Arctos: Community-driven innovations for managing biodiversity and cultural collections

Carla Cicero[1¶&*], Michelle S. Koo[1¶&], Emily Braker[2¶&], John Abbott[3&], David Bloom[4&], Mariel Campbell[5&], Joseph A. Cook[5,6&], John R. Demboski[7&], Andrew C. Doll[7&], Lindsey M. Frederick[8&], Angela J. Linn[9&], Teresa J. Mayfield-Meyer[10&], Dusty L. McDonald[10&], Michael W. Nachman[1&], Link E. Olson[9&], Dawn Roberts[11&], Derek S. Sikes[9,12&], Christopher C. Witt[5,6&], Elizabeth Wommack[13&]

[1] Museum of Vertebrate Zoology, University of California, Berkeley, California, United States of America

[2] University of Colorado Museum of Natural History, University of Colorado, Boulder, Colorado, United States of America

[3] Department of Museums Research and Collections and Alabama Museum of Natural History, University of Alabama, Tuscaloosa, Alabama, United States of America

[4] VertNet, Sebastopol, California, United States of America

[5] Museum of Southwestern Biology, University of New Mexico, Albuquerque, New Mexico, United States of America

[6] Department of Biology, University of New Mexico, Albuquerque, New Mexico, United States of America

[7] Denver Museum of Nature & Science, Denver, Colorado, United States of America

[8] New Mexico Museum of Natural History & Science, Albuquerque, New Mexico, United States of America

[9] University of Alaska Museum, University of Alaska Fairbanks, Fairbanks, Alaska, United States of America

[10] Arctos Consortium, 1000 Broadway, Suite 480, Oakland, California, United States of America

[11] Chicago Academy of Sciences, 2430 N Cannon Drive, Chicago, Illinois, United States of America

[12] Department of Biology & Wildlife, University of Alaska Fairbanks, Fairbanks, Alaska, United States of America

[13] University of Wyoming Museum of Vertebrates, University of Wyoming, Laramie, Wyoming, United States of America

¶ These authors contributed equally to this work. Other authors are listed in alphabetical order.

& Authors are current or former Officers or Board of Director members for the Arctos Consortium.

* Corresponding author. Email: ccicero@berkeley.edu

Abstract

Museum collections house millions of objects and associated data records that document biological and cultural diversity. In recent decades, digitization efforts have greatly increased accessibility to these data, thereby revolutionizing interdisciplinary studies in evolutionary biology, biogeography, epidemiology, cultural change, and human-mediated environmental impacts. Curators and collection managers can make museum data as accessible as possible to scientists and learners by using a collection management system. However, selecting a system can be a challenging task. Here, we describe Arctos, a community solution for managing and accessing collections data for research and education. Specific goals are to: (1) Describe the core elements of Arctos for a broad audience with respect to the biodiversity informatics principles that enable high quality research; (2) Highlight the unique aspects of Arctos; (3) Illustrate Arctos as a model for supporting and enhancing the Digital Extended Specimen; and (4) Emphasize the role of the Arctos community for improving data discovery and enabling cross-disciplinary, integrative studies within a sustainable governance model. In addition to detailing Arctos as both a community of museum professionals and a collection database platform, we discuss how Arctos achieves its richly annotated data by creating a web of knowledge with deep connections between catalog records and derived or associated data. We also highlight the value of Arctos as an educational resource. Finally, we present a financial model of fiscal sponsorship by a non-profit organization, implemented in 2022, to ensure the long-term success and sustainability of Arctos. We attribute Arctos' longevity of nearly three decades to its core development principles of standardization, flexibility, interdisciplinarity, and connectivity within a nimble development model for addressing novel needs and information types in response to changing technology, workflows, ethical considerations, and regulations.

Introduction

Museum collections are veritable treasure troves of objects and associated data that document biological and cultural diversity across spatial and temporal scales. In recent decades, national and global digitization efforts that promote free and open access to those records have unleashed exciting initiatives in both research and education [1-6]. Furthermore, community science efforts aimed at digitizing museum data have shown that entire communities can be engaged in enhancing the scientific value of museum collections [7]. The vast increase in data available through different platforms has revolutionized interdisciplinary studies in evolutionary biology, biogeography, epidemiology, cultural change, and human-mediated environmental impacts [2, 8-10]. Although collections data are increasingly accessible, initiatives for research, education, and policy benefit the most from carefully curated, high-quality information that comprehensively assembles and links everything that is known about objects in an extended network [11-13].

Collection information management systems range from simple spreadsheets to sophisticated relational databases. Fortunately, advances in informatics focused on biodiversity and cultural heritage have enabled broad-scale aggregation of museum data [14] from different sources through the development of global metadata standards (e.g., Darwin Core, https://dwc.tdwg.org; Dublin Core, https://dublincore.org; Getty Vocabularies, https://www.getty.edu/research/tools/vocabularies). Although these efforts have massively increased the *quantity* of data that are available, the *quality* of data depends strongly on local controls that standardize and improve the consistency of data *values* [11]. Efforts to standardize data benefit from community input, especially when diverse disciplines with varying perspectives are represented [15]. Likewise, FAIR Data Principles for scientific data management and

stewardship (Findability, Accessibility, Interoperability, and Reusability [16]) promote discovery and use of data through transparency, reproducibility, and reusability.

Museum curators and collection managers are faced with a bewildering number of challenges and choices when considering collections digitization, management, and data access. Although collections data are increasingly available online, not all collection management systems have the advanced infrastructure needed to integrate diverse data sets, broaden the scope of accessible data as new technologies become available, and examine complex interactions and processes [17]. Here, we describe Arctos (https://arctosdb.org), a community solution for managing and accessing collections data for research and education. Specific goals to: (1) Describe the core elements of Arctos for a broad audience with respect to the biodiversity informatics principles that enable high quality research; (2) Highlight the unique aspects of Arctos; (3) Illustrate Arctos as a model for supporting and enhancing the Digital Extended Specimen [12, 18]; and (4) Emphasize the role of the Arctos community for improving data discovery and enabling cross-disciplinary, integrative studies within a sustainable governance model.

A brief history of Arctos

The foundation of Arctos was set in 1996 when the Museum of Vertebrate Zoology (MVZ) at the University of California Berkeley developed an information management model for its collections ("MVZ Database Model", [19]). This model was unique at the time in its ability to integrate and manage data from diverse collection types in a single environment, to relate cataloged objects across different collections, and to track and promote access to researchers and educators. The model was implemented as a web-based system and renamed Arctos in 1999 at the University of Alaska Museum (UAM) as part of the Arctic Archival Observatory (National Science

Foundation grant DEB-9981915). The University of New Mexico Museum of Southwestern Biology (MSB) and the MVZ began using this Arctos platform in 2003 and 2008, respectively. In 2009, a separate installation of Arctos that eventually became MCZBase was established at the Museum of Comparative Zoology, Harvard University, as a centralized repository for its collections data. The Texas Advanced Computing Center (TACC) began working with Arctos in 2008, first to host media and later to host and provide database support and security for the entire shared system.

Arctos has averaged ~8-9% annual increase in collection records since its inception over 20 years ago (Fig 1A). Records served by Arctos are globally distributed (Fig 1B), and growth has been concomitant with diversification in the types of data served. Initially developed for vertebrate collections, the University of Alaska Museum became the first institution to add object records from cultural collections in 2014. Arctos now serves rich data across a spectrum of collections beyond vertebrates including archaeology, archives, art, botany, entomology, ethnology, geosciences, history, invertebrate zoology, meteoritics, parasitology, paleontology, teaching, and zooarchaeology. This breadth of data types, along with the accompanying expertise of curators who use Arctos for data management, fosters cross-disciplinary discussions and promotes scientific collaboration and integration.

Fig 1. Growth and geographic distribution of Arctos records.

(A) Growth of data in the Arctos Consortium showing total number of cataloged records by year (1999 through July 2023). Arctos has grown from ~614K records in 1999 to ~5 million records in 2023. (B) Global map of georeferenced localities in Arctos per 100 km2 grid, showing 777,380 spatially distinct localities for over 4 million georeferenced records worldwide with concentrations in North America and Alaska.

Core features of Arctos

Arctos is a community of museum curators, collection managers, researchers, and informatics professionals as well as a database platform for cutting-edge collection management. As such, it provides a full suite of features for governing, hosting, managing, and connecting collection object data, people, transactions, and other information relevant to collections-based research and education. Arctos is implemented in the relational database PostGRESQL/PostGIS controlled by a Virtual Private Database (VPD, [20]) with a Lucee-based web interface, which allows each collection to manage their data independently. Core versions of its software are released under the open-source license Apache 2.0. (https://github.com/ArctosDB).

As a collection management platform and data portal, Arctos provides a comprehensive solution for managing biological, educational, and cultural collections of all sizes (Fig 2) for museums, universities, state and federal agencies, and field stations. In addition, it functions as its own data aggregator and publisher with dynamic (non-static) object-based data housed in multiple collections and institutions. Because it is a hosted and entirely web-based service, individual collections do not need to spend time or financial resources installing or updating software, maintaining servers, responding to security threats, or coordinating backups. Furthermore, the centralized packaging and publishing of records to external data aggregators (e.g., VertNet, [14]; Global Biodiversity Information Facility [GBIF, 21]) frees collection staff from handling this often-cumbersome process. Arctos is supported through a combination of subscription-based fees, external grants, donations, and in-kind support in the form of personnel subsidized by its members. Subscriptions are on a sliding scale based on collection size and ability to pay, and fee waivers are granted to a small number of collections that lack funding support. The Arctos software development model of “release early, release often” means that it responds quickly as

research or collection management needs arise within the consortium. Furthermore, the data model can accommodate a wide variety of data types and values as new collections are added.

Fig 2. Snapshot of Arctos Collection Management System statistics.

Arctos Collection Management System statistics across all collections as of 1 July 2023.

The flexibility of Arctos, combined with its focus on displaying all that is known about a collection object through its integrated data ecosystem, provides a rich platform for scientific and cultural discovery. Its feature-rich components can be categorized into four core areas (Table 1) that are described more fully below: (1) Community; (2) Records; (3) Tools; (4) Connectivity.

Table 1. Summary of the core elements of Arctos.

Community	● Active forum for community discussion of needs and priorities ● Advisory Committee (Board of Directors) with fiduciary and strategic oversight ● Collaborative training, documentation, and proposals ● Executive Officers who oversee daily operations ● Members Council of collection representatives guiding development ● Monthly meetings to discuss issues and community needs ● Network of diverse expertise in managing different types of collections ● Peer mentorship for new and existing collections ● Shared vocabularies and authorities to improve data consistency and retrieval
Records	● Archives ● Artwork ● Cultural objects ● Biological specimens ● Earth science and paleontological materials ● Environmental samples ● Genetic and genomic resources ● Meteorites ● Microbiological samples ● Observational occurrences ● Physical and digital media

Tools	• Automated reminders and data-quality checks • Batch edits and updates of existing records • Collection level control over data management, permissions, access • Creation and management of projects, publications, and citations • Data entry and bulkloading of new records • Data migration and cleaning services • Export of records in standards-compliant formats • Generation of labels, reports, and invoices • Internal statistics and data quality notifications • Management of information about people and organizations • Management of transactions (accessions, loans, borrows, permits) • Object tracking and machine-readable labels integrated with catalog records • Semi-automated georeferencing, reverse geocoding, and mapping
Connectivity	• Agents and organizations • Biological and cultural interactions • Data publishers and aggregators • Digital Object Identifiers • External data repositories • Project management platforms • Taxonomic resources

Community

The Arctos community is a collaborative, self-governing consortium of collection professionals, information experts, researchers, and educators from diverse disciplines and institutions. As a community, Arctos' priority is to make research-grade collection data openly accessible and richly networked for multidisciplinary research and public understanding of natural and cultural history. It also serves as a repository and public portal for curated data associated with specimens under federal ownership (e.g., those collected on lands administered by agencies such as the U.S. National Park Service and Bureau of Land Management, among others), and supports or enhances the mission of government agencies at all levels. Community members share in the governance, policy writing, maintenance, and development of Arctos as a collection management platform and data portal (https://arctos.database.museum). In addition to sharing knowledge and expertise, participants form a network of peers that are available to mentor new

and existing collection representatives, create training modules (e.g., tutorials, https://arctosdb.org/learn/tutorial-blitz; webinars, https://arctosdb.org/learn/webinars; documentation, https://handbook.arctosdb.org), and discuss development needs and priorities. A shared data environment compels Arctos users to collectively manage controlled vocabularies across different nodes of the database (e.g., geography, taxonomy, agents, preparations, attributes) to promote data standardization and discovery. Consequently, proposed vocabulary or changes to the functionality of Arctos undergo a community decision-making process, ensuring that database developments are guided by Arctos users and reflect community needs. By integrating datasets across biological, geological, and cultural collections, Arctos brings together varied perspectives and data types that lead to innovative, integrative, and broadly beneficial new features and capabilities.

