DOI 10.4430/bgta0078

# The Geo-Seas seismic data viewer: a tool to facilitate control of data access

P. Diviacco and A. Busato

*Istituto Nazionale di Oceanografia e di Geofisica Sperimentale, Trieste, Italy*

(Received: May 31, 2012; accepted: July 20, 2012)

***ABSTRACT*** Data transparency could be used as a metaphor to describe how data owners can give controlled access to their data to end users. In a scientific environment and focusing on seismic data, this is a very delicate activity that needs to consider the desire for data owners to position themselves in the research community, possible legal issues and the commercial incomes that data could provide. In this perspective, this work focuses on the possibility of having various levels of access to data so that its opacity can be progressively removed according to the contractual status with the end user. Within the EU FP7 Geo-Seas project, an interactive web-based seismic data viewer tool has been developed in order to provide these features. End users work within secured shells contained in the viewer without downloading any data. This relieves the concerns of data providers of losing control of their data, fosters their collaborative attitude and ultimately increases the availability of data within the community.

**Key words:** seismic data viewer, data protection, data dissemination.

## 1. Introduction

The Geo-Seas [Geo-Seas] initiative is implementing an e-infrastructure of 26 marine geological and geophysical data centres, located in 17 European maritime countries, that will enable users to identify, locate and access pan-European, harmonised and federated marine geological and geophysical data sets, and derived data products, held by the data centres through a single common data portal.

Similar initiatives in many scientific research fields reveal the importance of data not only in the process of scientific reasoning and discovering, but also in its social aspects. Modern epistemology has undermined the theoretical framework of the traditional view of Science, shaping a new line of thought that connects Duhem, Lakatos, Latour and other authors [see Diviacco (2012) for further insight into this issue].

Following this view, Science is not a "cold", mechanical, "un-debatable" process, instead, it is profoundly conditioned by human factors, and since researchers are part of communities, as in Latour and Woolgar (1979) *"Science is a social construct"*.

In this perspective, researchers, institutions, funding agencies, but also instruments, tools and software contribute, as actors, to the success or failure of a research path.

Concurrent paths, schools and single scientists are committed to the development and successful strategic positioning in the community, of their visions and of themselves as an *"obligatory* *passage point"* (Latour and Woolgar, 1979). This can grant them the control of the evolution of ideas and allow them to enter new research projects.

This goal can be pursued whilst balancing at the same time (Diviacco, 2012):

- the need to preserve the scientist's original position;
- the need to be at the cutting edge of innovation;
- the need for other partners and their resources.

Where the control of data is an extremely important factor, this can be a very complex task.

Many studies have addressed this issue and demonstrated that this is a general problem that can be traced in many research fields. In the case of medicine, concerns about releasing data within large collaborative initiatives have been reported by Olson *et al.* (2008). Pratt *et al.* (2004) highlighted the different sharing attitudes of junior and senior, or higher ranking physicians. Nentwich (2003) reports concern for data transparency in fields that are close to economic applications and Kötter (2001) highlights the economic interests of private companies in molecular biology. Orlikowski (1992) highlighted the strong need for motivation in data sharing which, in the case of scientific research, means essentially the possibility of publishing; for example, in the case of high energy physics, Birnholtz (2006) reports on the practice of acknowledging all participants of a project as authors in publications. Ribes and Bowker (2009) report on the possibility that scientists can be ashamed of the quality of their data and thus unwilling to share it with their peers.

At the same time, but in the opposite direction, the Open Data movement pushes the idea that data should be freely available to everyone to use and republish as they wish without restrictions.

## 2. Data transparency/opacity

Data transparency or its converse, data opacity, is a metaphor used to denote the faculty that a data provider grants to an end user to access the knowledge contained in some data.

The access can be seen as moving through layers that progressively allows the end user to extract more and more information from the data. These layers, sorted from higher opacity (lower transparency) to higher transparency (lower opacity) are (Fig. 1):

- **data discovery:** where the end user identifies the data that could match his/her needs, generally from a geo-spatial/temporal perspective;
- **data browsing:** where the end user browses the characteristic parameters of the data sets he/she is focusing on, as for example sampling rate or processing history, to verify if the data could address the issues under consideration;
- **visualisation:** where the end user is first enabled to access the data, viewing them without the possibility to work on them and without having a copy of them;
- **usage:** where the end user is enabled to work remotely on the data but does not have a copy of them;
- **download:** where the end user obtains a copy of the data and therefore is restricted in its usage only by possible legal agreements with the provider.

In extreme cases data discovery itself can be made completely opaque to avoid that competitors could take advantage from knowing the activities of the provider, while on the other extreme, data can be released for download without restrictions, opening all the knowledge contained within data.

The first two layers are generally referred to as metadata (an abstraction that describes the data) and are intended to allow users to only select the data in which they are interested in, while data, of course, is where the actual observation resides.

