Neptune (NSB): working with a legacy database

Johan Renaudie
Museum für Naturkunde
2023-05-17

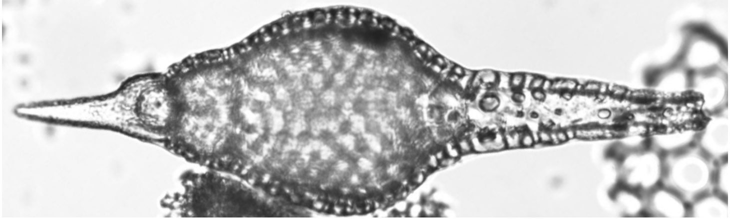
Leibniz-Gemeinschaft

Neptune SANDBOX Berlin

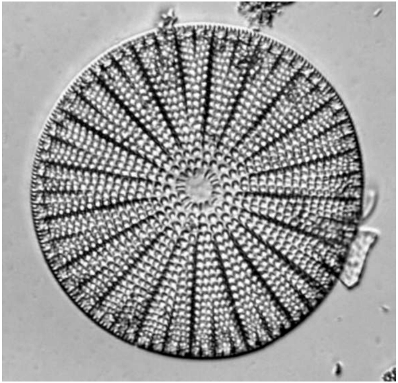
MUSEUM FÜR NATURKUNDE BERLIN

The microfossil record

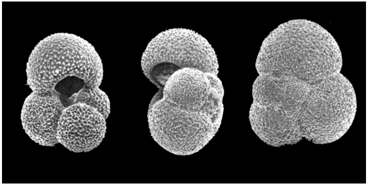
Radiolarians



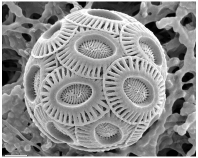
Diatoms

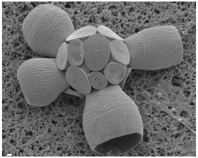


Planktonic Foraminifera



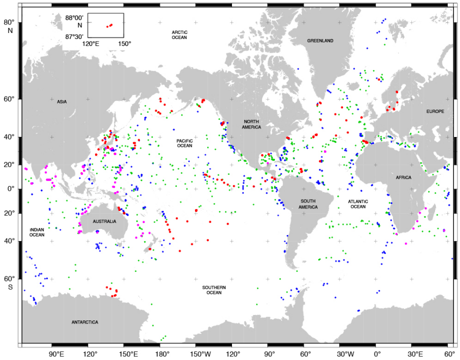
Calcareous Nannofossils





Deep-sea drilling projects

- DSDP (1966–1983), ODP (1983–2003) and IODP (2003–2013; 2013–2024)
- >1500 deep-sea sites drilled in all world's ocean
- In situ, mostly undisturbed marine sediments from late Jurassic to the Present
- Most deep-sea sites primarily composed on microfossils in a context of continous sedimentation



DSDP Legs 1–96 (•), ODP Legs 100–210 (•), IODP Expeditions 301–348 (•), IODP Expeditions 349–371 (•)

The history of the Neptune Database

The 1990s: ETH Zürich

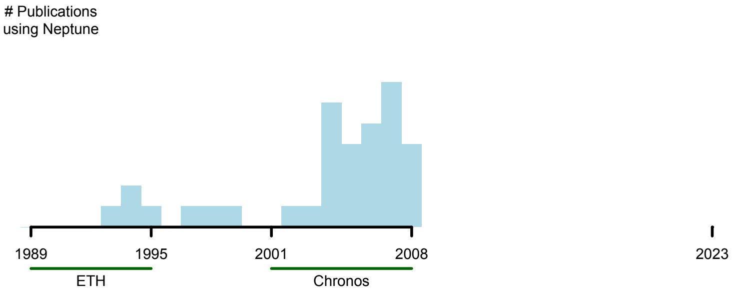
Goal: Mobilize the remarkable fossil record from DSDP-ODP into an occurrence database to test hypotheses on macroevolution.

- Micropaleontology group at the ETH in Zürich including H. Thierstein, C. Cervato, D. Lazarus.
- Technical limitation of the time (RAM 4Mb; HD $<$ 100Mb) drives DB design (no place for metadata!).
- Programmed in 4th Dimension. Available only offline.
- Content: 100 sites including ca. 300k occurrences.

The history of the Neptune Database

The 2000s: NSF-CHRONOS

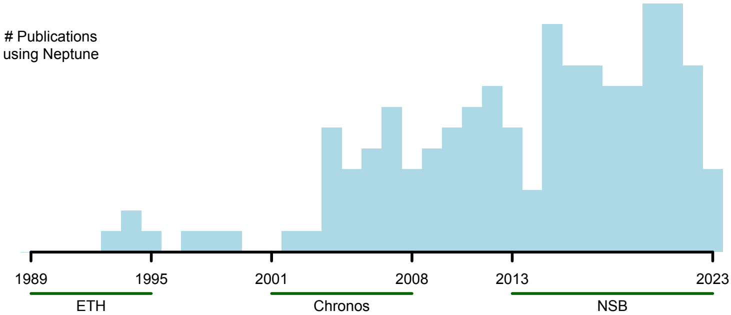
- Moved to Iowa thanks to NSF grant of Chronos project
- Development team including C. Cervato, P. Diver, D. Fils, B, Huber, M. Leckie, ...
- Moved to PostgreSQL. First online front-end.
- Content: ca. 500k occurrences.
- Funding stopped in 2008 and project was more or less abandoned.



