European Crop Wild Relative Diversity Assessment and Conservation Forum

PGR FORUM

www.pgrforum.org

Report

Workshop 3: *In situ* Data Management Methodologies

8.-10. September 2003, Prague, Czech Republic

Sabine Roscher, compiler

PGR Forum - EVK2-2001-00192

Fifth Framework Programme for Energy, Environment and Sustainable Development

Coordinated by: University of Birmingham, UK

0. Contents

1. Opening Session

Speakers:

Zdenek Stehno:	Welcome by Hosting Institute
Nigel Maxted:	Welcome by PGR-Forum Co-ordinator
Sabine Roscher:	Introduction to Workshop 3
Caroline Pollock:	Workpackage 2 update
José Iriondo:	Workpackage 4 update
Sónia Diaz:	Workpackage 5 update
Brigitte Laliberté:	Workpackage 6 update
Nigel Maxted:	Workpackage 6 update
Shelagh Kell:	Workpackage 1 Progress report

1.2 Introduction

1.2.1 The host

On behalf of the hosting institute Zdenek Stehno welcomed the workshop participants to Prague and the Research Institute for Crop Production.
The Czech Republic has a total land area of 7 887 000 ha. The agricultural land is 4 277 000 ha (arable land 3 075 000 ha, Hop-gardens 11 000 ha, Vineyards 16 000 ha , Permanent grassland, 966 000 ha, Non-agricultural land 3 610000 ha, and Forest land 2 639 000 ha) *15.8* % of the area of the Czech Republic are protected areas (1 244,4 000 ha).
Responsible for nature conservation is the Ministry of the Environment of the Czech Republic and the Agency for Nature Conservation and Landscape Protection of the Czech Republic, which is the technical organisation of the Ministry). One of the activities of the ‚National Programme on Plant Genetic Resources Conservation and Utilization‘ is Act No. 148/2003 on conservation and utilization of genetic resources of plants and micro-organisms important for food and agriculture. 4976 Crop wild relatives accessions are stored in the Czech gene bank (Wild *Triticeae* 1505, Aromatic and medicinal plants 771, Vegetables 745, Grasses 455, Fodder crops 436)

1.2.1 Project coordinating issues

The project coordinator Nigel Maxted welcomed the participants and reported on issues, that are closely related to the PGR-Forum (Marie Curie RTN, PGR IUCN SSC Group, PGR Society, Replacement for 1467 Grace FP6). That followed an explanation of the WP-6 deliverables:

- Project Web Page
- Project Database Designed
- Web-enabled of Project Database
- Project Database on CD
- Publication of Dissemination Conference Proceedings
- Project Newsletter 1-6

The coordinator also introduced a concept for the product exploitation and dissemination. It was suggested to produce a PGR Forum Publication Series by Cambridge University Press.
Plant Genetic Resources of Europe (WP1&2)

Management and Use of Plant Genetic Resources Data (WP3)
In Situ Conservation in Plant Genetic Reserves (WP4)
Species Loss, Genetic Erosion and Pollution of Plant Genetic Resources (WP5)
Dissemination Conference Proceedings (WP6)

1.2.3 Introduction to the workshop

Sabine Roscher introduced the objectives of the workshop:

- Define the data types that are required for *in situ* conservation of CWR
- Investigate data sources and issues of data quality and scale
- Establish data standards for in situ data management
- Investigate available data analysis tools and techniques
- Define a data(base) structure for management of in situ data
- Discuss issues of user needs, data security, IPR etc.

The results of workshop presentations and working group discussions will be compiled to produce the work package deliverables

- data(structure) for management of *in situ* data
- recommendations for data management
- recommendations for (spatial) data analysis tools, techniques
- data structure for management of *in situ* data

1.3 Progress report for WP 1

Shelagh Kell reported on the progress for WP1.
The objective of WP 1 is to create a European crop wild relative database, incorporating baseline biodiversity data with current conservation and threat status.
The deliverables are

- List of European crop wild relatives
- Agreed taxon conservation dataset
- European crop wild relative database
- Web enabled database available via project web site

The list of European CWR
To create the list of European CWR genus names from Euro+Med matching those from Mansfeld were selected and selected entries for those genera compiled.
The preliminary list contains 23,072 taxa, which represent the breadth of the project database.
Other parameters to consider are:

- Countries or country subunits
- Native / introduced / cultivated status
- E.g. 20,151 taxa considered “N” native; 4,757 taxa considered not native or status uncertain (ref. E+M incomplete draft)

The taxon conservation dataset

- Methodology for collation of trial conservation datasets was circulated to project participants
- CWR taxon data sheet modelled on CBSG CAMP TDS
- questionnaire responses received to date

- 57 taxa selected for in-depth data gathering
  Trial datasets collected for 21 taxa
- Working group discussions to continue post-WS3

The CWR Database
UoB has identified an individual who would be a potential candidate to carry out this work
Proposed strategy:
- A data entry module could be developed within the next 2 months
- A prototype would be developed and made available in Spring/early Summer 2004 (soon after WS4)
- The database would be tested and prepared for general release
- Version 1.0 would be available in time for WS5 (September 2004)
- The web-accessible version would be made available by June 2005

For more details on WP1 progress report see http://www.pgrforum.org/Documents/WS3%20Presentations/WS3%20WP%20Reports/WP1%20Progress%20Report.pdf

1.4 Workshop participants

Belgium,	Andre Toussaint
Colombia,	Andrew Jarvis
Czech Republic	Vojtech Holubec, Zdenek Stehno
Denmark	Kell Kristiansen
Fiji	Luigi Guarino
Finland	Juha Helenius
France	Dough Evans, Martine Mitteau
Germany	Frank Begemann, Helmut Knüpffer, Dirk Hinterlang, Olaf Nölle, Sabine Roscher
Greece	Stelios Samaras
Hungary	Attila Simon
Italy	Petra Engel
Italy	Brigitte Laliberte, Samy Gaiji, Thomas Metz
Latvia	Isaak Rashal
Lithuania	Juozas Labokas
Norway	Aasmund Asdal
Poland	Wieslaw Podyma
Portugal	Eliseu Bettencourt, Sónia Ricardo Dias
Romania	Silvia Strajeru
Russian Federation	Tamara Smekalova
Slovak Republic	Daniela, Benedikova
Spain	José M. Iriondo, David Draper, Mauricio Parra, Lori De Hond
Sweden	Dag Terje Endresen
The Netherlands	Theo J. L. Hintum van
United Kingdom	Shelagh Kell, Stephen Jury, Nigel Maxted, Maria Scholten, Caroline Pollock

2. Session A: Data sets, data types and data standards

Speakers:

Maria Scholten:	Data on CWR in the UK, a case study
Nigel Maxted:	In situ Models and Data Management
Sabine Roscher:	WG 1: Taxon based data, distribution data”
José Iriondo:	WG 2: Ecogeographic data
Doug Evans:	How can existing habitat conservation and classification projects help to conserve PGR?

The session was opened with a case study, demonstrating data acquisition and data analysis issues.

2.1 Data on CWR in the UK, a case study

Speaker: Maria Scholten

Data for Conservation Gap Analysis
To assess conservation *in situ* and *ex situ* assessment are needed. Only with both a Conservation Gap Analysis can be carried out.
Data types, availability and access

- Data required for *in situ* assessment:
  - Occurrence
  - Abundance
  - Trends-in-time
- Data types available in the U.K. :
  - Floristic surveys (Free – contract-based)
  - Vegetation and habitat surveys (cost based)
  - Studies on individual species (?)
  - Data from conservation areas (To be negotiated)

Data quality issues associated with floristic data
Data quality: Taxonomical issues:
Several taxonomic issues influence the quality of floristic data. The taxon unit can be different, e.g. aggregates, s.l. or s.s., subspecies, microspecies, hybrids. Also the different status (alien / native) has to be taken into account and last but not least the changes in taxonomical backbone, i.e. flora’s, causes a lot of difficulties for data handling and data quality.

- Example: Influence of Taxon units on data quality
- Influence of status (alien / native)
  by the example of *Allium schoenoprasum*
  native 18 [10x10 km squares]
  introduced 104 [10x10 km squares]

Data quality: Recording or data collecting

- Identification problems
- Scale: 10 x 10 km unit / 5 x 5 km unit / 1 x 1 km unit
- Recorder bias or variability

- Taxonomical and identification issues, e.g.:

Malus sylvestris s. l.
alien *M. domestica* (Apple)
- native *M. sylvestris sens. str.* (Crab Apple)

Mapped together as native because of:
- Widespread hybridisation
- Identification problems

- Sources of recording bias in floristic surveys
    - Unequal botanical expertise
    - Lack of standardised search methods
    - Uneven time spent recording
    - Uneven coverage of grid and biotopes
    - Unknowns:
      Complete coverage? Or Representative listing?

A case study of 66 PGR:
Approach
- List of 66 wild socio-economically important PGR
- (Secondary) distribution data-analysis
- Categorise species into frequency classes
- Select less common categories
- Analysis of occurrence at different geographical levels
- Analysis of long term trends
- Collect conservation data
- Gap-analysis

Categories of occurrence
Red Data Book categories:
- Rare: less than 16 10 x 10 km squares
- Scarce: between 16 – 100 10 x10 km squares
- Less common 1: between 100 10 x 10 km squares and 10% of all squares
- Less common 2: between 11 and 20% of all squares
- Common: between 21 and 50% of all squares
- Very common: more than 50% of all squares

Conclusions:
- Local floras suggest an overall lower occurrence at vice-county level than the national occurrence of 11%
- Looking at the ratio between 10 by 10 km squares and 1 by 1 km squares, a rather high local frequency is suggested in Dorset, Cumbria and W.Gloucestershire

Elements of the Biodiversity Action Plan for Wild Asparagus
- The elements of Biodiversity Action Plan for Wild Asparagus are
    - Current status
    - Current factors causing loss or decline
    - Current action

- Action Plan objectives and targets:
    - Maintain and enhance population
    - Create a reserve population
    - Establish new populations

- Establish an ex situ program

- Proposed Actions:
  - Policy and legislature
  - Sites safeguard and management
  - Species management and protection

Conclusions (Case study of 66 wild PGR in Great Britain)

- 12 out of 66 socio-economically important species are scarce to rare
- Among these 12, four are not covered by a conservation status
- 1 species has a Biodiversity Action Plan
- Most current *ex situ* conservation does not cover the genetic diversity of wild genetic resources
- Current *in situ* conservation does not cover the less common PGR
- Introduced PGR tend to fall out of existing conservation frameworks

