U.S. DEPARTMENT OF THE INTERIOR
U.S. GEOLOGICAL SURVEY

A Bibliographic Database of Geologic Hazards Publications in Northern California on the World Wide Web

Compiled by

Ronald C. Schott[1]

Open File Report 95-661

September 1995

[1] MS 977, Menlo Park, CA 94025

Abstract

A bibliographic database of publications relating to geologic hazards in northern California has been published on the World Wide Web (WWW) since August 7, 1995. The title of the WWW page from which the database is accessed is "Northern California Geologic Hazards" and its Universal Resource Locator (URL) is currently:

http://quake.wr.usgs.gov/CALHAZ/calhaz.html

The database can also be accessed by following links from the "Earthquake Info from the U.S.G.S. in Menlo Park, CA" page with URL:

http://quake.wr.usgs.gov/

Basic bibliographic reference information including title, author(s), source, year of publication, map information, report number, and availability have been compiled for over 1500 references. The database can be searched by both county and hazard type. The homepage of the database contains a brief description of its contents, contact information for the U.S.G.S. Library in Menlo Park, CA, and addresses from which U.S.G.S. and C.D.M.G. publications can be ordered. Following this background information are the primary links to the database. Finally, links are presented which allow forms-based updating of references, and comments to the maintainer.

Introduction

The purpose of this publication is to increase public awareness of research and planning efforts which aim to understand and mitigate the adverse effects of a variety of geologic hazards in this heavily populated, geologically active area. The information in this database is envisioned to be useful to public planners, those doing active research in the area, interested citizens, students, and businesses. The compilation and WWW formatting of this database was done by Ronald C. Schott as a volunteer to the U. S. Geological Survey Office of Earthquakes, Volcanoes, and Engineering (USGS-OEVE) during the summer of 1995.

Database Organization

The database is accessed via a treed series of links. At present the references have been grouped by general subject matter and by county. The general subject categories include:

* Seismic Zonation, Earthquake Shaking, and Liquefaction - Geologic Investigations
* Earthquakes - Public Planning, Response, and Policy
* Earthquake Triggered Landsliding
* Landslides, Debris Flows, and Slope Stability - Geologic Investigations
* Landslides - Public Planning, Response, and Policy
* Flooding and Erosion
* Erosion and Landsliding Triggered by Timber Harvesting
* Volcanic Hazards
* Coastal Hazards and Policy
* Geologic Hazards to Dams
* Geologic Maps and Fault Maps
* Landslide and Slope Maps
* Land Use Maps
* Miscellaneous Subjects

All northern and central California counties are included south to (and including) San Luis Obispo, Kern, and Inyo Counties. For each publication referenced in the database, some or all of the following information has been made available: title, author(s), source, year of publication, map information (including map type and scale), report number (particularly for USGS and CDMG publications), and availability. Future revisions of the database will include annotations summarizing the specific subject matter of individual references.

Preparation of the Database

Most of the reference information was derived from searches of the GeoRef CD-ROM bibliographic database published by the American Geological Institute (AGI). Additional references have been compiled directly from their individual bibliographic sources. What follows is a general account of the process by which the GeoRef citations were prepared for WWW publication.

First, Georef searches were conducted on relevant topics, and the search results downloaded onto computer disk. Downloads were executed employing the options for downloading "All Fields" and "Short Field Labels". The resulting text files are readable from most word processing programs (i.e. Microsoft Word). The next step was to reformat these text files to a "text-delimited" format so that they could be imported into a spreadsheet or database program. (Beware! Spreadsheet programs will truncate long individual records.) This was achieved by first replacing all double hard returns (which signify a break between two references) with an "odd character" (i.e. "|") which was not likely to be found in the downloaded text. Next, all remaining hard returns were replaced with tabs. Finally, all of the "odd character" (i.e. "|") were replaced with single hard returns. At this point each reference should have been separated by a hard return and the individual fields within a given reference by a tabs. The resulting file was in text delimited format with tabs as field delimiters and no text delimiters.

Unfortunately, GeoRef is not formatted uniformly throughout; consequently, not all records will have the same fields. Therefore, in order to most efficiently review and sort the large number of references, it became necessary to add empty fields to each record so that all records had all of the 25 data fields that GeoRef currently uses. This was a tedious and time consuming task. The author found that it was most efficiently accomplished in a spreadsheet program (i.e. Lotus 123R5) through a series of sort and move (cut and paste) operations. A talented programmer might be able to automate this task, as it is fairly repetitive, but it is at least a step beyond simple macro programming.

