
Online tools facilities for historical earthquake data investigation

Mario Locati <locati@mi.ingv.it>
Andrea Cassera <cassera@mi.ingv.it>

Abstract

SVG can be an easy and effective solution in the historical earthquake data investigation when it comes to a very specific need not achievable using general purpose software. With a basic knowledge of web languages (HTML, Javascript and PHP) and commonly available server side software (Apache and not spatially-aware MySQL) we succeeded creating an interactive web-mapping tool for macroseismic data and a time-saver gazetteer management tool.

Table of Contents

Introduction	1
Online macroseismic databases	2
Background experience with the Italian Macroscopic Database	2
Macroscopic data structure	4
Required functionalities	5
Development and technical details	6
2008 update	9
Final users feedback	9
Gazetteer management tool	10
Gazetteer structure	10
Required functionalities	11
Development and technical details	12
Conclusion and future development	14
References	14

Introduction

In our working group at the INGV, Istituto Nazionale di Geofisica e Vulcanologia, Sezione di Milano-Pavia we have to deal with the geographical representation and manipulation of data describing the macroseismic effects of historical earthquakes.

We first collect and store scientific papers and documents, called Studies, supplying such data covering a thousand years of earthquakes, then a record for each collected item is added to our earthquake inventory.

It often happens that a single earthquake is being studied by multiple Studies, so the second step is to make connection between already collected items and analyze the differences.

The third step is to extract from each Study the distribution in time and space of the effects of the described earthquakes. In this way each collected Study item produce the so called "intensity distribution" of an earthquake: a list of places cited, each with a degree scale of effect (for example the EMS98 [Grünthal et al., 1998] or Modified Mercalli scale [http://en.wikipedia.org/wiki/Mercalli_intensity_scale]). The whole collection of observed intensities become the so called "Macroscopic Intensity Database".

The fourth and last part is to sum all the retrieved data in calculated parameters, creating the so called "Earthquake Parametric Catalogue", a catalogue in which each entry represents a single earthquake with a series of parameters describing the event such as the number of observations, the maximum observed intensity, the epicentre coordinates, the calculated epicentral intensity and magnitude.

This process is extremely time consuming and often not much rewarding in term of research papers but is the fundamental input piece of the general seismic hazard assessment process.

The working group is involved with European Community research projects and shares the entire research process among researcher spread in different countries. We are also an Italian public financed research institute and we do have to spread knowledge to the Italian public.

The working group is involved in two major project:

- at European level with NA4, the Networking Activity 4, called Distributed Archive of Historical Earthquakes Data, part of the EU funded project NERIES, Network of Research Infrastructures for European Seismology [[http:// www.neries-eu.org/](http://www.neries-eu.org/)];
- at Italian level, with the Italian Macroseismic Intensity Database DBMI04 [Stucchi et al., 2007], a contribute to the S1 Project [<http://esse1.mi.ingv.it/>] [Montaldo et al., 2007] for improving and using the Italian seismic hazard map in collaboration with the Italian Civil Protection Department.

The goal of both projects is to make historical earthquake intensity data available to the public in digital form. In Europe data are still collected and archived in isolation by many of the seismological observatories, with varying criteria and degrees of commitment. So far we have completed the time-span 1000-1600 for the whole European region, and for Italy we cover until the year 2006.

In Europe only a three countries already have a consistent, available set of historical earthquake data, interpreted in terms of macroseismic "IDP", Intensity Data-Points: Italy with DBMI04, France with SISFRANCE [Scotti et al., 2004], and Switzerland with ECOS [Swiss Seismological Service, 2002]. In other areas, scattered data are available on paper, though the coordinates of the data-points are not easily obtainable and the place-names are missing in many cases. Homogenisation and consistent documentation of many studies and data across Europe is missing.

This is in stark contrast with the corresponding situation in the USA, where the centralised Earthquake Intensity Database [<http://www.ngdc.noaa.gov/hazard/intintro.shtml>] provides more than 150,000 IDP data for over 23,000 damaging events after 1638, with the whole data set freely available on the Internet. Also in South America the CERESIS [Giesecke et al., 2004] published on the Internet a unified Earthquake Intensity Database.

A series of software facilities were planned to achieve our mentioned projects goals and here we presents two of them:

- an historical earthquake intensity database web mapping tool;
- a gazetteer management tool.

In both tools we found that SVG demonstrate an extreme flexible simplicity.

