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

INTRODUCTION

During 2003/4 the Science Archives Sub-committee within the ICA Section on University and Research Institutions' Archives (ICA/SUV) initiated and carried out an international survey on the records of science. The survey was sent to all SUV members and also distributed to participants in the SUV annual seminar in Poland in September 2003. 18 different institutions in 12 countries answered the survey.

THE RESULTS OF THE SURVEY

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

The first three questions of the survey aimed to clarify definition of *science* in different countries, along with 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. In the UK *science* means "all sciences except social sciences [and humanities]", but additionally it is important to specify medicine, health and related sciences separately. In Australia the 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 primarily the 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 nationally 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 all those definitions and includes even private papers of a scientist within this term. The University of Illinois defines *records of science* as all records concerning a specific scientist and private papers. The Observatoire de la Côte d´Azur in France gives all these definitions and, in addition, *records of science* as those records created during the administration of science. In Brazil the term means records created during a scientific project as well as all records concerning a specific scientist including private papers. For Poland, Czech Republic, Iceland and Sweden it is interpreted as records created during the entire research activity. In Australia the term is even broader and also includes organisational records, results of scientific activity and the cultural context. In Canada it also denotes, as in Australia, records of scientific organisations. The UK gives the definition of *records of science* as all records relating to scientific activity.

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

As definitions of records of science vary, there is a major difference in rules applied to those records. Records of science considered as private papers of a scientist will demand quite different rules from those applied to the records created during a scientific project at a publicly-owned university. Poland (Warsaw School of Economics, Textile Museum), the University of Haifa in Israel, the University of Sydney in Australia and Swedish universities apply archival legislation and institutional or university policy to scientific material, while Iceland applies only archival legislation. The Czech Republic notes that such records are managed according to archival legisltation and the wishes or special requests of the donor; a similar situation occurs in Austria, where university policy applies in addition. Canada and Brazil apply both archival legislation and university policy as well as adhering to any rules given by the research team.

The Science and Technology Heritage Centre at University in Melbourne, in Australia and University of Michigan in the USA have the same practice and, in addition, compliance with the donor's wishes. The Archives of the Polish Academy of Sciences apply archival legislation, institutional policy and wishes of the donor; the Royal Academy in Belgium applies only institutional policy, and the University of Illinois in the USA applies institutional policy and wishes of the donor. The Observatoire de la Côte d´Azur in France applies archival legislation and requests of the donor. The Wellcome Library in UK mentions other rules such as the Data Protection Act.

To define the "owner" of records of science in a homogeneous way does not seems easy either. Scientific material is considered public property in the Czech Republic (Academy of Sciences), Iceland (University of Iceland), France (Observatoire Côte d´Azur), and Swedish universities. The University of Tel-Aviv, the Polish Academy of Sciences and the University of Michigan regard records of science as private property while the Royal Academy in Belgium considers them as institutional property. According to the University of Haifa and the Centro de Documentaçao Cientifica (CEDIC) in Brazil, they can be owned by the state, institution or can be private. In Austria and elsewhere (e.g. the University of Illinois) ownership may be unclear or unresolved.

In Australia ownership depends on copyright; there is a possibility the records may be public but ownership differs for each set of records. At the University of Sydney records created by university staff as a part of their employment are covered by legislation and are "state records". The situation is similar in Canada. Different rules apply for commercially-funded research. The Wellcome Library in UK specifies that ownership depends on scientist, institution or funding body.

Methods of collecting/ingest of the material to the archives and types of records

The answers given to the question regarding methods of "collecting" records of science exemplify the implementation of different rules applied across different countries and institutions. The transfer of scientific records in compliance with archival legislation only can be found at the Swedish universities; the Czech Academy of Sciences collects records in accordance with archival law, too, but also by encouraging scientists to donate, by purchase, and by donation. With the exception of purchase, the Observatoire Côte d´Azur acts in the same way.

Scientists or scientific groups are obliged by the university policy to transfer records to the archives at the University of Haifa; donation is a way of collecting at the universities of Tel-Aviv and Iceland and for the Royal Academy in Belgium. Archives persuade scientists to donate records at the Polish Academy of Sciences and at Science and the Australian Technology Heritage Centre (AUSTEHC), and at the University of Melbourne. Retention schedules are applied at the universities of Ottawa and of Sydney. The universities of Sydney and Illinois collect and encourage the donation of records, as do CEDIC and the School of Economics in Poland. The Textile Museum in Poland implements institutional policy and persuasion to take records into the archives. The Wellcome Library in UK responds to requests to "give a home" to records.

