

# A web-based, relational database for landslide, mud-debris flows and flood studies in northern Italy

Guido Nigrelli\*  
Maura Vergnano\*\*

\* Consiglio Nazionale delle Ricerche,  
Istituto di Ricerca per la Protezione  
Idrogeologica, Torino  
\*\* Università degli Studi di Torino,  
Dipartimento di Informatica, Torino

## Acknowledgments

We would like to express our gratitude to Dr. Ornella Turitto for many years devoted to BIBLIO, for helping in the development of OpenBIBLIO and for giving us support in the preparation of this manuscript.

## 1. Introduction

In its forty-two years of activity, the Research Institute for Geo-hydrological Protection of the Italian National Research Council (CNR-IRPI) has amassed a vast and varied collection of historical publications. The collection focuses primarily on landslides, mud-debris flows and floods but also includes earthquakes, avalanches and glacier front variations. The documentation covers a time span from the 1800s to the present, with some records dating back to the 18<sup>th</sup> and 17<sup>th</sup> centuries. The geographical area of reference (over 102,000 km<sup>2</sup>) comprises northern Italy and contiguous areas of

neighboring countries (Austria, France, Slovenia, Switzerland). Within this area are seven major hydrographic basins (Po, Isonzo, Tagliamento, Piave, Brenta, Bacchiglione, Adige) plus numerous sub-basins that drain into the Adriatic and the Tyrrhenian.

By the late 1980s, the collection had grown to over 10,000 resources, making consultation increasingly time-consuming and cumbersome. The solution lay with transferring the catalogued data for each resource into a computerized database. This was done using a dedicated program named BIBLIO written in DB4. The BIBLIO database was composed of a single table with 13 fields. Originally installed on a PC (Olivetti 486), it could carry out simple functions such as executing queries and printing out reports, thus shortening user time considerably (Di Nunzio *et al.*, 1990).

Some twenty years later, a completely new system had to be developed that would accommodate the greater mass of documents (over 17,000 by 2008) and could extend user accessibility. To respond to these needs, in the spring of 2009 plans began on creating a new data-

*In its forty-two years of activity, the Research Institute for Geo-hydrological Protection of the Italian National Research Council has amassed a vast and varied collection of historical publications. In this collection, the documentation on landslides, mud-debris flows and floods covers a time span from the 1800s to the present and consists of more than 18,000 bibliographic resources. The evidence collected was not stored with information systems. To respond to these needs, in the 2009 plans began on creating a new database called OpenBIBLIO that incorporated the data from bibliographic resources. The OB access is free on line, upon request for non-profit purposes. The aim is to provide a tool that may be useful for improving knowledge in this area of geohydrology and to furnish information for civil protection agencies.*

**Keywords:** Open source solutions, Database, Floods, Landslides, Italy.

**Un nuovo database finalizzato alla gestione di risorse bibliografiche inerenti frane, colate detritiche ed alluvioni nell'Italia settentrionale.** L'Istituto di Ricerca per la Protezione Idrogeologica del CNR, ha raccolto nel corso dei suoi 42 anni di attività una considerevole quantità di documentazione inerente frane, colate di detrito ed alluvioni, avvenute nell'Italia settentrionale dagli inizi del 1800 sino ad oggi. Per quanto riguarda le risorse bibliografiche archiviate, queste ammontano ad oltre 18000 e necessitavano di uno strumento informatico che ne rendesse più facile la reperibilità e la fruizione. Per questi scopi è stato realizzato un nuovo database denominato OpenBIBLIO. L'accesso ad OB è libero per finalità senza scopo di lucro, previa richiesta di autenticazione.

**Parole chiave:** Software open source, Database, Alluvioni, Frane, Italia.

**Un nouveau database pour la gestion des ressources bibliographiques inhérentes éboulements, laves torrentielles et inondations dans l'Italie Septentrionale.** L'Institut de Recherche pour la Protection Hydrogéologique (IRPI-CNR) a ramassé, dans le cours de ses 42 ans d'activité, une considérable quantité de documents se rapportant à éboulements, laves torrentielles et inondations, produites dans l'Italie Septentrionale entre début 1800 jusqu'à à aujourd'hui. Les ressources bibliographiques qui dépassent les 18000 sujets nécessitaient d'un moyen informatique afin de rendre plus facile l'organisation et l'utilisation des informations. On a exprès réalisé un nouveau database dénommé OpenBIBLIO. L'accès à OB est libre avec simple demande d'authentification.