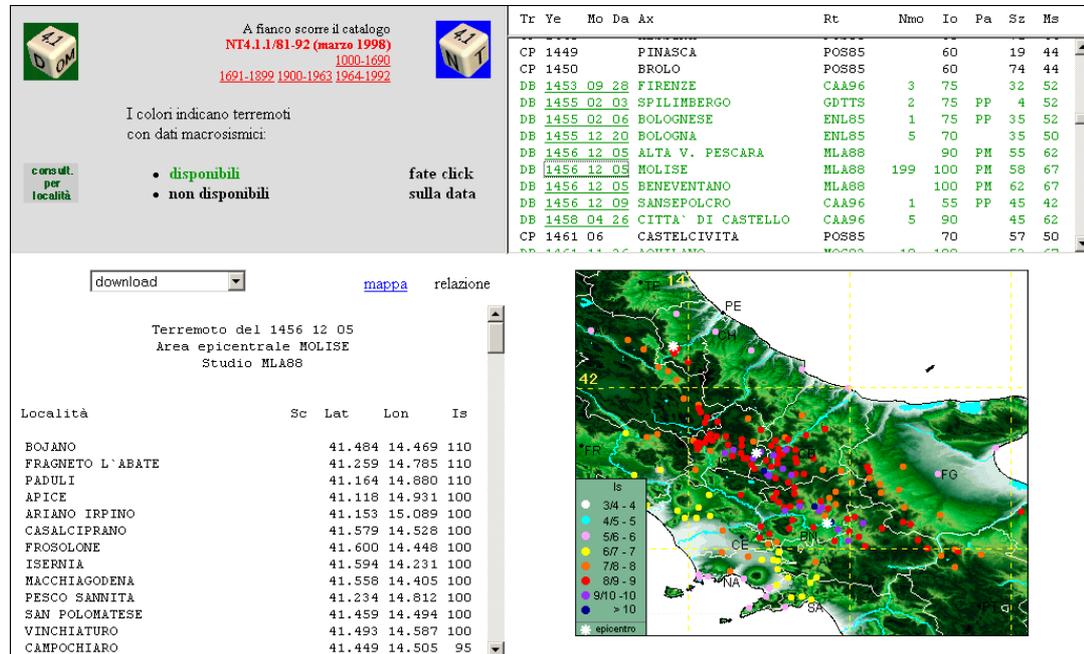
Online macroseismic databases

Background experience with the Italian Macroseismic Database

The working group has faced the needs for web-mapping publication back in 1997 with the first Italian Macroseismic Database called DOM4.1 [Monachesi and Stucchi, 1997]. At that time the words "web-mapping" or "web-GIS" still had to be invented and every solution had to be built from scratch. It is enough to have a look at the Wikipedia "Web mapping" [http://en.wikipedia.org/wiki/Web_mapping]

entry to see how well in advance the implementation was. Figure 1 shows how this early attempt looks like: a frameset and table based lightweight website offer the possibility to retrieve earthquake macroseismic maps intuitively.

Figure 1. The 1997's DOM4.1 web-mapping user interface.



Maps are simple bitmapped GIFs with an html map: each macroseismic intensity point is clickable. Neither zoom nor pan tool were possible at that time.

In 2004 a new Seismic Hazard Map of Italy [MPS Working Group, 2004] has been released by a task force in collaboration with the Italian Civil Protection Department that produced an amount of new or updated data, such as a new version of the Italian Parametric Earthquake Catalogue CPTI04 [CPTI04 Working Group, 2004] and the new historical earthquake intensity database DBMI04.

The design process of the second generation web mapping site for distributing the updated Intensity Database was started in 2003. Back in 2003 the web mapping technology attempts were greatly enhancing the user experience possibilities and both the commercial products and the open source world started to introduce promising products.

Existing Intensity Earthquake Databases published on the Internet were using different web-mapping software solutions:

- SISFRANCE uses an open source software: University of Minnesota published their MapServer [http://mapserver.gis.umn.edu/];
- ECOS, uses a commercial product: ESRI ArcIMS [http://www.esri.com/software/arcgis/arcims/].

Both solutions are "WMS", Web Map Service [http://en.wikipedia.org/wiki/Web_Map_Server] "OGC", Open Geospatial Consortium [http://en.wikipedia.org/wiki/Open_Geospatial_Consortium] compliant: a client software collect the user requests and calls a server that generate and sends back bitmapped maps.

We tested both software and selected none of them: they were over powered, over complicated for our purposes; huge computational power from the server-side point of view, uneasy customisation of the layout interface and web-accessibility make us deciding to internally develop our own solution. Apart from the developing point of view, resulting maps with both MapServer and ArcIMS are simple bitmaps and once downloaded the offline use is very limited: no high printing quality and no interaction is possible.

During the initial software solution investigation we encounter SVG and with very little effort we were able to produce:

- high quality graphic, resolution independent both on video and on printer;
- interactive maps with zoom and pan with just few line of easy-to-do scripting;
- "clever" maps: once the SVG maps has been created it is self sustained, no further remote server calls are needed anymore;
- hight accessibility potential, being SVG obejcts XML based;

Dealing with "clever" and self-sustained maps may not appear important today, but in our opinion is the warranty that in the future (we are talking about years) we will still be able to browse into our file archives without loosing information or loss of functionalities. From our point-of-view external resources based solutions, such as WMS services are, may cause archiving problems in long terms.

In 2003 the SVG choice were looking promising and after years we still are happy with it.

Macroseismic data structure

The macroseismic intensity data-points in DBMI04 are 58926 related to 1024 earthquakes out of 2550 listed in the CPTI04 catalogue, for which an intensity distribution has been assessed. A IDP is identified by: the place name, its geographical coordinates, and the macroseismic intensity assessed according to a macroseismic intensity scale.

Table 1. DBMI04 field structure

Parameter	Description	Type	Sub-type
NDBMI04	Record identifier	numeric	integer
NCPTI04	Earthquake identifier	numeric	integer
An	Year	numeric	integer
Me	Month	numeric	integer
Gi	Day	numeric	integer
Or	Hour	numeric	integer
Mi	Minutes	numeric	integer
Se	Seconds	numeric	integer
AE	Epicentral area	text	30
Rt	Root (original study code)	text	8
Np	Number of observed intensities	numeric	integer
Ix	Maximum observed intensity (MCS scale)	text	5
Io	Epicentral intensity (MCS scale)	text	5
LatEp	Epicentral latitude	numeric	6.3
LonEp	Epicentral longitude	numeric	6.3
Mw	Moment magnitude	numeric	1.2
Daw	Moment magnitude associated error	numeric	1.2
Loc	Place name	text	50

Parameter	Description	Type	Sub-type
Sc	Place special case code	text	2
LatIDP	Place latitude	numeric	6.3
LonIDP	Place longitude	numeric	6.3
Is	Place observed intensity (MCS scale)	text	5
Cou	Country code	text	2
Istat01	Italian municipality code (ISTAT 2001)	text	8
Pr	Italian province code	text	2
LocOr	Place name in the original study		50
LatOr	Place latitude in the original study	numeric	6.3
LonOr	Place longitude in the original study	numeric	6.3
IsOr	Place observed intensity in the original study	text	5

The 58926 intensity data-points in the database are supplied by different studies that used different formats with regards to: name standards, coordinates rounding, etc.

Thus implementing a homogeneous database required:

1. referring the macroseismic observations to places listed in a unique and standard reference directory;
2. correcting possible mistakes in the georeferencing of the IDP, deriving mostly from the original association of an observation coming from the primary source to a place with the same name but different geographical location.

The homogenization process might imply modification from what the original Study reports, therefore DBMI04 always reports two series of information: the original and the revised version. Out of the 58926 input records, over a thousand of IDP geo-referencing were modified.

The web-mapping tool helped the process thanks to its quick use.

The continuous feedback from end-users helped finding bugs and implementing new features. The system was tested both locally and remotely, being the working group distributed all around Italy.