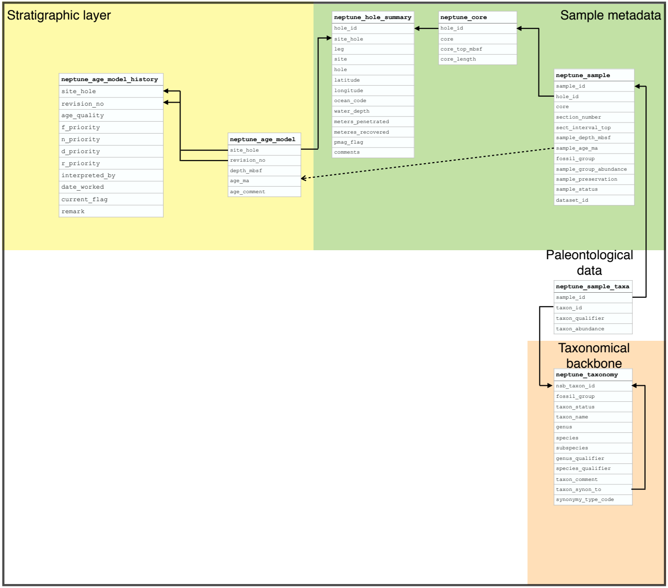
The history of the Neptune Database

The 2010s: Neptune Sandbox Berlin

- Moved to Berlin thanks to CEES funding (L. Liow and N. Stenseth).
- Expansion funded through EARTHTIME-EU ESF project (H. Pälike).
- Development team including D. Lazarus, P. Diver and me.
- New online presence (using python's Django module).
- Significant overhaul of structure to put back the metadata in the DB.
- Content: ca. 500 sites and ca. 750k occurrences.

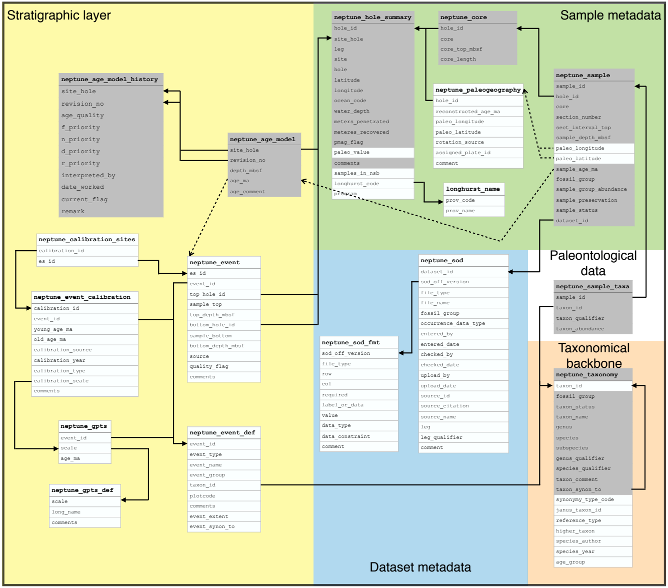


NSB Structure

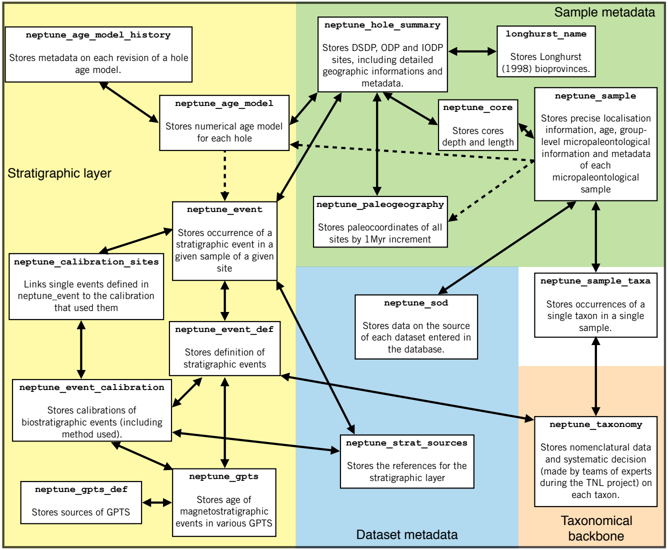


Neptune prior to NSB

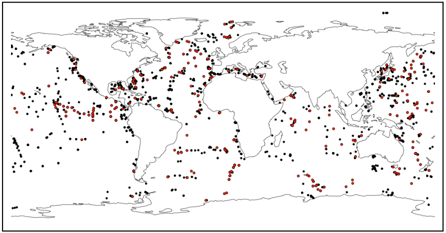
NSB Structure



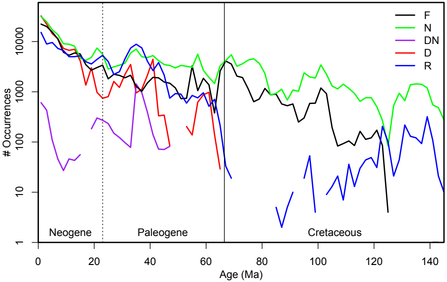
NSB Structure



NSB Content



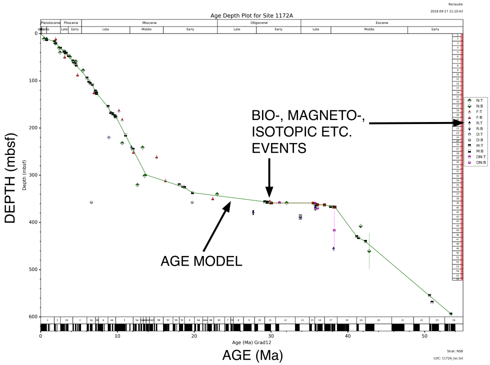
768 057 occurrences.
502 deep-sea drilling holes.
Mostly Cenozoic, but significant Cretaceous.
More carbonate than siliceous fossil data so far.



18 859 taxa names
for 5 microfossil groups
(R, D, PF, N, DN).

Synonymy resolved using TNL: international effort from IODP Paleontology Coordination Group.

