

Jerusalem Declaration

on

Data Access, Use and Dissemination for Scientific Research.

The tools and resources of Information Technology, robotics and the development of computer-assisted services have changed the nature of several disciplines. Specifically, several scientific disciplines once characterized by a low rate of data production have recently become disciplines with a huge rate of data production. Today a huge amount of data easily accessible in electronic form is produced by both research and, more generally, human activity.

Scientific research is both producing data and benefiting from the production of data performed by the society.

To move towards the objective of an optimal degree of access, use, data mining, conscious use, and dissemination of data the undersigning scientists present the following statements, promote the following actions and suggest the following policies:

For the social recognition of the problem:

-) Supranational organizations, States, Funding agencies and Research institutions should recognize that *information infrastructures* are, and will be even more in the future, essential resources for the progress of the scientific knowledge.
-) Supranational organizations should coordinate national projects collecting large databases to enlarge and harmonize the national databases in international information infrastructures.
-) Research societies should set up committees selecting protocols for the optimal access, use and dissemination of data produced by the research activity or by the society.
-) States, Funding agencies and Research societies should take into consideration (i) the “OECD Principles and Guidelines for Access to Research Data from Public Funding” endorsed by OECD Council on 14 December 2006, (ii) the results of the 1997 CODATA studies on effective access to data for scientific research purposes in the natural sciences and (iii) the principles pertaining to research and statistical uses of data held by government agencies proposed at the Bellagio Conference in 1977.

For the promotion of access and use of data in scientific research:

-) Data produced while performing publicly funded scientific research should be made available to all scientists willing to use them along with protocols of data recording and data assessment. Scientists using data produced by other research groups need to acknowledge the origin of data. Failures to do this should be considered as scientific misconduct.
-) Scientists should organize the data produced by their research activity in a way, which makes the access to their data simple for other scientists. Funding agencies should support research consortia, Internet based repository and programs

promoting data sharing both by promoting these aspects in research grants and with special grants only devoted to these aspects.

-) Private companies producing large quantity of data should be given public incentives to make their data accessible to the scientific community. To protect intellectual rights and legitimate priority exploitation data could be disseminated after a period of exclusive use.

-) Private companies operating under state license and producing large quantity of data should make their data available for research (in part or fully). A company should be made contractually responsible for data dissemination when it contracts the license by the State. The company could make data available after a significant period of time has passed from the time when events or information are recorded.

-) Public entities collecting and producing large quantity of data should make their data available for research (in part or fully). These entities should include in their aims data dissemination. These public institutions could make data available only in a coded or surrogated form and, if national policies require that, after a significant period of time has passed from the time when events or information are recorded for security or privacy reasons.

For the development of data access and data mining whilst ensuring privacy issues:

-) Research on data conveying private information on individuals, groups or business organizations should be performed by protecting the right to privacy of the agents investigated.

-) Data containing private information should be carefully managed while performing research projects to prevent any illegal use or dissemination of the information or any abuse of third parties. Scientists need to take appropriate measure to ensure the best practice of managing such kind of information.

-) Public data with sensitive information should not be made publicly available to researchers for security reasons. However, surrogated versions of these sets of data might be prepared by specialized research groups and made publicly available for research.

-) Supranational organizations, States, Funding agencies and Research institutions should promote the discussion about protocols and policies for wide distribution of data conveying privacy issues among scientists in a way that prevent the possibility of disclosure or exploiting of private information.

We are entering the era of a high rate of production of information of physical, biological, environmental, social and economic systems. The recording, accessing, data mining and dissemination of this information affect in a crucial way the progress of knowledge of mankind in the next years. Scientists should design, explore and validate protocols for the access and use of this information able to maximize the access and freedom of research and meanwhile protect and respect the private nature of part of it.

Jerusalem, September 2008.