



Diatoms.org: supporting taxonomists, connecting communities

Sarah A. Spaulding, Marina G. Potapova, Ian W. Bishop, Sylvia S. Lee, Tim S. Gasperak, Elena Jovanoska, Paula C. Furey & Mark B. Edlund

To cite this article: Sarah A. Spaulding, Marina G. Potapova, Ian W. Bishop, Sylvia S. Lee, Tim S. Gasperak, Elena Jovanoska, Paula C. Furey & Mark B. Edlund (2021) *Diatoms.org*: supporting taxonomists, connecting communities, Diatom Research, 36:4, 291-304, DOI: [10.1080/0269249X.2021.2006790](https://doi.org/10.1080/0269249X.2021.2006790)

To link to this article: <https://doi.org/10.1080/0269249X.2021.2006790>



This work was authored as part of the Contributor's official duties as an Employee of the United States Government and is therefore a work of the United States Government. In accordance with 17 U.S.C. 105, no copyright protection is available for such works under U.S. Law.



Published online: 11 Jan 2022.



Submit your article to this journal [↗](#)



Article views: 855



View related articles [↗](#)



View Crossmark data [↗](#)



***Diatoms.org*: supporting taxonomists, connecting communities**

SARAH A. SPAULDING ^{1*}, MARINA G. POTAPOVA ², IAN W. BISHOP ³, SYLVIA S. LEE ⁴, TIM S. GASPERAK⁵, ELENA JOVANOSKA⁶, PAULA C. FUREY ⁷ & MARK B. EDLUND ⁸

¹U.S. Geological Survey/INSTAAR, Boulder, CO, USA

²The Academy of Natural Sciences of Drexel University, Philadelphia, PA, USA

³Graduate School of Oceanography, University of Rhode Island, Narragansett, RI, USA

⁴U.S. Environmental Protection Agency, Office of Research and Development, Center for Public Health and Environmental Assessment, Washington, DC, USA

⁵Strange Attractor LLC, Boulder, CO, USA

⁶Department of Palaeoanthropology, Senckenberg Research Institute, Frankfurt am Main, Germany

⁷Department of Biology, St. Catherine University, St. Paul, MN, USA

⁸Science Museum of Minnesota, St. Croix Watershed Res. Station, Marine on St. Croix, MN, USA

Consistent identification of diatoms is a prerequisite for studying their ecology, biogeography, and successful application as environmental indicators. However, taxonomic consistency among observers has been difficult to achieve because taxonomic information is scattered across numerous literature sources, presenting challenges to the diatomist. Firstly, literature is often inaccessible because of cost or its location in journals that are not widely circulated. Secondly, taxonomic revisions of diatoms are taking place faster than floras can be updated. Finally, taxonomic information is often contradictory across literature sources. These issues can be addressed by developing a content creation community dedicated to making taxonomic, ecological, and image-based data freely available for diatom researchers. *Diatoms.org* represents such a content curation community, providing open, online access to a vast amount of recent and historical information on North American diatom taxonomy and ecology. The content curation community aggregates existing taxonomic information, creates new content, and provides feedback in the form of corrections and notices of literature with nomenclatural changes. The website not only addresses the needs of experienced diatom scientists for consistent identification but is also designed to meet users at their level of expertise, including engaging the lay public in the importance of diatom science. The website now contains over 1000 species pages contributed by over 100 content contributors, from students to established scientists. The project began with the intent to provide accurate information on diatom identification, ecology, and distribution using an approach that incorporates engaging design, user feedback, and advanced data access technology. In retrospect, the project that began as an ‘extended electronic book’ has emerged not only as a means to support taxonomists, but for practitioners to communicate and collaborate, expanding the size of and benefits to the content curation community. In this paper, we outline the development of *diatoms.org*, document key elements of the project, examine ongoing challenges and consider the unexpected emergent properties, including the value of *diatoms.org* as a source of data. Ultimately, if the field of diatom taxonomy, ecology, and biodiversity is to be relevant, a new generation of taxonomists needs to be trained and employed using new tools. We propose that *diatoms.org* is in a key position to serve as a hub of training and continuity for the study of diatom biodiversity and aquatic conditions.

Keywords: *website, online flora, taxonomic consistency, content curation community, North America*

Introduction

The discovery and documentation of life on the planet Earth remain a vast endeavour. The challenge is far too large for the small number of formally trained scientists and far too complex for traditional publication (Wilson 2003, Jetz et al. 2012, Parr et al. 2014, Meyer et al. 2015). The internet empowers extended, dispersed scientific communities to collaborate over broad spatial scales, nearly instantaneously (Borgman 2007). The *Encyclopedia of Life* (EOL), a model content curation

community (Rotman et al. 2012), is at the forefront of collaboration and provides open access and engagement of the public in science (Wilson 2003). EOL facilitates access to knowledge about Earth’s life, in all of its forms, and thus, makes science more egalitarian. Data on the many aspects of organisms (e.g., names, distribution, life cycle, food webs) can be gathered together and served via platforms such as EOL. Microorganisms such as diatoms, which require high-powered microscopes for observation, present additional challenges. However, knowledge of the

*Correspondence author: Sarah A. Spaulding E-mail: sspaulding@usgs.gov

Associate editor: Rebecca Bixby

(Received 27 January 2021; accepted 22 September 2021)

This work was authored as part of the Contributor’s official duties as an Employee of the United States Government and is therefore a work of the United States Government. In accordance with 17 U.S.C. 105, no copyright protection is available for such works under U.S. Law.

This is an Open Access article that has been identified as being free of known restrictions under copyright law, including all related and neighbouring rights (<https://creativecommons.org/publicdomain/mark/1.0/>). You can copy, modify, distribute and perform the work, even for commercial purposes, all without asking permission.

basic biology and diversity of diatoms continues to be limited, even by specialists, and even within North America. Yet, knowledge of diatoms has great utility in environmental sciences (Smol & Stoermer 2010), as does an EOL-type resource that is specific for diatom specialists.

Since Ehrenberg (1854) first recognized the distinct diatom flora of western North America, floristic zones of the eastern, southeastern and northwestern parts of the U.S. have been noted (Kociolek & Spaulding 2000, Potapova & Charles 2003, Morales 2005, Kociolek 2006, Ponader & Potapova 2007). Yet, while there are a number of recent regional floras, including the Arctic (Antoniades *et al.* 2009a), Cape Cod (Siver *et al.* 2005), Montana (Bahls 2005, 2010, 2013), Great Smoky Mountains (Furey *et al.* 2011), and the Atlantic Coastal Plain (Siver & Hamilton 2011), floras are limited in terms of geographic and taxonomic scope and become outdated because of the rapid pace of revisions in classification that follow advancements in understanding phylogenetic relationships (Nakov *et al.* 2014, 2015, Ruck *et al.* 2016a, b). Furthermore, it was estimated that 20–25% of the species within the U.S. have yet to be described by science (Potapova & Charles 2003), an estimate that has increased over the past decade (Bishop *et al.* 2017a). While a number of publications address uncovering and describing new species or documenting species distributions (e.g., Spaulding *et al.* 2002, Potapova & Ponader 2004, Potapova & Winter 2006, Antoniades *et al.* 2009b, Kumar *et al.* 2009, Kociolek & Thomas 2010, Potapova *et al.* 2019a, Stone *et al.* 2020), many more taxa remain to be described.

Furthermore, advances are being made to expand the research potential of specimens in programmes such as the Extended Specimen Network (Lendemer *et al.* 2020), which link data types including genetic, phenotypic, and environmental data. Taxonomic and ecological data are crucial for water quality monitoring and assessment programmes and biodiversity inventories, including those overseen by federal agencies [e.g., U.S. Environmental Protection Agency (U.S. EPA), U.S. Geological Survey (USGS)], state and local agencies, tribal governments, and academic researchers. Taxonomy is the framework upon which biodiversity science relies (Ebach *et al.* 2011, Sluys 2013), yet taxonomy and taxonomists are as vulnerable to extinction as the organisms they study. The decline of taxonomy as a field of study is well-documented (Lücking 2008, Wheeler 2010, Sluys 2013), and taxonomy often is portrayed as ‘a science in crisis’ (Agnarsson *et al.* 2007, Rodman 2007). If *diatoms.org* is a means to continue the science of taxonomy, that outcome would be an immense accomplishment.

