Archives of Science: An International Perspective and Comparison on Best Practices for Handling of Scientific Records

Renata Arovelius

Abstract

The main activity of a university is research and education. Universities are thus important creators of scientific records. But there is also a range of other institutions, both private and governmental, producing scientific material. Conditions for handling of this material will differ as the creators are different organisational bodies acting under different legal premises and internal rules. The understanding and interpreting of the term "science" will also differ due to different national traditions. Results from an international survey on records of science and a case study from Swedish universities illustrate this problem.

The new information technology has significant impact on the work of scientists and on the scientific environment. Science "became" digital and the amount of data and records produced, is increasingly larger. There is also growing scientific cooperation across countries, institutions and disciplines. The crucial issue for establishing of best practices for handling of scientific records created during the scientific process is appraisal and the reliability of digital records. The appraisal strategy must focus on raw and analyzed data; securing of the reliability demands establishing of a sufficient archival system. The OAIS Reference Model gives a good ground for an appropriate technical implementation.

Contents

1. Introduction:

The Subcommittee on Science Archives within the ICA/SUV initiated and carried out during 2003/2004 an international survey on records of science. The survey was sent to all SUV members and also handled out to participants in the SUV annual seminar in Poland, in September 2003.
18 different institutions in 12 countries answered the survey.[1]

2. The results of the survey

2.1 The terms "science", "research", and "records of science"

The first three questions of the survey aimed to make clear the extent of the definition of science in different countries, the term "research" and the term "records of science".
11 respondents (12 including Sweden) from Israel, Poland, Czech Republic, Austria, Canada, Iceland, Brazil, France define science as "all sciences", which means physical, life sciences, mathematics, engineering, philosophy, history, languages, social sciences etc.
In the USA the term science indicates "all sciences except social sciences and humanities". The Royal Academy of Belgium gives the same definition. The term science means in UK "all sciences except social sciences", but it is important to specify separately medicine, health and related sciences in addition.
In Australia this term points towards "all sciences excluding humanities and some social sciences", but not all social sciences. The interpretation of this term may also differ in relation to current organisation of the university.
The term research implies for the majority of respondents a detailed study of a subject that can be academic and non-academic. In Austria, UK, Belgium, Iceland, France and Sweden this term denotes however primarily academic activity of detailed study, analysis, writing and reporting of results.
The definition of records of science varies considerably not only at international level, but also national across different institutions. In Austria and Belgium records of science are interpreted as records created during a scientific project. The University of Tel-Aviv in Israel has the same definition, but the University of Haifa describes records of science as records created during a scientific project, the entire research activity and also records concerning a specific scientist. The University of Michigan has also all those definitions and includes even private papers of a scientist into this term. The University of Illinois defines records of science as all records concerning a specific scientist and private papers. The Observatoire de la Cote d´Azur in France gives all definitions and supplements with a definition of records of science as also records created during the administration of science. In Brazil this term means records created during a scientific project and also all records concerning a specific scientist including private papers. Poland, Czech Republic, Iceland and Sweden interpret this term as records created during the entire research activity. In Australia is this term even broader and includes also records of an organisation, results of scientific activity and cultural context. In Canada records of science denote also, like in Australia, records of scientific organisation. UK gives definition of records of science as all records relating to scientific activity.

2.2 External and internal rules available for handling of records of science and owner of records

As the definition of records of science varies, there is a major difference in rules applied on those records. Records of science considered as private papers of a scientist will demand quite

[1]See the separate combination of all answers

different rules than records of science measured as records created during a scientific project at a public owned university. Poland (Warsaw School of Economics, Textile Museum), University of Haifa in Israel, University of Sydney in Australia and Swedish universities apply archival law and institutional or university policy on scientific material, while Iceland only archival law. Czech Republic relates handling of records of science to archival law and request of the donor; the similar situation occurs in Austria, but also supplemented by the university policy. Canada and Brazil apply both archival law and university policy completed by rules stated by the research team. Science and Technology Heritage Centre at University in Melbourne, in Australia and University of Michigan in the USA have the same extension, but also completed by the request of the donor. Archives of the Polish Academy of Sciences apply archival law, institutional policy and request of the donor; the Royal Academy in Belgium applies only institutional policy, while University of Illinois in the USA institutional policy and request of the donor. The Obsevaroire de la Cote d´Azur in France applies archival law and request of the donor. Wellcome Library in UK mentions other rules as for example Data Protection Act.