2.2 Introduction *in situ* plant conservation and data management of European CWR

Speaker: Nigel Maxted

Nigel Maxted introduced to the workshop with a presentation on "In Situ Plant Conservation and Data Management of European Crop Wild Relatives"
According to Davis et. al.,(1995) the European Flora includes 12,500 native species. All are threatened by varying degrees of taxonomic and genetic erosion. Due to the huge number of a need to prioritise conservation actions is given. Prioritisation is associated with 'value' and there are many reasons for prioritising. The socio-economic valuation is part of the traditional PGR activities.
Genetic conservation is explicitly utilitarian (q.v. CBD, ITPGRFA, GSPC, EPCS) and there is an intimate link between conservation and use. Plant Genetic Diversity -> Conservation ->Utilisation
According to "Genetic material of plants which is of value as a resource for the present and future generations of people" (IPGRI, 1993)" are modern cultivars, breeding lines and genetic stocks, obsolete cultivars, primitive forms of cultivated plants and land races, weedy races, crop wild relatives (CWR) , and other wild species. The most urgent need for action is for land races and crop wild relatives.
A Crop wild relative is a taxon related to a species of direct socio-economic importance. Europe an important centre for crop wild relative diversity. Major crops: oats *(Avena sativa*), sugar beet (*Beta vulgaris*), apple (*Malus domestica*), annual meadow grass (*Festuca pratensis*), and white clover (*Trifolium repens*) have wild relatives in Europe. Minor crops are arnica *(Arnica montana*), asparagus *(Asparagus officinalis*), lettuce (*Lactuca sativa*), and sage (*Salvia officinalis*).
According to Hawkes et al. (2000) "Efficient and effective data management is critical for biodiversity conservation, but few recognise or acknowledge that efficient biodiversity conservation is so intimately linked with effective data capture, storage and retrieval. Although data management is integral to any conservation initiative, it is not always given appropriate credence … ". Painting et al. (1993) pointed out that "Data management, or documentation is required for all aspects of plant conservation and use, including taxonomy, inventory and monitoring, genetic diversity studies, population and pollination biology studies, in situ management practices, germplasm collection and storage, cultivation, trade and legislation, mapping, and literature reviews"

2.2 Working group 1: „Taxon based Data, Geographic Data and Information Networking

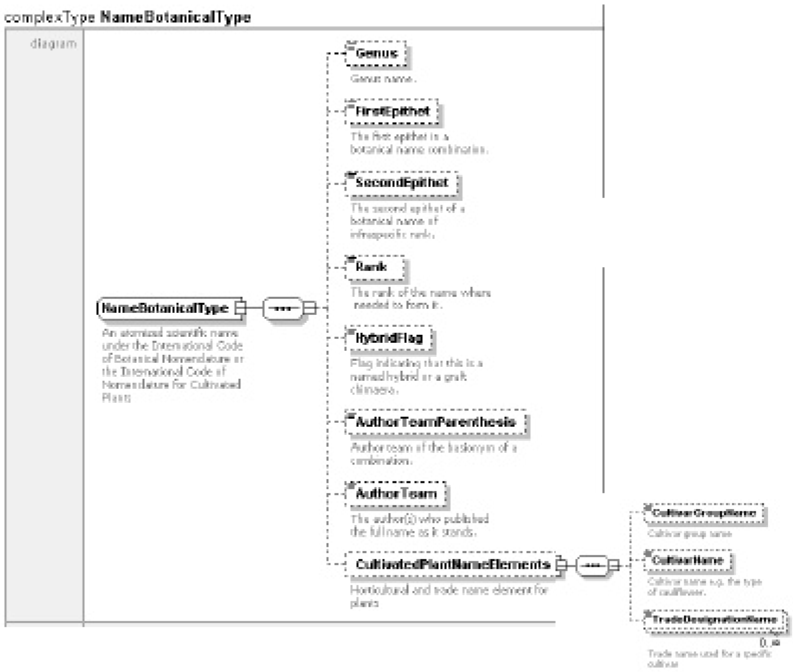
Speaker: (Sabine Roscher)

Working group 1 prepared proposals for
- standard on nomenclature
- standards and infrastructure for geographic information and geographical information systems (GI & GIS)
- standard and infrastructure for Phenology
- and investigated other standards for data types like life form, etc.

Results

a) Taxon based data

For the exchange of taxon based data the standards (xml-schema) of the TDWG and CODATA working group was introduced. These standards (ABCD-schema, Darwin Core) are accepted within the international user community - especially within the GBIF-[1]network - to share and access biodiversity data,.



b) Geographic data

For the in situ data management different „types“ of spatial data are needed:
- Distribution species (CWR)
- Locations populations (CWR)
- other „GIS-layers“ (abiotic and biotic factors) for analysis of spatial patterns

Those data are needed for different geographic dimensions or “scales”:
- regional level
- national level

[1] Global Biodiversity Information Facility (www.gbif.org)

- subnational level | provincial
- local level

A proposal for data types (Nomenclature, Geography, Population and other data types resulting form WS 1) was prepared by Sabine Roscher, José Iriondo and Shelagh Kell.

c) Other European project with relevance to PGR-Forum

As a basis for the use, adoption and creation of data standards *inter alia* following European project were investigated

- Project GI & GIS - Harmonisation and Interoperability

  - is carried out by the Joint Research Centre of the EC
  - aims at creation of a European GI policy and GI infrastructure
  - activities are *inter alia* INSPIRE, NATURA2000, CORINE land use / land cover

- Project Nature-GIS http://www.gisig.it/nature-gis/

  Objective
  - focal point to identify specific GI & GIS requirements for "Nature Conservation & Biodiversity" in the European Policies

  Activities
  - to identify user needs, to characterise the stakeholders in the field, survey of the scenarios for use of GI in this domain
  - data requirements, to individuate and specify the common kernel of GI content to describe protected areas
  - functional requirements, to identify the functional requirements for the management of GI and to define the framework for spatial databases for protected areas.

  Results
  - to define and realise web access to information on European protected areas.
  - to produce technical guidelines for data infrastructures for protected areas,
  - in order to contribute to geo-data access and exchange through standardisation of data infrastructures for protected areas.

- Project: European Phenology Network

  Objectives:
  - To facilitate integration and co-operation between existing phenological monitoring networks and to actively stimulate expansion of existing and creation of new monitoring networks
  - To improve the integration of, and access to phenological data in Europe in a systematic, structural and user-friendly way
  - To exchange knowledge between phenologists of different scientific disciplines (ecology, agriculture, human health) on tools and techniques used for phenological
  - monitoring, database development, (statistical) data analysis, model development, and impact assessment (background).
  - To demonstrate the wide variety of possible applications of phenological research and its benefits for ecology, agriculture and society (human health + education) ....

  Results:
  - As standard for phenology data the BBHC was identified.

a) Relevant standards identified:

- The United States (US) Federal Geographic Data Committee (FGDC)
- Content Standards for Digital Geospatial Metadata (CSDGM) (http://www.fgdc.gov)
- The Dublin Core Metadata Initiative (DCMI) (http://www.dublincore.org)
- The International Organization for Standardization (ISO) 19115 GI – Metadata (http://www.isotc211.org)
- For life form, habit and breeding system no standard is available jet. Solutions are under development, inter alia the 'dictionary' of botanical terms (Edinburgh Botanic Gardens).

b) Requirements for data access:

- Meta data (Definition: "information and documentation, which makes data understandable and shareable for users over time" (INSPIRE, 2002)
- Common Data Definitions
- Providing standard terms for data that is collected
- e.g. the Interpretation Manual of European Union Habitats
- Mapping of terms (concepts)
- e.g. match biotope type classification used on national level
- to other habitat classification like EUNIS
- Definition of data structure
  e.g. protected area is of type nature park, has a name, has a code, is represented by polygon(coordinates), has contact person, ...
- Definition of use cases
- Organisation of data flow

Results

- Structure for nomenclature data was accepted, the extension for CWR concerns like the Gene pool-concept has to be investigated
- Proposed meta data for structured description of distribution data was accepted
- Complete list of sources f. distribution data of CWR has to be compiled
- Data structure for detailed site / location and abiotic / biotic factors (GIS-layers) depends on the structure that is available at the local level. For the exchange of those data the so called footprint (e.g. coordinate or polygon) and the use of WebfeatureServices for GIS-layers is necessary.
- The use of site-specific fields of the crop descriptors should be promoted. By means of GIS those data can increasingly be derived from digital elevation models, soil informations systems etc.

2.4 Working Group 2: Ecogeographic data

Speaker: José Iriondo

As preparation for the discussion on data types following general information categories were proposed

- Taxonomy
- Species biology
- Geography

- Ecology
- Demography
- Threat assessment
- Conservation
- Characterisation
- Evaluation
- Utilisation

For the data type selection procedure it was proposed to list and structure data types in major groups (species and population).
Criteria for the decision process are

assess usefulness
- cost of gathering and processing information
- objectivity (measurable, clear definitions)

2.5 Habitat classification and the Common Database on Designated Areas (CDDA)

Speaker: Doug Evans

As the conservation of European Crop Wild Relatives is by nature closely related to the traditional nature conservation activities in Europe it is important to coordinate the efforts, especially in the field of documentation and data flow. Therefore one of the most important stakeholders on the European level is the European Environment Agency (EEA), that was created by EU Regulation in 1990. "The EEA aims to support sustainable development and to help achieve significant and measurable improvement in Europe's environment through the provision of timely, targeted, relevant and reliable information to policy-making agents and the public."

Occurrence of CWR and protected areas
Many crop wild relative species occur in protected areas - deliberately or accidentally -, but maybe they are not recognised by site managers.

- The European Environmental Agency runs a "Common Database on Designated Areas (CDDA)". This is a joint project between the Council of Europe, UNEP-WCMC and the EEA. National designations for protected areas are 44 695 sites in 48 countries. The size ranges from 0.005 ha to 97 200 000 ha and there are some 600 different types of sites.
- CDDA includes international designations for protected areas, e.g. Ramsar, UNESCO, Council of Europe (1 625 sites in 48 countries) and EU designations for protected areas (Natura 2000). Information about NATURA 2000sites at present is limited to site name, coordinate, area, site type and year of designation. One problem is that many sites have more than one designation.

The planned improvements for CDDA are to include the site boundaries, the main habitat types and links to more detailed national databases.

