



**Information and Communication
Technologies for**

Biodiversity Conservation and Agriculture

Lisa Maurer,
Klaus Tochtermann
(Eds.)

**SHAKER
VERLAG**

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche
Nationalbibliografie; detailed bibliographic data are available in the Internet
at <http://dnb.d-nb.de>.

Copyright Shaker Verlag 2010

All rights reserved. No part of this publication may be reproduced, stored in a
retrieval system, or transmitted, in any form or by any means, electronic,
mechanical, photocopying, recording or otherwise, without the prior permission
of the publishers.

Printed in Austria.

ISBN 978-3-8322-8459-6

Shaker Verlag GmbH • P.O. BOX 101818 • D-52018 Aachen
Phone: 0049/2407/9596-0 • Telefax: 0049/2407/9596-9
Internet: www.shaker.de • e-mail: info@shaker.de

**Lisa Maurer,
Klaus Tochtermann
(Eds.)**

**Information and Communication Technologies for
Biodiversity Conservation and Agriculture**

**Shaker Verlag
Aachen 2010**

Chapter 10 THE EUROPEAN *EX SITU* PGR INFORMATION LANDSCAPE

Theo van Hintum

Centre for Genetic Resources, The Netherlands (CGN), Wageningen University and Research
Centre, Wageningen, The Netherlands
theo.vanhintum@wur.nl

Frank Begemann

Federal Agency for Agriculture and Food (BLE), Bonn, Germany
frank.begemann@ble.de

Lorenzo Maggioni

Bioversity International, Maccarese, Rome, Italy
l.maggioni@cgiar.org

***Abstract:** The situation regarding the documentation of Plant Genetic Resources (PGR) maintained in *ex situ* collections in Europe has changed considerably over the last decades. For this reason an attempt is made to describe the current landscape. The systems that are used to manage the information involved are described, as well as the mechanisms and systems that exist to exchange this information, and the developments and challenges in this area. Apart from this technical description, also a functional description of the changing roles of these elements in the landscape in the light of international, technical and legal developments is given.*

Keywords: plant genetic resources documentations, European genebanks

1. INTRODUCTION

In Europe there are over 625 genebanks that conserve the plant genetic resources (PGR) of our crop plants. These PGR usually consist of seed samples, dried and frozen, or small pieces of tissue maintained *in vitro* in little tubes under slow growth conditions or even in liquid nitrogen. It is estimated that currently in Europe, we maintain a total of approximately 2 million of such samples. This creates many challenges to these genebanks individually, but also collectively, to properly select the material to include in the collections, and to keep it alive and available to users.

One of the major challenges turns out to be the proper documentation of the samples to make these resources visible and accessible to users and potential users.

Over the last decades, much has changed in the landscape of European *ex situ* plant genetic resources documentation, just like in any other sector working with information technology applications. This paper presents an overview of the current status and tries to look forward. It was prepared as a discussion paper for the European PGR community as organised in the European Cooperative Programme for Plant Genetic Resources (ECPGR), and received feedback from several players. This feedback was incorporated in the paper and it was made accessible to a larger audience.

2. COMPONENTS OF THE LANDSCAPE

The European PGR documentation landscape is built up from a number of elements. In the first place, there are the documentation systems at the institutional level where the original data are captured, stored and curated. Many of these data are channelled to the next (national) level, i.e., the National Inventories (NIs), and there is an even higher level of agglomeration, the European level, including EURISCO, which is the European central repository of passport data, and the crop based European Central Crop DataBases (ECCDBs).

2.1 SYSTEMS AT INSTITUTIONAL LEVEL

The basic elements in the European *ex situ* PGR documentation landscape are the documentation systems used by genebanks and other institutes or actors conserving and using PGR (see for example [Menting et al. 2007] - available on line). Nearly all genebanks have computerized the information about their germplasm. The way this is done varies widely and depends on the size of the collection, the available expertise and the level of integration in the larger PGR community. Small collections allow for simple documentation systems: Excel spreadsheets can do the job; however once the collections grow, and database management issues, such as data integrity (use only codes that exist), data security (distinguish different users with different rights) or data processing (select material for regeneration) start to become more important, more specialized computer software – database management software, such as MySQL, Access or Oracle – need to be used.

