

Schlussbericht des EFRE-Projektes

**“Effiziente Arbeitsabläufe und innovative Methoden
zur Erschließung und dauerhaften
Verfügbarmachung objektreicher
Spezielsammlungen am Beispiel der
entomologischen Sammlung des Museum für
Naturkunde Berlin”**

Projektnummer Inno 02-2013000365

Laufzeit 01.02.2013 - 30.09.2015

1	Einleitung	2
2	Personalentwicklung	2
3	Digitalisierungsstationen	4
3.1	Insektenkastenscanner	4
3.2	Einzeltiere	5
3.3	45° Schnellerfassungsmethode	7
3.4	ZooSphere	8
4	Workflows	9
4.1	QR-Codes.....	9
4.2	Digitalisierungs-Datenbank	11
5	Erfolgsindikatoren	12
5.1	Insektenkästen.....	12
5.2	Einzeltierdigitalisierung	12
5.3	Webseiten.....	13
5.4	Ausgerichtete Tagungen und Workshops	14
5.5	Präsentationen auf Tagungen.....	15
5.6	Teilnahme an Veranstaltungen	17
5.7	Publikationen	17
6	Domänenübergreifende Publikation der Digitalisate.....	18
6.1	Europeana Portal.....	18
6.2	Deutsche Digitale Bibliothek	19
7	Feedback.....	19
8	Ausblick	20
9	Zusammenfassung	22
10	Anhang.....	23
10.1	Abstracts des internationalen Workshops zur Massendigitalisierung	23
10.2	Abstracts der Konferenz Langzeitzugriff auf Sammlungs- und Multimediadaten ...	39

1 Einleitung

Das Museum für Naturkunde (MfN) ist das größte naturhistorische Museum Deutschlands sowie das Zentrum zur Erforschung der Vielfalt der Natur und ihrer Entwicklungsgeschichte. Mit seinen Sammlungen, zu denen weit über 30 Millionen Objekte zählen, und mit seinen Ausstellungen zählt das Museum für Naturkunde zu den bedeutendsten naturwissenschaftlichen Sammlungen der Welt.

Bisher wurde nur ein kleiner Teil der Sammlungsobjekte am MfN digitalisiert und in Datenbanken erfasst. Bei 30 Millionen Objekten ist dies nicht überraschend und gerade die entomologische Sammlung, mit vielen kleinen bis sehr kleinen Objekten, wurde bis jetzt nur marginal digital erfasst. Gerade die Digitalisierung solch umfangreicher Sammlungen mit kleinen, räumlich komplexen Objekten stellt eine der größten Herausforderungen bei der digitalen Erschließung und Verfügbarmachung naturkundlicher Sammlungen und generell des kulturellen Erbes dar. Der Umfang der entomologischen Sammlung des MfN, mit Sammlungsobjekten von wenigen Millimetern bis zu einigen Zentimetern, beziffert sich auf ca. 15 Millionen Sammlungsobjekte und umfasst damit die Hälfte aller Objekte.

Im Projekt sollen am Beispiel der entomologischen Sammlung des MfN 1) effiziente Arbeitsabläufe zur Massendigitalisierung auf dem Niveau von Sammlungseinheiten geschaffen, 2) innovative Methoden zur effizienten Erschließung der räumlichen Komponente kleiner Objekte entwickelt, 3) die Vernetzung von Einzelprojekten und Initiativen zum Thema Langzeitarchivierung unterstützt und der Erfahrungsaustausch und die Strategiebildung auf diesem Gebiet beschleunigt werden.

Projektziele sind 1) die Bereitstellung von Empfehlungen zu Workflows und Kostenkalkulationen zur Massendigitalisierung objektreicher Sammlungen sowie zu Methoden räumlicher Digitalisierung kleiner komplexer Objekte, 2) die digitale Erschließung und Verfügbarmachung von Teilen der Insektensammlung des MfN über Europeana und eines neu zu entwickelnden digitalen Fensters in die Sammlungen und 3) die Etablierung eines interdisziplinären Netzwerkes Berliner Institutionen zur Langzeitarchivierung.

2 Personalentwicklung

In der Projektlaufzeit (32 Monate) waren insgesamt 19 Mitarbeiter in unterschiedlicher Länge und in unterschiedlichen Zeiträumen im Projekt angestellt. Über die gesamte Projektlaufzeit waren neben dem Projektleiter, ein Visualisierer und ein IT-Entwickler angestellt. Für 1,5 Jahre konnten zwei Taxonomen angestellt werden, die direkt in den Sammlungen arbeiteten und Sammlungsteile für die Digitalisierung aufbereiteten. Der größte Teil der Mitarbeiter war allerdings direkt mit der Digitalisierung der entomologischen Sammlungen beschäftigt. Dies umfasste das Scannen der Insektenkästen, das Fotografieren von Einzeltieren und die Digitalisierung von Insekten, mit zugehörigen Etiketten, in einem Winkel von 45°.

Ein weiterer Mitarbeiter hatte die Aufgabe eine 3D-Digitalisierungsmethode zur qualitativ hochwertigen Erfassung von genadelten Insekten zu entwickeln. In Zusammenarbeit mit zwei Technikern des MfN konnte bis Projektende ein Tool entwickelt werden, mit welchem es gelang, die ersten Objekte in 3D zu digitalisieren und über eine eigens für diese Objekte entwickelte Webseite zu veröffentlichen.

In der zweiten Projekthälfte wurde eine Mitarbeiterin angestellt, um die Anbindung der generierten Bilddateien an die Deutsche Digitale Bibliothek und Europeana umzusetzen.

Mehrere Praktikanten wurden während des Projektes betreut und konnten einen Einblick in die Digitalisierungsaktivitäten des MfN erhalten.

Auch studentische Hilfskräfte wurden erfolgreich für die Digitalisierung gewonnen. Von den fünf angestellten studentischen Hilfskräften ist es uns gelungen zwei, nach Beendigung ihres Masterstudiengangs, als technische Mitarbeiter anzustellen und sie bis zum Projektende zu beschäftigen.

Aufgrund der hohen Zahl an Mitarbeitern war die Fluktuation besonders hoch (s. Abb. 1). Zwei studentische Hilfskräfte beendeten ihren Vertrag bereits vorzeitig und zwei technische Mitarbeiter schieden ebenfalls vorzeitig aus dem Projekt aus. Hinzu kam ein Mitarbeiter der für mehrere Monate Elternzeit seine Arbeitszeit reduzierte. Weitere Mitarbeiter beantragten eine Reduzierung ihrer Arbeitszeit und somit hatten bis zu 8 Mitarbeiter pro Monat eine Teilzeitstelle von 75% oder weniger. Zusammen mit dem verzögerten Beginn des Projektes führte dies dazu, dass immer wieder neue Stellen für Mitarbeiter ausgeschrieben werden mussten, um die verringerten Arbeitszeiten zu kompensieren. Das führte gerade im letzten Projektjahr zur Einstellung mehrere neuer Mitarbeiter. In den letzten Projektmonaten waren 15 Mitarbeiter und eine Praktikantin im Projekt tätig (s. Abb. 2).

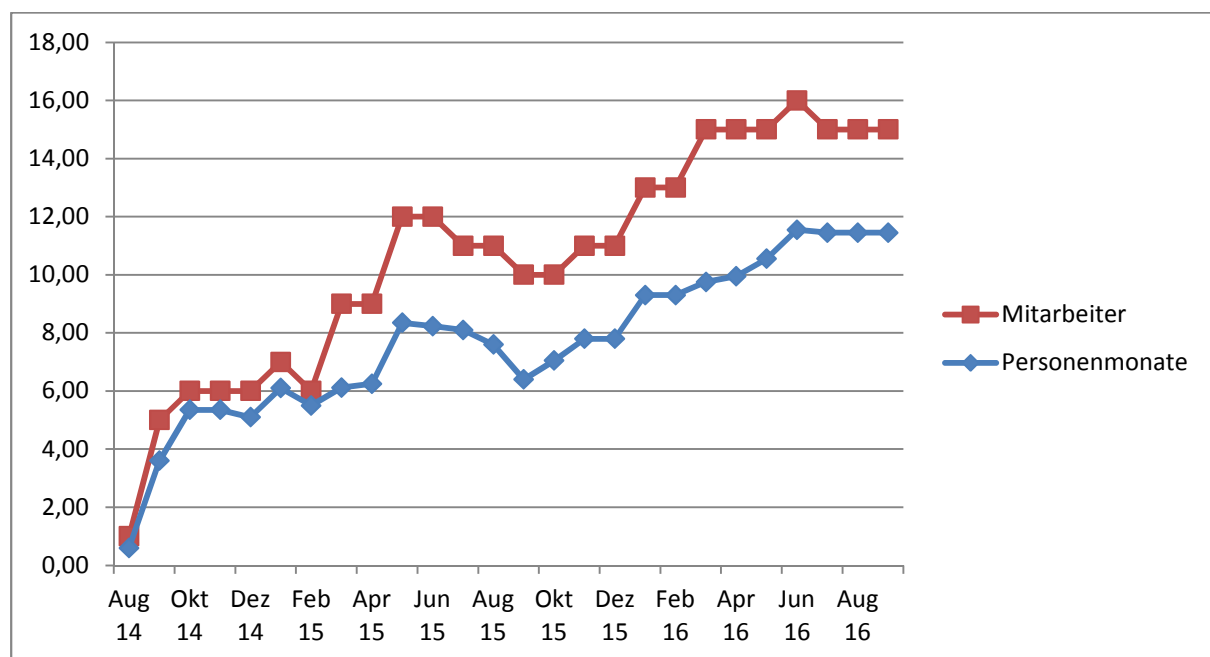


Abb. 1: Verlauf der Anzahl angestellter Projektmitarbeiter und ihrer Arbeitszeit, in Personenmonaten.



Abb. 2: Mitarbeiterfoto vom 28.09.2015. Von links nach rechts: Günther Korb, Alexander Kroupa, Birgit Jaenicke, Mélanie Turiault, Christina Kuhlmann, Sebastian Demtröder, Bernhard Schurian, Mirko Lange, Nina Kaiser, Karina Georgi, Kathrin Fahrngruber, Felix Maier. Fehlende Mitarbeiter: Falko Glöckler, Sascha Hentschel, Martin Pluta und Janetta Stroutchenkov.

3 Digitalisierungsstationen

3.1 Insektenkastenscanner

Der Scanner **SatScan®**-Collection der Firma Smartdrive wurde eigens zur Digitalisierung von großen Insektenkästen am Museum für Naturkunde Berlin angeschafft (s. Abb. 3).

Bei diesem System fährt eine Kamera mithilfe eines Gestänges über den Insektenkasten und erstellt senkrecht von oben 256 nur geringfügig überlappende Bilder. Diese Bilder werden anschließend von einem Programm (Analyse 64) zu einem einzigen Bild zusammengefügt – dieser Prozess wird als „stitchen“ bezeichnet. Die zusammengesetzten Bilder erreichen eine Größe von 400 MB mit 270 MegaPixel (s. Abb. 4). Ein einziger digitaler Insektenkasten kann bei einer Auflösung von 200 dpi auf eine Größe von 2,3 mal 1,8 m gedruckt werden.

Die Insektenkästen werden nach Sammlungen unterteilt und den einzelnen, entomologischen Kustodien des Museums für Naturkunde zugeordnet, aus denen sie stammen. Jeder Insektenkasten und jede Art in einem Insektenkasten bekommen eindeutige QR-Codes zugeordnet und werden in einer Datenbank erfasst (s. Kapitel 4.1).



Abb. 3: Blick in den Insektenkastenscanner mit einem Insektenkasten von Wegwespen (Pompilidae).



Abb. 4: Drei Beispiele digitaler Insektenkästen. Links: Kleinlibellen (Zygoptera); Mitte: Zikaden (Hemiptera); Rechts: Laufkäfer (Carabidae).

Im Internet sind alle gescannten Insektenkästen über das Kastenportal auf der EoS-Projekt-Webseite¹ für jeden frei zugänglich. Damit die großen hochauflösenden Bilder angezeigt werden können, nutzen wir den „krpano Viewer“². Diese Technologie ermöglicht ein schnelles Laden des hochaufgelösten Bildes und ein übergangloses Hineinzoomen in das Bild.

Experten können bei bestimmten Tiergruppen, wie z. B. den Libellen, wichtige Bestimmungsmerkmale im digitalen Insektenkasten erkennen und bereits auf diesen Aufnahmen Arten unterscheiden und Falschzuordnungen erkennen.

3.2 Einzeltiere

Für die **Specimen Digitalisierung** werden einzelne, einzigartige und besonders wichtige Tiere ausgewählt. In den meisten Fällen handelt es sich um primäre oder sekundäre Typen, welche in der taxonomischen Forschung eine bedeutende Rolle spielen.

¹ <http://eos.naturkundemuseum-berlin.de/>

² <http://krpano.com/>

Bevor die Tiere digitalisiert werden können, müssen die Etiketten von der Nadel genommen und neben dem Tier auf eine erhöhte Fläche in genau der vorgefundenen Reihenfolge ausgebreitet werden. Zu den bereits vorhandenen Etiketten wird ein Label mit einem Einzeltier-QR-Code hinzugefügt, welches das Tier eindeutig mit seinen Metadaten in der Datenbank verbindet. Die Individuen werden in eine weiße Schachtel auf eine Graukarte gesteckt, so dass sie sich mit den Labeln in einer Ebene befinden. Erst dann kann das Tier zusammen mit seinen Etiketten, einem Maßstab und einem Farbchart fotografiert werden.

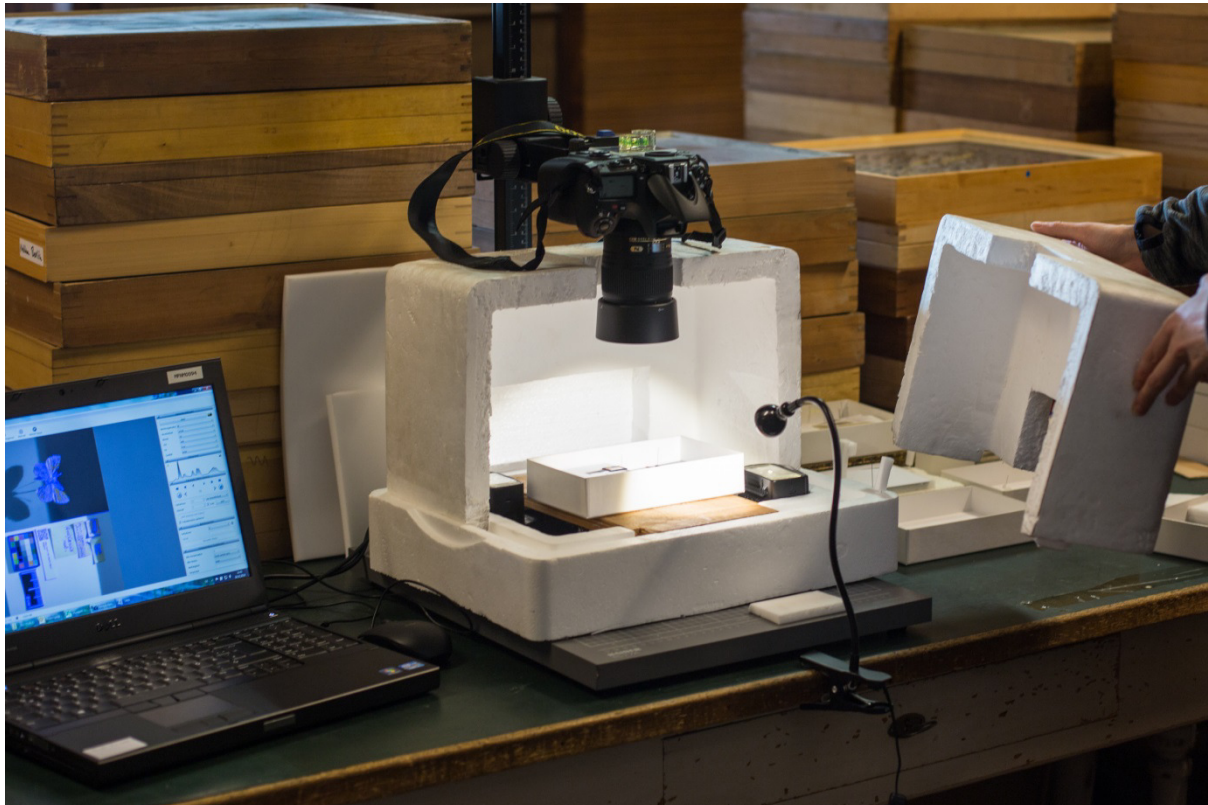


Abb. 5: Aufbau zur Digitalisierung von Einzelobjekten der entomologischen Sammlungen.

Für eine optimale, indirekte Beleuchtung wird die Schachtel mit dem Tier in eine Styroporbox gelegt, in der sich zwei Blitze befinden. Die Kamera wird mit einem Fernauslöser bedient, um Verwacklungen zu vermeiden (s. Abb. 5).

Zur Digitalisierung der Unterseite des Tieres wird das Tier in der Schachtel umgedreht, mithilfe von Knetmasse befestigt und fotografiert. Die Etiketten werden ebenfalls umgedreht, da sich teilweise auch auf der Unterseite Informationen befinden.

Im anschließenden Postprocessing werden die Bilder noch bearbeitet und beschnitten. Damit kann eine gleichmäßige Beleuchtung und ein ordentlicher Bildausschnitt gewährleistet werden.