To define the "owner" of records of science in a homogeneous way seems not to be easy either. The scientific material is considered public property in Czech Republic (Academy of Sciences), Iceland (University of Iceland), France (Observatoire Cote d´Azur), and Swedish universities. University of Tel-Aviv in Israel, Academy of Sciences in Poland and University of Michigan in the USA indicate records of sciences as private property – the University of Michigan on a condition that it is not a question of a private donation. The Royal Academy in Belgium considers records of sciences as the institutional property. According to the University of Haifa in Israel and the CEDIC (Centro de Documentaçao Cientifica) in Brazil records of science can be owned by the state, institution or can be private. In Austria and the USA/University of Illinois the ownership is unclear respectively unsolved.

In Australia/University of Melbourne the ownership depends on copyrights; there is a possibility the records may be public. In Canada the rules are similar, but the ownership may also belong to the university. According to the University of Sydney/Australia the ownership can differ for each set of records. Records created by university staff as a part of their employment are covered by legislation and are "state records". Different rules are valid for commercially funded research.

The Wellcome Library in UK specifies the ownership as depending on scientist, institution or funding body.

2.3 The ways of collecting/ingest of the material to the archives and types of records

The answers given to the question regarding ways of "collecting" records of science exemplify results of implementation of different rules applied in different countries and institutions. Delivery of scientific records in accordance with archival law only can be found at the Swedish universities; Academy of Sciences in Czech Republic collect records in accordance with archival law, too, but also by adopting persuasion of the scientist, purchasing and as a donation. The Observatoire Cote d´Azur in France implements the same ways with the exception of purchasing.

Scientist or scientific group is obliged by the university policy to deliver records of science to the archive at University of Haifa in Israel; a donation is a way for collecting at University of Tel-Aviv in Israel, University of Iceland and the Royal Academy in Belgium. Archives persuade scientists to donate records at the Polish Academy of Sciences and at Science and Technology Heritage Centre (AUSTEHC), at the University of Melbourne in Australia. In Canada, University of Ottawa and at the University of Sydney in Australia are in force rules in accordance with retention schedule or rules for disposal. The University of Sydney collect also records by persuasion and as a donation.

The University of Illinois in the USA, CEDIC in Brazil and School of Economics in Poland collect records by persuasion and as a donation.
The Textile Museum in Poland implements institutional policy and persuasion in its activity for ingest to the archives. The Wellcome Library in UK responds to the request to "give home" to the records.

There is a huge variety of different types of records depending on scientific discipline. However, the scientific process during a research activity may be described from the start to the finalization, in following steps: 1. setting of a hypothesis and choice of methods, collection of raw data, 2. analysis, explanation and interpreting of data, 3. valuation and critical review 4. reporting of results 5. inception for a new research activity. Types of scientific records can thus be listed, in general, as follows:
Projects description or plan, funding application, budget, contracts, correspondence with sponsors, application to ethical committees in some cases, methods description, raw data, analyzed data, research results.
Personnel records are a sole type of scientific records.
All those types of scientific records, and even other records distinctive for a particular discipline, can be found at Tel-Aviv University in Israel, Science and Technology Heritage Centre and the University of Melbourne in Australia, and also Swedish universities.
University of Illinois in the USA and University of Ottawa in Canada have all those types, too but with the exception of personnel records. Academy of Sciences in Czech Republic and Vienna University of Technology in Austria receive all types as well with the exception of applications to ethical committees. The Observatoire de la Cote d´Azur do not have this type of records either and does not keep reporting of the results in its archive.
Analyzed data, application to the ethical committees and methods description do not occur in the Archives of Polish Academy of Sciences; the University of Haifa does not keep any projects descriptions, raw and analyzed data, and applications to ethical committees.
University of Iceland does not have applications to ethical committees either and furthermore no raw and analyzed data. Funding applications, applications to ethical committees and analyzed data do not occur in CEDIC in Brazil, but the archive does keep additionally popular musical records.
The Royal Academy in Belgium keeps only methods descriptions, analyzed data and reporting of results. The Textile Museum in Poland receives to the archives projects descriptions, funding applications, correspondence, articles, exhibits manuscripts and personnel records. The Wellcome Library in UK takes care of correspondence, diaries, notes and photos.

