2012

An Approach to the Virtual Flora of Mongolia – From a Data Repository to an Expert System, http://greif.uni-greifswald.de/floragreif/

Jörg Hartleib
University of Greifswald, hartleib@uni-greifswald.de

Martin Schnittler
University of Greifswald

Sabrina Rilke
University of Greifswald

Anne Zemmrich
University of Greifswald

Bernd Bobertz
University of Greifswald

Erforsch. biol. Ress. Mongolei (Halle/Saale) 2012 (**12**): 287–295

An approach to the virtual flora of Mongolia – from a data repository to an expert system

http://greif.uni-greifswald.de/floragreif/

J. Hartleib, M. Schnittler, S. Rilke, A. Zemmrich, B. Bobertz, U. Najmi, R. Zölitz & S. Starke

Abstract

FloraGREIF is an internet accessible information system providing taxonomic, phytogeographic and ecological information on Mongolia's flora in terms of descriptions, high-resolution plant images and an interactive WebGIS application. Organised along an updated checklist of the approx. 3000 Mongolian vascular plants that serves as a taxonomic backbone, information is split into the taxon level, referring to plant species, and the record level, referring to record or a collected plant specimen. At the latter level, images of living plants, scans of herbarium sheets, habitat photos and further notes can be found. Both data levels are linked by the name of the respective plant species. An integrated WebGIS application enables the spatial visualisation at the taxon (distribution within phytogeographic units) and record level (dot maps). FloraGREIF is intended to be an internet accessible application tool for ecological and taxonomic research on Mongolia's vegetation, but can serve as well as a data repository, e.g. for data from relevés survey.

Key words: WebGIS, biogeography, determination key, taxonomy, biogeography, geobotanic, ecology

Introduction

FloraGREIF is a web application that aims at the aggregation of fundamental knowledge about the flora of Mongolia. Decades of research on taxonomy and plant ecology beginning in the 1960s resulted in extensive body of biodiversity information about Mongolia (GUBANOV & HILBIG 1989, PAVLOV et al. 2004, HILBIG 2006). Mostly published in Russian, this information comprises herbarium collections of vascular plants, repeated revisions of the countries' floristic checklists, determination keys for vascular plants of the country and revisions of certain taxa, a biogeographic division of the country and selected regions, vegetation classification and zoning for the entire country and particular regions. Only a few of this information is compiled in maps. Much of this data is hardly accessible because published in Russian language and scattered over a very few libraries of Western countries (ZEMMRICH et al. 2013).

Alternatively, the rapid development of digital data processing and storage supports the provision of extensive information in terms of text data, images, and interactive maps in worldwide accessible internet databases (CBIF 2006, USGS 2009). Against this background, the FloraGREIF web application has been developed to offer taxonomic, ecological and biogeographic information on Mongolia's flora to the World Wide Web public. Specific objectives of the FloraGREIF data platform include (1) accessibility of herbarium collections of vascular plants that are stored in Germany, (2) facilitation of online-comparisons of collected plant material for experts and the interested public, (3) conversion of spatial information on plants and vegetation into interactive application of WebGIS maps and finally (4) the establishment of a digital repository of existing biodiversity information in text, images and maps.

FloraGREIF organisation

The FloraGREIF system comprises a database of taxon data involving information about vascular plant species, a database presenting detailed information on records, a query algorithm to search at both levels, and an interactive WebGIS application to visualise spatial information such as

species distribution within the countries' floristic regions and localities for records assigned to a certain taxon (ZEMMRICH et al. 2013). In addition, an overview of literature, separately for cited and further biodiversity-related literature including major publications in Russian language (given with English translations of titles and keywords) is provided. Record data include high-resolution scans of herbarium sheets, sometimes additional close ups of details that are of diagnostic value, in vivo plant images as well as habitat images. Both data levels are linked by the name of respective plant species. Data quality and homogeneity are guaranteed by pre-processing steps such as revision of plant determinations from various sources through experts, adding further information to existing herbarium labels and, if necessary, georeferencing of sheets (including reconstruction of geographical coordinates and assignment to the administrative units Aimag (districts) and Sum (municipalities).