Jedes aufgenommene Tier wird über die Datenbank dem Insektenkasten zugeordnet, in welchem es sich befindet. Somit können alle Einzeltiere eines Insektenkastens auf der Webseite unter dem entsprechenden Insektenkasten aufgelistet werden. Bisher umfasst das digitalisierte Material Typen zweier Schmetterlingsfamilien, HesperIIDae (Dickkopffalter) und Lycaenidae (Bläulinge), sowie Typusmaterial der Libellensammlung (Odonata).



Abb. 6: Drei Beispiele digitalisierter Typen der Schmetterlingsammlung (Hesperiidae – Dickkopffalter). Links: *Hasora borneensis* (Elwes & Edwards, 1897); Mitte: *Tharides henricus* (Staudinger, 1876); Rechts: *Kedestes chacooides* Gaede, 1916.

3.3 45° Schnellerfassungsmethode

Mit einem **Dino Lite®** Digitalmikroskop werden Einzeltiere in einem 45°-Winkel fotografiert, um die Informationen zu digitalisieren, die sich auf mehreren Etiketten unter dem Tier befinden (s. Abb. 7). Ziel dieser Digitalisierungsstation besteht darin, die gesamten Informationen der Etiketten, zusammen mit dem Sammlungsobjekt, digital zu erfassen. Damit in einem nächsten Schritt diese Information, wie z. B. der Fundort, das Funddatum, der Sammler etc. vom digitalen Bild transkribiert und in einer Datenbank gespeichert werden kann.

Hier arbeiten zwei Mitarbeiter Hand in Hand: der Eine entnimmt die Tiere einzeln aus den Systemschachteln des Insektenkastens und steckt einen Platzhalter an dessen Stelle. Die Etiketten auf der Nadel werden so zurechtgerückt, dass sie aus einem 45°-Winkel gut zu lesen sind und das Tier in eine vorbereitete, kleine Schachtel auf eine Graukarte gesteckt. Ein Etikett mit einem Einzeltier-QR-Code wird auf dem Boden, so neben der Nadel platziert, dass es im 45°-Winkel ebenfalls gut lesbar ist. Der Zweite Mitarbeiter nimmt die so ausgestatteten Schachteln entgegen und stellt sie auf die Fotoposition, löst das Dino-Lite Digitalmikroskop aus und verbindet Foto, QR-Code und weitere Metadaten in der Datenbank. Dann wird die Schachtel wieder der Fotobox entnommen und dem ersten Mitarbeiter übergeben. Dieser faltet den Einzeltier-QR-Code und steckt ihn unter die bereits vorhandenen Etiketten des Tiers. Am Ende des Prozesses wird das Tier wieder an seinen ursprünglichen Platz im Insektenkasten gesteckt.



Abb. 7: Drei Beispiele digitalisierter Sammlungsobjekte, welche mit dem Dino Lite Digitalmikroskop aufgenommen wurden (*Sigara lateralis* (Leach)).

Auf diese Weise können in einer Stunde bis zu 45 Tiere von einem Zweier-Team digitalisiert werden. Trotz der Monotonie der Arbeit, erfordert dieser Arbeitsablauf höchste Konzentration, damit keine Tiere beschädigt werden, die Datenbankeinträge korrekt erfolgen und die Tiere an die richtige Stelle zurückgesteckt werden. Aus diesem Grund sollte

idealerweise nach einer halben Stunde, aber spätestens nach einer Stunde, das Team wechseln.

Tiere, bei denen Etiketten auch auf der Unterseite beschriftet sind, werden in Schachteln mit einem Spiegel fotografiert, so dass sowohl die Unter- als auch die Oberseiten der Etiketten auf dem gleichen Bild zu sehen sind.

Mithilfe dieser Schnellerfassungsmethode lassen sich viele Tiere und deren Etiketten in kurzer Zeit digitalisieren. Mit einer entsprechenden Datenerfassung (Transkription) könnten die gewonnenen und georeferenzierten Fundortdaten z. B. als Grundlage von Art-Verbreitungskarten genutzt werden oder auch Aufschluss über invasive Arten geben.

3.4 ZooSphere

Die ersten Entwicklungsschritte wurden in Zusammenarbeit mit der Freien Universität Berlin im Rahmen des TET-Projektes³ unternommen und dabei wurden digitale Objekte erstellt, welche sich in einer Ebene drehen lassen⁴. Das **ZooSphere** ist eine deutliche Weiterentwicklung dieses Ansatzes. Dies betrifft nicht nur die Hardware (s. Abb. 8), sondern vor allem die Software, welche es ermöglicht 3D sphärische Bilder kleiner Objekte (bisher ausschließlich von Insekten) automatisiert aufzunehmen. Die resultierenden ZooSphere-Objekte kann der Nutzer über eine Webanwendung betrachten und frei in alle Richtungen drehen.

Die Funktionsweise der ZooSphere-Hardware beruht auf zwei Motoren, welche das Objekt in zwei Achsen sowohl horizontal als auch vertikal um 360° drehen, während eine Kamera, die auf einem Stacking-Schlitten montiert ist, hochauflösende Bilder in unterschiedlicher Entfernung zum Objekt macht. Die Motoren, der Stacking-Schlitten und der Auslöser der Kamera werden von einem Computerprogramm gesteuert, so dass die Aufnahmen – rund 2.000 bis 6.000 Bilder pro Objekt – voll automatisch gemacht werden. Die Steuerungs-Software wurde im Rahmen des Projektes entwickelt.



Abb. 8: Blick in das ZooSphere-Tool mit dem Insekt in der Mitte und dem Kameraobjektiv im Hintergrund.

Durch ein anschließendes Stacking werden die Bilder einer Position zu einem einzigen Bild zusammengerechnet und es entstehen insgesamt zwischen 100 und 200 Bilder, je nachdem wie viele Positionen aufgenommen wurden. Jedes Bild einer Position zeigt das Objekt in

³ <http://didaktik.physik.fu-berlin.de/projekte/tet/>

⁴ <https://tetfolio.fu-berlin.de/Naturkunde/IBE/bieneIPAD/>

voller Tiefenschärfe. Für das Postprocessing wurde eine eigene Desktop-Anwendung, die ZooSphere Tools, entwickelt. Mithilfe dieser Software können z. B. die Bilder der einzelnen Positionen übereinander gelegt und einheitlich geschnitten werden oder eine automatische Umbenennung aller Dateinamen erfolgen. Über die ZooSphere-Webseite können alle ZooSphere-Objekte und die zugehörigen Metadaten, wie z. B. eine Karte mit dem Fundort des Objektes, taxonomische Klassifizierung oder auch eine Verbreitungskarte (ein Service welcher von GBIF⁵ zur Verfügung gestellt wird) der Art, abgerufen werden. Mit jedem aktuellen Browser kann der Nutzer zudem jedes ZooSphere-Objekt nach Bedarf drehen und vergrößern. Um dies zu ermöglichen wird das OpenSeadragon-Plugin verwendet⁶, ein webbasierter Viewer für hochauflösende zoombare Bilder und für die horizontale Drehung wurde eine Erweiterung entwickelt. Ohne dieses Plugin wäre eine flüssige Bewegung eines ZooSphere-Objektes nicht möglich, da jedes hochauflösende Bild erst komplett heruntergeladen werden müsste, bevor es angezeigt werden kann.



Abb. 9: Drei Beispiele mit ZooSphere digitalisierter Sammlungsobjekte. Links: *Onthophagus elegans* (Klug, 1932); Mitte: *Melipona paraguayaca* Strand, 1910; Rechts: *Eucorysses iris*.

Wissenschaftler und Laien auf der ganzen Welt wird mit dieser Entwicklung der Zugang zu kleinen Sammlungsobjekten des MfN ermöglicht. Somit müssen in Zukunft nur noch wenige der einzigartigen Typen verschickt und damit dem Risiko der Beschädigung oder des Verlustes ausgesetzt werden. Wissenschaftler selbst sparen sich die Kosten und den Aufwand einer langen Reise zu allen Museen mit relevantem Forschungsmaterial. Vorausgesetzt auch andere Museen werden in Zukunft die ZooSphere-Technologie einsetzen und ZooSphere-Objekte kostenfrei zur Verfügung stellen.

4 Workflows

4.1 QR-Codes

Bei der Digitalisierung von Sammlungsobjekten ist die Nachvollziehbarkeit der Verbindung zwischen den physischen Objekten und den digitalen Repräsentationen sehr wichtig. Das bedeutet konkret, dass mit einer Referenz auf dem Etikett eines Sammlungsobjektes ohne zeitaufwendige Recherchen auf die digitalen Objekte (Bilder, Texte, verknüpfte Informationen) zugegriffen werden konnte und umgekehrt. Es bedarf somit einer Schnittstelle vom Physischen zum Digitalen. Diese Herausforderung konnte in diesem Projekt sehr gut gelöst werden: Bei dem Digitalisierungsvorgang bekamen die physischen Objekte ein Etikett, das einen Quick-Response-Code (QR-Code) trägt, in dem der Identifikator enthalten ist. Ein QR-Code ist eine binäre 2D-Matrix und kann von Barcode- bzw. QR-Code-Scannern und den meisten Smartphones und Tablets problemlos eingelesen

⁵ <http://gbif.org/>

⁶ <http://openseadragon.github.io/>

werden. In dem Fall von Sammlungsetiketten haben QR-Codes die Vorteile, dass man den meist sehr kleinen Schrifttyp nicht entziffern und die Information nicht abtippen muss, sondern einfach den QR-Code einscannt und die digitale Information z.B. zum Aufrufen des zugehörigen Datensatzes in der Datenbank verwenden kann. Somit ist das eine ideale Schnittstelle zwischen den physischen Objekten und der digitalen Repräsentation, die bei der Digitalisierung entsteht.



Abb. 10: Beispiele von QR-Codes in der entomologischen Sammlung. Links: Kasten-QR-Code; Mitte: Art-QR-Code; Rechts: Individuum-QR-Code.

Im Fall der digitalisierten Insektenkästen wurden die QR-Codes auf mehreren Ebenen eingesetzt: Jeder Kasten, jede Individuengruppe einer Art und ggf. wichtige Einzelindividuen (z.B. Typusexemplare) tragen jeweils einen eigenen QR-Code (s. Abb. 10). Durch die Vergabe auf diesen Ebenen wird ein weiterer Vorteil der binären Matrizen auf den Etiketten deutlich: Die Relationen, die sich physisch ergeben, können ohne aufwendige manuelle Eingabe hergestellt werden, indem man in ein vorgefertigtes Formular (z.B. in Excel) die Identifikatoren aus den einfach gescannten QR-Codes einträgt. Beispielsweise enthält ein Insektenkasten mehrere Arten, diese sind mit mehreren Individuen dergleichen Art vertreten. Nun konnte einfach dokumentiert werden, welche Art und welches Individuum in welchem Insektenkasten steckt, wenn der QR-Code des Kastens einmalig gescannt wird und alle enthaltenen Arten und Individuen-Codes eingescannt werden. Dies ist gerade von Vorteil zur Dokumentation von Änderungen der physischen Zuordnung, denn durch das alltägliche Sammlungsmanagement werden regelmäßig Individuen von einem Insektenkasten in einen anderen umorganisiert. Die Anzahl der Optionen für ein technisch unterstütztes und somit effizienteres Sammlungsmanagement ist durch die QR-Code-Vergabe während der Digitalisierung nachhaltig gestiegen. So konnte den Kustoden und Sammlungspflegern eine einfache Möglichkeit gegeben werden, Listen von Inventarnummern in den

Sammlungsräumen durch Scannen der QR-Codes mit einem Smartphone zu erstellen. Dieser „Warenkorb“ kann dann direkt zur Selektion der Datensätze aus der Sammlungsdatenbank genutzt werden, um beispielsweise Anfragen nach Informationen abzuarbeiten, Leihscheine und Inventarlisten zu drucken und / oder weitere Informationen zu den Objekten in die Datenbank einzutragen.

Für den Digitalisierungsprozess selbst wurde dieses Prinzip ebenfalls genutzt: Digitalisierte Insektenkästen haben ein Bild, das nur solange aktuell ist, bis sich in dem Kasten selbst etwas ändert (z.B. neue Objekte hinzukommen). Somit zeigen die Digitalisate immer nur ein Abbild der Objekte zu einem bestimmten Zeitpunkt. Um Änderungen der bereits digitalisierten Kästen zu erkennen, wurden die durch das Sammlungspersonal veränderten Kästen mit einem orangenen aufgeklebten Punkt versehen, während die frisch digitalisierten Kästen mit einem grünen Punkt markiert waren. Mit Hilfe der über die QR-Code-Etiketten möglichen „Warenkorb“-Zusammenstellung konnten wir innerhalb kurzer Zeit eine Liste der Insektenkästen erstellen, die mit einem orangenen Punkt versehen waren, um eine Inventarisierung der nach dem Digitalisieren veränderten Kästen durchzuführen.

4.2 Digitalisierungs-Datenbank

Im Rahmen des Projektes wurde eine umfangreiche Datenbank erstellt, die die Digitalisierung als Erfassungs-Werkzeug begleitete. Sie diente vor allem dazu die Einzelschritte des Digitalisierungsprozesses abzubilden, so dass der aktuelle Status eines Objektes bis zur Archivierung des Digitalisates immer erkennbar war. Diese Meta-Information spielt bei der Koordination eines größeren Digitalisierungsteams in der Massendigitalisierung eine entscheidende Rolle, um den Arbeitsprozess personenbezogen zu dokumentieren. In Hinblick auf die Optimierung der Arbeitsabläufe während der Projekt-Laufzeit wurde auch die Datenbank fortlaufend weiterentwickelt und den Ansprüchen des Prozesses angepasst. Dazu gehören Optionen zur Suche von Daten, Schnelleingabemöglichkeiten, Autovervollständigungen, Plausibilitätsprüfungen sowie zur Erstellung von Berichten, Fortschrittsstatistiken und Etiketten (mit QR-Codes) für die physischen Objekte. Unsere im Haus eingesetzte Sammlungsdatenbank Specify 6 ist nicht für die Schnell- und Masseneingabe von Daten ausgelegt und hätte den Arbeitsfluss bei der Digitalisierung stark verlangsamt. Aus diesem Grund wurden die digitalisierten Inhalte (Artnamen, Typenstatus, Inventarnummern, ggf. Fundort oder biogeografische Region etc.) ebenfalls in der Digitalisierungsdatenbank abgelegt. Dabei wurden die Strukturen so kompatibel zur Sammlungsdatenbank aufgebaut, dass ein einfacher und automatisierter Transfer der Daten ermöglicht wurde. Nach dem Transfer der Daten stehen die digitalisierten Sammlungsinformationen dem normalen Sammlungsmanagement zur Verfügung, wobei die Meta-Informationen in der Digitalisierungsdatenbank die Digitalisierungshistorie abbilden.

Nachdem ein Sammlungsobjekt in der Digitalisierungsdatenbank angelegt und das Digitalisat erstellt wurde, startete ein halbautomatisierter Workflow, der zur Nachbearbeitung, Konvertierung, Archivieren und Publizieren der Bilder genutzt wurde. Gerade bei der Plausibilisierung bedarf es eines manuellen Eingriffs, durch den ggf. Digitalisate, die durch die Qualitätsprüfung fallen, wieder in den Anfangsprozess überführt werden konnten, um ggf. eine Nachjustierung (z.B. beim Scannen selbst oder in der Nachbearbeitung) zum Erreichen optimaler Ergebnisse zu ermöglichen. Der automatisierte Anteil dieser Prozesse lief dabei asynchron, so dass sich in dem gesamten Ablauf keine

Warteschlangen in der Bearbeitung bildeten. Der jeweilige Bearbeitungsstatus der Digitalisate wurde dabei recht einfach nachvollziehbar gemacht, indem eine Verzeichnisstruktur im Dateisystem aufgebaut wurde, die den Arbeitsschritten entspricht. Über die eindeutige Bezeichnung der Dateien kann man auch ganz unabhängig von der Datenbank erkennen, wo ein Bild in einem Prozess „hängt“. Das sorgt für eine erhöhte Transparenz für alle an der Massendigitalisierung beteiligten Projekt-Mitarbeiter.

Die Digitalisierungsdatenbank wurde in dem Digitalisierungsprozess nicht als Medienrepositorium genutzt. Lediglich die Referenzen zu den Digitalisaten wurden abgebildet, da die erstellten hochauflösenden Bilder in die im MfN genutzte Mediendatenbank Canto Cumulus eingeflossen sind. Hier werden die Bilder archiviert und über Backups gesichert. Gleichzeitig dient die Mediendatenbank mit einer Webschnittstelle als Medienrepositorium. So werden die statisch auf der Projekt-Webseite eingebundenen Digitalisate direkt aus dem Medienarchiv geladen, was eine zusätzliche Kopie von mehreren 100 GB auf dem Webserver überflüssig macht. Die dynamisch zoombaren Digitalisate werden allerdings mit dem Programm KrPano gerendert. Diese Objekte müssen nicht im Backup berücksichtigt werden, da sie in einem vorgeschalteten, automatischen Prozess aus der Mediendatenbank jederzeit neu generiert werden können.