Required functionalities

The website target are both the seismologist's scientific community and the general public.

DBMI04 presents an updated version of both the old DOM4.1 website functionalities: a query by earthquake and a query by place.

The main revision consists in the user possibility to interact with maps: zooming, panning and mapped features query are possible.

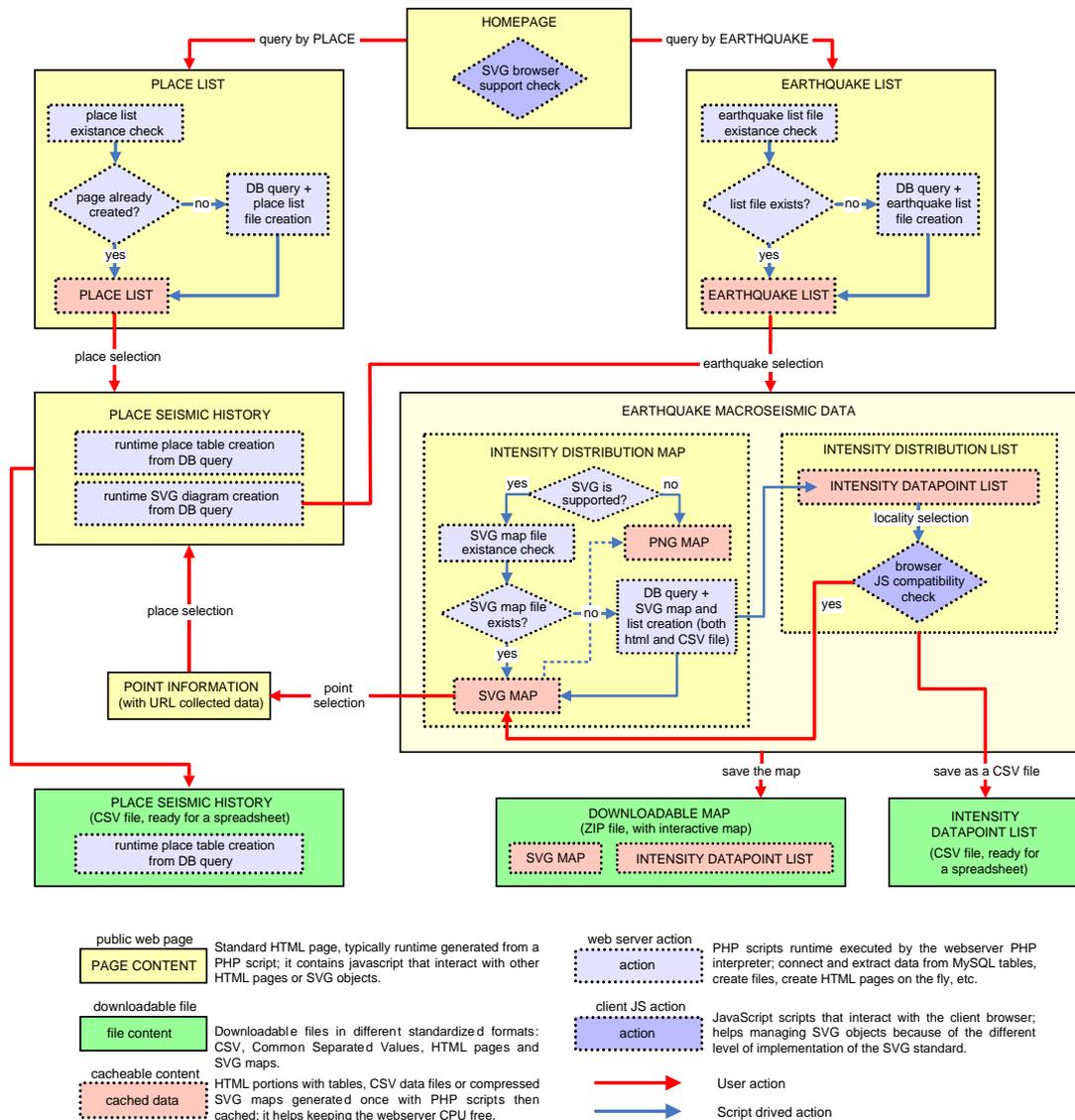
The used SVG technology offers by design the possibility of implementing an offline browsing with ease, allowing the future publication of the entire database in a cd-rom.

Responsiveness is a priority: the general user experience must not suffer slow down by passing from still bitmapped maps to interactive maps.

The whole system has to be developed in a way that the result is open to future expansions and revisions; for this reason has to follow well defined web standards.

Development and technical details

The website workflow scheme is as follow:



Being the lightweight footprint one of the main goals on both the server-side and the client-side we decided to cache every generated file on the file-system: this process decreases dramatically the computing power needed on the web server. On the client side SVG maps don't need any remote call, being every further user map interaction (zoom, pan, place search) directly done by the browser JavaScript engine.

The pre-caching include the bitmapped DEM, that is created in ESRI ArcMAP from SRTM [<http://en.wikipedia.org/wiki/SRTM>] public data exported as a GeoTIFF file. Internal PHP code directly encode into the SVG file the bitmap by using geographical information stored in the TFW file.

Other than speeding up the user experience, pre-generated SVG maps can be putted in a cd-rom and offline consultation is possible using only javascript.

In the early developing process we didn't cache files, and users lamented a lack of responsiveness between the earthquake selection and the browser displaying the map. An extensive check were done to identify factors that influence the visualisation speed on the client side:

1. map generation on the server, later solved with a cached pre-generation of maps;
2. map file transferring time, solved with a GZIP compression of the SVG file;
3. the complete DEM, Digital Elevation Model [http://en.wikipedia.org/wiki/Digital_elevation_model] of Italy was too big so we decided to crop it on the area of interest; for a potential later offline use of the SVG map we also decided to encode the DEM bitmap binary data directly into the code;
4. client hardware configuration: the SVG file is completely managed by the client computer, its CPU and its RAM must be up-to-date. We guarantee a satisfying use of the system starting from a CPU clocked at least at 1Ghz and a system equipped with a minimum RAM size of 512Mb;
5. client software configuration: we found that SVG enabled browsers are faster on the Microsoft Windows operating system; Apple OSX browser does appear a bit slower. Linux based browsers showed both a lack of stability and slowness but Mozilla Firefox is promising.