There is a huge variety of different types of records depending on the scientific discipline. However, the scientific process during a research activity may be described from start to finish, in the following steps:

1. setting a hypothesis and choice of methods, collection of raw data
2. analysis, explanation and interpretation of data
3. evaluation and critical review
4. reporting of results
5. inception of new research activity.

Types of scientific records can thus be listed, in general, as follows:

- project descriptions or plans
- funding applications
- budgets, contracts, correspondence with sponsors
- applications to ethical committees
- methods descriptions
- raw data, analyzed data, research results
- personnel records

All these types (and other records specific to a particular discipline), can be found at Tel-Aviv University, the Science and Technology Heritage Centre and the University of Melbourne in Australia, and in Swedish universities. The University of Illinois and the University of Ottawa in Canada have all these types, too with the exception of personnel records. The Czech Academy of Sciences and the Vienna University of Technology in Austria receive all types with the exception of applications to ethical committees. The Observatoire de la Côte d´Azur does 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 are not present in the Archives of Polish Academy of Sciences; the University of Haifa does not keep any projects descriptions, raw and analyzed data, or applications to ethical committees. The University of Iceland has neither applications to ethical committees either nor raw and analyzed data. The Royal Academy in Belgium keeps only methods descriptions, analyzed data and reporting of results. The Polish Textile Museum archives receives projects descriptions, funding applications, correspondence, articles, exhibits manuscripts and personnel records. The Wellcome Library holds correspondence, diaries, notes and photos.

Records created and kept electronically; and digital preservation policies

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 the archive yet (e.g. at the universities of Sydney and Michigan), even though the records are created electronically. AUSTEHC, the University of Ottawa and the Warsaw School of Economics, as well as Swedish universities, keep all listed types electronically. However this does not mean that their long term preservation is guaranteed.

Tel-Aviv University, Vienna University of Technology and CEDIC state that there are no electronic records it their archives. The Czech Academy of Sciences, The Royal Academy in Belgium, the Polish Academy of Sciences and the University of Haifa keep only research results electronically. According to the University of Illinois there may be electronic data in some projects. The University of Iceland does not keep electronic records; they are kept by a special data centre and scientists. The Textile Museum in Poland keeps interviews with artists electronically and the Wellcome Library correspondence between scientists.

As to whether there was an agreed policy on digital preservation, there was a “yes” from two institutions: the Polish Academy of Sciences and the University of Haifa. AUSTEHC works on behalf

of other institutions and consequently does not have its own policy for preservation. Discussions about future strategy and projects are in progress at CEDIC in Brazil and at Swedish universities.

The majority of respondents specified that the method of preservation is storage in native formats; however the University of Illinois also employs emulation and migration. Migration is also adopted at Swedish universities and in the Textile Museum in Poland. AUSTECH uses non-proprietary formats and open filing systems. At the University of Iceland it is a scientist who makes the decision about the preservation methods and formats.

Responsibility for appraisal decisions

As conditions for the management of scientific records differ considerably, the responsibility for appraisal is at different levels. At the Warsaw School of Economics and at Haifa University appraisal and disposal decisions are made by the university board. At the latter those decisions are made by a scientific group. Scientists are also responsible for appraisal at the University of Iceland, specifically a scientist conducting the research project. Archivists are involved in the appraisal process at Tel-Aviv University, the Textile Museum in Poland, the University of Illinois, the Czech Academy of Sciences in Czech Republic, Vienna University of Technology and the 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 the decision is made by the archivist in conjunction with the administrative committee and in Australia, including AUSTEHC, with scientists and department heads. At the SLU in Sweden the main responsibility is with the department head in conjunction with the archival function. When the State Records Act does not apply, archivists make s the decision on appraisal at the University of Sydney. At the University of Michigan, there are also state and federal laws regulating this process. At the Royal Academy in Belgium, it is the head of department who makes the decision.

According to the response of the Observatoire de la Côte d´Azur in France, the appraisal decision is too often haphazard. This may be the case in many institutions.

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 evident in archival practice.

