



„Vernetzung von Bilddaten im GBIF-D Verbund“

Tagungsort:

Botanischer Garten und Botanisches Museum Berlin-Dahlem
Königin-Luise-Str. 6-8, 14195 Berlin, Konferenzraum A105

Anwesend:

Carola Söhngen (DSMZ), Michael Martin (DSMZ), Dagmar Triebel (SNSB, IT Zentrum), Gregor Hagedorn (JKI), Alexander Kroupa (MfN), Lothar Menner (Senckenberg), Peter Grobe (ZFMK), Evelin Boos (SysTax), Maren Gleisberg (BGBM), Anton Güntsch (BGBM), Dominik Röpert (BGBM), Wolf-Henning Kusber (BGBM)

TOP 0 Begrüßung, Organisatorisches, Vorstellungsrunde

- Anton Güntsch gibt Überblick über die Aufgaben der IT-Kommission (wie im Antrag)
 - o Abstimmung und Abgrenzung der von den Knoteninstitutionen aufgebauten und angebotenen Services.
 - o Informationsaustausch über wichtige Entwicklungen bei den internationalen Initiativen (GBIF, LifeWatch, etc.) und Beurteilung derselben im deutschen Kontext
 - o Identifikation von möglichen gemeinsamen institutionsübergreifenden Entwicklungen
 - o Koordination von gemeinsamen Entwicklungen
 - o Erarbeitung von F&E-Themen für gemeinsame Projektanträge
 - o Koordination der öffentlichen Darstellung von Entwicklungen des GBIF-D Verbundes (Website, Konferenzen, etc.)
- Verabschiedung der Tagesordnung, bleibt wie zuvor zirkuliert.

TOP 1: Der zukünftige Bildserver des BGBM

Dominik Röpert stellt den Imageserver am BGBM vor (vgl. pdf der ppt anbei)

Diskussion:

- Anton Güntsch weist darauf hin, dass das primäre Interesse ist, dass die Knoten ihre Bilder selber verwalten und der Image Server lediglich der Visualisierung dient.
- Michael Martin regt an, in den Knoten nur Software zu verwenden (falls neue installiert wird), die mit dem Image-Server kompatibel ist.
- Peter Grobe wünscht sich Erfahrungsinput seitens BGBM für den Aufbau eines eigenen Imageserver-Systems am ZFMK.
- Erste Testläufe des Systems zur Anzeige von Bildern aus Heidelberg haben ergeben, dass der Caching-Prozess zeitaufwändig sein kann. Hier soll in weiteren Tests mit den GBIF-Knoten herausgefunden werden, bis zu welchem Datenvolumen ein Bildertransfer über das Internet sinnvoll ist. Denkbar wäre auch, einen jeweils initialen Transfer per Platte vorzunehmen und Aktualisierungen dann wieder per Internet.
- Das BGBM schreibt seit ca. einem Jahr Metadaten direkt in die Bilder und wünscht sich hier Standardisierung. Hierzu gibt es eine TDWG-Arbeitsgruppe, in der Metadatenelemente definiert und dokumentiert werden. Gregor Hagedorn gehört dieser Arbeitsgruppe an, die den Standard nun fast fertiggestellt hat.

- Welchen Vorteil hat es, Metadaten (Braunschweig) ans BGBM zu liefern? - Durchsuchbarkeit; Konkreter Aufbau der Suchanfrage kann zu einem späteren Zeitpunkt geklärt werden.
- Der Bildserver des BGBM ist auch in der Lage, Audio- und Video-Files zu speichern. Die jeweiligen Möglichkeiten für die Ausgabe müssen allerdings noch getestet werden.
- Metadaten sollten vor der Bildlieferung an das BGBM-System eingepflegt sein.
- Im Bildserver gibt es derzeit keine Versionskontrolle (z. B. im Fall von Lizenzänderungen), was problematisch bei Änderungen der jeweiligen Lizenzschemas werden könnte.