Querying by earthquake is done by both choosing the date on a list or by an interactive map.

Figure 2. DBMI04: query by earthquake example screenshot.

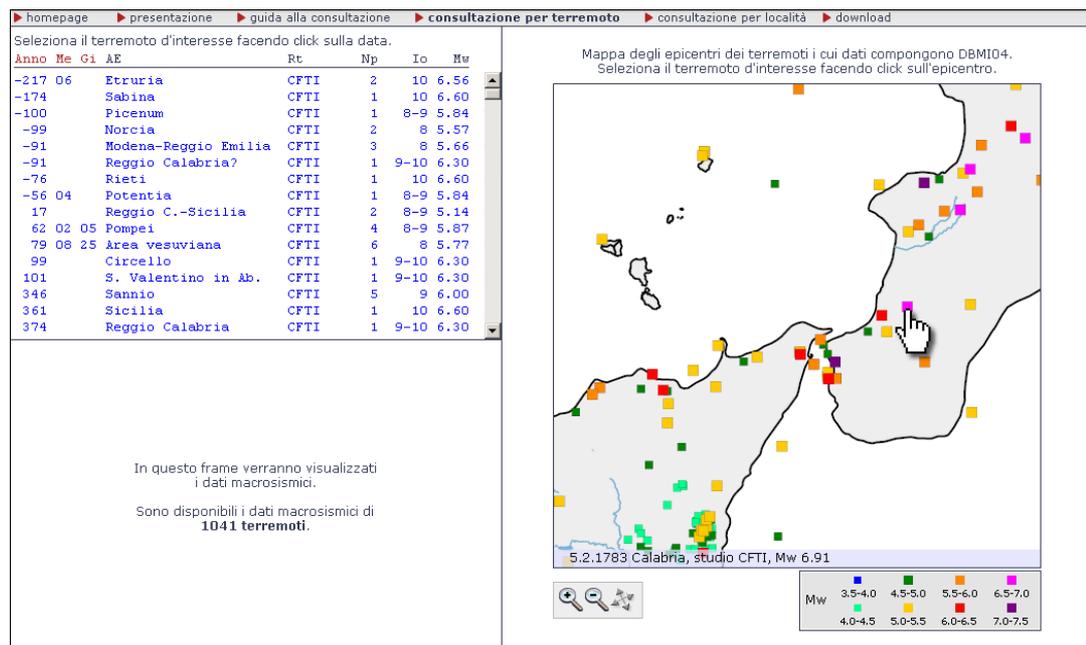
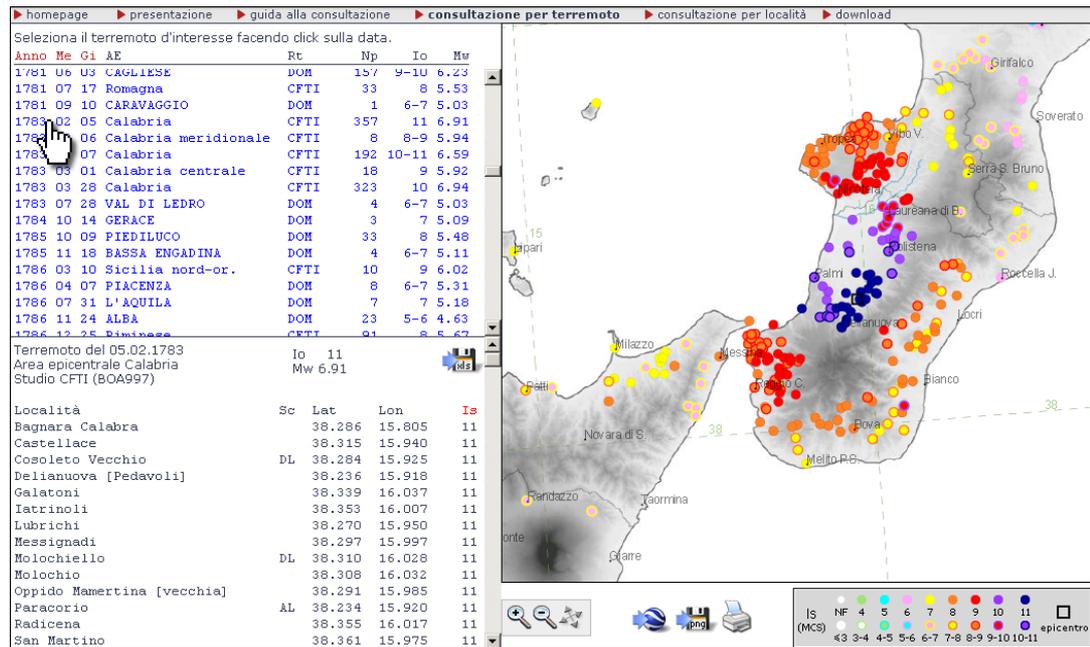


Figure 3. DBMI04: query by earthquake example screenshot.



Query by place, both by list in alphabetical order and by using an interactive map.

Figure 4. DBMI04: query by place example screenshot.

