OPEN ACCESS

CrossMark

Workshop Report

Workshop Report: Supporting inclusive and sustainable collections-based research infrastructure for systematics (SISRIS)

Andrea Weeks‡, Elizabeth Collins‡, Twanelle W. Majors§, Zack E. Murrell|, Deborah L. Paul¶, Matthew Sheik#, David P. Shorthouse¤, Shawn Zeringue-Krosnick«

‡ George Mason University, Department of Biology, Fairfax, Virginia, United States of America
§ Tennessee Technological University, Department of Chemistry, Cookeville, Tennessee, United States of America
| Appalachian State University, Boone, North Carolina, United States of America
¶ University of Illinois Urbana-Champaign, Prairie Research Institute, Illinois Natural History Survey, Species Profile Group, Champaign, Illinois, United States of America
# Denver Botanic Gardens, Denver, Colorado, United States of America
¤ Independent, Gatineau, Quebec, Canada
« Tennessee Technological University, Department of Biology, Cookeville, Tennessee, United States of America

Abstract

We created and delivered a workshop and symposium series for biologists at all career stages focused on the skills and practices needed to sustain natural history specimen attribution and citation. The name of the workshop and symposium series, SISRIS, reflected our ultimate goal of effecting community-level change by sharing skills and practices that can support inclusive and sustainable (collections-based) research infrastructure for systematics. We report here the rationale for SISRIS, its learning objectives for participants and its results, including the assessment of outcomes from three iterations of the workshop held in 2023. The SISRIS workshops and symposia were held in person at the annual meeting of the Association for Southeastern Biologists in Winston-Salem, North Carolina and Botany 2023 in Boise, Idaho. A stand-alone SISRIS workshop was held online later to accommodate individuals who were unable to travel to the in-person events.

Keywords

Bionomia, collections-based research, Extended Specimen Concept, natural history collections, SISRIS, systematics

Date and place

Workshops were held in three venues. Two in-person workshops followed by symposia were held at the annual meeting of the Association of Southeastern Biologists (Winston-Salem, North Carolina; 25 March 2023) and Botany 2023 (Boise, Idaho; 25 July 2023). An online, workshop-only event was held 6 August 2023 via Zoom to accommodate individuals who could not attend in person.

List of participants

Disseminating the workshop participant list is restricted by the Tennessee Technological University Institutional Review Board (IRB #2609). In accordance with stipulations for IRB exempt research, the project "Collaborative Research: Conference: Supporting inclusive and sustainable research infrastructure for systematics (SISRIS) by connecting scientists and their specimens" must ensure that any specific information regarding the identity of participants or their institutions are not shared in the dissemination of its results.

Introduction

Background

Scientific discoveries derived from natural history collections depend, first and foremost, on the people who create the foundation of preserved specimens. Their taxonomic expertise and continued contributions are essential to the long-term health of collections-based research. Yet, data indicate that this historical foundation is threatened despite continued calls to speed the discovery and description of species for a world that is rapidly losing them to extinction (Grace et al. 2021, Mabry et al. 2022). Recent studies show that rates of specimen collection are declining (Prather et al. 2004, Rohwer et al. 2022) and backlogs of unidentified specimens from understudied taxonomic groups are accumulating (Bendull 2022, Lewis and Budke 2022).

Scientists have advocated for the creation of better measures of professional productivity to strengthen the foundation of human resources needed to conduct biodiversity research in the 21st century (McDade et al. 2011, Thessen et al. 2019, Groom et al. 2020, Lagomarsino and Frost 2020). Currently accepted metrics that define professional success for taxonomists focus on the number of publications that they produce yet disregard the intellectual contributions they make to natural history collections in the form of expertly identified specimens. Activities of specimen collection, taxonomic identification and

annotation produce knowledge about biodiversity that is used by other scientists, akin to that found in published papers. Advocates argue that greater visibility of these activities could help sustain the taxonomic expertise needed to support biodiversity discovery and identification.

One way for taxonomists to document their research impact is to use 21st century digital scholarship tools to quantify their direct contributions to collections. With these tools that interact with major biodiversity informatics databases, collectors and determiners can also track the contributions of their specimens to subsequent scientific discoveries and thereby document greater professional recognition of their work. These tools include the web-based services provided by the Open Researcher and Contributor ID initiative (ORCID 2024), Wikidata (Wikidata 2024), the Global Biodiversity Informatics Facility (GBIF 2024), Bionomia (Bionomia 2024) and Zenodo (European Organization For Nuclear Research and OpenAIRE 2013, Shorthouse et al. 2022).

Central to these tools' functionality is the use of globally unique identifiers for the people who collect and identify scientific specimens. These are needed because ambiguity and inconsistency in the way a person's name is recorded as digitised metadata can limit the ability of informatics tools to make accurate attributions. For example, a collector could be referred to on specimen labels as "Helen Jones," "H. Jones," "Miss Jones" and "Mrs. John Smith" over the course of their life, yet informatics tools would regard these name strings as four different individuals. Linking separate name strings into one machine-interpretable knowledge graph through globally-unique identifiers creates a more complete representation of an individual's contributions and productivity. Additionally, for this purpose, Bionomia (Bionomia 2022) is an especially powerful informatics tool because it can be used by anyone to disambiguate names such as those in the example above.

Rationale

Building community awareness of the new informatics tools and attribution practices are essential to increasing the visibility of collections-based researchers and their scientific contributions. The goal is to establish a virtuous circle within the community of people who create or use natural history collections in their research. Each person would:

1. append to their own specimens and annotations their globally-unique identifier,
2. track and report their specimens' use in others' published research using automated informatics tools and
3. cite the specimens of others in their published research using new community best practices to return the courtesy of attribution.

