

„Nutzung von GBIF-Services“

Treffen der GBIF-D IT-Kommission am 13.09.2012 von 11.00 – 16.00 Uhr



Tagungsort:

Leibniz-Institut DSMZ-Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Inhoffenstraße 7 B, 38124 Braunschweig

Teilnehmer:

Carola Söhngen (DSMZ), Dagmar Triebel (SNSB), Gregor Hagedorn (JKI/SNSB), Wolfgang Reichert (SNSB), Alexander Kroupa (MfN), Lothar Menner (Senckenberg), Martin Müller (Senckenberg), Maren Gleisberg (BGBM), Anton Güntsch (BGBM), Peter Grobe (ZFMK)

Entschuldigt: Wolfram Fritsch (MfN), Michael Martin (DSMZ)

TOP 0 Begrüßung, Organisatorisches, Verabschiedung der Tagesordnung

Anton Güntsch begrüßt die Anwesenden Teilnehmer_Innen. Um allen Mitgliedern der IT-Kommission die Einblicke in die nachfolgenden Themen ermöglichen zu können, wird folgender Beschluss gefasst: Beginn der Besprechung wenn Teams aus München und Bonn auch vor Ort sind (es gab Probleme im ÖPNV). Um Zeit zu sparen erfolgt **TOP 4 – Neue Projekte und Anträge mit GBIF-Bezug** schriftlich.

TOP 1: Vorträge & Diskussion

Vorstellung GIS-Komponenten/Georeferenzierung in SeSAM, Senckenberg (L. Menner)

[Siehe ppt IT-Kommission_GIS.ppt \(2MB\)](#)

Nutzung von GBIF-Services in Workflow-Umgebungen, z. B. BioVel (A. Güntsch)

Anton Güntsch stellt eine Workflow-Entwicklung des BioVeL-Projektes (EU, 7. Rahmenprogramm, <http://www.biovel.eu/>) zur Bereinigung von Beleg- und Beobachtungsdaten vor. Der Workflow wurde auf der Taverna-Plattform (Uni Manchester, www.taverna.org.uk/) entwickelt und integriert 11 verschiedenen verteilte Services in einer konsistenten Oberfläche. Integrierte Services sind zum Beispiel GBIF Name Parser, GBIF Checklist Bank, GBIF Occurrences, Catalogue of Life, BioSTIF, Google Refine und InfoXY. In einer Life-Demonstration stellt Herr Güntsch vor, wie mit Hilfe des Workflows Occurrence Daten zu einer taxonomischen Gruppe (hier Schwarzwurzel) zusammengestellt und gereinigt werden können.

Anton Güntsch bietet an, zu dem Thema einen Workshop zu organisieren.

DWB GIS Editor (<http://www.diversityworkbench.net/Portal/DiversityGisEditor>) (W. Reichert)

[Siehe pdf Diversity_GIS_Editor.pdf \(3MB\)](#)

Mittagsimbiss

TOP 2: Weiterentwicklung ABCD (A. Güntsch)

Anton Güntsch stellt die aktuellen Weiterentwicklungen des ABCD-Schemas vor, wie sie derzeit am BGBM vorangebracht werden (s. pptx). Version 2.06 ist seit ca. sieben Jahren in Verwendung und hatte mehrere hundert Installationen weltweit. Über die Jahre haben sich eine Reihe von Korrektur- und Verbesserungswünschen angesammelt.

Für die Zukunft sind zwei grundsätzliche Strukturen denkbar:

- 1) ABCD-Extensions werden von Zeit zu Zeit in den Kern gezogen (das ist der ursprüngliche Ansatz)
- 2) Extensions werden beibehalten. In diesem Falle müssen aber Konzepte, die auch im Kern vorhanden sind aufgelöst werden.

Der BGBM verfolgt jetzt die 2. Strategie.

Stand der Arbeiten: ABCD Inventur (Wiki, BGBM intern) abgeschlossen, möglicherweise ist dies für externe zu öffnen. Nächster Schritt: Zusammenstellung der Feature Requests, Abarbeitung Bug-Liste, Minor Release 2.06d noch in 2012, dazu wird keine ABCD Neuinstallation erforderlich sein.

