

# Overview of historical biology literature digitisation projects – compiled by vifabio, the Virtual Library of Biology

Gerwin Kasperek

The retrospective digitisation of print library holdings has already become everyday business in many places. Apart from a few players who, with the investment of significant resources, are able to digitise the holdings of large libraries in a few years, many small and medium digitisation projects have been established, often focussed on narrowly defined segments of technical literature. It is difficult for users and also librarians to keep track of the various digitisation activities; to date a useful central directory does not exist. The Virtual Library of Biology – vifabio – is trying to address this serious deficiency with an overview of historical literature digitisation projects that are relevant to biology (Kasperek, 2010). A second approach addressing this deficiency is provided by BHL-Europe by improving the interoperability of European biodiversity digital libraries and making their content available in one central portal.

## 1. vifabio and the biology Internet Guide

The Virtual Library of Biology – vifabio (<http://www.vifabio.de/?lang=en>) – was established by the University Library Johann Christian Senckenberg in Frankfurt (Main), Germany, with many partners (Kasperek, 2007). No digitisation takes place at vifabio, however digitised items from major subject-relevant digitisation projects are in vifabio's Virtual Catalogue; these currently include the Biodiversity Heritage Library (<http://www.biodiversitylibrary.org/>), BioLib (<http://www.biolib.de/>) and AnimalBase (<http://www.animalbase.de/>), as well as those digitised items that are in UB Frankfurt's subject catalogue. The inclusion of all subject-relevant digitised holdings in a central catalogue – a central

record of each digitised item – has, until now, failed due to some obstacles. Considerably more advanced is the plan, within vifabio's framework, to create an overview of retrospective digitisation projects relevant to biology (Section 2). This takes place in vifabio's Internet Guide, which provides records of various types of Internet resources. Resources available on the Web that match a documented collection profile and its quality requirements are catalogued in a specially-constructed database with comprehensive metadata. 23 different options are available in the "Resource Type" metadata field (such as "Factual database", "Literature database" or "Personal website"). One of these options is dubbed "Access to historical literature" and is used when creating an overview of a digitisation project. When a digitisation project is indexed and described in the Internet Guide, its information can be researched and used in multiple ways: with vifabio's comprehensive search (Virtual Catalogue), with special search possibilities and browsing by subject in the Internet Guide; or with a special page, which was created in vifabio specifically to record historical biology literature digitisation projects (<http://www.vifabio.de/other/vifabio-dig.html?lang=en>)





## 2. Overview of existing digitisation projects

41 projects have been discovered by vifabio employees (as of February 2010; see addendum for a list of project names and URLs). The list does not claim to be exhaustive – there are assuredly many further initiatives that are only known in relatively small circles of professionals and do not appear in any library indices.

According to vifabio's subject-specific focus, only resources that are specific to biology or at least the higher areas of natural history or natural sciences with emphasis on biology, or multidisciplinary sources that can offer easy access to a biology section are catalogued.

A project like "Google Books" that offers only keyword search, does not offer a biology section and can give no estimation of the number of biology titles amongst its holdings can therefore not be included. In many cases, sensible demarcation is difficult.

Demarcation problems also stem from the fact that there is a smooth gradient from pure digitisation projects on one side and aggregation services on the other. Many projects offer users not only their own digitised items (that are in some cases not recorded elsewhere) but also records of digitised items that come from other projects (usually with hyperlinks to full text from external providers). The Overview in vifabio only covers projects that provide at least some own digitised items. In some exceptions, projects that do not offer their own digitised items but instead provide access to digitised items that are otherwise difficult to find may be included. Another condition for inclusion in the Overview is that a project or Web site does not present just a single or a few works. Projects in early stages that do not yet offer any digitised items, or retrospective digitisation from publishers that require payment or licensing are also not included. (Regardless of whether or not a digitisation project is included in the Overview, it is likely to be indexed in vifabio's Internet Guide as this

service is not limited to just digitisation projects.) Some important basic data of the included projects are summarised here.