Für die Darstellung und Suchbarkeit der Digitalisate auf der Projekt-Webseite wurde in die Digitalisierungsdatenbank eine Export-Routine entwickelt, die den aktuellen Datenstand der abgeschlossen digitalisierten Objekte auf eine webbasierte Datenbank transferierte. Die hierbei notwendigen Datentransformationen wurden automatisiert beim Export durchgeführt, damit die Daten das von der Projekt-Webseite vorgesehene Format hatten. Dabei wurde darauf geachtet, dass die Daten dennoch für eine Mehrfachnutzung jenseits der Projekt-Webseite geeignet sind. So wurde aus derselben Web-Repräsentation der Daten die Anbindung an domainübergreifende Datenportale (siehe Kapitel 6) ohne zusätzliche Exporte ermöglicht.

5 Erfolgsindikatoren

5.1 Insektenkästen

Die Digitalisierung der Insektenkästen am MfN konnten wir mit Erfolg abschließen. Es wurden 10.181 Kästen digitalisiert, das entspricht ca. 30% aller Insektenkästen bei uns am MfN, und auf unserer Projektwebseite veröffentlicht. Die genaue Zahl der Individuen, die sich in den gescannten Insektenkästen befinden, lässt sich nur näherungsweise bestimmen, denn es war nicht möglich und auch nicht vorgesehen im Rahmen des Projektes die Anzahl der Tiere pro Kasten zu erfassen. Wenn wir davon ausgehen, dass sich in 35.000 Kästen, 15 Millionen Objekte befinden, so haben wir mit der Digitalisierung der Insektenkästen etwa 4,5 Millionen Tiere digitalisiert.

5.2 Einzeltierdigitalisierung

Für die Digitalisierung von Einzeltieren wurden drei unterschiedliche Ansätze gewählt. Bei der ersten Methode wurde sowohl die Oberseite als auch die Unterseite jedes einzelnen Sammlungsobjektes fotografiert und nur in Einzelfällen kamen noch Seitenansichten und Frontalansichten hinzu (s. Kapitel 3.2). Mit dieser Methode haben wir uns im Rahmen des Projektes ausschließlich auf Schmetterlinge und Libellen fokussiert. Diese Objekte können

trotz ihrer räumlichen Struktur, eher als flache Objekte betrachtet werden und sind mit zwei Bildern i.d.R. gut zu dokumentieren und für taxonomische Arbeiten oftmals ausreichend. Außerdem bekommt jedes Insekt ein Etikett mit einem eindeutigen QR-Code. Es konnten 2.824 Sammlungsobjekte mit mehr als 6.000 Fotos digital erfasst werden.

Eine weitere Methode wurde im Rahmen des Projektes zur Digitalisierung von Einzeltieren entwickelt und eingesetzt. Hierbei werden die Tiere, zusammen mit den Etiketten, aus einem 45° Winkel fotografiert und ebenfalls mit einem eindeutigen QR-Code versehen (s. Kapitel 3.3). Mit dieser Methode wurden 7.176 Sammlungsobjekte digitalisiert, mit jeweils einem Bild.

Die dritte Methode, die zum Einsatz kam, wurde im Rahmen des Projektes entwickelt, befindet sich jedoch auch nach Abschluss des Projektes noch in der Optimierung. Es handelt sich hierbei um die 3D-Digitalisierung zur räumlichen Visualisierung von genadelten Sammlungsobjekten und hat von uns den Namen „ZooSphere“ erhalten (s. Kapitel 3.4). Für die Digitalisierung eines einzigen Sammlungsobjektes werden 2.000 bis 6.000 Fotos benötigt, in Abhängigkeit von der Größe des Objektes und der gewünschten Anzahl an Perspektiven. Da die Entwicklung dieser Methode noch nicht abgeschlossen ist und diese zudem zeitintensiver ist als die beiden zuvor genannten Methoden, konnten bisher erst ca. 50 Sammlungsobjekte digitalisiert werden. Der große Vorteil dieser Methode besteht allerdings in ihrer Bedeutung für die taxonomische Forschung, denn diese kann damit nicht nur effizienter werden, sondern sie wird gleichzeitig demokratisiert. Das heißt taxonomische Forschung ist nicht mehr abhängig von dem Budget das ein Wissenschaftler für seine Forschung zur Verfügung hat.

Insgesamt konnten mit diesen drei Digitalisierungsmethoden 10.075 Tiere digital erfasst und auf der Projekt- bzw. der ZooSphere-Webseite veröffentlicht werden.

5.3 Webseiten

Im Rahmen des Projektes wurden zwei Webseiten entwickelt. Zum einen die Projektwebseite⁷, zur Darstellung des Projektes nach außen und für die Veröffentlichung der im Rahmen des Projektes digitalisierten Sammlungsteile (s. Abb. 11). Hier finden sich alle digitalen Insektenkästen über die Zuordnung zu deren Sammlung. Die Fotos der einzelnen Sammlungsobjekte sind dem Insektenkasten zugeordnet, aus dem sie entnommen wurden. Darüber hinaus können Arten auch über ihren wissenschaftlichen Namen gefunden werden.

⁷ <http://eos.naturkundemuseum-berlin.de/>

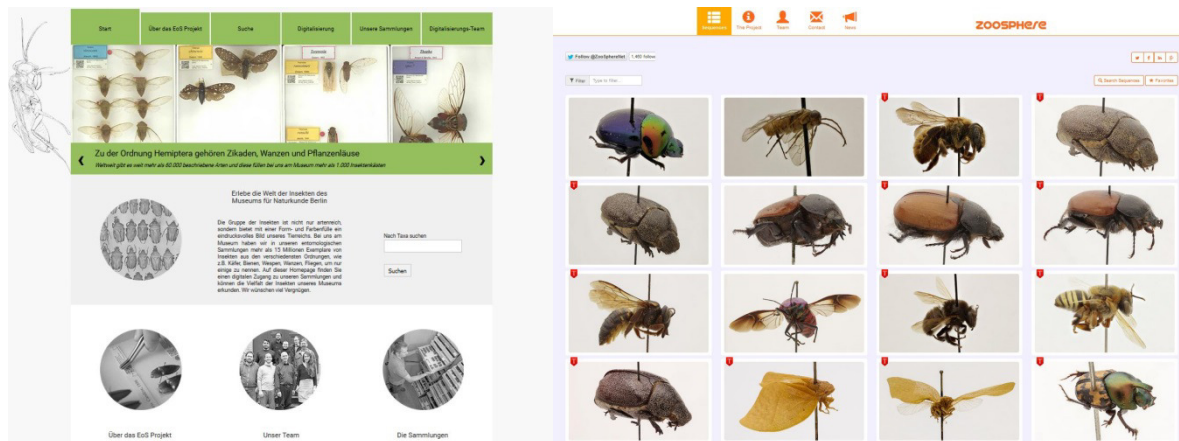


Abb. 11: Screenshots der im Rahmen des Projekts entwickelten Webseiten. Links: EoS-Projektwebseite und Insektenkastenportal; Rechts: ZooSphere-Webseite.

Die zweite Webseite wurde für die Präsentation der ZooSphere-Objekte erstellt⁸. Da diese Objekte ganz eigene Ansprüche an die Darstellung haben, musste für die ZooSphere-Objekte ein eigenes Verfahren für die Präsentation entwickelt werden. Die Herausforderung bestand darin die hochauflösenden Bilder online zur Verfügung zu stellen, ohne dass die großen Bilddateien sich negativ auf die Performance der 3D-Betrachtung auswirken. Jedes ZooSphere-Objekt kann bis zu 500 MB groß sein, in Abhängigkeit von der Anzahl der fotografierten Perspektiven (s. Abb. 11).

Die freie Verfügbarmachung der Sammlungsobjekte über das Internet trägt nicht nur zum physischen Schutz von wertvollem Kultur- und Wissenschaftsgutes bei (präventive Konservierung), sondern ermöglicht einen völlig neuen Zugang zur Sammlung für eine breite Bevölkerungsschicht und über Sprachbarrieren hinweg.

Für das gesamte Bildmaterial werden zwei Creative Commons Lizenzen verwendet. Die Bilder der Insektenkästen und der ZooSphere-Objekte werden unter der Creative Commons Lizenz CC0 1.0⁹ veröffentlicht und das hochauflösende Bildmaterial von Einzeltieren (i.d.R. von Typusexemplaren) steht unter der CC Lizenz CC-BY-SA 4.0¹⁰.

5.4 Ausgerichtete Tagungen und Workshops

5.4.1 Internationaler Workshop zur Massendigitalisierung

Am 9. September 2014 wurde der Internationale Workshop zur Massendigitalisierung im Rahmen der Tagung DigitalSpecimen von unserem Projekt organisiert und durchgeführt. Die Vortragenden kamen von 12 unterschiedlichen Institutionen aus sieben Ländern. Unser Projekt war mit drei Vorträgen vertreten. Der Workshop bot einen sehr guten Überblick über aktuelle, internationale Entwicklungen in der Sammlungsdigitalisierung im Naturhistorischen Bereich (Abstracts s. Anhang 10.1).

⁸ <http://www.zoosphere.net/>

⁹ <http://creativecommons.org/publicdomain/zero/1.0/deed.de>

¹⁰ <http://creativecommons.org/licenses/by-sa/4.0/>

5.4.2 Konferenz Langzeitzugriff auf Sammlungs- und Multimediadaten

Die Konferenz „Langzeitzugriff auf Sammlungsdaten und Multimediadaten“ (KoLa2015) wurde vom 24. Juni bis zum 26. Juni 2015 durchgeführt. Weitere Informationen finden sich auf der Tagungswebseite unter <http://kola2015.naturkundemuseum-berlin.de>. Fast 90 Personen aus ganz Deutschland haben an der Konferenz teilgenommen. Das Feedback der Teilnehmer war so positiv, dass überlegt wurde die Konferenz in zwei Jahren erneut durchzuführen, denn sowohl der Langzeitzugriff auf unsere Daten, als auch die Langzeitarchivierung wird uns als Thema noch lange begleiten. Die nächste Konferenz soll zusammen mit nestor¹¹ geplant und durchgeführt werden (Abstracts s. Anhang 10.2).

Die Eröffnungsveranstaltung fand am 24. Juni im Sauriersaal des MfN statt. Es nahmen bereits 60 Personen an dieser Veranstaltung teil und hörten den sehr spannenden Eröffnungsvortrag von Paul Klimpel mit dem Titel „Nachhaltigkeit, Zugang und Vernetzung“.

Die Konferenz fand an den nächsten beiden Tagen, 25. und 26. Juni, in den Räumlichkeiten der Leibniz Gemeinschaft statt. Sie gliederte sich in vier Sektionen:

- Weitergabe und Import von Daten und Metadaten
- Verwaltung und Pflege von Sammlungs- und Multimediadaten
- Backup von Sammlungs- und Multimediadaten (Speicherung und Langzeitarchivierung)
- Datenzugriff, Verfügbarkeit und Veröffentlichung

5.5 Präsentationen auf Tagungen

5.5.1 Oktober 2013: Herbsttagung der Fachgruppe Naturwissenschaftliche Museen im Deutschen Museumsbund e.V., Kassel (Naturkundemuseum im Ottoneum).

Kroupa, A.S., Glöckler, F., Schurian, B. & Maier, F. (2013) 200.000 Insektenkästen ins Netz - Mobilisierung und Verfügbarmachung deutscher entomologischer Sammlungen. (Vortrag).

5.5.2 März 2014: 33. Jahrestagung der Gesellschaft deutschsprachiger Odonatologen e.V., Kiel (Universität).

Turiault, M., Fahrngruber, K., Glöckler, G., Hentschel, S., Kroupa, A.S., Lange, M. & Maier, F. (2014) Wie kommt die Libelle ins Netz? Digitalisierung von Insektenkästen am Museum für Naturkunde Berlin. (Poster)

5.5.3 Juni 2014: The Society for the Preservation of Natural History Collections, Cardiff (Wales Millenium Center)

Kroupa, A.S., Pluta, M., Schurian, B. & Glöckler, F. (2014) ZooSphere - A tool for automated spheric image capturing and interactive 3D visualization of biological collection objects. (Vortrag)

¹¹ www.langzeitarchivierung.de/

5.5.4 September 2014: DigitalSpecimen, Symposium Mass Digitization, Berlin (Humboldt-Universität).

Kroupa, A.S., Glöckler, F., Hagedorn, G., Häuser, C., Schmidt, S. & Schurian, B. (2014) The Open Drawer Project - Providing free access to high resolution images of entomological collection drawers (digitization - participation - publication). (Vortrag)

Pluta, M. (2014) ZooSphere - Development of a software for automated spheric image capturing and interactive 3D visualization of biological collection objects. (Vortrag)

Glöckler, F. (2014) Stable URIs for mass digitization approaches, strategy and workflows. (Vortrag)

5.5.5 September 2014: Mit Netz und offenen Daten - Kulturgut digital, Berlin (Zuse-Institut).

EoS Team (2014) Insektenkästen 2.0 – Erfassung objektreicher Speziessammlungen (EoS) am Museum für Naturkunde Berlin. (Poster)

5.5.6 November 2014: Fachkonferenz Museum von Babel, Frankfurt am Main.

Kroupa, A.S. (2014) Vom Sinn und "Unsinn" der Digitalisierung entomologischer Sammlungen - Welchen Einfluss hat die Digitalisierung auf die Forschung? (Vortrag)

5.5.7 Januar 2015: Konferenz "3D ins Museum?", Berlin (Kunstgewerbemuseum).

Kroupa, A.S., Pluta, M., Schurian, B. & Maier, F. (2015) Development of a software for automated spheric image capturing and interactive 3D visualization of biological collection objects. (Vortrag)

5.5.8 Juni 2015: Konferenz Langzeitzugriff auf Sammlungs- und Multimediadaten, Berlin (Leibniz-Gemeinschaft).

Glöckler, F. (2015) Publishing digital representations of collection objects via stable URIs. Konferenz Langzeitzugriff auf Sammlungs- und Multimediadaten. (Vortrag)

Kroupa, A.S. (2015) Effiziente Arbeitsabläufe und innovative Methoden zur Erschließung und dauerhaften Verfügbarmachung objektreicher Speziessammlungen am Beispiel der entomologischen Sammlung des Museum für Naturkunde Berlin. (Vortrag)

Kuhlmann, C., Pluta, M., Maier, F., Schurian, B. & Kroupa, A.S. (2015) ZooSphere – globalizing taxonomic sharing of insects in 3D. (Poster)

5.5.9 Oktober 2015: Berliner Bibliothekswissenschaftliches Kolloquium, Berlin (Institut für Bibliotheks- und Informationswissenschaft)

Kroupa, A.S. (2015) Wie kommen 15 Millionen Sammlungsobjekte ins Netz? Und warum eigentlich? (Vortrag)

5.5.10 Oktober 2015: Fachgruppe Dokumentation des Deutschen Museumsbundes, Berlin (Konrad Zuse Institut)

Kroupa, A.S. (2015) Aktuelle Digitalisierung am MfN - ZooSphere-Ansatz (Vortrag)

5.6 Teilnahme an Veranstaltungen

5.6.1 Mai 2014: Lange Nacht der Wissenschaften, Berlin (MfN).

Bei der Langen Nacht der Wissenschaften am 10. Mai wurde die Internetseite des Projektes vor der Biodiversitätswand auf einer 5m x 5m großen Leinwand vorgestellt. Des Weiteren wurde das Tool ZooSphere präsentiert und live vorgeführt. Führungen hinter die Kulissen zeigten den Besuchern die Insektenkastendigitalisierung im SatScan-Raum.

5.6.2 Juli 2014: Coding da der Kultur-Hackathon Vinci, Berlin.¹²

Coding da Vinci ist ein Kultur-Hackathon, der Daten von Kultur- und Gedächtnisinstitutionen Entwicklern und Programmierern näher bringen möchte, um so neue Verwertungsmöglichkeiten zu schaffen. Das MfN stellte bereits beim ersten Hackathon im Jahr 2014 u.a. 50 Bilder von Insektenkästen zusammen mit deren Metadaten, zur Verfügung und stockte den Umfang für das Jahr 2015 auf¹³. Unter den Gewinnern des Hackathons 2014 war der Cyberbeetle¹⁴, ein Kunstprojekt. Hier wurden Bilder von Insektenkästen mit Bildern aus dem Herbarium des Botanischen Gartens und Botanischen Museums Berlins kombiniert und mit einem passenden Lied aus Tönen des Tierstimmenarchivs (MfN) untermalt. Das Projekt zeigt beispielhaft was bei der Kooperation zwischen Entwicklern- und Gedächtnisinstitutionen entstehen kann und inspiriert für neue Projekte. Die Daten verbleiben auf der Coding da Vinci Projektseite und stehen auch für zukünftige Projekte zur Verfügung.

5.6.3 Dezember 2014: MINT Lehrerkongress, Berlin (Urania Berlin e.V.).

Die Digitalisierung von ZooSphere-Objekten wurde live demonstriert und die Darstellung auf der ZooSphere-Webseite erläutert.

5.6.4 Juni 2015: Lange Nacht der Wissenschaften, Berlin (MfN).

Bei der Langen Nacht der Wissenschaften am 13. Juni wurden Führungen zur Insektenkastendigitalisierung organisiert und durchgeführt. Des Weiteren wurde das Tool ZooSphere präsentiert und live vorgeführt.