TOP 2 Bildverarbeitung bei den GBIF-D-Knoten [kurze Vorstellung jedes Teilnehmers]

- DSMZ Braunschweig, Carola Söhngen
 - Mikroskopische Aufnahmen
 - Digitalisierung der Myxobakterien-Sammlung derzeit problematisch
 - Keine zentrale Erfassung von Bilddaten
 - Speicherung in Braunschweig, Auslieferung via FSI
 - Geplant: DIA-Digitalisierung
 - Eigenes Portal, Bilddaten über FSI, Funktionalität scheint auszureichen
 - Frage: Kann man die Skala des Bildes mitgeben? DR: Ja und Messungen im Bild möglich wenn Umwandlungsfaktor bekannt
 - Volumen schwer abzuschätzen: Maximal 90000 Bilder, 6TB, über 3 bis 5 Jahre.
 - Ein Problem wäre Anschaffung einer zentralen Bildbearbeitungssoftware.
 - Grundsätzlich ist das DSMZ mit den Funktionalitäten des neuen BGBM Bildservers sehr zufrieden und möchte seine Bilder gerne über das System ausliefern.

Anton Güntsch weist darauf hin, dass der BGBM zwar in Einzelfällen das Hosting von Bilddatenbeständen übernehmen kann, dies aber als generellen Service aus Kapazitätsgründen nicht anbieten kann.

- SNSB München, Dagmar Triebel
 - Es gibt bereits einen einheitlichen Informationsfluss für Bilddaten, Bildserver, Webserver, links zu Datenbankserver, mit Archivfunktion im Rechenzentrum (Bandarchivierung). Die Bilder sind auch über das entsprechende Feld in ABCD an BioCASE/GBIF angebunden.
 - Aktuell beträgt das Gesamtvolumen an Bilddaten 13TB..
 - Insgesamt wird derzeit keine Notwendigkeit gesehen, das Berlin Bildsystem mitzubnutzen, da an den etablierten Informationsflüssen erste einmal nichts geändert werden soll.
- MfN Berlin, Alexander Kroupa
 - Verwenden Cumulus (derzeit nur hausintern genutzt und zugänglich).
 - Anbindung an Specify über die Katalognummer angebunden, Metadaten (welche?) werden in den Bildern gespeichert.
 - Die Bilder sind in einem geschützten Bereich gespeichert. Für die Auslieferung der Bilder durch den BGBM-Server muss ein Zugang zu diesem Bereich ermöglicht werden.
 - Derzeit ist sie Stelle eines Scan-Operators ausgeschrieben.
 - Geplant ist die Erfassung von Sammlungskatalogen (5-10.000 Seiten), die fast die ganze Sammlung abdecken.
 - Ebenfalls digitalisiert werden sollen Sammlungskästen mit Priorität auf besonders wertvolle Objekte (z.B. Prachtkäfer).

- Ebenfalls interessant für den Bildserver ist ein Bestand von digitalen Bildern von lebenden Objekten (ca. 10000).
 - GBIF Entomologie-Workshop am Vortag signalisierte, dass große Bereitschaft besteht, Bilddaten an GBIF zu liefern (z.B. Dresden, M. Nuss).
 - 20TB vorgesehen, derzeit im GB-Bereich, bei den Katalogen eher unklar
 - Überlegung, die handschriftlichen Sammlungskataloge (immer die gleiche Handschrift) über eine Citizen Science-Gruppe, die ehrenamtlich die Daten erfassen. Herr Güntsch weist auf Kooperation mit der FH Hannover im Bereich Bildverarbeitung hin.
 - Weiterhin ist ein Treffen mit einer Arbeitsgruppe der TU-Berlin zum Thema 3D-Digitalisierung geplant.
- Senckenberg Frankfurt, Lothar Menner
 - Bilder immer ein Problem wg. Lizenzen, kommerzieller Nutzen problematisch
 - In jedem Falle könnten ca. 1TB Herbarbögen (jeweils 200MB) über das BGBM-System ausgeliefert werden.
 - Auch hier wäre noch der beste Weg für den Datentransfer zu klären.
- Gregor Hagedorn
 - Stellt BioWikiFarm (www.biowikifarm.net/meta/Biowikifarm) vor, für das er sich Synergien mit den GBIF-D Bildsystemen wünscht.
 - In den „Offenen Naturführern“ sind zahlreiche Bilder (>100000) in weniger hoher Auflösung vorhanden, die an das Bildsystem angeschlossen werden könnten. Es handelt sich häufig um Beobachtungsdaten, allerdings fehlen meist die notwendigen Sammlungsdaten (z.B. Geo-Bezug).
- ZFMK Bonn, Peter Grobe
 - Die IT-Infrastruktur des ZFMK ist im Aufbau begriffen. Derzeit gibt es noch keine großflächige Sammlungsdigitalisierung.
 - Das ZFMK hätte Interesse daran, mit anderen Institutionen gemeinsame Strategien für die sichere Speicherung von Bilddaten zu implementieren.
 - Es gibt ebenfalls ein großes Interesse an der Entwicklung von Standards und Best Practices bei der Digitalisierung (z.B. für die Orientierung von zoologischen Objekten).
 - Ein großes Problem ist die Heterogenität der zu digitalisierenden Objekte.
 - Da das ZFMK für die Zukunft ein sehr großes Datenvolumen erwartet (>70TB), wird angestrebt, eine eigene Infrastruktur für Bilddaten im Web aufzubauen.