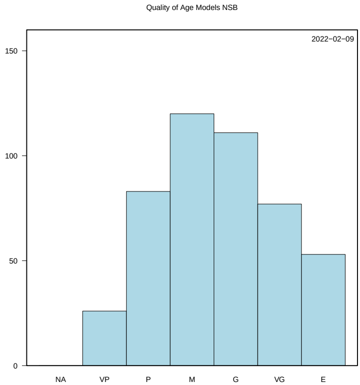
NSB Content



Continuous age vs depth functions (age models) for each section

28 774 stratigraphic events (including 5 130 calibrations for them)

Age models for 483 DSDP, ODP or IODP holes.



Age model quality vary but most "above average".

Website option to ignore poor age models set by default.

Web portal: nsb.mfn-berlin.de

Neptune Sandbox Berlin

About

Help

Search the database

Downloaded datasets

Recent Changes

Neptune SANDBOX Berlin

About Database Access History

About

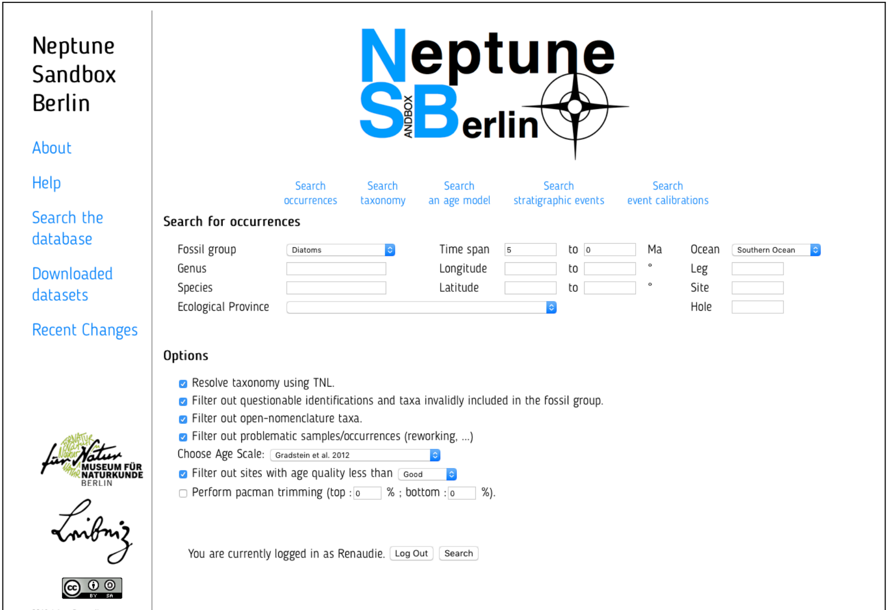
NSB is the current implementation of the Neptune database (Lazarus, 1994; Spencer-Cervato, 1999). It holds hundreds of thousands of occurrence records for thousands of marine plankton microfossil species from hundreds of deep-sea ocean drilling sections; a taxonomic name management list; age models for all sections; and the geochronologic data used to create these age models. NSB serves several distinct groups of users including microfossil taxonomists, evolutionary (paleo)biologists, and paleoceanographers. A selection of papers that have used Neptune/NSB data is given below, and a full list of all papers using, describing or mentioning the database is given here.
NSB also provides data services to the Mikrotax community catalog of microfossils and to the Geobiodiversity Database (GBDB).
NSB is free to use. User accounts are employed to maintain database security and provide feedback on user needs, and can be obtained simply with an email to one of NSB's managers (see here). The only obligation is to cite the database properly (references here) in any publications or public presentations.

Twenty selected papers using NSB

Lazarus, D. 1994. Neptune: a marine micropaleontology database. *Mathematical Geology*, 26(7):817-832.
Spencer-Cervato, C., Thierstein, H. R., Lazarus, D. B., and Beckmann, J. P. 1994. How synchronous are Neogene marine plankton events? *Paleoceanography*, 9:739-763.
Finkel, Z. V., Katz, M. E., Wright, J. D., Schofield, O., and Falkowski, P. 2005. Climatically driven macroevolutionary patterns in the size of marine diatoms over the Cenozoic. *Proceedings of the National Academy of Sciences of the United States of America*, 102(25):8927-8932.
Allen, A. P., Gillooly, J. F., Savage, V. M., and Brown, J. H. 2006. Kinetic effects of temperature on rates of genetic divergence and speciation. *Proceedings of the National Academy of Sciences of the United States of America*, 103(24):9130-9135.
Liow, L. H. and Stenseth, N. C. 2007. The rise and fall of species: implications for macroevolutionary and macroecological studies. *Proceedings of the Royal Society B*, 274(1626):2745-2752.
Muttoni, G. and Kent, D. 2007. Widespread formation of cherts during the early Eocene climatic optimum. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 253(3-4):348-362.
Rabosky, D. L. and Sorhannus, U. 2009. Diversity dynamics of marine planktonic diatoms across the Cenozoic. *Nature*, 247:183-187.
Cermeño, P. and Falkowski, P. G. 2009. Controls on diatom biogeography in the ocean. *Science*, 325:1539-1541.
Fils, D., Cervato, C., Reed, J., Diver, P., Tang, X., Bohling, G., and Greer, D. 2009. CHRONOS architecture: Experiences with an open-source services- oriented architecture for geoinformatics. *Computers and Geosciences*, 35(4):774-782.