2.4 Records created and kept electronically and policy for digital preservation.
Different archival traditions make it difficult to answer unequivocally if there are records of science kept electronically. In cases when records management is separated from archives management the answer often will be negative: those records are not in archive yet[2], even though the records are created electronically.
Science and Technology Heritage Centre in Australia, the University of Ottawa in Canada and Warsaw School of Economics, also Swedish universities, keep all listed types electronically. However it does not mean the long term preservation of those records is established.
Tel-Aviv University in Israel, Vienna University of Technology in Austria and CEDIC in Brazil state that none of records is in archive. The Academy of Sciences in Czech Republic, The Royal Academy in Belgium, Polish Academy of Sciences and University of Haifa in

[2] see the answers from University of Sydney and Michigan

Israel keep only the research results electronically. According to the University of Illinois in the USA there may be electronic data in some projects. University of Iceland does not keep electronic records of science; they are kept by a special data centre and scientists.
The Textile Museum in Poland keeps interviews with artists electronically and the Wellcome Library in UK correspondence between scientists.
The question about an agreed policy on digital preservation has been answered with "yes" by two institutions: Polish Academy of Sciences and the University of Haifa in Israel.
Science and Technology Heritage Centre in Australia works on behalf of other institutions and consequently does not have a policy for preservation. Discussions about future strategy and projects are undertaken at CEDIC in Brazil and at Swedish universities.
The majority of respondents specify methods of preservation as storage in native formats. The University of Illinois however applies also emulation and migration. Migration is also adopted at Swedish universities and in the Textile Museum in Poland. AUSTECH in Australia uses non-proprietary formats and open filing system. At the University of Iceland it is a scientist who makes the decision about the preservation method and format.

2.5 Responsibility for the decision on appraisal of records of science

As conditions for handling of scientific records and also types of records differ considerably, the responsibility for the appraisal will be put on different levels.
At Warsaw School of Economics in Poland and at Haifa University in Israel appraisal and disposal decisions are made by the university board. At the University of Haifa it is also a scientific group that makes those decisions. Responsible for appraisal at the University of Iceland is the scientists group too, respectively a scientist conducting the research project.
Archivists are involved in the appraisal process at Tel-Aviv University in Israel, the Textile Museum in Poland, the University of Illinois in the USA, at the Academy of Sciences in Czech Republic, Vienna University of Technology in Austria and Polish Academy of Sciences. At Tel-Aviv University, Vienna University of Technology and the University of Illinois, it is also the scientist who is involved in the appraisal process. At the University of Ottawa in Canada the decision is made by the archivist in conjunction with the administrative committee and in Australia/ AUSTEHC, with scientists and department heads. At the SLU in Sweden the main responsibility belongs to the department head in conjunction with the archival function. Archivist makes the decision on appraisal when the State Records Act does not apply at the University of Sydney in Australia. At the University of Michigan in the USA, there are also state and federal laws regulating this process. At the Royal Academy in Belgium, it is the chief of department, who makes the decision.
According to the answer from the Observatoire de la Cote D´Azur in France, the appraisal decision is too often a hazard. This situation may be significant for many institutions.