In the collaborative community model, all Arctos collections, regardless of size, are able to actively participate in development priorities and are given equal access to mentors, decision processes, programming aid, community discussion boards, and resources. Learning how to use and contribute to Arctos is a collaborative process which includes regular dialogue among data managers and users. This is especially beneficial to personnel who are less experienced or are interested in growing their knowledge of scientific tools for collection management. Although consensus-building presents its own challenges, the Arctos model encourages community-based solutions, workflow efficiencies, and data quality improvements, thereby advancing best practices in collection data management and data fitness for use [11, 22].

Arctos governance consists of an Advisory Committee and a Working Group composed of volunteer officers, institutional members, technical staff, and subcommittees focused on particular database functions (Fig 3). Regular meetings and online communications through GitHub enable the community to discuss specific issues, address questions or concerns, and

resolve problems. Community discussion often focuses on Arctos data that are shared among all collections (e.g., taxonomy, geography, preparations, people, and organizations). Because data standardization is a core tenet of Arctos, database features and enhancements are forged from input and discussion among Arctos users. New features requested by one collection and approved by the community result in a benefit to the community as a whole. Thus, Arctos can be highly responsive to emerging innovations and community needs.

Fig 3. Arctos Consortium organization chart.

The Arctos Working Group drives development priorities, responds to issues, engages in outreach, and produces documentation, among other activities. Strategic and financial planning are overseen jointly by the Arctos Advisory Committee and Arctos Officers. Committees within the Working Group focus on specific issues identified by the Arctos community and meet regularly or ad hoc depending on need (see https://arctosdb.org/contacts for details on all Arctos committees).

Shared data lead to positive benefits in efficiencies and data quality improvements. For example, locality georeferences created by one collection are available to other collection records from the same place, forming a gazetteer of vetted data and reducing redundant staff effort (Fig 4). As a case study, the MVZ acquired an orphaned bird collection in 2005 and was able to match over 60% of the records with georeferenced localities already in Arctos. Beyond gained efficiencies, data collated across Arctos institutions enable novel discoveries about people, organizations, and other parties (e.g., businesses, societies, museums, zoos, and government agencies, all of which are included as ‘agents’ in the database) and their contributions within and beyond Arctos (Fig 5). Biographical and statistical information in Arctos produces a holistic view of an individual’s career-long activities across institutions rather than partitioning that information by institution. It also provides opportunities to discover that names

associated with different collections are indeed the same person, thus allowing for reconciliation of name variants and corroboration of low-resolution identities while making the data richer and more complete. Finally, the ability to store and share biographical information about agents across collections is an important feature of Arctos, especially for cultural and archival records.

Fig 4. Example of how localities and associated georeferences may be shared in Arctos. Data managers can choose to apply georeferences for a specific locality (e.g., Locality 2) to their own cataloged records that may be from the same descriptive place (e.g., Locality 1) but are lacking coordinates and associated metadata.

Fig 5. Synopsis of agent activity summary in Arctos. Arctos provides a holistic view of research and curatorial contributions by people, organizations, and institutions (i.e., “agent activity”) to facilitate data relationships, attribution, and assessment metrics. Dynamic links associate agents with related agents (e.g., academic lineages, family members), external resources (wikidata, ORCID), collection activities (objects collected, prepared, or identified), curatorial work (transactions, edits, georeferences), and projects and products (publications, media, grants, expeditions).

Records

Arctos serves records on over five million specimens, objects, and observations (biological, archival, cultural, geological, and meteorological) curated by participating collections and institutions. The shared database environment allows curators to manage records that cross disciplines (Fig 6), such as objects with both biological and cultural materials or significance, artworks made of iron that intersect with geology, or fossils with mineral taxonomy. Collections can add a mixture of scientific names and/or associated taxa to the identifications of their

objects, thus allowing the records to be useful for ethnographic, geologic, and biodiversity research.

Fig 6. Example of an Arctos record that cross-links cultural and biological records. Arctos record from the University of Alaska Museum Ethnology and History Collection (https://arctos.database.museum/guid/UAM:EH:UA91-014-0001) showcasing a cultural object composed of biological materials that are cross-linked to biological records in Arctos.

Nearly one million media records (e.g., still images, sound recordings, video files) enrich the data in Arctos. Media may showcase objects and specimens before and after preservation [23-24] or function as a voucher for non-specimen observations. In addition, they can be key to increasing access to rare or fragile collection objects that are not typically loaned, such as eggs and nests [25-26], type specimens [27], and objects that are no longer available [28-29]. Increasingly, specimens are being used in high-resolution photogrammetry and three-dimensional scanning projects [30-32] that can be linked directly to the Arctos record. Finally, Arctos media that are linked to collecting events add value by documenting field work, landscapes, habitats, and people.

Tools

Collection management tools are the nuts and bolts of Arctos functionality, and are used for basic data entry, editing, and searching as well as to improve data quality and increase discoverability. Arctos reports generate transaction invoices, collection ledgers, and dry or wet labels. Different tools find duplicate agents, gaps in catalog numbers, records without parts, unreciprocated relationships between two cataloged items, and records that potentially correspond to GenBank (National Center for Biotechnology, NCBI) accessions but need

verification and linking, among other data quality issues. Collection contacts may receive specific reminders with notifications about loans that are due or permits that are expiring. Statistics generated across all of Arctos, or filtered by collection, provide summaries of specific data such as numbers of cataloged items, localities, georeferenced localities, collecting events, agents, media, publications, GenBank links, and specimen relationships.

Access to different tools depends on a user's training and role. For example, students and volunteers may be given access to enter data for a specific collection by choosing values from existing controlled vocabularies but may not create new taxonomies, higher geography, or names of people and organizations. The addition of new values to controlled vocabularies is limited to collection curators or focused Arctos committees. This hierarchy of permissions limits misspellings and duplication of values (e.g., the same person entered multiple ways), ensures that entries are verified (e.g., a country, state, or county is valid), and compels consistent values for certain fields (e.g., sex, preparations). Ultimately, the focus on standardizing data *values* leads to higher data quality [11] and increases discoverability for researchers and educators using specific criteria.

Below we describe how tools for the following core functions operate in Arctos: data entry and encumbrances; taxonomy and identifications; transactions; object tracking; and spatial data quality.

Data entry and encumbrances

New records are entered individually or through a variety of batch tools, validated by Arctos through a series of data checks and accepted by curators prior to data loading. Arctos also has implemented workflows to capture data in batches from digitized collection objects such as herbarium sheet labels. Once the data are added, they are immediately accessible online

unless a curator chooses to encumber (i.e., restrict access to) those records. Encumbrances may protect sensitive data such as sacred cultural information, collecting locations for fossils or endangered species, and archaeological resources. The Arctos community has developed guidelines for data redaction measures as required by paleontological and cultural collections to meet U.S. federal and private land regulations (e.g., Paleontological Resources Preservation Act of 2009, 16 U.S.C. § 470aaa 1-11; National Historic Preservation Act of 1966, Public Law 89-665; 54 U.S.C. 300101 *et seq;* and Archaeological Resources Protection Act of 1979, 16 U.S.C. 470aa-470mm; Public Law 96-95 and amendments). Likewise, encumbrances restrict usage of collection objects and data according to permit conditions and material transfer agreements (e.g. Nagoya Protocol [33]).

Taxonomy and identifications

Identifications within Arctos are treated separately from taxonomy, both during data entry and editing. Taxonomy refers to formal classification systems including cultural lexicons, and Arctos harvests data from external web services (e.g., Global Names Architecture, https://globalnames.org) while allowing for customized taxonomies. This provides both data quality control and the flexibility of collections to choose and modify their own taxonomy. Collection staff then apply those names to collection objects through a flexible identification module that allows for vernacular, regional, and Indigenous names, taxonomic uncertainties, biological realities (e.g., hybrids, intergrades, new species with temporary designations), multiple identifications, non-hierarchical identifications, and nomenclature from geological, archival, art, and cultural collections. For example, a hybrid specimen is identified by selecting two parental names from the Arctos taxonomy table, and a parka made of furs from different species is identified by multiple taxonomic names that comprise the components (e.g., parka, *Bos taurus, Rangifer tarandus*, *Canis lupus*, *Gulo gulo*; Fig 6). This can also be applied to egg

sets with nest parasites, which are identified by the taxonomic name of both the host and the parasite (e.g., *Melospiza melodia* and *Molothrus ater*, https://arctos.database.museum/guid/MVZ:Egg:609). Importantly, Arctos records the history of all changes to identifications; when new identifications and associated metadata (e.g., determiner, date, basis of determination) are added, old but invalid (i.e., erroneous or synonymous) identifications are retained and remain searchable. Arctos also allows batch identification updates, critical for management of entomology as well as other collections.

Transactions

Transactions include accessions, loans (outgoing collection material), borrows (incoming material from another collection), and permits, all of which are managed in Arctos. Accessions, loans, and borrows are collection-specific, with options for formatting transaction numbers depending on in-house curatorial practices. Permit metadata, on the other hand, are shared among collections and can be linked to accessions, loans, and borrows (the permits themselves are private unless a collection chooses otherwise). This facilitates the management of permits, material transfer agreements, memoranda of understanding, and other documentation for compliance with state, federal, and international (e.g., CITES, Nagoya) regulations [33]. It also allows specimens or objects acquired by one collection under its permit(s) to be accessioned into a different collection using the same permit(s). Likewise, loans sent by different collections can use the same institutional permit. In addition to capturing basic information about the transaction (i.e., persons and/or agencies involved, date, transaction number, nature and amount of material, remarks), transactions can be linked to media such as images or documents (including spreadsheets) that provide supporting information. Arctos also tracks the use of data *about* objects (e.g., metadata or media) through transactions such as data loans and media loans; users may request data records or object photographs rather than objects per

se, allowing more comprehensive documentation about how records, information, and associated elements are being used.

Projects and publications

A special feature of Arctos is the ability to link transactions to thematic web pages called "Projects" that summarize the contributions and use of cataloged records for activities, studies, or uses (e.g., field trips, digitization initiatives, traveling exhibitions, collections on state or federal lands). Projects also are used to document biosampling contributions by Indigenous communities for resource policy and decision-making as well as to highlight cultural object acquisitions, displays, and practices. Publications and media resulting from these activities are easily added to the project page. Further, publications may be cross-referenced to digital object identifiers (DOIs) and cataloged records may be added as citations, directly showing usage of individual records within or across collections. Thus, projects serve as the central hub that link all related data in Arctos to showcase the impact of collections by highlighting how researchers, educators, organizations, and others are using the platform for discipline-specific or interdisciplinary goals (Fig 7; examples in Table 2). Projects are automatically related to other projects based on shared objects, which enables deep-tracking of a collections' utility and products through time. They serve as a convenient hub for agencies, funders, and users to access up-to-date records associated with specific activities.

Fig 7. Synopsis of the Beringian Coevolution Project in Arctos.

The Beringian Coevolution Project (https://arctos.database.museum/project/51) showcases primary source data and collated products. The project included 28,941 cataloged records representing co-examined boreal mammals and associated parasites from two museums and resulted in 8 media objects, 216 publications citing 8,561 specimens, and a dynamic web of 153

related projects that either contributed records used by this project or that used specimens from this project to generate additional research outputs (293 subsequent publications with 69,087 citations).

Table 2. Examples of projects in Arctos that target different users.

Arctos project IDs have the base URL https://arctos.database.museum/project/ (e.g., https://arctos.database.museum/project/10000850).

