Fourteenth International Conference on Grey Literature
National Research Council, Rome, Italy 29-30 November 2012

Data sharing in environmental sciences: A survey of CNR researchers

Daniela Luzi*, Roberta Ruggieri#, Stefania Biagioni°, Elisabetta Schiano§

*CNR-IRPPS, #Senate of the Italian Republic,
°CNR-ISTI, §CNR- ISMAR

Consiglio Nazionale delle Ricerche
Istituto di Ricerche sulla Popolazione e le Politiche Sociali

ISTITUTO DI SCIENZA E TECNOLOGIE DELL'INFORMAZIONE "A. FAEDO"

ISMAR

Senato della Repubblica

Objectives

Analysis of researchers' attitudes in data sharing

- Are researchers willing to share data?
- At which conditions?
- What are their concerns?

Analysis of practices and research context

- How research activities are carried out?
- How research data are managed, diffused, preserved and re-used?

Capture perceptions, identify barriers and possible enablers to support open data sharing and contribute to the development of suitable information resources

SOURCE OF ANALYSIS:

CNR Institutes belonging to the Department of Earth and Environment

- data-intensive community
- multidisciplinary

Methods

- **Survey design**
  - ✓ Analysis of previous international surveys
  - ✓ Identification of the survey target group
  - ✓ Identification of ad hoc questions for CNR
  - ✓ Development of the questionnaire (40 questions)

- **On-line questionnaire submission via e-mail**
  - ✓ Survey period: 20 June – 15 September
  - ✓ Collection of anonymous responses

