

# Data sharing in environmental sciences: A survey of CNR researchers

*Daniela Luzi<sup>\*</sup>, Roberta Ruggieri<sup>#</sup>, Stefania Biagioni<sup>°</sup>, Elisabetta Schiano<sup>§</sup>*

*National Research Council - <sup>\*</sup>Institute for Research on Population and Social Policies (IRPPS),  
<sup>°</sup>Institute of Information Science and Technologies (ISTI), <sup>§</sup>Institute of Marine Sciences (ISMAR),  
<sup>#</sup>Senate of the Italian Republic, Italy*

## Abstract

The paper presents the results of a survey on researchers' attitudes and practices of data sharing in the area of Environmental sciences. It is based on an online questionnaire submitted to CNR researchers active in this disciplinary field, that has proved to be data intensive, collaborative and multidisciplinary. The study lies within the framework of other international analyses that consider this complex process exploring different aspects that may influence the propensity of a consistent and effective data release. Therefore, motivations, perceived barriers and enablers to data sharing are analysed together with the outline of research context and practices in this field.

## 1. Introduction

Today the free availability of research data is considered an important driver of innovation and of new scientific insights. Due to the increasing amount of data collected as well as to the variety of purposes, process of acquisition and formats this is not an easy task. It implies the development of policies that promote data curation and preservation, the recognition of the value of research data as "first-class publication", the enforcement of clear rules for open access, copyright and ownership. It is also necessary that the scientific community agree on the development and use of common interoperability standards related to data models, format and exchange protocols. Last but not least, it requires that suitable infrastructures be developed at national and international level considering discipline specificity.

There is now a vast literature devoted to the definition and importance of research data [Borgman, 2012, Kowalczyk & Shankar 2011]. Many studies consider the technical aspects of preservation and management [Tjalsma & Rombouts, 2010, Graaf et al. 2011], while official documents and whitepapers outline current changes in the research process and propose policies and infrastructure that can promote data sharing [Hey et al. 2009, NSF, 2005]. Moreover, various surveys have been carried out to explore researchers' practices and perceptions towards data acquisition, curation and preservation, focusing in particular on perceived barriers and enablers of data sharing. Some surveys conducted within European projects, have analysed attitudes and opinions of different stakeholders: researchers, data managers, publishers, funding organisations [PARSE.Insight, 2009, Dallmeier-Tiessen et al, 2012], as well as libraries, national and local governments [EU Directorate, 2012]. They rely on different methods in the collection of results: questionnaires, interviews, desk research. They gained insight into differences between disciplinary fields across various countries with the aim of developing roadmaps or setting up a participatory process for the construction of international e-Science infrastructures [Tenopir et al. 2011]. Among the surveys that were particularly focused on specific research areas (Pinowar, 2011, Milia et al. 2012), it is also worth mentioning the studies related to biodiversity that

combine the analysis of researchers' attitude (Enke et. al., 2012) with the evaluation of technical and information resources available in this multidisciplinary field [Bach et al., 2012, Bendix et al. 2012].

Most of these studies have a common vision on data lifecycle as closely connected with the research process, where data sharing "begins with good data practices carried out in all phases of the data lifecycle" (Tenopir et al., 2011). Moreover, researchers' propensity to data sharing largely depends on the research context, synthesized by Kim as the combination of technological infrastructure, institutional support and interpersonal interactions (Kim & Stanton 2012).

Based on this view our survey intends to analyse researchers' attitude in data sharing posing a particular emphasis on the exploration of research practices and context within the broad multidisciplinary field of Environmental sciences. In our vision the understanding of the complexity of data sharing embedded in a specific research environment can bring to the fore opinions, beliefs, concerns and practices that may contribute to the development of suitable information systems tailored on researchers' needs as well as to the introduction of policies that may promote their consistent and long- term diffusion.

## 2. Methods

Among the different CNR departments devoted to different disciplinary fields, we choose to analyse attitudes of researchers belonging to the Institutes of the Department of Earth and Environment because this research field has proved to be data intensive and multidisciplinary in nature. Moreover, in this area there are several initiatives both at international and CNR level that are promoting and setting up infrastructures for data sharing.

The survey makes use of a semi-structured questionnaire of 40 questions that consists of two main parts. Reflecting the survey hypotheses, the first one aims to gain insight into research practices that may influence data sharing. Based on the chosen target group we identified *ad hoc* questions to explore in particular:

- The general research context (research lines, types of funds, types of collaboration);
- Data acquisition (type or research carried out, data used, modes of data acquisitions and instrumentation);
- Data management (availability of standards, use of descriptive metadata, adoption of preservation procedures, presence of dedicated personnel for data management);
- Data re-use and availability (propensity of using data produced by others and related evaluation of its reliability; available resource to store own data, practices in data sharing).

The second part is specifically focused on capturing perceived barriers to as well as conditions that may motivate data sharing. This part contains a selection of questions submitted in large-scale international surveys (PARSE.Insight, 2009, Tenopir et al., 2011, Enke et. al., 2012) in order to explore commonality and differences in attitudes.

Additionally, respondents were asked for information on gender, age, length of CNR service and occupational position. Most of the questions are multiple choice, while two plain text answers were also included in the questionnaire to collect researchers' free opinion on this topic. Both questionnaire and survey data are available at: <https://easy.dans.knaw.nl/ui/datasets/id/easy-dataset:53292>.

The survey was Internet-based and the link was sent out via e-mail using LimeSurvey open source software that also supports invitations, reminders, and makes answers anonymous. The survey period was June-September 2012.

### 3. The sample

1087 questionnaires were sent to all researchers affiliated to the 13 CNR Institutes belonging to the Department of Earth and Environment. We included researchers with both permanent and temporary contracts and also external collaborators, who are generally researchers coming from universities that closely collaborate with CNR Institutes. The response rate was 48% (i.e. 523 responses) that can be considered satisfactory given the voluntary basis of the survey.