Topic	Arctos Project ID	Target User(s)
Alaskan Insect Pollinators	10000850	Alaska Department of Fish and Game, Researchers, National Science Foundation
Alaskan Plant Survey	1000054	National Park Service
Albuquerque BioPark & Museum of Southwestern Biology Specimen Repository Agreement	10002948	Researchers, Zoos
Art and Ethnology Collections purchased with funds from Rasmuson Foundation	10003033	Artists, Funder
Art Work Inspired by Life	10002855	Artists
Beringian Mammal and Parasite Coevolution	51 (also see Fig. 9)	Researchers, National Science Foundation

Center for Disease Control Hantavirus Survey in National Parks	10002373	U.S. Center for Disease Control, Epidemiologists, Researchers, National Park Service
Educational Collaboration between Art and Biology	10003671	Students
Support for Ornithology and Herpetology Collections	10003135	National Science Foundation
Mexican Wolf Recovery Program	1000071	U. S. Fish and Wildlife Service, Conservation Groups
Resurvey of Vertebrate Communities in California	10000047	Researchers, Agencies, National Science Foundation
Seal Specimens Hunted by Native Americans in Alaska	15	Alaska Native Harbor Seal Commission
Greenwood Wildlife Rehabilitation Center Salvaged Vertebrates	10004199	Agencies/governments (city, county, state, etc.), Researchers

Object tracking

Object tracking in Arctos acts as an independent module linked via barcodes to the catalog record and provides the capacity to track materials from collection and accession through cataloging, storage, and loans. For example, genetic resource collections can be organized in a hierarchy that associates barcoded cryovials with specific positions in barcoded boxes, racks,

413 freezers, and rooms within buildings (e.g., nested “containers”; Fig 8). Similarly, barcodes are
414 used to track the locations of dry specimens in cabinets or fluid-preserved specimens in jars on
415 shelves. Arctos can accommodate different types of scannable codes (e.g., a true barcode) or
416 non-scannable codes (e.g., cabinet numbers printed on labels) that are captured through the
417 container module, and these can be applied to both cataloged and non-cataloged (in process)
418 items. In addition to tracking the location of collection objects, container environments (e.g.,
419 relative humidity, temperature, ethanol concentration, Integrated Pest Management check) and
420 their history can be tracked to better monitor collections and document changes over time.
421 Containers in Arctos are broadly applicable to a variety of curatorial functions, including
422 management of accessions and loans, collection inventories and moves, conservation and
423 preservation of collection objects, and Integrated Pest Management practices.
424
425 **Fig 8. Schematic showing hierarchical object tracking of tissues in Arctos.**
426 In this example, the hierarchy shows where an institutional collection freezer is located and the
427 nested position of a cryogenic vial within a box in that freezer.

428 ## Spatial data quality
429
430 Geography in Arctos adds critical value to records as fundamental metadata and as a measure
431 of data quality. Adding spatial data allows records to be correlated with environmental and
432 geographic data, thus ensuring their usefulness beyond Arctos [34-36]. Over 75% of the more
433 than 850,000 localities in Arctos have associated georeferences that are available to Arctos
434 users and global data aggregators. Localities are shared among all collections in Arctos, which
435 brings advantages (Fig 4). For example, diverse taxa from the same expedition (e.g., snails,
436 fish, and salamanders collected in the same pond) or collected at different times from the same

location can share the same locality in Arctos. This saves georeferencing effort, ensures data consistency, improves discoverability, and stimulates cross-disciplinary integration.

With an emphasis on spatial accessibility and quality, Arctos has a suite of tools for mapping and describing spatial locations of collection objects. A plug-in for GeoLocate (https://geo-locate.org) facilitates semi-automated georeferencing while BerkeleyMapper (https://berkeleymapper.berkeley.edu) provides data visualization and spatial exploration tools. Arctos also uses Google Maps web services for automated data-quality checking, whereby reverse geocoding verifies if coordinates are in the correct higher geography (i.e., continent, country, state/province, county). Higher geographies are defined with polygons, and countries follow a spatially explicit authority (Database of Global Administrative Areas, https://gadm.org) which uses ISO standards (International Organization for Standards, https://iso.org). Polygons are not limited to higher geography but can be used to describe an object's locality instead of point coordinates. Finally, locality attributes add descriptive terms to a place (see https://handbook.arctosdb.org/documentation/geology.html), and localities can be verified and locked once checked by collectors or curators to preserve data integrity.

Connectivity

Arctos prizes connectivity, in which everything that is known about an object and its relationships, interactions, or derivatives can be displayed or made accessible. For this reason, Arctos integrates with a growing list of external data repositories and services (Table 3) that add value to its data records. This core feature makes Arctos a uniquely rich center of collection-related data and tools for the exploration and visualization of biological, geological, and cultural diversity in novel ways. For example, Arctos can integrate with any resolvable identifier, and was the first collection management system to develop reciprocal, dynamic connections

462 between specimen records and genetic data in GenBank. Dynamic linkages from GenBank
463 back to the Arctos record are created when submissions to NCBI involve referencing the
464 specimen voucher in the NCBI “specimen_voucher” field using three-part “Darwin Core Triplets”
465 (institution:collection:catalog number). As of July 2023, over 37,600 Arctos records are linked to
466 associated data in GenBank.
467

468 **Table 3. Arctos connections with external repositories, authorities, and databases.**
469 Links to these sites add value and relevance for collection transactions, curation, and record
470 management for both natural and cultural history collections. Some connections are reciprocal
471 while others integrate values from an authority. Connections may be made through web
472 services automatically or updated manually.
473

Category	Data Service or Repository	URL
Taxonomy & Legal Status	Global Names Architecture	https://globalnames.org
	Hey's Mineral Index	https://www.mindat.org/cim.php
	Integrated Taxonomic Information System	https://www.itis.gov
	Nickel-Strunz Mineral Classification	http://webmineral.com/strunz.shtml
	Species+	https://www.speciesplus.net
	World Register of Marine Species	https://www.marinespecies.org
	Nomenclature 4.0	https://www.nomenclature.info
	Getty Art & Architecture Thesaurus	https://collectionstrust.org.uk
People & Publications	Bionomia	https://bionomia.net
	Crossref	https://www.crossref.org
	ORCiD	https://orcid.org

	Library of Congress	https://loc.gov
	PubMED	https://pubmed.ncbi.nlm.nih.gov
	Wikidata	https://www.wikidata.org
Data Repositories	Barcode of Life Data Systems	https://boldsystems.org
	BugGuide	http://bugguide.net
	CalPhotos	https://calphotos.berkeley.edu
	Dryad	https://datadryad.org
	GenBank	https://ncbi.nlm.nih.gov/genbank
	iNaturalist	http://inaturalist.org
	IsoBank	https://isobank.org
	MorphoSource	https://www.morphosource.org
	National Center for Biotechnology	https://www.ncbi.nlm.nih.gov
	Information	https://sketchfab.com
	Sketchfab	https://en.wikipedia.org
	Wikipedia	
Data Aggregators	Biodiversity Information Serving Our	https://bison.usgs.gov
	Nation	https://pnwherbaria.org
	Consortium of Pacific Northwest Herbaria	https://www.gbif.org
	Global Biodiversity Information Facility	https://wiki.ggbn.org
	Global Genome Biodiversity Network	https://www.idigbio.org
	Integrated Digitized Biocollections	https://vertnet.org
	VertNet	https://scan-bugs.org
	SCAN Portal	https://swbiodiversity.org/seinet/
	SEINet Portal	https://vilda.alaska.edu/

	Alaska's Digital Archives	
Interactions	Global Biotic Interactions (GloBI)	https://www.globalbioticinteractions.org

A model for the multidimensional Digital Extended Specimen

Connectivity is a core principle of the Digital Extended Specimen network, which promotes interdisciplinary research into functional traits [37] and biological interactions [38], provides a critical foundation for global conservation efforts [38], and reflects the role that "next-generation collections" [17] play in advancing science and society. From its inception, Arctos as a data platform has been built on the "extended specimen network" concept - that is, linking physical objects to all of their derived data (especially web-accessible digital assets) and to third-party repositories for increased accessibility and discoverability [13, 18, 39].

Arctos achieves its richly annotated data by creating a web of knowledge with deep connections between catalog records and derived or associated data, and by using reliable published resources for globally shared information. Here we illustrate how the extended specimen in Arctos can become multidimensional (Fig 9). An Alpine Chipmunk (*Tamias alpinus*) was collected with a pinworm parasite (*Rauschtineria eutamii*) in Inyo County, California, in 2010, and the two specimens are accessioned in the Museum of Vertebrate Zoology (https://arctos.database.museum/guid/MVZ:Mamm:225308) and Museum of Southwestern Biology (http://arctos.database.museum/guid/MSB:Para:27057), respectively. Each specimen has its own extended data network with URL-based links to GenBank sequences, media, and a shared georeferenced collecting event. Within Arctos, the extended specimen networks are multiplied by several inter-collection connections: (1) the two specimens are explicitly related to each other with a biotic interaction of host and parasite, and these relationships are harvested

by the Global Biotic Interactions platform [40]; (2) they were cited in publications [41, 42] shared across collections; and (3) they share a collecting event with other specimens that may be important to ecological studies of parasites. Lastly, the chipmunk was collected as part of a state-wide effort to resurvey California biodiversity (Arctos project https://arctos.database.museum/project/10000244), placing occurrences for both the mammal and its associated parasite in a research context.

Fig 9. Multidimensional extended specimens in Arctos.

Records in Arctos include multidimensional extended specimens that share primary, secondary, and tertiary components and are directly related to each other. The example shown here includes a chipmunk (MVZ:Mamm:225308) and its parasite (MSB:Para:27057) that were collected and accessioned at separate Arctos institutions. They share host-parasite biotic interactions, the same collecting event, and primary source material such as field notes. Any updates to host or parasite metadata (e.g., identification, locality, date) are reflected and searchable in both records across institutions. Together, the chipmunk and parasite were used in graduate student research producing at least two publications and two dissertations that cited 569 specimens from four Arctos collections. The host is linked to one genomic sequence deposited at the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA), and the parasite is linked dynamically to three GenBank sequences at NCBI. In addition, the chipmunk was one of 10,987 vertebrate specimens and observations collected in the southern Sierra Nevada as part of the Grinnell Resurvey Project [41], which resulted in 19 more publications and 2,215 citations in Arctos.

The Arctos Entity

The premise of the Digital Extended Specimen revolves around an individual specimen or object with a single collecting event and links to its derivatives such as gene sequences, CT scans, isotope data, and media (Fig 9). However, the reality of many collections may be more complicated. An individual specimen may be split or composed of separate components, each of which may have different collecting events, preparators, preservation types, and other metadata. Arctos community discussion on how to address this challenge led to the creation of a network-wide “Entity” collection that acts to combine multiple component records sharing the same organism, object, or event ID into a single dashboard with a unique, shareable URL (Fig 10). These components are linked from the Arctos entity record to their own respective record via individual URLs.

For example, one entity record (https://arctos.database.museum/guid/Arctos:Entity:16) links the cataloged blood sample of a Golden Eagle (*Aquila chrysaetos*) chick banded at its nest in 2014 (MVZ:Bird:193216) with data from a radio transmitter device that tracked the individual’s last known location to Mexico in 2017; that observation was cataloged in Arctos as MVZObs:Bird:4792. Here, the coordinates for the original sampling locality are encumbered to protect the eagle nest. In another example (https://arctos.database.museum/guid/Arctos:Entity:134), a single endangered Mexican wolf (*Canis lupus baileyi*) was monitored through a federal conservation program with regular blood sampling at different places and times (e.g., MSB:Mamm:341613, MSB:Mamm:231704). Once moribund, the entire specimen was preserved and cataloged as MSB:Mamm:341614. Both examples illustrate how the Entity record functions to compile and unite multiple related occurrences or records of a single organism or collection object under one persistent identifier (Arctos base URL combined with the Darwin Core Triplet for the Entity record). The Entity identifier is passed to biodiversity data publishers via the Darwin Core organismID field, thus allowing aggregator portals and users to resolve these different records as the same individual.

The flexibility of this model allows additional samples and observations to be continually linked to their Arctos Entity record as more data are collected.