However, there is a lack of awareness and understanding by individuals at all career stages about these tools and practices, which poses a barrier to reaching the critical mass of people necessary to start and sustain new community standards of practice. Students ask, "What are these tools? And do I qualify to use them?". Professionals in early and mid-career stages lament, "Why should I join more science networking sites? I am too pressed

for time as it is". Senior-career individuals respond, "Why should I bother with new tools? My career has been productive without them".

In an effort to overcome the barriers to embracing biodiversity informatics best practices, we created a workshop and symposium series for early-career and established scientists in the skills and practices needed to sustain natural history specimen attribution and citation. The name of the workshop and symposium series, SISRIS, reflected our main goal of effecting community-level change by sharing skills and practices that can support inclusive and sustainable (collections-based) research infrastructure for systematics (Fig. 1). We believe that research infrastructure becomes more inclusive when the taxonomic contributions by all people are accounted for and acknowledged by the community as common practice. We also hold that research infrastructure will become more sustainable because better metrics of professional productivity can help ensure researchers receive appropriate credit from their employers. Such metrics also can be used to educate employers about the global impact of their natural history collections, which may encourage long-term institutional support overall.

Figure 1. doi

Promotional graphic for the SISRIS project used in online and print advertisements for the workshop and symposium. The QR-code leads to the SISRIS project site, https://github.com/aweeks3/SISRIS/. The graphic was created by Shawn Zeringue-Krosnick and Andrea Weeks and is shared under the Creative Commons licence CC BY 4.0.

Accurate attributions increase the research value of the specimens because they can be used to cross-reference collecting events, taxonomy or geographical details, to locate and connect related specimens across institutions and to uncover events in the history of science (Groom et al. 2020). Due to these important linkages, data about the people who collected and identified specimens are an essential part of the Extended Specimen concept (Webster 2017). The widespread use of unique identifiers for people, such as those used by Bionomia (Bionomia 2024), would be an important step in implementing this concept within biological collections' databases in the future, such as the Extended Specimen Network (Biodiversity Collections Network (BCN) 2019).

Few recent workshops to our knowledge have broached the challenge of changing the way collections-based research is measured, made visible and acknowledged by others at a community-level (Table 1). These few have trained groups of people on the use of Bionomia for tracking the contributions to biological research collections made by others and helped facilitate the transfer of tacit knowledge about collectors and identifiers. Only one workshop (NSF RAPID #2033973), which was delivered to bat collectors, has covered the entire process of specimen attribution and citation.

Table 1.

List of recent workshops that have incorporated one component of training, Bionomia. No prior workshop to our knowledge has matched the entire set of learning objectives of the SISRIS workshop/symposium series.

1. Bionomia, "People Identifiers Roundtrip to Collections Management Systems", 6 December 2023. Online, International.
2. Systematikdagarna, "Make Your Natural History Collections Available and Citable", Lund University, Sweden, 17 November 2023, In Person, European Union-based.
3. BioQUEST Biology and Math Educators Institute, "Revealing hidden figures within natural history collections through data sleuthing". 21 July 2022. Online, US-based.
4. Society for the Preservation of Natural History Collections, "People are unique, unique people are priceless". 9 June 2022. Edinburgh, Scotland.
5. iDigBio, "Using Wikidata to Capture and Share Information about People in Paleontology". 29-31 March 2022. Online, US-based.
6. Serie de webinars SiB Colombia, "Bionomía: una plataforma para visibilizar a las personas detrás de las colecciones biológicas". 15 April 2021, Online, Latin American-based.
7. NSF RAPID #2033973 Rapid Creation of a Data Product for the World's Specimens of Horseshoe Bats and Relatives, a Known Reservoir for Coronaviruses (Leachman 2020, Mast et al. 2021), "The "Who" in Collections: Revealing the Network of Collectors and Determiners of Bat Specimens". 1 December 2020. Online, US-based.

Aims of the workshop

Workshop format and learning objectives for participants

The half-day workshop format comprised four hours of instruction about the application of the web-based informatics tool Bionomia and the other software services that interact with this platform, including ORCID, GBIF, Wikidata and Zenodo (Table 2). The aim of the workshop was to address concerns held by participants at all career stages, with an emphasis on the practical benefits such tools can bring to them personally and to the biological research collections' community more broadly. Workshop-based activities were aligned with six major learning objectives (Table 3). The two major goals to be achieved through these objectives were to support participants' learning how to use these tools as part of their collections-based research activities and as a means of community-level change in how taxonomic expertise is acknowledged and rewarded. For the latter goal, participants practised attributing specimens to historical (deceased) collectors from an under-represented group. Workshop organisers assembled a list containing the names and

Wikidata profiles for over 720 historical female botanists from the United States who had herbarium specimen data potentially available for attribution on the Bionomia platform.

Table 2.

Workshop learning objectives and the activities that participants undertake to achieve them.

By the end of the workshop, participants will be able to explain and/or demonstrate that:

1. People and people-data are essential for advancing collections-based research.
    1. Participants will read Groom et al. (2020) prior to the workshop.
    2. Participants will be introduced to an example case study as part of introductory remarks for the workshop.
    3. Participants will take part in uncovering the specimens and identifications made by a historical collector from an under-represented group in science, chosen from a list generated by workshop organisers.
    4. Participants will be invited to listen to a symposium of related case-studies held before or after the workshop.
2. People-data are not available equitably, but web-based informatics tools are available to address this imbalance.
    1. Participants will be introduced to supporting data as part of the introductory remarks.
3. Web-based informatics tools raise the visibility of people who collect and identify natural history specimens.
    1. Participants will create their public ORCID profile.
    2. Participants will view a video about these tools prior to the workshop, as a primer on controlled vocabulary, global unique identifiers and Bionomia to understand the role of people-data in extending natural history specimens.
    3. Participants will practise using Bionomia to claim their own specimens and identifications.
    4. Participants will practise using Bionomia to claim specimens for one or more historical collectors from an under-represented group.
4. Web-based informatics tools track research contributions of people who collect and identify natural history specimens, both directly (e.g. the science they have authored personally) and indirectly (e.g. the science enabled by their specimens and annotations).
    1. Participants will learn how ORCID, Bionomia and Zenodo can track and report their direct and indirect research contributions made possible by their specimens and taxonomic identifications.
    2. Participants will learn how to cite their research contributions within biological research collections.
    3. Participants will learn about how to properly cite specimen datasets in their published papers to allow tracking by web-based informatics tools.
    4. Participants will use Bionomia to discover the science enabled by the specimens created by or identifications made by one or more historical collectors from an under-represented group.
5. Web-based informatics tools track the research impact of biological research collections and improve collections management.
    1. Participants will practise how to report the research impact of entire biological research collections using Bionomia and GBIF tools.
    2. Participants will listen to an example of how name-disambiguation improves collections management (e.g. Güntsch et al. (2021)).
6. Workshop participants can amplify this advocacy work after the workshop concludes.
    1. Participants will view instructors creating a Wikidata entry for a newly-uncovered historical collector.
    2. Participants will be given links to the GitHub repository of workshop documents.
    3. Participants will receive examples of tasks, small and large, that can advance effort as part of concluding remarks.
    4. Participants after the workshop will receive a cumulative report on the impact of participants' activities, alongside the two-month workshop survey for assessment purposes.
    5. Participants will be sent a copy of the workshop analysis paper when it is published.

Table 3.

The SISRIS workshop schedule of activities within a four-hour format.

- 08:00–08:10: Welcome and logistics
- 08:10–08:30: Introductory remarks
- 08:30–09:45: Hands-on: Participants disambiguate names of collectors from under-represented groups on the Bionomia platform.
- 09:45–10:15: Break
- 10:15–11:35: Hands-on: Participants document their own expertise and research contributions to biological research collections using Bionomia, learn how to use additional tools to augment their curriculum vitae and future publications and learn how to track collection-level research impact.
- 11:35–11:45: Concluding remarks about how to amplify this effort after the workshop
- 11:45–12:00: Hands-on: retrospective pre-test survey for assessment purposes.
- 12:00: Adjourn.

Symposium and workshop venues, participant solicitation and assessment

Workshops were held in three venues. Two in-person workshops, followed by symposia, were held at the annual meeting of the Association of Southeastern Biologists (Winston-Salem, North Carolina; 25 March 2023) and Botany 2023 (Boise, Idaho; 25 July 2023). Participants of in-person workshops were offered a stipend to defray the cost of travel (325 to 1000 USD), which was provided by a US National Science Foundation conference award (Grant No. 2247631 and 2247632). An online workshop-only event was held on 6 August 2023 via Zoom to accommodate individuals who could not attend in person. The workshops and symposia were advertised widely by the professional societies participating in the meetings and via social media. A promotional graphic that included a QR-code to the SISRIS project web page and registration site (Fig. 1; https://github.com/aweeks3/SISRIS) was used in all communications.

Workshop assessment was approved by the Tennessee Technological University Institutional Review Board and conducted using Qualtrics. Assessment activities took place during participant registration, immediately after each workshop event and 60 days afterwards to assess longitudinal effects. For the purpose of assessment, participants were categorised in one of two cohorts: junior-level (undergraduate students through post-doctoral researchers) or senior-level (early-career through retired). Attendance was used to assess the success of the symposia.

Key outcomes and discussions

Diverse workshop participants

Fifty-three workshop participants comprised two (4%) undergraduate students, five (9%) masters students, fifteen (28%) doctoral students, one (2%) post-doctoral researcher, twelve (23%) early-, thirteen (25%) mid- and four (8%) late-career professionals and one (2%) retired individual. The junior-level cohort (n = 23) was smaller than the senior-level

cohort (n = 30). The 53 participants included 51 unique individuals because two individuals participated in both in-person workshops. Amongst participants who reported their gender, race and ethnicity, 36 (68%) were female, 15 (28%) male and two (4%) non-binary. Of these individuals, 17 (32%) were minority (e.g. non-white) and three (6%) were Hispanic. Each workshop increased the geographic diversity of the participants, as determined by their institutional address or, if unspecified, home address (Fig. 2 A, B and C). In total, SISRIS workshop participants were drawn from 23 US States (Fig. 2 D).
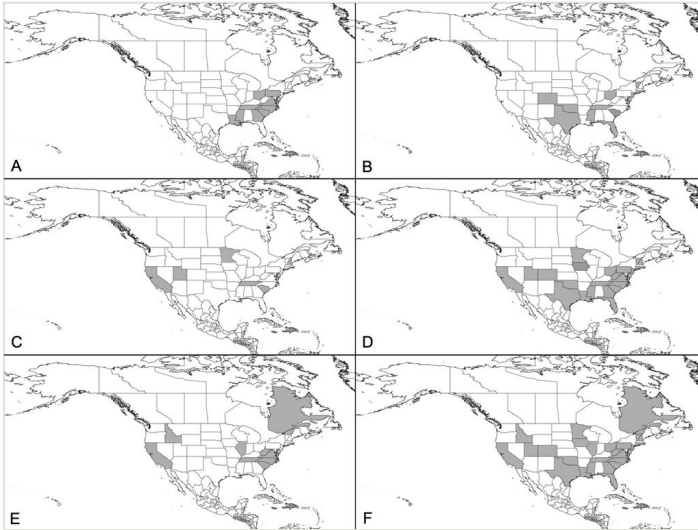

Figure 2. doi

Geographic affiliation of SISRIS workshop participants, symposium speakers and project personnel as determined by their institutional address or, if unspecified, home address. A) Participants of the SISRIS workshop held at ASB 2023 in Winston-Salem, North Carolina USA, 24 March 2023; (n = 20). B) Participants of the SISRIS workshop held at Botany 2023 in Boise, Idaho USA, 23 July 2023; (n = 27). C) Participants of the SISRIS workshop held online via Zoom on 6 August 2023; (n = 6). D) All SISRIS workshop participants, (n = 53). E) All invited speakers of the SISRIS symposia held at ASB 2023 and Botany 2023, (n = 14). F) All SISRIS workshop participants, symposium speakers and project personnel in 2023, (n = 59 different individuals).