5.7 Publikationen

Balke, M., Schmidt, S., Hausmann, A., Toussaint, E., Bergsten, J., Buffington, M., Häuser, C.L., Kroupa, A.S., Hagedorn, G., Riedel, A., Polaszek, A., Ubaidillah, R., Krogmann, L., Zwick, A., Fikáček, M., Hájek, J., Michat, M.C., Dietrich, C., La Salle, J., Mantle, B., Ng, P.K.L. and Hobern, D. 2013. Biodiversity into your hands - A call for a virtual global natural history 'metacollection'. *Frontiers in Zoology* 10: 55. doi:10.1186/1742-9994-10-55.

¹² Coding da Vinci: <http://codingdavinci.de/>

¹³ Daten des Coding da Vinci Kultur-Hackathon: <http://codingdavinci.de/daten/>

¹⁴ Cyberbeetle: <http://katihyppa.com/cyberbeetle/>

6 Domänenübergreifende Publikation der Digitalisate

6.1 Europeana Portal¹⁵

Das Museum für Naturkunde war Partner des im Februar 2014 erfolgreich abgeschlossenen Projektes OpenUp! – Opening up the natural history heritage for Europa¹⁶. Die innerhalb des Projektes etablierten Infrastrukturen¹⁷ bestehen weiter und werden von allen Partnern kontinuierlich genutzt um neue Inhalte auf Europeana zu veröffentlichen.

Auch ein Teil der gescannten Insektenkästen (5738, Stand: 20.07.2015) sind bereits über das Portal verfügbar und stehen den Nutzern unter der Lizenz CC0 zur Verfügung (s. Abb. 12). Die Veröffentlichung erfolgt prozessbedingt zeitverzögert. Bis zum Ende des Jahres sollen jedoch alle innerhalb des Projektes entstandenen Kasten- und Einzeltieraufnahmen an das Portal angebunden sein.

Europeana selbst stellt nicht die Bilder zur Verfügung, sondern kreiert nur ein Vorschaubild (s. Abb. 12a), das im Portal gezeigt wird. Die Originalbilder verbleiben bei den Institutionen und sind vom Portal aus über einen Link zu erreichen (s. Abb. 12b). Im Fall der Insektenkastenbilder wird auf das Insektenkastenportal¹⁸ verlinkt und die Bilder können in der höchstmöglichen, zoombaren Auflösung angeschaut werden. Beim Klick auf das Vorschaubild öffnet sich die Lightbox und das Bild kann in einer komprimierten Version vom Nutzer heruntergeladen werden.

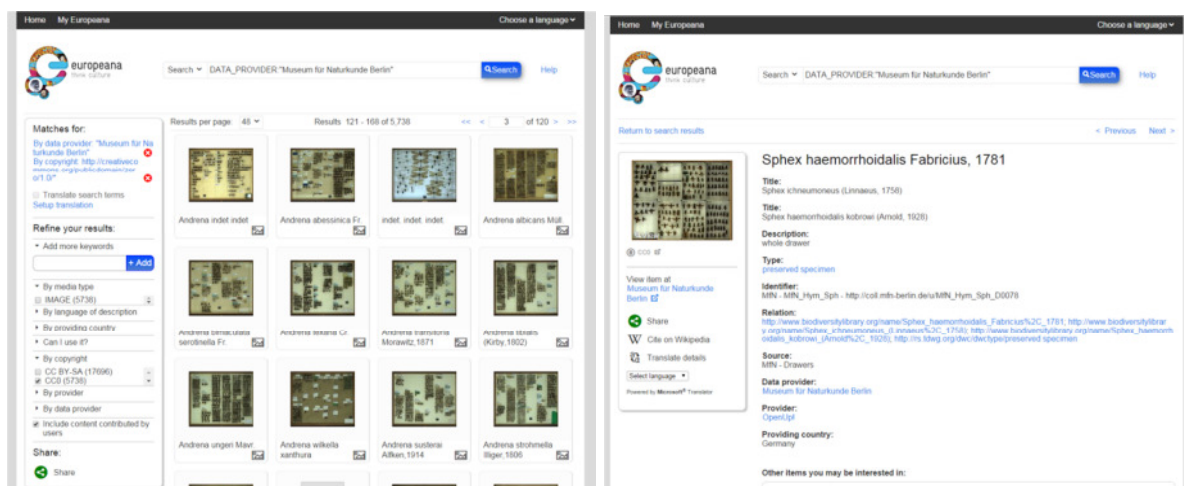


Abb. 12: Screenshot des Europeana Portals. Übersicht mit Vorschaubildern verschiedener Kästen (a), einzelner Kasten mit weiterführenden Informationen zum Bild, Links zu verfügbarer Literatur in der Biodiversity Heritage Library (unter „Relation“) und ein Link zum Insektenkastenportal (View item at), wo das Bild in hoher Auflösung angeschaut werden kann (b).

¹⁵ Europeana Portal: <http://www.europeana.eu/portal/>

¹⁶ OpenUp! Projekt : <http://open-up.eu/en>

¹⁷ Flowing smoothly into Europeana: the OpenUp! technical workflow: <http://open-up.eu/en/node/957>

¹⁸ Insektenkastenportal: <http://eos.naturkundemuseum-berlin.de/>

6.2 Deutsche Digitale Bibliothek¹⁹

Bereits im Jahr 2013 gab es erste Gespräche zwischen der Fachstelle Museum der Deutschen Digitalen Bibliothek (DDB) und dem MfN mit dem Ziel möglichst bald Inhalte des MfNs auch auf nationaler Ebene zu veröffentlichen. Der entsprechende Kooperationsvertrag wurde ebenfalls im Jahr 2013 von beiden Seiten unterzeichnet. Die DDB akzeptiert u.a. das Europeana Data Model (EDM) als Eingangsformat für Metadaten. Leider ist EDM ein sehr variabler Standard und seine Anwendung unterscheidet sich zwischen der DDB und Europeana. Wir bedienen aktuell nur das Europeana spezifische EDM. Erste Datensätze dieses Formats wurden bereits im Dezember 2014 an die DDB zum Abgleich geschickt und wir warten seitdem auf Rückmeldung seitens der DDB. Aktuell arbeitet die DDB an einem konzeptionellen Mapping der Daten und ein (Test) Import in das Portal ist für den Herbst 2015 geplant.

Neben EDM können wir der DDB unsere Daten auch in einem Domänspezifischen Format für Biodiversitätsdaten (ABCD²⁰) liefern. Es gab Überlegungen seitens der DDB diesen auch als Standard zu akzeptieren, eine finale Entscheidung wurde jedoch noch nicht getroffen.

Sobald die technischen Hürden behoben sind, werden die Kasten- und Einzeltierbilder auch in der DDB zur Verfügung stehen.

7 Feedback

Das Feedback von externen Wissenschaftlern zu unserem Projekt, welches das Ziel verfolgt sowohl die Insektenkästen als auch das Typenmaterial digital online zur Verfügung zu haben, war sehr positiv. Uns wurde von allen Seiten bestätigt, dass die online zur Verfügung stehenden Insektenkästen einen großen Vorteil bieten um einen Museumsbesuch zu planen, den Zustand von Typen bereits von Zuhause aus zu erkennen und zu entscheiden welches Material wirklich über eine Ausleihe benötigt wird.

Mit dem neuen Ansatz der räumlichen Visualisierung von genadelten Insekten sind wir auf breites Interesse gestoßen, sowohl bei Wissenschaftlern aber auch bei Museen. Wissenschaftler sind begeistert, dass sie innerhalb kurzer Zeit die qualitativ hochwertigen ZooSphere-Objekte für ihre Forschung zur Verfügung haben und in fast allen Fällen sind die Digitalisate für die taxonomische Forschung mehr als ausreichend. Museumsmitarbeiter sind bei allen Veranstaltungen, auf denen wir ZooSphere präsentiert hatten, an uns herantreten und haben sich danach erkundigt, wo die Hard- und Software von ZooSphere käuflich erworben werden kann und was sie kostet, bzw. ob sie Material digitalisieren lassen können.

Hier eine Auswahl von Rückmeldungen einiger Wissenschaftlern:

Natalia von Ellenrieder: *„I really look forward to seeing the Berlin type collection digitized - that's going to be extremely useful to all dragonfly researchers.“*

Bernd Kunz: *“Es ist eine Wahnsinnsache, von Zuhause aus eine Museumssammlung durchstöbern zu können. Chapeau.“*

¹⁹ Deutsche digitale Bibliothek (DDB): <https://www.deutsche-digitale-bibliothek.de/>

²⁰ Eine Einführung in ABCD: <http://wiki.tdwg.org/wiki/bin/view/ABCD/AbcdIntroduction>

Max Kasperek: „Bei dieser Angelegenheit möchte ich auch darauf hinweisen, wie wertvoll mir (und sicher vielen anderen auch) die digitale Erschließung Ihrer Sammlungskästen ist. Es ist dies eine einmalige Gelegenheit, sich vorab eine Übersicht über die Sammlungsbestände zu verschaffen, womit ich wertvolle Zeit gewinnen kann. Zudem sind die Aufnahmen von so hoher Qualität, dass einige Fragen zum Material bereits vorab geklärt werden können.“

8 Ausblick

Sowohl die Technik als auch die Workflows zur Fortsetzung der Digitalisierung der entomologischen Sammlungen sind etabliert und könnten bei einer erneuten Finanzierung sofort wieder eingesetzt werden. Die einzige Technologie die noch optimiert und damit weiterentwickelt werden muss ist diejenige der räumlichen Visualisierung, ZooSphere. Auch wenn mit dieser Methode jetzt täglich Sammlungsobjekte digitalisiert werden, kann und muss weitere Entwicklungszeit in die Optimierung investiert werden. Optimierungsbedarf besteht z. B. bei der Reduzierung der Aufnahmezeit für ein Objekt, weitere Verbesserung der Bildqualität, Vereinfachung des Workflows aber auch eine Vereinfachung des Postprocessings.

Da die ZooSphere-Technologie für eine große Aufmerksamkeit gesorgt hat, werden wir einen BMBF-Antrag im Rahmen der Förderrichtlinie VIP+²¹ vorbereiten. Der BMBF-Antrag hat das Ziel die Hard- und Software ZooSphere zur Vermarktung hinzuführen und mehrere Ansätze zu evaluieren. So soll z. B. eine Marktanalyse zur Evaluierung des Vermarktungspotentials der Hard- und Software erstellt werden. Die Etablierung eines On-Demand-Digitalisierungsservices sowohl für naturhistorische Sammlungsobjekte, aber ebenso für Sammlungsobjekte aus anderen Domänen, wird betrachtet. Inwieweit die ZooSphere-Objekte in Bildung und Lehre eingesetzt werden können, z. B. bei der Verwendung für taxonomische Bestimmungsübungen an Universitäten, wird ein weiterer Schwerpunkt des Antrags. Die Verwendung der digitalen Sammlungsobjekte im Ausstellungsbereich wird ebenfalls untersucht. Hier liegt der Fokus auf einer interaktiven Ausstellung digitaler Insekten in 3D.

Trotz der erreichten Projektziele, des positiven Feedbacks von Wissenschaftlern (s. Kapitel 7) und der deutlich gesteigerten Akzeptanz von Kustoden und Sammlungsmitarbeiter gegenüber der Digitalisierung am Ende des Projektes, ist die Nachhaltigkeit nicht gewährleistet. Dies liegt v.a. darin begründet, dass die entomologische Sammlung mit 15 Millionen Sammlungsobjekten eine extrem umfangreiche Sammlung ist und damit eine komplette Digitalisierung auf Objektebene äußerst zeit- und dementsprechend auch kostenintensiv ist. Eine Priorisierung ist deshalb unumgänglich und die Sammlung muss auf unterschiedlichen Ebenen digitalisiert werden, wie in dem Projekt bereits durchgeführt. Auch mit einer Priorisierung und einer optimalen Digitalisierungsstrategie bleibt die Sammlungsdigitalisierung durch die große Anzahl an Objekten und dem damit einhergehenden personellen Aufwand, kostenintensiv.

Um eine Vorstellung darüber zu bekommen wie hoch die Kosten wären, wenn weitere Teile der entomologischen Sammlungen digitalisiert würden, hier eine Hochrechnung (s. Tab. 1).

²¹ <https://www.bmbf.de/de/vip-technologische-und-gesellschaftliche-innovationspotenziale-erschliessen-563.html>

Die Digitalisierung der restlichen 25.000 Insektenkästen würde sich auf ca. 750.000 Euro belaufen und eine Person benötigt dafür ungefähr 20 Jahre. Beschränkt man sich bei den Einzelobjekten auf die für die taxonomische Forschung besonders wichtigen primären Typen, so werden nochmals 900.000 Euro nur für die Digitalisierung benötigt. Hier werden allerdings zwei Annahmen für die Berechnung zu Grunde gelegt, zum einen, dass wir bei uns am Museum 50.000 primäre Typen haben (dies ist nur eine grobe Schätzung) und die mehr oder weniger aus dem Bauch heraus getroffene Aufteilung der Typen auf die beiden Einzeltier-Digitalisierungsmethoden. Die zeit- und damit auch kostenintensivere Methode ZooSphere sollte nur bei solchen Sammlungsobjekten zur Anwendung kommen, bei denen zwei Bilder – von oben und unten – nicht ausreichend sind. Bei der Hochrechnung sind allerdings noch nicht die Kosten für IT-Kräfte, technische Betreuung der Digitalisierungsstationen und für die Koordination berücksichtigt. Diese Kosten sind abhängig von der Größe der Digitalisierungsteams sowie des Digitalisierungszeitraums und müssen dementsprechend den Kosten noch hinzugefügt werden.

Tab. 1: Geschätzte Kosten für die Digitalisierung der entomologischen Sammlungen bei unterschiedlichen Methoden und Objektklassen.

Digitalisierungsstation	Anzahl	Mannjahre	Kosten
SatScan	25.000 Insektenkästen	21	850.000 €
Einzeltiere	20.000 primäre Typen	3,5	140.000 €
ZooSphere	30.000 primäre Typen	19	750.000 €
Dinolite	15 Millionen Objekte	156	11 Millionen €

Legt man einen realistischen Zeitraum von vier Jahren fest, in welchem sowohl die Digitalisierung der restlichen Insektenkästen als auch der primären Typen durchgeführt werden soll, so ergibt sich ein benötigter Gesamtbetrag von ca. 2,4 Millionen Euro (s. Tab. 2). Dieser Betrag verteilt sich auf 14 Mitarbeiter, welche alle über den gesamten Zeitraum von vier Jahren angestellt sein müssen. Die Infrastruktur für die Durchführung eines solchen Vorhabens ist vorhanden und es müssen nur minimale Mittel für benötigte Geräte oder Instandhaltung eingeplant werden. So müssten z. B. weitere ZooSphere-Digitalisierungsstationen aufgebaut werden, um das Ziel von 30.000 digitalisierten Typen zu erreichen.

Tab. 2: Hochrechnung der Kosten für die Digitalisierung der restlichen Insektenkästen (25.000) und der 50.000 primären Typen mit zwei Digitalisierungsmethoden unter der Annahme diese Arbeiten innerhalb von vier Jahren abzuschließen.

	Mitarbeiter	Kosten
SatScan	5	850.000 €
Einzeltiere	1	140.000 €
ZooSphere	5	750.000 €
Scan-Operator	1	180.000 €
IT-Spezialist	1	200.000 €
Koordination	1	280.000 €
Σ	14	2.400.000 €

9 Zusammenfassung

Die Digitalisierung großer Sammlungsbestände stellt eine große Herausforderung sowohl aus organisatorischer als auch technischer Sicht dar. Eine Voraussetzung für die Digitalisierung ist die Verfügbarkeit der entsprechenden Technik und der technischen Infrastruktur. In vielen Fällen sind bereits eine gute Fotoausrüstung und eine optimale Beleuchtung der Objekte ausreichend, um mit der Digitalisierung zu beginnen. Damit allerdings so kleine Objekte wie Insekten effizient, gleichzeitig aber auch qualitativ hochwertig digitalisiert werden können, wird weiteres Equipment benötigt. Für die Erfassung der Insektenkästen wird ein spezieller Insektenkastenscanner verwendet und zur räumlichen Digitalisierung der kleinen Sammlungsobjekte musste eine neue Methode entwickelt werden.

Der Flaschenhals bei der Massendigitalisierung ist allerdings nicht die benötigte Infrastruktur, auch wenn diese in Einzelfällen (z. B. Insektenkastenscanner) eine größere Investition darstellt, sondern die notwendige Personalressource. Pro Digitalisierungsstation können in der Regel mehrere Personen arbeiten und in Extremfällen, wie z.B. beim Insektenkastenscanner, könnten 10 Mitarbeiter beschäftigt werden, bevor ein einziger Insektenkastenscanner ausgelastet ist. Denkt man einen Schritt weiter, um das Gerät noch effizienter zu nutzen, könnte in zwei Schichten gearbeitet werden und somit sogar 20 Personen beschäftigt werden. Neben den Mitarbeitern für die Digitalisierung werden auch IT-Fachkräfte benötigt, für die Betreuung der Datenbanken, der Webseiten und der benötigten Software. Die Digitalisierungsstationen müssen zudem technisch betreut werden, idealerweise von einem Fotografen, welcher für den optimalen Einsatz der Geräte verantwortlich ist. Damit die Arbeitszeit aller Mitarbeiter effektiv eingesetzt werden kann, ist eine vorausschauende Koordination sehr bedeutsam. Es hat sich bereits in diesem Projekt gezeigt, dass bei mehr als 10 Mitarbeitern immer wieder auf Veränderungen des Teams zu reagieren ist. Mitarbeiter beenden ihr Arbeitsverhältnis, andere Mitarbeiter nehmen eine Elternzeit in Anspruch oder möchten ihre Arbeitszeit aus anderen Gründen verkürzen und andere fallen wiederum krankheitsbedingt für längere Zeit aus (s. Kapitel 2).