Fig 10. Example of the Arctos Entity model.

The Arctos Entity model links diverse records to a single unifying record for increased discoverability. Here, Record A is a bird observation with an associated blood sample, Record B is a second observation of the same individual with associated radio telemetry data, Record C is the vouchered specimen, and Records D and E are endoparasites and ectoparasites, respectively, taken from the specimen.

An educational resource

From its beginnings, Arctos has spearheaded collection-based inquiries for undergraduate education because of its web accessibility and richly linked data [3]. Students interested in biodiversity, evolutionary dynamics, spatiotemporal variation, cultural heritage, responses to anthropogenic change, and other topics have access to an array of data and tools that can initiate and answer interdisciplinary questions. This is exemplified by educational platforms where Arctos-based modules are posted for open-access class exercises (Table 4). Arctos also is used as a live classroom tool at universities (e.g., “Natural History Museums & Biodiversity Science" at the University of California Berkeley; “Museum Practicum in Advanced Collections Management” at the University of Colorado Boulder; “Mammalogy” at the University of New Mexico) and has been important in training undergraduate and graduate students, post-baccalaureates, and postdoctoral researchers in museum curation and data management [43]. Additionally, collection staff have used Arctos to creatively integrate museum objects with artwork and public engagement in an effort to educate students and the broader community about their collections. For example, the Alabama Museum of Natural History collaborated with

572 the University of Alabama Fashion Archive to host a colorfest on social media that invited the
573 public to interact with museum objects in art projects (#ColorOurCollections,
574 http://library.nyam.org/colorourcollections; see Arctos project
575 https://arctos.database.museum/project/10003310). In another art-based public exhibit,
576 students, volunteers, and researchers spent a semester at the University of Wyoming Museum
577 of Vertebrates creating original art pieces inspired by natural history objects, which were then
578 displayed in a public show at the Berry Biodiversity Conservation Center (Art Inspired by Life,
579 see Arctos project https://arctos.database.museum/project/10002855). At the University of
580 Alaska Museum, staff in the Archaeology and Ethnology & History Departments work
581 collaboratively with local, regional, national, and international organizations to highlight cultural
582 items and Alaska Native heritage. As the Arctos network expands, so will its educational role in
583 promoting awareness of the rich legacy and potential of museum collections.
584
585 **Table 4. Examples of educational platforms with Arctos-based modules.**
586 Arctos modules aim to teach undergraduate students about biodiversity databases,
587 biogeography, evolutionary biology, and climate change, among other topics.
588

Sample Educational Topics	Educational Platform
Accessing Biodiversity Data Arctos User Tutorials Bat Species Distributions Biodiversity Assessment Climate Change	Advancing Integration of Museums into Undergraduate Programs (AIM-UP!) http://aimup.unm.edu https://arctos.database.museum/project/10003944
	Biodiversity Literacy in Undergraduate Education (BLUE)

Geographic Barriers Hemoglobin Function Island Biogeography Phylogenetics Plant Distribution Population Divergence Speciation and Gene Flow	https://biodiversityliteracy.com https://arctos.database.museum/project/10003955
	Quantitative Undergraduate Biology Education and Synthesis (QUBES) https://qubeshub.org

589

590 # A sustainable future for Arctos

591

592 To fulfill the responsibilities of managing and ensuring access to our natural and cultural
593 heritage data for current and future generations, museum administrators need to thoughtfully
594 plan for the financial viability and health of their collections. Unfortunately, museum staff are too
595 often overextended with diverse responsibilities and limited financial resources [44]. Another
596 facet of sustainability is the maintenance and development of technical infrastructure. Arctos
597 has a long track record as a model for “next-generation collections” and associated
598 interdisciplinary research that addresses current and future societal challenges [17, 45-47].
599 Arctos’ longevity of nearly three decades may be due in part to its core development principles
600 of standardization, flexibility, interdisciplinarity, and connectivity within a nimble development
601 model for addressing novel needs and information types in response to changing technology,
602 workflows, ethical considerations, and regulations [33, 48-50]. The sustainability and importance
603 of maintaining these networked and interconnected technologies ultimately becomes premised
604 on reliable funding. Despite the vital importance of these fundamental biodiversity digital
605 resources, financial sustainability remains an ongoing community issue [51]. Facing this reality,
606 the Arctos community sought to improve its financial model for the growing consortium of

607 independent and diverse institutions. The most practical solution was fiscal sponsorship by a
608 non-profit organization dedicated to supporting consortia like Arctos. This new business model,
609 implemented in 2022, allows Arctos to follow diverse sources of funding and support including
610 public and private grants, in-kind and volunteer assistance, fees for use, charitable donations,
611 and annual subscription fees [44, 52]. Our overall goal is to use fiscal sponsorship to guarantee
612 the success and sustainability of Arctos to ensure long term benefits to society and the
613 community of biodiversity scientists, cultural heritage stewards, and educators of all kinds.

614 Acknowledgements

615

616 We thank all members of the Arctos Working Group for their unflagging efforts to improve Arctos
617 and keep it an active, functioning, and engaged community and platform. We also thank the
618 generations of undergraduate and graduate students, post-baccalaureates, collection
619 managers, curators, and technicians who perform daily collection tasks using Arctos at member
620 institutions. The following individuals and collaborators have contributed invaluable expertise,
621 perspectives, and support that have helped to enrich and expand Arctos as both a data platform
622 and community: Stan Blum, John Deck, Jonathan Dunnum, Joyce Gross, Steffi Ickert-Bond,
623 Gordon Jarrell, Craig Moritz, Kyndall B.P. Hildebrandt, Barbara Stein, Lam Voong, Cam Webb,
624 John Wieczorek; Global Biotic Interactions (GloBI; Jorrit Poelen), Global Genome Biodiversity
625 Network (GGBN; Katharine Barker), Integrated Digitized Biocollections (iDigBio; Gil Nelson,
626 Deborah Paul, Erica Krimmel), and the Texas Advanced Computing Center (TACC; Chris
627 Jordan). We thank the National Science Foundation for funding specific to the development and
628 sustainability of Arctos (DBI-9630909, DBI-9876837, DEB-9981915, DBI-2034593, DBI-
629 2034568, DBI-2034577), as well as the Robert & Patricia Switzer Foundation for awarding
630 Arctos a Leadership Grant in 2023; additional grants from various sources have funded
631 collection-specific initiatives that resulted in Arctos improvements. Finally, we thank Community

Initiatives, especially Brandy Shah and Rose Cohen Westbrooke, for their guidance and expertise in our transition to fiscal sponsorship.

References

1. Peterson AT, Cicero C, Wieczorek JR. Free and open access to bird specimen data: Why? The Auk 2005;122: 987-990. doi: 10.1093/auk/122.3.987.
2. Sunderland ME. Computerizing natural history collections. Endeavour 2013;37: 150-161. doi: 10.1016/j.endeavour.2013.04.001.
3. Cook JA, Edwards SV, Lacey E, Guralnick RP, Soltis PS, Soltis DE, et al. Natural History Collections as Emerging Resources for Innovative Education. BioScience 2014;64: 725-734. doi: 10.1093/biosci/biu096.
4. Nelson G, Ellis S. The history and impact of digitization and digital data mobilization on biodiversity research. Philos Trans R Soc B. 2019;374: 20170391. doi: 10.1098/rstb.2017.0391
5. Bakker FT, Antonelli A, Clarke JA, Cook JA, Edwards SV, Ericson PGP, et al. The Global Museum: natural history collections and the future of evolutionary science and public education. PeerJ 2020;8: e8225. doi: 10.7717/peerj.8225.
6. Miller SE, Barrow LN, Ehlman SM, Goodheart JA, Greiman SE, Lutz HL, et al. Building natural history collections for the twenty-first century and beyond. BioScience 2020;70: 674-687. doi: 10.1093/biosci/biass069.
7. Hill A, Guralnick R, Smith A, Sallans A, Gillespie R, Denslow M, et al. The notes from nature tool for unlocking biodiversity records from museum records through citizen science. ZooKeys 2012;209: 219–233. doi: 10.3897/zookeys.209.3472.

8. Graham CH, Ferrier S, Huettman F, Moritz C, Peterson AT. New developments in museum-based informatics and applications in biodiversity analysis. Trends Ecol Evol. 2004;19: 497-503. doi: 10.1016/j.tree.2004.07.006.
9. Holmes MW, Hammond TT, Wogan GOU, Walsh RE, LaBarbera K, Wommack EA, et al. Natural history collections as windows on evolutionary processes. Mol Ecol. 2016;25: 864-881. doi: 10.1111/mec.13529.
10. Lacey EA, Hammond TT, Mast A, Guralnick RP, Monfils A, McCormack J, et al. Climate change, collections and the classroom: Using big data to tackle big problems. Evolution: Education and Outreach 2017;10: 2. doi: 10.1186/s12052-017-0065-3.
11. Cicero C, Spencer CL, Bloom DA, Guralnick RP, Koo MS, Otegui J, Russell LA, Wieczorek JR. Biodiversity informatics and data quality on a global scale. In: Webster MS, editor. The extended specimen: Emerging frontiers in collections-based ornithological research. Boca Raton: Studies in Avian Biology no. 50, CRC Press; 2017. pp. 201–218.
12. Webster MS, editor. The extended specimen: Emerging frontiers in collections-based ornithological research. Boca Raton: Studies in Avian Biology no. 50, CRC Press; 2017.
13. Lendemer J, Thiers B, Monfils AK, Zaspel J, Ellwood ER, Bentley A, LeVan K, Bates J, Jennings D, Contreras D, Lagomarsino L, Mabee P, Ford LS, Guralnick RP, Gropp RE, Revelez M, Cobb N, Seltmann K, Aime MC. The extended specimen network: A strategy to enhance US biodiversity collections, promote research and education. BioScience 2020; 70: 23-30. doi: 10.1093/biosci/biz140.
14. Constable H, Guralnick R, Wieczorek J, Spencer C, Peterson AT, The VertNet Steering Committee. VertNet: A new model for biodiversity data sharing. PLoS Biology 2010;8: e1000309. doi: 10.1371/journal.pbio.1000309.
15. Bates JM, Cicero C, Peterson AT, Wieczorek J, Zermoglio PF. Making bird specimen data more consistent, accessible, and informative. Abstract, Joint Meeting of the American Ornithological Society and Canadian Society of Ornithologists. 2021. Available from

https://americanornithology.org/wp-content/uploads/2021/08/AbstractBook_13August2021.pdf.

16. Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, et al. The FAIR guiding principles for scientific data management and stewardship. Scientific data 2016;3: 1-9. doi: 10.1038/sdata.2016.18.
17. Schindel DE, Cook JA. The next generation of natural history collections. PLoS Biol 2018;16: e2006125. doi: 10.1371/journal.pbio.2006125.
18. Hardisty AR, Ellwood ER, Nelson G, Zimkus B, Buschbom J, Addink W, et al. Digital extended specimens: Enabling an extensible network of biodiversity data records as integrated digital objects on the internet. BioScience 2022:72: 978-987. doi: 10.1093/biosci/biac060.
19. Ogle G, Koo MS. Database evolution. Berkeley: Berkeley Natural History Museums [Internet]. 2016. Available from https://bnhm.berkeley.edu/informatics/database-evolution.
20. Wikipedia. Virtual private database. 2021 Sep 5. Available from https://en.wikipedia.org/wiki/Virtual_private_database.
21. Robertson T, Döring M, Guralnick R, Bloom D, Wieczorek J, Braak K, et al. The GBIF Integrated Publishing Toolkit: Facilitating the efficient publishing of biodiversity data on the internet. PLoS ONE 2014;9: e102623. doi: 10.1371/journal.pone.0102623.
22. Viega AK., Saraiva AM, Chapman AD, Morris PJ, Gendreau C, Schigel D, et al. A conceptual framework for quality assessment and management of biodiversity data. PLoS ONE 2017;12 :e0178731. doi: 10.1371/journal.pone.0178731.
23. Simmons JE. Fluid preservation: A comprehensive reference. Lanham: Rowman & Littlefield; 2014.
24. Marquina D, Buczek M, Ronquist F, Łukasik P. The effect of ethanol concentration on the molecular and morphological preservation of insects for biodiversity studies. PeerJ 2021;9: e10799. doi: 10.7717/peerj.10799.