At the same time, trends to privatize and monetize science and scientific publishing, in general, are still strong (Jones *et al.* 2014, Ellwood *et al.* 2015, Heise & Pearce 2020) and such trends limit public access to science. When knowledge and data are more accessible to everyone, science becomes more equitable (Rotman

et al. 2012, Parr *et al.* 2014). Information on diatoms should not be held only within expensive journals, books, and scientific societies. In the case of U.S. Government-funded programmes, there are requirements to make data available online wherever practical and to make the information accessible for use by the public (OMB 2009, 2013). Meeting these requirements for diatom data collected with federal funding are indeed practical to the extent of making information fully discoverable in a convenient open format for a wide range of users. Furthermore, information can be more appealing to all types of students through the use of intentional design and compelling images.

In this paper, we provide an example of an open, online diatom flora following the model of a content curation community (Rotman *et al.* 2012). The flora is created through an aggregation of taxonomic, ecological, and distribution information that is curated by an Editorial Review Board. We outline the development of *diatoms.org*, document key elements, and examine ongoing challenges. With nearly one million page views each year, we consider the website a success. Yet, the project is not often cited as a source in journal articles and our contributors have not received a commensurate citation and professional recognition by academia. The purpose of this publication is to document the functioning of *diatoms.org*, its outcomes, and future potential.

Rationale

The primary audience of *diatoms.org* includes those who need to identify diatoms or those who seek detailed scientific information about their ecological roles in aquatic systems. This audience, which includes students and professionals in academic, consulting and public agency roles, may work as diatom analysts, academic professionals or graduate students in training. A typical diatom taxonomist spends time working at a microscope equipped with a digital camera and imaging software. Many of these users are tasked with identifying diatom species as data elements for research and monitoring projects, for which consistent taxonomic treatment is imperative. Consistent taxonomy across analysts and labs is particularly difficult because the information needed for identification (including images, dimensions, and descriptions) is dispersed across a large number of scientific publications and taxonomic volumes. Kociolek & Spaulding (2000) observed that taxon names applied to studies typically reflect the resources available to a particular person or laboratory rather than an accurate flora of the study region. In addition, the *diatoms.org* site serves as the standard for taxonomic accreditation by the Society for Freshwater Science genus certification programme (Alers-Garcia *et al.* 2021), with a species certification programme in progress.

The secondary audience, which includes people who spend little or no time at a microscope, includes scientists

or managers who wish to access the information associated with diatom species, such as the water conditions in which a species grows, or the definition of a term from the glossary. Students (primary and secondary school, university) and other public users form part of this audience. The site attracts over 70% of its visits from the lay public, such as people searching on the phrase ‘what is a diatom?’ Thus, *diatoms.org* opens the world of diatoms to a diverse, international audience by conveying biologic information in non-specialist terminology. Although we label this audience ‘secondary’, the outcomes from visits from this audience may be even more consequential. Biodiversity is crucial to society, and it is this audience that has the ability to effect species conservation and ecosystem management (Jetz et al. 2013, Meyer et al. 2015, Navarro et al. 2017).

The website project is intended to fill a gap in the formal scientific literature by serving diatom taxonomists, ecologists, and water quality managers. Each taxon page represents a synthesis of information that is otherwise time-consuming, or resource-prohibitive, for an individual user to obtain. The resource is based on (1) early efforts to prioritize design elements and (2) ongoing attention to the careful curation of content. Electronic floras have several advantages over traditional floras (i.e., books). Firstly, content can be continually improved through updates that incorporate new knowledge, expertise, and refinement of nomenclature based on systematics. When they are detected, errors can be corrected, and revisions to pages can be archived and tracked. The digital format also allows flexibility in the way that content can be displayed.

Names

Taxa included within *diatoms.org* were initially drawn from lists developed through the U.S. Geological Survey National Water Quality Assessment program. In the process of developing species pages for the project, however, nomenclatural errors and inconsistent identifications emerged in the list. Furthermore, the project began with a bias towards taxa reported from streams and rivers. As contributors have worked on projects on lakes, wetlands, and estuaries, the habitat coverage has expanded. Now, contributors use, improve, and update the list of taxa simultaneously with new taxon pages and updates to taxon pages.

Operating principles

The project adopted many of the principles outlined in the EOL for curating online taxonomic content (Parr et al. 2014). An open process was established to include contributors, develop taxon pages, develop site design, manage reviews, and respond to notification of errors. Editorial Review Board members regularly hold

meetings, workshops, and requests for feedback. Community listening sessions have included the biennial North American Diatom Symposium and International Society for Diatom Research (Edlund et al. 2019). The operating principles cover project oversight, taxon designation, specimen access, species concept, funding priorities, and scientific review.

Project oversight

The process for the inclusion of new pages and activities of the Editorial Review Board aims to be transparent and to avoid top-down management. Feedback from the scientific community, as well as forums for participation by water managers provide differing views on priorities and needs (Lee 2019).

Taxon designation

The project adheres to the International Code of Nomenclature (ICN) for algae, fungi, and plants (Turland et al. 2018) and does not bypass formal scientific publication in a peer-reviewed journal. Beginning in 2012, the ICN included electronic publication as constituting valid publication, under certain parameters:

29.1. Publication is effected, under this *Code*, by distribution of printed matter (through sale, exchange, or gift) to the general public or at least to scientific institutions with generally accessible libraries. Publication is also effected by distribution on or after 1 January 2012 of electronic material in Portable Document Format (PDF; see also Art. 29.3 and Rec. 29A.1) in an online publication with an International Standard Serial Number (ISSN) or an International Standard Book Number (ISBN).

While *diatoms.org* could serve as an outlet for formal taxonomic publication, it has not included new descriptions to date.

Specimen access

One of the requirements of *diatoms.org* is that light and/or scanning electron micrograph images are from specimens deposited in a public herbarium. Thus, anyone should be able to obtain the physical specimen of any image in *diatoms.org*. Applied taxonomic names should be validly published and correct. We recognize, however, that there will be errors in the application of names. Those errors are corrected through updates, as they are determined. For example, the taxon page for *Encyonema stoermeri* Spaulding, Pool & Castro was published (Spaulding et al. 2010, Spaulding 2010a), but the type had been previously published as *Encyonema temperei* Krammer (Krammer 1997). The scientific community reported the error, and an update was made to correct the name to *E. temperei*.

Species concept

We propose a common philosophy recognizing species as “morphologically distinct groups”, however, morphological species may not, in many instances, align with molecular concepts of species or phylogenetic species. We adopt a morphological species concept as an operational definition. In cases of questionable taxonomic affiliation, a ‘sort it out’ philosophy will prevail because it is easier to subsequently combine taxonomic errors resulting from splitting morphological specimens too finely, rather than being able to recover coarser groups. As we are learning, it is expensive and difficult to recover and rectify errors when different entities are combined that should have been treated separately (Lee *et al.* 2019).

Funding priorities

The selection of taxa for the development of taxon pages has been prioritized based on relative abundance, indicator status, and difficulty in identification for USGS and U.S. EPA monitoring programmes (Stoddard *et al.* 2005, Lee *et al.* 2019). To date, the funded development of content for species pages has followed two models, (1) direct funding of an expert for a specified group of taxa and (2) travel funding for contributors to work in collaboration, using the resources of the Diatom Herbarium at the Academy of Natural Sciences of Drexel University, Philadelphia, Pennsylvania, USA. Contributions have been made based on other motivations, including (1) personal research and exceptional dedication (e.g., the hundreds of species pages contributed by Loren Bahls, University of Montana), (2) production of pages by students, such as more than 100 species pages contributed by students in the *Ecology and Systematics of Diatoms* course at Iowa Lakeside Laboratory, and (3) outcomes linked to research grant requirements for broader impacts (Furey *et al.* 2011).

Scientific review

Members of the Editorial Review Board and recognized national and international specialists serve as referees for the online taxon pages. The review process follows the model of scientific peer-review, except that the referees are not anonymous. The curation of scientific content by contributors and reviewers is acknowledged throughout the site. The recognition is meant to support openness, to document contributions by reviewers, and for contributors, reviewers, and editors to share the responsibility for accurate information.

Framework

Ideally, the organizational framework and navigation of *diatoms.org* would be based on a consensus phylogenetic tree of diatoms. However, there are such large uncertainties around diatom relationships that current phylogenies

(Theriot *et al.* 2010, Williams & Kociolek 2010) are too broad to provide an effective organizational structure. As a result, the higher-level organization is artificially partitioned into morphological groups. Advances are being made in using phenotypic data from species descriptions to develop phylogenetic matrices (Cui *et al.* 2010, Endara *et al.* 2018) and we are testing computer-ready versions of *diatoms.org* species descriptions as a source of phylogenetic data (T. Nakov, pers. comm., University of Arkansas, 2019). Within the morphological group framework, the project implements visual and nomenclatural cues and a search function to aid practitioners in reaching a correct identification. Species pages serve as the core of the project and are organized in a specific manner.