Since the software only offers general functionalities to manage data, an application needs to be created that allows the actual storage of the specific information and the desired manipulations to create a true database management system (DBMS). This application requires, apart from other things, a data model that describes the elements of the information to be managed. In spite of the fact that the requirements on such a DBMS are similar, most genebanks in Europe developed local solutions and, as a result, the data models vary widely, although most are relational models. Some cover only passport data, many include genebank management information, others also include characterization and evaluation (C&E) data, data on the distribution and use of the germplasm, or other types of related information [Tab. 1]. Also the way these different domains are modelled differs between genebanks. Only the passport data models are usually rather similar; a flat table with a relatively standard set of fields describing identity, origin, etc. This structure is formalized in the Multi Crop Passport Descriptor (MCPD) format, in which data can be exported from any genebank documentation system (see [Hazekamp et al. 1997], also available on-line). C&E data are described in a wide variety of ways, ranging from an atomic model with one record per observation, to fixed tables with one column per trait and one record per accession. Also, the solutions for other data types are far from standard over genebanks.

Concerning the content of the DBMS, i.e. the data itself, hardly any standard coding systems, controlled vocabularies or ontologies are used in the PGR community. The only standard coding system commonly in use is the ISO coding system for country names (ISO-3166-1) using three letter codes. However, this list is usually extended with the codes for the countries which have been deleted since its first publication in 1974, such as 'SUN' for the 'Soviet Union', since these codes still occur in PGR documentations systems. Besides, FAO maintains a list of institute codes within the World Information and Early Warning System on Plant Genetic Resources (WIEWS) [FAO WIEWS 2009] as an attempt to offer a standard reference for institute identification. Neither standard taxonomic systems, nor ontologies for trait names, nor other standards are systematically used despite attempts by Bioversity International, UPOV and others to promote descriptor lists with traits per crop and other initiatives.

Currently, most European genebanks have a DBMS of some kind but, as indicated, these differ with regard to the database software used, their coverage of data domains, data models and coding systems used. Furthermore also the data quality, functionality and accessibility to outside users differ widely.

Data category	Interest for Use
Management data	Institute (internal)
Passport data	Institute, national, European and international levels, other users incl. breeders (external)
Characterization and evaluation data (C&E)	Breeders, researchers, other users incl. institute (external)
Data about distribution and use of germplasm	Institute (internal), national and international administration (pre-defined interest group)

table 10-1 Data categories and the level at which they are used

2.2 SYSTEMS AT NATIONAL LEVEL

If the information is properly collected and stored in a local DBMS, it is possible to extract it and combine it in other systems that do not so much have the objective to curate the information, but rather to analyze it and use it for other purposes such as administrative or coordination purposes. The most prominent systems at this level are the National Inventories (NI), compiling part of the information of the documentation systems of the respective collection holders and other germplasm providers. The NI not only documents PGR information at the national level, it also builds the interface to documentation systems at the regional and international levels such as EURISCO (see below). Data categories covered by the National Inventories are passport data and data about the distribution and use of germplasm. The type of material to include in the NI is at country discretion (in most cases it is a decision of the EURISCO National Focal Point (NFP): it ranges from accessions identified as available for exchange to the compilation of all available data in national institutions through the National Focal Points. Transparent terms of reference of what each NI is supposed to contain are currently not available in most cases.

Attempts to combine characterization and evaluation data (C&E data) from institutional databases at national levels have been made in a few countries but, due to the nature and complexity of these data and the very specific user needs, a generic approach to provide such data in a format really useful for breeders and other users

at national level have not (yet) been developed in a satisfactory way. Furthermore, researchers in the private and public sector would not necessarily need data according to country borders, but rather on a crop or trait basis which makes the development of such systems at the national level even more difficult to justify. However, the discussion about such a requirement seems useful.