Assessment data show that the workshops trained people who had significant financial barriers to obtaining this type of professional development. For instance, 51% of in-person workshop participants reported that the participation stipends were definitive in allowing them to attend the conference. Moreover, 22% of in-person workshop participants reported that the stipend was definitive in attracting them to participate in the workshop. The positive effect of the participation stipend on conference attendance was stronger amongst junior individuals (undergraduate students through post-doctoral researchers; 68%) as compared to senior individuals (early-career through retired; 35%). However, both groups reported equivalent levels of being attracted to participate in the workshop because of the stipend (22% and 23%, respectively).

Workshop outcomes

As our assessment data show, workshops were effective training experiences for participants. The workshop registration process alone improved the ability of participants to document their own contributions to research. For example, 26% of the workshop participants obtained their ORCID because of workshop registration activities. Amongst junior-level participants, this percentage was 43%, whereas it was 13% for senior-level participants.

At the end of the workshops, most participants reported substantial gains in their understanding of the use of biodiversity informatics tools (Fig. 3). For example, before the workshops, only 14% of participants reported average, high or expert knowledge of Bionomia, but immediately after the workshops, 96% of participants did. This change reflects a gain of 82 percentage points, overall. Participants showed similar gains for the other tools introduced during the workshops: GBIF (+25%), Wikidata (+69%), ORCID (+49%) and Zenodo (+29%). Moreover, participants reported high levels of motivation (> 95%; Fig. 4) to continue using these tools and sharing their knowledge with others, as measured immediately after the workshops and 60 days later. Knowledge gains made by the groups of junior and senior individuals were similar across tools, except ORCID and Zenodo. Junior individuals reported a gain of 68% and 40% for ORCID and Zenodo, whereas senior individuals reported a gain of only 35% and 21%, respectively. More junior individuals also reported a higher percentage of familiarity with Zenodo immediately after the workshop (45%) than did senior individuals (28%).
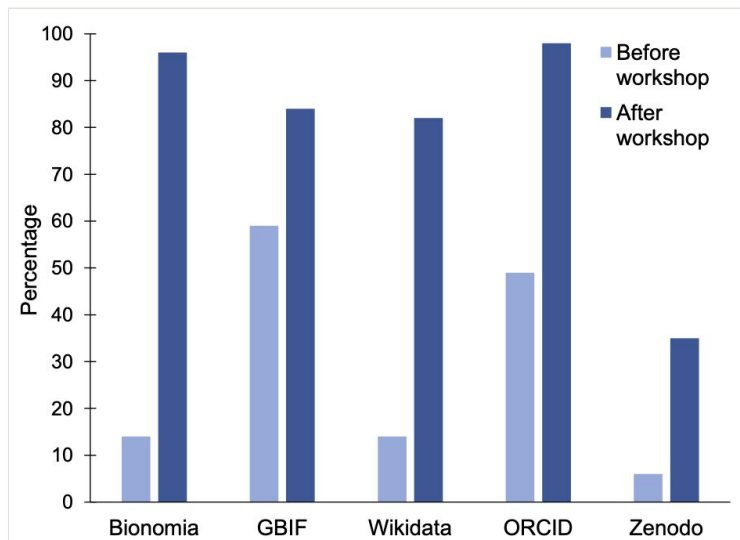

Figure 3. doi

The percentage of SISRIS workshop participants (n = 51) self-reporting "average" or "high" or "expert" familiarity with each software tool before the workshop and immediately after the conclusion of the workshop.
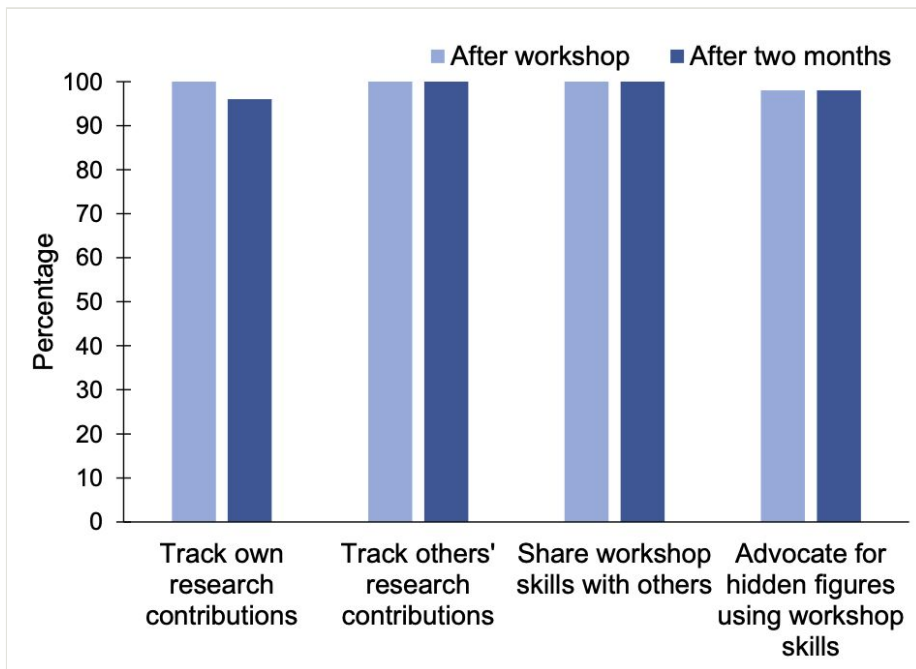
Figure 4. doi

The percentage of SISRIS workshop participants self-reporting "possibly" or "certain" in response to queries about if they may use the skills gained from the workshop in the future, as measured immediately after the conclusion of the workshop (n = 51) and two months later (n = 45).