Initial situation

Large herbarium collections of Mongolian plants exist in German universities and further research institutions such as the University Halle (HAL), the Leibnitz Institute Gatersleben (GAT), the Universities of Jena (JE), Osnabrück (OSBU), and Greifswald (GFW). These collections and their information are hardly accessible to the public through limited operational hours and loan restrictions for reasons of collection conservation. Moreover, determination books have long dichotomous keys, use sophisticated technical terms, and are rarely illustrated by informative pictures of the plants. They are published in limited editions and are usually cost-intensive. Only a notable exception is the volume on Mongolian plants of HAUCK & SOLONGO (2010). Further potential data sources involve plant images of excellent quality by experts, but are of limited accessibility. Additional barriers comprise the inconsistency of writing styles of localities and place names for sampling records taken in Mongolia.

Objectives and implementation

Currently, FloraGREIF team focuses on three main issues: developing a determination key, extending the WebGIS system in terms of content and interactive capabilities, and expanding the record data by screening available herbarium collections.

To implement this information, we have to carry out digitalisation, taxonomical data processing, georeferencing and spatial processing via GIS, and combining all information within one web portal.

To digitalise specimen information, the respective collections are screened. Representative specimens are selected and scanned with a Herbscan device, which was developed originally by a team of Royal Botanic Gardens, Kew (Kew Gardens 2006), with resolution of 1200 dpi, covering sizes up to A3 format. If necessary, additional close ups of diagnostic details like seeds or flower structures are provided, images are taken using a macro lens and bellow on a digital mirror reflex camera mounted on a microscope stand (fig. 1).

Labels of the herbarium sheets are stored in text sources in the database and completed by further spatial data stating provinces and counties. Wherever possible, the determination of the specimen was reviewed by an expert for the respective plant group.

The taxonomic backbone of the database is the vascular plant checklist of GUBANOV (1996, 1999). This list was updated for several taxa found or described as new for the country in recent years. Species names are entirely checked against the Names in Current Use (NCU) checklist (GREUTER et al. 1993) and the International Plant Names Index (IPNI 2012).

The information stored in the database is partitioned in two divisions: taxon and record data. Taxon data include a brief description of the species mainly based on diagnostic characters, information on habitat preferences, and distribution within the phytogeographic regions of the country based on GRUBOV (1955). The record data holds all specimen information, such as taxonomic affiliation, collector, determiner and, if appropriate, reviser, flowering status, the herbarium collec-

Fig. 1: Example of a scan of a herbarium specimen (left), macro image (center), and macro of living plant (right) of *Stellera chamaejasme*.

tion where the specimen is stored, locality, habitat, herbarium scans and field images. For images date and habitat where the image was taken, and photographer's name are given.

Information about the distribution of plant species can be found via the WebGIS application. Topographical maps, satellite images, and altitude tints give basic information about the region of interest within Mongolia. Geobotanic units based on GRUBOV (1955) and vegetation zones were digitalised and are provided as map layers (ZEMMRICH et al. 2013). For the functionality of the WebGIS as an interactive application tool, primary spatial data have to be converted into suitable formats following common standards, like Web Mapping Service (WMS) by Open Geospatial Consortium (OGC). The WMS creates maps on a server which can be submitted to any client application. FloraGREIF is dealing with the WebGIS application in two ways: It is used to display distribution maps inside taxon and record descriptions and as standalone WebGIS application embedded in a website, which creates dynamic maps according to the content currently available in the database.

Geographic data are handled by two GIS programs: Moskito-GIS (Moskito-GIS GmbH 2012) and Mapserver (open source, Mapserver.org 2012). The interface for managing the map, like navigation on the map, switching layers and so on, is provided mainly by the JavaScript library OpenLayers (open source, OpenLayers.org 2012). It is extended by custom-developed scripts for the needs of FloraGREIF.