**Mots-clés:** Open source solutions, Database, Inondations, Glissements de terrain, Italie.

base that incorporated the data from the old one. The new database was designed according to the latest technologies and called OpenBIBLIO (OB), which reflected its open, free and reusable characteristics. It was only later and by pure chance that we discovered that a far more important online database (“OpenBiblio, a library system that’s free”) carried the same name as ours. The following sections describe the database architecture, web interface and predefined queries.

## 2. Database architecture and technology

OB was designed to archive and manage a large amount of bibliographic resources in both paper and digital format. It serves two basic purposes: 1) to collect and store historical publications (articles, international, national and local journals, monographs, conference proceedings, technical reports) and maps (historical and thematic); 2) to provide access to information and products for use by geoscientists, researchers, decision makers, and students. The relational database’s modular architecture is extensible through tables (currently 17 tables, 11 of which already compiled so that users can select search specifications). Figure 1 illustrates the database’s schematic structure.

The *country* table (table name in italics) lists the countries to which each resource refers; the *region* table lists the geopolitical Italian regions (20); the *province* table lists the 110 provinces; and the *municipality* table lists the 8,092 municipalities. The tables are joined through a one-to-many relationship. The *riverbasin* table lists the seven hydrographic basins (in which the reported phenomena occur) and the sub-basins directly associated with them (total, 535 sub-basins).

The topics and phenomena users can associate with each resource are listed in the *key\_topic* table (total, 103 topics). This table is joined to the *biblio* table through the *jt\_key\_topic* table. The three junction tables (*jt*) serve to manage the many-to-many relationships between the two tables associated to them. The physical location of a resource in the collection is listed in the *office* and *institute* tables. The *resource\_type* table lists the type of support of a resource (paper, digital, paper-digital format). The *users* table lists the identification data of users and data entry operators, each of which with a respective login level. This table is joined to the *project* table through the *jt\_project\_user* table. The *project* table contains information on the research projects being carried out by the data entry operators. The *login* table contains access information on each user (login/logout date and time). Finally, the *temp* table temporarily stores the data entered in the so-called multiple fields, as will be described below.

The database access has five different login levels (lev) for each user assigned by the administrator: lev1 (search and print data); lev2 (lev1, plus export data); lev3 (lev2, plus access to the digital resource and enter/modify data); lev4 (lev3, plus modify resources entered by all users); lev5 (full control, database administrator).

The relational database and web application were developed on a Microsoft Windows Server 2003 operating system through open source solutions (Apache 2.2.4, MySQL 5.0.45, and PHP 5.2.4). The database is managed with phpMyAdmin, a free software tool written in PHP which handles the administration of MySQL over the world wide web. A Javascript jQuery framework that exploits Ajax technology was also integrated (Vergnano, 2009). Data persistence is guaranteed by a RAID 5 server configuration with three hard disks and by daily backups through an external network attached storage (NAS) unit.

Data migration from the old BIBLIO to the new OB database was performed in three phases: 1. Conversation of BIBLIO into a single spreadsheet file by means of an algorithm; record cleaning because the BIBLIO used alphanumeric characters which had become obsolete; creation of new fields, some of which were compiled using logical and mathematical functions; 2. Importing of the spreadsheet in MySQL using phpMyAdmin and entry of part of the data with the use of SQL queries; 3. Final revision with the use of a dedicated modify data interface.

## 3. Web interface

The web interface includes 133 pages that users can access through a simple and intuitive graphical user interface. OB was designed and developed following On-line Public Access Catalogue (OPAC, 2012) standards and requirements of the National Library Service (Servizio Bibliotecario Nazionale, SBN, 2012). All web site pages were validated according to World Wide Web Consortium (W3C) standards for XHTML 1.0 and CSS 2.1 specifications.

Each bibliographic resource is catalogued by compiling the 24 fields on the two web pages “inserimento” (insert resource) or “modifica” (modify resource). These 24 fields differ depending on how the data are entered. Four fields are already filled in when the page is opened or after data have been entered in other fields (id resource, user name, project leader). In the other 20 fields the data are entered by the users: selecting one option from a list displayed in a combo-box (3 fields: resource type, office, project title); performing a multiple choice selection which is temporarily saved in the *temp* table until all the fields have been filled in and the resource is