25. Stoddard MC, Yong EH, Akkaynak D, Sheard C, Tobias J, Mahadevan L. Avian egg shape: Form, function and evolution. Science 2017;6344: 1249-1254. doi: 10.1126/science.aaj1945.

26. Hauber ME, Kim CR, Goethe C, Hanley D. Self-referent phenotype matching is a poor predictor of egg rejection by American Robins. J Field Ornithol. 2020;91: 254–262. doi: 10.1111/jofo.12339.

27. Adams, TL, Strganac C, Polcyn MJ, Jacobs LL. High resolution three-dimensional laser-scanning of the type specimen of *Eubrontes* (?) *glenrosensis* Shuler, 1935, from the Comanchean (Lower Cretaceous) of Texas: Implications for digital archiving and preservation. Palaeontol Electronica 2010;13: 1T,11p. Available from https://palaeo-electronica.org/2010_3/226/index.html.

28. Hollinger RE, John E Jr., Jacobs H, Moran-Collins L, Thome C, Zastrow J, et al. Tlingit-Smithsonian collaborations with 3D digitization of cultural objects. Museum Anthropology Review 2013;7. Available from https://scholarworks.iu.edu/journals/index.php/mar/article/view/2173.

29. Brecko J, Mathys A. Handbook of best practice and standards for 2D+ and 3D imaging of natural history collections. Eur J Taxon. 2020;623. doi: 10.5852/ejt.2020.623.

30. Dubay SG, Fuldner CC. Bird specimens track 135 years of atmospheric black carbon and environmental policy. Proc Natl Acad Sci USA 2017.;114: 11321-11326. doi: 10.1073/pnas.1710239114.

31. Bardua C, Bon M, Fabre AC, Clavel J, Das K, Herrel A, et al. Size, microhabitat, and loss of larval feeding drive cranial diversification in frogs. Nat Commun. 2021;12: 2503. doi: 10.1038/s41467-021-22792-y.

32. Boyer D, Gunnell G, Kaufman S, McGeary T. Morphosource: Archiving and sharing 3-D digital specimen data. The Paleontological Society Papers 2016;22: 157-181. doi: 10.1017/scs.2017.13.

733 33. Zimkus B, Ford LS, Morris PJ. The need for permit management within biodiversity
734 collection management systems to digitally track legal compliance documentation and
735 increase transparency about origins and uses. Collection Forum 2021;35: 1-20. doi:
736 10.14351/0831-4985-35.1.1.
737 34. McGuire JL, Davis EB. Using the palaeontological record of Microtus to test species
738 distribution models and reveal responses to climate change. J Biogeogr 2013.;40: 1490-
739 1500. doi: 10.1111/jbi.12106.
740 35. Smith AB, Santos MJ, Koo MS, Rowe KMC, Rowe KC, Patton JL, et al. Evaluation of
741 species distribution models by resampling of sites surveyed a century ago by Joseph
742 Grinnell. Ecography 2013;36: 1017–1031. doi: 10.1111/j.1600-0587.2013.00107.x.
743 36. Wieringa JG, Carstens BC, Gibbs HL. Predicting migration routes for three species of
744 migratory bats using species distribution models. PeerJ 2021;9: e11177. doi:
745 10.7717/peerj.11177.
746 37. Guralnick RP, Zermoglio PF, Wieczorek J, LaFrance R, Bloom D, Russell L. The importance
747 of digitized biocollections as a source of trait data and a new VertNet resource. Database
748 2016;2016: baw158. doi: 10.1093/database/baw158.
749 38. Paton A. Biodiversity informatics and the plant conservation baseline. Trends Plant Sci
750 2009;14: 629-637. doi: 10.1016/j.tplants.2009.08.007.
751 39. Pauli JN, Newsome SD, Cook JA, Harrod C, Steffan SA, Baker CJO, et al. Opinion: Why we
752 need a centralized repository for isotopic data. Proc Natl Acad Sci USA 2017;114: 2997-
753 3001. doi: 10.1073/pnas.1701742114.
754 40. Poelen J, Simons J, Mungall C. Global biotic interactions: An open infrastructure to share
755 and analyze species-interaction datasets. Ecol Inform. 2014;24: 148-159. doi:
756 10.1016/j.ecoinf.2014.08.005.

757 41. Bi K, Linderoth T, Singhal S, Vanderpool D, Patton JL, Nielsen R, et al. Temporal genomic
758 contrasts reveal rapid evolutionary responses in an alpine mammal during recent climate
759 change. PLoS Genet 2019;15: e1008119. doi: 10.1371/journal.pgen.1008119.
760 42. Bell KC, Demboski JR, Cook JA. Sympatric parasites have similar host-associated, but
761 asynchronous, patterns of diversification. Am Nat 2018;192: E106-E119.
762 43. Hiller AE, Cicero C, Able MJ, Barclay TLW, Spencer CL, Koo MS, et al. Mutualism in
763 museums: A model for engaging undergraduates in biodiversity science. PLoS Biol 2017;15:
764 e2003318. doi: 10.1371/journal.pbio.2003318.
765 44. National Academies of Sciences, Engineering, and Medicine. Biological collections:
766 Ensuring critical research and education for the 21st century. Washington: The National
767 Academies Press. 2020. doi: 10.17226/25592.
768 45. Colella JP, Bates J, Burneo SF, Camacho MA, Bonilla CC, Constable I, et al. Leveraging
769 natural history biorepositories as a global, decentralized, pathogen surveillance network.
770 PLoS Pathog 2021;17: e1009583. doi: 10.1371/journal.ppat.1009583.
771 46. Hardisty A, Roberts D, and the Biodiversity Informatics Community. A decadal view of
772 biodiversity informatics: challenges and priorities. BMC Ecol. 2013;13: 16. doi:
773 10.1186/1472-6785-13-16.
774 47. Kagan JS. Biodiversity informatics: Challenges and opportunities for applying biodiversity
775 information to management and conservation. Northwest Nat. 2006;87: 80-85. doi:
776 10.1898/1051-1733(2006)87[80:BICAOF]2.0.CO;2.
777 48. Beaman RS, Cellinese N. Mass digitization of scientific collections: New opportunities to
778 transform the use of biological specimens and underwrite biodiversity science. Zookeys
779 2012;209: 7-17. doi: 10.3897/zookeys.209.3313.
780 49. Vermeylen S, Pilcher J. Let artifacts speak: Online museums and intangible cultural
781 heritage. Int J Intang Heritage 2009;4: 49-63. doi: 10.35638/ijih.2009..4.004.

782 50. Kremers H, editor. Digital cultural heritage. Cham: Springer Nature Switzerland. 2020. doi:
783 10.1007/978-3-030-15200-0.
784 51. Lin D, Crabtree J, Dillo I, Downs RR, Edmunds R, Giaretta D, et al. The TRUST Principles
785 for digital repositories. Scientific Data 2020;7: 144. doi: 10.1038/s41597-020-0486-7.
786 52. Johnson N, Druckenmiller ML, Danielsen F, Pulsifer PL. The use of digital platforms for
787 community-based monitoring, BioScience 2021;71: 452–466. doi: 10.1093/biosci/biaa162.

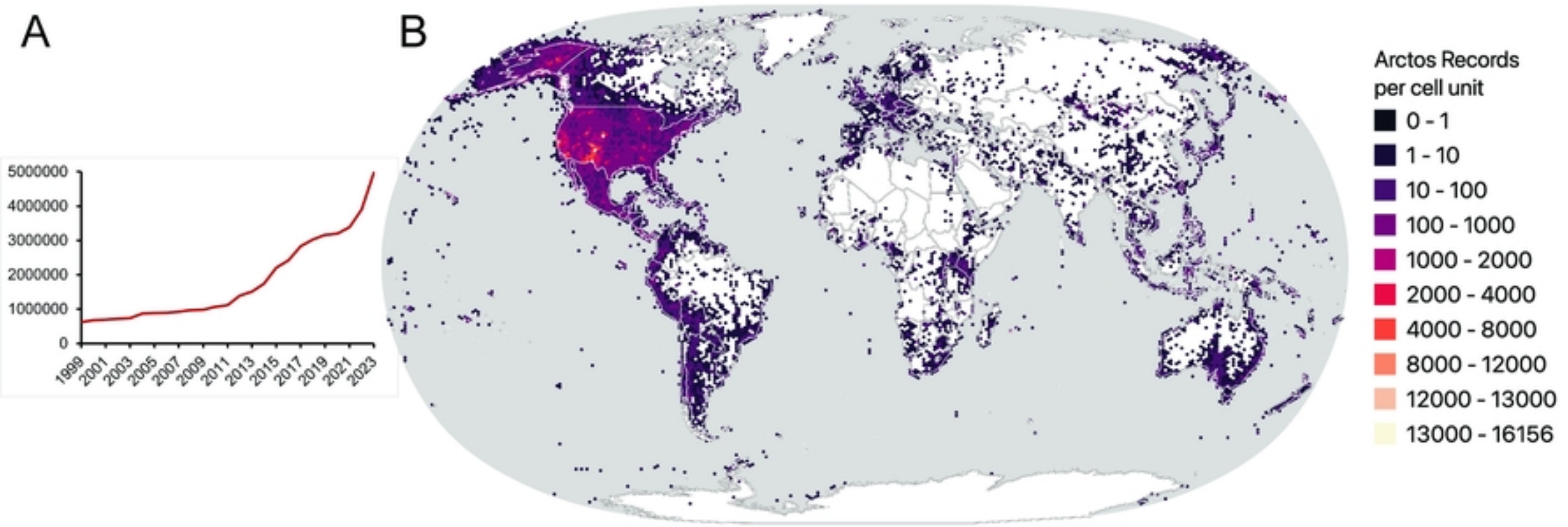
A
5000000
4000000
3000000
2000000
1000000
0
1999
2001
2003
2005
2007
2009
2011
2013
2015
2017
2019
2021
2023
B
Arctos Records per cell unit
0 - 1
1 - 10
10 - 100
100 - 1000
1000 - 2000
2000 - 4000
4000 - 8000
8000 - 12000
12000 - 13000
13000 - 16156