- CDDA has three main data categories, which constitute the main entry points for searching the database:
  - Species: Information about species and subspecies in Europe
  - Habitat types: Information about EUNIS habitat types classification and Habitat ANNEX I directive classification

- Sites: Information collected from various databases regarding sites

- Search results are presented as factsheets, *inter alia* with following items:
  - Scientific name, Author
  - Taxonomic Information (Taxonomic hierarchy, reference, source of information)
  - Synonyms (Scientific name, author) and Vernacular names,
  - Other Information like Geographic distribution, Populations, Trends referenced)
  - Habitat types populated by species (EUNIS Codes, ANNEX I Code, Habitat types name, Region, Abundance, Frequencies, Faithfulness, Species status)
  - Map of related sites for this species (Site code, Source data set, Country, Site name with coordinate, Altitude, Surface area, ...)
  - Ecological information about Fauna and Flora of the site
  - Ecological information of Habitats within the site

NATURA 2000 network

The NATURA 2000network is the most important network of nature conservation sites in Europe. The network comprises two categories of areas, the Specially protected area (SPA, Birds Directive) and the Special Areas of Conservation (SAC, Habitats Directive). The network has areas for 197 habitats, which are likely to increase to c. 225, and 483 plant species, which are likely to increase to c.575. The number of proposed Sites of Community Importance is 15500, which will be c. 13 % of the EU15 area.

Many CWR are included in Annex II or form important components of Annex I habitats e.g. *Quercus suber* forests (Annex I), *Brassica glabrescens*, *B. insularis*, *B. macrocarpa* (Annex II)
For NATURA 2000 sites a lot of ecological and socio-economic information is available and a requirement for monitoring and reporting on a regular cycle as well as the establishment of management plans is given.

Habitat classifications

Many different systems are in use at a variety of scales and developed for diverse reasons. The lack of a system usable across the EU led to the CORINE biotopes project, the Palaearctic classification and EUNIS.

- The EUNIS habitat classification is criteria based and hierarchic. Including marine and terrestrial habitats, natural as well as artificial habitats this classification is comprehensive. Links to other widely used classifications are also considered.
- Phytosociological classification is probably the most widely used approach. The problems with this is the disagreement between different groups and “Philosophical reservations” in some countries

- Important initiatives are:
- European Vegetation Survey, that provides a list of alliances and higher units (published April 2002)
- SynBiosis is a „information system for the evaluation and management of biodiversity among plant species, vegetation types and landscapes“. The aim is to bring together many years worth of vegetation data, e.g. 350 000 relevées in the Netherlands, c. 137 000 in France, c. 54 000 in the Czech Republic, estimated 750 000 relevées in Europe already in databases.
- ECOLAND (Pan European Forum for Countryside and Landscape Monitoring)
- The overall objective of ECOLAND is to provide an integrated assessment of change in habitats and biodiversity and the associated causes and impacts on the European landscape. A project called BioHab is underway.

Conclusion:
Many crop wild relative species occur in protected areas - deliberately or accidentally -, but maybe they are not recognised by site managers. *Inter alia* the list of Crop Wild Relatives will help to raise awareness for those species.

The EUNIS habitat classification is an important data type for the assessment of Crop Wild Relatives on the European level. A linkage between the EUNIS habitat classification and classifications that are used on the national and subnational level is needed to make the data comparable .

2.6 D09 Appropriate in situ genetic conservation data types identified

The definition of appropriate in situ conservation data types was based on the scientific requirements on the one hand but also taking into account data availability on the local and national level. By the example of Germany national mapping programmes (biotope surveys, national flora) were analysed and the mapping units evaluated with respect to their relevance for the conservation of Crop Wild Relatives. Based on the results of Workshop 1, a draft list of data types was discussed and presented during the workshop. After working group discussions and as follow-up of the workshop the list of data types was adopted. As a result the following list of data types was identified:

TAXON LEVEL CHARACTERISTICS	
Taxon nomenclature	TDWG Standard
ScientificName	
or ScientificNameAtomized	
Genus	
FirstEpithet	
SecondEpithet	
Rank	
HybridFlag	
AuthorTeamParenthesis	
AuthorTeam	
CultivatedPlantNameElements	
Principal Synonyms	
BotanicalType	
Vernacular name	
Taxon biological data	
Reproductive system	Sexual, vegetative
Breeding system	Allogamous, autogamous, auto-allogamous
Flower /plant sex structure	Hermaphrodite, monoecious, dioecious,...
Pollination	Wind pollinated, insect pollinated,...
Life form	Raunkiaer
Life span	
Habit	Herb, shrub, tree, climber
Taxon ecogeographical data	
Geographical distribution	Countries, country subunits
Approximate extent of occurrence	IUCN Definition
Area of occupation	IUCN Definition
altitudinal zone	Planar, colline, montane, alpine, nivale...

altitudinal range	(e.g. 350-500 m a.s.l.)
Soil description	e.g. shallow sandy acidic soils
Soil type	Reference to classification that is used
Soil texture	Reference to classification that is used
pH	Range
Climatic preference	Arctic, boreal, nemoral, meridional, tropic, austr. zone, anart
Climatic preference	more detailed subunits
Habitat	EUNIS category
Habitat status	Increase, decrease in area or quality
Vegetation	
Phenology	
Status	Native/introduced, archeophyte, neophyte...
Taxon population level data	
No. of locations or populations in which the taxon is distributed	
Overall population number	IUCN
Population trends	IUCN
Biotic interactions	Pollinators, parasites, diseases...
Taxon utilisation	
Uses/Ethnobotany	
Taxon threats	
Threats	IUCN authority file
Red List assessment	IUCN
Taxon conservation actions	
Legislation	
In situ: Protected area, *in situ* management plan, reintroduction, translocation…..	
Ex situ: Accesssions, time of collection, place of storage….	
Recovery plans	
Taxon documentation	
Herbarium	
Photo	
Illustration	
SPECIFIC SITE LEVEL DESCRIPTION	
Site location	
Country	CountryName
	Code (ISO2letter, ISO3letter)
AdministrativeUnit	Name
	Code
NamedArea	NamedAreaClass (Protected Area)
	NamedAreaCode (NSG345)
	AreaCodeStandard (National Coding System for Nature Res
	AreaCode
	DataSource
NearestNamedPlace	PlaceType (e.g. mountain)
	Name
	Relation to
CoordinatesLongLatDecimal	LongitudeDecimal
	LatitudeDecimal
	Spatial Datum
	AccuracyStatement
	CoordinateErrorDistanceInMeters

CoordinatesLongLatDMS	LongitudeDegree
	LongitudeMinutes
	LongitudeSeconds
	Direction
	LatitudeDegree
	LatitudeMinutes
	LatutudeSeconds
	Direction
	AccuracyStatement
	CoordinateErrorDistanceInMeters
CoordinatesUTM	UTMZone
	UTMEastWest
	UTMNorthSouth
	UTMMethod
CoordinatesGRID	GridCellSystem
	GridCellCode
	GridQualifier
URL to Map	(WebMappingService, WebFeatureService)
WMS	
Altitude (a.s.l.)	Upper - lower level
Site geomorphology	
Position in Relief /Geomorphology	
Aspect	N,S,E,W,NE,SE,SW,NW,open,closed
Slope	Measurement of the angle of slope
Site microclimate	
	Micro-climatic preference (max., min., mean temperatures, r
	(solar radiation, intensity of wind, distribution of precipitation
	Data source for micro-climatic reference and period of years
Site geology and soil	
Geology with effect on soil	
Geology / Statigraphy	
Soil type	Reference to classification that is used
Soil texture	Reference to classification that is used
Soil moisture regime	
Soil depth	
pH	
% organic matter, elements	
Salinity	
Derived soil evaluation	
Site habitat	
Habitat type	e.g. EUNIS
Vegetation	Association
Vegetation stratification	
Inventory of accompanying plant species	
Land use type(s)	
Land use intensity	
Anthropogenic effects	Hemeroby
Owner	
Site detail	Description
Location of population with regard to species distribution	e.g. site is at the border / center of areal of species
SPECIFIC POPULATION LEVEL DESCRIPTION	
Population size	

Boundaries	Polygon or Grid cell
Diameter	In the sense of Braun-Blanquet
Approximate area of occupancy	Measurement of area depends on scale
Population size (=Abundance)	
No. of mature individuals	
Population structure	
No. of subpopulations or microhabitats	
Spatial pattern of individuals	Uniform, aggregated, random (e.g. as in Morisita index)
Population density	
Isolation / Fragmentation	
Location of individuals and mapping	
Reproductive/ vegetative ratio	
Sex ratio for dioecious species	
Population dynamics	
Trends	As used in IUCN categories - estimates of decline / Fluctuat
Survey first or sequence, date	
Population management	
Site management plan	Detail of plan
Management interventions	Actual timing, frequency, duration of mowing, grazing, selec
Monitoring detail	Taxa monitored, location of quadrats and/or transects, frequ density, cover, frequency)
Biotic interactions	
Dominant vegetation	
Associatied vegetation	
Keystone species	
Dominant herbivores	
Asscoiated herbivores	
Grazing pressure	
Pollinators	
Seed dispersers	
Percentage tree cover	
Type of tree cover	
Ethnography	
Local ethnic group	
Language	
Traditional use of site	
Characterisation & Evaluation	
Preliminary characterisation	IPGRI DescriptorLists ?
Local threat (to specific population)	
Threat category	IUCN criteria also for that level? Red List assessment
Threat reason	Actual, IUCN/SSC Major threats authority file
Conservation measures (applied to specific population)	
Legislation	Level of site protection
In situ	Protected area, management plan, reintroductions, transloc
Ex situ	Detail of *ex situ* duplication, gene bank, number of accessio

3. Session B: *In situ* data management

Pesentations by

Wieslaw Podyma:Data management on the local scale
Vojtech Holubec: Quality standards for data capture
Dirk Hinterlang: Experiences from Biotope Survey – the Relevance for Mapping and Data Management of CWR
Olaf Noelle: Data Interoperability and GeoDataInfrastructure
Samy Gaiji: Legal and intellectual property issues to the practical implementation of an information infrastructure

3.1 Data structure for description of distribution data

D10 a) Appropriate data base structure identified

As a solution to the heterogeneity of data sources on national and local level, it was decided that first of all metadata for a structured description of databases are needed. The so called discovery metadata is a minimum amount of information that needs to be provided to explain the nature and content of the data source. The data comprehend items to describe