TOP 3 Abstimmungen zur weiteren Arbeit der IT-Kommission

Anton Güntsch drückt den Wunsch aus, dass die IT-Kommission fokussiert an bestimmten Themen arbeitet und versucht, im Rahmen der zeitlichen Möglichkeiten, Ergebnisse zu erzielen. Die Teilnehmer begrüßen dies.

- Kommunikation
 - Website: Technische Entwicklungen aus dem GBIF-D Verbund sollten auf gbif.de dargestellt werden.
 - Wiki: Sollte für konkrete Arbeiten eingesetzt werden.

- e-Mail-Liste: Wird benötigt. Dominik Röpert wird sich um die Einrichtung bei der FU kümmern
- e-Meetings (z.B. Skype) sollten genutzt werden, wenn es das jeweilige Thema zulässt und die Gruppe nicht zu groß wird.
- Mögliche Themen für zukünftige Treffen
 - Portalentwicklungen im GBIF-D Verbund
 - Digitalisierungstechniken und Workflows (hierzu wird von Triebel/Hagedorn ein Wiki eingerichtet). Ziel ist unter anderem ein Best Practice Handbook
 - Citizen Science Involvement
 - Standardisierungen bei der Digitalisierung und Bild-Metadaten
 - Lizenzen, IPR, Copyright. Gregor Hagedorn könnte hierzu einen ersten Vortrag anbieten.
 - Datenflüsse zu und in GBIF (international) und Identifikation von Fehlentwicklungen. Hierzu sollte ebenfalls ein Wiki zur Dokumentation eingesetzt werden.
 - Effiziente Erschließung von Textdaten (z.B. Kataloge)
 - Metadaten zu Objektgruppen (z.B. für Insektenkästen)

Es wurde beschlossen, dass das Thema „Portalentwicklung im GBIF-D Verbund“ Priorität hat, da es unter Umständen Einfluss auf die Arbeitspläne der Teilprojekte hat.

- **Termin und Ort für das nächste Treffen**

- Das nächste Treffen findet in Frankfurt im Januar oder Februar 2012 statt. Der BGBM wird einen Doodle-Poll dazu einrichten und Frankfurt bei der Organisation unterstützen.

- **Sprecher**

Anton Güntsch und Gregor Hagedorn werden als Sprecher und Stellvertretender Sprecher der IT-Kommission bestätigt.

TOP 4 Neue Projekte und Anträge mit GBIF-Bezug

Anton Güntsch stellt Projekte am BGBM vor:

- OpenUp (EU): Koordination BGBM, 1Mio. GBIF Multimedia-Objekte an EUROPEANA liefern
- VIBRANT (EU): NHM London koordiniert, Ausbau der Scratchpad-Infrastruktur und Schnittstellen zu anderen Plattformen
- BioVeL (EU): Universität Cardiff, Lifewatch spin-off, Workflows aus Services der Biodiversitätsforschung erstellen
- Annotationen (DFG LIS): Elektronische Annotation für Sammlungsobjekte
- reBinD (DFG LIS): Rettung bedrohter Primärdatenbestände

Dagmar Triebel/ Gregor Hagedorn berichten:

- Das Diversity Mobile-Projekt (DFG LIS) wird weitergehen; Wiki unter www.diversitymobile.net.

16:00 Ende des Treffens

Für das Protokoll: Maren Gleisberg, Anton Güntsch