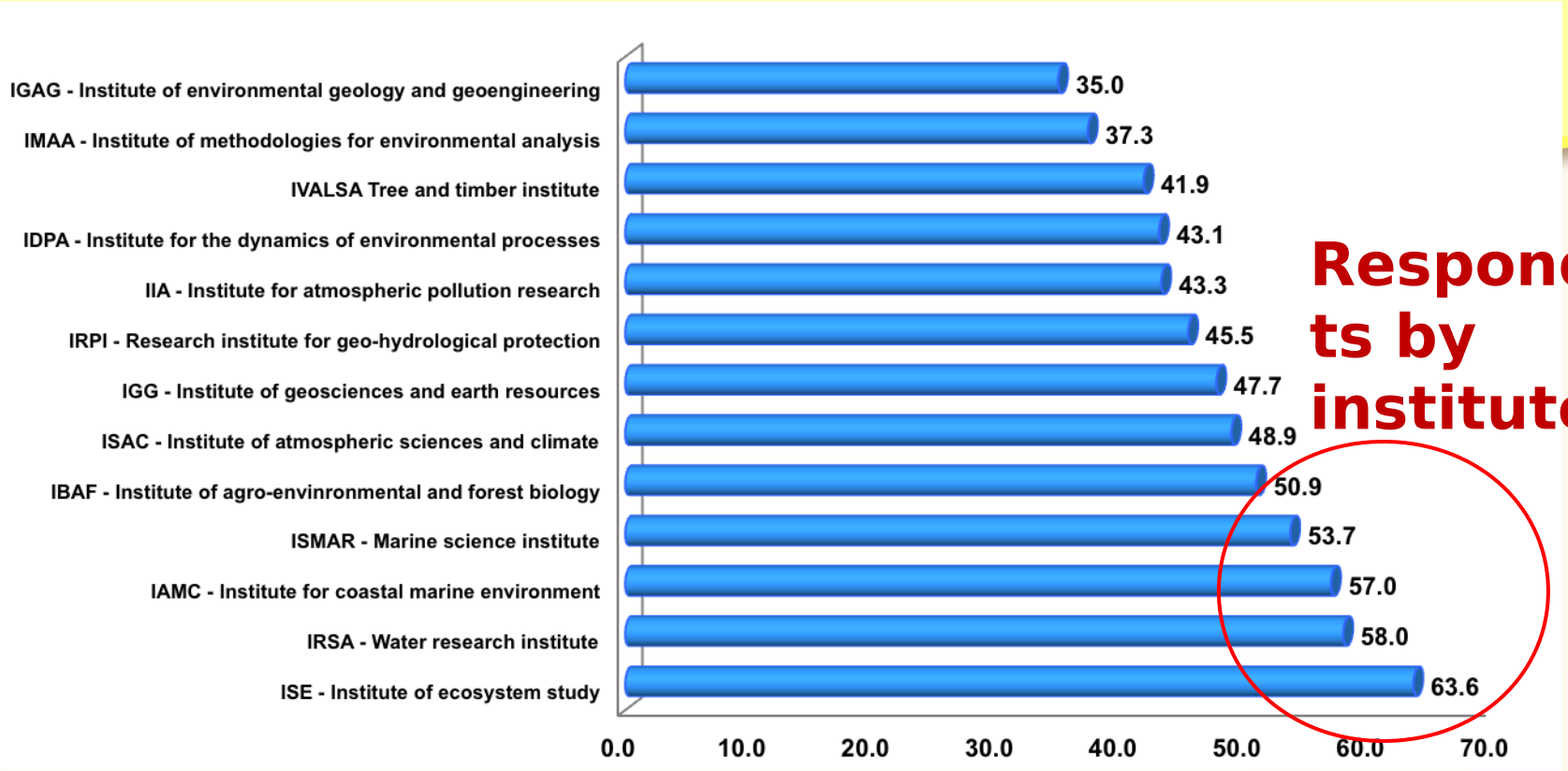
- **Data analysis**

The sample

1087 questionnaires sent

523 received

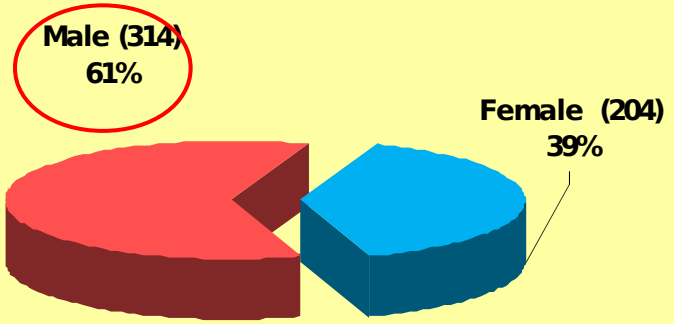
= 48%



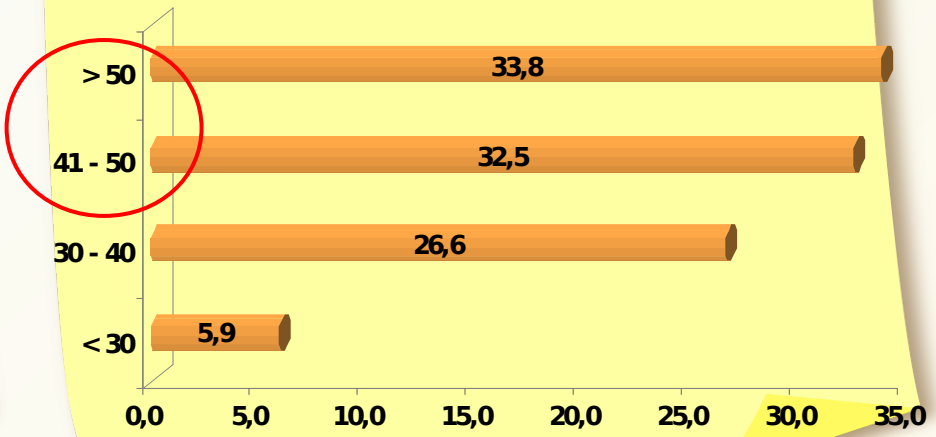
Respondents by institute

Respondents' profile

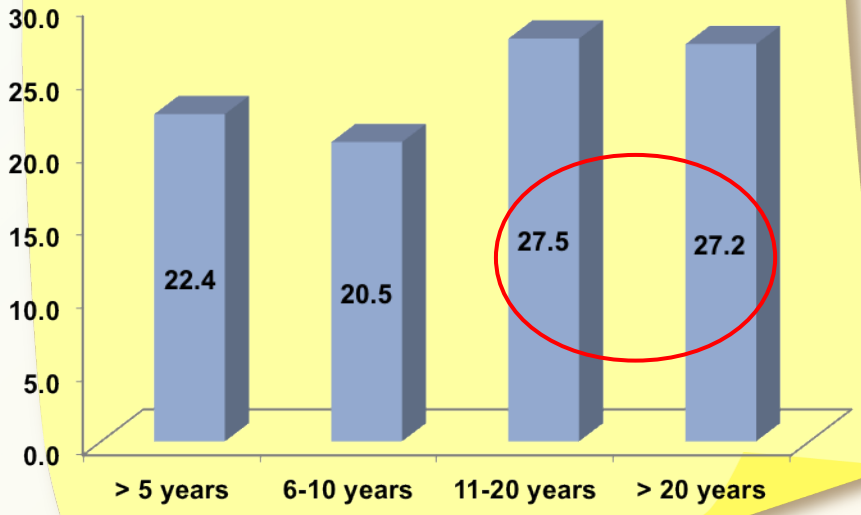
Gender



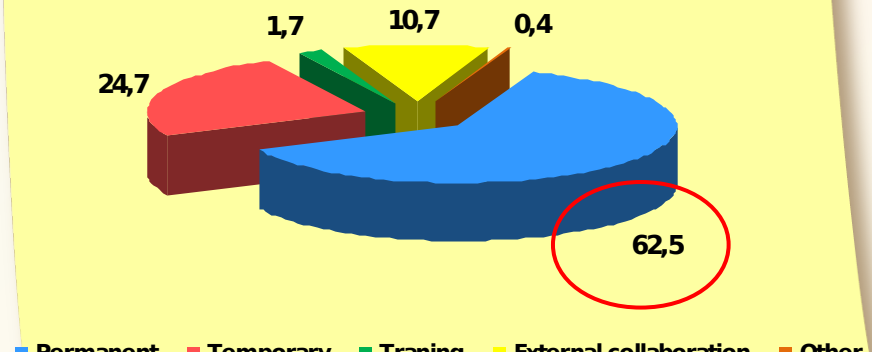
Respondent' age group



Length of service at CNR



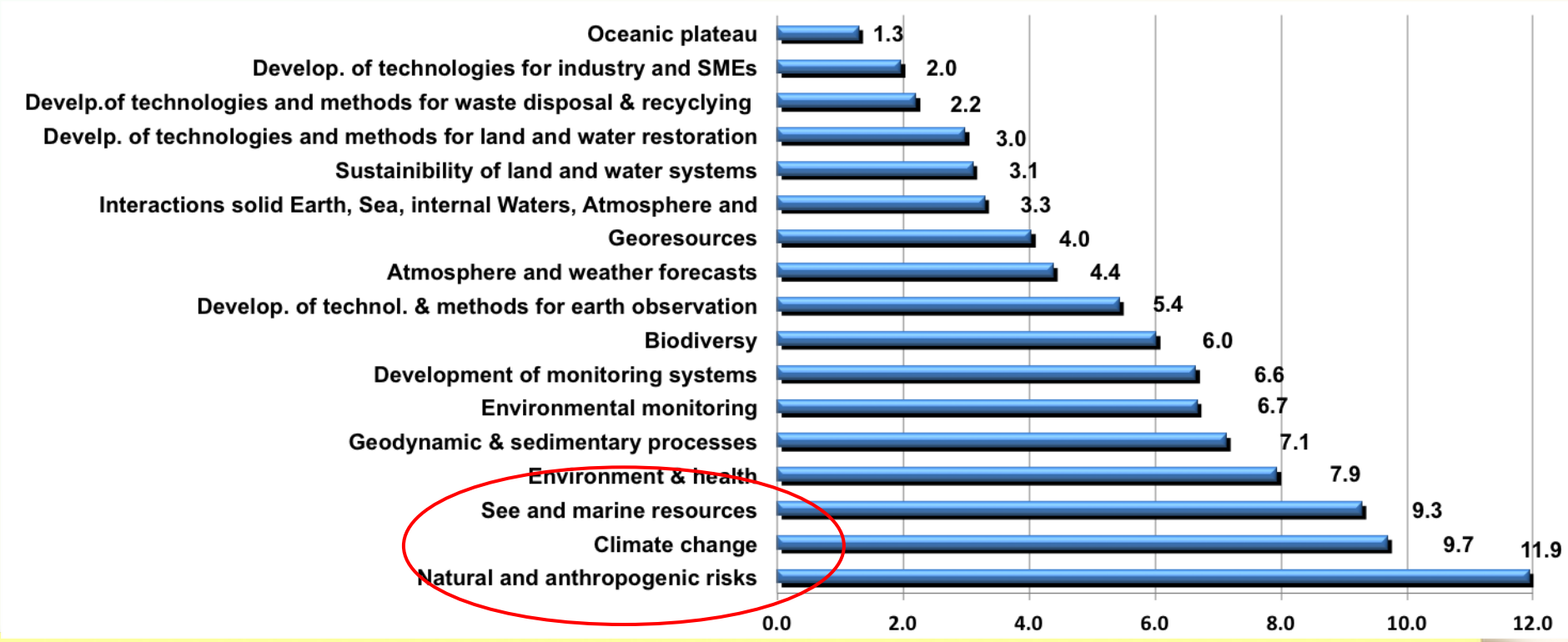
Type of contract



Results:

Research context and practices