Table 1: Key data on digitisation projects relevant to biology

No.	Name of project or portal	Classification	Number of own digitised items (titles/volumes only)	Number of other (external) records	Download of titles available
1	Aboca Museum - Bibliotheca Antiqua	580	<100	---	---
2	Anatomia 1522-1876 (Plates from the Thomas Fisher Rare Book Library)	590	<100	---	---
3	AnimalBase - Early Zoological Literature Online	590	<10,000	present	X
4	Audubon's Birds of America at the University of Pittsburgh	590	<10	---	---
5	Biblioteca Digital de Botánica	580	<100	---	X
6	Biblioteca Universidad Complutense de Madrid : Botánica	580	<1,000	---	---
7	Biodiversity Heritage Library	570	50,000	20,000	X
8	BioLib - Kurt Stübbers Online-Bibliothek	570	<1,000	---	(x)
9	Biologiezentrum - Österreichische biologische/erdwissenschaftl. Zeitschriften	570	<10,000	---	X
10	Botanicus Digital Library	580	<10,000	---	X
11	BUGZ : Bibliography of New Zealand Terrestrial Invertebrates - Online	590	<10,000	---	X
12	Complete Work of Charles Darwin Online	570	<1,000	---	X
13	Cyberliber: an Electronic Library for Mycology	570	<100	---	---
14	Digital Library of the Real Jardín Botánico de Madrid	580	<10,000	---	X
15	Digitale Bibliothek Braunschweig - Sachgebiet 580: Pflanzen (Botanik)	580	<1,000	---	X
16	Digitale Entomologische Information - Periodika	590	<1,000	---	(x)
17	Digitalisierte historische Gartenliteratur	580	<1,000	<1,000	X
18	Early Classics in Biogeography, Distribution, and Diversity Studies	570	<100	<1,000	(x)
19	E-Corpus : results for Dewey 500 Natural sciences & mathematics	570	---	<10,000	---
20	Gallica : Life sciences; biology / Botany / Zoology	570	<10,000	---	X
21	Guide to the plant species descriptions in seed lists, 1800 – 1900	580	<1,000	---	---
22	HELDA (University of Helsinki Digital Archive) : Fauna & Flora	570	<1,000	---	X
23	Herzogin Anna Amalia Bibliothek - Monographien Digital : Biologische Titel	570	<1,000	---	---
24	Humboldt Digital Library	570	<100	---	X
25	It's in the Blood! A Documentary History of Linus Pauling ...	590	<100	---	---
26	Jean-Baptiste Lamarck (1744-1829) : works and heritage	570	<100	---	(x)
27	Linnaean Correspondence	570	<100	---	---
28	Linus Pauling and the Race for DNA: A Documentary History	570	<100	---	---
29	Missouri Botanical Garden: Rare Books	580	<1,000	---	X
30	Nomenclator Zoologicus	590	<100	---	---
31	NYPL Digital Gallery: Nature Illustrated - Flowers, Plants, Trees, 1550-1900	580	(<100)	---	---
32	On-line catalogue of digitized books of Archbishop chateau in Kroměříž	580	(<100)	---	---
33	Oxford Digital Library : Key 19th century entomological literature	590	<100	---	---
34	Profiles in Science	590	<1,000	---	X
35	Sammlungen online zugänglicher botanischer Werke	580	<100	present	(x)
36	SEALS - swiss electronic academic library service : Naturwissenschaften	570	<1,000	---	X
37	Searchable Ornithological Research Archive (SORA)	590	<1,000	---	X
38	Service de la documentation University of Strasbourg - Sciences de la vie	570	<1,000	---	---
39	Tela Botanica : Ouvrages numérisés	580	<100	<100	X
40	United States Exploring Expedition, 1838 - 1842	570	<100	---	---
41	University of Wisconsin - The Ecology and Natural Resources Collection	570	<1,000	---	(x)

Explanation of parentheses in "Own digitised items" column: most works were only partly digitised, for example title pages and figures only; parentheses in "Download of titles" column: significant limitations in download possibilities, for example if only available for some items, if resources require licenses, or if no facsimiles are available. "Classification" according to DDC divisions (Dewey Decimal Classification): 570 Biology, 580 Botany, 590 Zoology.

The projects are organised by DDC subject groups (Alex & Heiner-Freiling, 2004) and the number of available digitised items in rough size classes; the latter can only be understood as an exploratory approach as there are no mandatory rules for units or quantification. Additionally, many projects do not reveal any explicit data on their holdings of digitised items, so only estimations were possible. Table 1 also indicates whether users are offered a complete download of individual digitised titles.

