Contents

Cover image- the main image on the front cover and the poster is of freshwater diatoms, Fragilaria flexura and Aulacoseira subarctica. ca. 850 calendar years old, from Two Yurts Lake, Kamchatka, Russia. Image by Ulrike Hoff, Uit, The Arctic University of Norway, Tromsø, and a winning entry in the 2015 TMS photomicroscopy competition.

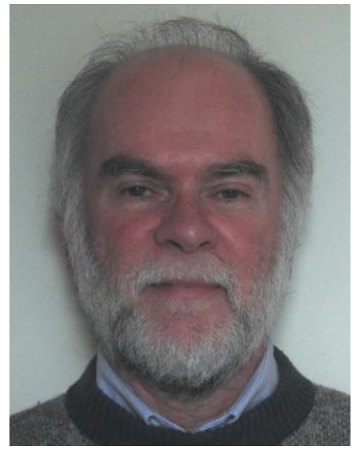
The Lyell meeting is organised via the Joint Committee for Palaeontology as a joint collaboration between the Geological Society, the Palaeontological Association, the Micropalaeontological Society and The Palaeontographical Society.

Oral Presentation Programme

Wednesday 9 March 2016	
09.00	**Registration & tea & coffee (Main foyer and Lower Library)**
09.30	**Opening comments:** Ken Johnson, (The Natural History Museum, London)
09.40	**Value and validity: non-marine ostracods and the challenges of compiling large useful databases** David Horne (Queen Mary University, London)
09.50	**The Neptune database – current developments and future plans** David Lazarus (Museum für Naturkunde, Berlin, Germany)
10.10	**Nannotax an online database of nannoplankton taxonomy – and the value of cybertaxonomy in palaeontology** Jeremy Young (University College London)
10.30	**Building an open-access online planktonic foraminiferal taxonomic database for future generations** Brian Huber (Smithsonian Inst, USA)
11.00	**Tea, coffee and refreshments (Lower Library)**
11.30	**eMesozoic: digitising historical palaeontological collections** Pip Brewer (The Natural History Museum, London)
11.50	**The Panama Paleontology Project Database** Paola G. Rachello-Dolmen (Texas A&M & Smithsonian Tropical Research Inst, USA)
12.10	**Challenges to 3-D digital data publication in palaeontology and the need for community-wide standards for data verification** Thomas Davies (University of Bristol)
12.30	**Lunch Break (lunch is not provided)**
14.00	**The fossil calibration database** Phil Donoghue (University of Bristol)
14.20	**Integrated management of biodiversity and paleontology data within GFBio and PANGAEA** Robert Huber (University of Bremen, Germany)
14.40	**Putting paleontology into WoRMS** Andreas Kroh (Natural History Museum, Vienna, Austria)
15.00	**Analysis of Mesozoic – early Paleogene terrestrial vertebrate diversity demonstrates the utility of the Paleobiology Database for deep time macroevolutionary research** Richard Butler (Lapworth Museum, Birmingham)
15.20	**Tea, coffee and refreshments (Lower Library)**

15.50	**The Geobiodiversity Database and its Taxonomy and occurrence data** Fan JunXuan (Nanjing Inst of Geology & Palaeontology)
16.10	**Synthesising Biostratigraphic Data; a proposed method for common practice and examples of uses** Monika Dlubak (Neftex Exploration Insights)
16.30	**Specimen-level mining: bringing knowledge back 'home' to the Natural History Museum, London** Ross Mounce (The Natural History Museum, London)
16.50	**GeoDeepDive: a cyberinfrastructure to support the construction and enhancement of literature-derived databases** Shanan E Peters (Wisconsin University, USA)
17.10 – 18.15	**Drinks reception (Lower Library)**

Oral Presentation Abstracts



Value and validity: non-marine ostracods and the challenges of compiling large, useful databases

David J. Horne
School of Geography, Queen Mary University of London, Mile End |Road, London E1 4NS, UK.
email: d.j.horne@qmul.ac.uk

Ostracods, diverse and widespread in marine and non-marine (continental) aquatic environments, have an excellent fossil record and are valuable palaeoenvironmental proxies. Calibration of ostracod species for palaeoclimate reconstruction utilises living distributional datasets in combination with modern climate datasets. Small regional datasets collected during relatively short time intervals may be internally consistent in terms of sampling methods, taxonomy and positional accuracy, but often fall short of capturing the full climatic ranges of taxa. Geographically more extensive, literature-based datasets are more likely to cover species' distributions in climate space but, since they represent collections by many people over many years, are prone to taxonomic inconsistency as well as poor locational precision and accuracy, and may even represent distributions blurred by shifts in response to contemporary climate change. Regional climatic datasets, too, span different time intervals. The inability to match climate data exactly to the interval during which distributional data were collected weakens an underlying assumption of palaeoclimate proxy methods: that species' distributions are at least approximately in equilibrium with climate. The Non-marine Ostracod Distribution in Europe (NODE) database, originally established to map the Pleistocene–Recent biogeography of sex and parthenogenesis, now contains c. 12,000 living and fossil (Quaternary) records matched to locality coordinates (latitude and longitude) and has supported quantitative palaeoclimate reconstruction applications such as the Mutual Ostracod Temperature Range (MOTR) method. Like NODE, equivalent databases for North America, such as the Delorme Ostracode Autecological Database (DOAD) for Canada and the North American Non-marine Ostracode Database (NANODe) for the USA, do not on their own encompass fully the climatic ranges of all the taxa they contain. The Ostracod Metadatabase of Environmental and Geographical Attributes (OMEGA) aims to achieve global coverage, capturing species' full climatic ranges, thus far focusing on NODE, NANODe and DOAD but with growing coverage of Asia, Australia and southern Africa. In combination with the WorldClim global interpolated climate dataset it will have the capability to facilitate Quaternary palaeoclimate reconstructions anywhere in the world where suitable fossil ostracod records are available. Realising the potential of such a global dataset requires harmonisation of the taxonomic schemes of the component regional databases as well as the checking, correction and validation of locality coordinates. Such challenges are time-consuming but ultimately rewarding and considerably more useful than the proliferation of web-based taxonomic accountancy that purports to audit biodiversity but often lacks adequate spatial resolution for meaningful biogeographical studies.

NOTES

The Neptune database - current developments and future plans

David Lazarus[1], Johan Renaudie[1], Cinzia Cervato[2] and Patrick Diver[3]

1-Museum für Naturkunde, Berlin, Germany;
2-Iowa State University, Ames, IA, USA;
3-Divdat Consulting, Wesley, AR, USA.
email: david.lazarus@mfn-berlin.de

Neptune is a database of marine microfossil data. It includes species taxonomy, occurrences of species in samples that are in sections, geologic age models for these sections, and the geochronologic data used to make these age models. The current form of this database is Neptune Sandbox Berlin, hosted at the Museum für Naturkunde, Berlin. As of this writing (9.11.2015) it contains 767K occurrences for 18K taxa names from 63K samples, taken from 433 sections, all so far from deep-sea sections recovered by the DSDP, ODP and IODP programs. The age models for these sections are based on ca 27K internally held bio- and magneto-stratigraphic events. Most data is for planktonic foraminifera, calcareous nannofossils, radiolarians and diatoms with a small amount of data for dinoflagellates. Coverage declines with increasing age, with maximum data densities of ca 10K occurrences/my per group in the late Neogene, ca 1K/my/group in the Paleogene, and a few hundred/my/group (foraminifera, nannofossils and radiolarians only) in the early Cretaceous. NSB is searchable either from its website (nsb-mfn-berlin.de) or by direct sql query. Accounts are provided on request. NSB is supported by several client-side apps and scripts, including an improved version of an Age-Depth Plot (ADP) app that allows age-modelling directly linked to NSB data tables, and an R package for local data analyses. The NSB system is mostly implemented in Python with a Postgres database backend, a design deliberately chosen to make the software easier to maintain even without professional programmer support.