Während der Projektlaufzeit konnten Arbeitsabläufe der Digitalisierung am MfN etabliert und optimiert werden. Darüber hinaus wurden auch neue Standards am MfN eingeführt, wie z. B. die Verwendung von QR-Codes auf unterschiedlichen Ebenen (s. Kapitel 4.1). Die damit einhergehenden stabilen Identifier (URIs) sollen in Zukunft in Publikationen für den Verweis auf Sammlungsobjekte zum Einsatz kommen. Ein erstes Beispiel ist die Publikation einer Neubeschreibung einer Wespenart²², in welchem sowohl auf Einzelobjekte²³ als auch auf den digitalen Insektenkasten²⁴ über die stabilen URIs verwiesen wird.

²² Michael Ohl, Volker Lohrmann, Laura Breitzkreuz, Lukas Kirschey und Stefanie Krause: *The Soul-Sucking Wasp by Popular Acclaim – Museum Visitor Participation in Biodiversity Discovery and Taxonomy*. PLoS ONE, 9, 4, e95068, 2014 [doi:10.1371/journal.pone.0095068](https://doi.org/10.1371/journal.pone.0095068)

²³ Z. B. http://coll.mfn-berlin.de/u/MfN_Hym_Amp_I00001.html

²⁴ http://coll.mfn-berlin.de/u/MfN_Hym_Amp_D0003

10 Anhang

10.1 Abstracts des internationalen Workshops zur Massendigitalisierung

10.1.1 *GB-3d Type Fossils Online: mass digitization of palaeontological collections*

Contessi, Michela. British Geological Survey, Nicker Hill, Keyworth, Nottingham, UK & Museo Civico di Scienze Naturali, Faenza, Italy. contessi.m@gmail.com

Howe, M. P. A. British Geological Survey, Nicker Hill, Keyworth, Nottingham, UK.

Harris, S. J. British Geological Survey, Nicker Hill, Keyworth, Nottingham, UK.

McCormick, T. British Geological Survey, Nicker Hill, Keyworth, Nottingham, UK.

Type fossils are essential to palaeontologists - they are the specimens that define a species. Without regular reference to these fossils, the concept of a species can “drift” over time, with the risk of making future research flawed in its understanding. Many of the UK type fossils were first described over a century ago, and with time, their present location and status are often uncertain. Today there are major repositories of type specimens in a number of larger museums across the UK, as well as a number of smaller museums, which may only hold one or two type specimens. This project aims, for the first time, to bring all these records together into a single, searchable web database making them available to anyone on the web as digital models, photographs and stereo anaglyph images. 3d-fossils.ac.uk aims to promote access to collections and information through digitization and free web delivery, targeting academic researchers, university and schools teaching, and the general public. All the material is available under a Creative Commons: Attribution – Share Alike – Non Commercial Licence.

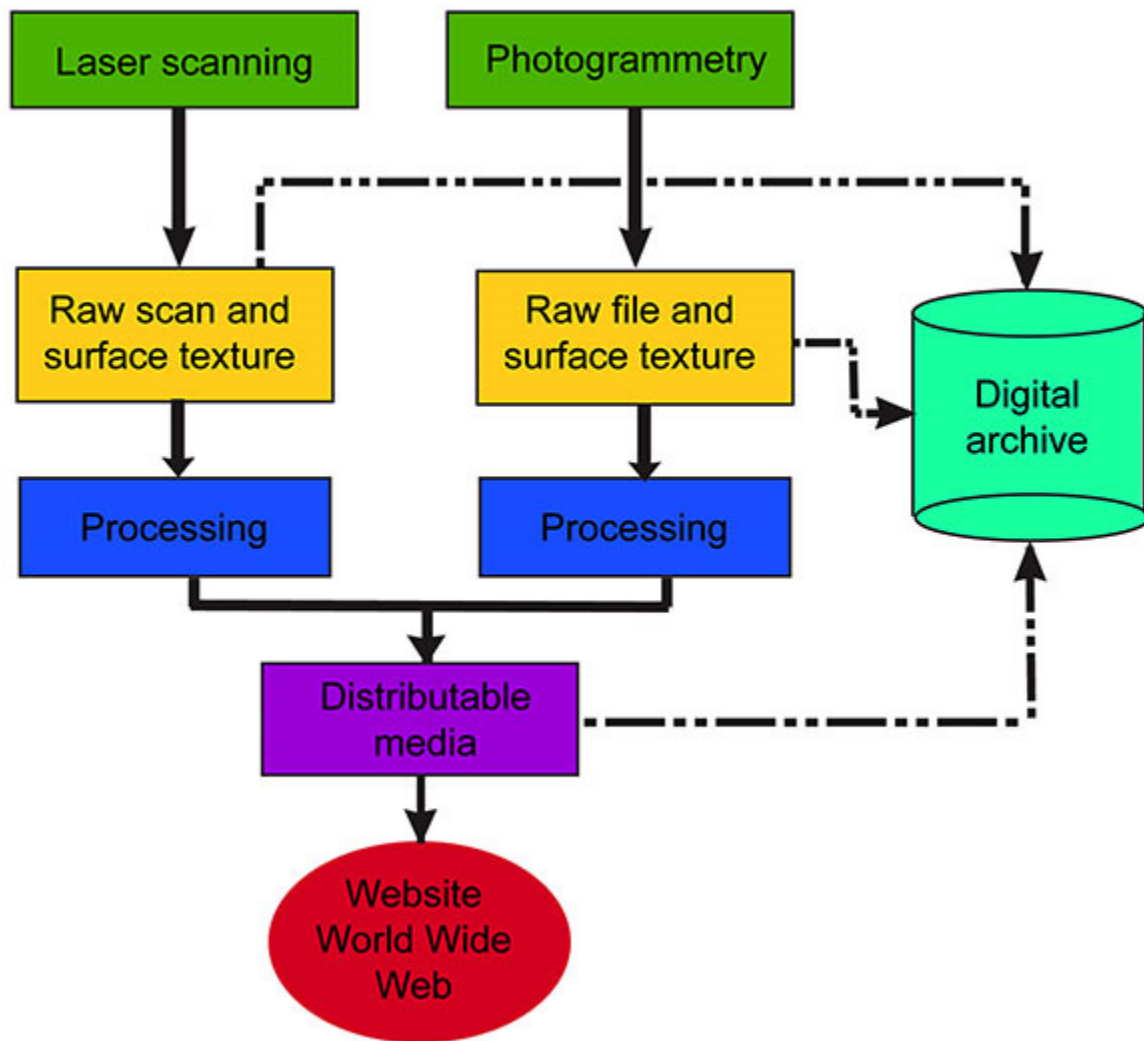
Three-dimensional models cover only an averaging 10% of specimens digitized during the project. This selection has been made taking into account: (1) the 21 months time available for the project completion; (2) recording no more than one specimen per species; (3) specific technical tools limitations. 3D models were acquired using the NextEngine™ HD Desktop 3D and ScanStudio™ HD PRO software (NextEngine™, 2008), and Agisoft PhotoScan PRO® 1.0 (Agisoft, 2006) photogrammetric software, both running on mid-high range Dell workstation and portable computer (Intel i7 processor, 8 GB RAM, Nvidia GeForce Graphic Card).

NextEngine™ HD scanner is a portable and economic (<3000 £) solution to laser scanning and it has been chosen as it represents a good compromise between price and performances. On macro mode (capturing field size 130x95 mm) the NextEngine™ HD has a maximum dimensional accuracy of 0.12 mm for 160K points/in² capture density and texture resolution of 400 DPI, whereas on wide mode (field size 340x255 mm) maximum dimensional accuracy is 0.38 mm for 22.5Kpoints/in² capture density and texture resolution is 150 DPI. Due to the large amount of data, limited post-processing time and in order to minimise errors, most specimens were acquired with a capture density of 17K points/in² (macro mode) and 2K points/in² (wide mode). Surface attributes setting (light-dark) had to be determined experimentally to give optimum colour fidelity and surface capture. Objects size have no present limit for the scanner, although objects smaller than 20 mm, or larger than scan field have been preferably digitized though photogrammetric techniques. All specimens requiring composite capturing were scanned using the NextEngine™ rotating platform; post-processing, alignment and merging into a single 3D model was performed

with ScanStudio™ HD PRO software, resulting in an average building time for finished 3D model of 1 hour.

Agisoft PhotoScan® is an image-based 3D modelling suite that allows creating accurate 3D models from digital pictures. Processing comprises four main stages: camera alignment, building sparse point cloud, building dense point cloud and building texture. Photos were taken with a 22 Mpx Canon EOS 5D camera; in order to have final outputs comparable to scanner results, building time for 3D model varied from 1 to 6 hours, depending of photos number. Photogrammetry is a relatively cheap and quick solution and it has been employed to integrate scanned 3D models when specimens size were too small (< 20 mm), too large (> 700 mm), or when inaccessible by NextEngine™ (e.g., specimens mounted on wall or sealed into cases).

The final 3D models were exported in Wavefront OBJ (Rose and Ramey, 1993) format that includes mesh, material and texture files (*.obj, *.mtl, *.jpg), and Polygon PLY binary format. Ply and obj formats have been chosen as final outputs because can be read by many computer-aided design (CAD) packages for visualisation and post processing. In particular: *.ply are light single files that comprise texture and are quick for visualisation proposes, whereas *.obj are delivered as compressed zip files, although larger than ply, obj are usually preferred for a number of 3D manipulation and post processes (videos and image rendering, animations, 3d printing). Raw ScanStudio™, Photoscan® files and processed 3D objects are archived within the British Geological Survey headquarters in Keyworth (Nottingham, U.K.) and are available for researchers enquiries (Fig.1).



Schematic flow chart illustrating the 3D digitizing process

References

- Agisoft LLC (2006) PhotoScan PRO, version 1.0.0. St. Petersburg, Russia (www.agisoft.ru).
 NextEngineTM, Inc. 2008. HD Desktop 3D scanner and ScanStudioTM. NextEngineTM, Incorporated. Santa Monica, California. (www.nextengine.com).
 Rose, L., and Ramey, D. 1993. Wavefront file formats, version 4.0 RG-10-004. Wavefront Technologies Inc., Santa Barbara, California.

10.1.2 The GB3D Fossil Types Online project (www.3d-fossils.ac.uk) – the inside story of assembling the world’s largest virtual fossil collection.

Howe, Mike P.A. National Geological Repository, British Geological Survey, Keyworth, Nottingham, UK. mhowe@bgs.ac.uk

McCormick, Tim. British Geological Survey, Keyworth, Nottingham, UK.

Harris, Simon J. National Geological Repository, British Geological Survey, Keyworth, Nottingham, UK.

Contessi, Michela. Museo Civico di Scienze Naturali, Faenza, Italy.

Since the origin of the British Geological Survey (BGS) in 1835, it has maintained collections of geological specimens, displayed first in the Museum of Economic Geology (1837 – 1851),

then the Museum of Practical Geology (1851 – 1935), and then the Geological Museum, South Kensington (1935 – 1985), before their transfer to Keyworth, Nottingham. In the Museum of Practical Geology (Jermyn Street, London), all the “crown jewels” were on display, but with each successive move, more specimens were confined to storage drawers.



Figure 1. Specimen-rich displays in the Museum of Practical Geology, Jermyn Street, London. Late 19th Century. BGS©NERC

The challenge to the collections management has always been to increase access and to ensure the collections are fully discoverable and available for research. The rapid development of the worldwide web in the late twentieth century provided the obvious mechanism, and the BGS launched an online GIS application with borehole and specimen data layers in 2000, together with a database of images. The BGS collections include over 30,000 type, figured and cited fossils, and these were an obvious priority.

Early attempts to raise grant funding for 3d digitisation failed, but the launch of the budget priced high resolution NextEngine laser scanner HD brought the technology within routine budgets and one was purchased in 2011. Early results underpinned a 21 month £691k grant from Jisc, which commenced in November 2011. The project partners (BGS; National Museum Wales, Cardiff; the Sedgwick Museum, Cambridge; the University Museum of Natural History, Oxford; and the Geological Curators' Group, representing a number of regional museums) were funded to produce an online database of their type British macrofossils, including new high resolution photographs and stereo anaglyphs, and approximately 10 – 20% of the 20,000 odd specimens with 3d digital models.

With less than one month's notice, resourcing proved problematic, particularly in view of the British Government recruitment restrictions then current. Other management challenges included the selection of formats and the establishment of standards and workflows, particularly difficult in the absence of any national and international consensus. Even two partners using the same commercial collections management databases had very different implementations and dictionaries. When dealing with organisations totalling over two and a half thousand years of combined history, the choice and selection of dictionaries or free text proved controversial. The timescale was particularly challenging. Laser scanning is a relatively slow process, and with the large number of specimens to scan during the project, the decision was taken to start scanning and photography immediately. In hindsight, we should perhaps have considered a short pilot project stage to test methods and procedures, and applied for a project extension. The progress of the project is documented in its blog (<http://gb3dtypefossils.blogspot.co.uk/>).

The screenshot shows the GB3D Type Fossils website interface. At the top, there is a navigation bar with links: Home, The project, Search, About fossils, Contributors, Downloads, Links, and Contact us. The main content area is titled "Fossil specimen : BRSUG 13487 - Holotype".

Taxonomy

Taxonomic Group	Kingdom Animalia
	Phylum Mollusca
	Class Cephalopoda
	Order Ammonoidea
Fossil Name	<i>Procerites twinloensis</i> Arkell
Type Status	Holotype
Name History	<i>Procerites twinloensis</i> Arkell

Provenance

Registration No.	13487
Collector	Fry, T.A. 1922
Donor or Associated Person	Name not available
Held By	University of Bristol Geology Collection
Comments	

Location

Country	England
Locality	Twinhoe Green Quarry, Nr. Wellow, Bath
Grid Reference	Not available

Stratigraphy

Rock Unit	Bathonian, Twinhoe ironshot Oolite
Biozone	Not available
Geological Age	Bathonian Age (Jurassic Period) (164.7 - 167.7 Ma B.P.)

References

References	Arkell, W. J. 1951-1958. English Bathonian Ammonites. Palaeontographical Society Monograph, 264 pp. Pls. 1-33
------------	---

High resolution photographs

Three high-resolution photographs of the fossil specimen are shown, including a stereoscopic pair and a single view.

Stereoscopic photographs

A stereoscopic pair of high-resolution photographs of the fossil specimen is shown.

3D scan

- [View scan in web browser](#) (Requires a web browser able to display HTML5 canvases. Internet Explorer 9+, Firefox, Opera, Chrome, and Safari support this. Note that Internet Explorer 8 and earlier versions do not support this functionality.)
- [Download zipped .ply file](#) (4961 kb)
- [Download zipped .obj file](#) (9381 kb)

Licence

This work is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License](#).

At the bottom of the page, there is a footer: "This site is hosted by the British Geological Society and copyright on the website is with the GB3D Type Fossils Online project partners. Responsibility for the content of the site lies with the GB3D Type Fossils Online project partners, not with the BGS."

Figure 2. Typical screen shot of an ammonite type within the GB3D fossil types database, including specimen metadata, thumbnail links to high resolution images and stereo-anaglyphs, and links to view or download the 3d digital models.

BGS@NERC



Figure 3. Screenshot of a high resolution stereo anaglyph. The image can be enlarged significantly further using the controls (top right). BGS©NERC

The project benefitted considerably from a number of recent developments not included in the original plan. The deployment of web browsers able to display HTML5 canvases enabled the inclusion of a 3d model viewer, and the rapid fall in price of 3d printers allowed the purchase of a MakerBot Replicator 2, which drew crowds whenever the project was demonstrated.