Metadata based data selection can be a rather complex task to perform especially when the re-use of data is different from that for which it was originally acquired. Data quality in these cases, for example, is an almost impossible parametre to capture. The border between metadata and data often needs, therefore, to be osmotic and in the quest to understand what data they need, users will often need to delve into the data themselves.

## 3. Data transparency modulation

Crossing the border between metadata and data is not a simple task. On one side, we have a technical problem related to the fact that two different worlds are addressed, and on the other side we have the much more challenging issue of providing a solution that could cope with the social and economic implications of data transparency.

These latter are related to the attempts of data owners to attract funding from commercial activities or improving their position in the scientific community, as mentioned in the introduction, asking to participate in new projects in exchange for the access to data.

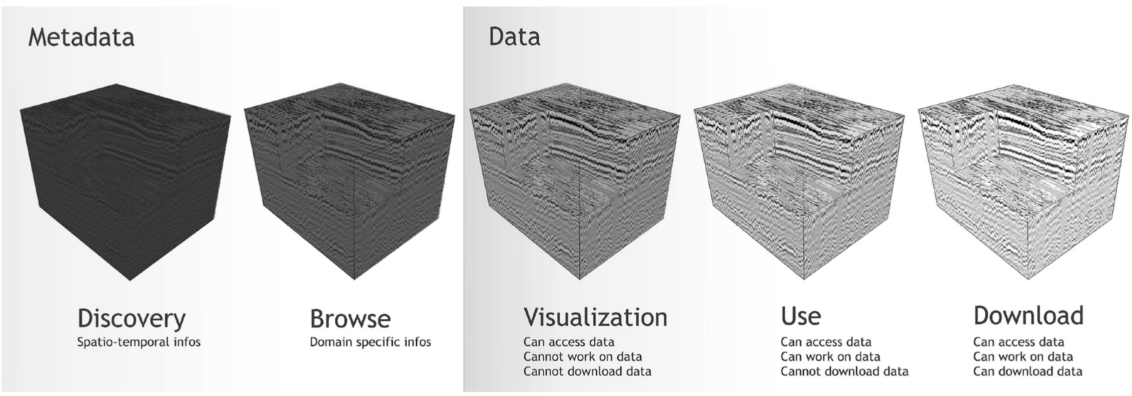


Fig. 1 - The metaphor of data transparency can be seen as increasing from discovery to browsing, visualisation, use and download [data from Wardell et al. (2002), courtesy of the authors].

In this, data providers should be able to modulate data transparency in order to define the depth which users are allowed to reach. Usually this is set within an agreement between the provider and the user and can change with time, depending upon how the collaboration evolves.

## 4. Legal issues in data transparency

The modulation of data transparency can be limited by legal issues, as for example in the case of geophysical data by mining exploration licensing regulations, or in the case of biomedical research by privacy issues, such as, for example, in the United States with the HIPAA federal regulation [HIPAA].

Legal aspects of data transparency can vary largely not only from field to field but also from nation to nation; it would be impossible to address here all the details of this problem.

In the EU, the 2003/98/CE directive [2003/98/CE] defines a non-obligatory general framework for the re-use of data of the public sector that takes into account the aim of avoiding cross-subsidies and introduces the idea that all member states should encourage public sector bodies to make documents and data available at marginal costs. If this holds in the case of information produced by most of the governmental sources it has a less direct application to scientific research, and in fact the 2003/98/CE directive explicitly mentions that it shall not apply in this case (Art.1, Clause 2.c). Focusing on geo-referenced data, the 2007/2/EC directive, better known as INSPIRE [INSPIRE], states that geographic information needed for good governance at all levels should be readily and transparently available. Its annexes defines the spatial data themes where this has to be applied, as land use or cadastre, while leaving as more “debatable” its application in other fields as, for example, Geophysics.

In the United States, Section 105 of the 1976 Copyright Act [US Copyright Act] prohibits the federal government from claiming protection in its publications. A large portion of the data and information thus produced in government programs automatically enters the public domain, year after year, with no proprietary restrictions. Much of the material that is not made available directly to the public can be obtained by citizens through requests under the Freedom of Information Act, although costs of dissemination can be not only marginal but also incremental (Office of the Management and Budget Circular A-130 [A-130]), which can create barriers to access, particularly for academic research.

In case strict legislation imposes open dissemination of the data, often, data transparency modulation is obtained by distinguishing raw data from a data product. In fact, with the latter being produced by the intellectual work of specialists, it is eligible for intellectual property protection.

This solution, however, does not satisfy all data providers because the availability of raw data generally allows end users to bypass completely the need of including providers in their research activities. This, as mentioned in the introduction, can, in fact, work against the need of the provider to position himself in the community as an obligatory passage point.

## 5. Examples of data policies

An example of practical implementation of a data licensing policy that supports scientific research whilst at the same time being limited by legal issues is that of the Antarctic Seismic Data Library System [SDLS] (Diviacco and Wardell, 2003). In this case, since the Antarctic Treaty requires that scientific data be exchanged and made freely available [Antarctic Treaty], seismic data are submitted to the SDLS within 4 years of collection and remain in the library under SDLS guidelines until 8 years after collection. Thereafter, the data are available from World Data Centres or equivalents for unrestricted use. In the domain of volcanology a similar approach has been used by the MULTIMO project (Carniel *et al.*, 2006) that offers open access to the data gathered during the project.