Visual and nomenclatural cues and search

Clearly, websites offer greater flexibility in the display of information than books. With *diatoms.org*, genera and species can be displayed by name, or by an iconic image, facilitating recognition based on the shape and size of the taxon. Genus member species may be filtered by valve length, stria density, or both, allowing practitioners to narrow a search to a smaller number of taxa. *Diatoms.org* also incorporates composite guide images that direct practitioners to the features that diagnose one taxon from all others. Finally, the *Compare* feature allows users to make direct, visual comparisons of similar taxa, as well as taxa that are often confused in surveys. The *Compare* feature allows a side-by-side examination of diatom images at a common magnification, a feature not possible with books, or other online floras.

The *Genus Considered* field communicates anticipated nomenclatural transfers. For example, *Navicula hodgiana* R.M. Patrick & Freese lacks the features of the genus *Navicula* Bory (Hamilton & Manoylov 2021) and instead, appears to be more aligned with the genus *Fallacia* Stickle & D.G. Mann in Round *et al.* *Genus Considered* field allows the taxon to be included in both *Navicula* and *Fallacia* genus pages until a formal nomenclatural act is published. As evolutionary relationships among diatoms are uncovered (Theriot *et al.* 2010, Nakov *et al.* 2014, 2015, Ruck *et al.* 2016a), nomenclatural changes follow (Ruck *et al.* 2016a, b, Rarick *et al.* 2017). However, in many instances, there is a time lag between the publication of phylogenies and nomenclatural change in the literature. The *Genus Considered* field allows a species to be included within its validly published genus, as well as listed under the ‘anticipated’ genus. Thus, the site allows the greater ability for practitioners to follow changing nomenclature. Furthermore, the *Search* feature allows users to find a taxon under nomenclatural synonyms. For example, *Luticola goeppertiana* (Bleisch) Mann ex Rarick *et al.* (Rarick 2017) can be searched as its current name, or its other validly published names (i.e., *Navicula goeppertiana*

(Bleisch) H.L. Smith, *Navicula mutica* var. *goeppertiana* (Bleisch) Grunow).

Species pages

The species page represents the core of the *diatoms.org* project, with sections dedicated to particular data formats. The current, correct, and complete scientific name of the taxon is included in the *Description* section, along with objective, or homotypic, synonyms. In the early development of the online flora, we learned that users sometimes knew a diatom by a name that had been applied incorrectly. That experience inspired the creation of a *Reported As* field as a way to acknowledge a previous record, even if a name other than the current name was used. For example, Kociolek (2011) noted that *Thalassiosira lacustris* (Grunow) Hasle had been reported as *Thalassiosira bramaputra* (Ehrenberg) Håkansson & Locker, a taxon that had not been verified in North America (Smucker et al. 2008). If a user enters ‘bramaputra’ in the search field, the species page for *T. lacustris* is returned. Thus, the *Reported As* field allows a user to search on a name that may be commonly in use, but which may not be correct, and still arrive at the species page showing the current, correct name.

The *Description* section includes a size series of high-resolution light micrographs. A minimum of five images of the taxon illustrates the size range and morphological variability from collections in North America, each with a reference by accession number to public herbarium. Images must be of good resolution, obtained on a high-quality microscope [minimum of 1.3 numerical aperture (NA) oil immersion objective and 1.3 NA condenser]. Contributors include metadata, including cell dimensions, herbarium slide number, collection location, and voucher slide number. A distinct symbol identifies images of a nomenclatural type (holotype, isotype, lectotype, isolectotype, and neotype). Scanning electron microscope images are included when possible, including the ongoing addition of images by other practitioners, pending approval of the original taxon page author.

Historically, as in foundational studies in limnology, European (primarily German) authors dominated the diatom literature and, arguably, continue to do so to the present day. In the absence of North American floras, European names were applied to North American species, often in error. In response to the previous blind application of European names, we strictly adhere to a rule that images and observations included in *diatoms.org* are based on North American specimens. Contributors provide written descriptions of the morphology and provide accurate dimensions of the taxa they observe and other verified North American records. An advantage of this flora is that the taxon dimensions may be updated when a practitioner documents a wider morphological range.

Species descriptions can expand or become modified to reflect new insights. Species descriptions include diagnostic features of the taxon, such as features of the striae, raphe, and valve shape. Contributors include notes on the specimens illustrated and how they compare to the range of variation of the taxon in North America. Contributors also include important nomenclatural history. For example, species pages for *Aulacoseira ambigua* (Grunow) Simonsen and *Aulacoseira granulata* (Ehrenberg) Simonsen includes examples of such nomenclatural notes (Potapova & English 2010a, b).

A diagnosis, or the features in which each taxon differs from closely related or similar taxa, is included in numbered elements of the *Guide* feature modelled on visual keys in popular bird identification guides. The diagnosis is further expanded in the *Compare* portion of each page. The intent is that these sections provide a concise diagnostic treatment of the key characters of each taxon, and those taxa with which it can be confused. The diagnosis guides practitioners on distinguishing closely related taxa, such as ‘Species 1 has an oval central area, while Species 2 has a rectangular central area’. Each diatom *Guide* includes a composite image with numbered points that direct the user to the named features. This feature, one of the most important parts of the page, is where the contributor informs the community on how to distinguish a species from similar, or often confused, species. The *Compare* section, another unique feature of *diatoms.org*, allows users to view similar species at the same magnification to facilitate visual comparison.

The *Autecology* section includes data on the basic information of the habits of each species. Although it was not emphasized at the outset of the project, ecological data are one of the most requested sections by users of the site. Since 2016, all new pages have included text or images in this section, and we continue to work to include ecological content in the balance of species pages. Additionally, the *Autecology* section includes tags that designate particular features of species such as size, motility, habitat types, attachment, and distribution. These tags were primarily conceived as a way to let analysts browse other taxa that may occur in the same region in lieu of a more complicated distribution map. Powerful means to combine or concatenate attributes may be used for browsing or searching. For example, a user may wish to construct a query such as ‘I wish to see taxa that are non-motile, unattached, and occur in California’. Other websites that are ‘metadata-heavy’, such as stock photography sites, were the inspiration for this faceted search. Efforts are being made to enhance these functionalities.

The *Original Description* section includes text in the language of original publication and iconotype, if available. These original descriptions and illustrations are often burdensome to locate in journal articles and books, or are even inaccessible outside of academic libraries and

databases. Illustrations of specimens of the nomenclatural type are especially valuable data and are marked and called out in image metadata. In some cases, the original description may be ambiguous (particularly for early authors, such as Ehrenberg and Kützing) and further work would be necessary to establish the identity of the type. In other cases, the original description may be clear but may not be a convincing match to the North American specimens. Thus, the *Description* text may note potential discrepancies for future revision.

Original species descriptions data are included in *diatoms.org* to aid in making species diversity information globally available to people for the purposes of teaching, scholarship, and research. Indeed, the ICN (Turland *et al.* 2018) states that effective publication of species includes widespread distribution. The *Citations & Links* section provides references to original descriptions and other relevant publications related to the taxon. When available, links to resources such as Index Nominum Algarum (Silva 2009), Diatombase (Kociolek *et al.* 2019), GenBank (Benson *et al.* 2013), and the Diatom New Taxon File at the Academy of Natural Sciences (Potapova *et al.* 2019b) are included.

Nomenclature, at least as relating to genus and species names, requires continuous updates and revisions to refine content, reflect nomenclatural changes, and correct errors (Vaidya *et al.* 2018). The *Update* section is reserved for changes to a page that would alter the concept of that taxon (e.g., addition or removal of particular images, change in size dimensions) or aspect of the name (e.g., transfer from one genus to another). Each update includes text describing the reason for the change, the date of the change, and the person who made the change. Previous versions of the page are archived and available within the *Practitioners* section of the site.

Discussion

Diatoms.org is built by practitioners, for practitioners. The project fills a gap in the resources that we, as diatom taxonomists, ecologists, and water quality managers, need in order to do our work. For the advanced practitioner, each taxon page represents a synthesis of information that is otherwise time-consuming, or resource-prohibitive, for an individual to obtain. *Diatoms.org* represents a growing scientific body of knowledge strengthened by revision, correction, and updates as science advances. Yet, the reference has its greatest reach for the general public, providing answers to basic questions and inviting people to learn more about microbial life on Earth.