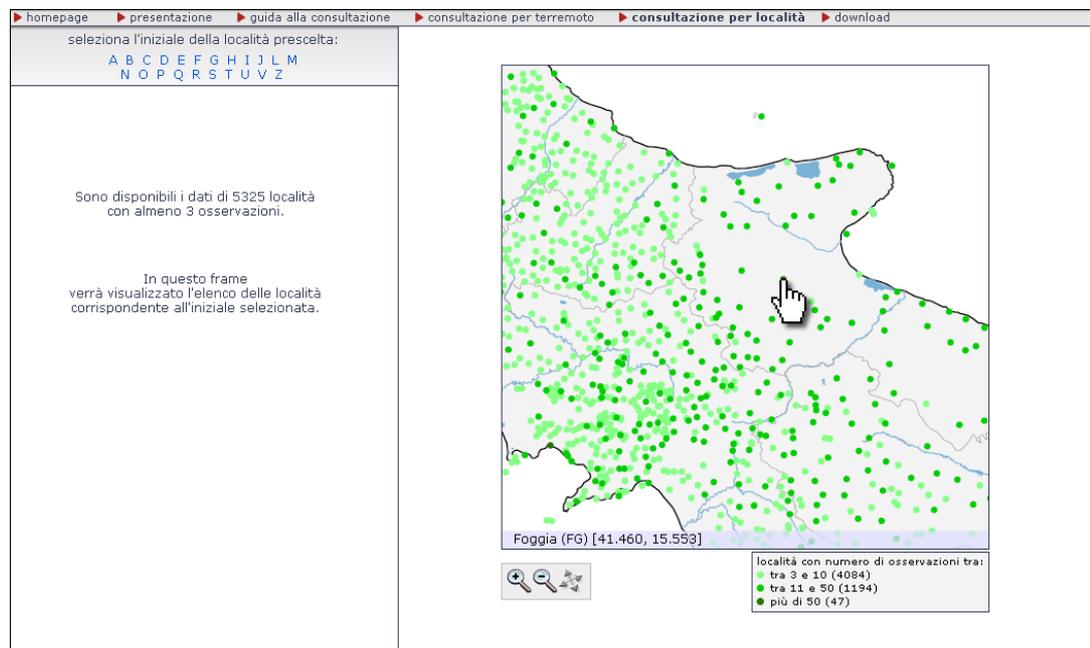
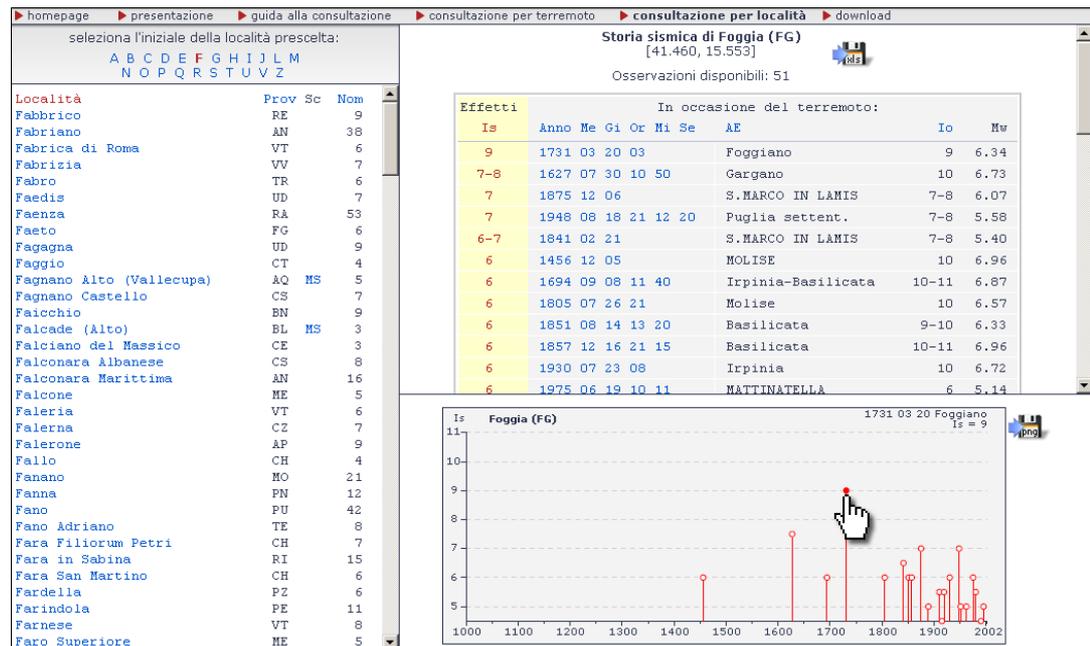


Figure 5. DBMI04: query by place example screenshot.



One of the problem that we faced was the used coordinate system. We chose the Universal Transverse Mercator (UTM) zone 32 based on the WGS84 ellipsoid for the whole Italian peninsula. The geographic coordinate conversion from latitude and longitude expressed in decimal degrees is done by internal code that uses equations USGS [Snyder, 1997] equations implemented in PHP.

Conversion of vector geographical elements in shapefile format to SVG code is done via the open source "shp2svg" utility (A. Neumann, 2007) [<http://www.carto.net/papers/svg/utls/shp2svg/>].

2008 update

This year a new version of the historical earthquake intensity database will be published and we are currently updating the web-mapping tool SVG code generator.

The most common criticism to the DBMI04 website implementation has been the dependency to a dying plug-in: the Adobe SVG Viewer. Nowadays this condition is out of date: browser such as Mozilla Firefox, Opera and Webkit based browsers have an internal SVG interpreter. The only browser that still does not support natively SVG is Microsoft Explorer in which a plug-in is still needed and luckily a new plug-in is surfacing for Explorer: the Renesis Player Browser [<http://www.examotion.com/>].

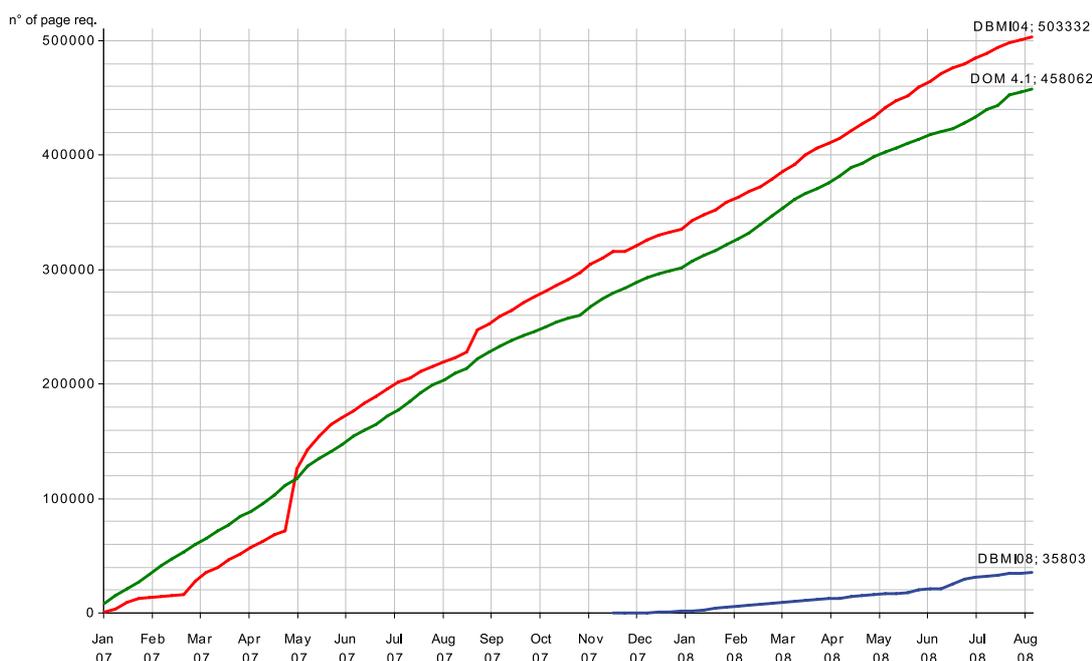
A revisione of our SVG generated code has been made and now zoom and pan function are possible in all browsers natively SVG aware. As a secondary results we found that responsiveness has increased as no time is spent on the client side loading the plug-in software: few seconds after the click and maps are flawlessly showed.