Considering these examples, we would like to highlight the importance of a continuous involvement of a funding agency that sponsoring the maintenance of the software and of the database avoids the obsolescence of the systems and keeps it at the centre of the designated scientific community.

Another example of data licensing is that of the Landsat system. This has been based on several considerations (Landsat Remote Sensing Policy act [LRSP]) made by the United States Federal Government, among which (I) the difficulty in full commercialisation of the Program, (II) the possible barrier to scientists in using the infrastructure due to its costs, and (III) the national interest in maintaining international leadership in satellite land remote sensing. These considerations led to the definition of a data policy where, albeit retaining its full ownership, the United States government provides un-enhanced data at the cost of fulfilling user requests, on the condition that such data are used for non-commercial purposes only, and charges a per image fee, license fee, or other such fee to non governmental entities operating ground receiving stations or distributing Landsat 7 data.

An example of fully open data policy is that of the Italian Istituto Nazionale di Statistica that offers freely its data assets under the Creative Common Licence [CC].

## 6. Data policy within the Geo-Seas project

As mentioned in the introduction, the Geo-Seas project is a collaborative initiative built upon the efforts of a large number of institutions from different countries. This means that the project has to cope with data policies that can vary from country to country, but can vary even among different institutions within the same country.

In practice, this means that trying to impose one single data policy is not only difficult but also not advisable, because some partners could feel uncomfortable with the imposed one and be tempted to leave their data outside the project's shared data space. This would trigger an "empty box syndrome" where an otherwise supportive infrastructure would remain a container without content (the data).

Geo-Seas, instead, has been committed to developing a data policy solution aimed at maximising the availability of data to the community at large, thereby ensuring that their maximum value can be realised and thus contribute to an increased understanding of the marine environment, but, at the same time preserving the rights and interests of providers.

It was therefore decided to try as much as possible to develop a solution that can be easily tuned to match the data policy requests of any partner.

Metadata is freely available without any condition, while data and products require (I) registration; (II) acceptance of additional conditions requested by the provider, which can depend also on the role of the user as, for example: commercial, non commercial, academic etc. [list C866 of BODC controlled vocabulary]; (III) acceptance of a license.

In this perspective, data is not directly made accessible to the end user and the level to which access is granted can be set by the provider. The solution is provided by a specific tool that has to mediate the requests of all actors whilst, at the same time be integrated with the rest of the Geo-Seas infrastructure.

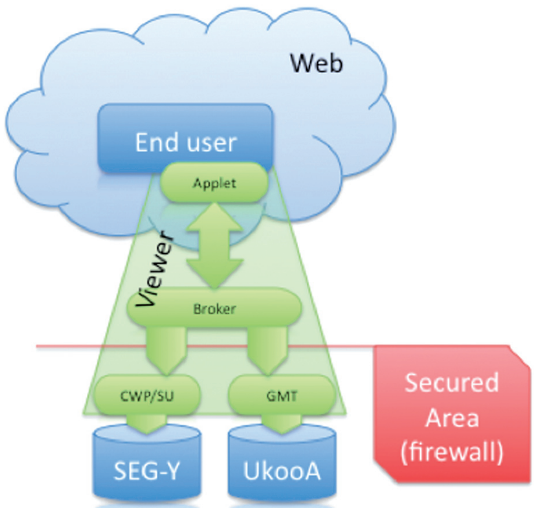


Fig. 2 - The actual data being within a secured area can be accessed only via the seismic viewer.

## 7. Extending an infrastructure

We had further constraints: the GeoSeas project is a sibling of the SeaDataNet European Commission Framework Programme 6 project [SeaDataNet]. Where the two projects overlap we decided to adopt several technical solutions developed within the latter, whilst new tools have been developed from scratch, or innovative ideas have been tested, where they diverge. The constraints that directly relate to data transparency are the Common Data Index (CDI) and the Download Manager/Request Status Manager (DM/RSM).

In SeaDataNet, the CDI provides an insight into the availability and spatio-temporal coverage of marine data archived at the connected data centres. It is an ISO19115 [ISO19115] metadata profile and ISO19139 [ISO19139] XML encoding designed to describe an individual ‘deliverable data object’: a file or group of files containing the data from a single instance of a feature type such as a profile, point time series or trajectory.

The Request Status Manager (RSM) service oversees the processing and administration of all requests for data sets. It handles for the users the communication from the SeaDataNet portal to the distributed data centres where the Download Manager (DM) handles all communications between the data provider and the RSM. After submitting data set requests via the CDI, users receive a confirmation e-mail of their data set requests and a link to the RSM service.

Within this paradigm, discovery is done only through metadata and data are eventually accessible only via downloading. Obviously, considering the premises outlined above this was not acceptable and a new vision had to be introduced.