1. The University of Michigan keeps scientific records as the part of its routine function to maintain the history of the University; the Observatoire de la Côte d´Azur holds archives on the scientific and non-scientific history of the Institution.

2. The Polish Academy of Sciences has the custody of private papers with some material dating between the 16th and 18th centuries. The Wellcome Library holds a wide range of materials ranging 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 Czech Academy of Sciences keeps records of scientific institutions and personnel papers; the Vienna University of Technology also holds personnel records, too along with 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; numeric data and administrative records are also represented in the collections.

3. The University of Haifa is in possession of social sciences and humanities data. The University of Sydney has records from a wide range of disciplines and the University of Iceland keeps raw data on earthquakes and data from sociology research; along with 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 keeps no records in the long term and has only a temporary custody of the scientific material.

MANAGEMENT OF THE RECORDS OF SCIENCE IN SWEDEN

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 the Principle of Public Access to Official Documents. This Principle is valid for scientific material during the research process itself, 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 has consequently applied in Sweden, with minor exceptions, for over 200 years.
The Freedom of the Press Act defines also the term "document" and "official document".[1] According to this definition a document may be any representation in writing or picture, but also any record which can be read, listened to or comprehended only by means of technical intermediary, i.e. e-mail, data base or microfilm. A document is official if it is in the keeping, has been received or drawn up by an authority. This means that both documents forwarded from and submitted to an institution are considered official. All documents created and generated or received during the activity of an authority/university can thus be described as: *official* - , but of course, many document are *not official* but just "working" documents, i.e. drafts, suggestions for a decision, the first version of a report.

Of course, access to official documents may be restricted to protect specific interests. Restrictions are always an exception to the general rule; they are specified in the 1980 Secrecy Act. Records of science can be secret too, but only if there is a legal authority for classifying them as such. According to the Secrecy Act a good grounds for keeping scientific records secret, even the results or the methods used, may be the commercial value, "commissioned" research or on security grounds. At the same time an applicant always has the 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.

Usually we divide records of science into four main groups:

1. administrative records (e.g. project plans and descriptions, funding 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. draft reports, excerpts, calculations or electronic records as a part of data processing)

4. all material relating to the results (final and other reports, publications and articles).

Rules for archiving

As the records of science created during different projects or as part of continuing research activity, *no matter how the project or the activity has been established*, are subject to the Principle of Public Access, the rules for registration and archiving according to the Archival Law[2] have to be applied to the records of science. These rules for archiving include the rules for disposal and guidelines for the decision making process of appraisal. In this context it is important to note that in Sweden we apply records management to archives management.

---

1 The Freedom of the Press Act, SFS 1949:105, section 2 (changed 2002:1049)
2 SFS 1990:782

The National Archives of Sweden issued a decree in 1999 on the disposal of scientific records.[3] The decision was a result of a previous investigation from 1997 into the legal status of the scientific material at universities in the light of the Principle of Public Access. The disposal decision has been in force since January 2000 and provides terms of reference for the appraisal process.

Implementation of the rules at SLU

The structure and organization of records and archives management at SLU is decentralized. Each department (around 40) has its own responsibility for the registration of official documents and for archiving (keeping archives). Each department has thus to implement the disposal instructions for scientific material made by the National Archives. The decree makes clear that it is up to the University to put disposal rules into practice and implement them on the terms approved. At the same time it makes also clear that no destruction should be carried out before there has been evaluation of the records. In the evaluation process, which involves the critical keep/destroy decisions, due attention always must be paid to:

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

There are three main reasons for the preservation of records of science

- administration and legal needs, which includes verification 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 descriptions and the methods used, and the results must always be kept for ever and are always excluded from destruction. At the same time, it is important to add that under the Swedish Archival Law, each transfer 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 authenticity, is destruction. And change in media, i.e. transfer from a database to paper, always means destruction. Therefore it is very important to consider appraisal in general and conversion or media change in particular as early as possible, which means as far back as the creation stage as an first step in the evaluation process.

The rules of disposal of scientific material and rules on the status of records of science with regard to the Principle of Public Access to Official Documents are relatively new and require a fair amount of work to put into practice. Information meetings for heads of the departments and researchers have thus become a way of communicating issues about the management and preservation of these records. In addition, special courses for postgraduate students and holders of scholarships from abroad are now a well-established instrument for dissemination to the research community of information about the rules.