Based on metrics for the *diatoms.org* audience, the project is already successful, not only in North America, but internationally. Data (via Google Analytics) show that over the two most recent years (June 2018–June 2020), nearly two million pages were accessed by 226,000 users

Table 1. Data tracked by Google Analytics for 1 June 2018–31 May 2020.

Total number of page views	1,886,000
Number of users (unique computers)	225,600
Number of sessions	424,780
Pages per session	4.4
Average session duration (minutes)	6.1

Table 2. Data tracked by Google Analytics for 1 June 2018–31 May 2020.

Country	Users	% Users
USA	99,112	44.0
India	16,762	7.4
Canada	12,809	5.7
UK	9138	4.1
China	5978	2.7
Japan	5583	2.5
Australia	5314	2.4
Indonesia	4667	2.1
Philippines	4265	2.0
Mexico	3722	1.7

(i.e., unique computers) (Table 1). The site visitation metrics are comparable, or greater, than the visitation of fifteen major botanical (primarily higher plant) web resources (Jones *et al.* 2014). Not only is site-traffic high, but a core set of users access information with high intensity. The average time spent on most websites is measured in a few seconds, while the average time spent on *diatoms.org* page is over six minutes, demonstrating that people engage with the content, i.e., they study it. The site is accessed most within the United States (44%), as compared to other countries (Table 2) with California having the greatest number of users within the United States (Table 3). The most commonly accessed page is the *diatoms.org/general/page*, followed by *diatoms.org/what-are-diatoms* (Table 4), indicating broad reach to a public audience.

Much of the site traffic of *diatoms.org* is by secondary school and college students, based on feedback from teachers. As students gain experience, they have the opportunity to contribute their knowledge, furthering the growth of *diatoms.org*. Students in the *Ecology and Systematics of Diatoms* course at Iowa Lakeside Laboratory have contributed over 100 pages since 2011. As a result, undergraduate and graduate students gain experience with peer-review and gain a citation for their resumé. The project serves as a pedagogical tool that integrates student involvement; students learn the practice of science and the dissemination of knowledge. Student created pages are an important contribution to the project. For example, the student page for *Gyrosigma acuminatum* (Kützing) Rabenhorst (Chabut 2014) was accessed over 2000 times in the past two years.

Table 3. Data tracked by Google Analytics for 1 June 2018–31 May 2020.

State	Users	Pages/session
California	12,446	2.9
Illinois	12,297	1.6
Texas	6224	4.3
Florida	6008	4.3
New York	5135	3.1
Washington	4006	4.5
Virginia	3909	3.0
Pennsylvania	3177	9.9
Ohio	3156	5.3
Massachusetts	3125	2.5
Georgia	2692	5.0
North Carolina	2562	2.5
Michigan	2372	6.6
Colorado	2343	6.0
New Jersey	2199	4.4
Minnesota	1864	5.5
Oregon	1828	3.9
Wisconsin	1519	4.7
Arizona	1316	4.9
Maryland	1312	2.6
Tennessee	1306	3.1
Louisiana	1244	5.8
Indiana	1151	6.8
Connecticut	1107	3.1
Iowa	1107	9.2
Missouri	1013	2.3
Alabama	914	2.4
South Carolina	911	3.0
Utah	850	11.2

Finally, the curation of diatom text, images, and associated data has led to impacts of *diatoms.org* beyond its use as a taxonomic flora. First, the project has led to increased use of the primary literature by exposing a large audience to original species descriptions. We, the Editorial Review Board, have noted increases in requests for original publications following the posting of species pages on *diatoms.org*. Second, the project provides detailed distribution records of species. For example, the project facilitated modelling of the hydrologic, climatic, and landscape processes that drive the distribution of *Didymosphenia geminata* (Lyngbye) M. Schmidt (Kumar et al. 2009, Spaulding 2010b). Third, the project has fostered taxonomic publication. For example, posting of *Planothidium biprorum* (M.H. Hohn & Hellerman) Lange-Bertalot led to the description of a new species, *P. incuriatum* Wetzel et al. (Wetzel et al. 2013). The taxon *Gomphonema caperatum* Ponader & Potapova was known under a provisional laboratory name for many years, but once it was described (Ponader et al. 2017) and published on *diatoms.org*, it was recognized in new regions by additional analysts. In another instance, the species page for *Planothidium rostratum* (Østrup) Lange-Bertalot was found to include a species that had not been recognized in North America (Wetzel et al. 2019). The new taxon, *P. potapovae*

Table 4. Data tracked by Google Analytics for 1 June 2018–31 May 2020.

Rank	Page	Page views
1	/what-are-diatoms	150,360
2	/genera	126,724
3	/home	69,670
4	/morphology	63,766
5	/species	42,718
6	/morphology/centric	34,000
7	/morphology/araphid	31,959
8	/genera/navicula	25,189
9	/morphology/monoraphid	24,598
10	/morphology/symmetric_biraphid	23,264
11	/genera/nitzschia	20,379
12	/morphology/symmetrical_biraphid	20,275
13	/genera/gomphonema	16,169
14	/morphology/nitzschoid	13,086
15	/genera/cymbella	12,818
16	/genera/fragilaria	12,768
17	/genera/achnanthidium	12,752
18	/morphology/asymmetric_biraphid	12,416
19	/genera/aulacoseira	12,368
20	/morphology/asymmetrical_biraphid	11,007
21	/practitioners	10,873
22	/genera/eunotia	9834
23	/genera/pinnularia	9829
24	/genera/cyclotella	9532
25	/glossary	9106
26	/morphology/eunotioid	8761
27	/genera/synedra	8547
28	/genera/surirella	8162
29	/genera/planothidium	8141
30	/genera/encyonema	8015
31	/morphology/epithemoid	7924
32	/genera/sellaphora	7801
33	/genera/cocconeis	7550
34	/morphology/surirelloid	7385
35	/genera/amphora	7132
36	/genera/achnanthes	7112
37	/genera/stephanodiscus	6671
38	/genera/grid	6173
39	/genera/stauroneis	6150
40	/genera/diatoma	6069
41	/genera/epithemia	5991
42	/genera/ulnaria	5986
43	/news/where-do-diatoms-live	5955
44	/about	5942
45	/genera/diploneis	5694
46	/news/what-is-diatomaceous-earth	5580
47	/genera/pseudostaurosira	5497
48	/news/how-big-are-diatoms	5310
49	/genera/staurosira	5208
50	/genera/staurosirella	4875

C.E. Wetzel & Ector, was described from the original *P. rostratum* species page. In the process of preparing content for species pages, issues concerning identification, nomenclature and systematic position of species prompted further publications. For example, students participated as co-authors in publications on nomenclatural changes and

species revisions (Spaulding *et al.* 2010, Rarick *et al.* 2017, Edlund *et al.* 2017, Williams *et al.* 2021).

Finally, news and project pages serve as public outreach and broader impacts for the National Science Foundation, National Park Service, U.S. EPA, and USGS funded efforts, with links to the taxa in each research programme. The taxa in several of these projects on paleolimnology and freshwater assessment (Spaulding *et al.* 2015, Bishop *et al.* 2017a, b, Tyree *et al.* 2020a, b) are supported by documentation in species pages. *Diatoms.org* conveys the relevance of diatoms through citizen science, actively engaging the public in contributing to the project (e.g., Edlund 2017).

Challenges and future directions

While *diatoms.org* now includes over 1000 species pages, the website has not reached a point of development to be able to serve as a primary resource for analysts. Many taxa remain to be documented, including some common taxa that are frequently encountered. Many of the taxa included to date are those with relatively few problematic taxonomic issues. Over the past several years, hundreds of taxon pages for common taxa have been initiated by contributors, but work was suspended when complex taxonomic issues were uncovered. Thus, we know that these additional pages require a greater amount of both time and expertise to complete. Perhaps our biggest challenge is to better support analysts by expanding the resource. A current goal is to expand the *diatoms.org* platform to better support analysts and researchers and train the next generation of taxonomists. Over the past five years funding has supported contributors' efforts to research taxa. Efforts made by 3–5 graduate students or postdoctoral researchers over the next five years would bring *diatoms.org* to the point of including 80% of the taxa encountered in North American freshwaters.