Considering the wide spectrum of data types that the Project was aiming to handle it was decided to focus first on seismic data and to later extend the solution to other data types.

## 8. The Geo-Seas seismic viewer

The Geo-Seas seismic data viewer is at the same time the connection and the separation between the end user and the actual data. Referring to Fig. 2, the data resides within a secured area while the end user can be anywhere on the web.

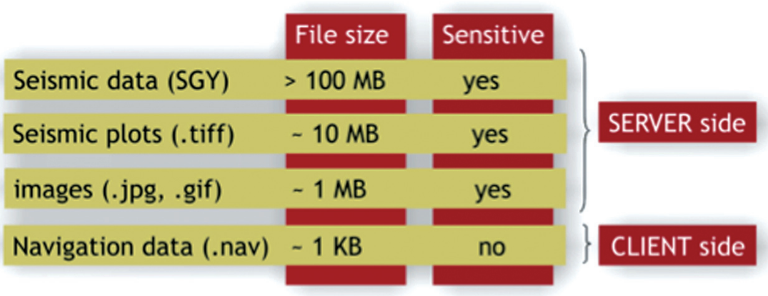


Fig. 3 - Sensitivity and dimension of various seismic data products.

The end user issues a request that can be handled only by the viewer. This latter, working on the actual data produces a result that is sent back to the end user for evaluation. Therefore, the viewer is the only bridge between the end users and the data, while the latter are kept away from the web.

To understand the functionalities that the viewer would need, the Geo-Seas team developed a questionnaire that was sent to a wide range of end users and providers and the results have been carefully analyzed.

At the same time, existing solutions with similar goals have been considered. Most of them are commercial and cannot be adopted within the project to avoid imposing further costs on partners. Some of these partners have developed internal solutions (Ifremer, OGS's SNAP system, GEUSS) but these do not have all the needed features, so that eventually it was decided to develop a new tool from scratch.

From the end user point of view, the questionnaire highlighted the need of tools that would allow the actual data to be visualised and some basic processing to be performed, mainly for quality control reasons. A map, representing interactively the position of a cursor in the viewing area of the seismic data should be made available, while panning and zooming this latter should produce consistent changes in the former.

From the data provider point of view, users should be identifiable singularly and it should be possible to apply limitations (see later) to the modalities of access to the data.

Visualisation of seismic data is not a trivial task; it has to take into consideration the fact that four data types can be found in this field (Fig. 3): (I) actual digital data in SEG-Y format [SEG-Y], (II) images scanned from paper sections generally as Tiff files, (III) other generic lower resolution images that can be available in formats as .jpeg or .png, (IV) the navigation file. These data types have different levels of sensitivity and different file sizes; SEG-Y files are very large and have to be protected, whilst navigation files are very light and generally considered non-sensitive.

When file sizes are large and data are sensitive, transferring large blocks of data sets back and forth through the web is not advisable, both for the risk of bandwidth overload and of possible unauthorised data use. It is therefore better to handle them within a server-side approach (Diviacco, 2005), where the original data-type would be transformed to one which can be handled more easily and securely, as, for example, converting the actual data to an image; in our case to pyramidal tiff files.

Pyramidal tiff files (Pitzalis *et al.*, 2006) wrap a sequence of bitmaps that each represents the same image at different resolutions and where images can be tiled. Each of these different resolution images can be associated with the action of zooming on data, while different tiles can be visualised upon panning the data when the section exceeds the dimension of the viewer window. The pyramidal tiff file is much lighter than the SEG-Y file and its content can be defined and limited upon the user access license. The same solution can be extended to other sensitive data, such as scanned paper section images and generic seismic data images.

On the contrary, if navigation data is very light and generally non sensitive, it would be better to render maps locally on the client browser, unloading the server and avoiding traffic on the net; in our case (Diviacco, 2006) coding geographical features and positioning using Scalable Vector Graphics (SVG).

SVG is an W3C XML based open standard file format for two dimensional vector graphics that allows to represent objects as vectors, not only as bitmaps. This allows web links, scripting functionalities and pop-up messages to be associated to each of them that will then be preserved when rendering the map. Using SVG graphics, objects, being vectors, do not suffer from pixelation and rendering, and being done directly at the client side, is very fast with any decent video card.

The proper solution for handling seismic data and its navigation at the same time should then be based on a mixed server-side client-side paradigm. How can this be practically implemented?

## 9. The seismic viewer engine

The Geo-Seas seismic viewer tool is perceived by users as a single tool, but actually it is composed of several bricks devised to work together (Fig. 2).

Upon connecting to the URL of the dataset [see Diviacco et al. (2011) for further insight into the Geo-Seas infrastructure metadata model and how it allows users to access data] the end user receives an applet that, running in the user's browser, collects the interactive requests of the user and sends them to a JMS broker. This understands the request and activates a service residing beyond the firewall in the secured area (as are the data themselves), which will process it. Only this service can access the data and respond directly to the broker.