The implementation of the FloraGREIF information system involves several techniques (see fig. 2). There are three databases hosted on different servers: taxon and record database server, image server, and geodata server. All data sources are combined through the web portal provided on a web server. The server system is protected against unauthorised access with a firewall. External user can interact with the information system via the website http://greif.uni-greifswald.de/floragreif/. Reproduction and publication of species descriptions and specimen in words and pictures is allowed with written permission of copyright holder only. The system supports a multi-user functionality. Users are assigned to different access groups. Regarding to this group, a user is allowed to edit or upload content.

Functionality

In its current state, taxon data, record data, and all connected images can be searched by species, genera, and family names. Advanced search criteria can be submitted to the system like growth form, endemic status, red list status, etc. Taxon query (fig. 3) provides basic information for

Table 1: Map layers provided by the WebGIS application. Overlay layers can provide additional information using query tool. On contrary, base layers present graphic background information only

Base Layer
Landsat Satellite Image
Orography
Physical Map
Political Map
Topographic Map

Overlays
Province (Aimag)
Geobotanical Units
Habitat Photos
District (Sum)
Records
Rivers and Lakes
Several Track Logs
UTM and Geographic Grid
Vegetation Zones

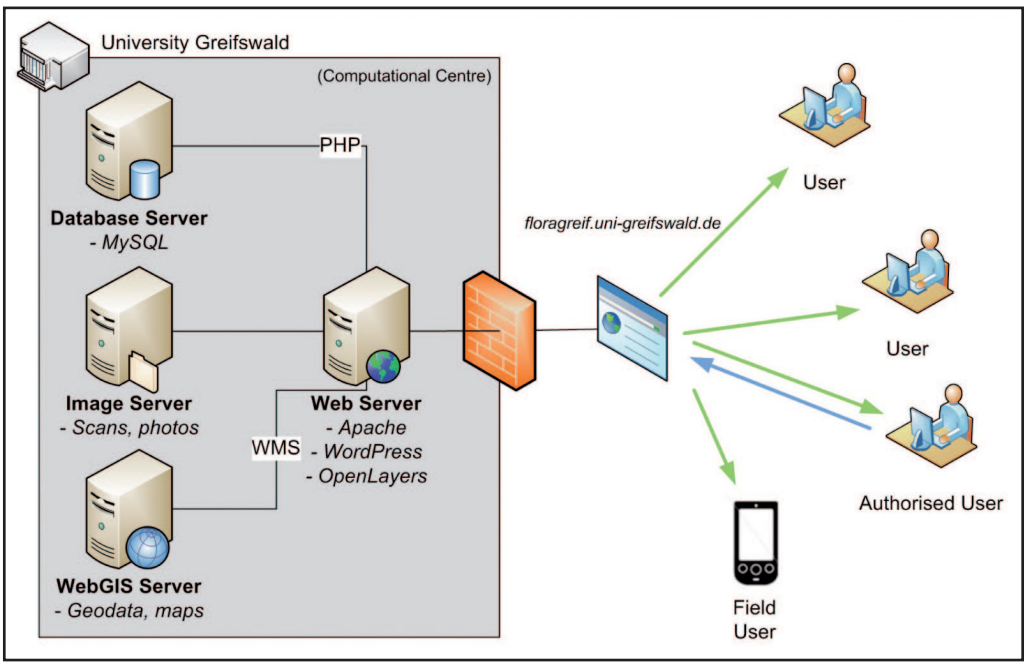

Fig. 2: Components and connectivity of the FloraGREIF information system. Specialised servers within the firewall-protected internal network of Greifswald University handle data and create dynamic web pages accessible by users. Users from outside interact with the information system via this web portal. Authorised users are able to add and modify data directly.

a species, like diagnostic characters, growth form, conservation status, habitat preferences, and distribution integrated as a WebGIS map. Record query (fig. 4) provides collection and repository information for a record or specimen, including habitat information similar to that found on the label of a herbarium sheet, and locality which is also shown on WebGIS map if coordinates are available. From all result pages, connected images can be accessed. Entire information provided is crosslinked internally and accessible via mouse click on displayed links. Herbarium scans are provided in full resolution. Usually, one scan consumes more than 200 Megabytes of disk space, an amount which cannot be loaded via web browser in reasonable time. Therefore the scanned

Fig. 3: Result of a taxon query. Beside taxonomic facts, species description, similar species, and other, species distribution is presented via the WebGIS application.

files are processed by a specific software (Zoomify Express 2012). It provides tiles of the scan with high quality at short loading time via Flash plugin.