Figure 1

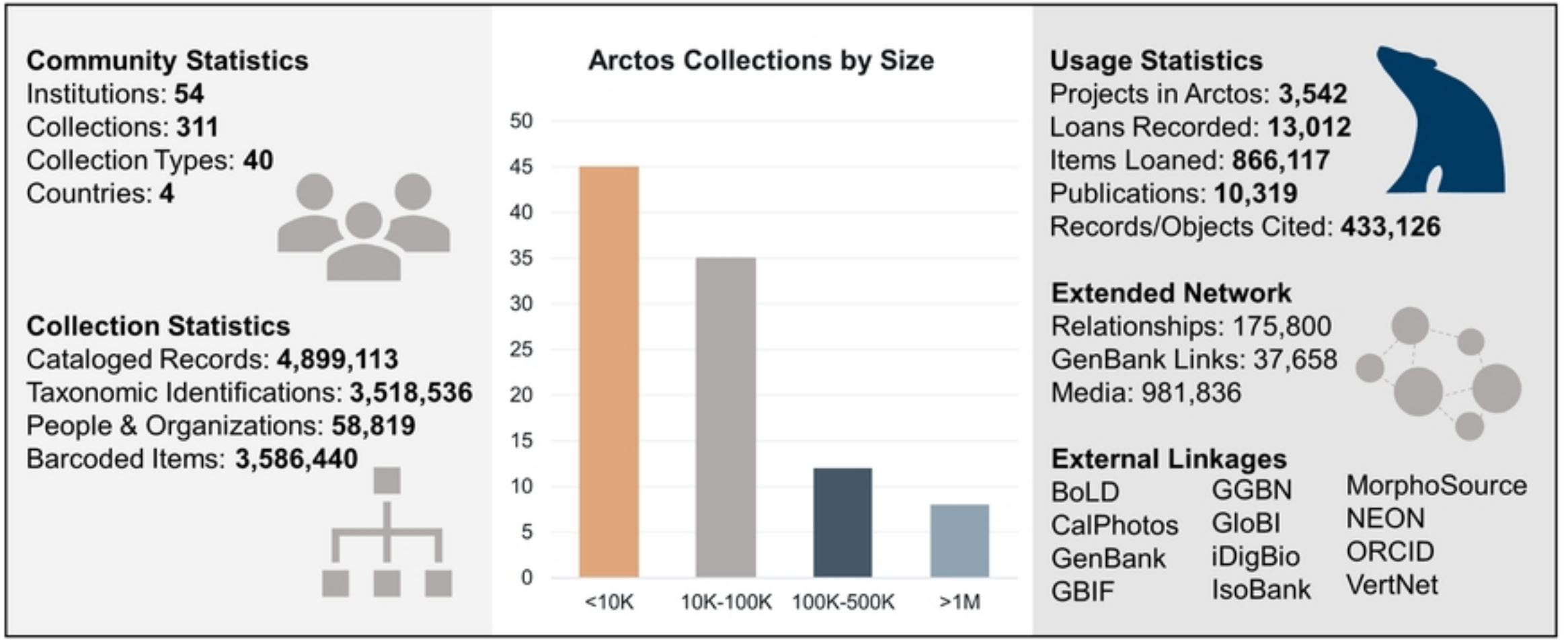
Community Statistics
Institutions: 54
Collections: 311
Collection Types: 40
Countries: 4
Collection Statistics
Cataloged Records: 4,899,113
Taxonomic Identifications: 3,518,536
People & Organizations: 58,819
Barcoded Items: 3,586,440
Arctos Collections by Size
50
45
40
35
30
25
20
15
10
5
0
<10K
10K-100K
100K-500K
>1M
Usage Statistics
Projects in Arctos: 3,542
Loans Recorded: 13,012
Items Loaned: 866,117
Publications: 10,319
Records/Objects Cited: 433,126
Extended Network
Relationships: 175,800
GenBank Links: 37,658
Media: 981,836
External Linkages
BoLD
CalPhotos
GenBank
GBIF
GGBN
GloBI
iDigBio
IsoBank
MorphoSource
NEON
ORCID
VertNet

Figure 2

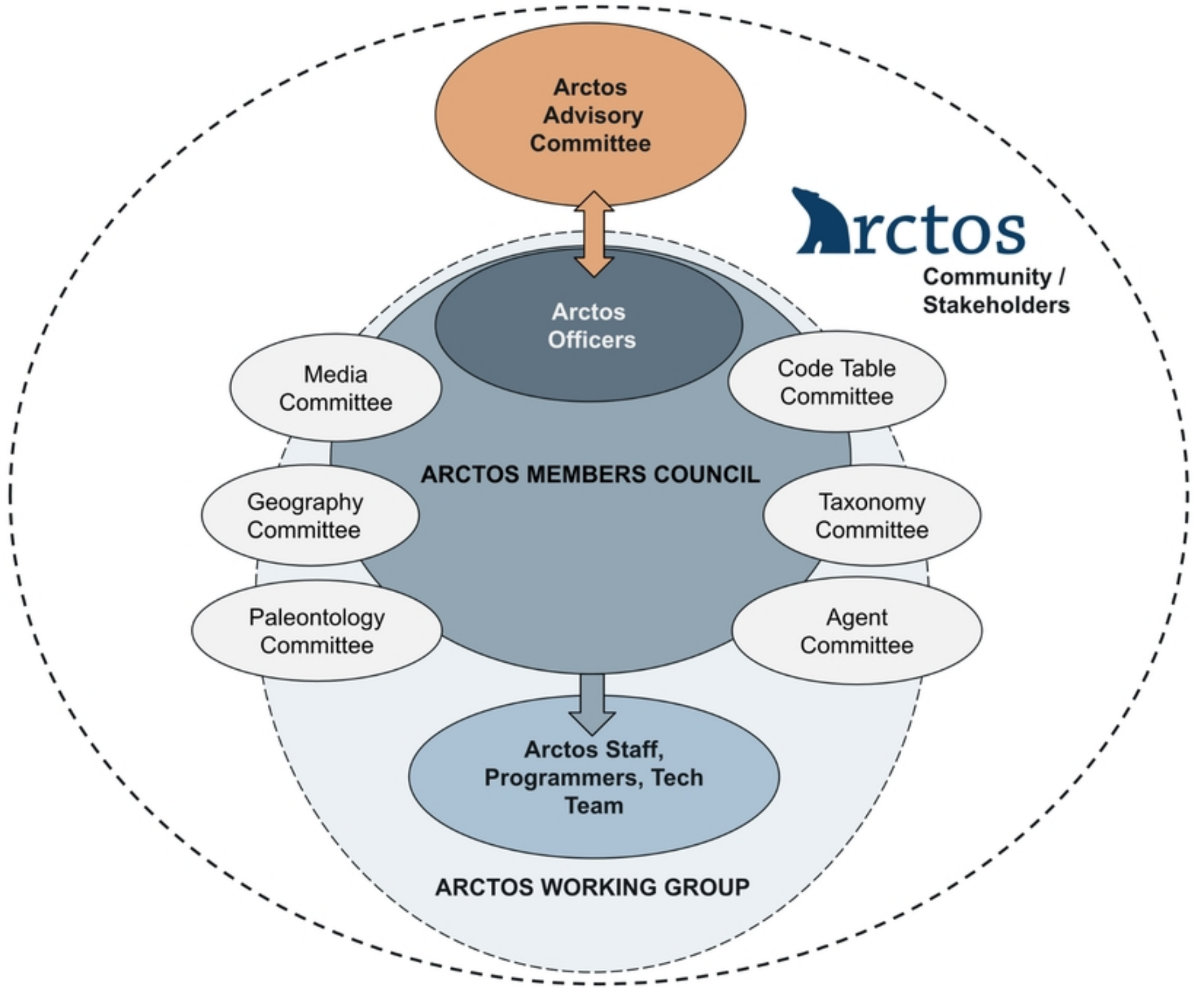

Figure 3

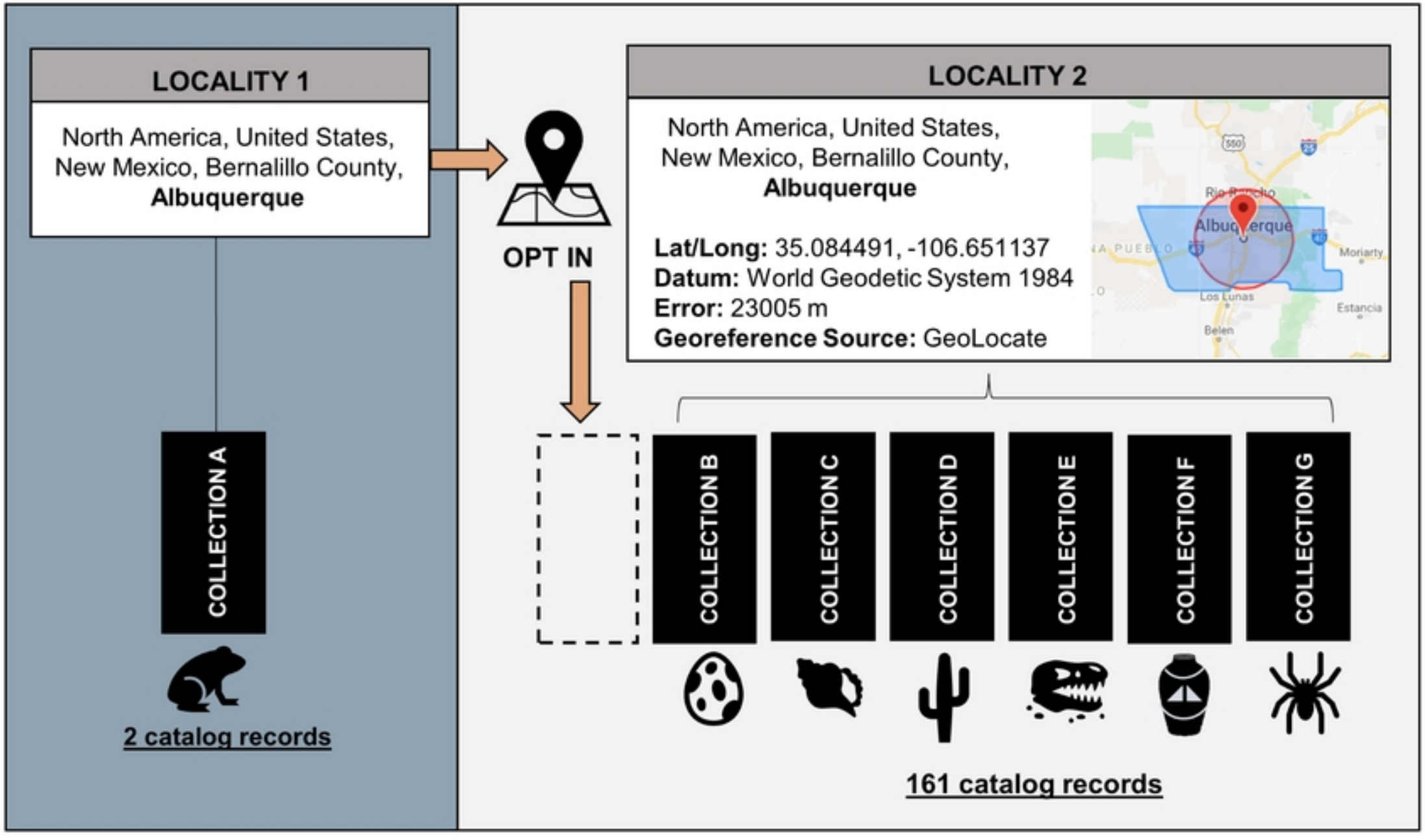
LOCALITY 1
North America, United States,
New Mexico, Bernalillo County,
Albuquerque
OPT IN
LOCALITY 2
North America, United States,
New Mexico, Bernalillo County,
Albuquerque
Lat/Long: 35.084491, -106.651137
Datum: World Geodetic System 1984
Error: 23005 m
Georeference Source: GeoLocate
Albuquerque
Rio Rancho
Moriarty
Los Lunas
Estancia
Belen
COLLECTION A
COLLECTION B
COLLECTION C
COLLECTION D
COLLECTION E
COLLECTION F
COLLECTION G
2 catalog records
161 catalog records

Figure 4

Agent Activity Summary
Name Variations (nicknames/aka, initials, birth names)
Remarks (biographical, curatorial)
Relationships (familial, collegial, academic)
Identifiers and Addresses (contact, URLs, wikidata, ORCIDs)
Collection Activity (e.g., collected, prepared, identified)
Curatorial Activity (e.g., transactions, edits, georeferences)
Products (media, publications, projects)
Project Participation (e.g., grants, expeditions, roles)