Participants' free-response answers to assessment questions reveal the scope of the insights that they developed as a consequence of attending the workshops. In response to the query, "What was the most surprising or interesting thing that you learned at the workshop?", three direct quotes illustrate these types of insights: "That I can get a DOI for my own collections that I can use to advocate [for] the importance of collections and the work [I] do with them (especially to admins!)"; "I did not realise the scope of SISRIS to be able to attribute to collectors that have passed. The breadth of this was incredible. For example, my undergraduate research was entirely based on the floristic work of Emma Jane Cole of Michigan, who is one of the women on the SISRIS list. If we had Bionomia during my undergraduate [time], our work would have been very streamlined. This is wonderful to add to the floristic/herbarium community"; and "How to create a Wikidata page for a person who is no longer alive, this will be a wonderful asset to our herbarium by being about to put information about our collections in a publicly accessible place rather than a document on one person's computer". A thematic evaluation of all free-response answers (Table 4) shows the full range of insights that participants gained as a consequence of attending the workshops.

Table 4.

Themes from free-response comments of workshop participants.

Queried at each workshop conclusion: What was the most surprising or interesting thing that you learned at the workshop?

1. Having a way to quantify and take credit for our collections and contributions to science, regardless if it results in a publication or not; conversely, how many people aren't getting the credit they deserve (includes finding publications that use your data).
2. Bionomia and data sleuthing techniques are empowering to learn more about people's history, their stories and their contributions, both living and dead.
3. Women had a greater role in natural hisotory collections than previously thought.
4. Connections between data sources (Bionomia, Wikidata, ORCID, GBIF).
5. Connections between specimens and biographical information – cross disciplinary.
6. The fact that even though the same datasets are shared at different herbaria, they have large differences in the amount of information available.
7. Learning about new tools to gather data on specimens (Wikidata, Bionomia, ORCID) and how the tools work together so well.
8. Learning how Bionomia can inform curators about their collections and the volume of data this can help with.
9. The fact that anyone can help illuminate hidden figures through citizen-science based efforts.
10. Interest in correcting mistakes in GBIF (name inconsistencies, importing titles) revealed via Bionomia.
11. People should cite ORCID as much as possible to get credit for their work.

Queried two months after each workshop: In retrospect, what was the most valuable thing that you learned from the workshop?

1. The fact that we can document our contributions to science in the same way we receive credit for publishing papers.
2. The importance of human contributions to natural hisotory collections and the research they support.
3. Increased awareness of hidden figures in natural history collections (herbaria).
4. Learning how to use essential new tools for the curation of natural hisotory collections.
5. It is extremely important to tie specimens correctly to their collectors and identifiers for proper attribution, need to use tools such as ORCID etc.
6. Learning about the availability of these tools, how to use them and how accessible they are to everyone.
7. The realisation of how important people-data are to understanding natural history collections and how hard it has been to access and how it can be used to give credit to hidden figures.
8. Learning how to create profiles in Bionomia and Wikidata to attribute specimens and give insights into the history of specific people.
9. The potential for using Bionomia in teaching: learning how to sleuth data and how these tools can be applied more broadly to teaching, curation etc.

In comparing junior-level participants with senior-level participants, additional themes were identified (Table 5). In responses from participants collected on the day of the workshops, both groups most frequently cited the value of gaining experience with ORCID, Bionomia and Wikidata, with 55% of all responses focused on that theme. Within that category, junior participants mentioned this theme 42% of the time, while senior participants mentioned it 58% of the time. Both groups of participants cited the value of increased attribution of individuals, though junior participants mentioned this theme more often (78%) than senior participants (22%). In contrast, senior participants mentioned the discovery of new biographical information more frequently (60%) than junior participants (40%). Senior participants mentioned the value of Bionomia as a means to document personal productivity at a greater frequency (80%) relative to junior participants (20%).

Table 5.

Themes from free-response comments by workshop participants that unite or distinguish junior and senior-level cohorts.

1.   Gaining experience with Bionomia, Wikidata and ORCID.
2.   Insights into a curator's home herbarium.
3.   Increasing attribution of the work of individuals.
4.   Discovery of new biographical information for individuals.
5.   Having a new mechanism to measure personal productivity.
6.   The importance of peer education.

The positive long-term effects of training are also revealed by participants' self-reporting measured 60 days after the workshops ended (Fig. 5). For example, 73% of participants continued using Bionomia after the workshops. Participants reported having "increased greatly" their use of Bionomia (53%), GBIF (22%), Wikidata (24%), ORCID (31%) and Zenodo (18%) as compared to before and after the workshops. Participants reported using Bionomia (51%), GBIF (31%), ORCID (44%) and Zenodo (13%) to support their own research; use of Wikidata in this way was not polled. Some interesting shifts in themes were observed in the 60 day free-responses of junior and senior participants. Both groups continued to cite their understanding of these tools most frequently (54%) amongst all responses. The relative number of junior and senior participants mentioning increased attribution of individuals and the discovery of biographical information remained consistent amongst responses taken the day of the workshops and those collected 60 days after. However, while senior participants mentioned the importance of Bionomia in measuring personal productivity more frequently than junior participants on the day of the workshop, this decreased to an equal split between the two groups 60 days after the workshops. Interestingly, only senior participants mentioned the importance of new insights into their institution's herbarium and the importance of peer education two months after the workshops. These two themes may be more relevant to those who have gained a greater perspective of collections-based research over a longer career trajectory.

