

Science for Society

Sharing research products in biosciences, issues and perspectives

Istituto Italiano di Antropologia

February 22, 2013

Department of Environmental Biology, Sapienza University of Rome, Anthropology building

The workshop entitled "Science for Society - Sharing research products in biosciences, issues and perspectives" took place at the Department of Environmental Biology-Sapienza University of Rome on 22th February 2013.

The meeting was attended by Giovanni Destro Bisol, Paolo Anagnostou and Marco Capocasa from Sapienza University of Rome and Istituto Italiano di Antropologia, Cecilia Carloni from DigiLab-Sapienza University of Rome, Ilaria Fava from CINECA, Pietro Greco (science journalist), Daniela Luzi from CNR-IRPPS and Fabio Parenti from the Istituto Italiano di Paleontologia Umana.

Pietro Greco opened the meeting with a discussion on the importance of open access for the dissemination of scientific knowledge. "Communicating everything to everyone", stated Greco, is the message for the promotion of a new science founded on the full, effective and free access to data and scientific knowledge.

After this introductory lecture, the debate has been focused on the "sharing behavior" in different research fields. Particular attention was dedicated to the open access strategies, data availability, management and reuse and barriers to data sharing. Participants presented their professional experience regarding data and knowledge sharing, explaining their point of view on how to contribute to an "open science for society".

Giovanni Destro Bisol closed the workshop presenting the initiative "Opening Science to Society", aimed at the development of a robust and efficient sharing of research data, supported by the Istituto Italiano di Antropologia.

Empirical investigations of data sharing in human population genetics

Paolo Anagnostou

Istituto Italiano di Antropologia

paolo.anagnostou@uniroma1.it

The achievement of a robust, effective and responsible form of data sharing is currently regarded as a priority for biological and bio-medical research. However, it has been argued that its possible advantages in terms of better exploitation of data and optimized use of resources may be counteracted by the time and economic costs required and by underlying ethical concerns. In this contrasting scenario, empirical evaluations of data sharing may be regarded as an indispensable first step in the identification of critical aspects and the development of strategies aimed at increasing availability of research data for the scientific community as a whole. Using an ad hoc designed procedure we analyzed a total of 543 mitochondrial and Y chromosomal datasets reported in 508 papers indexed in the Pubmed database from 2008 to 2011. A substantial portion of datasets (21.9%) was found to have been withheld, while neither strong editorial policies nor high impact factor proved to be effective in increasing the sharing rate beyond the current figure of 80.5%. Disaggregating our data by research field we show that the sharing rate for Forensic genetic datasets is elevated (86.06%), and higher than observed for Medical and Evolutionary genetics (63.72% and 79.25%, respectively). The results of the questionnaire based survey indicate ethical/privacy issues and data misuse as the main barriers to data sharing. In conclusion, we believe that developing a model based on cooperation and scientific rigor may contribute to a robust and efficient data sharing. Both are really needed to foster scientific progress and exploit the huge amount of data promised by next generation techniques.

Sharing data and knowledge in human genomic studies

Marco Capocasa
Istituto Italiano di Antropologia
marco.capocasa@uniroma1.it

Data sharing leads to many advantages for the progress of human genomics. However, making data freely available is also anything but simple to put into practice. Privacy and confidentiality violations may occur and secondary use of genomic data without previous donors' agreement cannot be excluded. These issues are common for all those who work with human subjects. Nonetheless, molecular anthropologists face the additional challenge to share their knowledge with the investigated communities. From the first contact to the final explanation of the results, researchers and participants mutually exchange information and publicly discuss the implications and interpretation of scientific outputs. This extended meaning of "sharing" can become an integral part of the experimental protocol in molecular anthropology. It is clear that anthropologists need to share data with the scientific community for the advancement of their discipline. But they also have the unique opportunity to share their knowledge, experience and feeling with non-experts and, thus, to help make science and society closer each other.

Research data in environmental sciences: A Survey of CNR researchers

Daniela Luzi
Consiglio Nazionale delle Ricerche. Istituto di Ricerche sulla Popolazione e le Politiche Sociali
d.luzi@irpps.cnr.it

The presentation provides 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. Results do not reveal significant differences compared to other international surveys. CNR researchers are aware of the importance of research data as shown by the rate of responses as well as by their opinions on the reasons for the availability and preservation of data. 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.

The European Commission towards publications and data sharing

Ilaria Fava
CINECA
i.fava@cineca.it

In 2008, the European Commission started a pilot on Open Access to research publications funded by the EC within the Seventh Framework Program. The OpenAIRE project (<http://www.openaire.eu>) puts in place the pilot, offering a one-stop-shop to take a look at the research produced in Europe. As a follow-up of OpenAIRE, the EC funded OpenAIREplus on research data, and on the ways to link data to publications. There is a strong feeling on research data openness and sharing at the European Commission, which is going to finalize its new research framework program (Horizon2020) making Open Access to publication the default value, and launching a pilot on OA to research data. Being involved in OpenAIREplus as National Open Access Desk, in order to understand current practices of research data production, management and sharing in Italy, CINECA decided to perform a survey on this topic among Italian researchers based in

Universities and research institutions. Although potential respondents would have been over 60000, the total number of responses is very low but not discouraging (e.g. respondents say they are aware of data sharing practices and trends). A report on the survey is in progress and will be delivered by the end of April, 2013.

The geographic database of Italian prehistory

Fabio Parenti
Istituto Italiano di Paleontologia Umana
scriptoriumparentii@gmail.com

Since 2003 the author, a prehistoric archaeologist, has planned and (partially) realized a relational database of the available scientific literature on Italian prehistory and palaeo-environmental sciences (Parenti 2007). *PreBiblio*, acronym of Prehistoric bibliography of Italy, will comprise a core of about 60,000 bibliographic references from more than 350 scientific journals and 5,000 books relating on earth sciences, archaeology, palaeoanthropology, palaeontology on every published site from lower Pleistocene (about 2.5 million years ago) till first Iron age (IX century B.C.). The expected number of sites all over Italian peninsula is about 25,000, with a total 700,000 links between references and sites. At the moment (winter 2013) the database has more than 15,000 localized sites from 6,000 references. All sites, at different levels of accuracy, will be geo-referenced on a national topographic map at 1:25,000 scale and visible on a web-GIS. A first version of the database will be available online thanks to a joint project with Digilab of Sapienza University of Roma till the end of 2013. The targeted stakeholders of the projects are researchers, students, planners.