2.3 SYSTEMS AT EUROPEAN LEVEL

2.3.1 EUROPEAN CENTRAL CROP DATABASES (ECCDBs)

Soon after the genebank documentation systems became computerized, scientists were tempted to combine the information of different systems and analyze the result to determine the coverage of the gene pool in the combined collections, but also to determine the redundancy between collections and try to coordinate activities of genebanks. This became apparent when some ECCDBs played an important role in the formulation and coordination of recent EU-funded projects.

Within the ECPGR crop working groups, creating a central crop database that combines the passport data of the collaborating institutes was a priority activity and a large number of such databases have consequently been set up. Creation and management of the ECCDBs was a voluntary 'input in kind' contribution of participating institutes or scientists (see for example [Stavelikova et al. 2002] or , [Germeier et al. 2003]). With the ongoing commercialization of crop science including privatization of institutes, this increasingly becomes an obstacle for their development and maintenance. In a recent review, we listed 62 ECCDBs in Europe covering most species maintained in European genebanks. Taken together, they comprise over 755,000 accessions. However, only 16 databases currently contain a limited number of C&E data. In conclusion, the ECCDBs vary widely with regard to their completeness, data quality, age of datasets, inclusion of C&E data, but also the possibility to search or download them via the web.

2.3.2 EURISCO

In an attempt to centralize all PGR passport information in Europe, irrespective of crop, a three year project called EPGRIS (European Plant Genetic Resources Information Infrastructure) that was partly funded by the EU was started in 2000. It identified, through the ECPGR Network, a National Focal Point (NFP) in each European country and supported the NFP, as far as possible, in creating or further developing a National Inventory as part of an aggregate system with passport data

from all germplasm collections that the country thought should be part of this system that was called EURISCO [EURISCO 2009]. On top of the National Inventories, a central database was created in which the information could be uploaded and thus combined in one large database with passport information of germplasm maintained in Europe under *ex situ* conditions: EURISCO. When the project finished in 2003, there was a database with nearly 900,000 accessions, by far the largest of its kind, compared to the US system [GRIN, 2009] and the system of the CGIAR [SGRP, 2008]. At that point the responsibility over the network of National Focal Points and the technical infrastructure of EURISCO was taken over by ECPGR, providing a permanent and transparent infrastructure of the PGR data flow in Europe. The current EURISCO, maintained by Bioversity International on behalf of ECPGR [EURISCO, 2009], has over 1.16 million accessions from 281 holding institutions in 38 participating countries. It has recently undergone a complete website re-design and update of the database technology, which has improved its user-friendliness and its use as a data-analysis tool. Active improvement of the quality of the data that it contains is ongoing and the frequency of updating by the NFPs should be increased.

3. CHANGES IN THE LANDSCAPE

With the recent developments with regard to the International Treaty on Plant Genetic Resources for Food and Agriculture (ITPGRFA) and the Convention on Biological Diversity (CBD) the need for new functionalities in the centralized systems has arisen. The European countries, almost all of them Contracting Parties to the ITPGRFA, will need to register the material they place in the Multilateral System (MLS), and some form of reporting of the germplasm transactions under the ITPGRFA will need to be organized. In a very cost-effective way with minimum additional changes, such registration could be integrated at all three levels, i.e., the institutional, national and European levels: the National Inventories receive such information from all institutes within the respective country and forward it via the agreed exchange format to EURISCO. This exchange format will need the addition of some fields to cover the additional information [figure 10-1]. Based upon this existing infrastructure, every National Inventory (and hence EURISCO) will be able to indicate which accessions can be considered part of the MLS. Adjustments of the EURISCO protocols need further development. For the second component, i.e. the reporting of germplasm transactions, a similar concept could be feasible; however, no solution has been developed yet as these reports still require more discussions with respect to their level of detail.

