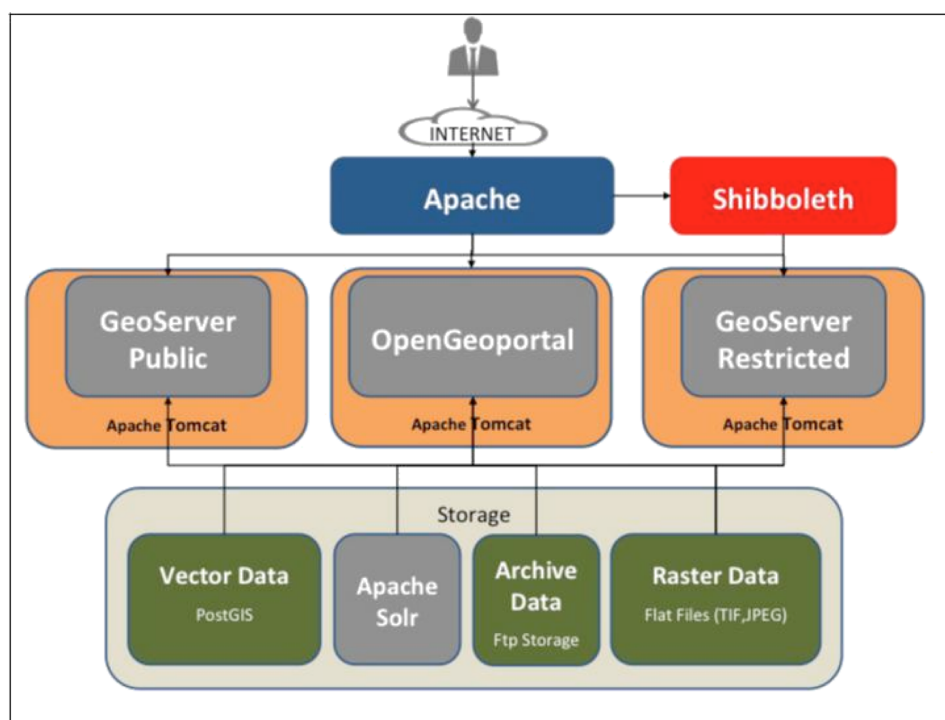


Figure 9: Geo Portal architecture for blob data.

3.1. Code improvement

To fill the solution described above an FTP client has been added in the OGPIngest and OGP controller to handling specific actions and trigger the movement of files from the upload to the storage and viceversa on users demand when they needs to download them asynchronously by email.

3.1.1. OGPIngest

A new bean for the FTP client is declared in the applicationContext.xml:

```
<beans:bean id="theFtpClient" class="org.OpenGeoPortal.Utilities.TheFtpClient">
  <beans:property name="ftpserver" value="${local.ftp.server}"/>
  <beans:property name="ftpport" value="${local.ftp.port}"/>
  <beans:property name="ftpusr" value="${local.ftp.usr}"/>
  <beans:property name="ftppwd" value="${local.ftp.pwd}"/>
  <beans:property name="ftppdir" value="${local.ftp.dir}"/>
</beans:bean>
```

Figure 10: OGPIngest application Context.xml with new bean.

And a corresponding properties section in the ingest.properties file:

```
local.ftp.server=192.168.56.102
local.ftp.port=21
local.ftp.usr=vagrant
local.ftp.pwd=vagrant
local.ftp.dir=/tmp
```

Figure 11: OGPIngest ingest.properties with ftp section.

3.1.2. Open Geoportal

A modified bean for the download method is declared in the *applicationContext.xml*:

```
<beans:bean id="downloadMethod.email.polimi"
  class="org.opengeoportal.download.methods.EmailDownloadMethodImpl"
  scope="prototype">
  <beans:property name="smtpuser" value="${smtpuser}" />
  <beans:property name="smtppassword" value="${smtppassword}" />
  <beans:property name="smtpfrom" value="${smtpfrom}" />
  <beans:property name="smtpsubject" value="${smtpsubject}" />
</beans:bean>
```

Figure 12: Open Geoportal applicationContext.xml with download bean.