If some seismic data processing is required, the service launches CWP/SU [CWP/SU], an open source vertical application. The data are processed and the output is converted to pyramidal tiff files.

If mapping is required the process launches GMT [GMT] which reads the UKOOA navigation files [UKOOA] of the seismic line and produces an SVG map.

Results are collected by the broker and sent back for visualisation to the applet.

It is important to highlight that the role of the applet in this loop is rather complex. In fact the results sent back from the broker are not static images only, but pyramidal tiff and SVG files that allow the user to perform zooming and panning on both. A new loop is triggered only when what is required is not contained within the results from the previous call.

Besides, the applet complexity is increased also by the fact that it embeds a module called Batik [Batik] that allows the user to visualise SVG files directly within the viewer. Up to a couple of years ago, most of the web browsers did not support SVG natively so that a specific plug-in was necessary. Adobe offered one that ran within Windows and Mac systems, but once most of the browsers started supporting SVG natively, the maintenance of the Adobe SVG plug-in software was discontinued. However, since the native support of SVG is not at same level for all the browsers, this creates quite a lot of confusion and the use of SVG based mapping is sometimes perceived as complicated. To keep usage of maps seamless it was therefore decided to embed directly into the viewer the tool to render locally the SVG maps.

## 10. Functionalities of the viewer

The active area of the viewer applet can be divided in three sections (Fig. 4):

- the Seismic section: where the seismic data is visualised and where by moving the cursor a box reports the position in time and trace number of the sample under the cursor;
- positioning: where an SVG map of the track of the seismic line is plotted;
- the toolbox: where it is possible to find all the buttons related to the processing and visualisation functionalities that the viewer applet offers.

A very important feature of the viewer is that the positioning and the seismic section area are synchronised, so that moving the cursor in one moves the cursor also in the other. This, of course, cannot be demonstrated here, while Fig. 5 tries to represent that the position of the cursor in the section and the circle on the track in the map are consistent.

The toolbox section contains all the buttons that allow the end user to access the functionalities of the viewer applet. As in Fig. 6 from left to right it is possible to identify the zoom/pan button that must be activated in browsing the section. The following dimmed button links to a set of collaborative functionalities that are currently under development but pertain to a wider scope than that reported in this paper. This button is currently disabled.

A very useful tool that the viewer applet offers is the “on the fly” gain control (Fig. 7) which is to the right of the dimmed button. This allows the moving of the mean value of the data in the colour scale, so that the contrast between high amplitude and low amplitude values can be changed.

This allows subtle signals to be highlighted even if they would have been hidden by larger events or noises, or to highlight larger signals only in order to have an overview of the main events in the data. To the right of the gain button is the “all section” presentation changing button, and the “boxed” area presentation changing button which allows wiggle plots to be created. These last two were separated to prevent users from accidentally producing large wiggle plots of the entire section.

The viewer applet allows multiple manipulation of the seismic data section presentation. Several colour scales can be used and a specific function is reserved for wiggle plotting. In this case, it is possible to select a portion of the section, dragging the mouse to select a box, to be submitted to the JMS broker. The selected area is reconstructed from the actual data and the resulting image is sent back to the applet but to a different layer. In fact, the viewer allows multiple images to be present at the same time in the viewer so that it becomes possible to switch from one to the other without the need to recreate them. These images are called layers, since each of them can be zoomed, panned and gained, while the first layer is able to represent all of them, overlaid one over the other, with the correct geometric and gain scaling.