Web portal: nsb.mfn-berlin.de

Occurrence search



Web portal: nsb.mfn-berlin.de

Taxonomy explorer

Neptune Sandbox Berlin

About

Help

Search the database

Downloaded datasets

Recent Changes

Neptune SANDBOX Berlin

Search occurrences | Search taxonomy | Search an age model | Search stratigraphic events | Search event calibrations

Search the database

Genus: Plannapus

Species:

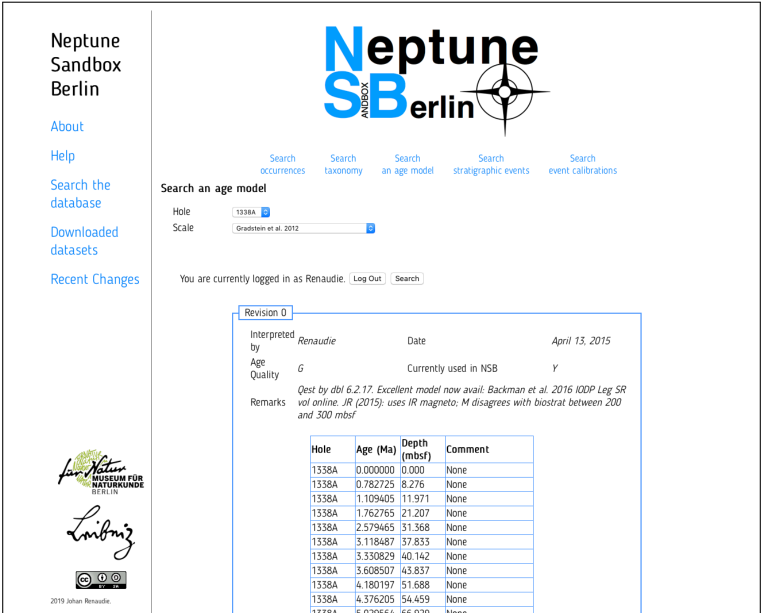
You are currently logged in as Renaudie. Log Out Search

Fossil Group	Taxon ID	Taxon Name	Taxon Status	Taxon Synonym to	Synonymy type	Author	Year	ID in Janus DB	Old NSB ID
R	2100165	Plannapus hornibrooki	V			O'Connor	1999		
R	2100166	Plannapus mauricei	V			O'Connor	1999		
R	2100167	Plannapus microcephalus	V			O'Connor	1999		
R	2003804	Plannapus microcephala	S	2100167		(Haeckel) O'Connor	1997		
R	2001115	Dicolocapsa microcephala	S	2100167	OBJECTIVE	Haeckel	1887	2740	RDICO0020
R	2100452	Plannapus sp. A	G						
R	2100052	Dicolocapsa microcephala (q)	Q	2001115					RDICO0021

MUSEUM FÜR NATURKUNDE BERLIN

Web portal: nsb.mfn-berlin.de

Age models



Neptune Sandbox Berlin

About

Help

Search the database

Downloaded datasets

Recent Changes

Search occurrences | Search taxonomy | Search an age model | Search stratigraphic events | Search event calibrations

Search an age model

Hole: 1338A

Scale: Gradstein et al. 2012

You are currently logged in as Renaudie. Log Out Search

Revision 0

Interpreted by	*Renaudie*	Date	*April 13, 2015*
Age Quality	*G*	Currently used in NSB	*Y*
Remarks	*Qest by dbl 6.2.17. Excellent model now avail: Backman et al. 2016 IODP Leg SR vol online. JR (2015): uses IR magneto; M disagrees with biostrat between 200 and 300 mbsf*		

Hole	Age (Ma)	Depth (mbsf)	Comment
1338A	0.000000	0.000	None
1338A	0.782725	8.276	None
1338A	1.109405	11.971	None
1338A	1.762765	21.207	None
1338A	2.579465	31.368	None
1338A	3.118487	37.833	None
1338A	3.330829	40.142	None
1338A	3.608507	43.837	None
1338A	4.180197	51.688	None
1338A	4.376205	54.459	None

Web portal: nsb.mfn-berlin.de

Event calibrations

Neptune Sandbox Berlin

About

Help

Search the database

Downloaded datasets

Recent Changes

Search occurrences | Search taxonomy | Search an age model | Search stratigraphic events | Search event calibrations

Search an event calibration

Event BOT | Scale Gradstein et al. 2012

You are currently logged in as Renaudie. Log Out Search

Event ID	Calib. ID	Type	Event Name	Group	Age min	Age max	Geographical Extent	Source	Calibration Type	Original Scale	Comments
1073	1210	BOT	Emiliania huxleyi	N	0.3		global	Berggren et al. 1985	M	Berg85	ODP Technical Note 24
1073	174	BOT	Emiliania huxleyi	N	0.29		global	Backman et al. 2012	O	Grad04	

All ages given here on Gradstein et al. 2012 scale.

This event has been found in the following sites:

Event ID	TOP				BOTTOM				Source event	Comment	
	Hole	Sample	Depth (mbsf)	Age (*)	Hole	Sample	Depth (mbsf)	Age (*)			
1073	101_626C	5-1,40	38.90		101_626C	5-CC	48.01		626C_fn_bstrat95		
1073	104_642B	2-6,85	13.15		104_642B	3-2,87	16.67		Donnally 1989		
1073	104_643A	1-2,50	2.00	0.278	104_643A	1-3,50	3.50	0.31	Donally 1989		
1073	104_644A	4-1,50	26.20		104_644A	4-2,50	27.70		644A_mfnr_bstrat95		
1073	105_646A	2-5,102	12.02	0.184	105_646A	2-6,104	13.54	0.207	Baldauf et al 1989		
1073	105_646B		12.03	0.182	105_646B		13.55	0.2	Baldauf et al 1989		
1073	105_647A	1-6,130	8.80	0.275	105_647A	1-7,14	9.14	0.285	647A_mfn_bstrat95		
1073	107_651A	8-CC	64.42		107_651A	9-1,69	69.49		651A_mn_bstrat95		
1073	107_653A	3-2,60	15.30		107_653A	3-2,120	15.90		Rio et al. 1990		Used for calibration of the event (see calibration No 174).
1073	107_653A		15.60		107_653A		15.60		Glaçon et al 1990		
1073	107_655A	1-2,20	1.70		107_655A	1-2,120	2.70		Müller 1990		
1073	108_657A		0.90		108_657A		3.30		657A_fn_bstrat95		
1073	108_658A		34.20	0.275	108_658A		43.70	0.318	Manivit 1989		
1073	108_658A		34.20	0.275	108_658A		43.70	0.318	658A_mfn_bstrat95		
1073	108_659A	1-5,130	7.30	0.241	108_659A	2-1,30	8.10	0.27	Manivit 1989		