And a corresponding json key/value file for configuring the download methods:

```
"polimi": [
  { "classKey": "layerDownloader.email.polimi",
    "accessLevel": ["public", "restricted"], "dataType": ["vector"], "outputFormats": ["shp"] }
  ,
  { "classKey": "layerDownloader.email.polimi",
    "accessLevel": ["public", "restricted"], "dataType": ["raster"], "outputFormats": ["geotiff"] },
  { "classKey": "layerDownloader.email.polimi",
    "accessLevel": ["public", "restricted"], "dataType": ["vector", "raster", "undefined"], "outputFormats": ["shp", "geotiff"] },
  ]
```

Figure 13: Open Geoportal ogpDownloadConfig with configured classes.

4. Metadata in Double Language

The Politecnico di Milano has different maps with various typologies and provenances, the most recent ones are accompanied by a metadata according to the INSPIRE guidelines

(<http://inspire.ec.europa.eu/documents/inspire-metadata-implementing-rules-technical-guidelines-based-en-iso-19115-and-en-iso-1>) and therefore compatible with the ISO 19139 standard

(<http://inspire.ec.europa.eu/id/document/tg/metadata-iso19139> and

<https://www.iso.org/standard/32557.html>), while the maps of the 80s and 90s are free of metadata.

For all the native digital cartography, in order to allow the widest possible dissemination of the information contained (Goodchild 2000), the choice was to include abstract, lineage, keywords and usage in double language: Italian and English. The choice was also supported by the institutional bodies of the Politecnico di Milano who see in the internationalization one of the goals to be reached in the short term. In fact, more and more foreign students come to study in Italy and need to find suitable material for the tasks assigned.

For recent cartographies, in accordance with the guidelines drawn up by the federated (Florence et al. 2015) and in agreement with the owners of the data that could benefit from this work, we have prepared new metadata (always in double language).

Metadata Download Metadata (XML) X

1980, Lombardia: Carta Tecnica Regionale, foglio B6A4 / 1980, Lombardia: Regional Base Map, tile B6A4

Abstract:

Rappresenta l'intero territorio regionale suddiviso in 708 quadranti. Il contenuto informativo è in scala 1:10.000. La Carta Tecnica Regionale è costituita da: 1) elementi ed entità di tipo geometrico: reticolato chilometrico, coordinate geografiche, punti quotati, curve di livello. Il valore di equidistanza tra le isoipse è di 50 metri per le direttrici, di 10 metri per le intermedie e di 5 metri per quelle ausiliarie, indicate a tratteggio; 2) elementi costitutivi del paesaggio naturale, quali il reticolato idrografico, i laghi, i rilievi, la vegetazione, etc.; 3) elementi costitutivi del paesaggio antropico, quali insediamenti, strade, ferrovie, canali, colture agricole, etc; 4) limiti amministrativi; 5) toponimi. / It represents the entire territory of the Lombardy Region divided into 708 tiles. The scale is 1:10000. The Regional Base Map is composed by: 1) geometric elements and entities: kilometric grid, geographic coordinates, elevation points, contour lines. The contour interval is 50, 10 and 5 meters respectively for reference, intermediate and auxiliaries lines, represented with dashed lines; 2) elements of the natural landscape such as hydrography, lakes, elevation, vegetation, etc.; 3) elements of the anthropic landscape such as urban settlements, roads, railways, canals, land use, etc.; 4) administrative boundaries; 5) toponyms.

Lineage:

La carta tecnica regionale è stata acquisita in formato digitale tramite scansione, mosaicata e georeferenziata per permetterne un uso nell'ambito dei sistemi informativi geografici. Questa banca dati va considerata come documentazione storica. / The Regional Base Map has been acquired in digital format by scanning and georeferencing the paper maps into a mosaic map which can be used in Geographic Information Systems. This dataset must be considered as historical documentation.

Purpose:

Questa base dati è indirizzata a studenti e ricercatori per rappresentazioni e analisi geografiche; può essere utilizzata per visualizzare, interrogare e produrre cartografie o come elemento per effettuare operazioni di overlay mapping con altre basi dati geografiche. / This dataset is intended for researchers, and students, for reference and mapping purpose, and may be used for basic applications such as viewing, querying, and map output production, or to provide a basemap to support graphical overlays and analysis with other spatial data.

Keywords:

Lombardia / Lombardy
Italia / Italy

Figure 14: Example of metadata visualization in double language.