There are 13 CNR Institutes that belong to the Earth and Environment Department. These institutes are different in size, ranging from 40 researchers to more than 100 and are organised in various research units located all over Italy. Their area of interest ranges from land and water ecosystems to climate change, from the use of resources to the monitoring of anthropogenic risks, from biodiversity to the development of methods and technologies for environment protection. The distribution of responses by Institute (tab 1) shows that five Institutes out of 13 reach a response rate higher than 50%. Most of them are in the area of marine sciences and water resources.

**Table 1 - Distribution of responses by Institute**

<i>Institutes</i>	<i>Questionnaires sent (No.)</i>	<i>Questionnaires received (No.)</i>	<i>%</i>
IAMC - Institute for coastal marine environment	93	53	<b>57.0</b>
IBAF - Institute of agro-environmental and forest biology	53	27	<b>50.9</b>
IDPA - Institute for the dynamics of environmental processes	58	25	<b>43.1</b>
IGAG - Institute of environmental geology and geo-engineering	103	36	<b>35.0</b>
IGG - Institute of geosciences and earth resources	132	63	<b>47.7</b>
IIA - Institute for atmospheric pollution research	67	29	<b>43.3</b>
IMAA - Institute of methodologies for environmental analysis	83	31	<b>37.3</b>
IRPI - Research institute for geo-hydrological protection	88	40	<b>45.5</b>
IRSA - Water research institute	69	40	<b>58.0</b>
ISAC - Institute of atmospheric sciences and climate	94	46	<b>48.9</b>
ISE - Institute of ecosystem study	55	35	<b>63.6</b>
ISMAR - Institute of Marine sciences	149	80	<b>53.7</b>
IVALSA Tree and timber institute	43	18	<b>41.9</b>
<b>Total</b>	<b>1,087</b>	<b>523</b>	

### 3.1. Respondents' profiles

An overview of the respondents' profile is given in table 2. The majority of respondents to the survey are male. They fall mainly into two age groups (from 41 to 50 and over 50 years old). The length of service at CNR is concentrated in two groups: from 11 to 20 years and over 20 years of CNR service. The majority of respondents have a permanent contract.

Table 2. - Respondents' profile

<i>Respondents' profile</i>		
	<b>No.</b>	<b>%</b>
<b>Gender</b>		
F	204	<b>39.4</b>
M	314	<b>60.6</b>
	<b>518</b>	
<b>Age</b>		
< 30	31	<b>5.9</b>
30 - 40	139	<b>26.6</b>
41 - 50	170	<b>32.5</b>
> 50	177	<b>33.8</b>
	<b>517</b>	
<b>Length of service</b>		
> 5 years	117	<b>22.4</b>
6-10 years	107	<b>20.5</b>
11-20 years	144	<b>27.5</b>
> 20 years	142	<b>27.2</b>
	<b>510</b>	
<b>Position</b>		
Permanent	327	<b>62.5</b>
Temporary	129	<b>24.7</b>
Training	9	<b>1.7</b>
External collaboration	56	<b>10.7</b>
Other	2	<b>0.4</b>
	<b>523</b>	

## 4. Research context

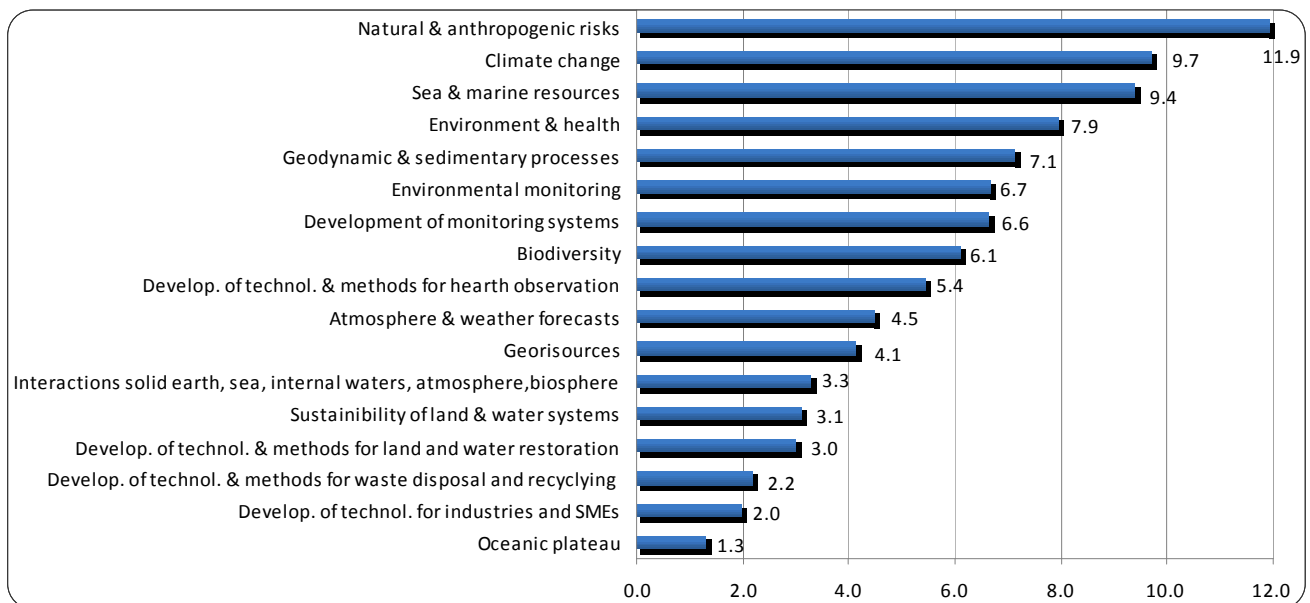
Data sharing does not simply represent an individual propensity, but it is influenced by socio-cultural, contextual and institutional factors. It has proved to vary from discipline to discipline and within disciplines, it depends on different factors: types of research and collaboration setting, types of data, modes of acquisition and handling strategies as well as human, technical and institutional support for long-term preservation, to mention but a few.

Therefore, the first part of the questionnaire is devoted to exploring general features of research practices to obtain a more detailed framework on how research activities are carried out in this area.