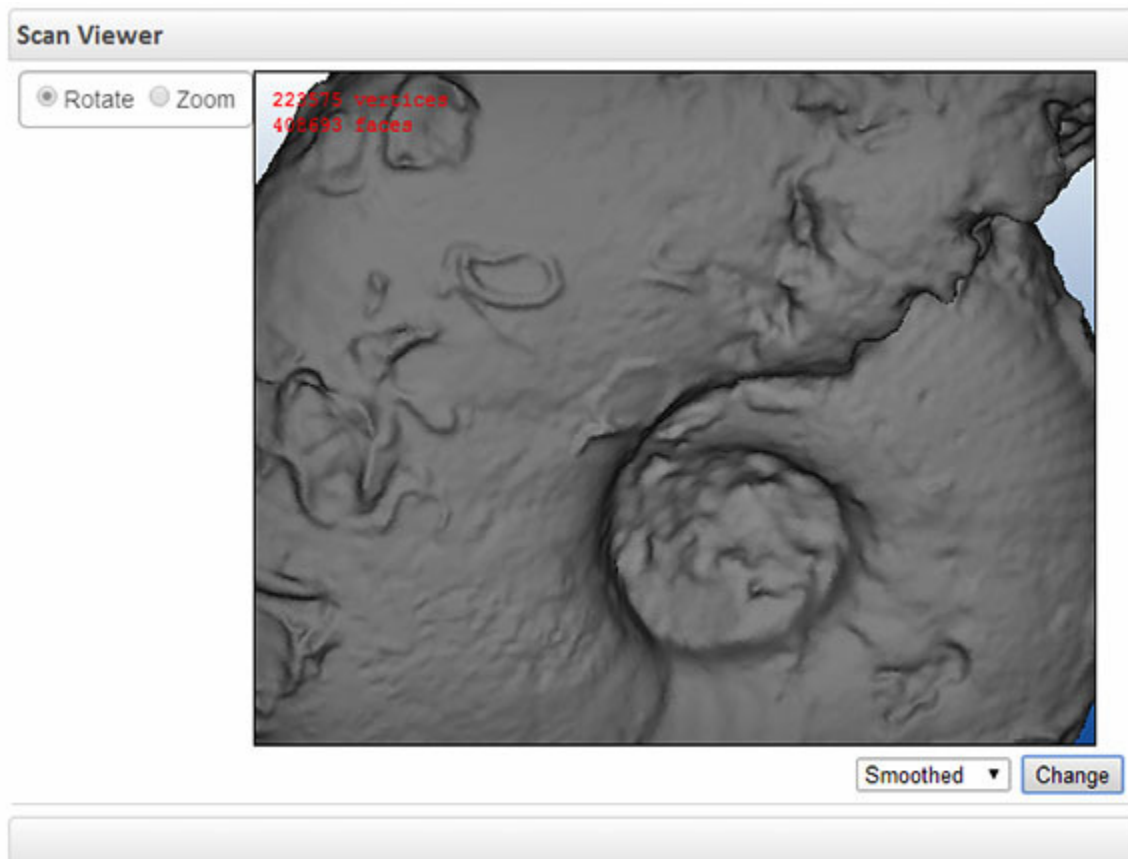


Figure 4. Screenshot of the webpage 3d digital model viewer. Can be set to point cloud, wireframe, smoothed, and textured, and the models can be zoomed and rotated. BGS©NERC

The project is now almost complete and the web portal (www.3d-fossils.ac.uk) is gaining popularity. Routine maintenance is provided by the UK National Geoscience Data Centre/National Geological Repository at BGS, and new specimens are continually being added. An API is undergoing beta-testing, and funding and partners are now being sought to extend the project internationally.

10.1.3 The 3D printer on the beach: Taking 3D fossils back into the field.

Harris, Simon J. National Geological Repository, British Geological Survey, Keyworth, Nottingham, UK. simhar@bgs.ac.uk

Howe, Mike P.A. National Geological Repository, British Geological Survey, Keyworth, Nottingham, UK.

McCormick, Tim. British Geological Survey, Keyworth, Nottingham, UK.

Contessi, Michela. Museo Civico di Scienze Naturali, Faenza, Italy.

“Fused-filament additive manufacturing”, or “3D printing” as it is more commonly known, is a much talked about technology amongst lovers of gadgets and all things technical. Media headlines regularly feature stories describing how the technology will one day be used to create everything from the houses we live in to replacements for our body organs.

Presented with an ever-growing library of 3D models and falling hardware prices, the Jisc funded 3D fossils team at BGS decided to purchase a relatively inexpensive desktop printer to see if 3D printing is the future for fossil replicas.

The team had the opportunity to experiment with the plethora of software available from across the internet, both open source and commercial, and to determine which models were most suited to the new manufacturing technology. Having completed the journey from physical object, through virtual model, and back to physical object again, we were able to compare finishing techniques with traditional cast finishing methods.

Finally we discuss the enormous outreach potential that the device created, and explain how we came to be operating our own printer (literally) a stone's throw from the English Channel!

10.1.4 Metadata extraction from digital specimen: a standardised software driven approach

Berendsohn, Walter G. Botanic Garden and Botanical Museum Berlin-Dahlem, Berlin, Germany. w.berendsohn@bgbm.org

Chaves-S., Fernando. Fraunhofer Institute of Optronics, System Technologies and Image Exploitation IOSB, Karlsruhe, Germany.

Kirchoff, A. Botanic Garden and Botanical Museum Berlin-Dahlem, Berlin, Germany.

Röpert, D. Botanic Garden and Botanical Museum Berlin-Dahlem, Berlin, Germany.

Steinke, K.-H. Hochschule Hannover University of Applied Sciences and Arts, Fachgebiet Ingenieurinformatik, Hannover, Germany.

Usländer, T. Fraunhofer Institute of Optronics, System Technologies and Image Exploitation IOSB, Karlsruhe, Germany.

On herbarium sheets, data like plant name, collection site, collector, barcode, accession number, etc. are found mostly on labels glued on the sheet. The data is thus visible on images taken of the specimen. Currently, they are mostly entered manually into collection databases. The StandAP-eHerb Project funded by the DFG (German Research Foundation) develops a standard process for (semi-) automatic detection of meta-data on Herbarium specimens to replace the time consuming manual data input as much as possible. Image processing software detects objects such as labels or barcodes on the digitized record and classifies them (fig.1, fig.2).



Fig. 1: Original image of a herbarium sheet

Text objects are transformed into structured information using text mining algorithms. For handwriting, author identification is attempted. The project evaluates and enhances existing software to comply with standard interfaces and integrate them into an open software architecture based on established IT standards. The software modules thus become available for work flow processing, in order to verify data quality, facilitate data discovery and enhance the application of collection data in research.

The project addresses a large proportion of scientific collections: approximately 22 million herbarium specimens exist as botanical reference objects in Germany, about 500 million worldwide.



Fig. 2: Object detection: different objects in the image

10.1.5 Image stacking: A semi-automatic approach allowing mass digitization of type specimens

Brecko, Jonathan. Royal Belgian Institute of Natural Sciences, Brussels, Belgium.
jonathan.brecko@naturalsciences.be

Mathys, Aurore. Royal Belgian Institute of Natural Sciences, Brussels, Belgium & Royal Museum for Middle Africa, Tervuren, Belgium.

Semal, Patrick. Royal Belgian Institute of Natural Sciences, Brussels, Belgium.

Scientific institutions like museums maintain large collections enabling present and future research. Thanks to the digitization of the collections, in most of these institutes, it enables researchers across the globe to see which collections might be interesting for their projects. However as most of these databases only provide descriptive information and/or metadata, it remains impossible to study these digitized specimens from a distance.

As the most precious parts of the collections, like type specimens, are the most requested ones by fellow researcher, these become, inevitably, the most handled ones. Unless a policy exists not to handle them at all, which unfortunately, prevents research. To make sure that collection material like type specimens remain documented for future research and can be studied through the internet, digitization is the key, 3D or 2D. The only challenge is to choose the right digitization method for the right material and or size (Mathys et al., 2013). Especially small specimens which are often found in insect and invertebrate collections tend to be difficult to digitize in 3D as fine structures can only be seen in μ CT recordings, which are still quite expensive pieces of equipment. However, 2D image might provide enough information to conduct for instance taxonomic research. Image stacking is the only way to capture enough detail in a single picture as the low depth of field of camera lenses, makes it almost impossible to get the complete object in focus, unless the aperture is stepped down. However this results in other aberrations as the optical resolution reduces due to the diffraction effect. Thanks to the large computational power of today's workstations, it is possible to do image stacking rather easily. The only remaining challenge is taking the individual pictures. Readily made commercial stacking columns do exist, but are too expensive to provide one to each department or research group. In this paper we will present the results of a low cost approach consisting of a DSLR camera attached to an automated macro rail with a custom build light tent. As there are no over- or underexposed parts on the resulting images, they are good enough to allow publication without the use of a post-processing software. More importantly, during the photo-shoot of one specimen another can be prepared, alcohol based or dry, for the next shoot. When similar specimens are digitized at the same time, light and aperture settings stay more or less the same, providing a fast and smooth workflow. The stacking of the images, done in Zerene Stacker, can easily be started as a batch process at the end of the working day or during the night. As the total package of this system can be purchased for around € 2.5k several digitization centers can be started in different department allowing faster digitization of the type material.



Figure 1: A stacked image of a museum beetle is an example of a small specimen pictured with our image stacking equipment.



Figure 2: A stacked image of a Venezuelan butterfly is an example of a medium sized specimen pictured with our image stacking equipment.

References

Mathys, A., Brecko, J. and P. Semal. 2013. Comparing 3D digitizing techniques: What are the differences? In proceeding of 2013 Digital Heritage International Congress, Marseilles, France, Vol 1.

10.1.6 Mobilising and making accessible The Australian National Collections, through mass digitisation in an interconnected digital world

Fisher, Nicole. CSIRO Ecosystem Sciences, Canberra, Australia. Nicole.fisher@csiro.au

Mass digitisation programs are rapidly transforming the landscape of data extraction from specimens in natural history collections, and facilitating a change in the way we think about, and manage, their specimens and data. CSIRO is the custodian of The Australian National Collections, which contribute to national and international biological knowledge. This presentation will focus on digitisation workflows, tools & techniques and explore the initiatives currently underway in the CSIRO National Collections. As well as challenges faced by the curator of collections, such as incorporating digitisation into the collection as part of standard practices, exemplifying the benefits in curatorial aspects and mobilising and making accessible these online digitised collections.

The CSIRO Australian National Insect Collection is the pre-eminent collection of Australian insects, maintained by CSIRO researchers, industry and government agencies. It is a leading figure in exploring new technologies to accelerate the rate of digitisation from collections, through development of digitisation workflows and whole drawer imaging. By building strength in an interconnected digital world thru collaborative efforts, an existing and important collaboration with iDigBio has led to the development of further initiatives based on the Whole-Drawer digitisation experience at the Australian National Insect Collection. This and other global initiatives that the CSIRO National Collection are involved in will be discussed.

10.1.7 The role of stable URIs for mass digitization approaches, strategy and workflows

Glöckler, Falko. Museum für Naturkunde - Leibniz Institute for Research on Evolution and Biodiversity, Berlin, Germany. Falko.Gloeckler@mfn-berlin.de

Kroupa, Alexander S. Museum für Naturkunde - Leibniz Institute for Research on Evolution and Biodiversity, Berlin, Germany.

Hagedorn, Gregor. Museum für Naturkunde - Leibniz Institute for Research on Evolution and Biodiversity, Berlin, Germany.

Schurian, Bernhard. Museum für Naturkunde - Leibniz Institute for Research on Evolution and Biodiversity, Berlin, Germany.

Since the dissemination of digital objects is a chance for natural history collections to open up the “hidden” treasures that exist behind the scenes of exhibitions the common method to present the digital representations is focused on web-based technology. In general, web-sites referenced by a web address are used to show images and metadata of these objects. But there is much more potential in that approach as we can understand the Unified

Resource Identifiers (URIs) not only as web address, but also as globally unique identifiers for the digital assets and even for the physical objects.

In biodiversity informatics a lot of discussions about object identifiers took place in the last two decades, but at the same time web technology established all services needed for the same demands.

In my talk I will present advantages of stable URIs in both physical and digital workflows which are closely related in (mass) digitization processes. I will especially focus on object handling, databasing and dereferencing by giving examples from experiences that have been made at the Museum für Naturkunde Berlin (MfN).

10.1.8 iDigBio: Facilitating Digitization on a National Scale

Nelson, Gil. Florida State University/iDigBio. gnelson@bio.fsu.edu

In 2011 the U.S. National Science Foundation (NSF) launched a 10-year, one-hundred-million-dollar initiative to digitize as many as one billion biological specimens held within approximately 1600 non-federal collections in the United States. NSF designated Integrated Digitized Biocollections (iDigBio), a collaborative between the University of Florida and Florida State University, as its national resource for facilitating and enabling digitization and data storage for the project. To date, 13 Thematic Collections Networks encompassing nearly 300 institutions and representing virtually every type of biodiversity collection are involved in the project. iDigBio has played a major role in providing an array of training and best practices workshops and encouraging collaboration and sharing among collections professionals throughout the United States. Part of this effort has resulted in the establishment of several national and international working groups, including the International Whole-drawer Digitization Interest Group (WDD), composed of representatives from Australia, Germany, The Netherlands, Sweden, the United Kingdom, and the United States. The group meets monthly to take up topics related to whole-drawer digitization of insects and fossils and hopes to build its membership to include representatives from any domain that stores collection objects in drawers or trays. This presentation will offer a brief history of iDigBio and a detailed overview of the activities of the WDD.

10.1.9 ZooSphere - Development of a software for automated spheric image capturing and interactive 3D visualization of biological collection objects

Pluta, Martin. Museum für Naturkunde - Leibniz Institute for Research on Evolution and Biodiversity, Berlin, Germany. martin.pluta@mfn-berlin.de

Glöckler, Falko. Museum für Naturkunde - Leibniz Institute for Research on Evolution and Biodiversity, Berlin, Germany. Falko.Gloeckler@mfn-berlin.de

Kroupa, Alexander S. Museum für Naturkunde - Leibniz Institute for Research on Evolution and Biodiversity, Berlin, Germany.

Schurian, Bernhard. Museum für Naturkunde - Leibniz Institute for Research on Evolution and Biodiversity, Berlin, Germany.

We developed a tool (ZooSphere) for automated image capturing from about 400 to 1500 positions around a specimen. A high resolution image with extremely high depth of focus is produced by using stacking technology in each position. The combination of all images is a spheric image sequence, ready for seamless animation.

The image sequences will be freely available in the World Wide Web and will be displayed by a web-based tool or a Java desktop application, both developed at the MfN Berlin. These applications allow to rotate, zoom and pan the view of the image sequences by mouse, keyboard and gesture control. The same software is used to create an interactive stereoscopic visualization (3D) of the object.

Future development includes joining of the ZooSphere images with micro-CT images of internal specimen structures and the creation of 3D models of the specimens by applying photogrammetric methods on the image sequences. For data exchange, a RESTful webservice will be implemented.

Web-Application: <http://zoosphere.biodiv.naturkundemuseum-berlin.de/>

10.1.10 High-performance individual-based digitisation of insect collections

Saarenmaa, Hannu. Digitalium, SIB labs, University of Eastern Finland, Joensuu, Finland.
www.digitalium.fi

Mononen, Tero. Digitalium, SIB labs, University of Eastern Finland, Joensuu, Finland.
www.digitalium.fi

Karppinen, Janne. Digitalium, SIB labs, University of Eastern Finland, Joensuu, Finland.
www.digitalium.fi

Sääskilähti, Mira. Digitalium, SIB labs, University of Eastern Finland, Joensuu, Finland.
www.digitalium.fi

Tegelberg, Riitta. Digitalium, SIB labs, University of Eastern Finland, Joensuu, Finland.
www.digitalium.fi riitta.tegelberg@uef.fi

We present a novel solution for mass-digitisation of individual pinned insects. We use a specially designed conveyer belt system, which was miniaturised from a big one used for herbarium sheets. On the imaging line, plastic pallets custom-built with 3D-printers carry the specimens to an imaging station. The insect is automatically imaged from above, and the labels attached to the pin are imaged from the side. Underside of the labels is visible through a mirror in the pallet. Each specimen is tagged with a unique ID and a web address. This basic metadata, which establishes the digital object also includes taxon name. Further metadata can be transcribed later from the images, and retrieved from the accession books which document the collection.

At the moment, imaging line is applied to objects between 0.5 – 30 mm in length, and the peak performance is around 500 specimens /day. An entire Coleoptera collection has already been processed this way, and work is starting on another about parasitic *Hymenoptera* and their hosts.

This system is being used to process collections donated to the Finnish Museum of Natural History. Digitalium receives collections, and disassembles them, driving them through the automatic digitisation line and placing them in units. After digitisation, the specimens and produced data will be delivered in taxonomical order to the museum.

10.1.11 *From specimens to drawers: mass digitisation efforts at the Bavarian State Collection of Zoology in Munich, Germany.*

Schmidt, Stefan. Bavarian State Collection of Zoology Munich, Germany.

hymenoptera@zsm.mwn.de

Spelda, Jörg. Bavarian State Collection of Zoology Munich, Germany.

Balke, Michael. Bavarian State Collection of Zoology Munich, Germany.

The Zoologische Staatssammlung in Munich (Bavarian State Collection of Zoology) is, with about 25 million objects, one of the top natural history collections, nationally and internationally. About 90% of the collection are insects, including 10 million Lepidoptera, 3-4 million Coleoptera, and about three million *Hymenoptera*, stored in more than 100,000 standard sized drawers (51 × 42 cm). The sheer amount of specimens deposited in natural history collections like the ZSM poses enormous challenges to any digitisation effort. To remedy this situation, the entomology section of the ZSM has been developing new technologies for the rapid digitisation of entomological collection drawers, and the subsequent extraction of relevant metadata from drawer images. The drawer scanner DScan is a prototype scanning machine and the foundation of a digitisation system that allows the fast and efficient digitisation of entomological (and other) drawers. Since its development started a few years ago, the focus has been on optimising the system for the requirements of on-demand-digitisation. Because the content of drawers and the arrangement of specimens in drawers changes through time, re-scanning of drawers needs to be as fast and as easy as possible, without sacrificing image quality. The resulting images should allow inspection of insect specimens at high resolution, thus minimizing the need for physical access to the collection. The level of image detail needs to be adjustable, e.g. in relation to the size of the specimens and the required level of detail. Our vision is to use state-of-the-art mass imaging technology for creating a web-accessible, natural history metacollection that provides access to tens of millions of museum specimens and makes them available for automated digital image analyses.