A series of cartographies, for copyright reasons, will never be available in digital format, but are already present in the University Catalogue. For this type of map, we followed the MIT guidelines to suit our needs. In this way the user can geographically search for this type of cartography even if cannot download them, but will have to go to the library personally. Technically, the catalog record is produced in Unimarc format (<https://www.ifla.org/publications/unimarc-formats-and-related-documentation>), converted to MARC21

(<https://www.loc.gov/marc/bibliographic/>) and extracted in xml format.

Through the tool MARCIngest (<https://github.com/gravesm/marcingest>) the MARC21 is inserted into the Solr instance allowing the user to easily search it. Within the catalogue card are inserted the geographic coordinates that allow to define the space occupied by the paper and to display its bounding box. The presence of the permalink allows the user to display the catalogue card directly inside the University Catalogue, where he will also find all the information necessary for his request.

Figure 15: Example of paper map searching.

5. Conclusions and Next Development

The realization of GeoData@Polimi is not a point of arrival for us, but a starting point. We intend to make new improvements to OGP shared with the Federation, perhaps even having a multilingual interface (not just a metadata).

In particular, our work will focus on the opportunity of:

- digitalize and georeferencing historical maps owned by the Libraries (in respect of copyright);
- implement the data on the portal with the federation of other Institutions and Universities (Italian and foreign);
- use the geoportal to spread the georeferenced data produced by our Departments for study and research purposes.

The Politecnico di Milano has a rich collection of historical cartography, buried in the archives of Engineering, part of illustrious collections that for various reasons are difficult to find by users. It is our intention to start a work on this material, in order to make it available not only to the Politecnico users. Harvard colleagues, who have a very high-class historical map collection, will have the opportunity to suggest the most suitable method for scanning and georeferencing heterogeneous and ancient maps such ours.

GeoData@Polimi is not the first Open Geoportal in the world, but is the first in Europe and in Italy. We hope that our work can arouse the interest of other nearby universities so we can expand the Federation and share new materials. Finding cartography is getting easier through the geoportals of the various Institutions: each one with its authentication, downloading, or service modes (Florance 2006). For students, used to work on different geographic locations, having a single point for searching is an advantage. We would like to share this way of searching cartography with the students from other University as well.

Other partner in the Federation are working on how to build and evaluate the user experience of an Open Source geoportal⁵ (Wolff and Parker 2014), we are available to integrate the results of this research with our reference experience and the feed back provided by our users. We are at the beginning, but we count on the practice done by the Federates in these years of work.

Others Universities in Europe are partner in great projects such as Old Maps Online (<http://www.oldmapsonline.org/>) or promote directly their work on open linked data (<http://lodum.de/>). It could be very interesting find a way to explore these experiences and add new data and new interoperability option between this innovative projects and GeoData.

Like every University, our Departments and Laboratories produce quality research whose results are disseminated through international journals. It is difficult, however, to give the same relevance and dissemination to geographic data that has been realized from those researches (Durante and Wang 2012). Which best tool can be GeoData@Polimi? In addition to searching for basic geographic data, we will also make available those unique data produced by our researchers. We will have to work in collaborations with the Departments and the Federation to build appropriate metadata, which unequivocally provide the value of innovation.

References

- Durante, K., Wang, Z., 2012. Creating an actionable assessment framework for discovery services in academic libraries. *College Undergraduate Libraries* 19(2-4), 215–228.
- Florance, P., 2006. Gis collection development within an academic library. *Library trends* 55(2), 222–235.
- Florance, P., McGee, M., Barnett, C., McDonald, S., 2015. The open geoportal federation. *Journal of Map Geography Libraries* 11, 376–394.
- Goodchild, M. F., 2000. Communicating geographic information in a digital age. *Annals of the Association of American Geographers* 90(2), 344–355.
- Kollen, C., Dietz, C., Suh, J., Lee, A., 2013. Geospatial data catalogs: Approaches by academic libraries. *Journal of Map Geography Libraries* 9, 276–295.
- Muratori, I., Samakovlija, M., 2015. Sharing geodata through university libraries: the case of politecnico di milano. Online, online; accessed May 31, 2017.
URL <http://eprints.rclis.org/27995/>
- Wolff, R., Parker, K., 2014. Usability testing: Open geoportal 2.0. place project. paper 9. Online, online; accessed May 31, 2017.
URL <http://scholars.unh.edu/place/9>