Dokumentation: Mittelfristig soll die ABCD-Dokumentation aus dem Schema herausgezogen werden. Ziel ist es, dass ABCD-Konzepte eine maschinenlesbare (RDF) Repräsentation bekommen. Es wäre zum Beispiel möglich, hierfür Semantic Media Wiki einzusetzen.

Einschub Gregor Hagedorn: Vibrant/GBIF arbeitet derzeit an neuem Semantic Tool

Frage Peter Grobe: Steht bei Semantic eine Ontologie dahinter? Wenn ja, wer macht diese? Dazu Anton Güntsch: Ja, und keiner macht die O. derzeit. Es sollen in einem ersten Schritt zunächst die Konzepte maschinenlesbar zur Verfügung gestellt werden.

Anmerkung Gregor Hagedorn: GBIF entwickelt eine Plattform, die als Repository für Terminologien (auch mehrsprachig) dient, die auch maschinenlesbar sein soll, Semantic Media Wiki soll zur Bearbeitung der Vokabulare verwendet werden.

Frage Peter Grobe: Wie wird ein neuer Standard in ABCD etabliert? Gibt's da schon einen Prozess? – Anton Güntsch: ABCD ist ein TDWG Standard, d.h. der ausgearbeitete Vorschlag wird weitergegeben an Executive Committee, von dort an Experten, dann kommt eine finale Abstimmung (es gibt interne/externe Review Phase). Für BiNHum wäre eine Extension sinnvoll.

[Siehe pptx ABCD_Weriternetwicklung.pptx](#)

TOP 3: Digitalisierungs-Wiki (Stand der Arbeiten / Perspektiven) & Website (M. Gleisberg)

http://wiki.gbif.de/web/Digitalisierungsstationen_nach_Institutionen

Derzeit zeigt sich ein heterogenes Bild an Informationen, da die einzelnen Bereiche nicht vollständig ausgefüllt sind. Ferner herrschen Unterschiede in der Informationstiefe, z.T. wird mit Bildmaterial gearbeitet und es fällt ein vergleichsweise geringer Anteil an Workflowbeschreibungen auf.

Frage in die Runde: Ist es möglich die Lücken bis zur Vorstellung des Wikis am 10.10. (Knotentreffen) zu stopfen?

Als ein Ergebnis der Website-AG, die sich im Mai am BGBM zu einem Brainstorming getroffen hat, wird die überarbeitete Startseite von GBIF.de durch Maren Gleisberg vorgestellt:

<http://www.bgbm.org/temp/gbif/default.html> Die Seite ist zurzeit noch nicht live geschaltet, es wird parallel noch eine Grundüberarbeitung der Navigation geben.

Konstruktive Rückmeldungen und Anmerkungen zur Website nimmt Maren Gleisberg gerne an.

TOP 4: Neue Projekte und Anträge mit GBIF-Bezug

Carola Söhngen, DSMZ:

Im Rahmen der DFG-geförderten Initiativen GFBio und iDiv plant die DSMZ die weitere, zunächst nationale Anbindung an die geplanten zentralen Plattformen. Des Weiteren steht eine Standardisierung des Datenaustauschs im Fokus. Zusätzlich ist auch im Rahmen dieser Projekte eine Fortführung der Mobilisierung vorhandener in house Daten geplant. Die Zukunft der in GBIF erfolgreich betriebenen Text und Data Mining und Semantic Web-Projekte ist außerhalb von GBIF jedoch völlig unklar

Anton Güntsch, BGBM:

Stand beim **DFG-Antrag GFBio (Datenzentrum)**: Das Projekt ist durch die Begutachtung gegangen. Die Gutachter haben hervorgehoben, dass sie die Ziele des Antrages voll unterstützen. Kritik wurde an (aus Gutachtersicht) unzureichender Darstellung der Methodik geübt. Im Hearing wurde dann beschlossen, dass der Arbeitsplan überarbeitet und für eine Projektlaufzeit 18 + 36 Monate ausgelegt werden soll. Neuer Abgabetermin ist Dezember 2012.