An unexpectedly strong outcome of training is the degree to which participants shared knowledge or skills they learned at the workshops with others afterwards (Fig. 6). For example, 93% of participants shared what they learned with at least one botanical colleague or student, with 26% of participants telling five or more people. Sixty-two percent of participants shared with at least one non-botanical colleague or student (9%, five or more people) and 67% of participants shared with at least one non-academic acquaintance (4%, five or more people). Percentages reported by junior and senior participants were broadly similar, except that more senior participants reported sharing knowledge with their non-botanical academic networks. For example, 70% of senior participants shared with at least one non-botanical colleague or student, whereas only 50% of junior participants did. Conservatively estimated, SISRIS participants disseminated workshop information to over one hundred individuals who did not directly attend the workshop themselves.
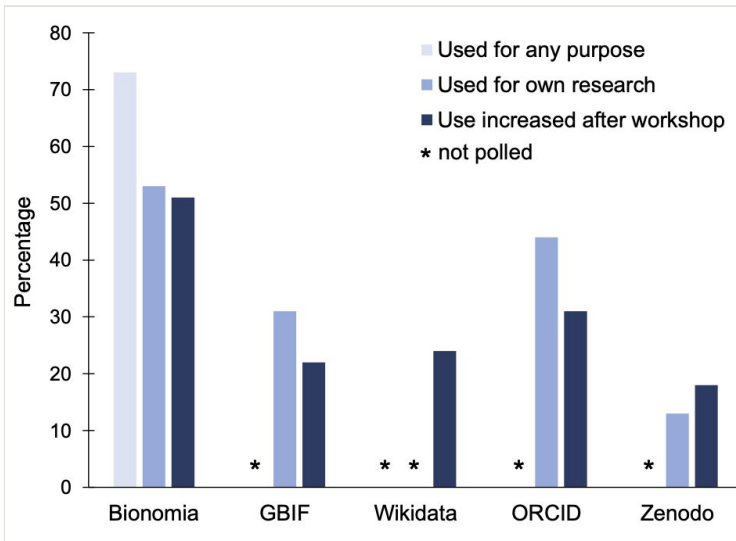
Figure 5. doi

The percentage of SISRIS workshop participants (n = 45) self-reporting use of software tools in the two months after the conclusion of the workshop.
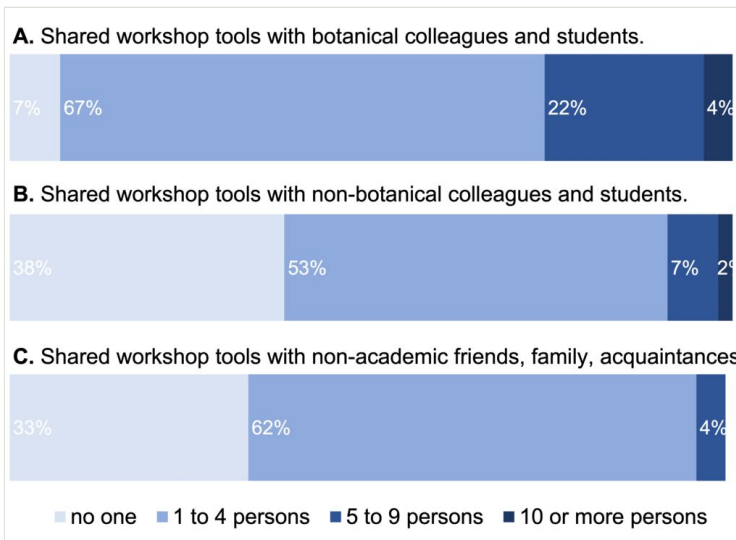

Figure 6. doi

The percentage of SISRIS workshop participants (n = 45) self-reporting that they shared knowledge and skills with individuals in their communities during the two months after the conclusion of the workshop.

During the workshops, participants improved the records for 90,492 herbarium specimens with attributions for 135 different collectors and identifiers using Bionomia, including 95 historical collectors from under-represented groups. In some cases, workshop participants

attributed their own specimens (e.g. herbarium specimens they created or identified) to quantify their contributions and to augment their institution's research impact. Senior individuals were more productive using Bionomia during the workshop than junior individuals. They worked on more collectors and/or identifiers on average (four vs. two people) and attributed more specimens on average (2923 vs. 643 specimens). Both senior and junior participants worked on the same number of under-represented collectors/identifiers (two), but senior participants were still more productive on average (1654 vs. 527 specimens). The causes underlying this performance difference may include senior participants' more extensive personal collections needing attribution, their greater experience in evaluating historical specimen labels and their adoption of bulk-attribution methods within Bionomia.

Based on feedback from the first workshop, we expanded the training module for creating Wikidata entries for deceased collectors or identifiers in subsequent workshops. A number of participants knew of other historical individuals that needed Wikidata Q-numbers in order to attribute their specimens within Bionomia. This turned out to be a well-received improvement in the following workshops (Table 4), as it was a common theme in response to the assessment question, "In retrospect, what was the most valuable thing that you learned from the workshop"?