- What – title and description of the data set
- Why – abstract detailing reasons for the data collection and its uses
- When – when the data set was created and the update cycles if any
- Who – originator, data supplier, and possible intended audience
- Where – the geographical extent based on latitude, longitude, co-ordinates, geographical names or administrative areas
- How – how it was built and how to access the data

ID-Metadata	Unique Identifier for the metadata set
Character set	Code to describe the character set of the dataset
Title	Title resp. name by which the cited resource is known
Alternate Title	e.g. a short name by which the cited resource is known
Responsible person name	The full name incl. title of the responsible person
Organisation name	The official name of the responsible organisation
Organisation role	Owner etc.
Organisation address	Address of the organisation, split into Street, city, postcode, country
Organisation link	on-line information for contact the organisation
Organisation contact	Split into phone, fax, email
Bounding Box	Split into four coordinates that define the bounding box, coordinates in decimal degrees.
Geographic Identifier	Name for the area (e.g. name of a province or landscape unit)
Reference System	Name of the reference system
Spatial representation type	e.g. vector
Dataset reference date	
Dataset reference type	e.g. creation
Time period begin	date and time for the content of the dataset
Time period end	date and time for the content of the dataset

Spatial resolution	
Dataset language	
Abstract	
Keyword	
Use Constrains	
Lineage of the dataset	general explanation of the data producer's knowledge about the lineage of a dataset
Metadata language	
Metadata timestamp	

The appropriate concrete data items and structure for a meta-database were compiled, based on the standard ISO 19115 Geographic Information / Geomatics Metadata of the International Organisation for Standardisation http://www.isotc211.org and the Dublin Core Metadata Initiative (DCMI). This approach is also compatible to the overall strategy of the initiative „Infrastructure for Spatial Information in Europe – INSPIRE" (http://www.ec-gis.org/inspire).

3.2 Data management on the local level

Speaker 1: Wieslaw Podyma
Speaker 2: Vojtech Holubec

Based on the two presentations of Wieslaw Podyma and Vojtech Holubec and the related discussions about in situ data management on the local level a basic consideration for data management was agreed: Data should be collected once and maintained at the level where this can be done most effectively. E.g. site and population specific data about management measures and an instant flow of monitoring data are best maintained at the local level, where knowledge about the site and the taxa is available.
Vojtech Holubec reported on which data types are already available resp. which is the basic set of descriptors for wild collected PGR:

- Identification data
- Taxonomic data
- Geographic data
- Ecological data (topography, slope, orientation, bedrock, drainage, soil type and texture, habitat/vegetation, disturbance)
- Population data
- Sampling method

In addition to the data that is are sampled for collected material monitoring data to document change e.g. in protected area are of high interest, *inter alia*

- Phytocenological mapping
- Card for locality
- Observation of threatened species populations
- Line transects
- Photo documentation

Conclusion
It was agreed that for the combination and integration of data from different sources across the local, national and international level an agreement on the most relevant data types of those listed in D09 is necessary, with flexibility to add additional more detail data from the local level.

This led to the question whether or not information technology can help to integrate heterogeneous data provided by the stakeholders.

3.3 Experiences from Biotope mapping, Data interoperability and Geo-data-infrastructure (SDI)

Speaker 1: Dirk Hinterlang
Speaker2: Olaf Noelle

Introduction
How to build a decentralised information infrastructure with the help of modern information technique like internet, XML and software interfaces was subject matter of the second part of section B, where online-demonstrations showed how to use XML and Geographic Information services to combine spatial and other data provided by different data holders.

The discussion and decision making could profit from experiences with the XML-based exchange format in North Rhine Westphalia for Landscape Information (LINFOS) that is technically implemented as OSIRIS (object related data processing in a Geo-Information-System, see workshop presentation by Dirk Hinterlang) and the presentation and online-demonstration of Data Interoperability and GeoDataInfrastructures (GDS) that were developed *inter alia* for the GeoDataInfrastructure-Initiative (GDI) of the state of NRW (see workshop presentation by Oliver Noelle).

Biotope Mapping
Regularly biotope mappings are an important source for data related to the conservation of CWR. In that context Dirk Hinterlang presented his experiences from biotope mapping that is arranged by the State Agency for Ecology, Land Reform and Forestry NRW (LÖBF). The LÖBF work together with approx. 40 so called „Biologische Stationen", which are non-governmental institutions of public support for landscape management, and the
North Rhine Westphalia Foundation.
For the integration of data that are gathered in a decentralised way the information system OSIRIS consists of different modules. The object classes comprise *inter alia*:

- Natura 2000 - pSCI
- nature reserves
- specific biotope protection
- landscape reserves
- biotope mapping
- biotope types
- in situ localities and distribution (animals/plants)
- relevées
- vegetation types
- springs and sources
- management indicator species
- geological objects of nature conservation interest
  compound biotopes
  address administration
- and many more

The object class biotope mapping comprises:
- spatial reference (GIS)
- geometrical exactness
- status of protection
- relations to other objects (spatially/sequentially/specific)
- biotope type –
  vegetation type –
  vegetation layer –
  list of species –
  frequency/status/piece of evidence etc.
- description of locality
- grade of endangering
- management proposals and execution
- literature
- time and address of recording
- etc.

Data exchange between the locally dispersed players is facilitated by means of xml and an “OSIRIS namespace”, which is used to define the different object classes.

Data interoperability and Geo-data-infrastructure (SDI)

While central maintained data repositories – EURISCO is a good example - have a lot of advantages like syntactical and semantically homogeneous data sets, the need for interoperability and access to distributed data providers was stated. Global data sets are often used for GIS analysis although better data with higher resolution are available, e.g. as elements of the above mentioned Spatial Data Infrastructure. It was shown how internet technique and recent standardisation processes under the OpenGIS-Consortium can help to make better use of the geo-data that are available. An example are so called Web Map Services (WMS) that produce maps of geo-referenced data. WMS provide a visual representation of geo-data (map) via the internet. With the operations GetMap and Get FeatureInfo invoked by URL parameters, a map image and information about particular features (e.g. topography, soil,..) are shown on a map.

OpenGIS® Specifications support interoperable solutions that "geo-enable" the Web, important specifications are *inter alia*:

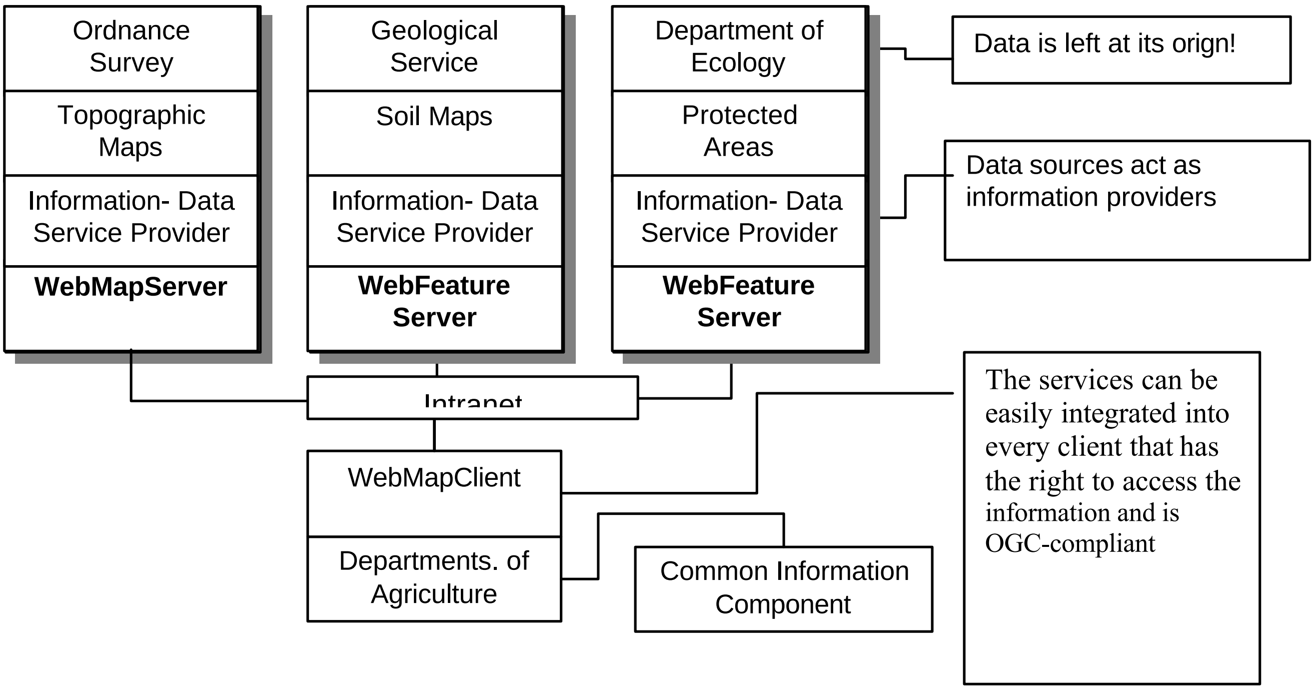
WebMapServices (WMS)
- produce maps

WebFeatureServices (WFS)
- allow clients the retrieval of geodata encoded in Geographic Markup Language (simple: vector data)

CatalogServices (CS)
- allows search for geoinfomation and geo-services

CoordinateTransformationServices (CTS)
- transforms coordinates between Spatial Reference Systems (SRS)

The Geography Markup Language (GML) is an XML encoding for the modelling, transport and storage of geographic information including both the spatial and non-spatial properties of geographic features. Service specifications (WFS) guarantee service-interoperability, whereas GML guarantees information-interoperability. The most challenging issue however is semantic interoperability that ensures that the content is understood in the same way within an interacting network of data providers. To tackle this issue formal ontologies for the PGR domain have to be developed.



Conclusions

D10 b) Appropriate data base structure identified

As a result of Session B the second major decision for *in situ* data management methodology was to use the OpenGIS consortium -software interfaces, XML (extensible markup language) and GML (geographical markup language) to build a decentralised information infrastructure for Crop Wild Relatives. For access and exchange of taxon related data an xml-schema was already introduced. This schema is based on a standard for data exchange (ABCD-schema) that is supported by the Global Biodiversity Information Facility (http://www.gbif.org). Additions for the requirements of Crop Wild Relatives were made.