Moreover, in the light of the climatic and environmental upheaval of earth's species diversity and ecological systems, there is an urgent need for *diatoms.org* to expand its reach to a diverse generation of young people, few of whom have access to microscopes. We regularly receive inquiries and requests to expand *diatoms.org* into a global resource by linking to other freshwater diatom floras and databases, as well as including marine and fossil taxa. We need to expand the roles for everyone, including students and citizen scientists, across geographic provinces (Ellwood *et al.* 2015, 2018). Expansion of roles, along with the expansion of content, requires greater coordination of resources. While infrastructure is in place that supports the connection of databases across global hubs (Meyer *et al.* 2015), a clear plan and funding are required to implement those connections. Finally, despite the immense impact of web content, there is little recognition by academia of the value of web contributions; the legitimacy of online curation remains to be established and matched

in accordance with its influence (Rotman *et al.* 2012, Parr *et al.* 2014). As the scientific value of online contributions becomes more recognized, contributors and editors should gain institutional recognition for activities such as site maintenance and editorial oversight of lineages, or groups, of taxa.

The further design could involve developing the ability for practitioners to create customized floras. For example, a user might require smaller floras of taxa from lakes in the northeastern US, fens in the midwestern US, or *Psammothidium* Bukhtiyarova & Round species across North America. Additionally, now that the number of species pages has surpassed a critical threshold of being required in task agreements (e.g., U.S. EPA National Lakes Survey), we have the opportunity to investigate diatom biology through meta-analyses of the site itself. For example, we can query not only geographic distributions but also size distributions of species. For example, the data now allows comparing the morphological range of species within a genus, or species that live in the most pristine streams, or downloading large sets of images for use in artificial intelligence. This opportunity to query *diatoms.org* allows us to begin using big morphological data. Are there underlying evolutionary, physiological, or ecological factors that drive these different allometric patterns? Using data contained within the website, we can explore the biogeography of species on a continental scale, improving the understanding of rates of evolutionary change, stochastic processes associated with dispersal limitation, and spread of invasive species.

Technological advances themselves are likely to offer opportunities for scientific progress. For example, in combination with *diatoms.org*, data from national lake surveys, such as U.S. EPA National Lake Survey and National Ecological Observatory Network (neonscience.org), allowed analysis of species distributions and environmental modelling (Kumar *et al.* 2009). Additionally, image recognition software may facilitate automated diatom identification (Spaulding *et al.* 2012, Bishop *et al.* 2017b, Cristóbal *et al.* 2020, Kloster *et al.* 2020), allowing taxonomists to focus on uncovering biodiversity. Through development of associated voucher floras (Tyree *et al.* 2020a, b) and improved coordination of analysts, *diatoms.org* can streamline workflows and encourage productive outcomes. As a result, there is potential for understanding diatom biogeography with verified data that has not been possible before.

Diatoms.org provides a hub for the toolset that diatomists need to work efficiently. For many groups of organisms, monitoring is compromised by taxonomically biased efforts (Navarro *et al.* 2017), a problem that has been solved for diatoms (Bishop *et al.* 2017a, Tyree *et al.* 2020a, 2020b). *Diatoms.org* adapted rapidly to communicating developments in analysis, taxonomy, harmonization, and taxonomic certification by hosting the widely viewed online seminars of the *Diatom Web*

Table 5. Consideration of issues across (1) books and other non-digital literature, (2) *diatoms.org*, and (3) the future of *diatoms.org*.

Issue	Books	<i>Diatoms.org</i>	Future of <i>diatoms.org</i>
Audience	Narrow range of specialists and experienced professionals	Wide range – lay public to experienced professional	Broader public education; enhanced community involvement through online teaching and discussion
Cost to Audience	Very high (US \$100-1200) per book	No cost; important to keep no cost to promote fair science	No cost; important to keep no cost to promote fair science
Cost to Public	No cost	In-kind content by contributors, federal and academic support	Support from user communities through grants; adoption by a professional society; or establishment of a non-profit foundation
Access	Limited to those with academic or federal affiliations through libraries	Open access through the internet	Expand workshops, courses, and discussions to all experience levels of audience
Contributors	Specialists; experienced professionals	Specialists; experienced taxonomists	Increase training for students; continue and expand participation through Diatom Web Academy workshops
Reviewers	Not reviewed; limited editorial review	Scientific and editorial standard met through peer review; reviews are not anonymous to further responsibility for content	Expansion of review process to include early-career scientists to review content as part of their training
Advancement of Science	High at time of publication, potential decrease in impact over time	Value increases with each contribution; updates are timely so that practitioners can stay informed	<i>diatoms.org</i> itself provides data that advances science; data are served through queries; functions to drive scientific inquires; opens new avenues of inquiry; the project provides a means to continue taxonomic training and expertise into the future
Content Flexibility	Fixed to time of publication	Can be updated or corrected with new information	Incorporation of displays in response to practitioner need
Display Flexibility	Fixed to time of publication	Flexible in terms of updates, search, organized by morphology (text or shape), filter by size or striae	With advances in uncovering diatom relationships, phylogeny can become the organizational structure
Number of Taxa	Fixed to time of publication	Continually added but currently limited by funding	Expand with support from user communities through grants; expand through the Society for Freshwater Research Taxonomic Certification Program
Taxon Comparison	"flipping" between pages of interest	Selection fixed by editors	Develop further flexible format for practitioners to view taxa of interest
Aids to Identification	Descriptive text, comparison tables	Descriptive text, iconic images, guide images, comparison feature	Users add their own images; access to high-resolution scans of entire microscope slides
Taxonomic Coverage	Limited to geographic region or taxonomic group, or both	Limited to North America; freshwater	Practitioners and funders include 2000–3000 North American taxa; include worldwide freshwater, distribution; expand to marine genera and species

Academy (<https://diatoms.org/news/diatom-web-academy-1>), an especially valuable role in fostering community and communication during the global pandemic. In addition, the project provides an archive of diatom data and information in a way that the larger scientific community can view, verify, and utilize content to infer patterns and processes, using statistical and meta-analyses. The activities supported and promoted by *diatoms.org* provide a means for documentation of a chain of inference forming a taxonomic standard of practice. Thus, *diatoms.org* provides a structured approach for working with diatoms as data elements. Next steps should include linkage of images and voucher floras (Bishop *et al.* 2017a, Tyree *et al.* 2020a) to standardized formats, such as those of the Humboldt Core for biological inventories (Guralnick *et al.* 2018).

The project began as a version of an extended electronic book in the form of *Diatoms of the United States*. Over time, the interaction of contributors, content, users, courses, students, workshops, reviewers, artists, and online seminars blossomed into something else (Table 5). By way of these interactions, facilitated by strong a visual format, the project fosters human relationships and allows new means for the international diatom community to learn from one another. We think that this ‘something else’ can be expressed as a type of relational ecology, and that there is a strong link among human relationships, the content on the site, and the engagement of a community. Indeed, it is the emergence of human connections that adds value to the endeavour of *diatoms.org*. This relational ecology encompasses a broad swath of research objectives and social behaviour. The establishment of an emergent intelligence opens doors to continued conversations and to interactions that ‘poke’ at science to further its advancement. The project needs to remain nimble enough to take advantage of new technologies to address issues, including the human interplay between the parts and the whole of the project. It may be that the original core of the project, the species and genus pages, are actually the *least* of *diatoms.org*. Across the various pages, there is now enough data to provide science outcomes from querying the data. *Diatoms.org* is now, in itself, a source of data.

Acknowledgements

David Lubinski developed the original concepts, structure, and technical design. Over a decade’s worth of students in the *Ecology and Systematics of Diatoms* course at Iowa Lakeside Laboratory served to stimulate the teaching aspects of the project. The authors appreciate the ongoing support of Jane Shuttleworth and Friends of Lakeside Lab. Past members of Editorial Review Board (Sam Rushforth, Rex Lowe, and Pat Kociolek) were instrumental in advancing the foundation of the web flora. Frequent, dedicated, undercover contributors (Loren Bahls, Rob Kimmich, Rosalina Stancheva, and Ionel Ciugulea) work quietly and productively. The authors thank anonymous reviewers who provided both feedback and new ideas that greatly improved the manuscript. Students at the University of Colorado (Alec Camp, Tanya Hannis, and Janey Le) cleaned up endless files. The project

could not have been completed without Diane McKnight, Meredith Tyree, Nicholas Schulte, David Burge, Teofil Nakov, Mitchell Kimbrough, Nicolas Bottari, and Michiko Swiggs. Eugene F. Storermer provided the inspiration to take on a seemingly impossible mission.