Symposium outcomes

The symposia comprised nine and ten invited speakers at the in-person conferences, respectively (Table 6). In total, 14 different individuals from the United States (eight States and the District of Columbia) and Canada (Fig. 2 E) presented different perspectives regarding the application of people-data in herbarium collections for advancing collection management, specimen curation, research and education. At least 140 conference attendees were part of the in-person audiences. All Botany2023 SISRIS presentations were broadcast live online as well as recorded for asynchronous viewing as part of the virtual conference programme, which remained accessible after the event. Facilitated discussions during and after the symposia revealed that the community is receptive to embracing new biodiversity best practices, many of which were new to audience members.

Conclusions

The enthusiastic response to the SISIRIS initiative by workshop participants, symposium speakers and audience members indicates that it tapped into a need within the community to learn, to share and to discuss how people-data facilitate specimen-based research. For example, we had to expand the length of each symposium to accommodate a greater number of speakers than originally anticipated due to the widespread interest amongst researchers we approached to present their work. SISRIS events included people from the majority of US States (Fig. 2 F). Double-digit learning gains by both junior and senior workshop participants for basic digital scholarship tools, such as ORCID, Wikidata and Zenodo and biodiversity informatics tools, such as Bionomia and GBIF, reveal a continuous

need for these types of professional development training events in the future. Given that the majority of workshop participants would not have attended the in-person scientific conferences without the assistance of the participation stipend and one-fifth of participants only attended the workshop because of it, improving community-level awareness of the need for this type of professional development training and access to it will be a continuing challenge.

Table 6.

The SISRIS symposia speakers and their presentation titles.

Association of Southeastern Biologists 2023:

- Deborah L Paul (University of Illinois). "Discovering together "who dunnit?": building a borderless knowledge community".
- David Shorthouse (Bionomia). "Connecting scientists and their specimens through Bionomia".
- Shawn Zeringue-Krosnick (Tennessee Tech University), Olubunmi Aina, Adania Flemming, Molly Phillips, Jennifer Kovacs, Siobhan Leachman and Makenzie E. Mabry. "Digital data sleuthing and storytelling as tools to engage students with the unsung heroes of natural history collections".
- Andrea Weeks (George Mason University). "An exception proves the rule: Lena Artz (1891–1976) and her legacy of botanical specimens".
- Carol Ann McCormick (The University of North Carolina). "Delving into the lives of herbarium collectors".
- Daniel Koenemann (Claflin University), Janelle Burke. "Prolific plant collectors and the HBCU legacy: Charles Parker as a case study".
- Pamela Puppo (Marshall University), Lori Thompson and Lindsey Harper. "Uncovering hidden figures at the Marshall University Herbarium through inter-disciplinary collaborations".
- Robin Lewis (Spartanburg Community College), Hannah Bendull and Jessica Budke. "The State of Herbarium Backlogs: Perspectives from Bryophyte Collections".
- Zack Murrell (Appalachian State University), Michael Denslow, Herrick Brown, Andrea Weeks and Shawn Krosnick. "Strategies to enhance and sustain an established cyberinfrastructure".

Botany 2023

- David Shorthouse (Bionomia). "Connecting scientists and their specimens through Bionomia".
- Deborah L Paul (University of Illinois). "Discovering together "who dunnit?": building a borderless knowledge community".
- Erica Krimmel, Holly Little. "People in the biodiversity knowledge graph and their roles in building the data connections we need".
- Katie Pearson (Symbiota Support Hub), Ed Gilbert, Jenn Yost, Greg Post and Nico Franz. "Even more hidden figures in biodiversity specimen collections: Tracking agent contributions in Symbiota".
- Herrick Brown (University of South Carolina), Zack Murrell, Michael Denslow, Andrea Weeks and Shawn Krosnick. "Strategies to enhance and sustain an established cyberinfrastructure".
- Jessica Budke (University of Tennessee-Knoxville), Hannah Bendull and Robin Lewis. "The State of Herbarium Backlogs: Perspectives from Bryophyte Collections".
- Rebecca Dikow (Smithsonian Institution), Jenna T. B. Ekwealor, William J. B. Mattingly, Michael G. Trizna, Elizabeth Harmon, Torsten Dikow, Carlos F. Arias, Richard G. J. Hodel, Jennifer Spillane, Mirian T. N. Tsuchiya, Luis Villanueva, Alexander E. White, Madeline G. Bursell, Tiana Curry, Christelle Inema and Kayla Geronimo-Anctil. "Let the records show: attribution of scientific credit in natural history collections".
- Andrea Weeks (George Mason University). "An exception proves the rule: Lena Artz (1891–1976) and her legacy of botanical specimens".
- Janelle Burke (Howard University), Dominique Pittman-Kidd and Daniel Koenemann. "Prolific plant collectors and the HBCU legacy: Charles Parker as a case study".
- Shawn Zeringue-Krosnick (Tennessee Tech University), Olubunmi Aina, Adania Flemming, Molly Phillips, Jennifer Kovacs, Siobhan Leachman and Makenzie E. Mabry. "Digital data sleuthing and storytelling as tools to engage students with the unsung heroes of natural history collections".

The strongest outcome of the SISRIS initiative is the degree to which workshop participants once trained have continued to share their new knowledge with others. The long-tail effect of training was an intended goal of the workshop and is one that, in our opinion, has the best chance of building community-level awareness of 21st century informatics tools and attribution practices. After the last round of formal assessments that documented widespread dissemination on the part of participants 60 days after the workshops, one alumna has created a scientific presentation about the impact of Bionomia training (Struwe and Struwe 2023) and a group of three alumni have begun organising a similar symposium for the 2024 annual meeting of the Society of Herbarium Curators in collaboration with the Historical Section of the Botanical Society of America at Botany 2024 in Grand Rapids, Michigan. Two other alumni have trained groups of herbarium curators and plant collectors using techniques learned at the SISRIS workshops in West Virginia (Gregg 2024) and South Africa (Morgan Gostel, pers. comm.). Additionally, at least one biography of a historical collector has been written using research techniques developed in the workshop (Weeks 2024). The community of collections-based researchers has always had an interest in the history of science, but we attribute this contemporary enthusiasm in large part to the growing realisation of the research potential contained by the burgeoning digital data now available about collections and their collectors.