Another request to the old DBMI04 interactive maps was the missing scale bar which, for technical reasons was not implemented. We found this apparently easy-to-be-done request a quite tricky task because of the fact that SVG maps width and height are specified as a 100% of the available window frame area. We decided to invest time on solving this technical problem and now we have a working "overlay" layer which is not influenced by zoom, pan or the browser window resize.

Final users feedback

Feedback on the map consultation is considered the key factor: both the internal team and the public are encouraged to post a comment. Obviously we first evaluate advices from the seismologists' scientific community, but the public also influence our decisions.

Figure 6. DBMI: websites access statistics



Visitors of our web-site are mainly focus on the content, so comments are related to it; all the minor technical problems pointed out will be resolved with the final publication of the 2008 updated version in fall.

Gazetteer management tool

The historical earthquake intensity database stores a list of observed effects reported in specific places. The 2008 version of the Italian intensity database contains around 15000 different places, mainly in Italy.

Working with such a variety is only possible using a focused and well structured Gazetteer [Hill L., 2006]. Given the absence of an official Gazetteer for Italy, the Working Group started compiling its own more than ten years ago. The starting point was an already existing Gazetteer that offered both a good quality geographical co-ordinate resolution and reliable place names.

After years of integrations the complexity of the Gazetteer became unwieldy. Crosscheck, entry search or updating became not easy.

We therefore decided to invest a short time developing a centralised software always up-to-date for the whole Working Group.

Gazetteer structure

The Gazetteer (internally called "DIRip", Directory of places In Progress) contains common geographic field such as place name, coordinates, national, regional and local administrative information and is focused on the history of places.

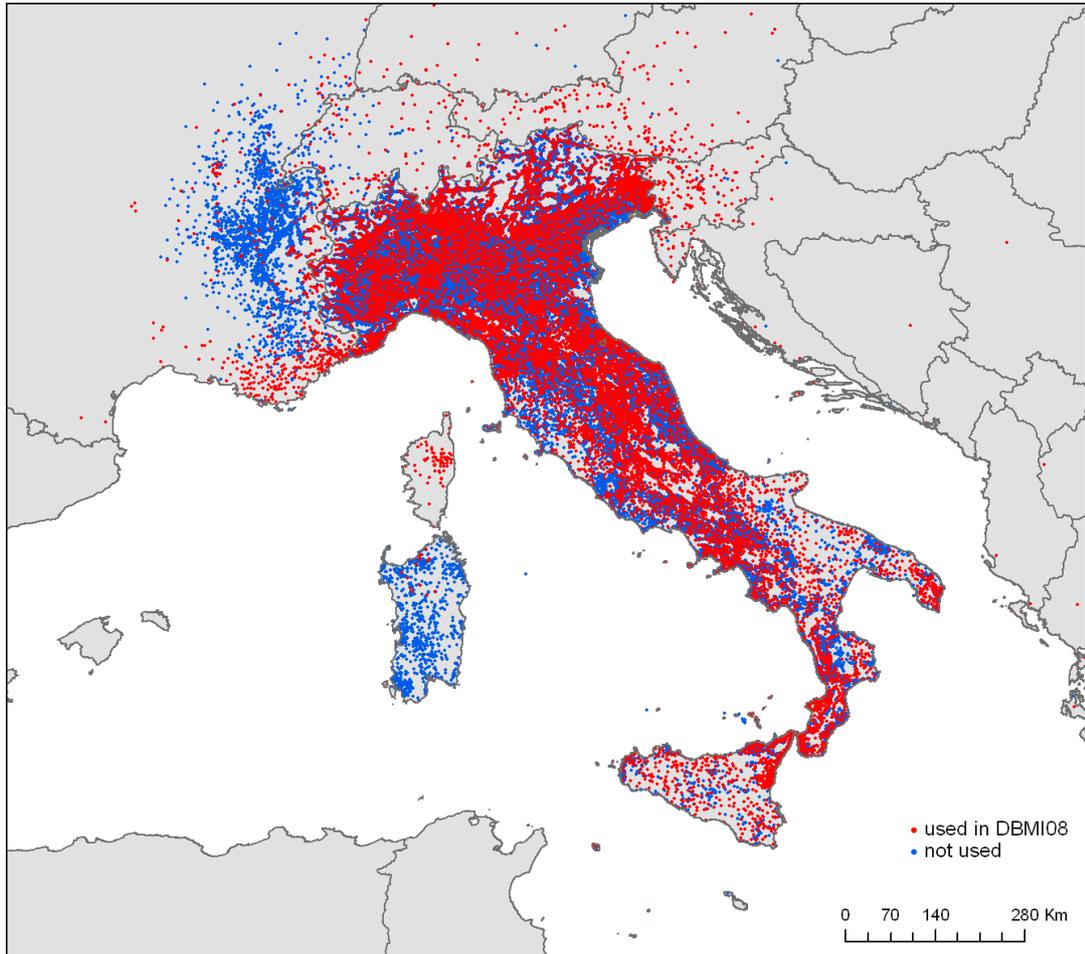
Other than simple place names and coordinates DIRip keeps trace of:

- abandoned places specifying why and when they were abandoned;
- place absorbed in others;
- small settlements, single and isolated buildings such as lighthouses, castles or monasteries;

- territories such as entire regions or valleys (vague damage description into original historical sources);
- historical place names and their related use time-span;

Nowadays DIRip contains 74963 places (71551 in Italy) of which 14690 are used in the DBMI version 2008.