The WebGIS application provides several functionalities. Various map layers can be displayed; some of them are semi-transparent and offer combinations of additional geographical information. Via mouse click attribute information are available for a map location and can be retrieved (fig. 5). In addition, geographic coordinates can be obtained for any point (hold control key and left mouse click). This feature facilitates the georeferencing of sampling places when no geographical coordinates are available.

Currently (March 2012), about 5800 records are represented by locality data and either digitalised herbarium specimens (1339), field images of the plants, including close-ups (733), or a combination of both (see fig. 6 for details).

Fig. 4: Result of a record query. Scans of herbarium sheets and other images are shown as thumbnails and can be displayed in full size via the Flash Zoomify plugin.

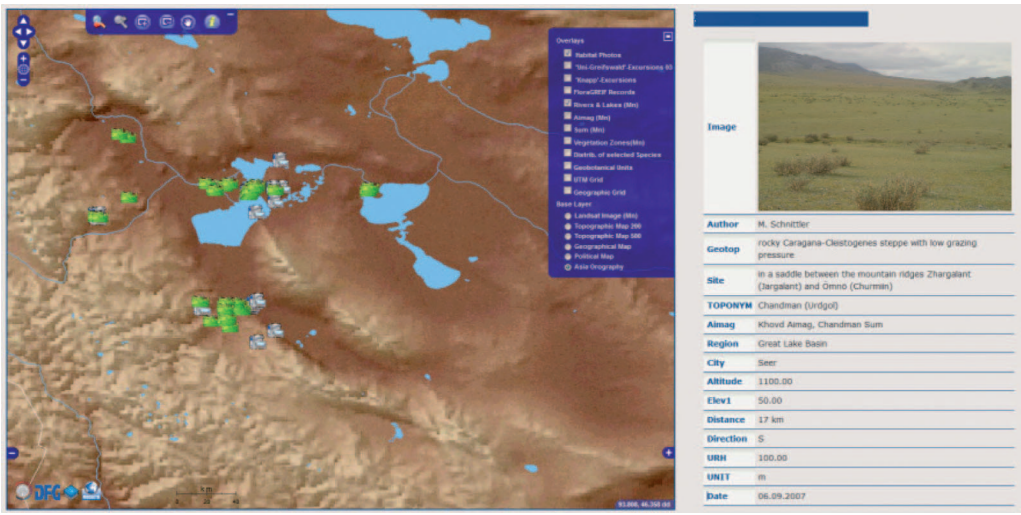

Fig. 5: Example of an interactive map created by the WebGIS application. A habitat image is underlaid by an elevation map (left side). Via mouse click at the map a request was sent to the GIS Server. The result, here including a habitat photo, is shown to the right.