Neptune was originally developed at the ETH Zürich in the 1990s to provide a tool for taxonomic syntheses and biodiversity dynamics studies. It was ported to standard sql data structures and the internet by the Chronos project (Ames, Iowa) in the early 2000s, and was largely rewritten and substantially expanded in content at the MfN in the last few years. Numerous people and organisations have contributed to this system's development over the years, including many ODP and IODP supported micropaleontologists who have provided comprehensive, authoritative lists of species names and synonyms, and funding organisations such as SNF, NSF, DFG, ESF and CEES (Oslo).

Neptune has mostly been used for paleobiologic research. Topics include paleobiodiversity dynamics and testing evolutionary ecologic models. There have also been a few paleoceanographic studies, concentrated on global syntheses of the Cenozoic silica cycle. With the recent addition of geochronologic event data and the improved age-modelling tool ADP, NSB will also be of greater interest to biostratigraphers and paleoceanographers. Proposals have been submitted recently to better network NSB with several other databases, including community taxonomic catalogs such as Nannotax; the main paleoceanographic data repository: Pangea; and oceanographic databases hosted by NCEI (NOAA, USA).

NOTES

Nannotax an online database of nannoplankton taxonomy- and review of the value of cybertaxonomy in palaeontology

Jeremy R. Young, Paul R. Bown, Jacqueline A Lees
Earth Sciences, University College London, WC1E 6BT, UK
email: jeremy.young@ucl.ac.uk

Nannotax is an online database of nannoplankton taxonomy (ina.tmsoc.org/Nannotax3) - including both living coccolithophores and the full range of calcareous nannofossils. We have been developing it since 2007 with the broad ambition of providing an authoritative guide to the biodiversity and taxonomy of nannoplankton. This has been aided by a few factors, including (1) Nannoplankton, like many plankton groups, have relatively low diversities but high abundances and wide biogeographic ranges and so play major roles in modern geology, especially palaeoceanography and industrial biostratigraphy. (2) The three of us have been collaborating over a long period, have complimentary expertise and a strong interest in taxonomy. (3) There has been substantial previous work on cataloguing and monographing nannoplankton which we were able to directly draw upon, without copyright problems. (4) We have been able to obtain funding support for the work, initially from the EU EDIT (European Distributed Institute for Taxonomy) and subsequently from a UK Natural Environment Research Council knowledge exchange grant. As a result development of the site, has largely lived up to our ambitions, although there is still much scope for improvement. It currently includes basic details of about 3000 taxa illustrated by 20,000 images and supported by a catalog of original descriptions of ca 5000 species (including many synonyms). In addition >500 illustrated tables outline the content of all higher taxa and allow rapid navigation through the site. More recently we have used occurrence data from the Neptune database to add histograms of stratigraphic occurrence frequency to pages.

So, the site is probably currently the most comprehensive web database available for any substantial group of fossils. As such it can be used as a case study to determine if such efforts are productive and worthwhile. Obvious questions include, is the site actually used? Is the usage serious? Is there evidence of changes in behaviour as a result of access to the site? Related questions include which types of features do users want? Are there dangers of users over-relying on websites and failing to engage with the original references? Do users want both printed monographs and online sources? Is online monography a purely a service task or academically worthwhile itself? How can user contributions be facilitated? Can scholarship standards be maintained online? In the talk we will attempt to address these questions using personal experience, web-usage data from Google Analytics, anecdotal evidence from colleagues, online comments and other feedback.

NOTES

Building an open-access online planktonic foraminiferal taxonomic database for future generations

Brian T. Huber[1], Paul R. Bown, Jeremy R Young

1: Smithsonian Institution, Washington DC, USA

2: University College London, Gower St, WC1E 6BT, UK
email: huberb@si.edu

The largest expert database on planktonic foraminiferal taxonomy was developed for the *CHRONOS* Database System (www.chronos.org/) and it has been extensively used by specialists since 2005. It includes >1100 species from the Mesozoic, Paleocene and Eocene, and ca. 3000 images (mostly with 3 or more views of a single specimen). Once the foraminiferal database became available online participants of planktonic foraminiferal working groups added numerous species records, images and other content, including species' original descriptions, notes on ecology and taxonomy, a standardized morphological character matrix with 20 fields, original and new illustrations of type-images (holotypes, paratypes, hypotypes) with additional images for many taxa, and searchable age/stage/biozone assignments.

Premature loss of funding for the *CHRONOS* initiative prevented completion and programing updates to the foraminiferal database and it is now in jeopardy of becoming permanently inaccessible. We now have a plan to extract the database and associated images from the *CHRONOS* host server and migrate it to the *MySQL Nannotax* platform (http://ina.tmsoc.org/Nannotax3), which is widely used as an open-access taxonomic resource for the study of calcareous nannofossils. Building this new "*Planktonic Foraminifera@Nannotax*" database with an open access architecture will allow continual expert improvements, and the built-in capability of community input will ensure that the database is constantly improved and updated for future generations of end-users. Members of the expert working groups will be notified each time a new public comment is added and the cumulative comments will be visible to all users, which will help identify species concepts that may be controversial and unstable.

The project will forge an international partnership between planktonic foraminiferal specialists with outstanding taxonomic expertise and calcareous nannofossil workers who developed *Nannotax.* Further, this initiative provides the opportunity to rejuvenate the planktonic foraminiferal working groups that provided the expertise used to decide synonymy and content for the earlier version of the database.

Meetings of the Mesozoic and Neogene Planktonic Foraminiferal Working Groups are essential to resolving important debates on synonymy and phylogeny and will provide the impetus to populate the database with updated species diagnoses, synonymies, descriptions, and phylogenetic relationships. Password-protected access by working group members will allow multi-user uploads of content, and open-access program coding will ensure that the database can be continually updated and improved by the working group participants. This renewed synergy will ultimately lead to publication of taxonomic atlases for the Jurassic through Cretaceous and the Neogene planktonic foraminifera, which have been in need of revision for several decades. Updates of content published in the new taxonomic atlases and other information added to the *Planktonic Foraminifera@Nannotax* database will vastly improve the accuracy and coverage of taxonomic and age-range information, provide additional search fields and capabilities that will aid in species identifications, and will serve as the primary resource for planktonic foraminiferal taxonomy and biostratigraphy for present and future generations of academic and private industry end-users.