3.4 Legal and intellectual property issues to the practical implementation of an information infrastructure

Speaker: Samy Gaiji

For designing a legal and intellectual property framework four major steps can be identified:

3.4.1 Definition of mission / objectives
3.4.2 Identification of risks
- a) Within the Infrastructure…
  - Members -> Infrastructure
  - Infrastructure -> Members
- b) Between Infrastructures…
- c) Between Infrastructure and Users

3.4.3 Design an IPR protection

3.4.4 Monitoring (1-3 years)

The design process is iterative and the steps (1) 2 to 4 are repeated after a monitoring period to adopt and optimise the framework. The identification of risks should be accomplished within the infrastructure, between different infrastructures and between the infrastructure and the users.

3.4.2 a) Identification of risk within the infrastructure

Example: copyright infringement
A dataset provided by member X of the network is subjected to copyright (i.e extract of Microsoft Encarta). Microsoft engaged legal claim against Y; Y engaged legal claim again X but no signed information transfer agreement were signed.

Example: roles infringement
A dataset provided by member X to Y is received by Z for data compilation. Z found some latitude and longitude errors and decided to modify the original dataset. X noticed the changes and complained to Y. No roles and responsibilities agreement were signed between Y and Z in relation to data integrity…

3.4.2 b) Identification of risk between infrastructures

Example: infringement of IPR
Y downloaded large datasets from X over its Internet site. Y incorporated this new information "as is" into its database. Y is a commercial company that sells each CD-ROM at 800 US$ and claims that it contains key information from X… X has never mentioned on its site that "the user will not claim ownership of, or seek to obtain IPR over the materials" downloaded.

3.4.2 c) Identification of risk with users

Example: commercial lost
Based on information in X, the species *A* had particular characteristics for cancer treatment. The company Y decided to invest 1 M$ in research. However after 1 year, the company Y discovered that the species name *A* was misspelled by X. Y initiated legal claim against X for commercial lost due to inexact information.

Example: absence of proper citation
Based on information in X, the researcher Y published a series of articles in Plant Genetic Resources and won and award from the "international community" and a check of 10,000 US$. No proper citation of X was made and Members of the network X decided to leave the infrastructure due to lack of credit/recognition of their work.

3.4.3 Design of IPR protection

The main elements of an IPR umbrella are:

a) Clearance mechanism and information policy
b) Roles and responsibilities
c) Attribution statement and citation
d) Terms of use
e) Disclaimers
f) Copyright notifications
g) Confidentiality on user inquiries

3.4.3 a) Clearance mechanism and information policy

Assuring there is a clearance mechanism used by the Members in acquiring/gathering and managing the data which is replicated to the Central platform, then:

- The management of information is a responsibility for each member.
- The information policy only applies to the core set of data that is replicated by the Members on to the Central platform.
- Even where the databases are linked to the Central platform each member may set its own terms of use, disclaimers etc. at the entry points to their information systems.
- It would be virtually impossible to impose and monitor different restraints or conditions associated with the disclosure or use of information on the Central platform.

Example of IPR element
The Central platform should only contain data which can be made publicly available; and can be used within the Network "terms of use" without limitation or restraint.

If information cannot be made publicly available then it should NOT be replicated on the Central platform (or it should be part of ANOTHER platform with different "terms of use"...)

3.4.3. b) Roles and responsibilities

Clarifying the roles and responsibilities between the Members and the Organization in charge of the Central platform, and agreeing on assurances to be given reflecting these roles.

Examples for the role of Members:

- Acquire/gather/manage the data
- Attend to the accuracy of the data
- Attend to the currency of the data
- Replicate the data into the Central platform at appropriate intervals
- Manage any authorizations or consents required in relation to using or replicating the data
- are the ultimate authoritative source for data confirmation and corroborative information

Examples for the role of the Central platform:

- Compiles Members data
- Provides public access to the data
- Facilitates and promotes access to the source of the data (Members)
- Manipulates and/or reformats the data to increase accessibility and utility

- Manages the tools and means on which the data is supplied to users
- Manages the legal basis for access to the Central platform and use of the data (disclaimers, copyright notifications, terms of use etc.)
- Manages responses to any abuses of the collective information at the Central platform level

It is recommended that the roles and responsibilities be formalized by way of the following assurances. It is proposed that these assurances be by an exchange of letters of agreement between the Organization in charge of the Central platform and each Member.

3.4.3 c) Attributions Statement and Citation

It is important to acknowledge that the Central platform is the result of the collective efforts of the Members. An attribution statement is an opportunity to state this publicly and in a manner satisfactory to all participating Members. Similarly, users should be encouraged to attribute the source of the information when using the Central platform by using a simple citation.

Example
SINGER is the genetic resources information exchange network of the International Agricultural Research Centres of the Consultative Group on International Agricultural Research (CGIAR). SINGER provides common access to information concerning the collections of genetic resources held by the CGIAR Centres and is a gateway to the independently managed information systems of the CGIAR Centres. SINGER would not be possible without the significant effort and valuable contributions of the CGIAR Centres.
This resource should be cited as: CGIAR SINGER.

3.4.3 d) Terms of use

Minimizing the risk of claims arising by inserting terms of use in the Central platform, which also has the benefit of making appropriate uses clear. “Terms of use” statements establish what are acceptable and non-acceptable uses of the Central platform data, and identify any specific requirements of the this Platform and the Members with respect to such use.

Examples:
The User may only access and use the Materials for the purposes of scientific research, breeding, genetic resource conservation and/or the sustainable management of genetic resources.

The Materials (including the results of searches conducted by the User) may not be sold or distributed for profit, including any commercial publication, reproduction, transmission or storage in a retrieval system of the Materials, in whole or part, without X’s prior written consent.
Any commercial use of the Materials is strictly prohibited without the express written consent of X.

Where the User wishes to publish, reproduce, transmit or store the Materials, in whole or part, for the purposes permitted by these terms and conditions, the User will ensure that the Materials are not modified or altered in any way and must include the copyright notice

3.4.3 e) Disclaimers

Minimizing the risk of claims arising. Disclaimers are designed to minimize or reduce liability associated with claims made by users in contract or negligence. It is impossible to definitively state how effective the disclaimers will be. Such disclaimers are traditionally disfavoured in the law of many jurisdictions.

Examples:
To the extent permitted by applicable law, X and the Members disclaim all warranties, guarantees or representations, express or implied, concerning the content, accuracy, completeness, fitness for purpose or other use of the Information stored in or made accessible
by this web site.

The User accepts the responsibility of verification and corroboration of the complete data (including provenience, taxonomic and ecological information) in consultation directly with the relevant Members.

The designations employed and the presentation of the Material stored in or made accessible by this web site do not imply the expression of any opinion of …

X and the Members shall not be liable for any losses or damages whatsoever, whether in contract, tort or otherwise, from the use of, or reliance on, the Information, or from the use of the Internet generally.

3.4.3 f) Copyrights Notifications

Copyright notifications are to create a means of control of the information by including notifications asserting copyright in the component parts of the Central platform. The most important function of the copyright notification is to assist in protection of the material.

Copyright notifications are a public assertion of rights, and are important to:

- Inform users of the material being the subject of copyright
  (placing on them an obligation to observe the exclusive rights of the copyright owners)
- Remove any opportunity to argue the limited defense of "honest infringement"
- In some jurisdictions the use of the copyright convention sign is recognized as an essential prerequisite to copyright protection of the material (pursuant to the Universal Copyright Convention).

3.4.3 g) Confidentiality on user inquiries

It is necessary to create a mean to protect users because:

Example
Some users of the Central platform may wish to ensure that the nature of their Inquirer on the database be kept confidential.

They may be concerned, for example, that by tracking their Enquirer an interested observer may be able to gain an insight as to the research being conducted by the user, or anticipate activities of the user. There may be political, academic or commercial reasons as to why the user wishes such confidentiality.

Conclusion
Based on the mission and the objectives of the information infrastructure a clear idea of the “product” has to be developed. The phase of the subsequently risk assessment is import and has to cover the relation between the Member and the Infrastructure, the relation to other Infrastructures and the relation between the User and the Infrastructure. A letter of agreement on the one hand and legal notice and license agreements on the other hands are the instruments to protect against the identified risks.

4. Session C: *In situ* analysis techniques
D11 In situ genetic conservation data management and analysis techniques identified

Speakers:

Luigi Guarino:	Spatial analysis methodologies for conservation of crop wild relatives
Mauricio Parra:	GIS Assessment of in situ and ex situ conservation of Lupinus spp. in the Iberian Peninsula
David Draper:	The role of GIS in the management of wild threatened populations of Narcissus
Andy Jarvis:	Tools for spatial analysis of PGR – State of the art and future needs –
Sabine Roscher:	Visualisation tools

4.1 Spatial analysis methodologies for conservation of CWR

Speaker: Luigi Guarino

The aim of in situ genetic conservation data management is to support the development and monitoring of conservation strategies for target taxa. A comprehensive overview of aspects related to in situ data management and analysis techniques was given by Luigi Guarino.
In general
- species-specific (Taxonomy, Ecology, Reproductive biology, Ethnobotany, Uses)
- population-specific (Locality (passport), Characterization, evaluation (genetic diversity), Genetic erosion)
- area-specific (Physical environment, Biotic environment, Human influence)

are the main data categories that are needed. Those data have to be compiled come from different data sources like documentation systems for *ex situ* collections, *in situ* reserves and others. It was pointed out, that the necessary precision of data and the appropriate analysis technique depends on the question that has to be solved.

Typical spatial questions in relation to in situ genetic management are:
- Where *can I* find a particular species*?*
- Where *can I expect to find germplasm with* specific adaptations*?*
- Where *are the* "hotspots" *of diversity?*
- Where *is the risk of* genetic erosion *or extinction highest?*
- Where *are the* gaps *in collections?*

Problems in using geo-referenced data in PGR documentation systems are:
- Data often missing
- Data often in hard-to-use form (e.g., degrees, minutes, seconds)
- Data are sometimes not very precise
- and are sometimes incorrect

Conclusion
It was concluded that following data analysis needs following input data to generate the listed resulting products:
Input
- Accession passport data
- Herbarium label data
- Administrative maps
- Present and future climate surfaces
- Protected area maps
- Environmental maps (altitude, vegetation etc.)
- Human impact maps

Output
- Species distribution maps
- Information on adaptation of germplasm
- Diversity hotspot maps
- Areas of genetic erosion
- Gaps in collections

Benefit: Better spatial targeting and prioritisation for conservation and use
As crucial for the enhancement of in situ conservation data management the clear definition of the right questions (spatial and of other kinds) was identified.