Figure 5

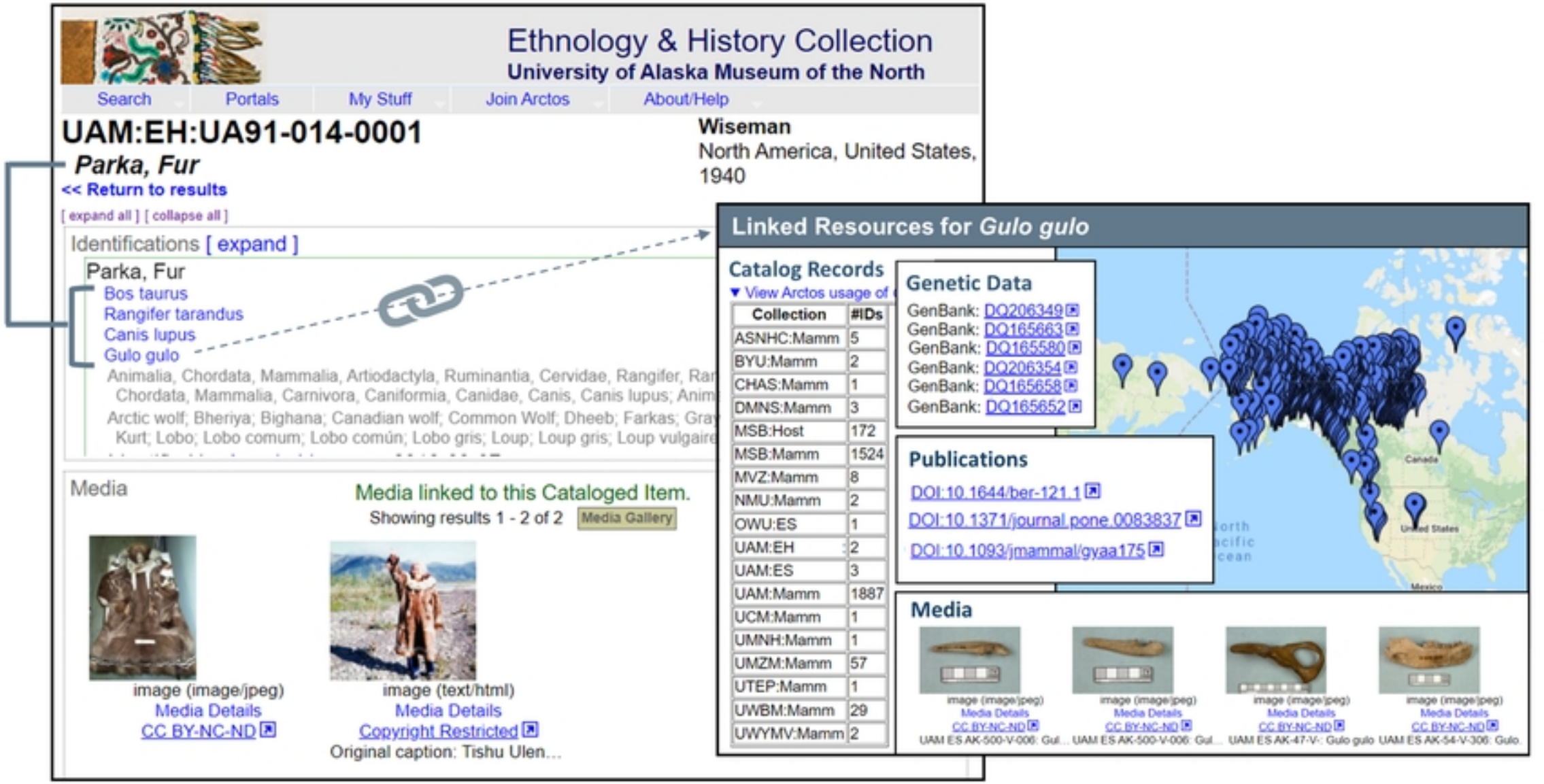
Ethnology & History Collection
University of Alaska Museum of the North
Search
Portals
My Stuff
Join Arctos
About/Help
UAM:EH:UA91-014-0001
Parka, Fur
<< Return to results
[expand all] [collapse all]
Identifications [expand]
Parka, Fur
Bos taurus
Rangifer tarandus
Canis lupus
Gulo gulo
Wiseman
North America, United States,
1940
Media
Media linked to this Cataloged Item.
Showing results 1 - 2 of 2
Media Gallery
image (image/jpeg)
Media Details
CC BY-NC-ND
image (text/html)
Media Details
Copyright Restricted
Original caption: Tishu Ulen...
Linked Resources for Gulo gulo
Catalog Records
View Arctos usage of
Collection
#IDs
ASNHC:Mamm 5
BYU:Mamm 2
CHAS:Mamm 1
DMNS:Mamm 3
MSB:Host 172
MSB:Mamm 1524
MVZ:Mamm 8
NMU:Mamm 2
OWU:ES 1
UAM:EH 2
UAM:ES 3
UAM:Mamm 1887
UCM:Mamm 1
UMNH:Mamm 1
UMZM:Mamm 57
UTEP:Mamm 1
UWBM:Mamm 29
UWYMV:Mamm 2
Genetic Data
GenBank: DQ206349
GenBank: DQ165663
GenBank: DQ165580
GenBank: DQ206354
GenBank: DQ165658
GenBank: DQ165652
Publications
DOI:10.1644/ber-121.1
DOI:10.1371/journal.pone.0083837
DOI:10.1093/jmammal/gyaa175
Media
image (image/jpeg)
Media Details
CC BY-NC-ND
UAM ES AK-500-V-006: Gul...
UAM ES AK-500-V-006: Gul...
UAM ES AK-47-V-: Gulo gulo
UAM ES AK-54-V-306: Gulo.

Figure 6

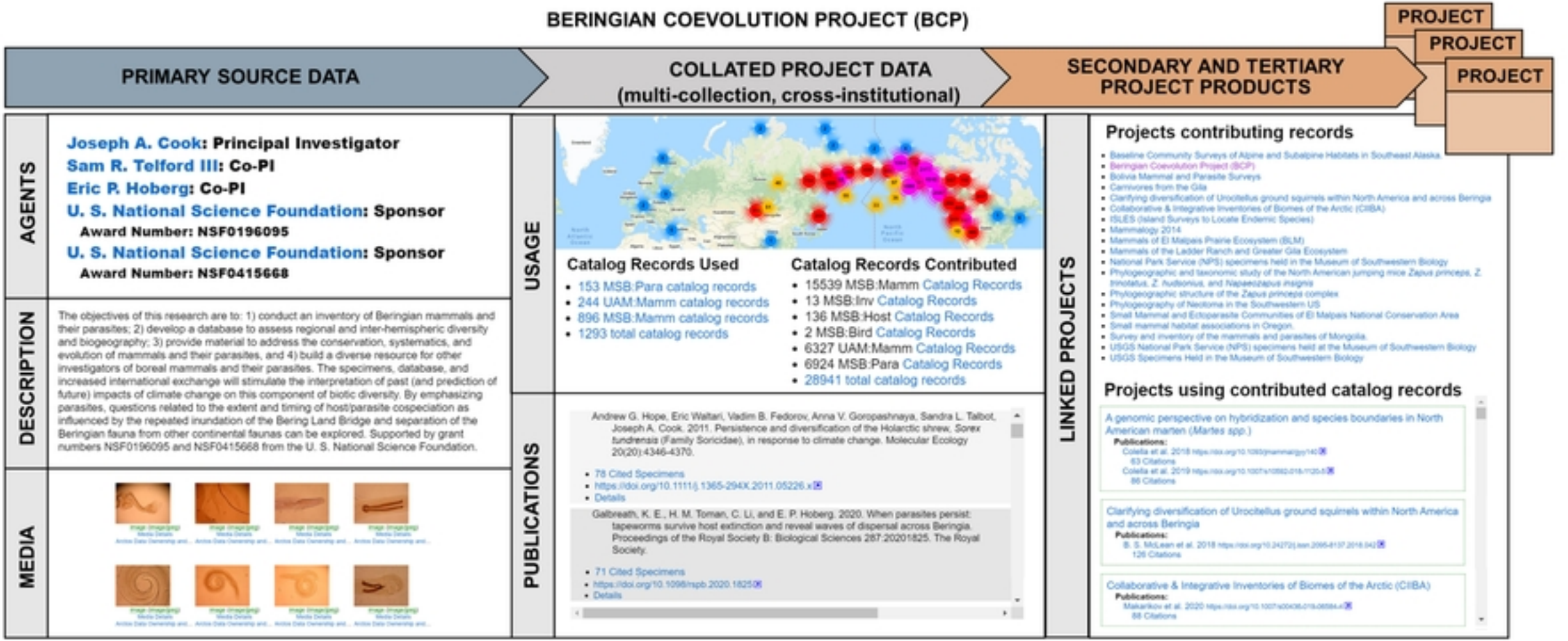
BERINGIAN COEVOLUTION PROJECT (BCP)
PROJECT
PROJECT
PROJECT
PRIMARY SOURCE DATA
COLLATED PROJECT DATA
(multi-collection, cross-institutional)
SECONDARY AND TERTIARY
PROJECT PRODUCTS
AGENTS
Joseph A. Cook: Principal Investigator
Sam R. Telford III: Co-PI
Eric P. Hoberg: Co-PI
U. S. National Science Foundation: Sponsor
Award Number: NSF0196095
U. S. National Science Foundation: Sponsor
Award Number: NSF0415668
DESCRIPTION
The objectives of this research are to: 1) conduct an inventory of Beringian mammals and their parasites; 2) develop a database to assess regional and inter-hemispheric diversity and biogeography; 3) provide material to address the conservation, systematics, and evolution of mammals and their parasites, and 4) build a diverse resource for other investigators of boreal mammals and their parasites. The specimens, database, and increased international exchange will stimulate the interpretation of past (and prediction of future) impacts of climate change on this component of biotic diversity. By emphasizing parasites, questions related to the extent and timing of host/parasite cospeciation as influenced by the repeated inundation of the Bering Land Bridge and separation of the Beringian fauna from other continental faunas can be explored. Supported by grant numbers NSF0196095 and NSF0415668 from the U. S. National Science Foundation.
MEDIA
USAGE
Catalog Records Used
153 MSB:Para catalog records
244 UAM:Mamm catalog records
896 MSB:Mamm catalog records
1293 total catalog records
Catalog Records Contributed
15539 MSB:Mamm Catalog Records
13 MSB:Inv Catalog Records
136 MSB:Host Catalog Records
2 MSB:Bird Catalog Records
6327 UAM:Mamm Catalog Records
6924 MSB:Para Catalog Records
28941 total catalog records
PUBLICATIONS
Andrew G. Hope, Eric Waltari, Vadim B. Fedorov, Anna V. Goropashnaya, Sandra L. Talbot, Joseph A. Cook. 2011. Persistence and diversification of the Holarctic shrew, Sorex tundrensis (Family Soricidae), in response to climate change. Molecular Ecology 20(20):4346-4370.
78 Cited Specimens
https://doi.org/10.1111/j.1365-294X.2011.05226.x
Details
Galbreath, K. E., H. M. Toman, C. Li, and E. P. Hoberg. 2020. When parasites persist: tapeworms survive host extinction and reveal waves of dispersal across Beringia. Proceedings of the Royal Society B: Biological Sciences 287:20201825. The Royal Society.
71 Cited Specimens
https://doi.org/10.1098/rspb.2020.1825
Details
LINKED PROJECTS
Projects contributing records
Baseline Community Surveys of Alpine and Subalpine Habitats in Southeast Alaska.
Beringian Coevolution Project (BCP)
Bolivia Mammal and Parasite Surveys
Carnivores from the Gila
Clarifying diversification of Urocitellus ground squirrels within North America and across Beringia
Collaborative & Integrative Inventories of Biomes of the Arctic (CIIBA)
ISLES (Island Surveys to Locate Endemic Species)
Mammalogy 2014
Mammals of El Malpais Prairie Ecosystem (BLM)
Mammals of the Ladder Ranch and Greater Gila Ecosystem
National Park Service (NPS) specimens held in the Museum of Southwestern Biology
Phylogeographic and taxonomic study of the North American jumping mice Zapus princeps, Z. trinotatus, Z. hudsonius, and Napaeozapus insignis
Phylogeographic structure of the Zapus princeps complex
Phylogeography of Neotoma in the Southwestern US
Small Mammal and Ectoparasite Communities of El Malpais National Conservation Area
Small mammal habitat associations in Oregon.
Survey and inventory of the mammals and parasites of Mongolia.
USGS National Park Service (NPS) specimens held at the Museum of Southwestern Biology
USGS Specimens Held in the Museum of Southwestern Biology
Projects using contributed catalog records
A genomic perspective on hybridization and species boundaries in North American marten (Martes spp.)
Publications:
Coletta et al. 2018
63 Citations
Coletta et al. 2019
86 Citations
Clarifying diversification of Urocitellus ground squirrels within North America and across Beringia
Publications:
B. S. McLean et al. 2018
126 Citations
Collaborative & Integrative Inventories of Biomes of the Arctic (CIIBA)
Publications:
Makarkov et al. 2020
88 Citations