NOTES

eMesozoic: digitising historical palaeontological collections

Philippa Brewer, Lil Stevens, Emma Bernard, Sandra Chapman, Lorna Steel, David Smith and Molly Clery

The Natural History Museum, Cromwell Road, London, SW7 5BD, UK
email: p.brewer@nhm.ac.uk

The Natural History Museum London (NHM) is embarking on the boldest digitisation project of its type in the world. Our collection comprises 80 million objects and is widely considered as one of the most internationally important natural history resources because of its taxonomic, geographical and historical breadth. Now that digital technologies offer the opportunity to unlock this treasure trove of information at scale, we endeavour to image and transcribe data from millions of our specimens, making our collections openly available to global scientific, entrepreneurial, academic and public audiences.

eMesozoic is a pilot project within this context. The primary aim of the project is to develop digitisation workflows by focussing on 18,000 vertebrate specimens from the British Mesozoic. These specimens were selected as they present a wide range of challenges to mass digitisation including complex historical registration practices, varying levels of documentation, information in hard copy and electronic formats, and a wide range of specimen sizes, types and storage methods.

The project utilised web-based forms linked to the NHM's collection management system (KE-EMu). One of the key aims was to trial the use of Master records to aid transcription of specimen data. By filtering only high quality records on taxonomy, sites and stratigraphy through to our transcription apps we aimed to dramatically reduce the need for quality checking and allow data to be pushed quickly to the web. Another implication of this methodology is the potential to share master records between digitisation programmes, thus reducing the burden on individual digitisation programmes to clean and verify specimen data.

The eMesozoic project has already produced a wealth of data and lessons learned. Major issues encountered include dealing with legacy issues, the critical importance of pre-digitisation collection preparation, the need for multiple workflows to encompass the wide range of specimen sizes and types, and the slow speed (and therefore high cost) of digitising palaeontological specimens. Positive outcomes include the successful use of Master records to aid transcription, the positive impact on working relationships and newly shared practices within the department, the development of apps which can be used outside digitisation programmes, and of course, the specimen data itself. Overall eMesozoic has been a very positive experience, and we have gathered valuable data on estimating the time, effort and resources required to digitise historical palaeontological collections on a much larger scale.

NOTES

The Panama Paleontology Project Database

Rachello-Dolmen, P.G.[1,2] O'Dea, A.[1], Grossman, E. L.[2]

1. Smithsonian Tropical Research Institute, PO Box 0843-03092, Panama City, Panama.
2. Dept. of Geology & Geophysics, Texas A&M University, College Station, TX 77843.
email: pgrachellod@gmail.com

The Panama Paleontology Project (PPP) broke new ground in the rigorous quantitative analysis of evolutionary and ecological change over geological time in a highly biodiverse tropical setting. Many advances were made but most studies within the PPP have focused on a single taxonomic group or a distinct metric of environmental or ecological change. Our aims is to produce a synthetic analysis of temporal and spatial changes across entire biological communities using the combined efforts of over 100 workers in the PPP over the last 30 years. To this goal we have designed the PPP Database (PPPD). This relational database will compile quantitative taxonomic, ecological and functional biotic data within stratigraphic and geochronological context for comparison with environmental data as derived from high-resolution geochemistry ($\delta^{18}O$ and $\delta^{13}C$ isotopes). PPPD users will have rapid access to a large range of standardized information; will allow downloads and direct analyses to be performed on its contents, interactive access for browsing the contents and formulating specific queries. The PPPD is linked to R to facilitate exploration of patterns of biotic and environmental change in tropical marine ecosystems over geological time.

NOTES

Challenges to 3-D digital data publication in palaeontology and the need for community-wide standards for data verification

Thomas G. Davies, Imran A. Rahman, Stephan Lautenschlager, John A. Cunningham, Emily
J. Rayfield and Philip C.J. Donoghue

School of Earth Sciences, University of Bristol, Bristol, UK
email: thomas.davies@bristol.ac.uk

The generation and analysis of three-dimensional (3-D) digital models in palaeontology has facilitated significant advances in the ways in which fossil specimens can be interrogated to inform on past life. Tomographic techniques (e.g. X-ray computed tomography) and surface- based methods (e.g. photogrammetry) provide high-resolution data that have opened up avenues of study that were previously impossible, providing unprecedented insights into the anatomy, preservation and functional morphology of extinct species. Publishing 3-D data openly has enormous value in facilitating synthetic studies of multiple taxa and enabling the use of such data by educators, and having mechanisms for this is fundamental to the continued strength of the discipline as a whole; not least because of current and proposed research council policies on data accessibility. Challenges to open data publication include: lack of standardised principles on type and number of files to be published, total file sizes and big data issues, outdated institutional policies on data ownership, questions over embargo periods, and investment in preparation and curation of files, including accurate recording of metadata.

We offer a perspective on the expectations that should be held of studies working with these types of data in order to maintain scientific integrity by allowing verification of results, as well as enabling the full utilization of the data that has been collected. We propose that raw tomographic data be published as full resolution .tif image stacks, that any three-dimensional models analysed in a study be published in a freeware compatible format such as .stl, and that all files should be accompanied by detailed metadata records. Furthermore, publication of working or project files, either as freeware compatible (ideally) or proprietary formats (otherwise), would allow full verification of the approach used to transform original images into 3-D models. Ultimately, a single archive modelled on equivalents from other disciplines (e.g. GenBank) would provide an extremely valuable resource to the community, perhaps integrated alongside databases of a variety of other types of palaeontological data. However, significant discussions are required to reach that point. In the intervening time, a variety of accessible repositories are already available to palaeontologists and there is thus no tenable excuse for the majority of researchers not to publish the underlying data upon which any conclusions are drawn.

NOTES

The fossil calibration database

Philip Donoghue
on behalf of NESCent Working Group: Synthesizing & Databasing Fossil Calibrations

School of Earth Sciences, Life Sciences Building, Tyndall Avenue, Bristol BS8 1TQ, UK;
email: phil.donoghue@bristol.ac.uk

Traditionally, the fossil record provided the only insight into the antiquity of evolutionary lineages, however, this role has been usurped entirely by the molecular clock. This is because the fossil record provides only direct insight into the minimum age of clades and their degree to which these minima approximate the true time of divergence will vary with the quality of the rock and fossil records. However, molecular data alone can only establish the length of branches in evolutionary trees; if these are to be calibrated to geological time prior information on at least the approximate age of clades is necessary. Calibration can be achieved by using legacy rates of molecular evolution, from geological evidence for the age of biogeographic constraints on clade origin or, most commonly, fossil evidence. Minimum age constraints are easy enough to establish, but molecular clock calibrations must encompass the true time of divergence and so it is necessary also to establish maximum constraints on clade ages which can be based on a careful analysis of the relationship between stratigraphic evidence and sedimentary facies or, more usually, by simply fitting an arbitrary parametric distribution to the minimum fossil age constraint to express some visceral perception of the degree to which the oldest fossil record of a clade approximates its true antiquity. Unfortunately, there is rarely any objective means of choosing among such prior probability distributions and yet they have an undue impact on the resulting molecular clock estimates. Furthermore, fossil calibrations have been poorly justified and have not kept up with the development of the geological timescale. This is not surprising since molecular clock analyses are usually performed by molecular biologists whereas fossil calibrations require expert knowledge of fossil groups, phylogeny, fossilization, and the rock record.