4.2 GIS Assessment of in situ and ex situ conservation of Lupinus spp. in the Iberian Peninsula

Speaker: Mauricio Parra

Case study I
For in situ conservation it is important to know which populations occur in protected areas. The presentation given by Mauricio Parra showed how GIS can help to assess the possibilities of conserving crop wild relatives in protected areas[2] [1]. By the example of Lupinus spp. in the Iberian Peninsula a GIS-based assessment of in situ and ex situ conservation was carried out.
The basic questions the case study dealt with were:
- How are the six *Lupinus* species distributed in the Iberian Peninsula?
- Where are the *Lupinus* "hot spots" of species richness located?
- What proportion of the species' populations is represented in ex situ conservation?
- Where are the gaps?
- How many Sites of Community Importance (SCIs) contain *Lupinus* species?
- How many *Lupinus* species are conserved within these SCIs?

Methodology
Data on the occurrence of *Lupinus* were compiled from various sources. In addition to passport data also herbarium specimens were compiled. The point data were superimposed to a UTM 1x1 km grid based distribution map. For the overlay with protected areas a SCI layer from the NATURA 2000 network was used. The map overlay was carried out by means of MapInfo.

[2] Parra-Quijano, M., D. Draper, and J.M. Iriondo Crop (2003): Assessing in situ conservation of Lupinus spp. In Spain through GIS. Crop wild relative, Issue 1, page 8-9

Result
As a result it could be shown that "a third of the *Lupinus* populations are potentially protected, although most SCIs do not presently provide effective protection due to a lack of specific management. Just 1,47 % of the existing SCIs matched with the *Lupinus* locations at the UTM 1x1 km grid level." [1]

4.3 The role of GIS in the management of wild threatened populations of Narcissus

Speaker: David Draper

Case study II
The important role of GIS for the management of wild populations was also shown by David Draper, who presented a case study on wild Narcissus in Spain. GIS was used for following conservation related activities:
- Selecting important areas for conservation
- Definition of vulnerable areas
- Population management
- Selecting suitable places for local cultivation
- Optimisation of actions/ efforts
- Minimisation of infrastructure impact

The practical management aspect to which a solution had to found was how to find a new location with the best suitability for the translocated population, that was one of the only two localities of this species in Portugal and under threat because anthropogenic influence (dam construction).
To solve this problem, spatial factors like e.g. spatial distance to the origin, distance to rivers, located outside dams, having appropriate land use, located inside protected areas, impact of global change were combined within a GIS to calculate suitability indexes. In combination with field validations of the model, the GIS-based methods could be used to suggest new suitable locations for the population.

4.4 Tools for spatial analysis of PGR distribution - State of the art and future needs Data analysis techniques

Speaker: Andy Jarvis

The focus on the presentations and discussion was on spatial data analysis techniques and models, general GIS-based packages for visualisation, analysis and management of spatial data like ArcInfo, WorldMap or DIVA were not subject-matter of the workshop.

Following the presentation of Andy Jarvis, the basic approach to species prediction is like following:

- For the prediction of species distribution training data of point collections where the species has been observed are used.
- With associated variables which are identified as (potentially) drivers of the distributions as input parameters a statistical model is run.
- The resulting product is a continuous surface of the predicted distribution.

Choosing the appropriate method for predicting the distribution of an organism is often difficult. The methods vary depending on many factors and no universal solution can be given. However the principle points to consider are:

- Knowledge about the biology/geography of the species
- Knowledge about the principal factors determining the distribution
- Proper scale and availability of data on the factors at that scale
- Precision for the prediction of the distribution
- Which biological level (gene, trait, species etc.) is of interest

Limits to the application of statistical methods are mainly determined by the input data. Data on plant genetic resources are typically point observations of species. These data are neither complete nor random sampled, but are Geographically biased (e.g. close to roads), and only based on presence – rarely any data on species absence are recorded.

The number of points that are needed to predict distribution depends on representativeness of data. So it is relevant whether or not the data points cover the full environmental gradient that the species inhabits. Also it has to be taken into account that different methods are more robust than others with limited data points. Known from experience could be said that a minimum of 10 points are needed, but ideally more that 30 should be used. Generally speaking the more points, the better.

The drivers of distribution (Climate, Soils, Land-use/habitat, Biology/ecology (pests, pollinators, migration etc.), Geography (barriers etc.), Topographic (aspect, slope, wetness), Spatial heterogeneity (for diversity), Evolutionary history) are scale depended. For example climate data are mainly used on the macro scale, topography on the meso scale and while biological data like pollinators or pest are more related to the micro scale. During the presentation a brief overview of data availability was given

Altitude and Topographic derivatives

- Global datasets
  - GTOPO30 – 1km resolution
  - SRTM "Shuttle Radar Topography Mission" elevation data – 30m - 100m resolution

Altitude data provides basic information that is often needed to produce climate surfaces and calculate many other derivatives like aspect, slope, solar radiation receipt, drainage network.

Climate

Climate data are interpolated from climate stations. For the interpolation other variables are needed (co-variable interpolation). For this purpose normally used is altitude, but also other data are sometimes taken into account.

- Global datasets (1km grids) for monthly means and basic variables like minimum and maximum temperature, and precipitation are available. Other factors, including length of dry season and Potential Evapo-Transpiration (PET) can be calculated from these data.
- Higher resolution surfaces can be produced given sufficient climate stations and understanding of the driving processes

Soils

- Soil type are non-continuous data and incompatible with some models
- Ph and soil texture are continuous data, but often not available

- Global level:
  - FAO Soil Map of the World
  - Coarse scale, many problems
- National level:
  - Often different classification systems between countries – makes multi-country analyses impossible
  - Not freely available for many countries
- Derivates:
  - Soil moisture
  - Soil texture (more advanced modelling)

Land-use/Habitat

- USGS Global 1km Land Cover data
  - Good at showing the big picture
  - Low resolution provides little help on the ground
- Global/continental/national ecosystem/ecoregion maps
- LANDSAT + other satellite imagery gives higher resolution
  - NDVI useful in locating specific habitats
  - Multi-temporal analyses to analyse best time to collect

Methods and tools for predicting species distribution

a) FloraMap

Method

Principal components analysis coupled with a probability model
Cluster analysis also can refine result and identify different adaptations in organism

Input Datasets

- Built-in climate surfaces of 36 monthly means of temperature, precipitation, and diurnal range in temperature
- 10 minute resolution for all of tropics, Europe and North America, 2.5 minute Asia, and selected 30 second climate grids for specific countries/regions

Result

- Probability surface of finding the species, ranging from 0.001 – 1
- PCA report with loadings and scores in order to further analyse controlling factors
- Climate data at each observation point

Pros:

- Built-in global climate datasets, with differing resolution depending on area
- Output provides full range of probability of finding species
- Simple to use
- PCA report hints at contributing climatic factors defining the distribution
- Cluster analysis can further refine predicted distribution to account for within species variation
- PCA methodology throws out input variables with little predictive capacity

Cons:

- Restricted in input datasets – must be FloraMap produced climate grids
- Few options for model parameterisation provided greater knowledge of species

b) GARP - Genetic Algorithm for Rule-set Production
http://www.lifemapper.org/desktopgarp/Default.asp?Item=1&Lang=1

Method
Genetic algorithm performs series of random runs examining the "fitness" of each grid cell based upon input data at point of observation.
Input Datasets
- 6 minute surfaces of precipitation, mean temp., wetness, aspect, slope, elevation
- User defined datasets can be imported from Arc ASCII format

Result
- Series of maps for each user defined iteration with presence (1)/absence (1)
- Table report for each iteration of model parameters, no. of points lying inside distribution, distributional statistics

Pros:
- Has built-in system to analyse predictive accuracy using split-sampling
- Means of automating procedure for numerous species
- Choice of model rules: Atomic, Range, Negated Range, Logistic Regression, and flexibility in model parameters
- Complete flexibility in choice of input data, though built-in data is poor

Cons:
- Provides little insight as to the controlling factors in the distribution
- Only predicts presence/absence – no range of probabilities
- Tends to over predict

Predicting species distribution with "Domain"
http://www.cifor.cgiar.org/scripts/default.asp?ref=research_tools/domain/index.htm

Method
Generates maps of similarity or distance from the conditions at known points of presence, using the Gower metric method.
Input Datasets
- User defined datasets imported from Arc ASCII format
- Will handle non-continuous data, such as soil type

Result
- Map of similarity, from 100 (found at each presence point) down to 0 (if found completely different). Usual result ranges from 70 – 100.

Pros:
- User has flexibility in choice of input data
- Deals with both continuous and non-continuous data
- Very easy to use, and based on simple but understandable concepts

Cons:
- Result is dependent on choice of input data
- Does not provide clear boundaries between predicted presence/absence
- Doesn't provide insight as to the drivers of species distribution
- Statistical method not particularly robust to small input sample sets

c) General Methods and Tools:
DIVA http://diva-gis.org/program.htm

Description
General system designed to visualise spatial distribution of biological collections.

Datasets included
- Global climate and topography at 10 minute resolution
- User defined datasets imported from variety of formats (BIL, IDRISI etc.)
- EcoCrop species information
- Downloadable gazetteers of town names for the globe
- Downloadable administrative zones for the globe
- Downloadable land-use datasets for the globe
- Downloadable population datasets for the globe
- Supports satellite imagery downloaded from MrSID image server

Specific Functions
- Map the locations of sites where populations of plant or animal species were observed, and of different characters that may have been recorded for these populations.
- Make grid maps of the distribution of biological diversity, and identify "hotspots" and areas that have complementary levels of diversity.
- Extract climate data for localities points, and predict the presence of species based on climate, using the BIOCLIM or DOMAIN models
- Grid manipulation (calculator, aggregate, neighbourhoods)
- Projection of lat/lon data
- Spatial autocorrelation
- Histograms and scattergrams of gridfiles
- Visualization of satellite images
- Prediction of crop adaptation (Ecocrop)

d) Methods and Tools: Others

Generalized Linear Models (GLMs)
Logistic regression and multiple linear regression – David Draper
Pros: Flexible modelling approach for advanced users, easily understood and communicated methods, transparent results
Cons: No software yet uses this method, requires pseudo-absence data, requires statistical experience, time-consuming especially for beginners, requires GIS package (IDRISI, ArcView)

Ecological niche factor analysis (BIOMAPPER - www.unil.ch/biomapper/)
Pros: Very complete and flexible package for predicting distribution, choice of methods, choice of input data, built-in validation statistics (Kappa + ROC)
Cons: Inputting data is difficult and requires prior GIS knowledge and software, sometimes not user friendly, requires continuous data
BioClim - http://cres.anu.edu.au/outputs/anuclim/doc/bioclim.html
http://www.wiz.uni-kassel.de/model_db/mdb/bioclim.html
Bounding box method

Validation of results
Difficult to have confidence in presence/absence results

Statistical methods for validation
- Kappa (simple presence-absence matrix)
- Receiver operating characteristic (ROC)
- …also Bootstrap and cross-validation

MANEL et. al. (2001) find kappa to be well correlated with ROC, and much easier to calculate.