#### 4.1. Research lines

Researchers were asked to provide a percentage of time dedicated to a set of research lines described in the website of the CNR Department of Earth and Environmental Sciences. Figure 1 shows the topics on which researchers concentrate their work (multiple answers were allowed). Researchers are generally involved in more than one research line, a relevant percentage on them deals with Natural and anthropogenic risks (11.9%), Climate change (9.7%) and Sea and marine resources (9.4%).

**Fig. 1. - Distribution of the research lines carried out by CNR researchers**



#### 4.2. Funds

When asked to provide a percentage of funds received in carrying out their research activities, the majority of researchers (47.8%) reported that they receive national funds, 28.9% rely on EU and/or international projects and 14.9% on national and international private funds. On average, only 4.4% of researchers reported that their work is directly funded by CNR.

#### 4.3. Collaboration

Two questions in the survey were focused on researchers' collaboration habits. The first one asked, whether they usually work as a single researcher, in small (max 3 persons), medium (from 3 to 7 persons) or in large groups (more than 8 persons). The majority of researchers work in a medium size (47.7%) group.

When asked how often and on which occasions they collaborate with multidisciplinary groups, 42% reported that they always do so in international projects and with colleagues of the same Institute (36.9%). Working in multidisciplinary groups occurs sometimes with other CNR institutes (56.2%) and with other Italian institutions and/or Universities (61.6%).

#### 4.4. Data acquisition

As data sharing is part of data lifecycle, a set of questions was devoted to exploring types of data used, how they are acquired and managed. A prerequisite of data sharing is that data are acquired following defined procedures, is associated with proper metadata, so that data are interpretable and properly reusable. Therefore a set of questions was focused on the type of research carried out, types of data used, how data are acquired, as well as information on measurements and instrumentation.

We first asked researchers to provide a percentage of time dedicated to theoretical and/or experimental research in order to gain insights into the type of research most frequently carried out in this field. On average the majority of CNR researchers (77%) carry out experimental research that generally implies the collection as well as an intensive use of data.

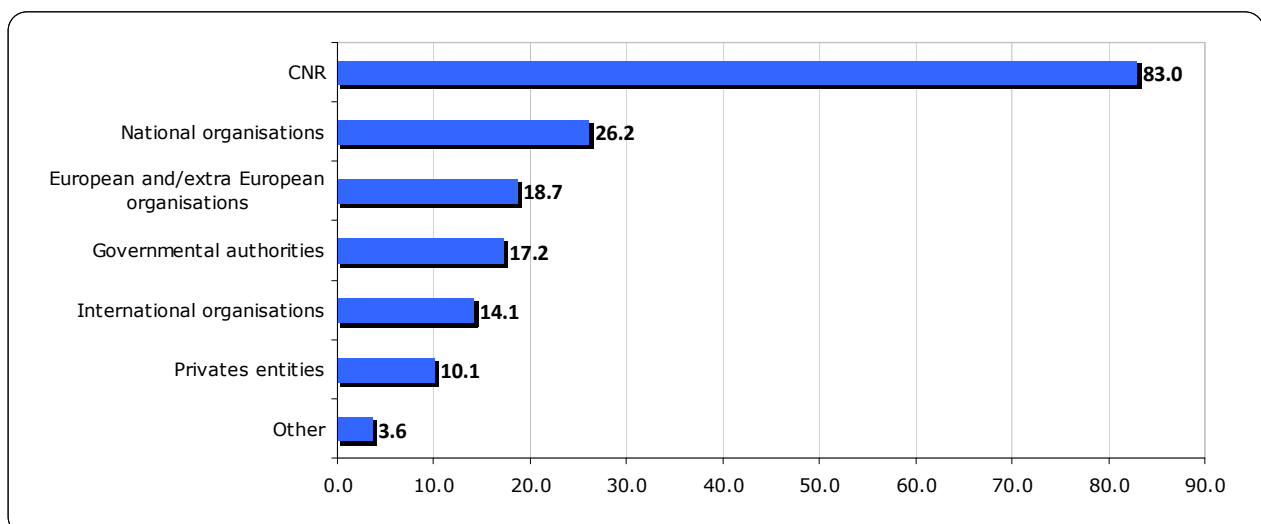
**Tab. 3. - Type of data used in the analysis of land, sea, internal waters, atmosphere and biosphere**

	<i>Biological</i>	<i>Chemical</i>	<i>Physical</i>	<i>Geological</i>
Land	18.5	31.0	31.9	45.1
Sea	27.0	30.0	30.8	26.2
Internal waters	22.6	36.5	30.2	27.3
Atmosphere	6.5	27.3	42.4	13.4
Biosphere	28.3	24.3	21.2	16.4

Researchers were asked to indicate the type of data used when they analyse phenomena related to land, sea, internal waters, atmosphere and biosphere. Multiple answers were allowed. Table 3 shows that CNR researchers more frequently use geological data related to the study of land (45.1%) as well as physical data related to the Atmosphere. Data gathered in the analysis of sea, internal waters and biosphere tend to be almost equally distributed among biological, chemical, physical and geological data. This multidisciplinary approach is confirmed by some researchers, who specified in the variable “other” that they use biogeochemical, geo-morphological, geophysical or geo-mechanical data. A small percentage of researchers (0.6%) indicate in “others” that they use remote sensing data.

Moreover, 21% of researchers also use demographic data to carry out their research activities.