The Fossil Calibration Database (http://fossilcalibrations.org), is a new open-access resource providing fossil calibrations to the community that have been established and justified by experts and evaluated by peer review. Calibrations accessioned into this database are based on individual fossil specimens and follow best practices for phylogenetic justification and geochronological constraint. The associated Fossil Calibration Series, a calibration-themed publication series at *Palaeontologia Electronica*, serves as a peer-reviewed pipeline to the database.

NOTES

Integrated management of biodiversity and paleontology data within GFBio and PANGAEA

Robert Huber
MARUM, University of Bremen, Bremen, Germany
email: rhuber@uni-bremen.de

Proper handling of organism names and their classification is one of the most common problems for earth science and paleontological data management. Taxon names, along with geographic constraints, are often used as a simple starting point for researchers to find relevant data sets. Therefore, taxon names are a key element for metadata describing environmental data. Based on a large, interdisciplinary user community, PANGAEA (jointly operated by the MARUM, Bremen and the AWI, Bremerhaven) has safeguarded a substantial multidisciplinary collection of environmental data during the last decades including a broad spectrum of biological as well as paleontological data. The biological content of PANGAEA includes diversity, abundance and distribution data of organisms as well as a large amount of taxon related data such as isotope measurements.

However, favoring a multidisciplinary and generic architecture capable to store data from many scientific domains, PANGAEA has not implemented specialized ontologies such as taxonomic hierarchies or synonymies in the past. Within the **German Federation for the Curation of Biological Data** (GFBio) the integration of specialized taxonomic ontologies is a key objective for PANGAEA. Our approach is largely based on semantic analysis of PANGAEA's content, the integration of tools provided by other GFBio partners as well as the usage of external 'authority' ontologies such as WORMS.

GFBio brings together national key players providing environmentally related biological data and. The overall goal is to provide a sustainable, service oriented, national data infrastructure facilitating data sharing and stimulating data intensive science in the fields of biological and environmental research. The federation will build on proven data archiving infrastructures and workflows such as those of PANGAEA for environmental data and the resources of Germany's major natural history collection data repositories. The new infrastructure will improve and integrate these existing components within a common technological and organizational framework.

The planned infrastructure will address data management requirements of a large range of stakeholders including individual scientists, natural history collections as well as larger research networks. It will be capable to handle highly interdisciplinary heterogeneous data. Special consideration will be given to genome data, ecological and environmental data as well as collection related data. Services provided by the GFBio will cover the full life cycle of research data from field or real time data acquisition to long term archiving and publication as well as analysis and re-use of these data.

NOTES

WoRMS goes Palaeontology – integration of extinct taxa in the Aphia database

Andreas Kroh
WoRMS Steering Committee Member & Taxonomic editor (Echinoidea)
Natural History Museum Vienna; Burgring 7, 1010 Vienna, Austria
email: andreas.kroh@nhm-wien.ac.at

Coverage of WoRMS increasingly extends to the fossil realm, with taxa flagged as "fossil only" becoming more and more commonly entered by taxonomic editors. The reasons for doing this are manifold. Some fossil species were included because they are type-species of genera containing extant taxa and some fossil genera have been entered because at some point in taxonomic history extant species have been attributed to them. Additionally, the naming of fossil and extant taxa (in zoology) is governed by the same set of rules (the International Code of Zoological Nomenclature). Having fossil species names in WoRMS could thus prevent accidental creation of junior homonyms.

It is unlikely that all fossil marine species will ever be entered in WoRMS and this clearly is not a priority of the project. In some groups, especially small ones, however, a decent level of coverage might be reached. Currently there are 19 phyla within the WoRMS Aphia-database which contain one or more fossil species. The editors of three phyla – Foraminifera, Mollusca and Echinodermata are actively working on the fossil species within their groups. These three groups act as 'test-cases' for the module used to capture fossil range data within WoRMS (currently data from this module is only displayed on the GSD webpages of these groups, but not on the main site).

Clearly WoRMS does not intend to replicate the effort of other Projects like the *Paleobiology Database* and records stratigraphic ranges rather than individual occurrences. The stratigraphic resolution for many fossil taxa, particularly such that have never been re-described since their first use, is often low. Without detailed knowledge of the geology of the region a particular taxon occurs in it often is difficult to correlate outdated stratigraphic terms with the current international classification. WoRMS therefore aims at a resolution down to stage level, which according to Zalasiewicz et al. (2004) are the smallest globally applicable chronostratigraphic units of the Phanerozoic. The mean duration of stages is more or less constant at about 5 Ma (Odin & Odin, 1990), with a variation between around c. 2.5 to 10 Ma. Since species ranges have a similar order of magnitude, stages are appropriate units to capture these distributions. The hierarchical system implemented in the fossil module of WoRMS allows capturing stratigraphic ranges at less detailed resolution as well.

As with any information within WoRMS, stratigraphic range data is tied to sources (publications). Two text fields associated with each stratigraphic record allow to record: A) the verbatim citation of the stratigraphic term actually used within the original source (e.g. outdated stage names) and B) notes relevant to the correlation of a given record.

References

Odin, G.S. & Odin, C. (1990). Échelle numérique des temps géologiques, mise à jour. Géochronique, 35: 12-21.

Zalasiewicz j., Smith, A., Brenchley, P., Evans, J., Knox, R., Riley, N., Gale, A., Gregory, F.J., Rushton, A., Gibbard, P., Hesselbo, S., Marshall, J., Oates, M., Rawson, P. & Trewin, N. (2004). Simplifying the stratigraphy of time. Geology, 32/1: 1-4.

NOTES

Analysis of Mesozoic–early Paleogene terrestrial vertebrate diversity demonstrates the utility of the Paleobiology Database for deep time macroevolutionary research

Richard J. Butler[1], Roger B. J. Benson[2], Philip D. Mannion[3]
[1]School of Geography, Earth and Environmental Sciences, University of Birmingham, Birmingham
[2]Department of Earth Sciences, University of Oxford, Oxford
[3]Department of Earth Science and Engineering, Imperial College, London
email: R.Butler.1@bham.ac.uk

The Paleobiology Database (PBDB) is a non-governmental, non-profit public resource for paleontological data. It has been developed over close to 20 years by an international group of palaeobiological researchers with the aim of providing global, collection-based occurrence and taxonomic data for organisms of all geological ages. More than 380 scientists from five continents have participated in the development of the PBDB, entering >57,000 references, >175,000 collections, and >330,000 taxa by January 2016. Nearly 250 official PBDB publications have been published, and many other publications and research projects have utilised PBDB data. Here, we provide an overview of the structure of the PBDB and the types and quantities of data that it contains. We demonstrate the utility of the PBDB for researchers through examples of our recent and active research projects that use the PBDB to examine diversity patterns among Mesozoic and Cenozoic vertebrates.