2.6. Types of scientific collections in different archives and institutions

The scientific material kept in the archives varies, as the interpreting of the term "records of science" differs and various traditions are manifested in the archival practice.
1). The University of Michigan keeps scientific records as the part of the routine function to maintain the history of the University; the Observatoire de la Cote d´Azur is in the possession of archives on scientific and non-scientific history of the Institution.
2). The Polish Academy of Sciences has custody of private papers with some material dated from 16^{th} - 18^{th} century and the Wellcome Library of a wide range of materials from manuscripts dating from late antiquity to modern archives of individual scientists and institutions. The Polish Central Textile Museum keeps historical collections about exhibits and also personnel records; the Academy of Sciences in Czech Republic keeps records of scientific institutions and personnel papers; the Vienna University of Technology the

personnel records, too and also manuscripts and publications. The same situation occurs at the Warsaw School of Economics and at Tel-Aviv University. Personal papers in different disciplines are kept at the University of Illinois; also numeric data and administrative records are represented in the collections.
3). The University of Haifa is in possession of data from social sciences and humanities. The University of Sydney has records from a wide range of disciplines and the University of Iceland keeps raw data on earthquake and data from sociology research; also personal papers. The Swedish universities are obliged by law to keep all types of scientific records if the records are official documents and fulfil the criteria for preservation.
4). AUSTEHC in Australia does not keep records in the long term and has only a temporary custody of the scientific material.

3. Handling of records of science in Sweden

3.1 The Principle of Public Access to Official Documents and scientific records

Almost all Swedish universities are public authorities. As a consequence, records of science created during the research process at Swedish universities are subject to implementation of the Principle of Public Access to Official Documents. This Principle is valid for scientific material during the research process itself, which means, even before a research project is finalized and the results published.
The Principle of Public Access was formulated for the first time in the 1766 Freedom of the Press Act and consequently applied in Sweden, with small exceptions, for over 200 years. The Freedom of the Press Act defines also the term "document" and "official document". [3] According to this definition a document may be any representation in writing or picture, but also any record which can be read, listen to or otherwise comprehended only by means of technical aids, i.e. e-mail, data base or microfilm. A document will be official if it is in the keeping, has been received or drawn up by an authority. It means that both: documents forwarded from and submitted to an institution are considered official.
Of course, the access to official documents, so even records of science, may be restricted in aim to protect specific interests, even though the general rule is the public access. Restrictions are always an exception to the general rule; they are specified in a special act which is the 1980 Secrecy Act. At the same time the applicant has always right, guaranteed by the Freedom of the Press Act, to appeal against a decision by an authority not to give access to an official document. All documents created and generated during the activity of an authority/university can thus be described as: *official* (so, kept by an authority, received or drawn up), - and official documents can be public or secret, as mentioned above - , and, of course, *not official* as a lot of documents are just "working" documents, i.e. drafts, suggestions for a decision, first version of a report.

Usually we divide records of science into four main groups:

1. administrative records (e.g. project plans and descriptions, means applications, contracts, correspondence with sponsors etc.)
2. raw data or primary records: all information used for scientific processing (e.g. surveys, laboratory or fields books, radiographs, even soil samples or increment cores)
3. analyzed data ("working" data as e.g. reports draft, excerpts, calculations or electronic records as a part of data processing)
4. all material reporting on the results (all kind of reports, final reports, publications and articles).

[3] The Freedom of the Press Act, SFS 1949:105, section 2 (changed 2002:1049)

Records of science can be secret too, but only if it is a legal authority for classifying as secret. According to the Secrecy Act a good ground for keeping scientific records secret, even results or method used, may be the commercial value, "commissioned" research or security reasons.

3.2 Rules for archiving

As the records of science created during different projects and also during the research activity carried on continuously, *no matter how the project or the activity have been founded* , are subject to the Principle of Public Access, the rules for registration and archiving according to the Archival Law[4], have to be applied to the records of science. And the rules for archiving include the rules for disposal and guidelines for decision making process of appraisal. In this context it is important to notice that we apply in Sweden records management on archives management.