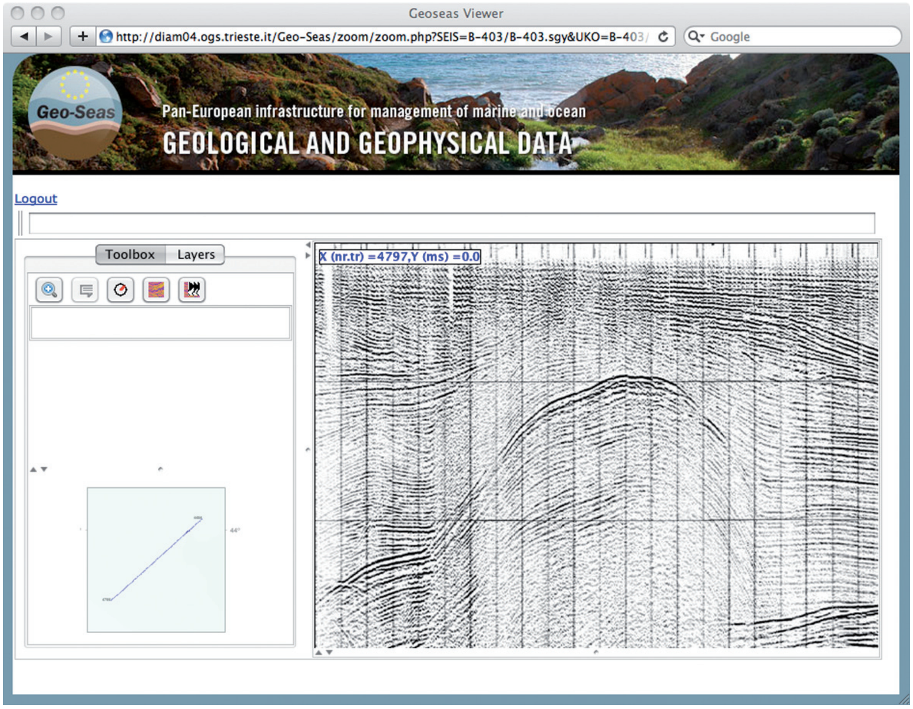


Fig. 4 - The seismic viewer active area can be subdivided in three sections: toolbox, Positioning and the seismic section itself.

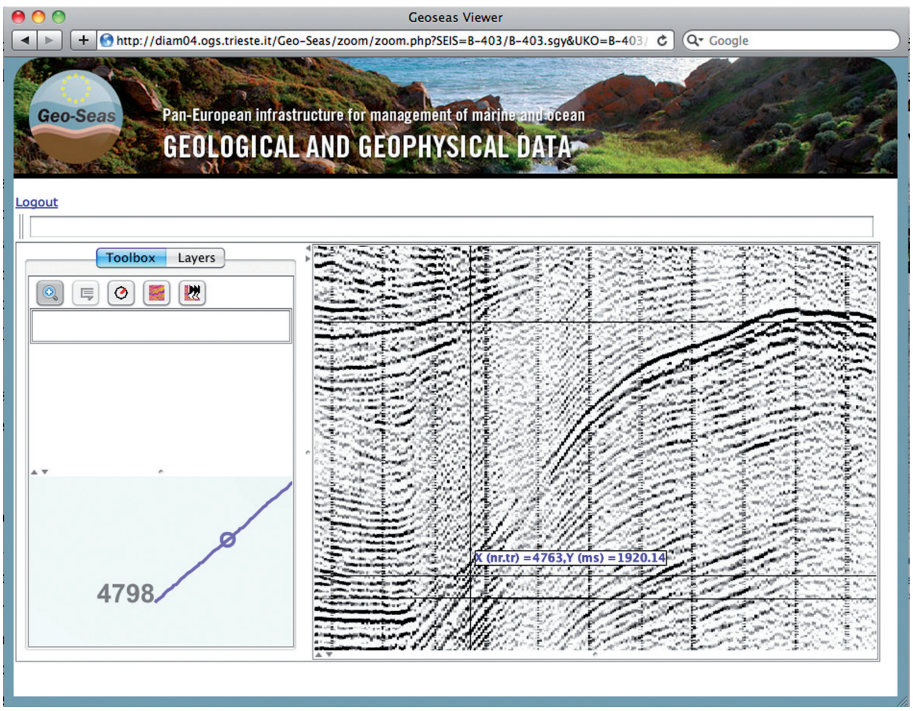


Fig. 5 - Positioning and the seismic section are synchronized so that the circle on the track represents the position of the cursor in the seismic section.

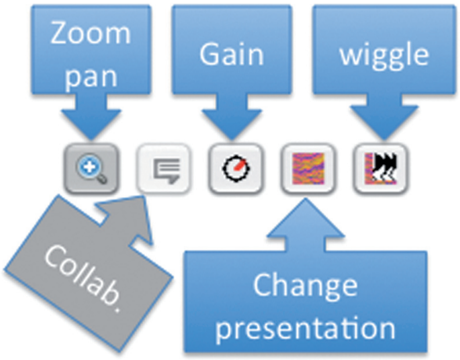


Fig. 6 - The seismic viewer applet's toolbox.

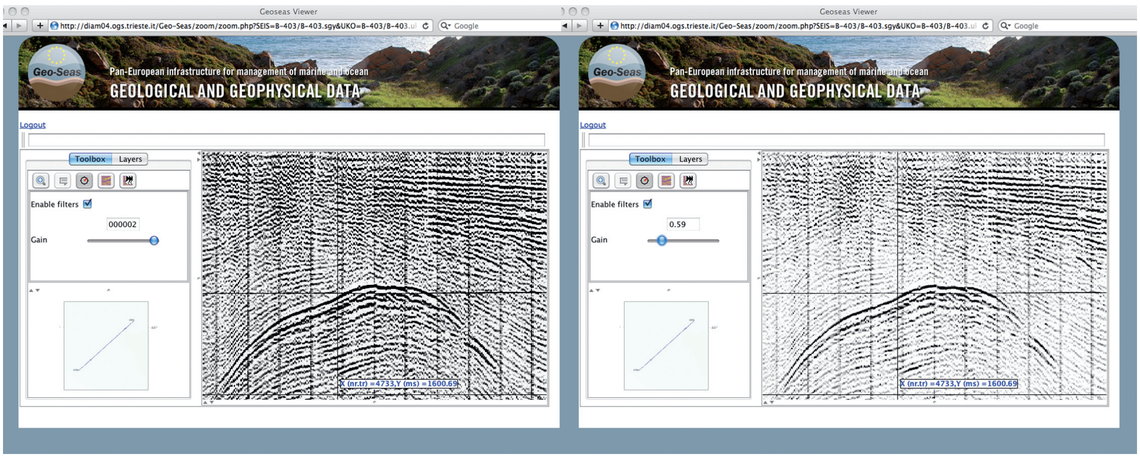

Fig. 7 - The Gain tool allows the highlighting of subtle features that would be hidden otherwise.