In the last five years, the parallel development in the European scene of ECCDBs and of EURISCO has, on some occasions and for many reasons, obscured their complementary role and has sometimes given the impression of the co-existence of two redundant or competitive data gathering systems. Both users and database managers have also, in some cases, complained about a confusion of roles of the existing tools. In the new PGR information landscape, all the elements to be developed through network collaboration need to be clearly identified for their unique and essential role. The expectations towards the ECCDBs are gradually shifting from passport data gathering and compilation, towards data analysis (allow crop specific searches/visualizations, identification of Most Appropriate Accessions, duplications and gaps, definition of core collections, management needs, etc.). More than on the database, the focus and the challenge is on the database manager, who can also catalyze crop groups activities, such as improving data quality, supporting activities related to the European Genebank Integrated System (AEGIS) [ECPGR 2009] and promoting the expansion of the database towards becoming a more complex crop portal. EURISCO should further strengthen its role as a 'one stop shop' window on accession level passport as well as C&E data of European *ex situ* collections.

The ITPGRFA foresees, as one of its supporting components in Article 17, a global information system to facilitate the exchange of information, based on existing information systems. The aforementioned information infrastructure with its institutional, national and European levels is certainly an essential contribution to this global system under the ITPGRFA. Furthermore, an initiative to create a global PGR accession level information system (ALIS) is being promoted by a tri-partnership funded proposal, including the Global Crop Diversity Trust. The intention is to have a common entry point to a distributed structure in order to enable users to identify and locate from world-wide available sources the germplasm most suitable for their needs. Existing regional catalogues, such as EURISCO and SINGER, as well as other regional systems, are expected to play their role in providing germplasm data to the global level. This can also be an opportunity for a concerted exercise to develop new standards and to increase their use at the global level.

MLS AND AEGIS REGISTRY STATUS DESCRIPTORS FOR EURISCO

34*. MLS Status

(MLSSTAT)

The coded status of an accession with regard to the Multilateral System (MLS) of the International Treaty on Plant Genetic Resources for Food and Agriculture. Provides the information, whether the accession is included in the MLS.

0 – not part of the MLS

1 – part of the MLS

If the MLS status is unknown, the field stays empty

35*. AEGIS Status

(AEGISSTAT)

The coded status of an accession with regard to the European Genebank Integrated System (AEGIS).

Provides the information, whether the accession is conserved for AEGIS.

0 – not part of AEGIS

1 – part of AEGIS

If the AEGIS status is unknown, the field stays empty

figure 10-1 Proposed additions to the current EURISCO exchange format to accommodate registration in the MLS and in AEGIS

* consecutive EURISCO descriptor number, pending on the decision to include these new descriptors.

4. TECHNICAL DEVELOPMENTS AND CHALLENGES

Information and Communication Technology (ICT) is the most prominent technology shaping and changing the world. What impact is it having and will it have on PGR documentation, conservation and use?

4.1 QUALITY OF INFORMATION

The size of computers has decreased and the access to the systems has increased. In the 70's and 80's the computer was located in one cooled room, and operated by a few technicians. Now, every desk has its computer and we all can use a wide range of software applications. Creating, managing and using a DBMS has become very simple as compared to 25 years ago. The main result of this development is that the experts able to interpret, complement and correct the data are much closer to the data, and able to make them more complete and more reliable.

Apart from the miniaturization, the ICT revolution has also brought the Internet that, apart from other things, increases the visibility of the data. Genebanks either have their own website providing access to the data or the data are available via the National Inventories, EURISCO and/or the ECCDBs. As a result of this increased visibility genebank curators will be more critical regarding the quality of the data and should try to make them more complete and more reliable. One of the areas to be completed is the geographic data.

4.2 TYPES OF INFORMATION

Currently, information exchange concentrates on passport information. This information is highly relevant for genebank management purposes, mainly for optimizing the composition of collections. Germplasm users however, are usually more interested in the traits of the germplasm. Obviously passport information can be, and is, used as a proxy for trait information: an accession from Siberia is likely to be cold tolerant, a modern malting barley variety from Japan is likely to have a high malting quality, but many germplasm users would like to be able to directly select for the resistant, high yielding, drought tolerant, early maturing, bright green, or sweetest accession. There is a range of technical problems involved. The information is rarely available in a computerized form in the genebanks and, if it is, the interpretation of the information is difficult since the phenotype is largely shaped by the environment; an early barley accession from Syria might be very late or not get to flowering in Norway. Another problem is the lack of standardization of the name of the trait and, more importantly, the methods of measurement in terms of scale (cm vs. 1-9), experimental design (number of replicates), treatment (inoculation, etc.), and many other factors that will influence or even determine the score and its reliability. The bottom line is that at the moment, the ability exists to bring together many sets of results from individual experiments. However, combining them in one phenotypic score per accession and trait is very difficult unless the traits are