**Fig. 2. - Distribution of respondents to the question “The data you are working on come mainly from instrumentation managed directly by ...”**



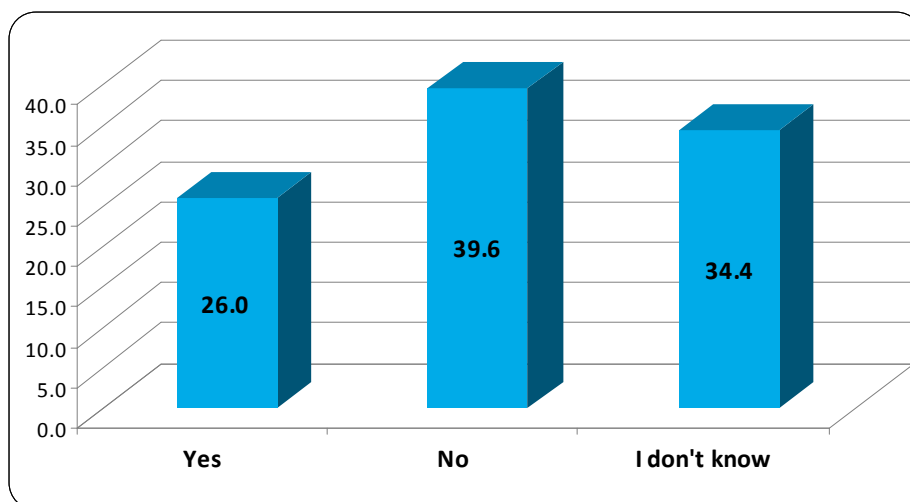
When asked whether they take measurements directly by themselves, or use measurements taken by others or alternatively use both, the majority of CNR researchers (53%) reported that they use measurements directly taken by themselves and/or by their research group, while 8.2% use measurements taken by others and 38.8% use both. Moreover, data are mainly acquired from both laboratory work and in the field (53.8%), while 32.5% of CNR researchers collect data from fieldwork alone.

A multiple answer was allowed to indicate who manages the instrumentation used. Figure 2 shows that an overwhelming majority of CNR researchers (83%) obtain data from instrumentation directly managed by CNR, while 26.2% of them also use data taken from instrumentation managed on the basis of agreements with other national organisations.

#### 4.5. Data management

The use of standards facilitates data sharing, while re-use and evaluation of data also depends on the metadata associated to the data acquired. Therefore, a set of questions aimed to explore different aspects of data management, such as the availability of standard of researchers' community of reference, use of descriptive metadata in their current research practice, data management plan in place in their Institutes and presence of trained staff that may support data curation.

**Fig. 3. - Distribution of respondents to the question "Does your community of reference use standards to manage data?"**

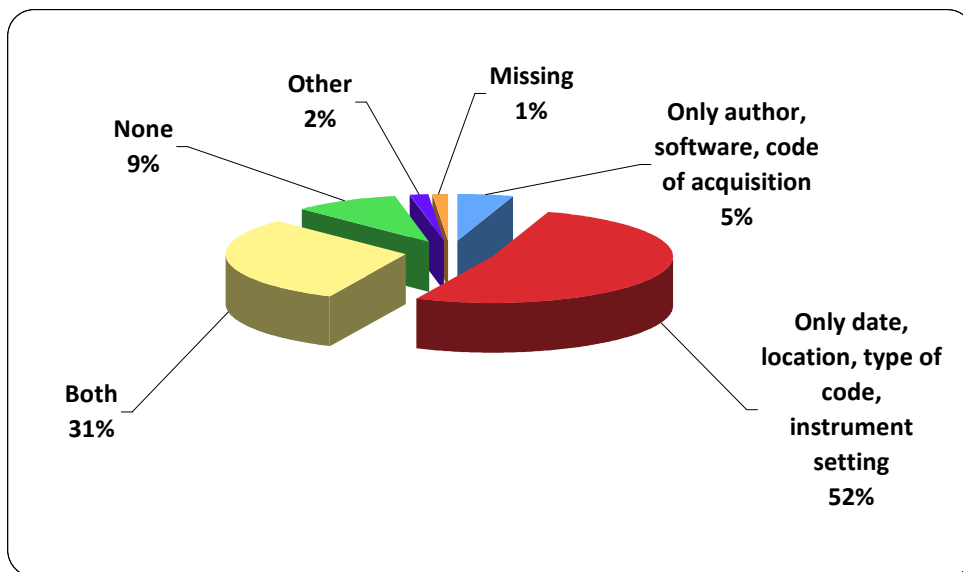


When asked about the use of standards, a high percentage of researchers reported that their community of reference doesn't use standards (39.6%), while 26% of them don't know about the use of standards in their research field (fig. 3).

The remaining 26% of researchers that answered positively to this question also specified the standard they more frequently use. Many of them use a set of standards specific to the type of data and infrastructure of reference for their work. Here a brief overview of the standard more frequently indicated. Many mention the European initiative INSPIRE (Infrastructure for Spatial Information in the European Community) that established a general framework for Spatial Data

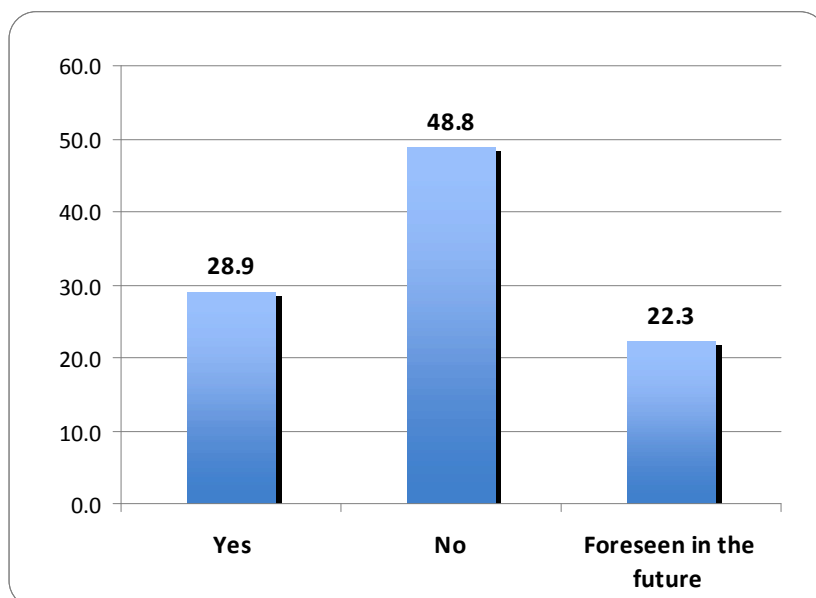
Infrastructure (SDI) together with ISO19115 (Geospatial metadata) as well as the standard developed by the Open Geospatial Consortium (OGC). Others rely on the SEG Y standard file format developed by the Society of Exploration Geophysicists for storing geophysical data, or on NetCDF (Network Common Data Form), an open standard for sharing array-oriented scientific data, and/or on ISO/WMM (World Meteorological Organization) to standardize meteorological data.