Figure 7. DIRip content: places used in DBMI08 (in red) and not used (in blue).



Required functionalities

Historical earthquake studies presents a list of places found on the original sources; places are often described with vague, historical or local names. Historians are rarely GIS users or aware of instruments that can help finding coordinates and their resulting Studies often do not provide coordinates at all. In such cases its our job to identify the right place and assign coordinates. The process of recognizing place references in text and associating geospatial coordinates with them is called "geoparsing" [Hill L., 2006].

One of the most tricky problems while working with places is the name disambiguation process: many places have the same name or a similiar one. When ambiguous names refer to well know towns the problem can be easily solved; as regard small villages, or worst, villages abandoned centuries ago, the disambiguation can be very difficult.

The described tool is aimed at

- finding the geographical position of places, helping the disambiguation process with extra information;

- speeding up the geographical check of preliminary earthquake intensity distribution coming from digitized Studies;
- export the list of places in a well-formatted file for the historical earthquake intensity database;

The tool is not open to public access, is built and maintained only for the working group team.

The user interface is particularly friendly, winning and snappy as it must justify itself instead of using other existing software solution.

Development and technical details

Few issues have been faced with this tool, first of all the internal code method to update the SVG map without the need of reloading the entire page.

The real-time content retrieve via the getURL/postURL and XMLHttpRequest (A. Neumann, 2007) [http://www.carto.net/papers/svg/network_requests/] combined with the SVG ECMAScript capability of changing the DOM solve the crucial speed issue. Thanks to it SVG maps are updated instantaneously at each request.

Figure 8. DIRip: search by place name.

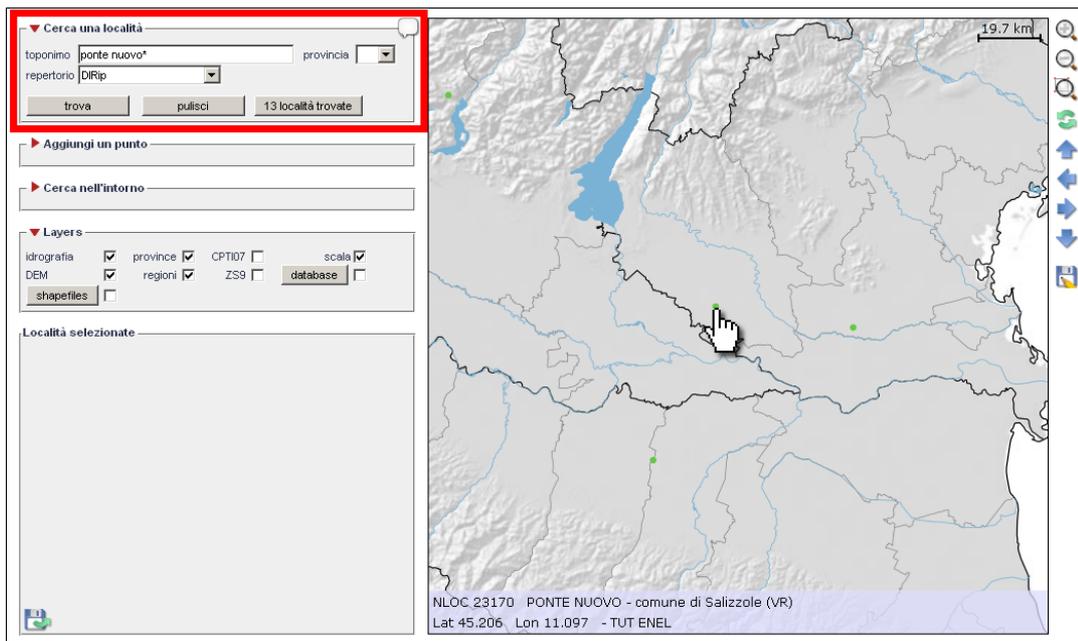
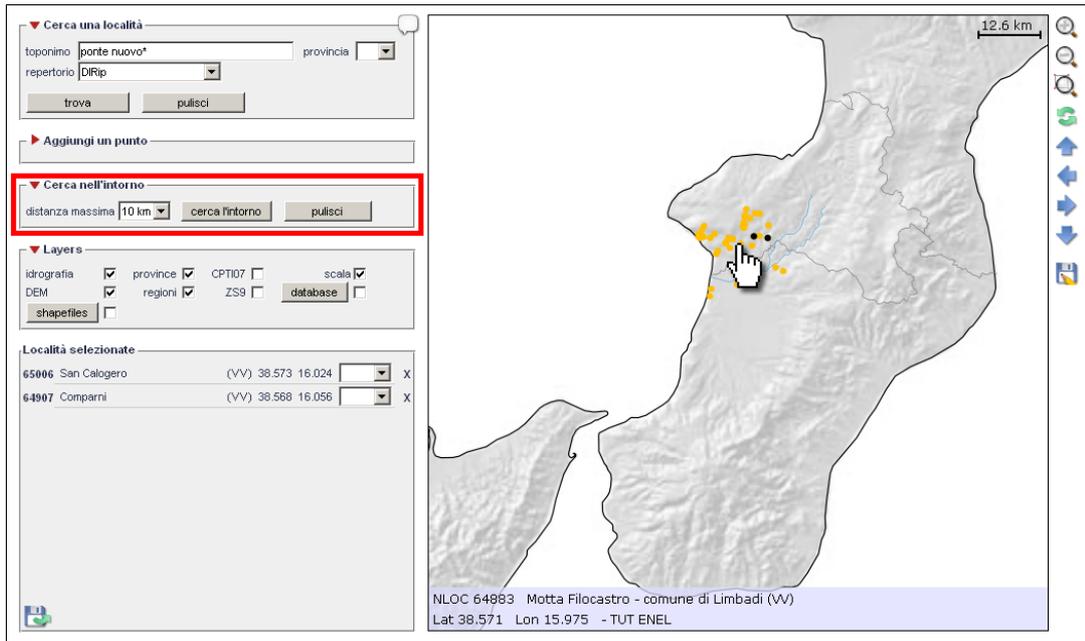


Figure 9. DIRip: search by area of interest.