The SISRIS project will hold its final, grant funded workshop 16 June 2024 at the Botany 2024 conference in Grand Rapids, Michigan. Future work will include creating a white-paper about the need for continuing professional development regarding the opportunities presented by the Extended Specimen concept, including people-data and disseminating this model of training to other domain-specific groups of collections-based researchers, curators and students.

Acknowledgements

Funding program

Grant title

Conflicts of interest

Bionomia is a project developed and maintained by co-author David P. Shorthouse. It does not form part of his official duties as Biodiversity Data Manager with Agriculture and Agri-Food Canada.

References

- Bendull H (2022) Herbarium backlogs: Challenge or opportunity? Zenodo https://doi.org/10.5281/zenodo.6593503
- Biodiversity Collections Network (BCN) (2019) Extending U.S. biodiversity collections to promote research and education. American Institute of Biological Sciences.
- Bionomia (2024) Bionomia. https://bionomia.net/. Accessed on: 2024-4-11.
- European Organization For Nuclear Research, OpenAIRE (2013) Zenodo: Research. Shared. Zenodo https://doi.org/10.25495/7gxk-rd71
- GBIF (2024) Global Biodiversity Information Facility. https://www.gbif.org/. Accessed on: 2024-11-04.
- Grace O, Pérez-Escobar O, Lucas E, Vorontsova M, Lewis G, Walker B, Lohmann L, Knapp S, Wilkie P, Sarkinen T, Darbyshire I, Lughadha EN, Monro A, Woudstra Y, Demissew S, Muasya AM, Díaz S, Baker W, Antonelli A (2021) Botanical monography in the Anthropocene. Trends in Plant Science 26 (5): 433-441. https://doi.org/10.1016/j.tplants.2020.12.018
- Gregg KB (2024) West Virginia herbarium curators annual meeting. The Vasculum 19 (1): 8-9.
- Groom Q, Güntsch A, Huybrechts P, Kearney N, Leachman S, Nicolson N, Page RDM, Shorthouse DP, Thessen AE, Haston E (2020) People are essential to linking biodiversity data. Database 2020 https://doi.org/10.1093/database/baaa072
- Güntsch A, Groom Q, Ernst M, Holetschek J, Plank A, Röpert D, Fichtmüller D, Shorthouse DP, Hyam R, Dillen M, Trekels M, Haston E, Rainer H (2021) A botanical demonstration of the potential of linking data using unique identifiers for people. PLOS ONE 16 (12). https://doi.org/10.1371/journal.pone.0261130
- Lagomarsino L, Frost L (2020) The central role of taxonomy in the study of neotropical biodiversity. Annals of the Missouri Botanical Garden 105 (3): 405-421. https://doi.org/10.3417/2020601
- Lewis R, Budke J (2022) Bryophyte specimen organization and storage systems: A comparative assessment of staff practices and user preferences. The Bryologist 125 (2): 222-247. https://doi.org/10.1639/0007-2745-125.2.222
- Mabry M, Zapata F, Paul D, O'Connor P, Soltis P, Blackburn D, Simmons N (2022) Monographs as a nexus for building extended specimen networks using persistent identifiers. Bulletin of the Society of Systematic Biologists 1 (1). https://doi.org/10.18061/bssb.v1i1.8323
- McDade L, Maddison D, Guralnick R, Piwowar H, Jameson ML, Helgen K, Herendeen P, Hill A, Vis M (2011) Biology needs a modern assessment system for professional productivity. BioScience 61 (8): 619-625. https://doi.org/10.1525/bio.2011.61.8.8
- ORCID (2024) ORCID. https://orcid.org/. Accessed on: 2024-4-11.

- Prather LA, Alvarez-Fuentes O, Mayfield M, Ferguson C (2004) The decline of plant collecting in the United States: A threat to the infrastructure of biodiversity studies. Systematic Botany 29 (1): 15-28. https://doi.org/10.1600/036364404772974185
- Rohwer V, Rohwer Y, Dillman C (2022) Declining growth of natural history collections fails future generations. PLOS Biology 20 (4). https://doi.org/10.1371/journal.pbio.3001613
- Shorthouse D, Goodwin DZ, Samanta Orellana K (2022) Zenodo enables a new workflow for collectors of natural history specimens. Zenodo https://doi.org/10.5281/zenodo.6761722
- Struwe L, Struwe I (2023) Bionomia, a personal testimony. Systematikdagarna, Lund, Sweden. November 27, 2023. URL: https://rutgers.mediaspace.kaltura.com/media/BionomiaA+Ingemar+Struwe+and+Lena+Struwe+Oct2023+recording/1_x7opvc16
- Thessen A, Woodburn M, Koureas D, Paul D, Conlon M, Shorthouse D, Ramdeen S (2019) Proper attribution for curation and maintenance of research collections: Metadata recommendations of the RDA/TDWG working group. Data Science Journal 18 (1). https://doi.org/10.5334/dsj-2019-054
- Webster MS (2017) The extended specimen. In: Webster M (Ed.) The Extended Specimen: Emerging frontiers in collections-based ornithological research. CRC Press.
- Weeks A (2024) In Memoriam: Lena Artz (1891–1976), a hidden figure of 20th-century southeastern U.S. botany. Castanea 88 (2). https://doi.org/10.2179/0008-7475.88.2.297
- Wikidata (2024) Wikidata. https://www.wikidata.org/wiki/Wikidata:Main_Page. Accessed on: 2024-4-11.