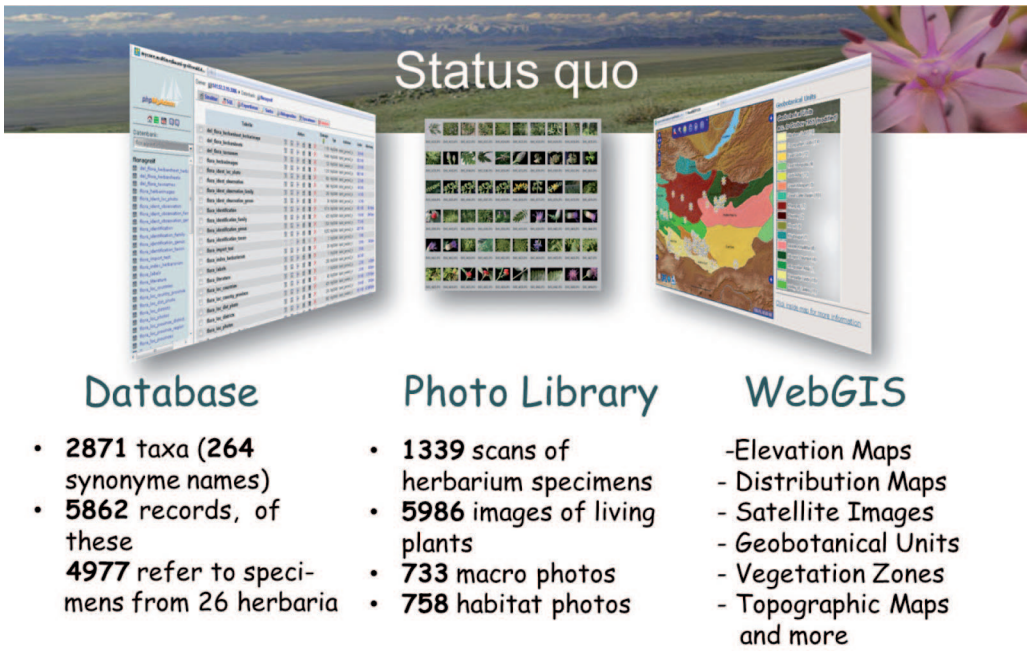

Fig. 6: Current content of the FloraGREIF information system; status March 2012.

Outlook

Beside the extension of the database, further activities will focus on three topics.

First, we are developing the information system into an expert system, creating a computer-aided identification tool based on easily accessible plant characteristics. Our aim is to lead the users to the level of a genus or a species group, where they can easily compare with the specimen in question, using digitalised specimen information. Since printed determination books are often out of print, require special knowledge of botanical terms, or keys ask for characteristics that cannot be easily observed in the field, a computer aided key would reach significantly more users. In addition, we will provide upload functionality for records and images of interested users. After that, the information is open to be revised by experts.

Second, the WebGIS application will be improved. A geographical dictionary (gazetteer) is under development. It connects regional toponyms with geographic coordinates and will allow users to search places by name and display their location at the map. Through OGC web feature service (WFS) an attribute filter will provide the option to display records or other vector data which have specific attribute values.

Third, the system should be translated from HTML4 standard to a level that conforms to HTML5. This will improve compatibility and functionality of the FloraGREIF system on mobile devices such as WebPads, Handhelds or SmartPhones.

Acknowledgment

We are grateful to all partners, participants, and friends supporting the FloraGREIF project. Without them the establishment of FloraGREIF would not have been possible. For access to their collections, we owe thanks to the following herbaria with comprehensive collections on the Mongolian Flora: Halle (U. Braun, E. Jäger), Gatersleben (K. Pistrick), Jena (A. Zündorf, K.-F. Günther). For revision of specimen, we are grateful to the following experts: H. Freitag (*Ephedra*, Suaeda, Kassel), N. Kilian (Liguliflorae, Berlin), M. Maier-Stolte (*Ephedra*, Kassel), R. Wisskirchen (Polygonaceae, Bonn), N. Friesen (*Allium*, Osnabrück), P. Kuss (Pedicularis, Zürich), E. Raab-Straube (*Saussurea*, Berlin), D. Podlech (Fabaceae, München), P. Hanelt (*Papaveraceae*, Gatersleben), H. Scholz (*Eragrostis*, Berlin), K.-F. Günther (Apiaceae, Jena). During the project we cooperated with the Mongolian Academy of Sciences, Ulaanbaatar (M. Urgamal, I. Tuvshintogtokh), the University of Khovd (D. Oyuunchimeg), and the Botanical Garden and Botanical Museum Berlin (W.G. Berendsohn). For technical assistance we finally wish to thank E. Dauber and K. Münse, Greifswald.

Literature

CBIF (Canadian Biodiversity Information Facility) (2006): Canadian biodiversity information facility, Government of Canada. – Accessed December 2011; http://www.cbif.gc.ca.