The views expressed in this article are those of the authors and do not necessarily represent the views or the policies of the U.S. Environmental Protection Agency. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

Funding from the U.S. EPA Office of Water and Office of Science and Technology supported the direct development of content by contributors, as well as through formal taxonomic workshops. The U.S. Geological Survey under Cooperative Agreement #G15AC00104 with the Institute of Arctic and Alpine Research (INSTAAR) at the University of Colorado supported web design and development.

ORCID

Sarah A. Spaulding  <http://orcid.org/0000-0002-9787-7743>

Marina G. Potapova  <http://orcid.org/0000-0001-5143-5213>

Ian W. Bishop  <http://orcid.org/0000-0003-2197-3197>

Sylvia S. Lee  <http://orcid.org/0000-0003-3855-7085>

Paula C. Furey  <http://orcid.org/0000-0002-2352-4427>

Mark B. Edlund  <http://orcid.org/0000-0002-0227-1117>

References

- AGNARSSON I., KUNTNER M. & PATERSON A. 2007. Taxonomy in a changing world: seeking solutions for a science in crisis. *Systematic Biology* 56: 531–539.
- ALERS-GARCIA J., LEE S.S. & SPAULDING S.A. 2021. Resources and practices to improve diatom data quality. *Limnology & Oceanography Bulletin* 30: 48–53. doi:10.1002/lob.10433
- ANTONIADES D., DOUGLAS M.S.V. & SMOL J.P. 2009a. Biogeographic distributions and environmental controls of stream diatoms in the Canadian Arctic Archipelago. *Botany* 87: 443–454.
- ANTONIADES D., HAMILTON P.B., HINZ F., DOUGLAS M.S.V. & SMOL J.P. 2009b. Seven new species of freshwater diatoms (Bacillariophyceae) from the Canadian Arctic Archipelago. *Nova Hedwigia* 88: 57–80.
- BAHLS L. 2005. Northwest Diatoms: a photographic catalogue of species in the Montana Diatom Collection, with ecological optima, associates, and distribution records for the nine Northwestern United States. *Northwest Diatoms, Volume 2*. The Montana Diatom Collection, Helena. 464 pp.
- BAHLS L. 2010. *Stauroneis* in the Northern Rockies: 50 species of *Stauroneis sensu stricto* from western Montana, northern Idaho, northeastern Washington and southwestern Alberta, including 16 species described as new. *Northwest Diatoms, Volume 4*. The Montana Diatom Collection, Helena. 172 pp.
- BAHLS L. 2013. *Encyonopsis* from western North America: 31 species from Alberta, Idaho, Montana, Oregon, South

- Dakota, and Washington, including 17 species described as new. *Northwest Diatoms, Volume 5*. The Montana Diatom Collection, Helena. 46 pp.
- BENSON D.A., CAVANAUGH M., CLARK K., KARSCHMIZRACHI I., LIPMAN D.J., OSTELL J. & SAYERS E.W. 2013. GenBank. *Nucleic Acids Research* 41 (Database issue): D36–42. doi: [10.1093/nar/gks1195](https://doi.org/10.1093/nar/gks1195)
- BISHOP I.W., ESPOSITO R.M., TYREE M. & SPAULDING S.A. 2017a. A diatom voucher flora from selected southeast rivers (USA). *Phytotaxa* 332: 101–140.
- BISHOP I.W., SPAULDING S.A. & KROTH P. 2017b. Life cycle size dynamics in *Didymosphenia geminata* (Bacillariophyceae). *Journal of Phycology* 53: 652–663.
- BORGMAN C.L. 2007. *Scholarship in the digital age: information, infrastructure, and the internet*. Cambridge, MA: The MIT Press.
- CHAPUT M. 2014. *Gyrosigma acuminatum* [online] Diatoms of North America. Available from: https://diatoms.org/species/gyrosigma_acuminatum [Accessed 27 June 2021].
- CRISTÓBAL G., BLANCO S. & BUENO G. (Ed.). 2020. *Modern trends in diatom identification: fundamentals and applications. Developments in Applied Phycology 10*. Springer International Publishing. doi: [10.1007/978-3-030-39212-3](https://doi.org/10.1007/978-3-030-39212-3)
- CUI H., BOUFFORD D. & SELDEN P. 2010. Semantic annotation of biosystematics literature without training examples. *Journal of the American Society for Information Science and Technology* 61: 522–542.
- EBACH M.C., VALDECASAS A.G. & WHEELER Q.D. 2011. Impediments to taxonomy and users of taxonomy: accessibility and impact evaluation. *Cladistics* 27: 550–557.
- EDLUND M.B. 2017. Citizen Scientists Lend a Hand [online] Diatoms of North America. Available from: https://diatoms.org/news/citizen_scientists_lend_a_hand [Accessed 23 April 2019].
- EDLUND M.B., BURGE D.R.L. & SPAULDING S.A. 2017. The transfer of *Navicula cuspidata* var. *obtusata* to *Craticula* (Bacillariophyceae). *Notulae Algarum* 19: 1–2.
- EDLUND M.B., JÜTTNER I. & SPAULDING S. 2019. Are websites science? And other topics, an international discussion [online]. Diatoms of North America. Available from: <https://diatoms.org/news/an-international-discussion> [Accessed 23 April 2019].
- EHRENBERG C.G. 1854. *Mikrogeologie. das Erden und Felsen schaffende Wirken des unsichtbar kleinen selbständigen Lebens auf der Erde*. Leopold Voss, Leipzig.
- ELLWOOD E.R., DUNCKEL B.A., FLEMONS P., GURALNICK R., NELSON G., NEWMAN G., PAUL D., RICCARDI G., RIOS N., SELTMANN K.C. & MAST A.R. 2015. Accelerating the digitization of biodiversity research specimens through online public participation. *BioScience* 65: 383–396. doi: [10.1093/biosci/biv005](https://doi.org/10.1093/biosci/biv005)
- ELLWOOD E.R., KIMBERLY P., GURALNICK R., FLEMONS P., LOVE K., ELLIS S., ALLEN J.M., BEST J.H., CARTER R., CHAGNOUX S., COSTELLO R., DENSLOW M.W., DUNCKEL B.A., FERRITER M.M., GILBERT E.E., GOFORTH C., GROOM Q., KRIMMEL E.R., LAFRANCE R., MARTINEC R.L., MILLER A.N., MINNAERT-GROTE J., NASH T., OBOYSKI P., PAUL D.L., PEARSON K.D., PENTCHEFF N.D., ROBERTS M.A., SELTZER C.E., SOLTIS P.S., STEPHENS R., SWEENEY P.W., VON KONRAT M., WALL A., WETZER R., ZIMMERMAN C. & MAST A.R. 2018. Worldwide engagement for digitizing biocollections (WeDigBio): The biocollections community’s citizen-science space on the calendar. *BioScience* 68:112–124. doi: [10.1093/biosci/bix143](https://doi.org/10.1093/biosci/bix143)
- ENDARA L., CUI H. & BURLEIGH J.G. 2018. Extraction of phenotypic traits from taxonomic descriptions for the tree of life using natural language processing. *Applications in Plant Sciences* 6(3): e1035
- FUREY P.C., LOWE R.L. & JOHANSEN J.R. 2011. *Eunotia Ehrenberg* (Bacillariophyta) of the Great Smoky Mountains National Park, USA. *Bibliotheca Diatomologica* 56: 1–134.
- GURALNICK R., WALLS R. & JETZ W. 2018. Humboldt Core – toward a standardized capture of biological inventories for biodiversity monitoring, modeling and assessment. *Ecography* 41: 713–725. doi: [10.1111/ecog.02942](https://doi.org/10.1111/ecog.02942)
- HAMILTON P. & MANOYLOV K. 2012. *Navicula hodgsoni* [online] Diatoms of North America. Available from: https://diatoms.org/species/navicula_hodgsoni [Accessed 14 September 2021].
- HEISE C. & PEARCE J.M. 2020. From open access to open science: the path from scientific reality to open scientific communication. *SAGE Open* 2020: 1–14. doi: [10.1177/2158244020915900](https://doi.org/10.1177/2158244020915900)
- JETZ W., MCPHERSON J.M. & GURALNICK R.P. 2012. Integrating biodiversity distribution knowledge: toward a global map of life. *Trends in Ecology & Evolution* 27: 151–159. doi: [10.1016/j.tree.2011.09.007](https://doi.org/10.1016/j.tree.2011.09.007)
- JONES T., BAXTER D., HAGEDORN G., LEGLER B., GILBERT E., THIELE K., VARGAS-RODRIGUEZ Y. & URBATSCH L. 2014. Trends in access of plant biodiversity data revealed by Google Analytics. *Biodiversity Data Journal* 2: e1558. doi: [10.3897/BDJ.2.e1558](https://doi.org/10.3897/BDJ.2.e1558)
- KLOSTER, M., LANGENKÄMPER, D., ZUROWIETZ, M., BESZTERI, B. & NATTKEMPER, T.W. 2020. Deep learning-based diatom taxonomy on virtual slides. *Scientific Reports* 10: 14416. doi: [10.1038/s41598-020-71165-w](https://doi.org/10.1038/s41598-020-71165-w)
- KOCIOLEK J.P. & SPAULDING S.A. 2000. Freshwater diatom biogeography. *Nova Hedwigia* 71: 223–241.
- KOCIOLEK J.P. 2006. Some thoughts on the development of a diatom flora for freshwater ecosystems in the continental United States and a listing of recent taxa described from U.S. freshwaters. *Proceedings of the California Academy of Sciences* 57: 561–586.
- KOCIOLEK J.P. & THOMAS E.W. 2010. Taxonomy and ultrastructure of five naviculoid diatoms (Class Bacillariophyceae) from the Rocky Mountains of Colorado (USA), with the description of a new genus and four new species. *Nova Hedwigia* 90: 195–214.
- KOCIOLEK P. 2011. *Thalassiosira lacustris* [online] Diatoms of North America. Available from https://diatoms.org/species/thalassiosira_lacustris [Accessed 20 February 2019].
- KOCIOLEK J.P., BALASUBRAMANIAN K., BLANCO S., COSTE M., ECTOR L., LIU Y., KULIKOVSKIY M., LUNDHOLM N., LUDWIG T., POTAPOVA M., RIMET F., SABBE K., SALA S., SAR E., TAYLOR J., VAN DE VIJVE, B., WETZE, C.E., WILLIAM, D.M., WITKOWSKI A. & WITKOWSKI J. 2019. DiatomBase [online]. Available from: <http://www.diatombase.org> [Accessed 23 April 2019].