**Fig. 4. - Distribution of respondents to the question: "What type of additional information do you generally associate with data you have collected/analysed?"**



When data are collected and/or analysed, 52% of CNR researchers provide metadata related to the date of collection, information on location, type of code used and instrument setting. 9.4% of researchers associate data with additional information on the author, software, code of acquisition, while 30.6% associate both types of the above-mentioned metadata. Only 9.4% do not associate any type of metadata to data gathered or analysed (fig. 4). The addition of descriptive metadata is an encouraging result as it makes research data more easily interpretable and reusable, thus more accessible and better suited for preservation.

**Fig. 5. - Distribution of respondents to the question: "Does your Institute have specific"**

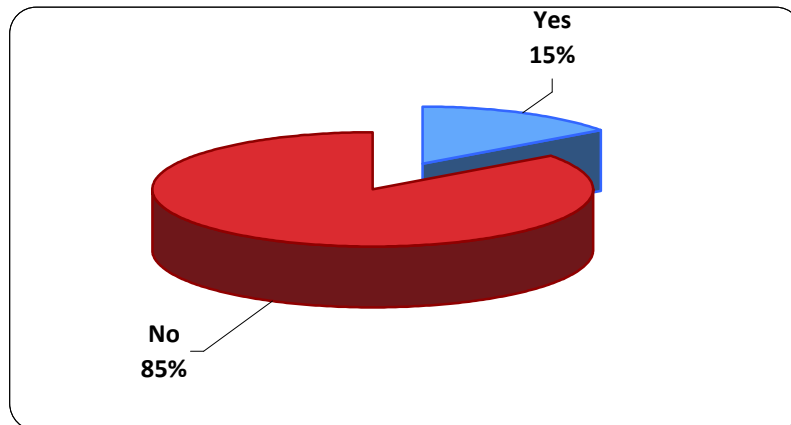




### *procedures for data preservation in place?"*

When asked whether specific procedures for preservation are set up by their institute, 28,9% of researchers reported that these procedures are in place in their institutes, while 22,3% reported that these procedures are going to be set up in the future (fig. 5).

**Fig. 6. - Distribution of respondents to the question: "Is there anyone in your Institute who is specifically trained to manage data?"**



The presence of personnel specifically trained to manage data is reported by 15.4% of researchers, while the majority of them answered that there is no one in their institute that is in charge of this task (fig. 6).

Out of 79 researchers that reported on the presence of personnel dedicated to data preservation, 60 indicated the type of personnel that carry out this task. Generally they are IT experts that manage local databases, GIS, digital images. Many researchers mention technicians or researchers that carry out this task, and only in few cases do respondents refer to a data manager, that is the emerging professional skill often mentioned in official documents on data management and preservation. One respondent reports that data preservation is carried out by the same person who manages the Institutes' publications, probably a librarian.

#### *4.6. Data re-use and availability*

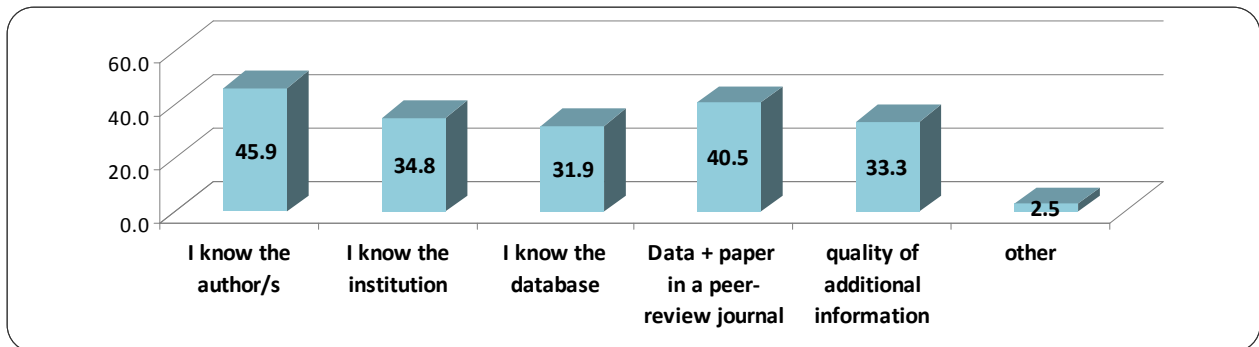
This group of questions aims to ascertain whether researchers use data produced by others, in which field, along with the criteria they apply to consider data reliable. Generally the use of data generated by others is associated with the propensity of sharing researchers' own data, in the hypothesis that this could represent a mutually coherent behaviour. Results of other surveys (PARSE.Insight, 2009, Tenopir et al., 2011, Enke et. al., 2012) generally showed a lower percentage of data sharing when compared with the re-use of data generated by other researchers.

59% (= 307) of researchers indicate that they use data produced by others. Among them, 43% re-uses data in the same disciplinary field, while a similar percentage of researchers re-use data coming both from the same disciplinary and from cross-disciplinary fields.