Web portal: nsb.mfn-berlin.de

Search for events in sections

Neptune Sandbox Berlin

About

Help

Search the database

Downloaded datasets

Recent Changes

Search occurrences | Search taxonomy | Search an age model | Search stratigraphic events | Search event calibrations

Search for stratigraphic events

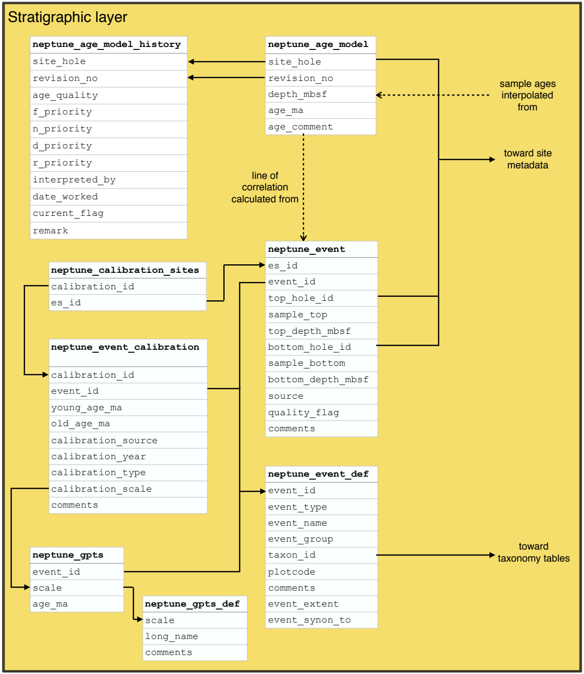
Hole: 320-1333A

Scale: Gradstein et al. 2012

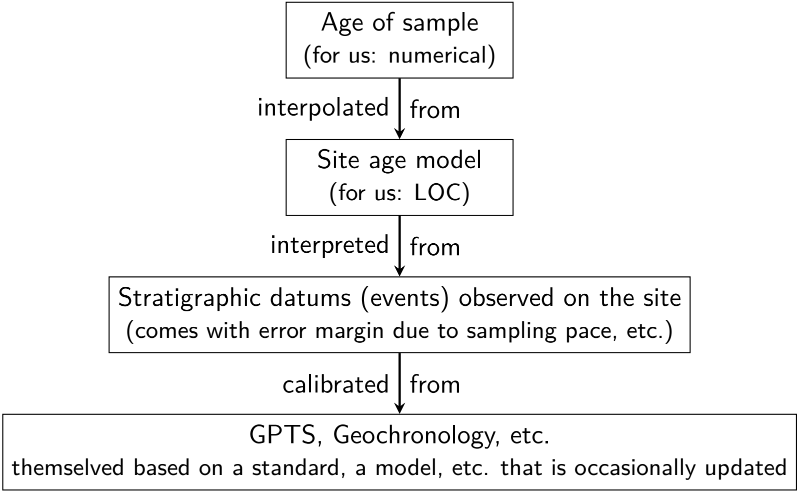
You are currently logged in as Renaudie. Log Out Search

	Type	Event Name	Sample (Top)	Sample (Bottom)	Depth mbsf (Top)	Depth mbsf (Bottom)	Age min	Age max	Source event	Source calibration	Comment
M	TOP	C6An.1n	320_1333A-1-1,55	320_1333A-1-1,65	0.55	0.65	20.04		Expedition 320/321 Scientists, 2010	Gradstein et al. 2012	
M	BOT	C6An.1n	320_1333A-1-1,110	320_1333A-1-1,115	1.10	1.15	20.21		Expedition 320/321 Scientists, 2010	Gradstein et al. 2012	
M	TOP	C6An.2n	320_1333A-1-2,50	320_1333A-1-2,60	2.00	2.10	20.44		Expedition 320/321 Scientists, 2010	Gradstein et al. 2012	
R	BOT	Stichocorys delmontensis	320_1333A-1-2,104	320_1333A-1-4,104	2.54	5.54	20.6		Expedition 320/321 Scientists, 2010	Kamikuri et al. 2012	Used for age model revision 0
										Kamikuri	

MUSEUM FÜR NATURKUNDE BERLIN


Stratigraphic layer
neptune_age_model_history
site_hole
revision_no
age_quality
f_priority
n_priority
d_priority
r_priority
interpreted_by
date_worked
current_flag
remark
neptune_age_model
site_hole
revision_no
depth_mbsf
age_ma
age_comment
sample ages interpolated from
toward site metadata
line of correlation calculated from
neptune_event
es_id
event_id
top_hole_id
sample_top
top_depth_mbsf
bottom_hole_id
sample_bottom
bottom_depth_mbsf
source
quality_flag
comments
neptune_calibration_sites
calibration_id
es_id
neptune_event_calibration
calibration_id
event_id
young_age_ma
old_age_ma
calibration_source
calibration_year
calibration_type
calibration_scale
comments
neptune_event_def
event_id
event_type
event_name
event_group
taxon_id
plotcode
comments
event_extent
event_synon_to
toward taxonomy tables
neptune_gpts
event_id
scale
age_ma
neptune_gpts_def
scale
long_name
comments

Keeping track of stratigraphic information



Age comes with lots of baggage as a whole chain of metadata: any link in the chain can (and will) change, and the changes need to ripple to the end product.