New problems and joint concerns about electronic records and the long term preservation of electronic scientific records have also required new projects on digital preservation. Digital preservation is the most important task after the appraisal process; it must be taken care of as soon as possible and must build on cooperation with records creators. The problem includes questions of appropriate funding and continuity, two factors which can never be guaranteed by a department, a unit subject to frequent changes in structure and organization. A separate system, independent of the department, is needed for the management of scientific records,. The department is, for other reasons, no longer a stable base for the long term preservation of primary data. Joint projects, where different units work together, and longitudinal field research projects attached to a particular department which has only the function of a "research hotel", and international cooperation, are more and more common. The organizational issues

[3] RA-FS 1999:1

and the responsibility for digital preservation of scientific electronic records were the subject of a faculty project carried out in January 2003.[4] The project steering group consisted of members of the faculty administration and of the IT-department, scientists and the university archivist.

The project group agreed that the issue of digital preservation cannot be separated from the question: who at the university has the responsibility for long term preservation? And as the university structure is subject to frequent organizational changes, and as the project itself is the records creator (and not a department), it should be the project leader who is 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 partners 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 clear 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 head of the department for evaluation and disposal according to archival rules. The head of department makes the decision on destruction and is responsible for the systematic transfer of records to the central university archives. Every decision on disposal must be well documented and reasons for destruction always given. The university archivist should always be contacted before the disposal is carried out.

A solution for the question of responsibility and the establishment of practices for the management of scientific records are only a part of the archival system. Another important issue is the technical solution for the long term preservation of electronic records. Creating all data in such a way that it is possible to read it, to understand and to transfer it to other media throughout the whole time it is preserved, must always be the starting point for long term preservation. Records format is a crucial problem in all future digital preservation planning. And - as the creation of digital information is very rapid and its 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 in each case with regard to the physical durability of the carrier and the logical durability of the information. The migration plan must also be a part of the university policy on digital preservation, and should include statements on security and secrecy, accessibility, authenticity and control of digital information.

As the preservation of electronic scientific records is a concern of all universities in Sweden, a joint project has been initiated. A project group consisting of archivists from four universities and librarians from two further other universities with experience on electronic publishing and metadata standards has been established. The main tasks for the group are: to consider standards, technical and organizational solutions and to set up a pilot project on digital preservation. The project was initiated in May 2004 at the Swedish University of Agricultural Sciences. Six different scientific databases 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. The results of the project and suggestions for future strategy are due for report in June 2005.

TOWARDS A MODEL OF BEST PRACTICE

Different archival rules and traditions affect the types of scientific records transferred to the archives, also on strategies for “ingest”. However, even though the types of records differ, the concern remains the same: how to resolve the long term preservation of digital scientific records and how to establish suitable guidelines for appraisal. There are therefore two crucial issues in establishing of best practices for scientific records: the evaluation process and the reliability of the electronic records of science.

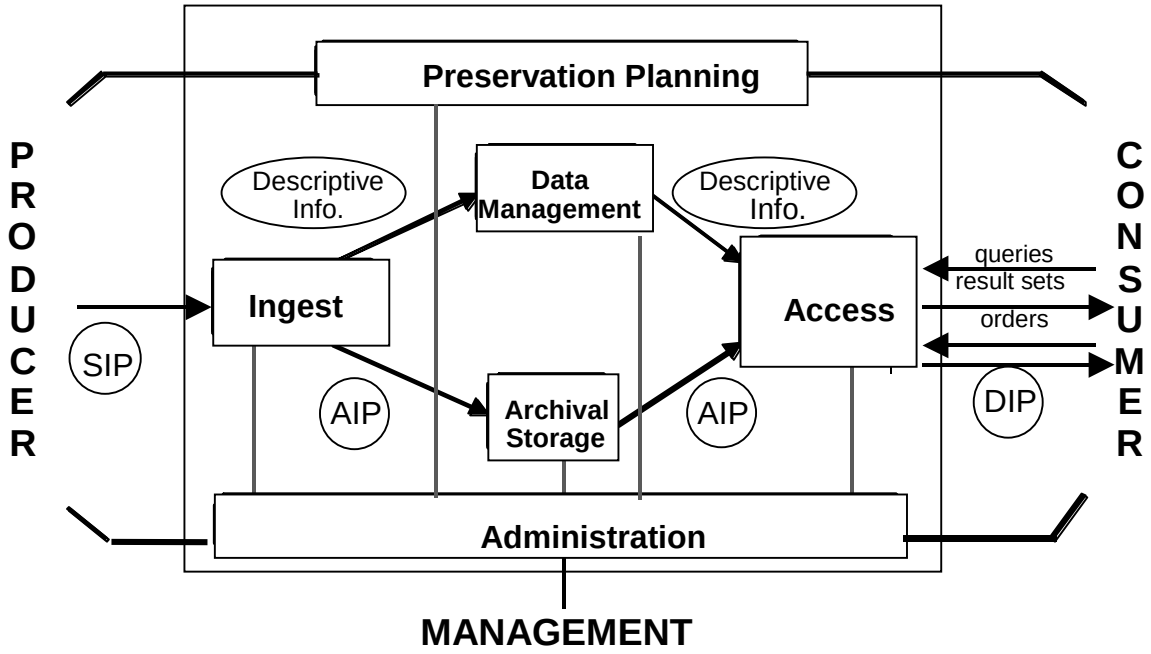
4 igital preservation of scientific records, Faculty of Forestry, SLU, 13 January 2003