The National Archives of Sweden had decreed 1999 on disposal on scientific records.[5] The decision was a result of a previous analysis from 1997 on the legal status of the scientific material at the universities in the light of the Principle of Public Access. The disposal decision has been valid since January 2000 and gives terms of reference for the appraisal process.

3.3 Implementation of the rules at SLU

The structure and organization of records and archives management at SLU is decentralized. Each department (after the last re-organization of the university we are fully 40) has own responsibility for registration of official documents and for archiving (archives keeping). Each department has thus to implement the disposal decision on scientific material made by the National Archives. The decision makes clear, that it is the University that put the disposal rules into practice and finalizes the cancellation on the terms approved. At the same time it makes also clear that no cancellation should be effected before the evaluation of the records is done. In the evaluation process, which involves the critical keep/destroy decisions, a due attention always must be paid to following aspects:

- records value for the actual research area, but also to other disciplines
- significance for history of science and culture, for individual history and also if particular scientific records are of great interest to the public.

Preservation of records of science must thus always satisfy three main reasons of preservation.

- needs of administration and legal needs, which includes verifying of research results
- scientific needs (access to raw data from previous projects)
- historical needs.

The records of science which contain data about the aim of the project, project description and methods used, and the results must always be kept for ever and are always excluded from cancellation. At the same time, it is important to add that according to the Swedish Archival Law, each transferring of data or conversion to other carrier which causes loss of information, loss of the possibility to compile the information, loss of search possibilities or of the authenticity, is cancellation. And change in media, i.e. transfer from a data base to paper, always means cancellation. Therefore it is very important to consider the appraisal in generally and conversion or media change in particularly as early as possible, which means already at the creation stage as an early step in the evaluation process.

The rules of disposal of scientific material and rules on status of records of science in regard on the Principle of Public Access to Official Documents are relatively new and demand quite a good portion of work for putting them into practise. Information meetings for chief of the departments and researchers have become accordingly a way of communicating the issue of handling and preservation of records of science. Also special courses for postgraduate

[4] SFS 1990:782
[5] RA-FS 1999:1

students and holders of scholarship from abroad are now a well-established instrument for dissemination of knowledge about existing rules to the research community.
New problems and join concern connected to electronic records and long term preservation of electronic scientific records have also stipulated new projects on digital preservation.
Digital preservation is after the evaluation process very important task which must be taken care off as soon as possible and which has to build on cooperation with records generators. This problem includes questions of appropriate funding and continuity, two factors which never can be guaranteed by a department, a unit which is a subject to frequent changes in structure and organization. This also demands quite another system for handling of scientific records, independent of the department. The department is, for different reasons, not any longer a stable basis for long term preservation of primary data. Joint projects, where different parts work together, and also longitudinal field research projects attached to a particular department with the function of a "research hotel" and wide-spread international cooperation, are more and more common. - The organizational issues and the responsibility for digital preservation of scientific electronic records were subject to a faculty project carried out in January 2003.[6] The project steering group consisted of members of the faculty administration and IT-department, scientists and university archivist.
The project group has agreed that the issue of digital preservation can't be separated from the question: who at the university has the responsibility for long term preservation. And as the university structure is the subject to frequent organizational changes, it should be thus the project itself that is the records creator and not, as it is today, the department. The project leader should be responsible for transferring of records to the central archives.

The project group has also recommended a new praxis, a responsibility agreement for documentation and preservation between different parts in the project or research activity, as a condition for cooperation when scientists from other universities or organizations are involved. In similar way an agreement should be reached when different departments are working together. The responsibility for documentation and preservation must be evident from the beginning.
At the same time it is also clear that it is still the individual scientist or scientific group that is responsible to the chief of the department for evaluation and disposal according to archival rules. Chief of the department makes the decision on cancellation and is responsible for systematic transferring of records to the central university archives. Every decision on disposal must be well documented and reasons for cancellation always given. The university archivist shall always be contacted before the disposal is carried out.