Dealing with legacy data

Quality Control over Data Acquisition

- Elimination of duplicates

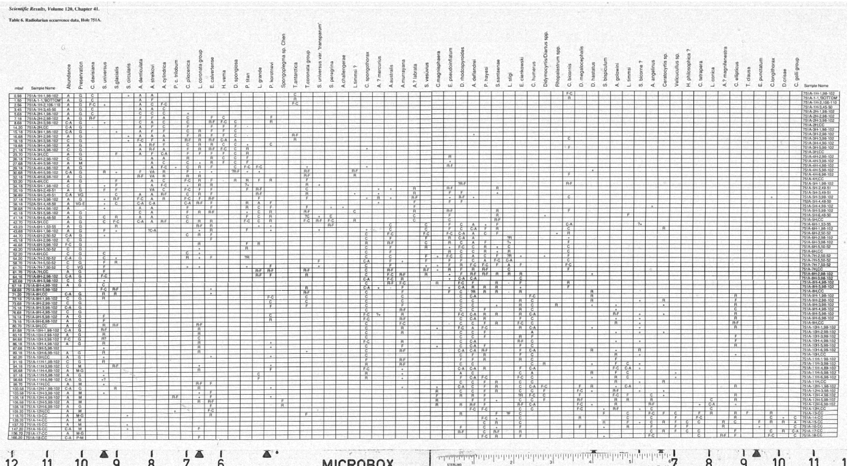
  How to identify duplicates based solely on site + fossil group info?

Dealing with legacy data

Quality Control over Data Acquisition

- Elimination of duplicates
- Finding old faulty entries due to manual entry

Difficult to find errors when original source is not known
Errors frequent due to manual entry



Dealing with legacy data

Quality Control over Data Acquisition

- Elimination of duplicates
- Finding old faulty entries due to manual entry
- Dealing with design choices linked to past technological restrictions

  No space on HD = no metadata
  Arbitrary field restrictions (e. g. Core Numbers, Codes that used to be unambiguous but are not anymore, etc.)
  Changing standards (e. g. coded absence)

Dealing with legacy data

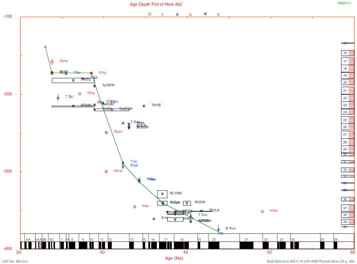
Quality Control over Data Acquisition

- Elimination of duplicates
- Finding old faulty entries due to manual entry
- Dealing with design choices linked to past technological restrictions
- Updating outdated taxonomy and stratigraphy

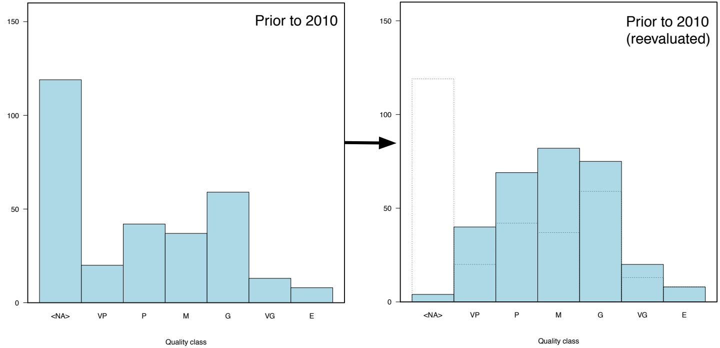
For taxonomy: TNL (see Lazarus et al. 2015 Zootaxa)
For stratigraphy: update GPTS ... and then redo all the age models

Quality Control

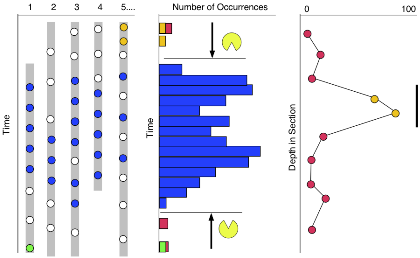
Age model assessment



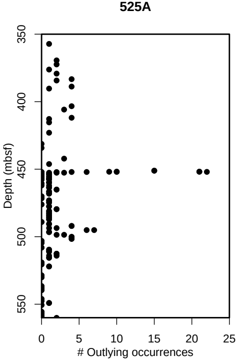
- New quality assessment of old Chronos-era age models

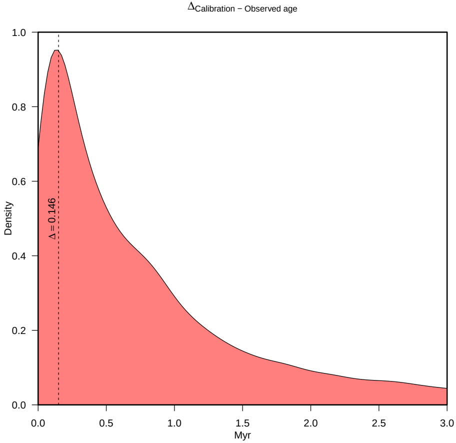


Quality Control
Pacman

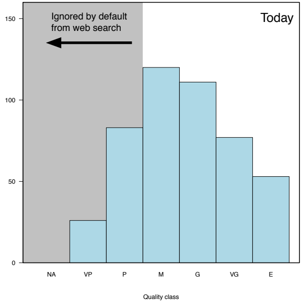


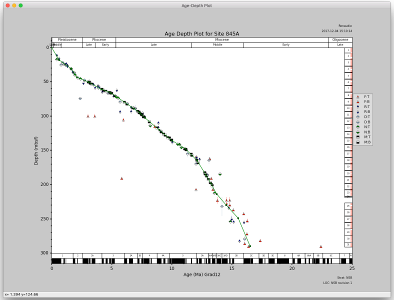
- New quality assessment of old Chronos-era age models
- Outlier detection using e. g. PacMan analysis (Lazarus et al. 2012)