Conclusion
Unless absence can be confirmed, no validation can give a definitive answer on model success

Model comparison, Comparison of Predictive Models

Method	**Prediction Accuracy (% validation points predicted correctly)**
GARP (present)	0.91
FloraMap (> 0.5 probability)	0.89
DOMAIN (> 99% similar)	0.53
BIOCLIM (optimal/sub optimal)	0.70

NB Potential caveats include:
- Different predictive variables and at different scales used
- Example for 1 species only – models likely to perform differently depending on datasets (no. points, representativeness, distribution, characteristics etc.)

Example: The exploration and *ex-situ* conservation of *Capsicum flexuosum*
- Uncommon species of wild chilli, found in Paraguay and Argentina
- 18 known registers of the plant
- 2 germplasm accessions conserved in the USDA
- Genetically unknown
- Found in an area undergoing high levels of habitat loss

Results
- One plant found with few seeds, where previous herbarium record was taken
- First accession conserved *ex situ*
- 1 plant found, with few seeds
- new collections of *C. flexuosum*
- 160 seeds conserved *ex situ*

Conclusions
- Choice of species distribution modelling approach depends on many factors
- Spatial analysis of PGR is now an established field, and one capable of providing relevant information to conservation projects
- Confidence levels in results should be measured, though these methods require further development in the coming years

- Field data collection should record sites where species not found, with clear description of habitat and effort made searching

4.5 Visualisation tools

Speaker: Sabine Roscher

Scientific visualisation can help to understand the complex interaction between plants and their environment because visual cognition is the most adequate sense for understanding complex information. With the raising availability of complex data and data types in geography and biology and the steadily increasing power of hard- and software scientific visualisation is a powerful method, developed and used by disciplines like geology and landscape planning. Especially the handling of the 3rd dimension (space) and the 4th dimension (time) is getting more and more important. Furthermore simultaneous illustrations and analysis of different topics like temperature in combination with precipitation and growth (5th dimension) is relevant for questions related to the interaction of plants and their environment.
It was stated that the application of scientific visualisation in the context of PGR is absolutely under-utilised.

Following scientific principles are identified as relevant for the development of adequate visualisation tools:
- Informatics (Development and allocation of basic technology)
- Geoinformatics (thematic and spatial referencing, "Interface human being - machine", bridge a gap)
- Cartography (provide basic principles for visualisation)
- Cognition science (consideration of cognition of noticeability)

Expected benefits from application of scientific visualisation are:
- Discovery of new interrelations between plants and their environment
- New techniques for interaction and demonstration allow a more intuitive and fast understanding of multidimensional processes (genes, organism, landscape, space and time)
- Enhancement of outreach and capacity building by portraying spatial information on Crop Wild Relatives quickly and easily for most users, requiring only map reading skills without knowledge about GIS-tools.

5.0 Summary outcomes

5.1 European crop wild relative information infrastructure

- A preliminary conceptual model for an information infrastructure for European crop wild relatives (CWRs) is proposed (Figure 1)3
- The European CWR taxon database is at the core of the information infrastructure and linked to n number of external data sources

5.2 European crop wild relative database

- Database of taxonomic and crop data curated by UoB during the lifetime of PGR Forum, to contain genus, species, related crop, gene pool / taxon group
- Version 1.0 to be made available for download from the PGR Forum project intranet as soon as possible post WS3
- Subsequent versions to be made available during the remaining lifetime of the project4
- XML schema for taxonomic data to be produced by Sabine Roscher, Theo van Hintum, Shelagh Kell and Nigel Maxted
- Structure to comply with the GBIF (Global Biodiversity Information Facility) standards where possible (see http://www.gbif.org/)
- Nomenclature to be updated from Euro+Med: technical issues, timing and practicalities to be discussed by taxon list working group
- Crop names to be updated from Mansfeld's World Database of Agricultural and Horticultural Crops: technical issues, timing and practicalities to be discussed by taxon list working group
- Data verification process via PGR Forum participants and other country focal points to be initiated
- Data sources and linkages to be investigated
- Taxon list working group to coordinate the above activities: Shelagh Kell, Stephen Jury, Helmut Kn?pffer, Sabine Roscher, Nigel Maxted, Brian Ford-Lloyd

5.3 European crop wild relative database user requirement analysis

- Coordination to re-circulate request for potential user requirements within PGR Forum
- Coordination to collate and analyse user requirements and formulate examples of use cases
- Document reviewed and approved by Steering Committee, Advisory Board and Stakeholder Panel
- CWR database user document to be circulated within PGR Forum
- PGR Forum participants to liaise with potential CWR database users in their country to gain feedback on summary CWR database user document and to solicit further user requirements
- Coordination to circulate regional potential CWR database users and any key national user not contacted by the local PGR Forum members to gain feedback on summary CWR database user document and to solicit further user requirements
- Produce final list of CWR database user needs
- Timescale for these activities to be agreed among PGR Forum participants

---

[3] The implementation of the infrastructure in terms of timing and resources may not be feasible within the lifetime of the project (also see point 6.)

[4] See the following documents for further details on the development of the taxon list: "Workpackage 1: European crop wild relative assessment progress report" and "Development of a list of European crop wild relative taxa"

5.4 European crop wild relative data types and data sources

- Sabine Roscher, José Iriondo + Coordination to refine current draft of CWR data types
- Second level refinement to be carried out according to user requirements identified in stage 3 above
- Draft list of CWR data types to be circulated within PGR Forum
- Sabine Roscher to coordinate investigations into XML data sources and potential linkages
- Data types, sources and data standards to be discussed via web site bulletin board[5]

5.5 European crop wild relative case studies

- CWR information infrastructure to be tested and applied using case study taxa
- Coordination to formulate and circulate revised proposal for establishing the selection of case study taxa[6]
- Finalise list of CWR taxa selected to present exemplar case studies for CWR conservation[7]
- Investigate linkages with ECP/GR re. hosting of project products following project completion, including potential linkage with elements of the ECP/GR documentation and information network, such as EURISCO and the ECCDBs (European Central Crop Databases)
- Investigate technical, legal and IP aspects of implementation
- Data sources and analysis methodologies to be identified and considered at all stages
- Identify and communicate with data providers during development

[5] UoB to provide and configure software for bulletin board; discussions to be chaired by José Iriondo and Sabine Roscher, and moderated by Shelagh Kell

[6] Building on the taxa chosen and data collated for selected taxa pre-WS3

[7] List to include a range of taxa suitable to illustrate the application of the IUCN Red List Categories (Workpackage 2)

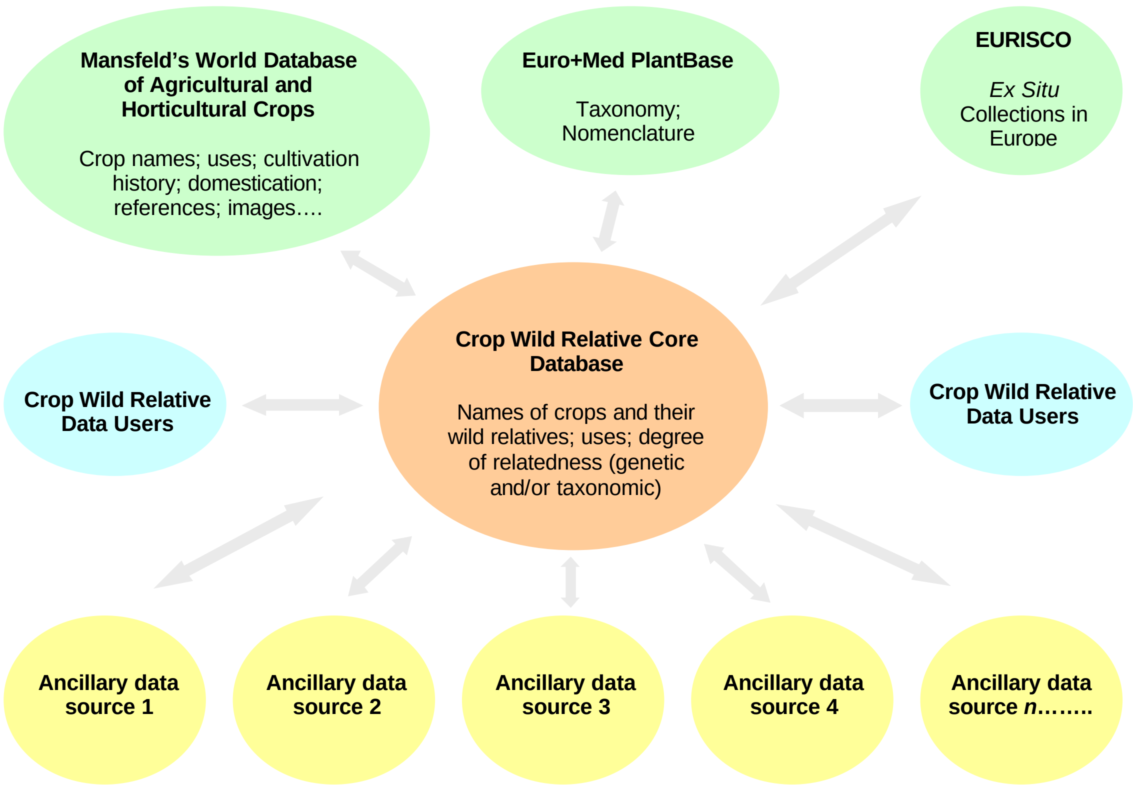


Figure 1: Proposed conceptual model for a European Crop wild relative information infrastructure. The European CWR taxon database is at the core of the information infrastructure and linked to *n* number of external data sources. Some existing data sources are indicated. The two-way arrows indicate the reciprocal nature of the system.