When researchers are re-using data produced by others they consider data reliable if they know the authors (45.9%) and when data are associated with peer-reviewed journals (40%) (fig. 7). Answers reported in the variable "Other" help to give a more complex picture of data reuse. Some researchers reported that they apply procedures of quality control and validation; others consider

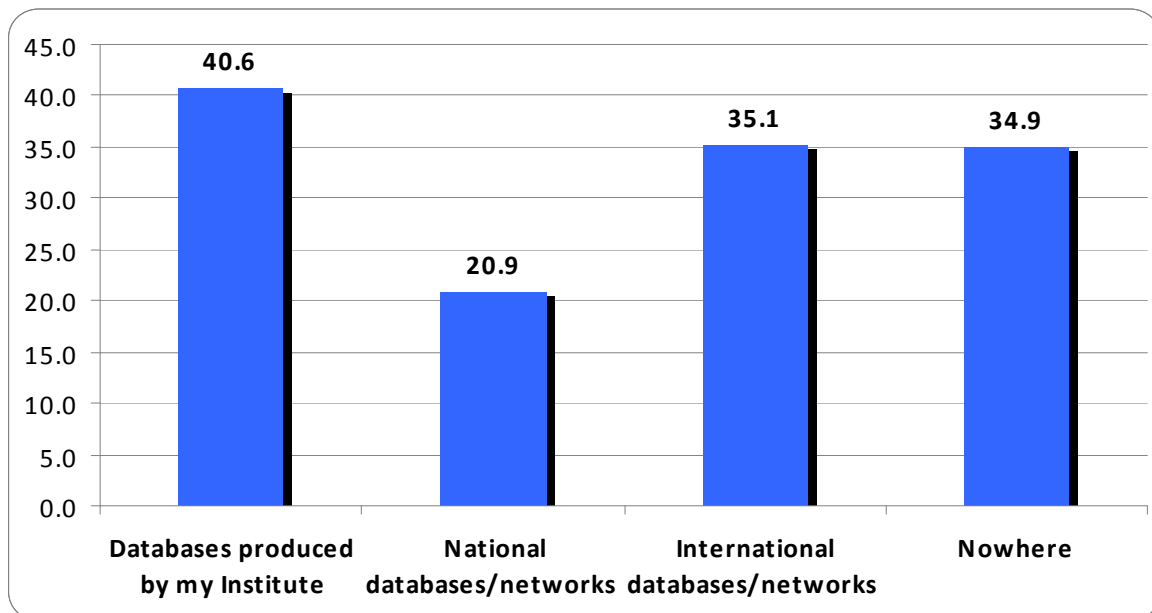
the experimental method adopted as well as methods and instrumentation used to collect data. This highlights that data reuse may not always be a straightforward process.

**Fig. 7. - Distribution of respondents to the question: "What reassures you that the data produced by others is reliable?"**



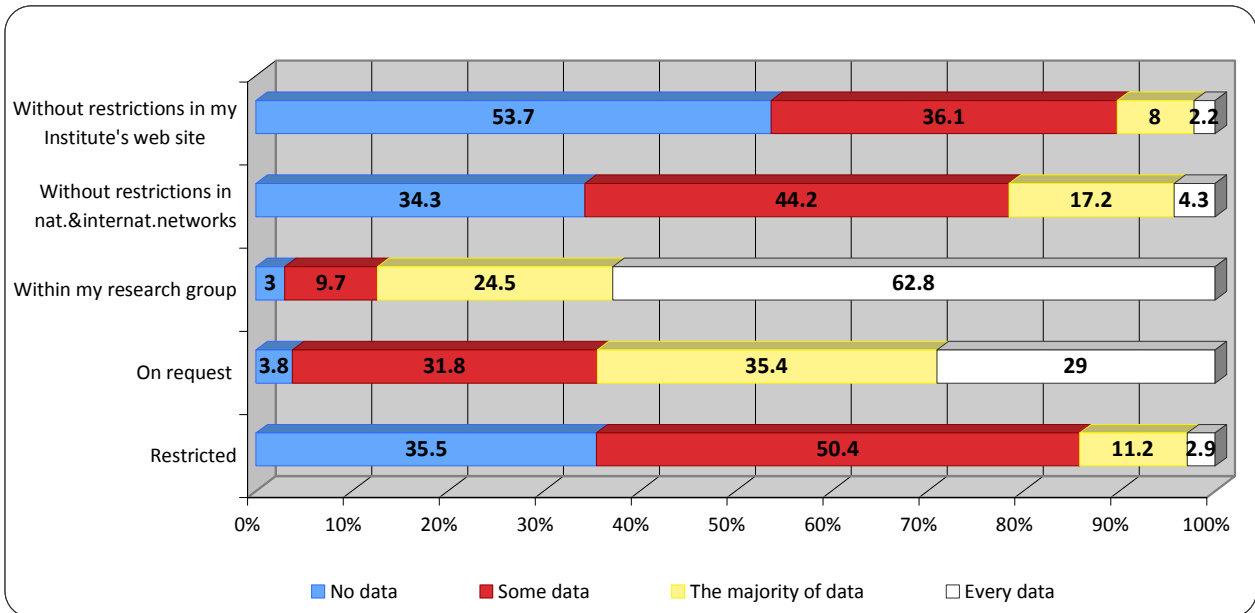
The willingness to share data also depends on the availability of databases or infrastructures where researchers can deposit their research data. For this reasons we asked whether there are databases or networks where they can deposit their data in their disciplinary field.

**Fig. 8. - Distribution of respondents to the question: "In your disciplinary field are there archives where your research data can be stored?"**



More than 40% of researchers store their data in databases produced by their institutes, 35% in international databases and 20% in national databases, while for 34% of researchers there are no databases where their data can be submitted (fig. 8).

**Fig. 9. - Distribution of respondents to the question: “ Data from your current research is available to everyone without restrictions “**



Turning to data made available by CNR researchers, we can generally say that they make a selection of the data that they share (fig. 9). At least some data are available without restriction in the Institute’s website (36.1%), or in national and international networks (44.2%). Of course all data are available within their research groups (62.8%). It is interesting to note that when data are requested, CNR researchers declare they do make them available, only 3.8% of them report that no data are available on request. A small percentage of researchers report that either all their data or the majority are restricted (2.8% and 11,2).

## 5. Researchers’ attitude

This part of the questionnaire intends to explore researchers’ opinions on the role played by research data, perceived obstacles and enablers to data sharing. As previously mentioned this part of the questionnaire is also based on other surveys carried out at international level, so that differences and communalities with the international context can be compared.