## 11. Data transparency modulation

As mentioned previously, from the data provider's point of view it is very important to control the access to the data so that different end users could be granted different data transparency levels for the same data set. Since we are discussing the viewer here, we are not addressing metadata anymore but the access to the actual data, then how can this access be limited in practice?

There are essentially two ways:

1) watermarking: this corresponds to the superposition of some text or logo to the images sent back from the broker, in order that any user might want to use them they will need to acknowledge the source;
2) lower resolution: this corresponds to limiting the number of levels in the pyramidal files sent by the broker, in order that the sampling of the displayed seismic data is lower that the original one.

The first case is rather oriented towards scientific publications; in fact any end user that might want to publish the data without permission will find it impossible to remove the text or logo from the image. The second case is more oriented toward end users that might want to directly work on the data. Without the proper resolution they will not be able to obtain what they need.

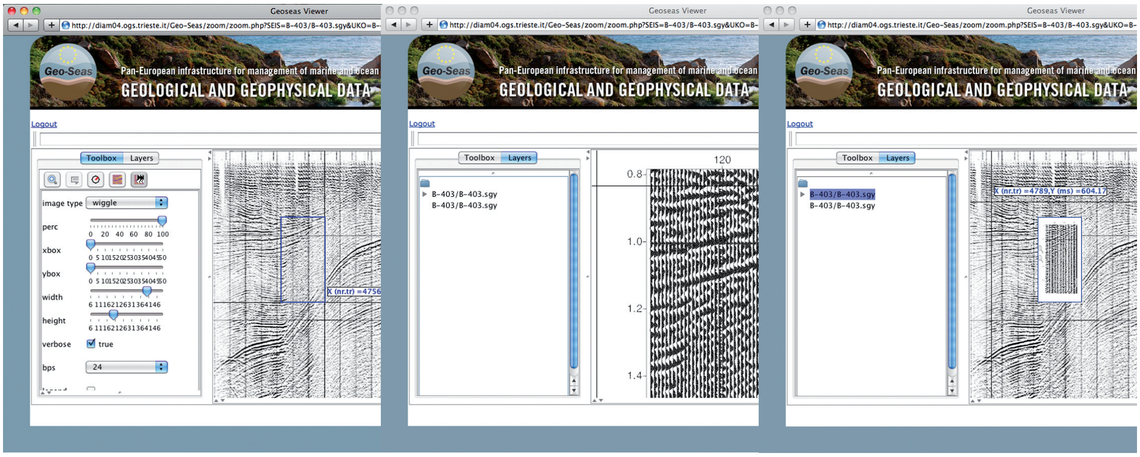


Fig. 8 - Changing the presentation of a portion of the seismic section: selection of the area (blue box in the leftmost image), the wiggle image sent back by the broker and referred to as a layer (centre) and the box overlaid on the full section.

A third option exists, where a limitation in transparency is lifted for a limited period of time. This case is generally not common or well accepted, since it grants end user freedom without control.

In the perspective of transparency modulation, each user has to be identified, and therefore has to log in upon connection while a database loads the matrix that links identities, permissions and data sets, that has been set by the data provider.

At the moment, within the Geo-Seas project, identification, transparency modulation and the possibility to orchestrate all this process from a single point of access so that the end user would perceive all the federated system as a single data space is currently under study and will be the subject of a specific oncoming work.

## 12. Concluding remarks

Data availability is recognised as one of the most important factors in the advancement of a scientific community. Several factors can influence the wish to open a data provider's archives and need to be carefully considered to avoid resulting in a generic retentive mental habit. These concerns can be removed by providing data owners with means to control end users accessing data and avoid downloading. This is the goal of the Geo-Seas seismic viewer tool, which keeps sensitive data unreachable from outside while acting as a bridge between them and the end users. Very soon, these latter will be identified one by one, so that to each of them a different data access policy could be assigned, and a progressive modulation of data transparency could be obtained controlling watermarking, resolution and the availability of functionalities.

**Acknowledgements.** The authors would like to thank Nigel Wardell for his revisions and helpful and constructive comments.

## REFERENCES


Birnholtz J.P.; 2006: *What does it mean to be an author? The intersection of credit, contribution, and collaboration in science*. J. Am. Soc. Inf. Sci. Tech., **57**, 1758-1770, doi:10.1002/asi.20380.