Research context



- ✓ On average research is funded by national public funds = **47.8%** and EU projects =**28.9%**
- ✓ Research activities are mainly carried out in medium size groups (3 to 7 persons) = **47.7%**
- ✓ **42%** of researchers always work in multidisciplinary groups in International projects, and **37%** with other colleagues of the same Institute

Data acquisition

Mainly experimental (77%)

	Biological	Chemical	Physical	Geological
Land	18.2	30.8	31.9	44.2
Sea	26.4	29.8	30.6	25.8
Internal waters	22.4	36.3	30.0	26.8
Atmosphere	6.5	27.3	42.1	13.4
Biosphere	28.3	24.3	21.0	16.4

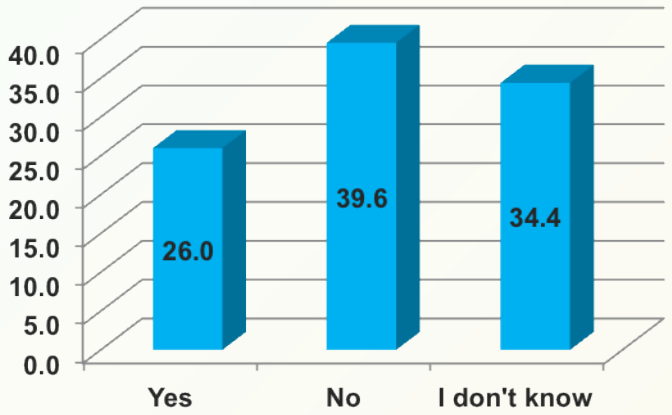
21.2% of respondents used also demographic data

How data are acquired:

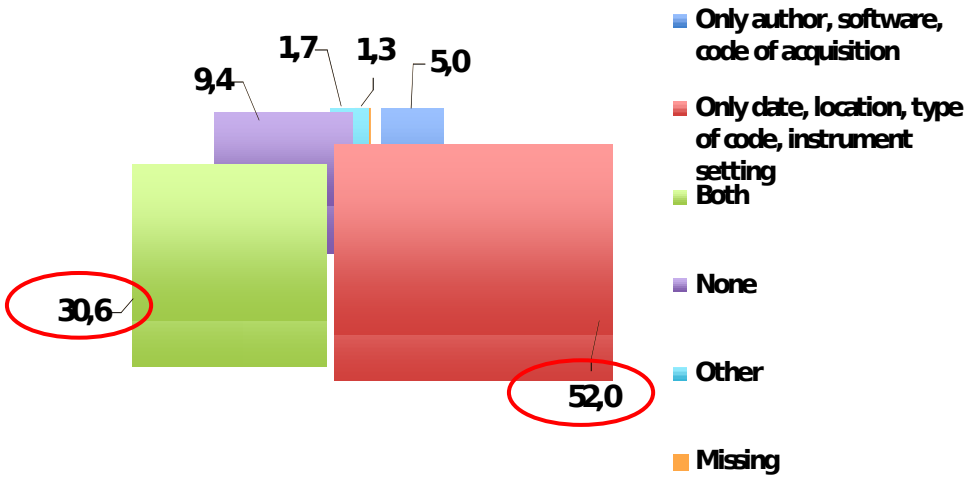
✓Measurements are mainly taken directly by researchers = **53%**
✓ Data are mainly acquired both in laboratory and in the field = **53%**
✓ Using instrumentation directly managed by CNR = **83%,** but also by other national organizations **(26%)**