We quantify patterns of vertebrate standing diversity on land during the Mesozoic–early Paleogene interval, in order to test opposing models of expansionist versus constrained diversification. We applied sample-standardization to a global fossil dataset, derived from the PBDB and compiled by ourselves and other PBDB members, containing 27,260 occurrences of 4,898 non-marine tetrapod species. Our results show a highly stable pattern of Mesozoic tetrapod diversity, underpinned by a weakly positive, but near-zero, long-term net diversification rate over 190 million years. Diversity of non-flying terrestrial tetrapods less than doubled over this interval, despite the origins of exceptionally diverse extant groups within mammals, squamates, amphibians and dinosaurs. Therefore, although major groups of modern tetrapods have Mesozoic origins, rates of Mesozoic diversification inferred from the fossil record are slow. An apparent four-fold expansion of species richness after the Cretaceous/Paleogene (K/Pg) boundary deserves further examination in light of potential taxonomic biases, but is consistent with the hypothesis that global environmental disturbances such as mass extinction events can rapidly adjust limits to diversity by restructuring ecosystems.

NOTES

The Geobiodiversity Database and its Taxonomy and Occurrence Data

JUNXUAN FAN[1], XUDONG HOU[2], JIAOYANG[1], QING CHEN[2]

1 State Key laboratory of Palaeobiology and Stratigraphy, Nanjing Institute of Geology and Palaeontology, Chinese Academy of Science, Nanjing 210008, China
2 Key Laboratory of Economic Stratigraphy and Palaeogeography, Chinese Academy of Sciences, Nanjing 210008, China
email: fanjunxuan@gmail.com

The Geobiodiversity Database (GBDB, http://www.geobiodiversity.com) is an integrated system for the management and analysis of stratigraphic and paleontological information. Its goal is to facilitate regional and global collaborations focused on regional and global correlation, quantitative stratigraphy, systematics, paleogeography and paleoecology. It became the formal database of the International Commission on Stratigraphy (ICS) in August 2012 and that of the International Palaeontological Association (IPA) in February 2015, and produces comprehensive and authoritative web-based stratigraphic and palaeontologic information services for global geoscientists, educators and the public. As of January 2016, over 13,000 sections, 70,000 collections and 358,000 fossil occurrences of over 100,000 taxa from around the world have been digitized and compiled into the GBDB system (Table 1). GBDB provides the capability of integrating different interpretations to the same paleontological and stratigraphic content. Many visualization and analytical tools has been developed or integrated to make the database more useful as a scientific and educational tool, such as geographical visualization (GeoVisual 1.0), stratigraphic visualization (powered by TS Creator), panoromic views of outcrops (GeoPano 1.0), quantitiative stratigraphic correlation (CONOP 9, SinoCor 4.0 and UA) and paleogeographic reconstruction (PaleoGIS online 1.0).

Table 1. Records of 10 major fossil groups in GBDB as of January 2016.

Major fossil group	Number of taxa	Number of occurrences
Brachiopoda	14,417	58,311
Trilobita	11,133	35,152
Anthozoa	10,441	23,218
Graptolithina	7,954	69,327
Crustacea	7,429	13,527
Foraminifera	7,279	47,565
Cephalopoda	5,519	12,702
Bivalvia	5,321	15,780
Conodonta	4,258	36,176
Gastropoda	2,881	4,686

NOTES

Synthesising Biostratigraphic Data; a proposed method for common practice and examples of uses

Monika Dlubak, Mike Bidgood, Pollyanna Maksymiw, Chris Robson and Amy Taylor
Neftex Exploration Insights, Abingdon, Oxfordshire, UK
email: Monica.Dlubak@halliburton.com

"*Synthesising paleontological occurrence data and taxonomy into useable databases and web-systems will be one of the major challenges for palaeontology over the next couple of decades*".

Dealing with vast amounts of public domain paleontological (biostratigraphic) data – often across multiple platforms and a sometimes bewildering array of formats – in a coherent way is indeed challenging. Issues such as the locating of the data and evaluating its quality, relevance, provenance and the format of delivery are only some of the hurdles that need to be overcome before any meaningful work may be carried out. The next challenge is to find a "universal framework" upon which biostratigraphic events may be hung, and with what confidence. The third major challenge is to standardise nomenclature which requires a practical mechanism for dealing with synonyms. Once biostratigraphic events are placed within a standard framework, a method is needed to usefully interrogate, filter and visualise the data.
Neftex Exploration Insights has developed such a workflow over the past 10 years. Using data from almost 6,000 public domain sources ("schemes") with global distribution, we have calibrated over 275,000 individual biostratigraphic "events" (species tops & bases and biozone bases) into a coherent stratigraphic framework.

Our framework comprises standard biozonal schemes: planktonic foraminifera (Cenozoic); Tethyan ammonites (Mesozoic); conodonts (Permian and Devonian); benthic foraminifera (Carboniferous); graptolites (Silurian and Ordovician); trilobites, archeocyathans, trace fossils and others (Cambrian and Precambrian). Each local bioevent or biozone is calibrated to a position within a standard biozone which is in turn calibrated to the latest Geological Timescale (Gradstein et al., 2012). This allows the database to react dynamically to timescale changes.
Taxa used in schemes retain the original author's names but these names are managed in a Species Dictionary. This enables the spacial and temporal searching and display of taxa irrespective of synonym used. Results can be viewed via web delivery or by importing datapacks into proprietary packages such as TSCPro© or StrataBugs©.
Calibrated data can be used for identifying regional and global sequence surfaces, characterising systems tracts and refining Gross Depositional Environment maps. Synthesising data sets from individual basins or regions allows a rapid evaluation of all publically-available biostratigraphic data saving significant work-years for explorationists.

NOTES

Specimen-level mining: bringing knowledge back 'home' to the Natural History Museum, London

Ross Mounce, Aime Rankin
Natural History Museum,Cromwell Road, London, SW7 5BD, UK
email:Ross.Mounce@nhm.ac.uk

The right to read is the right to mine.

Since mid-2014, the UK has had a specific copyright exception that permits researchers to do text and data mining on resources they have legitimate access to, provided it is for non-commercial research purposes. We have been taking advantage of this new exception to systematically examine the last 16 years' worth (2000-2015) of research, published across many hundreds of journals, constituting over a million full text articles in total. This is something that simply cannot be done with traditional 'title & abstract' searches. Unrestricted full text search is a new paradigm for literature synthesis.

Initially, we are mapping where Natural History Museum, London specimens are cited in the literature; linking these mentions up with the official specimen record on the NHM Data Portal (http://data.nhm.ac.uk/). We leverage both the CrossRef Metadata API, as well as the NHM Data Portal API for specimen lookup. With over 20,000 specimens found so far, we can demonstrate specimen-usage through time, the top 10 most cited NHM specimens, and comparisons between taxonomic groups. This approach also highlights what the NHM Data Portal specimen records are missing: many specimens are found mentioned in the literature, for which no information exists on the portal. This 'absence' information could help prioritize digitization efforts towards research-valuable specimens, that have and are being used in recent research. Finally, we suggest ways which in specimens should (and should not!) be cited in modern, digital research papers.