Once a suitable degree of uniformity had been imposed upon the database, each record was examined individually and assigned a county and subject field. This was done primarily on the basis of information in the title and descriptor fields provided by GeoRef. Ideally this will be checked in the future against each individual reference. Also, at this stage numerous vagaries of the GeoRef data were ironed out; for example, all USGS publications were checked to make sure that both report numbers and availability were uniformly and accurately portrayed. After assigning the county and subject fields and cleaning up the data, it was transferred to a database program (i.e. Microsoft Access). The

database program was employed to query the assembled database and output tables containing title, author, source, year of publication, map information, report number, and availability of each publication. The query simultaneously sorted by subject area, year of publication (most recent first), and author's last name. Such queries were run for each county and for the entire database as a whole. The resulting tables were exported into text delimited files (again with tabs as field delimiters and no text delimiters). The text delimited files were next opened in a word processing program (i.e. Microsoft Word). The following macro was used to convert the delimited text file into Hyper-Text Markup Language (HTML) format:

```
Sub MAIN
EditReplace .Find = "^p", .Replace = "^p^p", .WholeWord = 0, .MatchCase = 0,
.Format = 0, \
.ReplaceAll = 1
EditReplace .Find = "^t", .Replace = "^p", .WholeWord = 0, .MatchCase = 0,
.Format = 0, \
.ReplaceAll = 1
EditReplace .Find = "TI: ", .Replace = "<P><LI><B>TITLE: </B>", .WholeWord
= 0, \
.MatchCase = 0, .Format = 0, .ReplaceAll = 1
EditReplace .Find = "AU: ", .Replace = "<BR><B>AUTHOR(S): </B>",
.WholeWord = 0, \
.MatchCase = 0, .Format = 0, .ReplaceAll = 1
EditReplace .Find = "SO: ", .Replace = "<BR><B>SOURCE: </B>",
.WholeWord = 0, \
.MatchCase = 0, .Format = 0, .ReplaceAll = 1
EditReplace .Find = "YR: ", .Replace = "<BR><B>YEAR: </B>", .WholeWord =
0, \
.MatchCase = 0, .Format = 0, .ReplaceAll = 1
EditReplace .Find = "MP: ", .Replace = "<BR><B>MAP: </B>", .WholeWord =
0, \
.MatchCase = 0, .Format = 0, .ReplaceAll = 1
EditReplace .Find = "RN: ", .Replace = "<BR><B>REPORT NUMBER:
</B>",WholeWord =0, \
.MatchCase = 0, .Format = 0, .ReplaceAll = 1
EditReplace .Find = "AV: ", .Replace = "<BR><B>AVAILABILITY: </B>",
.WholeWord = 0, \
.MatchCase = 0, .Format = 0, .ReplaceAll = 1
End Sub
```

Once in HTML format, the text was cut and pasted into the generic HTML file that follows:

```
<HTML>
<HEAD>
<TITLE>##COUNTY## Geologic Hazards</TITLE>
</HEAD>

<BODY>
<A NAME="TOP"></A>
<CENTER><H1>##COUNTY## Geologic Hazards</H1></CENTER>
```

```
<P>
References are organized by year of publication within each subject area.  Click on
a specific subject area to jump down.
</P>

<P>
<UL>
<LI><A HREF="#AEQ">Seismic Zonation, Earthquake Shaking, and
Liquefaction - Geologic Investigations</A>
<LI><A HREF="#BEP">Earthquakes - Public Planning, Response, and
Policy</A>

:

<LI><A HREF="#NMI">Miscellaneous Subjects</A>
</UL>

<P>
<H2><A NAME="AEQ">Seismic Zonation, Earthquake Shaking, and
Liquefaction - Geologic Investigations</A>-##COUNTY##</H2>
<OL>
## References ##
</OL>
<A HREF="#TOP">Back to Top</A>

<P>
<H2><A NAME="BEP">Earthquakes - Public Planning, Response, and
Policy</A>-##COUNTY##</H2>
<OL>
## References ##
</OL>
<A HREF="#TOP">Back to Top</A>


<P>
<H2><A NAME="NMI">Miscellaneous Subjects</A>-##COUNTY##</H2>
<OL>
## References ##
</OL>
</P>
<A HREF="#TOP">Back to Top</A>

<P>
<H3>Return to <A HREF="CALHAZ/counties.html">County List</A>.</H3>
<H3>Return to <A HREF="CALHAZ/calhaz.html">Hazards Home
Page</A>.</H3>
</BODY>
</HTML>
```

Finally, the completed files were transfered to the USGS earthquake information server.

Suggestions for Future Work

Numerous aspects of this database and its presentation allow room for future improvement. Some of the following suggested improvements are ideas that were originally planned but have not yet been incorporated due to time constraints; others have manifested themselves during the evolution of its design and construction.

One of the most useful of all possible additions to the current database would be a field which summarizes the topic and scope of each publication. Addition of such a summary category will of necessity be a time consuming task, but it will greatly enhance the utility of the database. Additional ordering information (for publications from sources other than the USGS and CDMG) would be another addition that would make the database more user-friendly. Finally, in order to maintain its relevance, updates of the current information will be necessary from time to time. The current forms-based update procedure requires a maintainer to examine each update, reformat it, and add it to each relevant county or subject file. It is possible that some or all of this process could be programmed to be executed automatically.

Three aspects of the structural setup of the current database warrant refinement. First, many references that are listed under multiple counties (particularly in the San Francisco Bay area) are more relevant in a regional, rather than a county-specific sense. It would seem reasonable therefore to create separate regional categories for such references and leave only the references of a more county-specific nature in the individual county files. Second, some hazards categories that are only regionally relevant are currently listed for all counties. Coastal and volcanic hazards categories have been removed from uneffected counties, but landslide and earthquake hazard categories need to be reviewed regarding their relevance to a number of counties, particularly those of the Great Valley. Finally, addition of a keyword search feature would make the utilization of the database more flexible and user friendly.