Data management

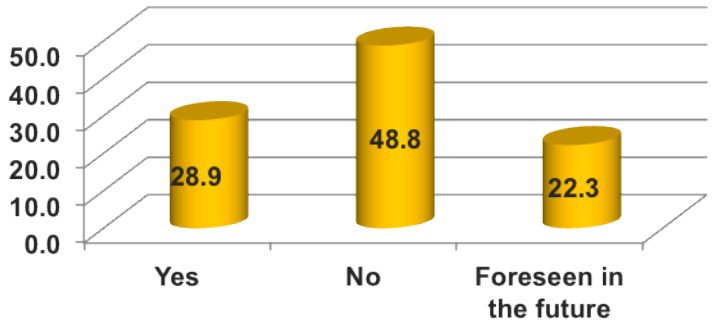
Does you community of reference use standards to manage data?



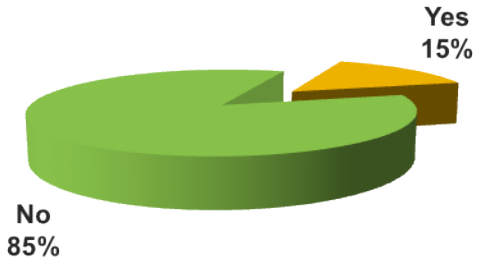
What type of additional information do you associate with data collected/analysed by yourself?



Does your institute have set up procedures for data preservation?



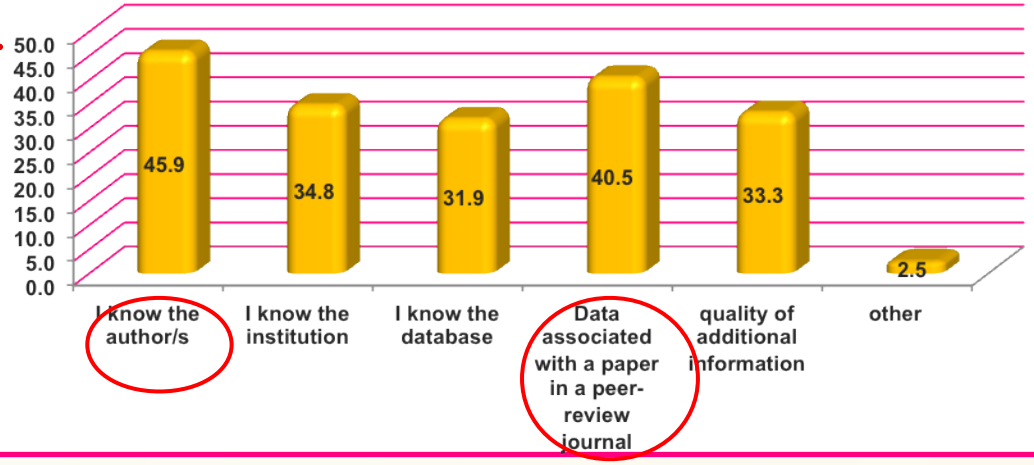
Is anyone in your Institute specifically trained to preserve data?