NOTES

GeoDeepDive: a cyberinfrastructure to support the construction and enhancement of literature-derived databases

Shanan E. Peters[1], Ian A. Ross[2], John Czaplewski[1], Miron Livny[2]
[1]Department of Geoscience, University of Wisconsin-Madison
[2]Computer Sciences Department, University of Wisconsin-Madison
email: peters@geology.wisc.edu

Literature-derived databases are a critical component of modern paleobiological research. Like all data from any source, literature-based databases are subject to errors and data quality concerns, but they are nonetheless useful for addressing many fundamental questions, ranging from biotic effects of global climate change to calibrations of molecular clocks. The Paleobiology Database (https://paleobiodb.org), a non-profit public paleontological resource assembled over some 20 years by an international team of paleontologists, is among the most impactful of databases of its kind. Despite demonstrable scientific successes, the PBDB has many limitations, most of which are related to the effort required to manually find and keystroke new data and the difficulty involved in extending the database to include new information, such as specimen-based data and geological information.

Here we describe GeoDeepDive (https://geodeepdive.org), a cyberinfrastructure designed to support the discovery and acquisition of data in the published scientific literature. The first component of this infrastructure involves rate controlled, secure fetching and storage of original digital documents (in particular, PDFs) and associated bibliographic information. The second component involves automated linguistic and layout parsing of documents using a variety of software tools and a High Throughput Computing (HTC) system. New tools can be rapidly deployed on all documents as they become available. An ElasticSearch implementation enables full-text search and facilitates pre-indexing of documents using external databases (e.g., PBDB taxonomic names accessed via the Application Programming Interface). Basic library information and full-text search results are publicly exposed via the GeoDeepDive API. All of these components combine to help characterize documents and to generate raw datasets which can be used to develop customized applications, such as those built using the DeepDive machine reading system (http://deepdive.stanford.edu).

To date, our infrastructure has harvested nearly 600,000 documents from Elsevier, Wiley, and the USGS, and more publishers will be sourced in the coming year. PaleoDeepDive, an application written to populate the PBDB, serves as a useful starting point for accelerating the pace of PBDB data acquisition, but the biggest advances will come from creatively rethinking what a literature-based paleontological database is in the era of machine reading and learning systems coupled to a dynamic digital library and HTC resources. As a simple example, we identify museum specimen numbers and their associated text and figure descriptions in the literature and relate them to potential PBDB occurrences.

Poster List

MolluscaBase – World Register of Marine Species goes fossil
Simon Schneider et al
CASP, University of Cambridge

A new era has begun: phenomenal discoveries of fossilized fleshes in Kerman, Iran
Mr. Mohsen Tajrobekar

Lagerstätten and counting taxa: a murky issue
Walker, F. M.[*], Woods, M. A., Dunhill, A. M., Benton, M. J., Newell, A. J.
[*]*School of Earth Sciences, University of Bristol.*

Poster Abstracts

MolluscaBase – World Register of Marine Species goes fossil

Simon Schneider[1], Ruud A. Bank[2], Rüdiger Bieler[3], Phillipe Bouchet[4], Wim Decock[5], Stefanie Dekeyzer[5], Serge Gofas[6], Andreas Kroh[7], Bruce Marshall[8], Thomas A. Neubauer[7], Eike Neubert[9], Gary Rosenberg[10], André F. Sartori[4], Aina Trias-Verbeeck[5], Leen Vandepitte[5], Bart Vanhoorne[5]

1 CASP, University of Cambridge, West Building, 181A Huntingdon Road, Cambridge CB3 0DH, UK; email: simon.schneider@casp.cam.ac.uk.
2 Chopinlaan 21, 9603 AM Hoogezand, The Netherlands.
3 Field Museum of Natural History, 1400 South Lake Shore Drive, Chicago, IL 60605-2496, USA.
4 Muséum National d'Histoire Naturelle; Systématique, Adaptation, Evolution, 55, Rue Buffon, 75231 Paris Cedex 05, France.
5 Flanders Marine Institute, VLIZ – InnovOcean site, Wandelaarkaai 7, 8400 Oostende, Belgium.
6 Universidad de Málaga; Faculdad de Ciencias, Departamento de Biología Animal, Campus de Teatinos, 29071 Málaga, Spain.
7 Naturhistorisches Museum Wien, Burgring 7, 1010 Wien, Austria.
8 Museum of New Zealand Te Papa Tongarewa, 55 Cable Street, PO Box 467, Wellington, New Zealand.
9 Naturhistorisches Museum Bern, Bernastrasse 15, CH-3005 Bern, Switzerland.
10 Drexel University, Academy of Natural Sciences, 1900 Benjamin Franklin Parkway, 19103 Philadelphia, Pennsylvania, USA.

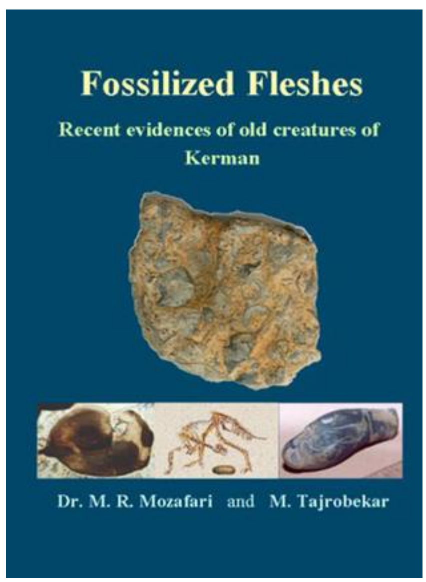
MolluscaBase (http://www.molluscabase.org) is a taxonomically oriented database which aims to provide an authoritative, permanently updated account of all molluscan species, marine, freshwater and terrestrial, recent and fossil. The database was launched on February 6-7, 2014 at the Flanders Marine Institute (VLIZ) at Ostende, host institute of the World Register of Marine Species (WoRMS). Currently, MolluscaBase contains almost 120,000 verified species names. More than 45,000 of these are valid names for Recent marine Mollusca, which are estimated to represent approximately 95% of all valid species. Beyond catching up with the missing 5%, and keeping up with the yearly established new species, the molluscan component of WoRMS is currently being expanded to contain freshwater and terrestrial Mollusca, estimated on ca 28,000 valid species. Adding the aspect of time to the current database, and acknowledging the fossil record as the source of extant diversity, it has been agreed that the list should be extended to include all fossil mollusc taxa. Almost 6,000 accepted names of fossil mollusc species have already been entered. A solid estimate of the number of fossil mollusc species is not available yet. MolluscaBase is a strictly taxonomic database, relying only on published sources. As such, it does not compete with occurrence-based initiatives, like, e.g., the Paleobiology Database, but rather supplements these platforms. WoRMS as a whole has had more than 1,100,000 unique visitors in 2014, and has currently more than 2,700 citations in Google Scholar. Moreover, it is linked to various biodiversity initiatives and repositories and is thus an important component of global biodiversity infrastructure, providing quality control and taxonomic backbone. Going fossil, the existing network shall be extended to link with global, regional or taxon-specific palaeontological databases. The fossil component of MolluscaBase is still in its starting phase, and needs your support! Data input in MolluscaBase is performed by a limited number of taxonomic editors and various corresponding contributors, who may provide datasets to the database management team. If you are a taxonomist, specializing in any group of fossil molluscs, and want to become a contributor to MolluscaBase, let us know! If you are holding your own database and want to join or link it to MolluscaBase, let us know! Please contact the WoRMS team at info@marinespecies.org, to find your role in MolluscaBase!