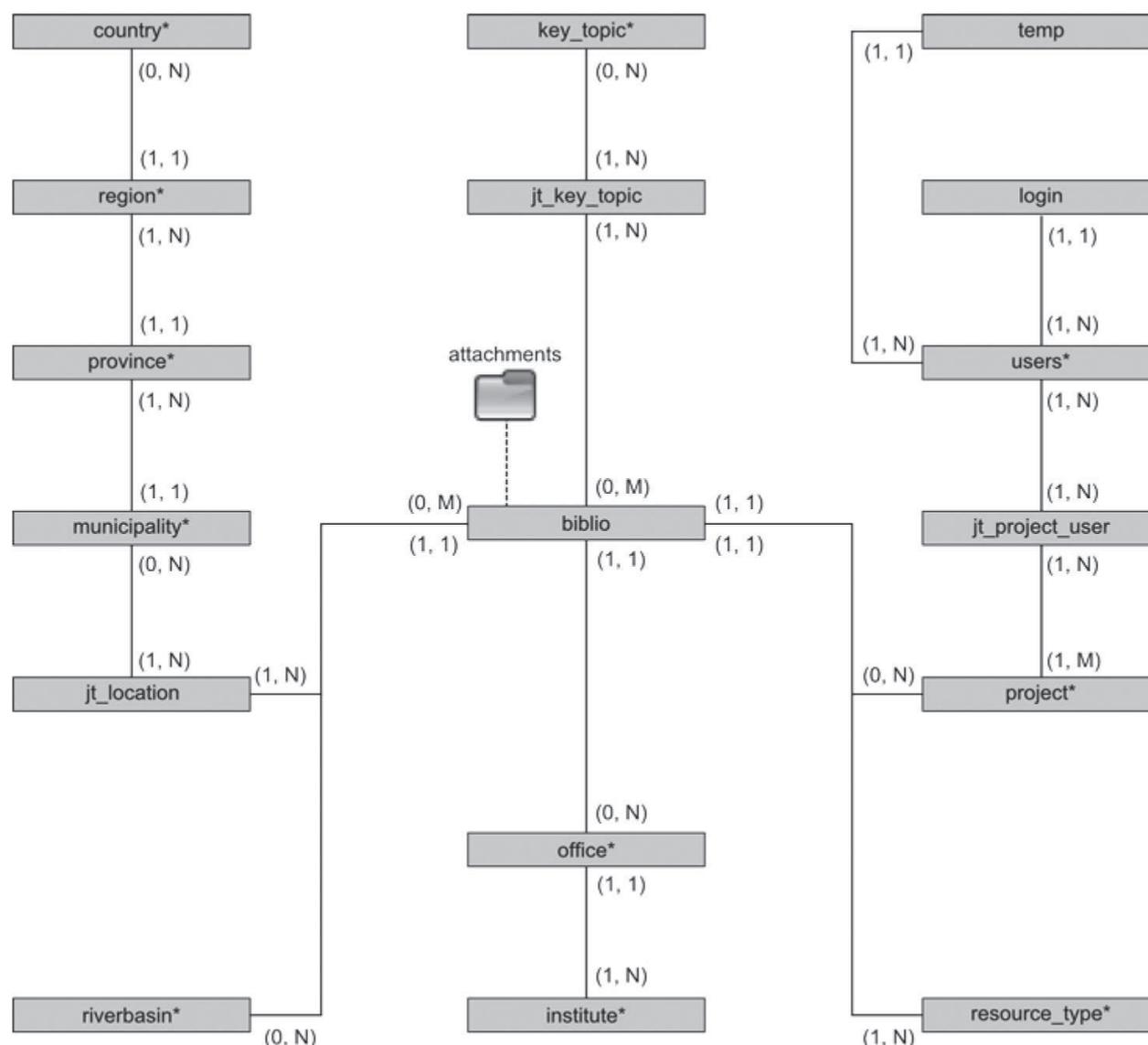


Fig. 1. Schematic structure of the database: tables (squares) and links (lines). Relationship between two fields of two tables: (1, 1), one-to-one; (1, N), one-to-many; (0, N) and (0, M), many-to-many with junction table (jt). The asterisk denoted the prefilled tables. *Struttura schematica del database, con in evidenza le tabelle (rettangoli grigi) ed i relativi collegamenti (linee nere). Le relazioni fra due tabelle sono: (1, 1), tipo uno ad uno; (1, N), tipo uno a molti; (0, M), molti a molti con le indispensabili tabelle di collegamento (jt). L'asterisco indica le tabelle precompilate.*

inserted in the database (5 fields: key topic, municipality, province, region, country); performing an optional choice which is automatically activated by the Javascript library (6 fields: the 5 preceding fields plus river basin); dialoging in the conventional file manager window (1 field: upload file); manual data entry (10 fields: other data). All commonly used formats for managing text or image files are supported by the database during file uploading. The downloaded files are automatically saved in the attachments cartel on the website (Fig. 1).