Figure 7

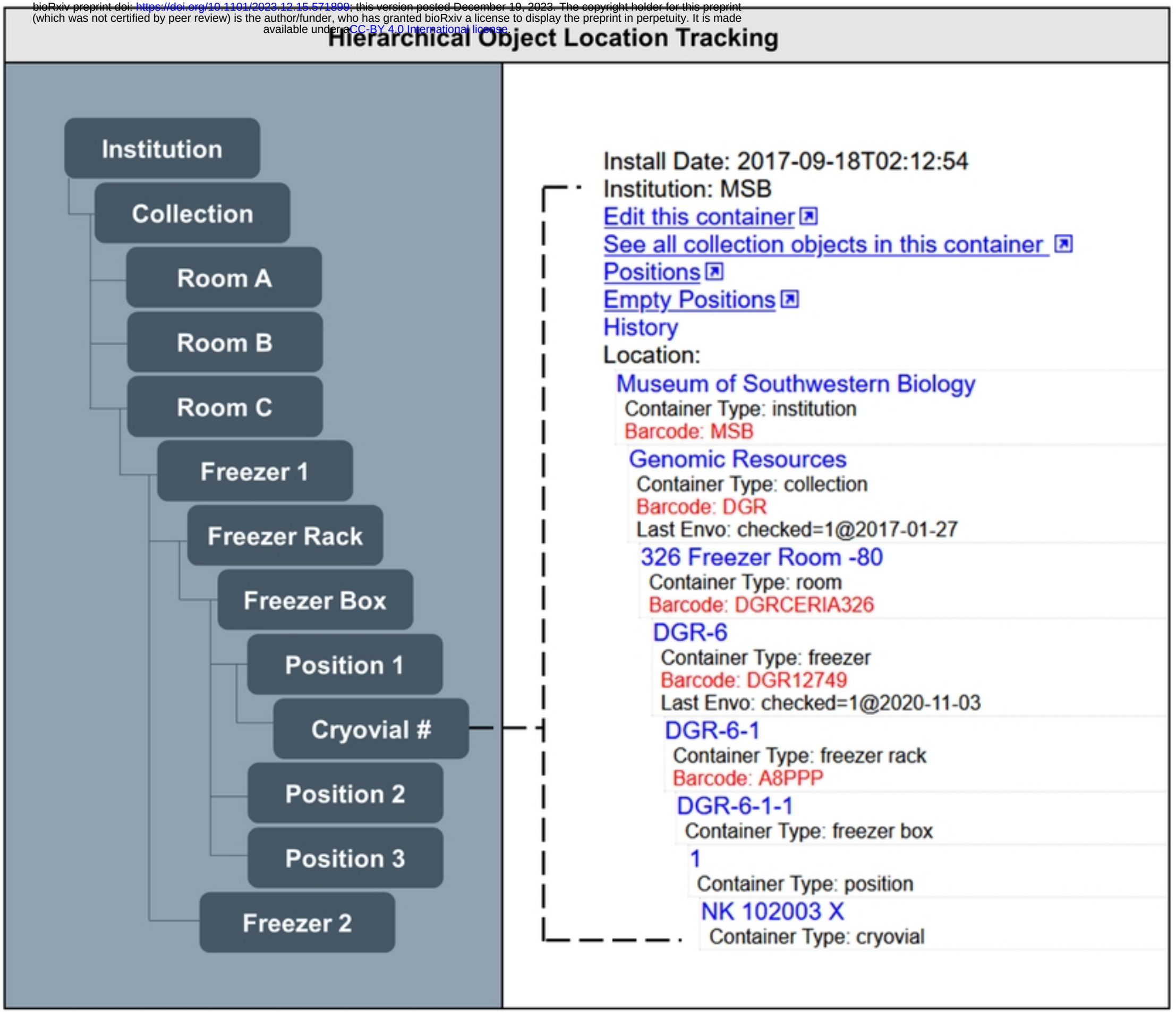

Figure 8

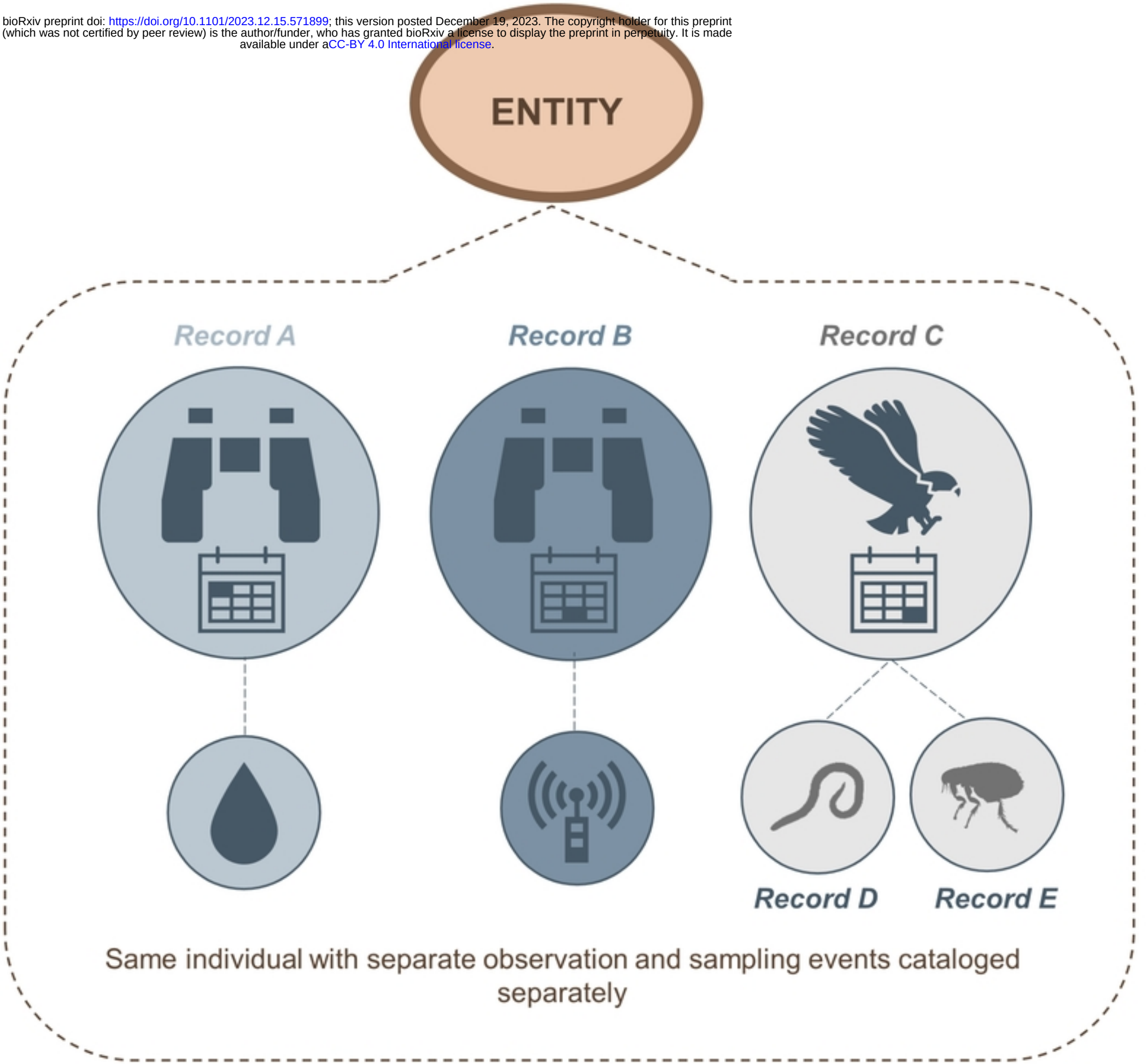

Figure 10

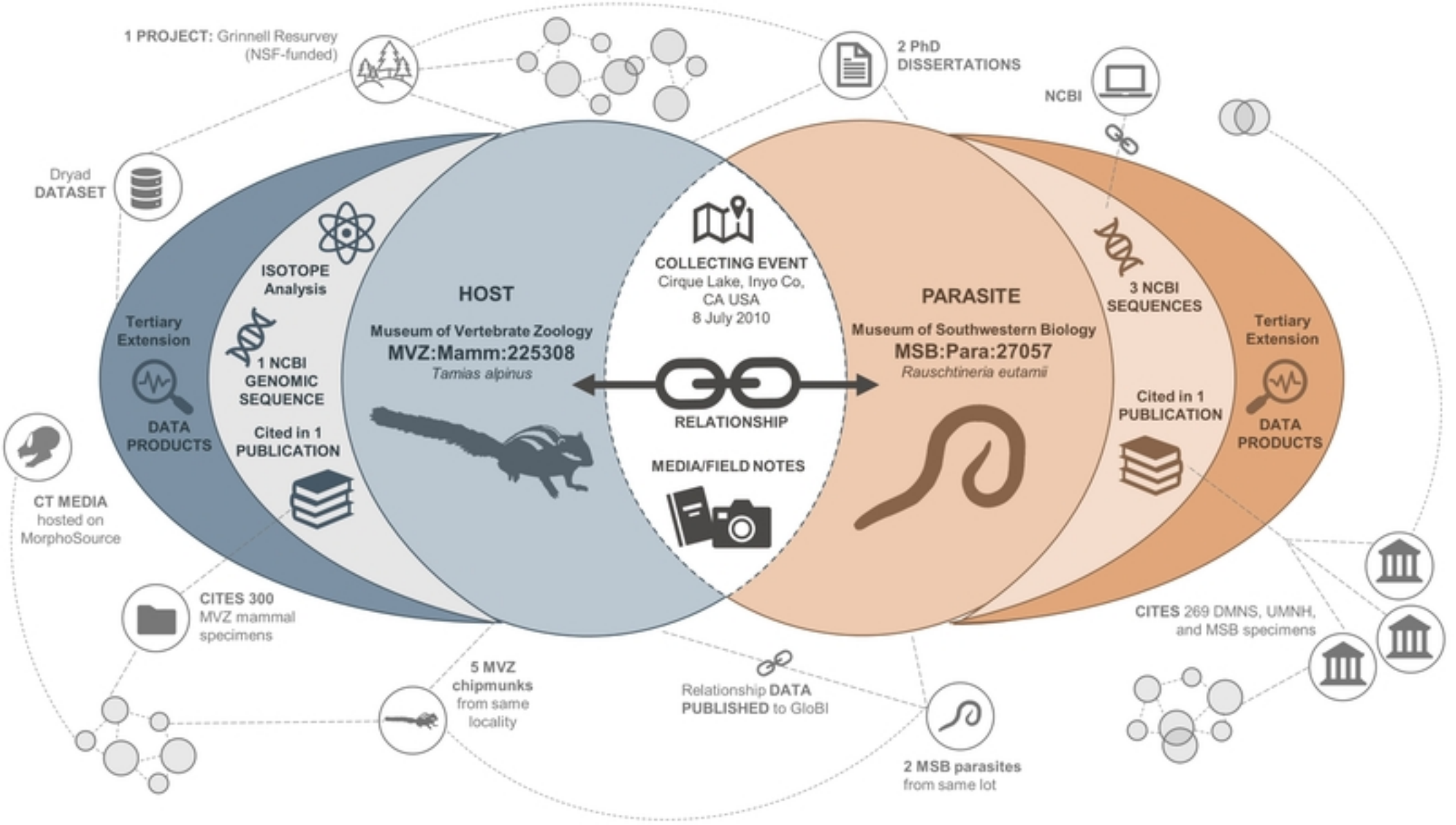
1 PROJECT: Grinnell Resurvey (NSF-funded)
2 PhD DISSERTATIONS
NCBI
Dryad DATASET
Tertiary Extension
DATA PRODUCTS
ISOTOPE Analysis
1 NCBI GENOMIC SEQUENCE
Cited in 1 PUBLICATION
HOST
Museum of Vertebrate Zoology
MVZ:Mamm:225308
Tamias alpinus
COLLECTING EVENT
Cirque Lake, Inyo Co, CA USA
8 July 2010
RELATIONSHIP
MEDIA/FIELD NOTES
PARASITE
Museum of Southwestern Biology
MSB:Para:27057
Rauschtineria eutamii
3 NCBI SEQUENCES
Cited in 1 PUBLICATION
Tertiary Extension
DATA PRODUCTS
CT MEDIA hosted on MorphoSource
CITES 300 MVZ mammal specimens
5 MVZ chipmunks from same locality
Relationship DATA PUBLISHED to GloBI
2 MSB parasites from same lot
CITES 269 DMNS, UMNH, and MSB specimens

Figure 9