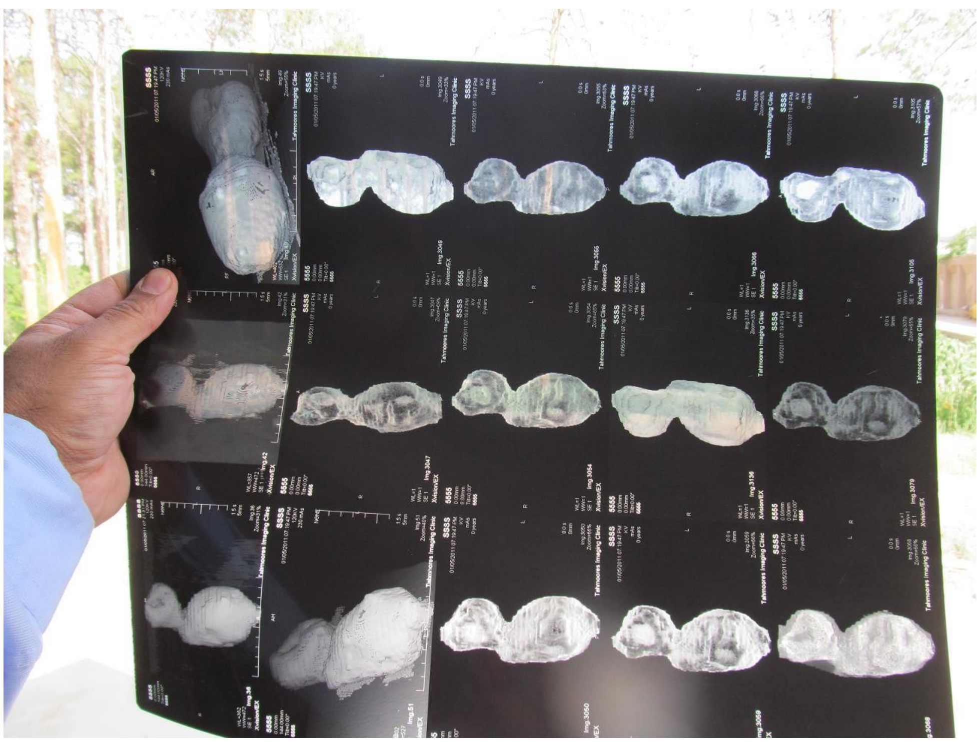
A new era has begun: phenomenal discoveries of fossilized fleshes in Kerman, Iran

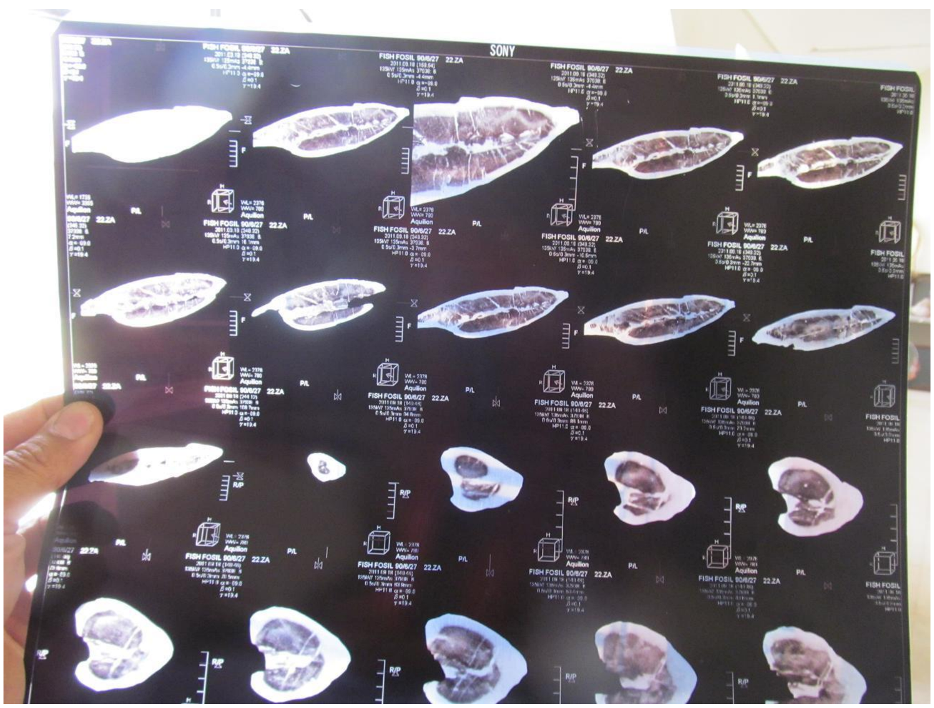
Mr. Mohsen Tajrobekar

Modern paleontology theory teaches us that fossils are traditionally the remains of the harder parts of organisms such as bones, teeth, and shells. For years, paleontologists have suggested that fossils are not usually developed from other parts of organisms such as flesh, skin and internal organs because these parts are decomposed over time and therefore destroyed. New research on fossils found in Kerman in south-eastern Iran, however, has proven otherwise. The climatic features of the Kerman region (i.e. very low humidity and a distinctive geological setting) have allowed the preservation of fossils for hundreds of millions of years. These unique fossils include a large amount of fossilized eggs as well as plant and animal flesh. Computerized tomography (CT) scans performed in Kerman have shown that fossils can be formed from softer textures of organisms such as skin, flesh, internal organs, or even the food that was eaten by an animal.

Researchers Dr. M. R. Mozafari and Mr. M. Tajrobekar have detailed these discoveries in their book *Fossilized Fleshes: Recent evidences of old creatures of Kerman* and would be very interested to present their research at the Lyell conference on Paleoinformatics in March 2016 and particularly to discuss in more detail their methodology of using CT scans to analyze these unique specimens.






SONY
FISH FOSIL 90/6/27

Lagerstätten and counting taxa: a murky issue

Walker, F. M.

School of Earth Sciences, University of Bristol. Life Sciences Building, 23 Tyndall Avenue, Bristol BS8 1TQ.
email: fiona.walker@bristol.ac.uk

Heterogeneous preservation in the fossil record is a problem for palaeobiodiversity studies. Lagerstätten, deposits with elevated preservation potential, can falsely inflate estimates of relative richness within the time bins which contain them. Diversity peaks in multiple taxa have been attributed to the lagerstätten effect. The lagerstätten effect is most extreme when there is a large range of preservation potentials across rock units, and when lagerstätten are heterogeneously distributed, both temporally and geographically. Here, we investigate how lagerstätten add to our knowledge of vertebrate diversity by counting the number of genera which only occur in lagerstätten, considering abundance distributions, and then comparing these measures between different regions. For this, we have combined taxonomic information from The Paleobiology Database with stratigraphical information obtained from a literature search, and for the UK, information from stratigraphical framework reports produced by the British Geological Survey. We find that, with the exclusion of formations which have previously described as lagerstätten in the literature, vertebrate diversity patterns in the UK are not dramatically altered. This case study will be compared with information from different regions, all of which have been sampled to differing extents by palaeontologists. However, this method throws up problems with defining lagerstätten, a fluid and subjective concept, as binary.

Burlington House
Fire Safety Information

If you hear the Alarm

Alarm Bells are situated throughout the building and will ring continuously for an evacuation. Do not stop to collect your personal belongings.