Twenty-two predefined parametric queries were created to enable users to query the database and retrieve relevant information (Fig. 2). The parametric queries can be broadly grouped into: a) direct queries performed

on bibliographic data (e.g., extract the list of resources associated with a specific author) and on other key topics (e.g., extract all resources dealing with landslides, floods, etc.); b) geographic location queries (e.g., extract the list of resources mentioning a selected municipality or hydrographic basin); c) management queries (e.g., extract the list of resources catalogued or modified by a data operator). The results report list can be printed out or exported in a portable document format (pdf) or CSV format by pressing a specific query button. CSV is a simple file format widely supported by consumer, business and scientific applications.

The OB database is available at <http://dbirpi.to.cnr.it/irpibiblio/index.php> (Italian version). Currently it

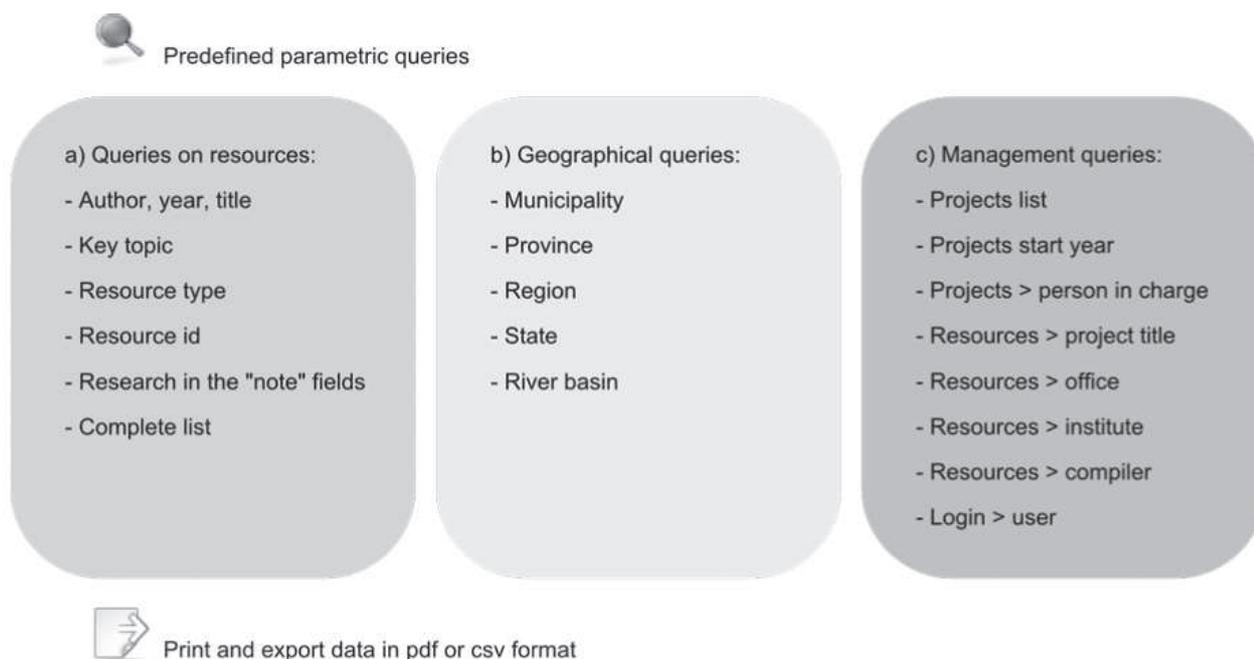


Fig. 2. Predefined parametric queries, print and export of the data. The symbol > denotes "related to".  
*Tipologie di interrogazioni predefinite presenti, stampa ed esportazione dei risultati. Il simbolo > indica "relativo a".*

contains about 18,000 bibliographic resources. Access is free upon request for non-profit purposes. The aim is to provide a tool that may be useful for improving knowledge in this area of geohydrology and to furnish information for civil protection agencies.

## References

DI NUNZIO, F., TURITTO, O., (1990). Banca dati bibliografici su frane ed inondazioni nell'Italia settentrionale (A bibliographic database on landslides and floods in northern Italy). Proceedings of the 2nd Workshop "GIAST, Informatica e Scienze della Terra", Sarnano, Italy, 126-129.

OPAC, (2012). On-line Public Access Catalogue. Istituto Centrale per il Catalogo Unico (Central Institute for the Union Catalogue of Italian Libraries and for Bibliographic Information), Roma. <http://opac.sbn.it/>, [accessed February 15, 2012].

SBN, (2012). On-line Public Access Catalogue. Servizio Bibliotecario Nazionale (National Library Service), Roma. <http://www.iccu.sbn.it/>, [accessed February 15, 2012].

VERGNANO, M., (2009). Sistema gestionale di trattamento delle pubblicazioni scientifiche inerenti la protezione geo-idrologica (An open source system for the management of scientific publications on geohazards). Unpublished B.Sc. Thesis, Università degli Studi di Torino, Torino, 78 pp.