- KRAMMER, K. 1997. Die cymbelloiden Diatomeen. Eine Monographie der weltweit bekannten Taxa. Teil 1. Allgemeines und. *Encyonema Bibliotheca Diatomologica* 36: 1–382.
- KUMAR S., SPAULDING S.A., STOHLGREN T.J., HERMANN K.A., SCHMIDT T.S. & BAHLS L.L. 2009. Potential habitat distribution for the freshwater diatom *Didymosphenia geminata* in the continental US. *Frontiers in Ecology and the Environment* 7: 415–420.
- LEE S.S. 2019. Forging effective use of diatoms [online] Diatoms of North America. Available from: <https://diatoms.org/news/forging-effective-use-of-diatoms> [Accessed 23 April 2019].
- LEE S.S., BISHOP I.W., SPAULDING S.A., MITCHELL R. & YUAN L.L. 2019. Taxonomic harmonization may reveal a stronger association between diatom assemblages and total phosphorus in large datasets. *Ecological Indicators* 102: 166–174.
- LENDEMER J., THIERS B., MONFILS A.K., ZASPEL J., ELLWOOD E.R. ET AL. 2020. The extended specimen network: A strategy to enhance US biodiversity collections, promote research and education. *BioScience* 70: 23–30.
- LÜCKING R. 2008. Taxonomy: a discipline on the brink of extinction—are DNA barcode scanners the future of biodiversity research? *Archives des Sciences* 61: 75–88.
- MEYER C., KREFT H., GURALNICK R. & JETZ W. 2015. Global priorities for an effective information basis of biodiversity distributions. *Nature Communications* 6: 8221. doi: [10.1038/ncomms9221](https://doi.org/10.1038/ncomms9221)
- MORALES E.A. 2005. Observations of the morphology of some known and new fragilarioid diatoms (Bacillariophyceae) from rivers in the USA. *Phycological Research* 53: 113–133.
- NAKOV T., RUCK E.C., GALACHYANTS Y., SPAULDING S.A. & THERIOT E.C. 2014. Molecular phylogeny of the Cymbellales (Bacillariophyceae, Heterokontophyta) with a comparison of models for accommodating rate variation across sites. *Phycologia*, 53: 359–373.
- NAKOV T., GUILLORY W.X., JULIUS M.L., THERIOT E.C. & ALVERSON A.J. 2015. Towards a phylogenetic classification of species belonging to the diatom genus *Cyclotella* (Bacillariophyceae): transfer of species formerly placed in *Puncticulata*, *Handmannia*, *Pliocaenicus* and *Cyclotella* to the genus *Lindavia*. *Phytotaxa* 217: 249–264.
- NAVARRO L.M., FERNÁNDEZ N., GUERRA C., GURALNICK R., KISSLING D., LONDOÑO M.C., MULLER-KARGER F., TURAK E., BALVANERA P., COSTELLO M.J., DELAUAUD A., EL SERAFY G.Y., FERRIER S., GEIJZENDORFFER I., GELLER G.N., JETZ W., KIM E-S., KIM H. ET AL. 2017. Monitoring biodiversity change through effective global coordination. *Current Opinion in Environmental Sustainability* 29: 158–169. doi: [10.1016/j.cosust.2018.02.005](https://doi.org/10.1016/j.cosust.2018.02.005)
- OMB 2009. *Open government directive. OMB Memorandum M-10-06. December 8, 2009.* Office of Management and Budget. Executive Office of the President of the United States. Washington, DC. pp. 11. [online]. Available from: <https://www.whitehouse.gov/sites/whitehouse.gov/files/omb/memoranda/2010/m10-06.pdf> [Accessed 16 September 2021].
- OMB 2013. *Open data policy-managing information as an asset. OMB Memorandum M-13-13. May 9, 2013.* Office of Management and Budget. Executive Office of the President of the United States. Washington, DC. pp. 12. [online]. Available from: <https://www.whitehouse.gov/sites/whitehouse.gov/files/omb/memoranda/2013/m-13-13.pdf> [Accessed 16 September 2021].
- PARR C.S., WILSON N., LEARY P., SCHULZ K.S., LANS K., WALLEY L., HAMMOCK J.A., GODDARD A., RICE J., STUDER M., HOLMES J.T. & CORRIGAN R.J. 2014. The Encyclopedia of Life v2: providing global access to knowledge about life on earth. *Biodiversity Data Journal* 2, e1079. doi: [10.3897/BDJ.2.e1079](https://doi.org/10.3897/BDJ.2.e1079)
- PONADER K. & POTAPOVA M. 2007. Diatoms from the genus *Achnantheidium* in flowing waters of the Appalachian mountains (North America): ecology, distribution and taxonomic notes. *Limnologica – Ecology and Management of Inland Waters* 37: 227–241.
- PONADER K.C., POTAPOVA M.G., DESIANTI N., HAMILTON P.B., LAVOIE I. & CAMPEAU S. 2017. *Gomphonema caperatum* sp. nov., *G. obstipum* sp. nov. and similar taxa from rivers of North America. *Diatom Research* 32: 59–73.
- POTAPOVA M. & CHARLES D.F. 2003. Distribution of benthic diatoms in U.S. rivers in relation to conductivity and ionic composition. *Freshwater Biology* 48: 1311–1328.
- POTAPOVA M.G., CIUGULEA I. & MINEROVIC A. 2019a. The novel species *Navicula eileeniae* (Bacillariophyta, Naviculaceae) and its recent expansion in the Central Appalachian region of North America. *Plant Ecology and Evolution* 152: 368–377.
- POTAPOVA M. & ENGLISH J. 2010a. *Aulacoseira ambigua*. [online] Diatoms of North America. Available from: https://diatoms.org/species/aulacoseira_ambigua [Accessed 26 August 2021].
- POTAPOVA M. & ENGLISH J. 2010b. *Aulacoseira granulata*. [online] Diatoms of North America. Available from: https://diatoms.org/species/aulacoseira_granulata [Accessed 26 August 2021].
- POTAPOVA M.G., MINEROVIC A.D., VESELÁ J. & SMITH C.R. (Ed.) 2019b. Diatom New Taxon File at the Academy of Natural Sciences (DNTF-ANS), Philadelphia. [online] Available from: <http://dh.ansp.org/dntf> [Accessed 26 August 2021].
- POTAPOVA M.G. & PONADER K.C. 2004. Two common North American diatoms, *Achnantheidium rivulare* sp. nov. and *A. deflexum* (Reimer) Kingston: morphology, ecology and comparison with related species. *Diatom Research* 19: 33–57.
- POTAPOVA M. & WINTER D. 2006. Use of nonparametric multiplicative regression for modeling diatom habitat: a case study of three *Geissleria* species from North America. In: *Advances in Phycological Studies* (Ed. by N. OGNJANOVARUMENOV & K. MANYOLOV), pp. 319–332. Pensoft Publishers, Sofia-Moscow.
- RARICK J. 2017. *Luticola goeppertiana* [online] Diatoms of North America. Available from: https://diatoms.org/species/luticola_goeppertiana [Accessed 26 August 2021].