Both those issues determine the possibilities of future access. As a result, a concerted approach is needed if the relevant scientific material is to be saved for the future.

The five steps of the scientific process were described above. The majority of survey respondents keep records reflecting this process. We also noted that, typically, they retain certain types of records: data about the aim of the project, project description and methods used, and the results document the research activity which should be kept for ever.[5] The appraisal process should thus focus on the evaluation of raw or primary data, and in particular cases, the analyzed data. Primary material makes possible the verification of results and is of central importance for other scientists and for the public interest. Many scientific projects are also dependant on re-use of raw data; their value for future research and for history must consequently be considered. Various disciplines will have different needs, but in general, those two criteria for evaluation have to be applied: verification of results and reuse for scientific and historical needs – even though it is often difficult to say what future users will need.

OAIS Reference Model

According to the Open Archival Information System (OAIS) Reference Model[6] an archival system, for electronic records, comprises ingest, storage, access and preservation planning. It includes hardware and software, and persons responsible for acquisition, preservation and dissemination. The concept and the main ideas of the model are illustrated as follows:



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

DIP = Dissemination Information Package

5 ifferent national and institutional rules may dictate archiving of the results. In many countries it is the library and not archive that is responsible for keeping of these records.

6 ee OAIS, Producer-Archive Interface Methodology Abstract Standard, CCSDS-661.0-R-1, Red Book, December 2002

The model[7] 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. Any discrepancy or specific regulations, however, must be well fixed in the metadata. In general, metadata is central to long term preservation. There are three types of metadata are necessary for archival purposes:

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

Information is always represented in some way. For digital information it is the way in which data are articulated that provides the understanding of it. Furthermore, for archiving digital information contextual information about the information has to be preserved, (i.e. preservation metadata): all those three types mentioned above. Those types may correspond, and they usually do, to archival description.

The OAIS Reference Model illustrates this as follows:

- AIP (Archival Information Package) = CI (Content Information) + PDI (Preservation Description Information).
- Content Information (CI) consists of Data Object and Representation Information.
- Preservation Description Information consists of four information elements: provenance, reference, fixity and context.[8]

Specific rules and conditions for managing records of science will then be shown in the Preservation Description Information. Consequently, it is of great importance to find a proper level for the description of digital scientific material that guarantees understanding in the future, but also understanding outside a particular research community. This description must then be accompanied by appropriate guidelines for appraisal.

The OAIS generic model with a suitable technical application and an appropriate level of description seems to be a good step in the effort towards the long term preservation of digital records of science.

Suggestion for best practice

It is not possible to find a “one size fits all” solution. Each scientific discipline has its own particular needs. However, it is possible to give some general suggestions for best practice:

- 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 verification 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 appropriate formats are specified
- incorporate specific rules for access and find a proper level for preservation description information

Renata Arovelius

---

7 bid, p. 4-57

8 Reference Model for an Open Archival Information System (OAIS), CCSDS 650.0-B-1, Blue Book, January 2002, p. 4-34