Before asking on data sharing practices and perceptions, we considered it important to let researchers express their opinions on reasons for the availability and preservation of data. We proposed a list of nine well-known statements (7 of which were taken from the Parse Insight survey) and asked whether they consider these reasons very important, important, slightly important or not important.

Almost all assertions of this self-evident list of reasons are considered very important or important by the majority of researchers. If we analyse how they ranked their importance, it emerges that researchers find that data availability and preservation foster the process of science (56.8%) and that it also enhances the transparency of research (53.9% very important and 40.7% important).

**Table 3 - Distribution of respondents to the question: “In your opinion for which reasons is it important to make research data available and preserve it?”**

	<i>Very important</i>	<i>Important</i>	<i>Not very important</i>	<i>Not important at all</i>	<i>Missing</i>
The availability of data enhances the transparency of research results	53.9	40.7	3.8	0.6	1.0
When research is publicly funded, data should be available to anyone	50.7	38.6	7.5	2.1	1.1
The availability of data fosters the progress of science (new research is based on pre-existing knowledge)	56.8	38.2	3.6	0.4	1.0
It is a means to validate the results obtained	40.2	43.6	12.6	1.1	2.5
Existing results can be re examined	34.6	46.3	14.5	2.7	1.9
It can promote collaboration among different fields	39.6	45.5	12.8	0.8	1.3
It has a potential economic value	19.1	39.8	34.2	4.2	2.7
Research data are unique	19.9	40.0	26.8	9.8	3.6
The availability of data reduces the duplication of research efforts	36.9	35.9	18.9	6.3	1.9

Another reason to make data available and preserve them is that research is publicly funded and therefore should be made available to everyone (50.7% very important and 38.6% important). The economic value of data (4.2% not important at all) together with the assertion that data are unique (9.8%) is regarded as the least important reasons for availability and preservation. These two values are not surprising, as also in the Parse Insight project the survey obtained the same results. In the case of CNR researchers these values are balanced against the rate given as important (respectively 39.8% and 40%). Moreover, the Parse Insight survey found out that opinions on the very important and important reasons depended on the disciplinary field of the respondents. CNR researchers consider that data availability and preservation can stimulate the advancement of science like researchers in Humanities, Life sciences, Physical Sciences and Socio-cultural sciences.

### *5.1. Obstacles to data sharing*

When asked on the obstacles of data sharing (table 4), we obtained a more homogenous distribution of responses, especially if we compare this question with the previous one. If we consider both the very important and important values we can notice a common agreement on some obstacles felt by the majority of CNR researchers (where sometimes the important value prevails on the very important one). These are: lack of technical support (41.9% important, 31.4% important) lack of standards (46.3% important, 25.8% very important), but also the fact that data are not evaluated like papers in scientific journals (37.5% very important, 31.5% important).

**Table 4. - Distribution of respondents to the question: In your opinion what are the main obstacles to data sharing?**

	<i>Very important</i>	<i>Important</i>	<i>Not very important</i>	<i>Not important of all</i>	<i>Missing</i>
Lack of funds	31.4	30.6	27.9	5.4	4.8
Lack of standards	25.8	46.3	18.9	3.1	5.9
It requires too much time	16.1	38.0	32.7	7.8	5.4
Difficulties in adoption of standard	13.0	38.4	33.7	8.4	6.5
No technical support	31.4	41.9	16.4	3.6	6.7
There are no archives to submit to	23.3	37.3	23.9	9.2	6.3
Procedures of data sharing are too complicated	10.7	33.8	38.4	10.7	6.3
Loss of data control	19.9	31.4	30.4	12.6	5.7
Data may be misused and/or misinterpreted	22.8	35.6	25.4	10.5	5.7
Data are not evaluated like papers in scientific journals	37.5	31.5	20.8	5.0	5.2
Loss of exclusivity of the work	26.4	29.4	30.0	8.6	5.5

When we look at the least important perceived barriers, the statement related to the too complicated procedures receives the majority of responses, but the differences with the researchers that consider it important are not so high. At this stage of the analysis we could say that CNR researchers perceive a relevant number of different barriers as being rather important. It would be interesting to further analyse these perceptions combining these results with other variables of the questionnaire.

### *5.2. Enablers of data sharing*

The last multiple-structured question summarizes some of the issues already investigated, but specifically asked on the conditions required to submit data to an open archive. A major consensus on very important facilitators of data deposit is evident here (table 5). The majority of researchers find very important to have the possibility to update data after submission (60.2%), to know who is using them, when and for which purpose (53.5%), to be contacted if data are used (52%). All these responses are related with a clear wish to keep control over their own data also after submission. Another very important factor that may encourage researchers to deposit is the availability of simple procedures for submission (52.6%) as well as receiving the same evaluation as in the case of publications.

**Table 5. - Distribution of respondents to the question: “What condition would you require to submit your research data to an open archive?”**

	<i>Very important</i>	<i>Important</i>	<i>Not very important</i>	<i>Not important at all</i>	<i>Missing</i>
I will be able to update data after submission	60.2	30.8	4.2	1.7	3.1
I will be able to delete data	31.2	33.7	22.0	7.3	5.9
I know who is using data, when and for which purpose	53.5	27.5	11.7	3.6	3.6
Be contacted if someone wants to use my data	52.0	30.2	12.0	2.3	3.4
Receive a formal acknowledgment	35.4	36.1	20.3	4.0	4.2
Be reassured about long-term data preservation	38.6	39.8	13.4	3.3	5.0
Simple procedures to deposit data	52.6	37.1	5.5	0.6	4.2
Receive additional funds	24.7	39.2	27.5	4.4	4.2
Receive the same evaluation received for publications	41.1	37.5	13.0	3.8	4.6

## 6. Conclusions

Summarising some of the main results of the survey, CNR researchers in the field of environmental sciences tend to work in collaboration, often involved in multidisciplinary projects within the same institutes and with external organisations. They mainly carry out experimental research, use different types of data, gathered directly by themselves or by their research group in both laboratory and field work, using instrumentation directly managed by CNR.