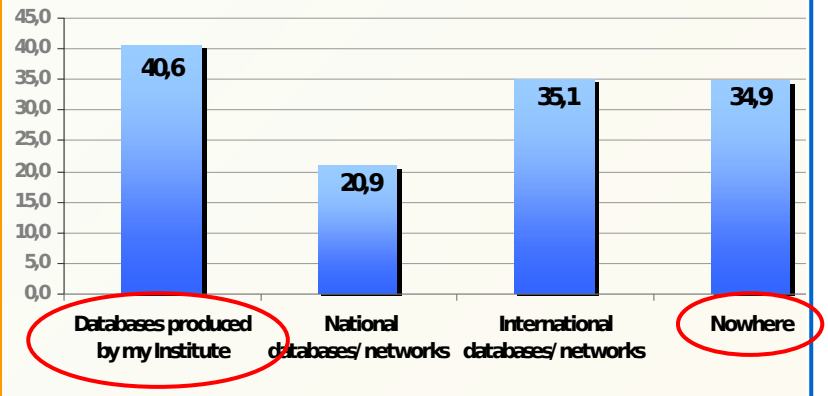
Data use

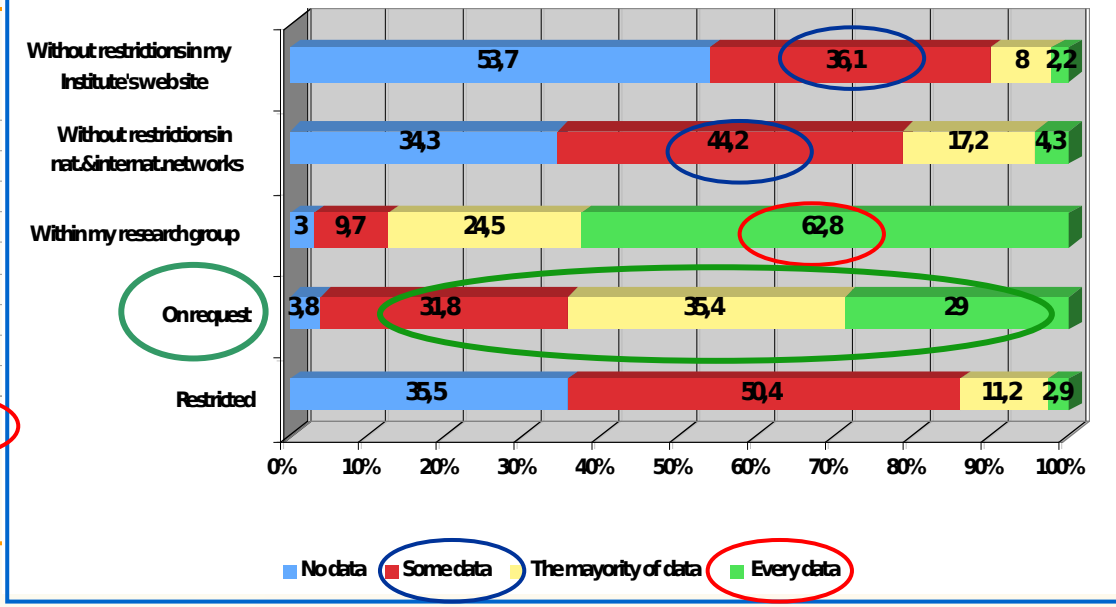
- 59% uses data produced by others, mainly in his/her disciplinary field (= 43%)

Data produced by others are considered reliable when..



Data availability





Results:

Researchers' attitudes

In your opinion for which reasons is it important to make research data available and preserve them?

	Very important	Important	Not very important	Not important of all	Missing
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

In your opinion what are the main obstacles of data sharing?

	Very important	Important	Not very important	Not important of all	Missing
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

What conditions would you require to submit your research data to an open archive?

	Very important	Important	Not very important	Not important of all	Missing
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

Conclusions

High level of awareness

- 48% of voluntary responses to the questionnaire;
- Data sharing and preservation are very important as it fosters the progress of science and enhances the transparency of research
- Publicly funded research should be available to anyone

Willingness to share correlated with actions

- 2.2% are not, 69.8% depending on the circumstances
- The majority of researchers store their data in local, national and international databases, but they tend to share only a fraction of them

Obstacles are connected with data management and Concerns are mainly related to the formal acknowledgment of time consuming data sharing

Motivations: a friendly environment to deposit data related to data storage, management and re-use