10.1.12 *Unlocking the vaults: the Natural History Museum London collections digitisation programme*

Smith, Vince. Natural History Museum London, UK. *vince@vsmith.info*

Achieving sustainable management of ecosystems depends upon the availability of accurate information about the variation of biodiversity across space and time. Natural history collections are a vital source of these data. These collections hold billions of specimens collected over three centuries, each witness to past ecological conditions and historic biogeographic distributions. This presentation will give an overview of the Natural History Museum London's collections digitisation programme, including ambitious plans to digitise 20 Million specimens over the next few year. This talk will cover the industrial scale digitisation activities we are undertaking to unlock the vaults of our collection and the tools we are developing to map, monitor and understand the natural world.

10.1.13 *Virtual collection of insects of E.-F. Germar – creation, curation and perspectives*

Holovachov, Oleksandr. Department of Zoology, Swedish Museum of Natural History, Box 50007, SE-104 05, Stockholm, Sweden. *oleksandr.holovachov@nrm.se*

Insects are dominant part of natural history collections and their digitization is currently one of the actively discussed topics among curators. There are a number of different digitizing methods and technologies, including the 'whole-drawer imaging', which offers very fast rate of digitization at a cost of collecting only part of visual information about specimens. Nonetheless, this technique proves to be extremely useful when digitizing old, historically and scientifically important, but uncatalogued and inaccessible collections, such as the collection of Hemiptera and Homoptera of E.-F. Germar, which is stored in the Zoological Museum of the Ivan Franko University of Lviv in Ukraine. Ten drawers include a total of 4500 specimens of over 1800 species collected from all over the world during the first half of the 19th century. About 10 % of the specimens are type or putative type material described by E.-F. Germar, G. Herrich-Schäffer, E. Taschenberg and other entomologists. Unfortunately, due to political reasons, this collection was inaccessible for the scientists during entire 20th century, and still remains difficult to work with, mainly due to the fact that specimens are not permitted to be mailed. Therefore, it was important to make this collection publicly accessible, which we did using our own modification of the whole-drawer imaging technique. Our setup included a consumer grade digital SLR camera fitted with a macro lens and attached to an immobile copy stand. Drawers with insects were moved manually along the pre-determined grid pattern, and all camera and lighting settings were set manually to ensure maximum consistency between the shots. Between 80 and 130 original photographs were stitched to create final ultrahigh resolution images, which are now displayed on the web site of the Zoological museum (http://zoomus.lviv.ua/en/germar_collection/) of the Ivan Franko National University of Lviv. As a result, scientists from all over the world can see the entire content of the collection, examine dorsal view of all specimens in details, assess their preservation quality and suitability for further research, verify the identification and read most of the label data. Experts who studied these images confirmed that these images often can provide sufficient amount morphological information. Whole drawer images also allow researchers to verify that specimens are correctly identified and that their preservation is satisfactory. This becomes increasingly important for the collections that lack databases or catalogues available online or in printed form, and especially for those collections, which are being curated by personnel without sufficient expert knowledge. Virtual collection of insects of E.-F. Germar made it easier for curators, who are not trained entomologists, to distribute new information to the scientific community. We think the future development of whole-drawer imaging technologies should not only focus on the increase of speed and effectiveness of the process itself, but also on its affordability to a broader customer base (scientifically and geographically) and flexibility of the system. Users must be able to design their own photographic system depending on the desired final quality and resolution of images; magnification and speed of the process; and the cost of the setup. Flexibility of whole-drawer imaging system would broaden their customer base and increase its use among museums, leading to faster and more efficient digitization of biological collections.

10.2 Abstracts der Konferenz Langzeitzugriff auf Sammlungs- und Multimediatdaten

10.2.1 Vorbereitungen der Langzeitarchivierung von AV-Materialien

Merle Friedrichsen. Technische Informationsbibliothek Hannover (TIB).

merle.friedrichsen@tib.uni-hannover.de

Die Weitergabe und der Import von Daten und Metadaten setzen eine gewissenhafte Vorbereitung voraus. Dabei stellen sich u.a. Fragen nach den vorhandenen Materialien, welche rechtlichen Bestimmungen damit verknüpft sind, wie diese die Konzeption des Imports beeinflussen und welche Metadaten erfasst werden sollen. Um diese Fragen im Vorfeld zu klären, sind eine Bestands- und Prozessanalyse notwendig. Am Beispiel von AV-Materialien im Bestand der TIB wird ein Überblick über die Themen gegeben, mit denen man sich im Vorfeld des Imports beschäftigen sollte. Diese reichen von den Abgabeformaten bis hin zur Zertifizierung.

10.2.2 buildm – Erarbeitung eines deskriptiven Metadatenschemas für digitale Gebäudeinformationen im EU-Projekt DURARK

Michael Panitz, Technische Informationsbibliothek Hannover (TIB). michael.panitz@tib.uni-hannover.de

Zum Auffinden gespeicherter oder gar archivierter Daten sind beschreibende Metadaten notwendig. Das von der EU im 7th Framework Programme geförderte Projekt „DURABLE ARchitectural Knowledge“ beschäftigt sich mit der Digitalen Archivierung dreidimensionaler Daten aus dem Bereich der Architektur. Da in dieser Domäne keine beschreibenden Metadaten vorliegen, wurde innerhalb des Projekts auf verwandten Metadatenschemata basierend das Schema „buildm“ entwickelt. Der Vortrag beschreibt den Weg von den ersten Ideen, über tiefer gehende Analysen bis hin zu der ersten veröffentlichten Version des Metadatensatzes „buildm“.

10.2.3 Von Metadata zu Big Data (Praxisbericht)

Denir Velagic, SVA GmbH. denir.velagic@sva.de

Langzeitarchivierung stellt Gedächtnisinstitutionen bei der Aufbewahrung digitalisierter Daten und Dokumente vor viele Herausforderungen. Das Archivgut muss sicher, unverändert und wiederauffindbar aufbewahrt werden. Dabei ergeben sich z. B. im Zusammenhang mit sog. „Legacy Systems“ regelmäßig Herausforderungen bzgl. der Datenmengen. Oft fehlt es an geeigneten Suchmöglichkeiten im Umgang mit Meta- und Volltextdaten. Wir werden darlegen, dass bei diesen Massendaten der Einsatz moderner Technologien aus den Bereichen der Suchmaschinen, des automatisierten Metadatenmanagements und portablen Datenstandards der richtige Ansatz ist, die Aufgaben zu bewältigen. Wir erläutern wie Metadaten als Basis der komfortablen, intelligenten und effektiven Recherchemöglichkeiten genutzt werden können.

10.2.4 AnnoSys-Annotationsystem und -Repository unterstützen die Verwaltung und Pflege von Sammlungs- und Multimediatdaten

Wolf-Henning Kusber, Lutz Suhrbier, Anton Güntsch & Walter G. Berendsohn.
Botanischer Garten und Botanisches Museum Dahlem, Freie Universität Berlin.
w.h.kusber@bgbm.org

Das DFG-geförderte Projekt AnnoSys (<https://annosys.bgbm.fu-berlin.de/>) entwickelt Verfahren zur softwaregestützten Annotation sammlungsbezogener Metadaten. In der aktuellen Ausbaustufe sind naturwissenschaftliche Sammlungsdaten annotierbar, die über den Biological Collection Access Service (BioCAsE) im Internet publiziert sind. AnnoSys

besteht aus einem Nutzerinterface sowohl für Datenportalnutzer als auch für Kuratoren, einem Datenrepositorium zur langfristigen Archivierung von Annotationen und den dazugehörigen originalen Sammlungsdaten sowie einem abonnierbaren Benachrichtigungssystem (Email). Das Repositorium basiert auf einem generischen Datenmodell, welches neben ABCD für eine Vielzahl weiterer Anwendungsgebiete und Datenschemas adaptierbar ist, und wurde anhand des W3C Open Annotation-Standard implementiert.

10.2.5 Publishing digital representations of the collection objects via stable URIs

Falko Glöckler. Museum für Naturkunde Berlin. falko.gloeckler@mfn-berlin.de

In natural history museums it has been generally understood that digitization of collections is a modern way to inventory, manage, analyze, present and share collection objects. Current ambitions in collection digitization tend towards full digital data management. This is expected to be the most efficient way to handle huge amounts of records which will be processed in very different workflows in collection management systems. Nevertheless, the connection to the physical object must not get lost.

In the Museum für Naturkunde Berlin (MfN) we pursue the strategy to reference both our physical and digital objects with stable URIs in our digitization efforts and to create physical interfaces to the digital representation of those objects by using QR-code labels. In parallel, these identifiers can be used for publishing our digital collection on the internet and for providing information for the semantic web.

In my talk I will present this generic approach and the included concepts for the digital collection management and data management with focus on long term availability and publication.

10.2.6 Wissenschaftliche Daten aufwerten – nachhaltige Sammlungsintegration von biologischen Daten parallel zum Forschungsprozess

Tilo Henning, Norbert Kilian & Patrick Plitzner. Botanischer Garten und Botanisches Museum Berlin-Dahlem, Freie Universität Berlin. t.henning@bgbm.org

Die Dokumentation von Belegen wird hauptsächlich sammlungsseitig thematisiert, aber viele Sammlungsobjekte sind zunächst Untersuchungsobjekte im Forschungsprozess. Auf der Basis der EDIT Plattform haben wir deshalb eine Lösung entwickelt, die es erleichtern soll, bereits im Forschungsprozess Untersuchungsobjekte in adäquat dokumentierte Sammlungsobjekte zu transformieren. Sie vereinigt drei Funktionalitäten: Ein Datenrepositorium mit Export-Services, einen Editor zur Erschließung, Verknüpfung und Verwaltung der Daten heterogener Sammlungsobjekte (Belege, Bilder, Primärdaten-Files) sowie ein Informationsportal zur Darstellung und Bereitstellung der Daten.

10.2.7 Kooperative Erfassung und Verwaltung von Daten in virtuellen Netzwerken

Dominika Pienkos¹, Sebastian Klarmann¹ & Frank Dührkohp². ¹Programmfabrik GmbH, ²Verbundzentrale des Gemeinsamen Bibliotheksverbundes (VZG).
dominika.pienkos@programmfabrik.de

In einer vielfältigen Museumslandschaft ist man mit unterschiedlichen Anforderungen an die Objekterschließung und die Datenstrukturen konfrontiert. Die Lösung einer zentralen Erfassung und Verwaltung von Daten ermöglicht es, den Inventarisierungsgrad von Sammlungen voranzubringen und bietet auch kleineren Einrichtungen die Möglichkeit der musealen Objektverwaltung auf hohem technischen Niveau. kuniweb und KENOM sind zwei exemplarische Projekte der Verbundzentrale des Gemeinsamen Bibliotheksverbundes (VZG), die in Kooperation mit der Programmfabrik erfolgreich durchgeführt und bisweilen verstetigt werden konnten. Die webbasierte Objekterfassung wurde technisch mit easydb realisiert und wird im Rahmen einer Verbundinventarisierung von den beteiligten Partnern zur musealen Objekterfassung eingesetzt. Das flexible Datenmodell ermöglichte die Abbildung von Daten unterschiedlicher Einrichtungen und die Berücksichtigung von bereits bestehenden Strukturen in einem kollaborativen Netzwerk. Über Portale sind der Zugang und die Verfügbarkeit der wissenschaftlich aufbereiteten Inhalte sichergestellt.

10.2.8 Effektives Medien- und Sammlungsmanagement im Einklang - Praxisanwendung im Museum für Kommunikation

Günther Gromke. CDS Gromke e.K. gg@cds-gromke.com

Für die Verwaltung Ihres Sammlungsbestandes verwendet die "Museumsstiftung Post und Telekommunikation" mit ihren vier Hauptstandorten in Berlin, Frankfurt, Bonn und Nürnberg die Software MuseumPlus. Mit der Zeit ist der Anspruch an die Verwaltung, Bereitstellung und einfache Nutzung von Mediendaten stark angestiegen. Vor diesem Hintergrund wurde das Digital Asset Management System Cumulus mit Hilfe der Firma CDS Gromke e.K. eingeführt. Über eine mehrstufige Integration beider Systeme können seitdem die Mitarbeiter der Sammlungen die Vorteile beider Lösungen nutzen.

Im Vortrag wird der entstandene Mehrwert dieser Integration vorgestellt und es wird dabei vor allem auf die Vielzahl an Möglichkeiten des DAM Systems in den Bereichen Kommunikation und Weiterverwendung von Mediendaten eingegangen.

10.2.9 Digital und analog – Strategien zur nachhaltigen Bewahrung und Verfügbarmachung des Bergbauerbes im Deutschen Bergbau-Museum Bochum

Stefan Przigoda. Deutsches Bergbau-Museum Bochum/Montanhistorisches Dokumentationszentrum (montan.dok). stefan.przigoda@bergbaumuseum.de

Das Deutsche Bergbau-Museum bewahrt in seinem Montanhistorischen Dokumentationszentrum (montan.dok) einen einzigartigen Fundus an dinglichen, schriftlichen und (audio-) visuellen Kulturgütern zur Geschichte des deutschen Bergbaus. Kennzeichnend ist die enge Verzahnung zwischen den klassischen Dokumentationssparten Sammlung, Archiv und Bibliothek bei Erschließung und Verfügbarmachung dieser Dokumente. Welche Konsequenzen ergeben sich daraus für die Verwaltung, Pflege, Weitergabe und vor allem für die Langzeitarchivierung der verschiedenen sammlungsbezogenen Daten? Der Vortrag skizziert aus Sicht des montan.dok als einer mittelgroßen Einrichtung die Herausforderungen sowie praktische und strategische Lösungsansätze für die nachhaltige Bewahrung und Nutzbarkeit des digitalen Kulturerbes.

10.2.10 *Nutzungsorientierte Erschließung von Sammlungen am Beispiel des Tierstimmenarchivs des Museums für Naturkunde Berlin*

Karl-Heinz Frommolt. Museum für Naturkunde Berlin. karl-heinz.frommolt@mfn-berlin.de

Bei der Erschließung von Sammlungen wird man in der Regel mit dem Problem konfrontiert, dass riesige Datenbestände zu erfassen sind und Prioritäten für die Sicherung von Sammlungsbeständen und deren Dokumentation zu setzen sind. Das Tierstimmenarchiv des MfN Berlin gehört mit mehr als 120000 Tonaufzeichnungen weltweit zu den größten Tierstimmensammlungen. Im Verlaufe von 10 Jahren konnten bereits 85% der analogen Tonträger gesichert werden. Hier wird eine Erschließungsstrategie vorgestellt, die darauf abzielt, Digitalisate und Annotation möglichst schnell für wissenschaftliche Fragestellungen zugänglich zu machen. Dies beinhaltet die komplette Digitalisierung der Audioaufzeichnung, minimale Dokumentation zu Übersichtszwecken und detaillierte Bearbeitung und Dokumentation „on demand“.

10.2.11 *KISS - Keep IT Still Stored*

Wolfram Fritzsich. Museum für Naturkunde Berlin. wolfram.fritzsich@mfn-berlin.de

Mit der zunehmenden Digitalisierung von Sammlungsobjekten entstehen langfristige, aber gleichzeitig auch sehr unterschiedliche Anforderungen an die IT-Infrastrukturen von Institutionen wie das Museum für Naturkunde Berlin. In diesem Vortrag wird aus technischer Sicht – beginnend mit der historischen Entwicklung des SAN (Storage-Area-Network bzw. Speichernetzwerk) im Jahre 2011 bis heute – von den Erfahrungen im Bereich der Speicherung von Daten des MfN berichten. Es wird ein Überblick zu verschiedenen Methoden gegeben, die Speicherkapazitäten zu erweitern und neue Ablagesysteme in ein Netzwerk bereitzustellen. Des Weiteren werden in einem Vergleich zwischen lokalen und externen Diensten die wichtigen Anforderungen der Dienstdefinition für Langzeitzugriff und -speicherung am MfN dargestellt.

10.2.12 *OAIS-konforme Langzeitarchivierung von Multimediadaten*

Alexander Herschung, startext GmbH. Alexander.Herschung@startext.de

Langzeitarchivierung digitaler Inhalte erfordert mehr als nur reine Speicherung. Dauerhafte Zugänglichkeit und Verfügbarkeit erfordern andauernde Anstrengungen. Das OAIS Referenzmodell stellt einen etablierten Rahmen für die hierfür erforderlichen Maßnahmen zur Verfügung.

Der Einsatz von OAIS-konformer Software hat bisher in der Regel Projektcharakter und ist dementsprechend teuer.

Die startext GmbH hat nun auf Basis eines kostenfreien OpenSource-Speichersystems ein OAIS-konformes Softwaremodul entwickelt. Dieses richtet sich als ready-to-install/ready-to-use Paket an solche kulturbewahrenden Institutionen, die jetzt eine kostengünstige, handhabbare und betriebsbereite Lösung wünschen.

10.2.13 *APPLAUSE - Digitalisierung von Astronomischen Photoplatten*

Harry Enke. Leibniz-Institut für Astrophysik Potsdam. henke@aip.de

Das Leibniz-Institut für Astrophysik Potsdam (AIP) digitalisiert - im Verbund mit der Hamburger Sternwarte und der Dr. Remeis-Sternwarte in Bamberg - Archiv-Bestände historischer astronomischer Photoplatten, um diese als Kulturgüter zu erhalten. Die Langzeitreihen astronomischer Beobachtungen liefern zudem wertvolle Informationen, die mithilfe moderner astronomischer Datenverarbeitungsmethoden gesichert werden können.