Leave the building via the nearest and safest exit or the exit that you are advised to by the Fire Marshall on that floor.

Fire Exits from the Geological Society Conference Rooms

Lower Library:
Exit via main reception onto Piccadilly, or via staff entrance onto the courtyard.
Lecture Theatre
Exit at front of theatre (by screen) onto Courtyard or via side door out to Piccadilly entrance or via the doors that link to the Lower Library and to the staff entrance.
Main Piccadilly Entrance
Straight out door and walk around to the Courtyard.

Close the doors when leaving a room. **DO NOT SWITCH OFF THE LIGHTS**.

Assemble in the Courtyard in front of the Royal Academy, outside the Royal Astronomical Society.

Please do not re-enter the building except when you are advised that it is safe to do so by the Fire Brigade.

First Aid

All accidents should be reported to Reception and First Aid assistance will be provided if necessary.

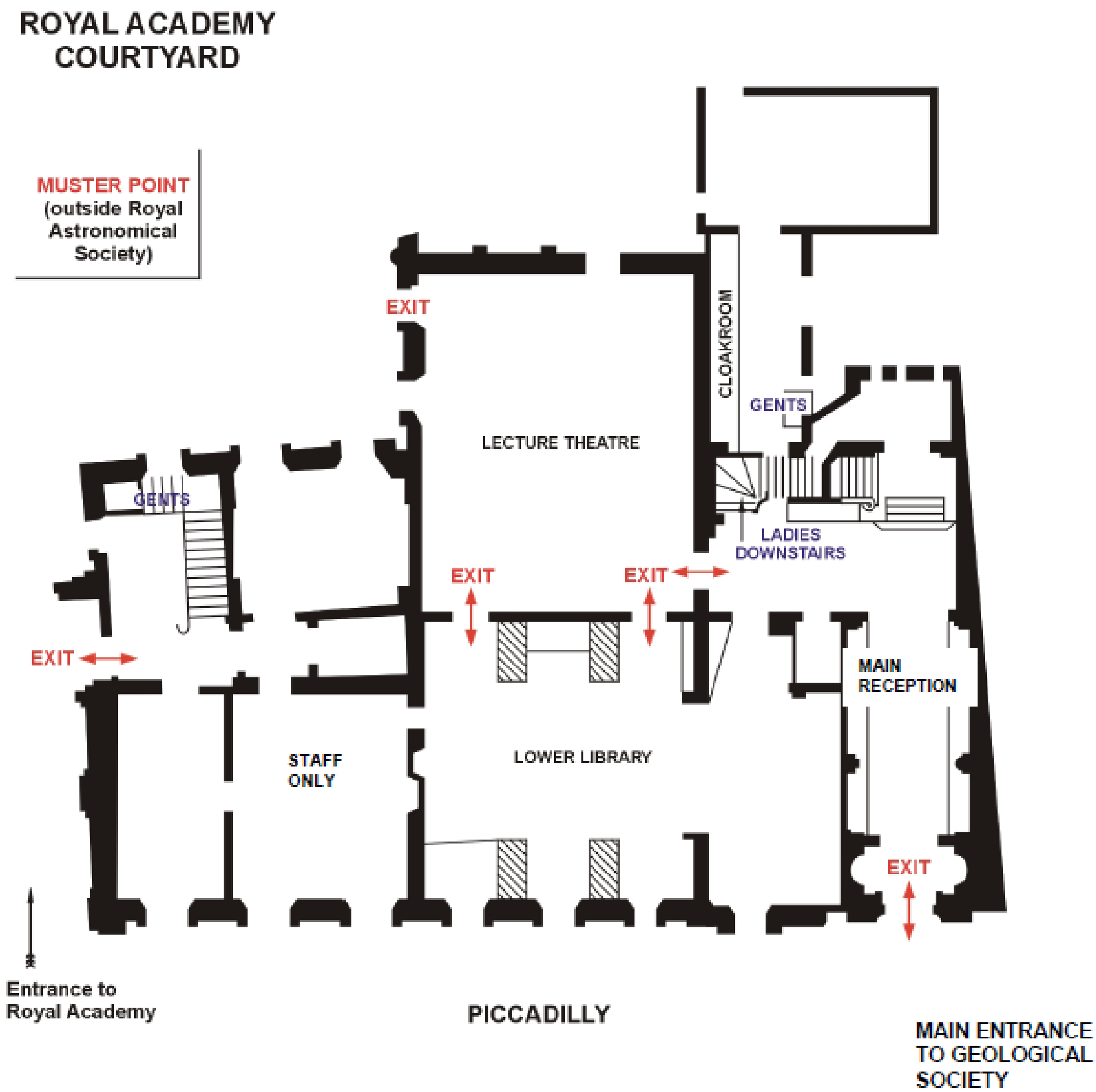
Facilities

The ladies toilets are situated in the basement at the bottom of the staircase outside the Lecture Theatre.

The Gents toilets are situated on the ground floor in the corridor leading to the Arthur Holmes Room.

The cloakroom is located along the corridor to the Arthur Holmes Room.

Ground Floor Plan of The Geological Society, London, Piccadilly



The Geological Society
serving science & profession

2016 Geological Society Conferences

16 March	GSL London Lecture – Deadly Volcanic flows: understanding pyroclastic density currents	Burlington House
22 - 24 March	Rifts III: Catching the wave	Burlington House
13-15 April	East Africa: From Research to Reserves	Burlington House
20 April	GSL London Lecture – The Water Book	Burlington House
27 -28 April	Janet Watson Meeting 2016: The Future of Hydrocarbon Exploration	Burlington House
18 May	GSL London Lecture – What Coal Mining Hyrdogeology tells about the real risk of Fracking	Burlington House
23-25 May	Arthur Holmes Meeting 2016: The Wilson Cycle: Plate Tectonics and Structural Inheritance during Continental Deformation	Burlington House
26-27 May	Palaeozoic Plays of Northwest Europe	Burlington House
2-3 June	William Smith Meeting 2016: Glaciated Margins: The Sedimentary and Geophysical Archive	Burlington House
10 June	Groundwater in Fractured Bedrock Environments: Managing Catchment and Subsurface Resources	Queen's University, Belfast, Northern Ireland
15 June	GSL London Lecture – Groundwater and its Global Significance	Burlington House
20-21 June	Martian Gullies and their Earth Analogues	Burlington House
7-9 September	Mesozoic Resource Potential in the Southern Permian Basin	Burlington House
14 September	GSL London Lecture: A little goes a long way: researching ash clouds and abrupt climate change	Burlington House
27-29 September	Rain, Rivers and Reservoirs	Burlington House
12 October	GSL London Lecture – Water on Mars	Burlington House
2-3 November	Operations Geology Conference: Bridging the Gaps	Burlington House
9 November	GSL London Lecture – Climate change and Antarctica: the great ice sheet in the past, present and future	Burlington House
9 November	GSL Nottingham Careers Day	British Geological Survey, Nottingham
23 November	GSL Edinburgh Careers Day	Our Dynamic Earth, Edinburgh
7 December	GSL London Lecture – Waking the Giant: how a changing climate triggers earthquakes, tsunamis and volcanoes	Burlington House