GREUTER, W.; BRUMMIT, R.K.; FARR, E.; KILIAN, N.; KIRK, P.M.; SILVA, P.C. (1993): NCU-3. Names in current use for extant plant genera. – Koeltz Scientific Books, Regnum Vegetabile. vol. **129**: 1–1464.

GRUBOV, V.I. (1955): Konspekt flory Mongolskoj Narodnoj Respubliki. – Trudy Mong. Kommiss. 67 (fett). – Moskva – Leningrad.

GUBANOV, I.A. (1996): Konspekt flory Vneshney Mongolii (sosudistye rasteniya). – Valant. [Conspectus of flora of Outer Mongolia (vascular plants)].

GUBANOV, I.A. (1999): Dopolneniya i ispravleniya k “Konspektu Flory Vneshney Mongolii (Sosudistye Rasteniya). – Turczaninowia (3): 19–23. [Additions and corrections to the “Conspectus of flora of Outer Mongolia (vascular plants)”].

GUBANOV, I.A.; HILBIG, W. (1989): Entwicklung und Stand der Erforschung der Gefäßpflanzenflora der Mongolischen Volksrepublik. – Erforsch. biol. Ress. MVR **7**: 127–146 Halle/Saale.

HAUCK, M.H.; SOLONGO, Z.B. (2010): Flowers of Mongolia. – Telgte, Biermann; 325 pp.

HILBIG, W. (2006): Der Beitrag deutscher Botaniker an der Erforschung von Flora und Vegetation in der Mongolei. – Feddes Rep. **117**: 321–366.

IPNI (2012): The International Plant Names Index. – Accessed in July 2012; http://www.ipni.org.

Kew Gardens (2006): Royal Botanic Gardens. Herbarium digitisation. – Accessed July 2012; http://www.kew.org/data/herb_digitisation.html.

Mapserver.org (2012): Open source web mapping. – Accessed in July 2012; http://Mapserver.org.

Moskito-GIS (2012): Moskito-GIS GmbH. – Accessed in July 2012; http://www.moskito-gis.de.

Openlayers.org (2012): OpenLayers: Free Maps for the Web. – Accessed in July 2012; http://openlayers.org.

PAVLOV, D.S.; SHAGDARSUREN, O.; KAMELIN, R.V.; ULZIIKHUTAG, N. (2004): 35 years of the activity of the joint Russian-Mongolian complex biological expedition RAS and MAS. – Arid Ecosystems, vol. **10** (Iss. 24–25): 13–16.

RILKE, S.; NAJMI, U. (2011): FloraGREIF – Virtual Guide and Plant Data Base as a Practical Approach to the Flora of Mongolia. – Willdenowia **41** (2): 371–379.

USGS (2009): NAS – Non-indigenous Aquatic Species. – U.S. Department of the Interior and U.S. Geological Survey (USGS). – Accessed in December 2011; http://nas.er.usgs.gov.

ZEMMRICH, A.; SCHNITTLER, M.; HARTLEIB, J.; BUSCH, M.; BOBERTZ, B.; RILKE, S.; ZÖLITZ, R. (2013): FloraGREIF – An internet-based tool for biogeographical research in Mongolia. – Folia Geobotanica **48/1**, DOI: 10.1007/s12224-012-9149-6.

Zoomify Express (2012): Zoomable Web Images. – Accessed in July 2012; http://www.zoomify.com/express.htm.

Addresses:

Martin Schnittler
Sabrina Rilke
Anne Zemmrich
Susanne Starke
University of Greifswald
Institute of Botany and Landscape Ecology
Grimmer Straße 88
D-17489 Greifswald

Ulrike Najmi
University of Greifswald
Computer Centre
Felix-Hausdorff-Str. 12
D-17487 Greifswald

Jörg Hartleib*
Bernd Bobertz
Reinhard Zölitz
University of Greifswald
Institute of Geography and Geology
Friedrich-Ludwig-Jahn-Str. 16
D-17487 Greifswald

e-mail: hartleib@uni-greifswald.de

* Corresponding author