- RARICK J., WU S., LEE S.S. & EDLUND M.B. 2017. The valid transfer of *Stauroneis goeppertiana* to *Luticola* (Bacillariophyceae). *Notulae Algarum* 29: 1–3.
- RODMAN J.E. 2007. Reflections on PEET, the partnerships for enhancing expertise in taxonomy. *Zootaxa* 1668: 41–46.
- ROTMAN D., PROCITA K., HANSEN D., PARR C.S. & PREECE J. 2012. Supporting content curation communities: The case of the Encyclopedia of Life. *Journal of the American Society for Information Science and Technology* 63: 1092–1107. doi: [10.1002/asi.22633](https://doi.org/10.1002/asi.22633).
- RUCK E.C., NAKOV T., ALVERSON A.J. & THERIOT E.C. 2016a. Phylogeny, ecology, morphological evolution, and reclassification of the diatom orders Surirellales and Rhopalodiales. *Molecular Phylogenetics and Evolution* 103: 155–171.
- RUCK E.C., NAKOV T., ALVERSON A.J. & THERIOT E.C. 2016b. Nomenclatural transfers associated with the phylogenetic reclassification of the Surirellales and Rhopalodiales. *Notulae Algarum* 10: 1–4. (17 August 2016).
- SILVA P. 2009. Index Nominum Algarum [online] University Herbarium, University of California, Berkeley. Available from: <http://ucjeps.berkeley.edu/INA.html> [Accessed 25 August 2021].
- SIVER P.A., HAMILTON P.B., STACHURA-SUCHOPLES K. & KOCIOLEK J.P. 2005. Diatoms of North America. The Freshwater Flora of Cape Cod. *Iconographia Diatomologica* 14: 1–463.
- SIVER P.A. & HAMILTON P.B. 2011. Diatoms of North America: The Freshwater Flora of Waterbodies on the Atlantic Coastal Plain. In: *Iconographia Diatomologica Annotated Diatom Micrographs* (Ed. by H. LANGE-BERTALOT), Vol. 22. A.R.G. Gantner Verlag 916 pp. <https://www.tandfonline.com/doi/abs/10.1080/0269249X.2011.639573>
- SLUYS R. 2013. The unappreciated, fundamentally analytical nature of taxonomy and the implications for the inventory of biodiversity. *Biodiversity and Conservation* 22: 1095–1105.
- SMOL J.P. & STOERMER E.F. 2010. *The diatoms: applications for the environmental and earth sciences*, Second Edition. Cambridge University Press, Cambridge 667 pp.
- SMUCKER N.J., EDLUND M.B. & VIS M.L. 2008. The distribution, morphology and ecology of a non-native species, *Thalassiosira lacustris* (Bacillariophyceae), from benthic stream habitats in North America. *Nova Hedwigia* 87: 201–220.
- SPAULDING S.A., KOCIOLEK J.P. & DAVIS D.R. 2002. A new diatom (Bacillariophyceae) genus with two new species from New Mexico, USA. *European Journal of Phycology* 37: 135–143.
- SPAULDING S. 2010a. *Encyonema temperei*. In *Diatoms of North America*. Available from: https://diatoms.org/species/encyonema_temperei [Accessed 23 April 2019].
- SPAULDING S. 2010b. *Didymosphenia geminata*. In *Diatoms of North America*. Available from: https://diatoms.org/species/didymosphenia_geminata [Accessed 27 June 2021].
- SPAULDING S.A., POOL J., CASTRO S. & HINZ F. 2015. Species within the genus *Encyonema* Kützing, including two new species *Encyonema reimeri* sp. nov. and *E. nicafeii* sp. nov. and *E. stoermeri* nom. nov., stat. nov. *Proceedings of the Academy of Natural Sciences of Philadelphia* 160: 57–71.
- SPAULDING S.A., JEWSON D., BIXBY R., NELSON H. & MCKNIGHT D.M. 2012. Automated measurement of diatom size. *Limnology and Oceanography: Methods* 10: 882–890.
- STODDARD J., PECK D., OLSEN A., LARSEN D.P., VAN SICKLE J., HAWKINS C., HUGHES R.M., WHITTIER T.R., LOMNICKY G., HERLIHY A., KAUFMANN P.R., 2005. Environmental Monitoring and Assessment Program (EMAP): Western streams and rivers statistical summary. U.S. Environmental Protection Agency, Office of Research and Development, Western Ecology Division, EPA 620/R-05/006, Corvallis, Oregon.
- STONE J.R., EDLUND M.B., STREIB L., QUANG H.H. & MCGLUE M.M. 2020. *Stephanodiscus coruscus* sp. nov., a new species of diatom (Bacillariophyta) from June Lake, California (USA) with close affiliation to *Stephanodiscus klamathensis*. *Diatom Research* 35: 339–351.
- THERIOT E.C., ASHWORTH M., RUCK E., NAKOV T. & JANSEN R.K. 2010. A preliminary multigene phylogeny of the diatoms (Bacillariophyta): challenges for future research *Plant Ecology and Evolution* 143: 278–296.
- TURLAND, N. J., WIERSEMA, J. H., BARRIE, F. R., GREUTER, W., HAWKSWORTH, D. L., HERENDEEN, P. S., KNAPP, S., KUSBER, W.-H., LI, D.-Z., MARHOLD, K., MAY, T. W., MCNEILL, J., MONRO, A. M., PRADO, J., PRICE, M. J. & SMITH, G. F. (EDS.) 2018: *International Code of Nomenclature for algae, fungi, and plants (Shenzhen Code) adopted by the Nineteenth International Botanical Congress Shenzhen, China, July 2017*. Regnum Vegetabile 159. Glashütten: Koeltz Botanical Books. doi: [10.12705/Code.2018](https://doi.org/10.12705/Code.2018)
- TYREE M.A., BISHOP I.W., HAWKINS C.P. & SPAULDING S.A. 2020a. Reduction of taxonomic bias in diatom species data. *Limnology and Oceanography: Methods* 18: 271–279. doi:[10.1002/lom3.10350](https://doi.org/10.1002/lom3.10350)
- TYREE M.A., CARLISLE D.M. & SPAULDING S.A. 2020b. Diatom enumeration method influences biological assessments of southeastern USA streams. *Freshwater Science* 39: 183–195.
- VAIDYA G., LEPAGE D., GURALNICK R. & FRANZ N. 2018. The tempo and mode of the taxonomic correction process: how taxonomists have corrected and recorrected North American bird species over the last 127 years. *PLOS ONE* 13(4): e0195736. doi: [10.1371/journal.pone.0195736](https://doi.org/10.1371/journal.pone.0195736)
- WETZEL C.E., VAN DE VIJVER B., HOFFMANN L. & ECTOR L. 2013. *Planothidium incuriatum* sp. nov. a widely distributed diatom species (Bacillariophyta) and the type analysis of *Planothidium biporumum*. *Phytotaxa* 138: 43–57.
- WETZEL C.E., VAN DE VIJVER B., BLANCO S. & ECTOR L. 2019. On some common and new cavum-bearing *Planothidium* (Bacillariophyta) species from freshwater. *Fottea* 19: 50–89.

- WHEELER Q. 2010. What would NASA do? Mission-critical infrastructure for species exploration. *Systematics and Biodiversity* 8:11–15.
- WILLIAMS D.M., SPAULDING S.A. & BISHOP I. 2021. Studies on type material from Kützing's diatom collection IV: The basionym, author and type of *Tetracyclus rupestre*. *Phytotaxa* 498: 44–50.
- WILLIAMS D.M. & KOCIOLEK J.P. 2010. Towards a comprehensive diatom classification and phylogeny (Bacillariophyta). *Plant Ecology and Evolution* 143: 265–270.
- WILSON E.O. 2003. The encyclopedia of life. *Trends in Ecology & Evolution* 18: 77–80.