Im Projekt **BiNHum** ist der BGBM zuständig für das WP 3 "Netzwerkentwicklung und Datenqualität", (Koordination: Güntsch/Berendsohn)

Provider-Software Konfiguration bei den Partnern, Entwicklung einer ABCD-Erweiterung für BiNHum, Aufbau einer BiNHum Index-Datenbank als Grundlage für das Portal, Bereitstellung von Datenqualitätsservices, Technische Dokumentation im Projekt-Wiki.

Dagmar Triebel, SNSB:

DFG-Projekt: **Integriertes Datenmanagement für die digitale Erschließung zoologischer und paläontologischer Sammlungsbestände (IDES)** - Start: Juli 2012

Projektpartner und Verantwortliche: IT-Zentrum der Staatlichen Naturwissenschaftlichen Sammlungen Bayerns (Dr. Dagmar Triebel, Koordination des Projektes); Jura-Museum Eichstätt (Dr. Martina Kölbl-Ebert); Bayerische Staatssammlung für Paläontologie und Geologie (Dr. Markus Moser und Dr. Oliver Rauhut); Zoologische Staatssammlung München (Dipl.-Biol. Dirk Neumann und Dr. Ulrich Schliewen); Staatssammlung für Anthropologie und Paläoanatomie München (Dr. Henriette Obermaier);

Datenkuratorische Aufgaben, Projektbetreuung: Ausbau Diversity GIS editor, DWB offline Umgebungen, Replikation; Schnittstellen zu internationalen Portalen GBIF, FishBase and FishNet 2.; Importroutinen

Ziel bei Projektende: rund 50.000 Datensätze (Fische rezent und fossil) für GBIF-D

DFG-Projekt: **Erschließung/Aufbereitung vorhandener digitaler Objektdaten, Anpassung etablierter Datenbanksysteme und Entwicklung eines Datenportals. Projekt: Biodiversitäts-Netzwerk des Humboldt-Rings BiNHum** - Start September 2012

IT-Zentrum der Staatlichen Naturwissenschaftlichen Sammlungen (Dr. Dagmar Triebel, Koordination WP 4) WP 4: Ausbau Diversity Workbench als Virtuelle Arbeitsumgebung für die Sammlungen des Humboldt-Rings

Installation von DWB-Umgebungen an 3-5 Sammlungen; Ausbau der Managementtools, Datenbanken und Clients z. B. DiversitySamplingPlots und DiversityScientificTerms; Ausbau von Importroutinen; Ausbau von DWB Updatemechanismen und Registries; Exportschnittstellen z.B. zu IMDAS; Optimierung der xml-Exporten aus DWB für GBIF und BiNHum-Portal; Anbindung von Webservices; Projektende: verbesserte IT-Infrastruktur am SMNK, SNSB und ZFMK u. a.

Alexander Kroupa, MfN:

Ein neues Projekt mit GBIF Bezug ist **EU BON**, bei welchem GBIF als Projektpartner eingebunden ist. Drei weitere Projektanträge laufen über **i4Life**. Dabei geht es um neue „Global Species Databases“ für den Catalogue of Life. Hier besteht allerdings nur ein indirekter Bezug zu GBIF, da der CoL bei GBIF als taxonomischer Backbone verwendet wird.

14:25 TOP 5: Abstimmungen zur weiteren Arbeit der IT-Kommission

Der Bericht IT-Kommission auf Knotenkoordinationentreffen am 10.10.2012 (MfN, Berlin) wird von Dagmar Triebel in Vertretung für Anton Güntsch vorgenommen. Neben einer Übersicht zu den heute behandelten „GBIF-Services“ wird das Wiki vorgestellt werden.

Thema, Termin und Ort für das nächste Treffen werden festgelegt:

Thema: „Datenflüsse zu und in GBIF (international) und Identifikation von Fehlentwicklungen“

Anton Güntsch recherchiert potenzielle (internationale) Gäste

Termin: 6. März 2013, SNSB München, vrsl. wieder 11:00 bis 16:00 Uhr

TOP 6: Sonstiges

Kein Sonstiges

15:00 Ende des Treffens

15:00 – 16.00 Optional: Führung durch die DSMZ

Für das Protokoll: Maren Gleisberg, Anton Güntsch