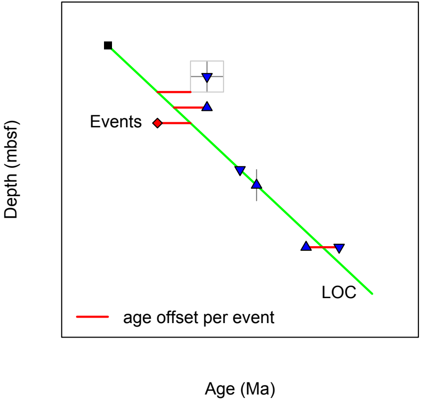
Quality Control

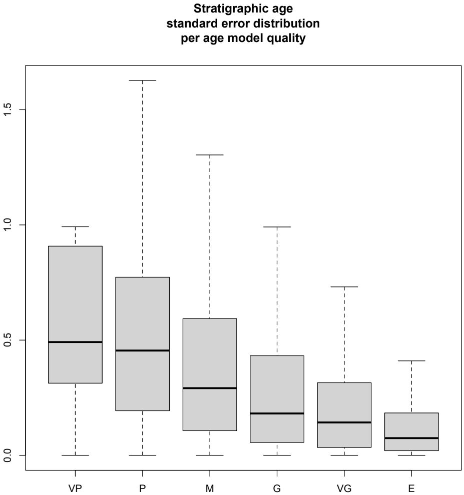




- New quality assessment of old Chronos-era age models
- Outlier detection using e. g. PacMan analysis (Lazarus et al. 2012)
- Selected undated holes containing the larger amount of samples
- Re-did main offenders by using modern calibrations and newly published statigraphic events (including astrochronology)
- Ported legacy software ADP (Age-Depth Plot) from Basic and Java to Python to facilitate workflow
- $>$350 new/revised age models since 2014; including $>$150 since 2020.

Error estimates on age models

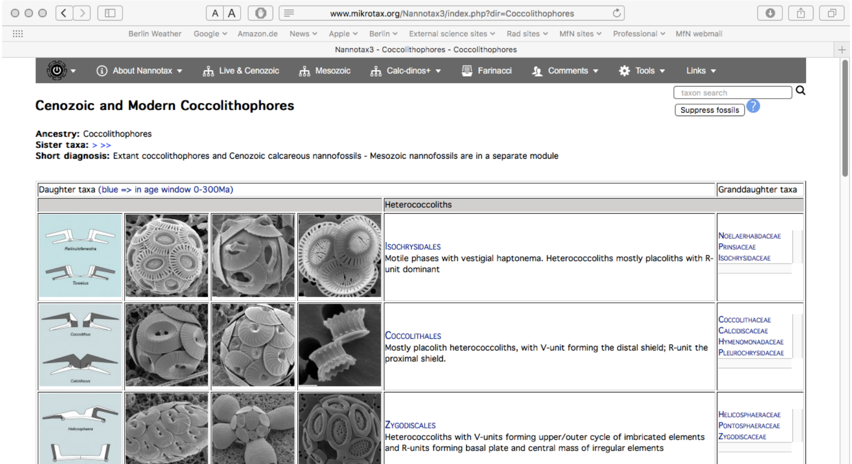




Age model quality estimate qualitative but match quantitative estimates: **VP**: LOC poorly constrained; **P**: median error ca. ± 0.45Myr; **M**: ca. ± 0.30Myr; **G**: ca. ± 0.20Myr; **VG**: ca. ± 0.15Myr; **E**: ca. ± 0.075Myr.

Other related projects

Mikrotax



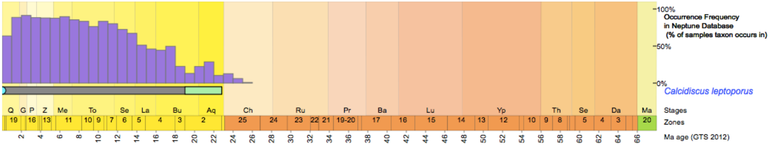
Other related projects

Mikrotax

Geological Range:
Last occurrence (top): **Extant** Data source: present in the plankton (Young et al. 2003)
First occurrence (base): within **NN2** zone (19.00-22.82Ma, base in Aquitanian stage). Data source: Young (1998)

Plot of occurrence data:

- Range-bar - range as quoted above, pink interval top occurs in, green interval base occurs in.
- Triangles indicate an event for which a precise placement has been suggested
- Histogram - Neptune occurrence data from DSDP and ODP proceedings. Interpret with caution & read these notes
- Taxon plotted: *Calcidiscus leptoporus*, synonyms included - *Calcidiscus leptoporus; Calcidiscus leptoporus f. rigidus; Calcidiscus leptoporus hol; Calcidiscus leptoporus subsp. centrovalis; Calcidiscus quadriforatus; Coccosphaera leptopora; Cyclococcolithina leptopora; Cyclococcolithus leptoporus;* Parent: C. leptoporus group



References:

Backman, J., (1980). Miocene-Pliocene nannofossils and sedimentation rates in the Hatton-Rockall Basin, NE Atlantic Ocean. *Stockholm Contributions in Geology*, **36**: 1-91.

Bartolini, C. & Pirini, C., (1969). *Decouverte de nannoplancton calcaire dans les gres de Ponsano, Miocene Moyen, Toscane, Italie.* In: Bronnimann, P. and Renz, H.H. (Editors), Proceedings of the First International conference on Planktoni Microfossils, Geneva 1967. E. J. Brill, Geneva, pp. 81-88.

de Kaenel, E. & Villa, G., (1996). Oligocene-Miocene calcareous nannofossil biostratigraphy and paleoecology from the Iberian Abyssal Plain. *Proceedings of the Ocean Drilling Program. Scientific Results*, **149**: 79-145.

Gartner, S., (1967). Nannofossil species related to *Cyclococcolithus leptoporus* (Murray and Blackman). *University of Kansas Paleontological Contributions*, **Paper 28**: 1-4.