environmentally independent. Since this exercise is so important, many people are working on it, but no simple solution can be offered, and it is therefore unlikely that national or European systems will go beyond offering 'raw' C&E data together with comprehensive documentation of experimental environments and used observation methodologies, i.e. the original data sets per experiment. The exchange format, however, could be much more standardized so that the time required to study the data should be spent on biology rather than informatics.

Other types of information, currently becoming available on a large scale, such as molecular marker data and, to a lesser extent information about qualitative trait loci (QTL) and genes, are relatively more simple to provide and should be made available to the user as soon as they become available to genebanks. This will require some innovation but, unlike C&E data, it is relatively straightforward.

4.3 SERVICES

Thanks to the online access to genebanks, it is now possible in a number of genebanks to better search databases, including the C&E data, look at pictures of the material and order the material on the spot. Material Transfer Agreement (MTA) requirements can be dealt with using click-wrap procedures, and the automated processing of the seed request allows minimal processing time. This is currently possible in only a few genebanks in Europe, although the required software is not that complex. In principle, any genebank could provide these services to its users, if not today, then certainly tomorrow.

4.4 VIRTUAL GENE BANK

As a result of the increased access to the data, the data become more widely distributed and accessible via several interfaces. The distance between data source, storing and curating the data and the user interface, giving access to the data, has increased and will continue to do so. Soon, (and in some genebanks this is already implemented), the DBMSs will provide direct access to their data to other computers on the Internet. These 'so-called' web-services will therefore allow creating interfaces such as the National Inventories and EURISCO by directly harvesting the data from their original locations, enabling the user to search many databases simultaneously. (For an introduction into web services see [Alonso et al 2004]). At the moment these services, as far as they have been implemented in genebanks, are still rather limited in their functionality, and only give the possibility to view a fixed set of passport data, but this can and will be developed further to give the PGR user access to all

information(s) he needs, and the possibility to order material without knowing (or caring) where it is maintained. This virtual genebank is very close, it is just a matter of implementation. The technology is available and already in use in some genebanks in Europe.

For example, the data of the Centre for Genetic Resources, The Netherlands (CGN) or the Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Germany are accessible via the institutional websites, via the National Inventories, via numerous ECCDBs, via EURISCO and via the Global Biodiversity Information Facility (GBIF). They can be searched on-line and downloaded from several locations in several formats, and they are provided as a web-service. Material can be ordered on-line and the MTA can be signed at CGN with a click wrap by any authorized person.

The function of a virtual genebank will also offer new opportunities for EURISCO when the European Genebank Integrated Systems (AEGIS), (which then will be *de facto* the virtual European genebank), is formally established. It should then be possible to search for material which is part of AEGIS, in other words, part of the 'European collection'.

4.5 CROP PORTALS

The breeding community will benefit from a further development of ECCDBs into user-oriented crop portals providing access to information much beyond the present C&E data, i.e. all data useful for research and breeding of a particular crop. Standards here seem even less likely to be agreed upon in the near future, but a generic approach should be sought to investigate such comprehensive user needs, based upon pilot crop portals to be developed. This will be an extremely useful exercise to facilitate not only access to the material but rather speed up research and breeding itself. Such crop portals may not necessarily stick to national or European borders and may rather draw from existing information sources worldwide. This implies that the creation of these portals should be coordinated with other players in the world, such as the International Agricultural Research Centres (IARCs) of the CGIAR, some of which are developing such portals already for their mandate crops. The development of such crop portals could be an extremely useful service offered by European specialists such as the ECCDB managers especially for crops which the IARCs do not focus on and where crops have their origin in Europe or European breeding has a deep interest.

