

Opening Science to Society, a new initiative of the *Istituto Italiano di Antropologia*

Giovanni Destro-Bisol, Marco Capocasa, Paolo Anagnostou & Pietro Greco

Istituto Italiano di Antropologia, Roma, Italy

e-mail: isita@isita-org.com

The processes of production and dissemination of scientific knowledge are being radically changed by the advent of the digital era. To what extent this ongoing revolution might foster the advancement of science depends on our ability to make information available to the whole scientific community. A form of efficient data sharing is, therefore, a necessary first step to better exploit the potential of new scientific discoveries and achievements. In line with this corollary, in 2011 the *Journal of Anthropological Sciences* started a policy which encouraged the depositing of data in the Anthro-digit^{data} (Anagnostou & Destro Bisol, 2011). This online archive has been created with the aim to store and disseminate original results relating to studies published in the *JASs*, but which also accepts primary datasets used in papers published in other journals (and hence validated) that have yet to be deposited elsewhere.

As a further step towards the adoption of an open science philosophy, the *Istituto Italiano di Antropologia* is now launching *Opening Science to Society*. This is an initiative which aims to deal with scientific, educational and public aspects of data sharing through a collaborative effort of evolutionary anthropologists and researchers from other disciplines.

The *Opening Science to Society* web site is a workspace to be shared with all those who believe that data sharing is an important means to advance scientific progress and to open science to society. At present, it gives access to: (i) a brief synopsis of the initiative; (ii) updated information about our ongoing activities; (iii) a forum for the discussion of scientific, educational and ethical aspects; (iv) an updated list of articles

concerning data sharing; (v) numerous links to scientific and educational resources.

The activities carried out in the framework of the project focus on three main aspects.

You can't manage what you don't measure

Knowing to what extent and in what ways data are actually shared is a fundamental step in identifying barriers to data sharing, evaluate their impact and, finally, design strategies which better suit scientific practices (Piwowar, 2011; Tenopir *et al.*, 2011).

We recently presented an overview of databases for mtDNA and Y chromosome polymorphisms in human populations in the *Journal of Anthropological Sciences* (Congiu *et al.*, 2012). This work had two aims: to provide researchers and practitioners with practical information which may help them optimize the use of these tools and to identify and discuss their strengths and weaknesses. Our work highlighted some positive aspects such as: i) the usefulness of secondary databases to complement the primary ones (e.g. GenBank); ii) the access they offer to types of polymorphisms that have yet to be implemented in primary databases (e.g. Y-chromosome microsatellites); iii) the accurate quality control provided by some of these tools.

Based on the scrutiny of a total of 543 datasets reported in papers published between 2008 and 2011, we have analyzed the various ways in which data are shared or withheld in studies of human genetic variation (Milia *et al.*, 2012). We observed an unexpectedly high percentage of

withheld datasets (21.9%), and estimated that approximately 30% of research funding is used to produce withheld data, a finding which is of particular interest in the discussion regarding Science and Society.

In a more recent paper based on the same dataset (Anagnostou *et al.*, 2013), we proposed that the elevated sharing rate of forensic genetic datasets (86.06%) might be explained by two factors. The proximate reason could be identified in the effective editorial policies of the two main journals in the field (*Forensic Science International: Genetics* and *International Journal of Legal Medicine*). The cooperative efforts to develop common standards and achieve full reproducibility of genotyping techniques in the forensic genetic community could instead be regarded as to the remote reasons. Further work is currently in progress.

We are now focusing on data sharing in ancient DNA studies, a research field where complying with gold standards for data quality is particularly challenging and necessary (Anagnostou *et al.*, in preparation).