There is not a diffuse use of standards, but researchers who use them apply different types of them, according to the data they are working on. Nevertheless, data collection is often associated with descriptive metadata that represent a pre-requisite for data reusability and interpretation as well as for preservation. It is also encouraging that a relevant number of researchers rely on procedures for data preservation already set up in their institutes or foreseen in the future. This process is generally carried out by the researchers themselves, as the majority of them do not have any support from specifically trained data managers.

Despite the use of data produced by others, CNR researchers tend to share only a fraction of data they produce. Generally they are more willing to share data on request, keeping control on whom is using their data and for which purposes.

A relevant number of obstacles are perceived by CNR researchers as rather important: lack of technical support, lack of standards, no formal recognition of practices of data sharing, but also lack of funds, fear of losing the exclusivity of their work. These perceptions are worth further analysis, combining these results with other variables of the questionnaire. Conditions required to

submit research data to open archives concern both technical and policy-related aspects that confirm a clear wish to keep control over research data even after submission as well as the provision of simple procedures for submitting them. Doubtless a further motivation is that data sharing is evaluated the same way as publications are.

Generally the high rate of responses received to the questionnaire as well as researchers' opinions on the importance of research data indicate a high level of awareness and an encouraging willingness to share data that should be further strengthened by the introduction of policies and the development of infrastructures tailored to researchers' needs.

## References

Bach Kerstin, Schäfer Daniel, Enke Neela, Seeger Bernhard, Gemeinholzer Birgit, Bendix Jörg (2012). A comparative evaluation of technical solutions for long-term data repositories in integrative biodiversity research. *Ecological Informatics*, 11: 16-24.

Bendix Jörg, Nieschulze Jens, Michener William K. (2012). Data platforms in integrative biodiversity research. *Ecological Informatics*, 11: 1-4.

Borgman Christine L. (2012). The Conundrum of sharing research data. *Journal of the American Society for Information Science and Technology*, 64 (6): 1059-1078.

Dallmeier-Tiessen Sunje, Darby Robert, Gitmans Katrin, Lambert Simon, Suhonen Jari, Wilson Michael (2012). Compilation of results on drivers and barriers and new opportunities. ODE Project (Opportunity for Data Exchange) URL: [www.ode-project.eu/ode-output](http://www.ode-project.eu/ode-output)

EU Directorate - General for Research and Innovation (2012) Online survey on scientific information in the digital age. URL: [http://ec.europa.eu/research/science-society/document\\_library/pdf\\_06/survey-on-scientific-information-digital-age\\_en.pdf](http://ec.europa.eu/research/science-society/document_library/pdf_06/survey-on-scientific-information-digital-age_en.pdf)

Graaf Maurits van der, Waaijers Leo (2011). KE Knowledge Exchange Primary Research Data Working Group. A Surfboard for Riding the Wave: Towards a Four Country Action Programme on Research Data. URL: <http://www.voced.edu.au/content/ngv48428>>

Enke Neela, Thessen Anne, Bach Kerstin, Bendix Jörg, Seeger Bernhard, Gemeinholzer Birgit (2012). The user's view on biodiversity data sharing – Investigating facts of acceptance and requirements to realize a sustainable use of research data. *Ecological Informatics*, 11, September 2012, pp. 25-33.

Hey Tony, Tansley Stewart, Tolle Kristin (2009). The Fourth Paradigm: Data-Intensive Scientific Discovery. s.l. Microsoft Cooperation. URL <http://research.microsoft.com/en-us/collaboration/fourthparadigm/contents.aspx>

Kim Youngseek, Stanton Jeffery M. (2012). Institutional and individual influences on scientists' data sharing practices. *Journal of Computational Science Education*, 3 (1), June 2012

Kowalczyk Stacy, Shankar Kalpana (2011), Data sharing in the sciences. *Ann. Rev. Info. Sci. Tech.*, 45: 247–294. doi: 10.1002/aris.2011.1440450113

Michener William K., Allard Suzie, Budden Amber, Cook Robert B., Douglass Kimberly, Frame Mike, Kelling Steve, Koskela Rebecca, Tenopir Carol, Vieglais David A. (2012). Participatory design of DataONE-Enabling cyberinfrastructure for the biological and environmental sciences, *Ecological Informatics*, 11, September 2012, pp. 5-15

Milia Nicola, Congiu Alessandra, Anagnostou Paolo, Montinaro Francesco, Capocasa Marco, Sanna Emanuele, Destro Bisol Giovanni (2012). Mine, yours, ours? Sharing data on human genetic variation. PLoS ONE 7(6): e37552.

National Science Foundation (2005). Long-lived Digital Data Collections: Enabling Research and Education in the 21<sup>st</sup> century. URL: <http://www.nsf.gov/pubs/2005/nsb0540/start.jsp>

PARSE.Insight (Insight into issues of Permanent Access to the Records of Science in Europe). D3.4 Survey Report. URL: <http://www.parse-insight.eu/publications.php>

Pinowar Heater A. (2011). Who shares? Who doesn't? Factors associated with openly archiving raw research data. PLoS ONE, 6 (7) e18657.

Tjalsma Heiko, Rombouts Jeroen (2010). Selection of research data. Guidelines for appraising and selecting research data. Dans Studies in Digital archiving, 6. The Hague and Delft. SURFfoundation, Data Archiving and Networked Services (DANS), 3TU.Datacentrum. URL: <http://www.dans.knaw.nl/en/content/categorieen/publicaties/dans-studies-digital-archiving-6>

Tenopir Carol, Allard Suzie, Douglass Kimberly, Aydinoglu Arsev Umur, Wu Lei, Read Eleanor, Manoff Maribeth, Frame Mike (2011). Data sharing by scientists: Practices and Perceptions. PLoS ONE 6(6): e21101. URL: <http://www.plosone.org/article/info:doi/10.1371/journal.pone.0021101>

## **Acknowledgment**

We thank Cristiana Crescimbene (CNR-IRPPS) for the technical support in the questionnaire submission and in the management of the LimeSurvey in all phases of the survey. We are also grateful to our CNR colleagues that voluntarily answered this questionnaire.