4.6 RELATIONSHIP EURISCO – ECCDBS

The difficulty to develop and agree on a European-wide generic approach for C&E data is an important factor that has hindered the development of EURISCO towards being the central repository of PGR data in Europe. A solution is currently being discussed that involves a simple format for exchanging C&E data, linked to the passport data already in EURISCO, and guaranteeing the interpretability of the data, without any additional standardization. If this format is adopted and applied, EURISCO will cover passport, C&E data and data about distribution and use of germplasm, as well as serving as the European platform for registration and reporting for the ITPGRFA. Furthermore, EURISCO could develop interfaces at crop level to bridge over and directly link to the respective crop portals.

PGR data flow in Europe

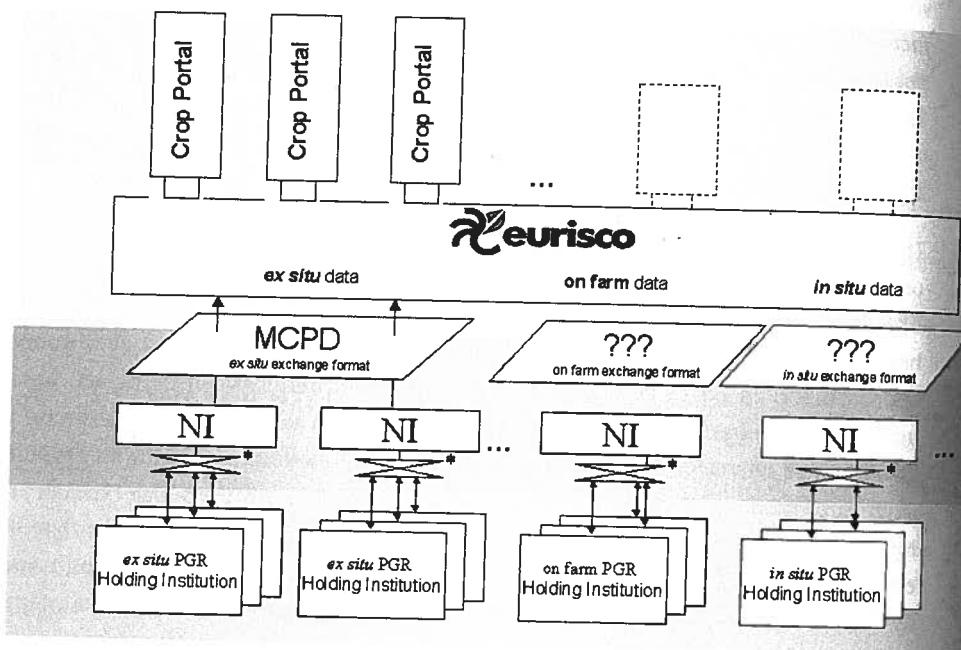


figure 10-2 Vision of the future PGR data flow in Europe

This would be a transparent and very user-friendly cooperative approach for the information needs under EURISCO, for the breeding research sector and other users' groups. ECCDB managers could take up this as a challenge and make use of their mixed expertise on ICT and crops in developing such user-oriented crop portals. The traditional role of ECCDBs focussing on C&E data storage would therefore convert into a much wider scope and include all information required by the breeders, including that on molecular markers, QTLs, genes or even relevant patents. Hence, the traditional database component covering passport and C&E data would be less important within such a crop portal approach and might even be given up in future for the benefit of a more user-oriented approach focussing on the needs of the breeding sector [figure 10-2].

It will be possible to fully pursue this scenario provided that EURISCO becomes the most reliable source of passport data. Currently, passport data from institutions in many countries may directly enter the ECCDBs and not EURISCO. This type of data flow generates discrepancies and uncertainty about the actual status and size of the available germplasm resource in Europe. There is a need to further strengthen the role of National Inventories and ensure that national systems provide all the relevant passport data to EURISCO through a clearing house mechanism.