In February 2010, 41 historical literature digitisation projects relevant to biology were included in the vifabio Overview. Of those, 13 projects are specialised in botanical literature and 10 in zoological literature; the remaining 18 involve biology in the broadest sense. With regard to the number of titles or volumes digitised by a project, the Biodiversity Heritage Library (BHL) is by far the largest. Six further digitisation



projects belong to the size class of 1,000 to 10,000 digitised items. Of these, the Botanicus Digital Library is practically a precursor to BHL; its digitised content is entirely contained in the latter. The remaining five large digitisation projects are predominantly resident in Europe, specifically in Germany (AnimalBase), France (Gallica), Austria (Biologiezentrum), and Spain (Digital Library of the Real Jardín Botánico of Madrid), or in New Zealand (BUGZ). AnimalBase, Gallica, and Biologiezentrum are already active content providers for BHL-Europe and their digital content is expected to be available within the next 1.5 years over the multilingual BHL-Europe portal. Linz Biologiezentrum's project to digitise Austrian biological and earth science journals is distinguished because not only are volumes or journals being almost completely digitised on a large scale, but are also accessible at the article level, so that users can conveniently download individual items as (searchable) PDF documents. With other projects that digitise large volumes of journal literature, comparable convenient access to individual articles is rarely possible. Roughly three quarters of all projects included (33 of 41) have digitised, and made available, fewer than 1,000 items each. Approximately half of all projects – including the previously-mentioned seven big projects – offer users the ability to download digitised entire titles or volumes as PDF; this feature is desired or required by a large proportion of biologists working in science (Tenopir, 2003; Kasperek, 2008 and Hull et al. 2009). The tendency is for PDF downloads to be missing from small projects rather than larger projects or the so-called mass digitisations. Sometimes, projects with a rather limited number of particularly rare works make them only available in special "Viewers" (see "boutique digitisation" in Milne, 2008). Approximately one fifth of all projects include not only their own digitised items in their web presence, but also items from other projects with which they share subject, geographical

or administrative similarities. BHL takes a lead position here as well, after the holdings in the BHL portal were fortified with 15,000 titles through cooperation with the Internet Archive in the second half of 2009. The fact that many digitised items do not only appear on the Web site of the project that produced them, but also in the Web sites of other projects, can lead to problems with duplicates when creating comprehensive indexing systems. This issue of comprehensive indexing is addressed by BHL-Europe as one key aspect of the project workplan. The Global Reference Index to Biodiversity (GRIB) currently available as a prototype under <http://grib.gbv.de> (see also this issue of the BHL-Europe Newsletter) is expected to be able to manage duplicates and search across various catalogues at the same time.

### 3. An outlook

For now, biologists and librarians must accept the fact that the efforts to make biological literature available digitally are distributed amongst many individual projects and the search for a particular title can be very complex. In conjunction with the Web pages presented earlier, vifabio – the Virtual Library of Biology – offers a focal point that can be used to gain an overview of the various digitisation activities. On the one hand, vifabio offers an overview of digitisation projects; on the other hand, users can find records of digitised titles from several important digitisation projects in the Virtual Catalogue (... but it is by no means all titles from all relevant projects). BHL-Europe, in a more comprehensive approach, will provide a Global References Index to Biodiversity (GRIB), expected to be available in spring 2011. The GRIB will bring together bibliographic records from many more European libraries, including records for digitised items, thus providing the opportunity to search for a particular title in only one web interface. For other libraries or owners of subject-related Web sites there is the possibility, with little technical effort, to re-use the data about



digitisation projects held in vifabio. That is to say, build them into their own Web sites. New entries will automatically appear in the embedded list, broken links will be quickly repaired by vifabio editors, so that Web site owners no longer need to react themselves. This form of re-use is realised through the “myBioLinks” service, which can also make any other excerpt from the Internet Guide available to external Web site owners (<http://vifabio.de/about/vifabio-info.html#myBioLinks>).