-v-more-v-

Calcidiscus leptoporus compiled by Jeremy R. Young, Paul R. Bown, Jacqueline A. Lees *viewed:* 22-10-2017

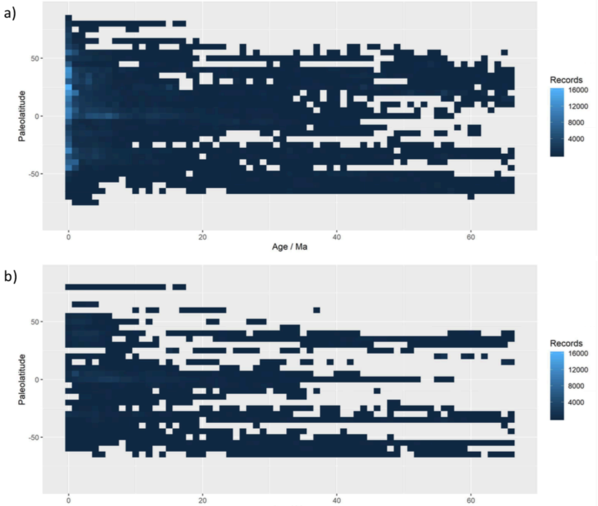
Taxon Search: Orthorhabdus Submit **Advanced Search**

AphiaID: 235923 Nomenclatural data on WoRMS

Other related projects

Triton

Database of Cenozoic Planktonic Foraminifera.
NSB + Land Sections + Surface sediments + ...
New ML-generated age models
500k records of PF (=4x NSB content)



Fenton, I.S., Woodhouse, A., Aze, T., Lazarus, D.B., Renaudie, J., Dunhill, A.M., Young, J.R., Saupe, E.E. (2021) Triton, a New Species-Level Database of Cenozoic Planktonic Foraminiferal Occurrences. *Scientific Data* 8.

Example of studies using NSB

To date, 152 publications using Neptune
Include micropaleontology case studies, paleobiology, theoretical biology, paleoceanography, marine geology, etc.

Thermal niches of planktonic foraminifera are static throughout glacial–interglacial climate change

Gwen S. Antell[a,1], Isabel S. Fenton[a], Paul J. Valdes[b], and Erin E. Saupe[a,1]
[a]Department of Earth Sciences, University of Oxford, OX1 3AN Oxford, United Kingdom, and [b]School of Geographical Sciences, University of Bristol, BS8 1SS Bristol, United Kingdom

Edited by Nils Chr. Stenseth, University of Oslo, Oslo, Norway, and approved March 12, 2021 (received for review August 12, 2020)

Out of the extratropics: the evolution of the latitudinal diversity gradient of Cenozoic marine plankton

Nussaïbah B. Raja and Wolfgang Kiessling
GeoZentrum Nordbayern, Department of Geography and Geosciences, Friedrich-Alexander University Erlangen-Nürnberg, Loewenichstr. 28, 91054 Erlangen, Germany
NBR, 0000-0002-0000-3944; WK, 0000-0002-1088-2014

Article
Neogene burial of organic carbon in the global ocean

https://doi.org/10.1038/s41586-022-05413-6
Ziye Li[1,2,3], Yi Ge Zhang[2], Mark Torres[4] & Benjamin J. W. Mills[5]
Received: 14 February 2022

SCIENCE ADVANCES | RESEARCH ARTICLE

ECOLOGY
Diversity dependence is a ubiquitous phenomenon across Phanerozoic oceans

Valentin Rineau[1]*, Jan Smyčka[1], David Storch[1,2]

Palaeontology The Palaeontological Association www.palass.org
[Palaeontology, 2022, e12615]

Diversity dynamics of microfossils from the Cretaceous to the Neogene show mixed responses to events

by KATIE M. JAMSON[1,3,*], BENJAMIN C. MOON[1] *and* ANDREW J. FRAASS[1,2,3]
[1]Palaeobiology Research Group, School of Earth Sciences, University of Bristol, Wills Memorial Building, Queens Road, Bristol, BS8 1RJ, UK; katiejamson@uvic.ca, ben@bcmoon.uk, andyfraass@uvic.ca
[2]The Academy of Natural Sciences of Drexel University, 1900 Benjamin Franklin Parkway, Philadelphia, PA 19103, USA
[3]*Current address*: School of Earth & Ocean Sciences, University of Victoria, Bob Wright Centre A405, Victoria, BC V8W 2Y2, Canada
*Corresponding author

Typescript received 10 May 2021; accepted in revised form 14 March 2022

AGU ADVANCING EARTH AND SPACE SCIENCE
Geochemistry, Geophysics, Geosystems
RESEARCH ARTICLE
10.1029/2022GC010667

The History of Cenozoic Carbonate Flux in the Atlantic Ocean Constrained by Multiple Regional Carbonate Compensation Depth Reconstructions

Adriana Dutkiewicz[1] and R. Dietmar Müller[1]
[1]EarthByte Group, School of Geosciences, University of Sydney, Sydney, Australia

Key Points:

- Carbonate carbon flux and storage has been computed for the Atlantic Ocean spanning the entire Cenozoic at 0.5 m.y. intervals

To-Do list

- Rest API – allow easy automatized remobilization of database content. Current Beta API limited to age models.
- Benthic Microfossils (benthic forams and ostracods for instance) – need taxonomic experts for synonymies.
- Land Sections: some time period are under represented in in situ deep sea sediments for some of the microfossils (e. g. Paleocene and Early Eocene diatoms) but present in land sections.
- Mirror site & database
- Versioning of DB through Adam's `chronoverse` package?

Additional information

Access to the Database:
Website: `http://nsb.mfn-berlin.de`

NSBcompanion, R package: `http://github.com/plannapus/NSBcompanion`

NSB_ADP_wx: `http://github.com/plannapus/nsb_adp_wx/releases`

Renaudie, J., Lazarus, D.B., Diver, P. (2020) NSB (Neptune Sandbox Berlin): an expanded and improved database of marine planktonic microfossil data and deep-sea stratigraphy. *Palaeontologia Electronica*, 23(1):a11.

Leibniz-Gemeinschaft

CEES
Centre for Ecological and Evolutionary Synthesis

European Science Foundation

MUSEUM FÜR NATURKUNDE BERLIN