Carniel R., Di Cecca M. and Jaquet O.; 2006: *A user-friendly, dynamic web environment for remote data browsing and analysis of multiparametric geophysical data within the MULTIMO project*. J. Volcanol. Geotherm. Res., **153**, 80-96, doi:10.1016/j.jvolgeores.2005.08.005.

Diviacco P.; 2005: *An open source, web based, simple solution for seismic data dissemination and collaborative research*. Comput. Geosci., **31**, 599-605

Diviacco P.; 2006: *Integrating Scalable Vector Graphics (SVG). In a web based seismic data Portal*. In: EAGE 2006, Vienna, Austria.

Diviacco P.; 2012: *Addressing conflicting cognitive models in collaborative E-Research: a case study in exploration geophysics*. In: Collaborative and Distributed E-Research: Innovations in Technologies, Strategies and Applications, IGI Global press, doi:10.4018/978-1-4666-0125-3.ch012.

Diviacco P. and Wardell N.; 2003: *Reprocessing and dissemination of Antarctic Seismic Data*. Terra Antarctica, **9**, 133-136.

Diviacco P.; 2011: *Marine seismic metadata for an integrated European Scale Data Infrastructure: the Fp7 Geo-Seas Project*. Boll. Geof. Teor. Appl., **53**, 243-252, doi:10.4430/bgta0051.

Kötter R.; 2001: *Neuroscience databases: tools for exploring brain structure-function relationships*. Philos. Trans. R. Soc. London, Ser. B, **356**, 1111-1120, doi:10.1098/rstb.2001.0902.

Latour B. and Woolgar S.; 1979: *Laboratory life: the construction of scientific facts*. Sage Publ., Beverly Hills, 296 pp.

Nentwich M.; 2003: *Cyberscience: research in the age of the internet*. Austrian Academy of Science, Vienna, Austria, 569 pp.

Olson J.S., Ellisman M., James M., Grethe J.S. and Puetz M.; 2008: *Biomedical informatics research network (BIRN)*. In: Olson G.M., Zimmerman A. and Bos N. (eds), Scientific Collaboration on the Internet, MIT Press, Cambridge, MA, USA, pp. 221-232.

Orlikowski W.; 1992: *Learning from notes: Organizational issues in groupware implementation*. In: Proc. CSCW 1992, pp. 362-369.

Pitzalis D., Pillay R. and Lahanier C.; 2006: *A new concept in high resolution internet image browsing*. In: Proc. ELPUB2006 Conference on Electronic Publishing, Bansko, Bulgaria, available http://elpub.scix.net/data/works/att/249_elpub2006.content.pdf

Pratt W., Reddy M.C., McDonald D.W., Tarczy-Hornoch P. and Gennari J.H.; 2004: *Incorporating ideas from computer-supported cooperative work*. J. Biomed. Inf., **37**, 128-137, doi:10.1016/j.jbi.2004.04.001.

Ribes D. and Bowker G.; 2009: *Between meaning and machine: learning to represent the knowledge of communities*. Inf. Organ., **19**, 199-217, http://www.sis.pitt.edu/~gbowker/pub- lications/Ribes%20Bowker%20-%20Between%20 Meaning%20and%20Machine.pdf.


## WEB REFERENCES


[2003/98/CE ] http://ec.europa.eu/information_society/policy/psi/docs/pdfs/directive/psi_directive_en.pdf

[A-130] http://www.whitehouse.gov/omb/circulars_a130

[Antarctic Treaty] http://www.ats.aq/e/ats_keydocs.htm

[Batik] http://xmlgraphics.apache.org/batik/

[CC] http://wiki.creativecommons.org/Data

[CWP/SU] http://www.cwp.mines.edu/cwpcodes/

[Geo-Seas] http://www.geo-seas.eu/

[GMT] http://gmt.soest.hawaii.edu/

[HIPAA] http://www.hhs.gov/ocr/privacy/

[INSPIRE] http://inspire.jrc.ec.europa.eu/

[ISO19115] http://www.iso.org/iso/catalogue_detail.htm?csnumber=26020

[ISO19139] http://www.iso.org/iso/catalogue_detail.htm?csnumber=32557

[list C866 of BODC controlled vocabulary] http://vocab.ndg.nerc.ac.uk/client/vocabServer.jsp : list C866

[LRSP] http://geo.arc.nasa.gov/sge/landsat/15USCch82.html
[SDLS] http://sdls.ogs.trieste.it/SDLS
[SeaDataNet] http://www.seadatanet.org/
[SEG-Y] http://www.seg.org/resources/publications/misc/technical-standards
[UKOOA] http://www.epsg.org/exchange/p1.pdf
[US Copyright Act] http://www.law.cornell.edu/copyright/copyright.act.chapt1a.html

*Corresponding author:* Paolo Diviacco
Istituto Nazionale di Oceanografia e di Geofisica Sperimentale
Borgo Grotta Gigante 42c, 34010 Sgonico (TS), Italy
Phone: +39 040 2140380; fax: +39 040 327307; e-mail: pdiviacco@inogs.it