Educating for the future

The importance of education of trainees and young researchers does not seem to have been given appropriate consideration in the current debate on open science. We are working on a set of educational modules through which students,

while getting to know the pros and cons of data sharing, may understand its connections with models of scientific progress, become aware of the importance of climate of research and get acquainted with ethical principles and issues. In the course of the lessons, students may get a more complete view of the research lifecycle and learn to distinguish the role played by the different actors of scientific production through the discussion of barriers to data sharing and of strategies to reduce withholding. As a final outcome, we expect students to be able to make a synthesis of different sources of information, developing a vision of science which goes beyond the contents of their own research field and the traditional boundaries of biological education.

Bridging science and society

We intend to promote an active participation of the public in the discussion concerning scientific data sharing. Such involvement should go beyond the return of results to study participants (Knoppers *et al.*, 2006; Bredenoord & van Delden, 2012) and cover the various aspects of the research lifecycle. We consider this action line as an important opportunity to favour dialogue among researchers, policy makers and the public (Marris & Rose, 2010). To this aim, conferences, public debates and on-line forums will be organized. These activities will be integrated with the publication of popular papers. A volume of the

Info on the web

www.isita-org.com/Anthro-Digit/data.htm

Anthro Digit^{data}, an online repository for datasets from papers published in the Journal of Anthropological Sciences.

<https://sites.google.com/site/openingsciencetosociety/>

The web site of the "Opening Science to Society" initiative.

<https://sites.google.com/site/scientificdatasharing/Presentations>

Site of the meeting "Scientific data sharing: an interdisciplinary workshop" (Anagni, Italy, 2-4 september 2013), with slides and videos of presentations.

magazine “Scienza e Società” dedicated to various aspects of open Science and data sharing is now in preparation under our supervision (P.G.).

In the framework of *Opening Science to Society*, we recently organized “Scientific data sharing: an interdisciplinary workshop” (Anagni, Italy, 2-4 september). This meeting was designed to provide researchers from different disciplines (e.g. Anthropology, Sociology, Bioethics, Psychology and Archeology) with an opportunity to start a discussion regarding the foundations of Open Science, barriers for data sharing in scientific practice and the implications of the increasing diffusion of open data in the complex relationships between science and society. A meeting paper summarizing all the presentations and relative debates is now being prepared (Destro Bisol et al., in preparation).

References

- Anagnostou P. & Destro-Bisol G. 2011. Anthropological data, an online resource for anthropological data sharing. *J. Anthropol. Sci.*, 89: 211.
- Anagnostou P., Capocasa M., Milia N. & Destro-Bisol G. 2013. Research data sharing: lessons from forensic genetics. *Forensic Sci. Int. Genet.*, 7: e117-119.
- Bredenoord A.L. & van Delden J.J.M. 2012. Research ethics in genomics research: feedback of individual genetic data to research participants. In J. Schildmann, V. Sandow, O. Rauprich & J. Vollman (eds): *Human Medical Research*, pp. 127-136. Springer, Basel.
- Congiu A, Anagnostou P, Milia N, Capocasa M, Montinaro F & Destro-Bisol G. 2012. Online databases for mtDNA and Y chromosome polymorphisms in human populations. *J. Anthropol. Sci.*, 90: 201-215.
- Knoppers B.M., Joly Y., Simard J. & Durocher F. 2006. The emergence of an ethical duty to disclose genetic research results: international perspectives. *Eur. J. Hum. Genet.*, 14: 1170-1178.
- Marris C. & Rose N. 2010. Open engagement: exploring public participation in the biosciences. *PLoS Biol.*, 8: e1000549.
- Milia M., Congiu A., Anagnostou P., Montinaro F., Capocasa M., Sanna E. & Destro-Bisol G. 2012. Mine, yours, ours? sharing data on human genetic variation. *PLoS One*, 7: e37552.
- Piwowar H. 2011. Who shares? Who doesn't? Factors associated with openly archiving raw research data. *PLoS One*, 6: e18657.
- Tenopir C., Allard S., Douglass K., Aydinoglu A.U., Wu L., Read E., Manoff M. & Frame M. 2011. Data sharing by scientists: practices and perceptions. *PLoS One*, 6: e21101.