Appendix 1: Materials - example for use case

Use case: Distribution map (CWR-Network)
Actors: Primary actor: biodiversity researcher
Description: Biodiversity Researcher requests a distribution map for a given species, optionally restricted to some geographic region and/or period of time. The CWR Network locates relevant observation records and generates a map. There may also be user options to restrict records mapped to some classes of data rather than others.

Basic Flow

1. Biodiversity Researcher submits request including species name and optional geographic region and/or period of time.
2. CWR Network generates request to Find Global Observations for the given request, specifying only minimal detail (which can be supplied from within the index itself).
3. CWR Network constructs map showing location of all observations found.
4. CWR Network returns map to Biodiversity Researcher.

Post-conditions

None

Alternative Flows

1a. Biodiversity Researcher may submit request for a taxon of a rank higher than the species level, or for multiple species. Subsequent mapping steps may distinguish different species by different visual indicators.
2a. In the future other classes of biodiversity data held within the network may also be useful for improving the quality of the mapping. These other sources could also be queried at this point.
3a. Additional processing will be required to reflect the variable accuracy associated with each georeference data-point.
3b. The results may be returned as a GIS layer rather than an image.

Appendix 2: Material - Example meta data spatial data

Title	
ShortTitle	
Responsible partiy organisation name	
Reponsible party role code	
Responsible party individual name	Name of contact person
PostalAdress	
City	
PostalCode	
Country	
URL_Institution	
Email	
Fax	
Phone	
West bounding coordinate	
East bounding coordinate	
North bounding coordinate	
South bounding coordinate	
GeographicExtentName	
Temporal extent date/time form	
Temporal extent date/time to	
ResolutionLevelCode	
LanguageOfDataSetCode	
Abstract	
ProgressCode	Status
DegreeOfDataCapture	Erfassungsgrad
ThemeCode	e.g. CWR
Keywords	Thesaurus
Keywords	Free Text
UseConstrains	
LineageStatement	SourceOfInformation
QualitativeNarrativeReport	Data Quality
DistributionFormat	DataFormat
DistributionMedia	
URL	
LanguageOfMetaDataCode	
MetaDataDate	

Appendix 3, Final Agenda

PGR Forum **Workshop: In situ Data Management Methodologies**		
Day 1	Monday 8 Sept.	
09:00	**Opening Session**	
	Welcome by Hosting institute	(Zdenek Stehno, 5 min)
	Welcome by PGR-Forum Co-ordinator	(Nigel Maxted, 5 min)
	Introduction to Workshop 3	(Sabine Roscher, 10 min)
	WP2 update	(Caroline Pollock, 10 min)
	WP4 update	(José Iriondo, 10 min)
	WP5 update	(Sónia Diaz, 10 min)
	WP6 update(Conference)	(Brigitte Laliberté, 10 min)
	WP6 update(Publications)	(Nigel Maxted, 10 min)
10:30	Coffee Break	
11:00	WP1 Progress report - Taxon list development - Collation of trial conservation datasets - Project database	(Shelagh Kell, 30 min)
	WP1 Deliverables: Discussion Session	
	Session A: Data sets, data types and data standards	
	Data on CWR in the UK, a case study	(Maria Scholten, 20 min)
12:30	Lunch	
14:00	*In situ* Models and Data Management	(Nigel Maxted, 20 min)
	Report WG1: "Taxon based data, Distribution data"	(Sabine Roscher, 20 min)
	Report WG2: "Ecogeographic data"	(José Iriondo, 20 min)
	How can existing habitat conservation and classification projects help to conserve CWR?	(Doug Evans, 20 min)
15:30	Coffee Break	
16:00	Parallel sessions on data types and data(base)structure WG 1 „Distribution data, taxon based data" WG 2 „Ecogeographic data" WG 3 „Habitats" WG 4 „On Farm"	
	Presentation of WG results	(WG1, WG2, 15 min each)
	Discussion	
	Finalisation of data types and draft data structure	
18:00	Closing first day	

Day 2	Tuesday 9 Sept.	
09:00	Presentation of WG results "Habitat" and "On-Farm"	(WG3, WG4, 15 min each)
	Session B: *In situ* data management	
	Data management on the local level	(Wieslaw Podyma, 20 min)
	Quality standards for data capture	(Vojtech Holubec, 20 min)
	Experiences from Biotope Survey - the Relevance for Mapping and Data Management of CWR -	(Dirk Hinterlang, 20 min)
10:30	**Coffee Break**	
	Data Interoperability and GeoDataInfrastructure	(Olaf Noelle, 60 min)
	Discussion, recommendations for data management	
12:30	**Lunch**	
14:00	**Session C: *In situ* analysis techniques**	
	Spatial analysis methodologies for conservation of crop wild relatives	(Luigi Guarino, 40 min)
	GIS Assessment of *in situ* and *ex situ* conservation of Lupinus spp. in the Iberian Peninsula	(Mauricio Parra, 15 min)
	The role of GIS in the management of wild threatened populations of Narcissus	(David Draper, 15 min)
	Discussion	
15:30	Coffee Break	
16:00	Tools for spatial analysis of PGR – State of the art and future needs –	(Andy Jarvis, 40 min)
	Visualisation tools	
	Discussion, recommendations (spatial) data analysis tools and techniques	
17:00	Closing second day	
Day 3	Wednesday 10 Sept.	
9:00	**Session D: Summary of the days before, perspectives**	
	Legal isssues and intellectual property to the practical implementation of an information infrastructure for crop wild relatives	(Samy Gaiji, 20 min)
	Discussion: Data use, user needs, Use Cases	
10:30	**Coffee Break**	
16:00	Presentation of draft data structure for *in situ* data	(Sabine Roscher)
	Overlap with exchange format *ex situ* data	
	Discussion	
	Adoption of draft data(base)structure	
	Future steps	
	Discussion	
	Results, Summary, Closing	
13:30	**Lunch**	
15:00	Excursion to Praha **(~3 hours)**	

Appendix 4, List of Workshop Parti cipants

COUNTRY	GIVEN_NAME	FAMILY_NAM	Institution	EMAIL
Belgium	Andre	Toussaint	Gembloux Agricultural University	baudoin.jp@fsagx.ac.be
Colombia	Andrew	Jarvis	International Plant Genetic Resources Institute Columbia	a.jarvis@cgiar.org
Czech Republic	Vojtech	Holubec	Research Institute of Crop Production	holubec@vurv.cz
Czech Republic	Zdenek	Stehno	Research Institute for Crop Production	stehno@vurv.cz
Denmark	Kell	Kristiansen	Danish Institute of Agricultural Sciences	kell.kristiansen@agrsci.dk
Fiji	Luigi	Guarino	International Plant Genetic Resources Institute	LuigiG@spc.int
Finland	Juha	Helenius	University of Helsinki	juha.helenius@helsinki.fi
France	Dough	Evans	European Environmental Agency	evans@mnhn.fr
France	Martine	Mitteau	Bureau des Ressources Génétiques	martine.mitteau@inapg.inra.f
Germany	Frank	Begemann	German Centre for Documentation and Information in Agriculture	begemann@zadi.de
Germany	Sabine	Roscher	German Centre for Documentation and Information in Agriculture	roscher@zadi.de
Germany	Helmut	Knüpffer	Institute for Plant Genetics and Crop Plant Research	knupffer@ipk-gatersleben.de
Germany	Dirk	Hinterlang	Landesanstalt f. Ökologie, Bodenordnung und Forsten NRW	dirk.hinterlang@loebf.nrw.de
Germany	Olaf	Nölle	University of Münster	noello@afi.uni-muenster.de
Greece	Stelios	Samaras	Greek Gene Bank	kgeggb@otenet.gr
Hungary	Attila	Simon	Institute for Agrobotany	jensen@grrobot.rcat.hu
Italy	Petra	Engel	Instituto Sperimentale per la Frutticoltura	f.grassi@mclink.it
Italy	Brigitte	Laliberte	International Plant Genetic Resources Institute	B.Laliberte@cgiar.org

Italy	Samy	Gaiji	International Plant Genetic Resources Institute	s.gaiji@cgiar.org
Italy	Thomas	Metz	International Plant Genetic Resources Institute	t.metz@cgiar.org
Latvia	Isaak	Rashal	University of Latvia	izaks@email.lubi.edu.lv
Lithuania	Juozas	Labokas	Institute of Botany	labokas@botanika.lt
Norway	Aasmund	Asdal	Norwegian Crop Research Institute	aasmund.asdal@planteforsk.no
Poland	Wieslaw	Podyma	National Centre for Plant Genetic Resources	w.podyma@ihar.edu.pl
Portugal	Eliseu	Bettencourt	Instituto Nacional de Investigação agrária e das Pescas	e.bettencourt@meganet.pt
Portugal	Sónia Ricardo	Dias	Instituto Nacional de Investigação agrária e das Pescas	soridi@net.sapo.pt
Romania	Silvia	Strajeru	Genebank of Suceava	genebank@assist.ro
Russian Federation	Tamara	Smekalova	N.I.Vavilov Research Institute of Plant Industry	s.shuvalov@vir.nw.ru
Slovak Republic	Daniela	Benedikova	Research Institute of Plant Production	benedikova@vurv.sk
Spain	José M.	Iriondo	Universidad Politécnica de Madrid	iriondo@ccupm.upm.es
Spain	David	Draper	Universidad Politécnica de Madrid	ddraper@fc.ul.pt
Spain	Mauricio	Parra	Universidad Politécnica de Madrid	senderos@bio.etsia.upm.es
Spain	Lori	De Hond	Universidad Politécnica de Madrid	optima@tiscali.es
Sweden	Dag Terje	Endresen	Nordic Gene Bank	dagterje@ngb.se
The Netherlands	Theo J. L.	Hintum van	Centre for Genetic Resources the Netherlands	Theo.vanhintum@wur.nl
United Kingdom	Shelagh	Kell	University of Birmingham	s.p.kell@bham.ac.uk
United Kingdom	Stephen	Jury	University of Reading	s.l.jury@reading.ac.uk
United Kingdom	Nigel	Maxted	University of Birmingham	N.Maxted@bham.ac.uk
United Kingdom	Maria	Scholten	University of Birmingham	GUEST655@bham.ac.uk

United Kingdom	Caroline	Pollock	IUCN-The World Conservation Union	Caroline.Pollock@ssc-uk.org