To summarize, the new roles of EURISCO and the ECCDBs will ensure better quality data, a higher service level to users and finally a decoupling of maintenance and interfaces, not only with regard to information but also to material. This will bring us, the PGR community, a lot closer to the ideal where, together, we jointly conserve and make accessible PGR for future generations, especially for research and breeding.

5. THE NEXT STEPS

What do we need to do to make this development go more smoothly?

5.1 STANDARDS

To allow an easier exchange of information and the implementation of web-services, more and better standards are needed. This applies to all structures and terms used in PGR documentation; data models and ontologies for all PGR documentation domains. For example, it will be critical to ensure the wide adoption of Life Science Identifiers (LSID) to uniquely identify each information object [Page 2008]. This is not only relevant to Europe but to the entire world, and requires more than a few experts

solving the issue. It requires a process involving the PGR documentation community (similar to EPGRIS), where Europe could play a leading role, given the state of technology and the number of actors. However, there has been hardly any attempt to maintain the EPGRIS community or create a new genebank documentation community that could be used as a platform for identifying problems, developing standards or creating ownership of existing standards and solutions. Obviously, in this regard, there is a natural role to play for Bioversity International and ECPGR.

5.2 TECHNOLOGY

All the technology which is required exists. However, this technology needs to be made more accessible. The first and easy step is to proceed with prototype implementation examples in a few technologically advanced genebanks and share the lessons learned and the applied technology, allowing an easier implementation in other genebanks.

However, to make implementation more cost effective, it would be desirable to have a few model systems that can be copied and adjusted to the local requirements. Now every genebank or other germplasm provider has to invest deeply in developing a documentation system. It would be better if we could use each other's solutions and software. Open source technology supports this approach. Currently, an initiative of the Secretariat of the ITPGRFA and the Global Crop Diversity Trust aims at developing, in partnership with the United States Department of Agriculture (USDA) and Bioversity International, an open source genebank documentation system (GRIN-Global, mainly for use in developing countries) that could serve as a kernel around which to develop software that can be used by many, considerably reducing the development costs.

5.3 CAPACITY BUILDING

Despite the increased access to hardware and software, lack of technical and personal capacity remains a problem in optimally using the available technology and in applying existing solutions. The status of the documentation specialist in genebanks is often lower than that of the curators and, as a result, most PGR meetings in Europe do not involve documentation specialists, whereas the area of PGR documentation can be seen as the most important with regard to innovation and integration of genebank activities. There is a clear need for teaching and training materials, teaching workshops, staff exchanges and other capacity building activities if the European PGR community is to achieve maximum gain from the ongoing ICT

revolution. These activities should be as open as possible for all working in the field of genetic resources information and documentation, including National Focal Points and ECCDB managers.

5.4 COORDINATION

Some of the innovations and steps described above will be driven autonomously by regular genebank or national programs. Some of these will benefit from coordination and collaboration, others will not happen at all without coordination and additional funding. On a European level, there is the ECPGR Documentation and Information Network, with a very limited budget, and the collaboration platform EPGRIS3, a self-funded initiative [EPGRIS3 2009]. This will not be sufficient to optimize impact. On a global level, the Global Crop Diversity Trust, and CGIAR-centered programs, such as the Generation Challenge Programme (GCP) [Bruskiewich et al. 2008] and the Global Public Goods Programme (GPG2), are making major steps forwards. However, Europe could benefit more from these programmes if higher priorities were given, and more capacity could be made available to PGR documentation at all levels, institutional, national and European.

6. CONCLUDING REMARKS

The objective of the genebank community is to optimize the PGR conservation and use. The current European collaboration in ECPGR serves this purpose, allowing better coordination of the activities of, and collaboration between, individual genebanks. AEGIS is an exciting initiative that could bring this coordination and collaboration to the next level. However, as a saying in the computer world goes 'garbage in - garbage out'. The quality of the combination of elements will not be better than the quality of the individual elements. Therefore we need to search for possibilities to increase the quality of the elements, the building blocks of the European PGR community. This has many elements, and obviously the management of the PGR material in local genebanks is the first element to focus on. Defining quality standards, and setting up quality management and assessment systems, is the first priority and a huge challenge to the community. We hope to have shown by the above that improving PGR documentation, with all the listed aspects, is a good second in terms of potential impact on collaboration in Europe. Options exist. The technology is available. EURISCO is functioning, and EPGRIS shows that it is possible to collaborate. If all the players do their part, the next decennium will bring a drastic

innovation in the way we document our valuable resources and offer our services to the user.