The solution for the question of responsibility and establishing of practices for handling of scientific records is only a part of the archival system. Another important issue is the technical solution for the long term preservation of electronic records.
Creating of all data in such a way that it is possible to read it, to understand and to transfer it to other media during the whole time of the preservation period, must always be basis and starting point for long term preservation. Records format is a crucial problem for all planning for future digital preservation. And - as the generating of digital information is very rapid and the longevity very short -, it is important to implement a strategy with a detailed plan for migration and appropriate routines for documentation. The frequency for migration or transfer rates must be adapted to each case in regard on physical lasting of the carrier and logical lasting of the information. The plan for migration must also be a part of the university policy

[6] Digital preservation of scientific records, Faculty of forestry, SLU, 13 January 2003

in relation to responsibility for digital preservation, issues about security and secrecy, accessibility, authenticity and control of digital information.
As the digital preservation of electronic scientific records is a join concern for all universities in Sweden, a join project has been initiated. A project group consisting of archivists from four universities and librarians from two other universities with experience on electronic publishing and metadata standards has been established. The main tasks for the group are: to look after standards, technical and organizational solutions and to set up a pilot project on digital preservation. In May 2004 such a project has been initiated and is taking place at the Swedish University of Agricultural Sciences. Six different scientific data bases will be tested within a technical solution called DSpace, developed by the Massachusetts Institute of Technology in the USA and Hewlett Packard. Scientists, IT-people, librarians and archivists are working together, side by side. Results of the project and suggestions for future strategy shall be reported in June 2005.

4. An attempt for a model on best practices.

Different archival rules and traditions have impact on types on scientific records delivered to the archives, also on strategy for "ingest". However, even though the types of records will differ, the join concern remains the same: how to solve the long term preservation of digital scientific records and how to establish suitable guidelines for appraisal.
There are therefore two crucial issues for establishing of best practices for scientific records: the evaluation process and the reliability of the electronic records of science. Both those issues decide on future access. As a result, a concerted approach is needed if the relevant scientific material has to be saved for the future.

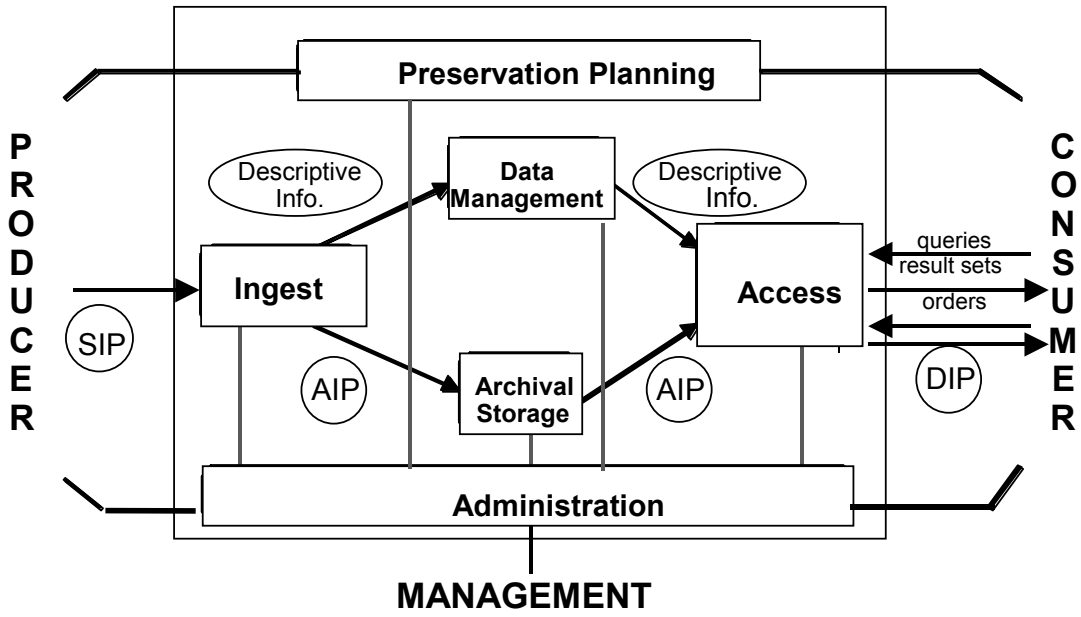
The scientific process has been described above in following steps: 1. setting of a hypothesis and choice of methods, collection of raw data, 2. analysis, explanation and interpreting of data, 3. valuation and critical review 4. reporting of results 5. inception for a new research activity. - The majority of respondents to the survey keep records reflecting this process. And as types of records as: data about the aim of the project, project description and methods used, also the results document the research activity they should be kept for ever.[7]
The appraisal process should thus focus on the evaluation of raw or primary data, in particular cases, also analyzed data. Primary material makes the verifying of results possible and has a central role for other scientists and for the public control. Many scientific projects are also depending on re-use of raw data; the value for future research and for the history must consequently be considered. Various disciplines will have different needs, but in general, those two criteria for evaluation have to be applicable: verifying of the results and reuse for scientific and historical needs, even though it is often difficult to say what future users will need.