Based upon this content update mechanism we develop each of the requested features: query by place name and query by area of interest. The query by name is a normal text query to the MySQL server; the query by geographical area is done via a great-circle distance equation [http://en.wikipedia.org/wiki/Great-circle_distance] translated in a simple SQL query.

Figure 10. DIRip: external shapefile loader.



The tool has a vector shapefile loader module based on the freely available PHP class "Shapefile Reader" (J. Gonzalez, 2004, revised by D. Granqvist, 2008) [<http://www.phpclasses.org/browse/package/1741.html>].

Conclusion and future development

This article presents our successful experience using the SVG graphic language to solve very focused tasks. Starting from thinking the final user experience, the technical implementation just followed as an irrelevant (well, sort of) detail.

Of course our implementation has a series of limitations, all coming from the fact that the system cannot communicate with other server; the web-mapping tool for example can only output files (SVG maps with high quality printing resolution, bitmapped images, Google Maps kml files and Excel tables), but this is what the working group does need. WMS /OGS compliant solutions such as ArcIMS or MapServer have no plus value in our daily job, and their customization comes with the price of complexity. Our implementation is fast, inexpensive and needs only software that is commonly available on research centre web servers (Apache and an geographically unaware database management system).

Our software solutions are not general purpose application, they are not meant to be flexible or easily adaptable to other purposes: we have very specific needs and we created very specific solutions. This is a very common practice in the scientific research world because of the extremely resource limitations both from the economic and from the human resource point of view. These tools are nothing else than small pieces inside the scientific core research, they are not meant to be sold to the public.

Entering an European Commission funded project forced us to generalize much more our implementation and a major revision is planned. European partners are asking us to deliver the web mapping software adapting it to their specific needs. One of the most requested feature is to output a complete static website: a series pre-generated html files that doesn't need any server side scripting language nor online database interaction; this request comes from the extremely restrictive research centres policies in terms of web server use.

An open source license is planned and documentation is in preparation.

Future development is concentrated on study and experimenting web-service based interfaces plugged into our online infrastructure.

References

- [CPTI04 Working Group, 2004] Paolo Gasperini, Romano Camassi, Carmen Mirto, Massimiliano Stucchi. "CPTI04, il Catalogo Parametrico dei Terremoti Italiani, versione 2004". Istituto Nazionale di Geofisica e Vulcanologia, 2004. <http://emidius.mi.ingv.it/CPTI04/>
- [Giesecke et al., 2004] Alberto Giesecke, Augusto Antonio Gómez Capera, Ilaria Leschiutta, Elena Migliorini, Leandro Rodriguez Valverde. "The CERESIS earthquake catalogue and database of the Andean Region: background, characteristics and examples of use". *Annals of Geophysics*, 47. Istituto Nazionale di Geofisica e Vulcanologia, 2004. <http://www.ceresis.org/>
- [Grünthal et al., 1998] Gottfried Grünthal. "European Macroseismic Scale 1998, EMS-98". Centre Européen de Géodynamique et de Séismologie. European Seismological Commission, 1998. http://en.wikipedia.org/wiki/European_Macroseismic_Scale
- [Hill, 2006] Linda L. Hill. "Georeferencing: the Geographic Associations of Information". The MIT Press, Massachusetts Institute of Technology, 2006. <http://www.informaworld.com/smpp/content~content=a790613367~db=all~jumptype=rss>
- [Monachesi and Stucchi, 1997] Giancarlo Monachesi, Massimiliano Stucchi. "DOM4.1, an intensity database of damaging earthquakes in the Italian area". Consiglio Nazionale delle Ricerche, 1997. <http://emidius.mi.ingv.it/DOM/>
- [Montaldo et al., 2007] Valentina Montaldo, Carlo Meletti, Francesco Martinelli, Massimiliano Stucchi, Mario Locati. "On-line seismic hazard data for the new Italian building code". *Journal of*

Earthquake Engineering, 11. Taylor and Francis, 2007. <http://www.informaworld.com/smpp/content~db=all~content=a777589695~tab=content>

[MPS Working Group, 2004] MPS Working Group. “Redazione della mappa di pericolosità sismica prevista dall'Ordinanza PCM 3274 del 20 marzo 2003. Rapporto Conclusivo per il Dipartimento della Protezione Civile” . Istituto Nazionale di Geofisica e Vulcanologia, 2004. <http://zonesismiche.mi.ingv.it/>

[Scotti et al., 2004] Oona Scotti, David Baumont, Grégory Quenet, Agnès Levret. “The French macroseismic database” . *Annals of Geophysics*, 47. Istituto Nazionale di Geofisica e Vulcanologia, 2004. <http://www.sisfrance.net/>

[Snyder, 1997] John P. Snyder. “Map Projections: A Working Manual” . U.S. Geological Survey, 1997. <http://pubs.er.usgs.gov/usgspubs/pp/pp1395>

[Stucchi et al., 2007] Massimiliano Stucchi, Romano Camassi, Andrea Rovida, Mario Locati, Emanuela Ercolani, Carlo Meletti, Paola Migliavacca, Filippo Bernardini, Raffaele Azzaro. “DBMI04, il database delle osservazioni macrosismiche dei terremoti italiani utilizzate per la compilazione del catalogo parametrico CPTI04” . *Quaderni di Geofisica*, 49. Istituto Nazionale di Geofisica e Vulcanologia, 2007. <http://emidius.mi.ingv.it/DBMI04/>

[Swiss Seismological Service, 2002] Domenico Giardini, Donat Fäh. “ECOS, the Earthquake Catalog of Switzerland” . Swiss Seismological Service, 2002. <http://histserver.ethz.ch/>