7. ACKNOWLEDGEMENTS

The authors would like to acknowledge the valuable input of Eliseu Bettencourt, Sónia Dias, Lothar Freese, Christoph Germeier, Helmut Knüpffer, Markus Oppermann and Andreas Stephanik in the discussions leading to this document.

8. REFERENCES

- [Alonso et al 2004] Alonso G, Casati F, Kuno H, Machiraju V (2004). *Web Services: concepts, architectures and applications*. Springer, 354 p.
- [Bruskiewich et al. 2008] Bruskiewich R, Senger M, Davenport G, Ruiz M, Rouard M, Hazekamp T, Takeya M, Doi K, Satoh K, Costa M, Simon R, Balaji J, Akintunde A, Mauleon T, Wanchana S, Shah T, Anacleto M, Portugal A, Ulat VJ, Thongjuea S, Braak K, Ritter S, Dereeper A, Skofic M, Rojas E, Martins N, Pappas G, Alamban T, Almodiel R, Barboza LH, Detras J, Manansala K, Mendoza MJ, Morales J, Peralta B, Valerio R, Zhang Y, Gregorio S, Hermocilla J, Echavez M, Yap JM, Farmer A, Schiltz G, Lee J, Casstevens T, Jaiswal P, Meintjes A, Wilkinson M, Good B, Wagner J, Morris J, Marshall D, Collins A, Kikuchi S, Metz T, McLaren G, van Hintum T (2008). *The Generation Challenge Programme Platform: Semantic standards and workbench for crop science*. *International Journal of Plant Genomics*, vol. 2008, Article ID 369601, 6p.
- [ECPGR 2009] *A Strategic Framework for the Implementation of a European Genebank Integrated System (AEGIS). A Policy Guide*. European Cooperative Programme for Plant Genetic Resources (ECPGR). Bioversity International, Rome, Italy. 23p.
- [EPGRIS3 2009] *The European PGR Information Infrastructure phase 3 (EPGRIS3)*, <http://www.epgris3.eu/>, date of consultation 2009-06-05.
- [EURISCO 2009] *The EURISCO Catalogue*, <http://eurisco.ecpgr.org>, date of consultation 2009-06-05.
- [FAO WIEWS 2009] *World Information and Early Warning System (WIEWS) on Plant Genetic Resources for Food and Agriculture*, <http://apps3.fao.org/wiews>, date of consultation 2009-06-05.
- [Germeier et al. 2003] Germeier CU, Frese L, Bücken S (2003) *Concepts and data models for treatment of duplicate groups and sharing of responsibilities in genetic resources information systems*. *Genetic Resources and Crop Evolution* 50: 693-705.

[Hazekamp et al. 1997] Hazekamp T, Serwinski J, Alercia A (1997). Multicrop passport descriptors. In: Lipman E., Jongen M.W., van Hintum Th., Gass T. and Maggioni L. compilers. Central Crop Databases: Tools for Plant Genetic Resources Management. Pp 35-39 and 75-78. International Plant Genetic Resources Institute, Rome, Italy/CGN, Wageningen, The Netherlands.

[Menting et al. 2007] Menting FBJ, van Veller MGP, van Hintum TJL (2007). Genis data dictionary. CGN report / Centre for Genetic Resources (Ext. rep. 2007/08). Wageningen: Centre For Genetic Resources (GCN)

[Page 2008] Page RDM (2008). Biodiversity informatics: The challenge of linking data and the role of shared identifiers. Briefings in Bioinformatics 9: 345-354.

[Stavelikova et al. 2002] Stavelikova H, Boukema IW, van Hintum TJL (2002). The International *Lactuca* Database. Plant Generic Resources Newsletter 130: 16-19.