Über das Web-Archiv APPLAUSE (Archives of Photographic PLates for Astronomical USE), das vom AIP gehostet wird, sind die Daten frei zugänglich und mit einer CC0-Lizenz abrufbar.

Der erste Datenrelease publiziert insgesamt 17 Terabyte an Datenmaterial, das über Datenbankkataloge effizient zugänglich ist. Die digitalisierten Platten erfassen 98,9 Prozent des gesamten Himmels.

Es werden die Verfahren zur Erstellung der Kataloge und deren Publikation vorgestellt.

10.2.14 Cross-Domain-Verfügbarkeit von biologischen Sammlungs- und Multimediatdaten in Europeana

Gisela Baumann, Wolf-Henning Kusber, Jörg Holetschek, Anton Güntsch & Walter G. Berendsohn. Botanischer Garten und Botanisches Museum Berlin-Dahlem, Freie Universität Berlin. g.baumann@bgbm.org

Europeana verknüpft digitale Objekte aus europäischen Museen, Bibliotheken und Archiven, um sie für verschiedenste Nutzergruppen zu erschließen. Die Zentraleinrichtung Botanischer Garten und Botanisches Museum Berlin-Dahlem (BGBM) ist verantwortlich für die Anbindung biologischer Sammlungsdaten mithilfe des Open Up!-Workflows und des Biological Collection Access Service (BioCAsE).

Das BGBM wird für Europeana weitere naturhistorische Sammlungen für die Einbindung von Multimediatdaten in eine Digitale Service Infrastruktur (DSI) gewinnen, das BioCAsE-Netzwerk stärken sowie Rechte- und Lizenzfragen für eine langfristig gesicherte Datenbereitstellung klären. Durch eine verbesserte Verfügbarkeit entstehen für Bürger, Institutionen und Kulturschaffende ungeahnte Möglichkeiten, naturkundliche Objekte zu erforschen und kreativ in einen neuen Kontext zu stellen.

10.2.15 Das Biodiversitätsnetzwerk des Humboldt-Rings (BiNHum): Ein gemeinsames Portal für naturkundliche Sammlungen

Joachim Holstein¹, Peter Grobe², Anton Güntsch³, Hubert Höfer⁴ & Dagmar Triebel⁵.

¹Staatliches Museum für Naturkunde Stuttgart, ²Zoologisches Forschungsinstitut und Museum Alexander Koenig, ³Botanischer Garten und Botanisches Museum Berlin Dahlem, Freie Universität Berlin, ⁴Staatliches Museum für Naturkunde Karlsruhe, ⁵Staatliche Naturwissenschaftliche Sammlungen Bayerns. joachim.holstein@smns-bw.de

BiNHum (<http://wiki.binhum.net>) besteht aus 5 naturkundlichen Forschungssammlungen, die im Humboldt-Ring kooperieren: SMNK, SMNS, ZFMK, SNSB & BGBM.

Das DFG-Projekt hat eine Laufzeit von 3 Jahren und beinhaltet:

- Datenrettung und -mobilisierung (SMNS, SMNK & Uni Ulm)
- Einführung neuer Technologien zur Datenrecherche und Entwicklung des Portals (ZFMK)

- Daten-Harvesting und Qualitätssicherung (BGBM)
- Bereitstellung und Anpassung der Diversity Workbench (DWB) als virtuelle Umgebung für BiNHum (SNSB).

Das BiNHum-Datenportal bietet umfangreiche Metadaten sowie Bilder und Audio-Files (<http://demo.binhum.net/>). Es enthält aktuell 2 Mio. Einträge naturkundlicher Sammlungsobjekte.

10.2.16 Die Jena Collection of X-Ray-Movies – Von der Forschung zur Sammlung zum Onlinezugang

Bettina Hesse¹, Rommy Petersohn¹, Carsten Senf², Klaus Küspert¹, Michael Lörzer³ & Martin S. Fischer¹. ¹Friedrich-Schiller-Universität Jena, ²Inverso GmbH & ³Thüringer Universitäts- und Landesbibliothek Jena. bettina.hesse@uni-jena.de

Seit 1993 wird die Funktionsmorphologie und Evolution tierischer Bewegungsabläufe am Institut für Spezielle Zoologie und Evolutionsbiologie mit Hilfe der Röntgenvideographie untersucht. Die in dieser Zeit entstandenen über 50.000 Filme wurden 2014 zur ‚Jena Collection of X-Ray-Movies‘ zusammengefasst. Sie sollen nun systematisch in einer maßgeschneiderten Datenbank erfasst und vor allem der Öffentlichkeit zugänglich gemacht werden. Mangels Infrastrukturen zur Forschungsdatensicherung und –veröffentlichung war es ein großer Glücksfall, dass diese Ziele nun im Rahmen einer Bachelorarbeit und in Zusammenarbeit mit der Thüringer Universitäts- und Landesbibliothek angegangen werden konnten.

10.2.17 GFBio-Fachdatenzentren an deutschen Sammlungen – Langzeitzugriff auf biologische Forschungsdaten

Dagmar Triebel¹, Peter Grobe², Anton Güntsch³, Gregor Hagedorn⁴, Joachim Holstein⁵, Carola Söhngen⁶, Claus Weiland⁷, Tanja Weibulat¹. ¹Staatliche Naturwissenschaftliche Sammlungen Bayerns, IT-Center, ²Zoologisches Forschungsinstitut und Museum Alexander Koenig, ³Botanischer Garten und Botanisches Museum Berlin-Dahlem, Freie Universität Berlin, ⁴Museum für Naturkunde Berlin, ⁵Staatliches Museum für Naturkunde Stuttgart, ⁶Leibniz-Institut DSMZ-Deutsche Sammlung von Mikroorganismen und Zellkulturen, ⁷Senckenberg Gesellschaft für Naturforschung. triebel@bsm.mwn.de

Das DFG-Infrastrukturprojekt „German Federation for Biological Data (GFBio)“ bringt eine Reihe von Akteuren aus dem Bereich Kuration biologischer Daten, Informationsmanagement und Forschung zusammen. Unter den 18 Partnern des Konsortiums spielen die sieben beteiligten naturwissenschaftlichen Sammlungen eine zentrale Rolle. Sie bilden zusammen mit PANGAEA und dem „European Nucleotide Archive“ den Verbund zur Langzeitkuration, Archivierung und Bereitstellung von Daten und digitaler Information aus DFG-Forschungsprojekten. Der Vortrag erläutert Charakteristika und Spezialisierung der Fachdatenzentren an Sammlungen. Ihre angestrebte Rolle als nach ISO Standard anerkannte domänenspezifische Datenzentren und digitale Archive wird beschrieben.

10.2.18 *Towards developing an ontology-controlled engine for semantic content management applications*

Lars Vogt¹, Roman Baum², Philipp Bhatt², Peter Grobe², Christian Köhler² & Sandra Meid².

¹Institut für Evolutionsbiologie und Zooökologie, Universität Bonn & ²Zoologisches Forschungsinstitut und Museum Alexander Koenig. lars.m.vogt@googlemail.com

An ontology is commonly used for representing knowledge. Here, we present our idea to use it for controlling a self-describing semantic content management application. A set of ontologies function as an ontology-based language that is used to describe templates and control the engine (i.e. middleware, Java libraries) that coordinates the overall operation of the application. With this language we define all actions and functions and thus control the behavior of the application, including the basic composition of its graphical user interface. We present examples of how the language is used to describe and therewith define basic actions for a tuplestore knowledge base application like for instance Semantic•Morph•D•Base (emorphology.morphdbase.de).

10.2.19 *Entwicklung von Standards zur Digitalisierung flüssigkonservierter Arthropoden Difa*

Joachim Holstein & Axel Steiner. Staatliches Museum für Naturkunde Stuttgart.
joachim.holstein@smns-bw.de

Das DFG-Projekt hat eine Laufzeit von 3 Jahren (Beginn 1.9.2014) und beinhaltet:

- Tests handelsüblicher Digitalkameras mit Makroobjektiven und entsprechendem Zubehör als Alternative zu kostspieligen Komplettsystemen
- Standardisierung der Aufnahmebedingungen von nasskonservierten Sammlungsobjekten (Fixierung in der Flüssigkeit, standardisierte Ansichten für die jeweiligen Gruppen, Beleuchtung)
- Entwicklung der Workflows zur Vorbereitung und Prozessierung der Objekte
- Möglichkeiten und Grenzen von 3D-Visualisierungen
- Datenfluss und nachhaltige Archivierung sowie Einbindung in Forschungsnetzwerke
- Erstellung eines „Manuals for best practice“ zur Digitalisierung flüssigkonservierter Arthropoden.

10.2.20 *Das Medien-Repository der Humboldt-Universität zu Berlin*

Frank Kühnlenz & Katarzyna Biernacka. Humboldt-Universität zu Berlin.
frank.kuehnlenz@cms.hu-berlin.de

Das Medien-Repository der HU-Berlin richtet sich an Mitarbeiter, die damit ihre Arbeitsgruppen beim Speichern und Organisieren von Mediendateien unterstützen. Web-offene Präsentationen sind ebenso möglich wie internes, kollaboratives. Das Medien-Repository ist organisiert in Projekte, die voneinander autark ein bestimmtes Thema haben. Projekte werden vom Computer- und Medienservice der HU-Berlin als Betreiber eingerichtet und danach vom auftraggebenden HU-Mitarbeiter selbständig verwaltet. Das Medien-Repository basiert auf dem webgestützten, open-source Digital-Asset-Management-System ResourceSpace. Es wurde erweitert, um beispielsweise Mandantenfähigkeit, HTTP-Streaming von Audio- und Videodateien per HTML5, sowie die direkte Verlinkbarkeit von Mediendateien zur Einbettung in andere Systeme.

10.2.21 *Preservation Workflow Architecture for Research and Cultural Heritage Data*

Marco Klindt & Kilian Amrhein. Zuse-Institut Berlin (ZIB). klindt@zib.de

Das Poster motiviert die Architektur und das zugrundeliegende Datenmodell zur OAIS konformen Langzeitarchivierung von digitalen Objekten. Der beschriebene Workflow ist agnostisch hinsichtlich der Herkunft von Daten und verarbeitet sowohl Forschungs- und Kulturerbedaten. Die Implementierung integriert die Open-Source Werkzeuge Archivemata, Fedora/Islandora und iRods.

10.2.22 *BacDive - Mikrobielle Diversität digital verfügbar machen*

Adam Podstawka, Carola Söhngen, Dorothea Gleim, Boyke Bunk & Jörg Overmann. Leibniz-Institut DSMZ-Deutsche Sammlung von Mikroorganismen und Zellkulturen. adam.podstawka@dsmz.de

Obwohl in der Regel unsichtbar für das menschliche Auge sind Bakterien und Archaeen allgegenwärtig in Boden, Wasser und sogar in unserem Körper. Dennoch im Vergleich zu ihrer großen Artenvielfalt, ist die Informationen über ihre Funktionen und Interaktionen mit höheren Taxa sowie ihre Bedeutung im Ökosystem noch weitgehend unerforscht. BacDive - The Bacterial Diversity Metadatabase (<http://bacdive.dsmz.de/>) bietet biodiversitätsbezogene Informationen zu Mikroorganismen, wie Sammlungsbeschreibungen, Forschungsdaten in Verbindung mit den jeweiligen Resultaten aus Genom- und Metagenomanalysen. Ziel ist die Mobilisierung, Digitalisierung, Bündelung und Bereitstellung dieser Daten sowie deren langfristige digitale Verfügbarkeit zu sichern.

10.2.23 *Coding da Vinci - der Kulturhackathon*

Anja Müller¹, Helene Hahn², Stephan Bartholmei³ & Barbara Fischer⁴. ¹Servicestelle Digitalisierung Berlin - digis, ²OpenKnowledge Foundation Deutschland, ³Deutsche Digitale Bibliothek & ⁴Wikimedia Deutschland.

Digitalisiert - und nun?

Coding da Vinci ist der erste Kultur-Hackathon (Entwicklertag) in Deutschland, der Entwickler/innen, Designer/innen und Gamer/innen zusammenbringt, um in Kooperation mit Kultureinrichtungen aus offenen Daten und eigener Kreativität neue Anwendungen, mobile Apps, Spiele und Visualisierungen umzusetzen.

Ziel von Coding da Vinci ist neben dem Etablieren und Vernetzen einer technikaffinen und kulturbegeisterten Community, auch das kreativen Ausschöpfen der technischen Möglichkeiten, die in unserem digitalen Kulturerbe stecken. Diese Möglichkeiten können neue Erlebnisformen schaffen, die Kulturgüter interaktiver und attraktiver machen. Wir setzen uns deswegen für die freie Verfügbarkeit und Nutzbarkeit von Kulturdaten ein und stellen sicher, dass sie den kreativen Menschen als Rohmaterial für ihre Ideen zur Verfügung stehen.

Coding da Vinci hatte ein fulminantes Auftaktwochenende im April. Die Ergebnisse werden am 5. Juli 2015 öffentlich im Jüdischen Museum Berlin präsentiert werden.

10.2.24 *digiS - Servicestelle und Förderprogramm Digitalisierung Berlin*

Anja Müller, Kilian Amrhein & Marco Klindt. Servicestelle Digitalisierung Berlin – digiS.
anja.mueller@zib.de

Die Servicestelle Digitalisierung Berlin (digiS) koordiniert seit 2012 das landesweite Förderprogramm zur Digitalisierung von Objekten des kulturellen Erbes. digiS entwickelt gemeinsam mit weiteren Projektpartnern am Zuse-Institut Berlin (ZIB) technische Services zur Sicherung der Langzeitverfügbarkeit der Daten auf. digiS ist ausserdem Mitveranstalter des Kulturhackathons "Coding da Vinci". Wer sich über das Berliner Förderprogramm und die besonderen Herausforderungen und Mehrwerte der Verfügbarmachung von Kulturdaten informieren möchte, ist herzlich an unsere Posterstände eingeladen.

10.2.25 *IDES: Making fish specimen data and images from natural history collections publicly accessible*

Tanja Weibulat¹, Martin Ebert², Martina Kölbl-Ebert², Markus Moser³, Dieter Neubacher¹, Dirk Neumann⁴, Henriette Obermaier⁵, Wolfgang Reichert¹, Veronica Sanz¹, Markus Weiss¹ & **Dagmar Triebel**¹. ¹IT Center of the Bavarian Natural History Collections, ²Jura-Museum Eichstätt, ³Bayerische Staatssammlung für Paläontologie und Geologie, ⁴Zoologische Staatssammlung München, ⁵Staatssammlung für Anthropologie und Paläoanatomie.
trieb@bsm.mwn.de

The project IDES (<http://ides.snsb.info>) is running a high performing distributed information infrastructure. It focuses on fossil and recent fishes from Europe and adjacent seas. The natural history collection objects like dried collections of skeletons, wet collections, micro- and macrofossils, DNA and tissue collections, otoliths and fish scales are photographed using specialised imaging stations. All images are managed together with the contextual data in a Diversity Workbench (DWB) network. New working processes are established and DWB client applications are adapted to IDES requirements. Until now, more than 28,000 data and 49,000 images are managed. The majority of data are already accessible via the GBIF and BiNHum portals. In future, the data will be available via GFBio (<http://www.gfbio.org>).

10.2.26 *The BioCAsE Provider Software for Publishing Biodiversity Data*

Jörg Holetschek, Anton Güntsch & Walter G. Berendsohn. Botanischer Garten und Botanisches Museum Dahlem, Freie Universität Berlin. j.holetschek@bgbm.org

International networks such as the Biological Collection Access Service (www.biocase.org) and the Global Biodiversity Information Facility (www.gbif.org) share a vision of free and open access to the world's primary biodiversity data, linking together natural history and species occurrence data from a large number of databases worldwide. They offer data portals for searching and browsing this unified data pool and web services that allow for a selective retrieval of data subsets.

The BioCAsE Provider Software can be used to connect data to these networks. It exposes occurrence records as a web service that can be queried through the BioCAsE protocol and that returns ABCD documents (Access to Biological Collections Data, an XML schema for storing rich occurrence data). Once published with BioCAsE, the data can be shared with any network that supports BioCAsE providers – from global networks like GBIF and the

GGBN (Global Genome Biodiversity Network) to smaller, special interest networks like the German Virtual Herbarium VH/de or BiNHum (Biodiversitätsnetzwerk des Humboldt-Rings) – or aggregated first and then fed into networks not supporting BioCASE, like the European Virtual Library, Europeana.

10.2.27 *ZooSphere – globalizing taxonomic sharing of insects in 3D*

Christina Kuhlmann, Felix Maier, Martin Pluta, Bernhard Schurian & Alexander S. Kroupa
Museum für Naturkunde Berlin. ZooSphere@mfn-berlin.de

Entomological collections, particularly the type specimens of the collections, are the fundament of taxonomic research. The necessity to examine these type specimens in the process of identifying new species is the most significant bottleneck. It is both time consuming and costly, because the type specimens needed for a taxonomic work are usually not stored in one single museum, but scattered in museums across the world. Furthermore, sending type specimens to specialist for examination via regular mail, exposes them to risks of damage and loss.

Under this aspect we developed the tool ‘ZooSphere’, which enables us to achieve this ambitious goal. The tool ‘ZooSphere’ positions the specimen and uses stacking technology at every requested perspective of the capturing algorithm. These ZooSphere-objects are freely available in the web (<http://zoosphere.net>) and this web application features an intuitive and interactive rotation and zooming of ‘ZooSphere’ objects in high resolution.