#### Addendum

Name and URL of projects to digitise historical literature relevant to biology

- 1 Aboca Museum - Bibliotheca Antiqua  
<http://www.abocamuseum.it/bibliothecaantiqua/home.asp>
- 2 Anatomia 1522-1876: Anatomical Plates from the Thomas Fisher Rare Book Library  
<http://link.library.utoronto.ca/anatomia/>
- 3 AnimalBase - Early Zoological Literature Online  
<http://www.animalbase.de/>
- 4 Audubon's Birds of America at the University of Pittsburgh  
<http://digital.library.pitt.edu/a/audubon/>
- 5 Biblioteca Digital de Botânica  
<http://bibdigital.bot.uc.pt/>
- 6 Biblioteca Universidad Complutense de Madrid : Serie Temática de libros - Botánica  
<http://alfama.sim.ucm.es/greco/d-digital.php?idMateriaDioscorides=Bot%E1nica>
- 7 Biodiversity Heritage Library  
<http://www.biodiversitylibrary.org>
- 8 BioLib - Kurt Stübers Online-Bibliothek  
<http://www.biolib.de/>
- 9 Biologiezentrum - Österreichische biologische/erdwissenschaftliche Zeitschriften  
[http://www.biologiezentrum.at/de/bz/pages.php?page\\_id=428](http://www.biologiezentrum.at/de/bz/pages.php?page_id=428)
- 10 Botanicus Digital Library  
<http://botanicus.org/>
- 11 BUGZ : Bibliography of New Zealand Terrestrial Invertebrates - Online  
<http://www.bugz.org.nz/>
- 12 Complete Work of Charles Darwin Online  
<http://darwin-online.org.uk/>
- 13 Cyberliber: an Electronic Library for Mycology  
<http://www.cybertruffle.org.uk/cyberliber/>
- 14 Digital Library of the Real Jardín Botánico of Madrid  
<http://bibdigital.rjb.csic.es/>
- 15 Digitale Bibliothek Braunschweig - Sachgebiet 580: Pflanzen (Botanik)  
<http://bib1lp1.rz.tu-bs.de/docportal/browse/dcc/5/58>
- 16 Digitale Entomologische Information - Periodika  
[http://www.zalf.de/home\\_zalf/institute/dei/php/archiv/journal.php](http://www.zalf.de/home_zalf/institute/dei/php/archiv/journal.php)
- 17 Digitalisierte historische Gartenliteratur  
<http://historischegaerten.de/Gartenbaubuecherei/digilinks.html>
- 18 Early Classics in Biogeography, Distribution, and Diversity Studies  
<http://people.wku.edu/charles.smith/biogeog/>
- 19 E-Corpus : results for Dewey500 Natural sciences & mathematics  
[http://www.e-corpus.org/search/search.php?search=search&subject\\_id=2443](http://www.e-corpus.org/search/search.php?search=search&subject_id=2443)
- 20 Gallica : Life sciences; biology / Botany / Zoology  
<http://gallica.bnf.fr/SearchDewey?d=5&sd/57&n=15&p=1&lang=en>
- 21 Guide to the plant species descriptions published in seed lists from Botanic Gardens, period 1800 – 1900  
<http://www.nationaalherbarium.nl/seedlists/>
- 22 HELDA (University of Helsinki Digital Archive) - Special collection : Fauna & Flora  
<http://helda.helsinki.fi/handle/10138/14068>
- 23 Herzogin Anna Amalia Bibliothek - Monographien Digital : Biologische Titel  
[http://ora-web.swkk.de/digimo\\_online/digimo.Volltextsuche2?vc=botan+plant+zool+biol&op=oder](http://ora-web.swkk.de/digimo_online/digimo.Volltextsuche2?vc=botan+plant+zool+biol&op=oder)
- 24 Humboldt Digital Library  
<http://www.avhumboldt.net/>
- 25 It's in the Blood! A Documentary History of Linus Pauling, Hemoglobin, and Sickle Cell Anemia  
<http://osulibrary.oregonstate.edu/specialcollections/coll/pauling/blood/index.html>
- 26 Jean-Baptiste Lamarck (1744-1829) : works and heritage  
<http://www.lamarck.cnrs.fr/?lang=en>