4.1 OAIS Reference Model

According to the OAIS (Open Archival Information System) Reference Model [8] an archival system, when talking about the electronic records, comprises ingest, storage, access and preservation planning, and includes hard- and software, and persons responsible for acquisition, preservation and dissemination. The concept and the main ideas of the model are illustrated as follows:

---

[7] Different national and institutional rules may dictate archiving of the results. In many countries it is the library and not archive that takes care of keeping of those records.
[8] See OAIS, Producer-Archive Interface Methodology Abstract Standard, CCSDS-661.0-R-1, Red Book, December 2002



SIP = Submission Information
AIP = Archival Information
DIP = Dissemination Information

DIP = Dissemination Information Package

The model[9] is very useful for all digital archives, and especially for records of science; it can be implemented regardless of types of records or specific legislation on access. This discrepancy or specific regulations, however, must be well fixed in metadata. Metadata in general will be central for long term preservation:
There are three types of metadata feasible to distinguish for the archival reasons:

- descriptive, both content and context metadata
- structural
- administrative

Information is always represented in some way. For digital information it is the way of how data are articulated that provides the understanding of it. Furthermore, for archiving of the digital information you need contextual information about the information that has to be preserved, in other words, preservation metadata: all three types mentioned above. Those types may correspond, and they usually do, to the archival description.
According to the OAIS Reference Modell it can be illustrated as follows:
AIP (Archival Information Package) = CI (Content Information) + PDI (Preservation Description Information).
Content Information consists of Data Object and Representation Information. Preservation Description Information consists of four information elements: provenance, reference, fixity and context.[10]

Specific rules and conditions for handling of records of science will then be manifested in the Preservation Description Information. Consequently, it is of great importance to find a proper level for the description of digital scientific material that guaranties understanding in the

[9] the same, p. 4-57
[10] Reference Model for an Open Archival Information System (OAIS), CCSDS 650.0-B-1, Blue Book, January 2002, p. 4-34

future, but also understanding outside the particular research community; this description has to be accompanied by appropriate guidelines for appraisal.
The generic model of OAIS with a suitable technical application and an appropriate level of description seem to be a good stride in the efforts at long term preservation of digital records of science.

4.1 Suggestion for best practices:
It is not possible to find a solution with "one size fits all"; each scientific discipline has its own needs and particulars. However, it is possible to perceive some general suggestions for best practices:

- *start the dialog with the records creator as early as possible*
- *initiate the evaluation process as early as possible in the life cycle of digital records*
- *monitor value for the actual research area and for other disciplines*
- *monitor importance for history of science and culture*
- *observe needs for verifying of research results*
- *document the aim of the project or research activity, methods used and the results*
- *take into account particular needs of different disciplines: analyzed data may be as important as raw data for re-use*
- *establish an archival information system based on the OAIS Reference Model*
- *set up submission agreements, where sufficient formats are specified*
- *"incorporate" particular rules for access and find a proper level for preservation description information*