- 27 Linnaean Correspondence  
<http://linnaeus.c18.net/>
- 28 Linus Pauling and the Race for DNA:  
A Documentary History  
<http://osulibrary.oregonstate.edu/specialcollections/coll/pauling/dna/index.html>
- 29 Missouri Botanical Garden: Rare Books  
<http://www.illustratedgarden.org/mobot/rarebooks/>
- 30 Nomenclator Zoologicus  
<http://www.ubio.org/NomenclatorZoologicus/>
- 31 NYPL Digital Gallery: Nature Illustrated -  
Flowers, Plants, and Trees, 1550-1900  
[http://digitalgallery.nypl.org/nypldigital/explore/dgexplore.cfm?topic=science&collection=NatureIllustratedFlo&col\\_id=184](http://digitalgallery.nypl.org/nypldigital/explore/dgexplore.cfm?topic=science&collection=NatureIllustratedFlo&col_id=184)
- 32 On-line catalogue of digitized books  
of Archbishop chateau in Kroměříž  
<http://digi.azc.cz/main.php?lng=2>
- 33 Oxford Digital Library : Key 19<sup>th</sup> century  
entomological literature  
<http://www2.odl.ox.ac.uk/gsd/cgi-bin/library?site=localhost&a=p&p=about&c=munahi01&ct=0&l=en&w=iso-8859-1>
- 34 Profiles in Science  
<http://profiles.nlm.nih.gov/>
- 35 Sammlungen online zugänglicher botanischer Werke  
<http://www.bgbm.org/bgbm/library/botlit-www.htm>
- 36 SEALS - swiss electronic academic library  
service : Sammlung Naturwissenschaften  
[http://retro.seals.ch/digbib/browse5\\_4](http://retro.seals.ch/digbib/browse5_4)
- 37 Searchable Ornithological Research Archive (SORA)  
<http://elibrary.unm.edu/sora/>
- 38 Service de la documentation University of  
Strasbourg - Digital old books - Sciences de la vie  
<http://num-scd-ulp.u-strasbg.fr:8080/view/subjects/570-590.html>
- 39 Tela Botanica : Ouvrages numérisés  
[http://www.tela-botanica.org/page:ouvrages\\_numerises](http://www.tela-botanica.org/page:ouvrages_numerises)
- 40 United States Exploring Expedition, 1838 - 1842  
<http://www.sil.si.edu/DigitalCollections/usexex/follow-01.htm>
- 41 University of Wisconsin - The Ecology and  
Natural Resources Collection  
<http://digicoll.library.wisc.edu/EcoNatRes/>

## References:

- Alex, H. & Heiner-Freiling, M. (Red., 2004):  
DDC-Sachgruppen der Deutschen  
Nationalbibliografie: Leitfaden zu ihrer  
Vergabe. 85 S. Leipzig (DDB).
- Hull, D., Pettifer, S.R. & Kell, D.B. (2009):  
Defrosting the digital library: bibliographic  
tools for the next generation web. In: PLoS  
Computational Biology 4 (10): 100-204.
- Kasperek, G. (2007): Aufbau einer Virtuellen  
Fachbibliothek für Biologie - vifabio im  
Entstehen. In: ABI-Technik, Zeitschrift  
für Automation, Bau und Technik im Archiv-,  
Bibliotheks- und Informationswesen 27: 78-95.
- Kasperek, G. (2008): Literaturbezogene  
Arbeitsweisen von Wissenschaftlern in der  
Biologie. 50 pp. + Appendix. [With English  
abstract] (Berliner Handreichungen zur  
Bibliotheks- und Informationswissenschaft,  
Heft 223.) [<http://www.ib.hu-berlin.de/~kumlau/handreichungen/h223/>, Download 25.01.2008].
- Kasperek, G. (2010): Eine Übersicht von  
für die Biologie relevanten Projekten  
zur Digitalisierung historischer Fachliteratur.  
In: Bibliotheksdienst 44: 448-460.
- Tenopir, C. (2003): Use and users of electronic  
library resources: an overview and analysis  
of recent research studies. 72 pp.  
[<http://www.clir.org/pubs/reports/pub120/pub120.pdf>, Download 04.05.2004] Washington,  
D.C. (Council on Libr. & Inf. Resources).
- Milne, R. (2008): From "boutique" to mass  
digitization: the Google Library Project  
at Oxford. In: Earnshaw, R. & Vince, J. (eds.):  
Digital convergence - libraries of the future:  
3-9. London (Springer)
- The author  
Dr Gerwin Kasperek, University Library Johann  
Christian Senckenberg, Bockenheimer Landstr.  
134-138, 60325 Frankfurt am Main, Germany  
E-Mail: [g.kasperek@ub.uni-frankfurt.de](mailto:g.kasperek@ub.uni-frankfurt.de)