



Executive Summary

In development.

Mission: *The “Documenting Marine Biodiversity Through Digitization of Invertebrate Collections” (DigIn) Thematic Collections Network (TCN) improves accessibility and sustainability of marine biodiversity data by digitizing and propagating specimen information (including data capture, georeferencing, and data linkage), by building a community of marine biodiversity experts, and by delivering education and outreach from DigIn products.*

Vision: *A national community of marine biodiversity researchers and collections-based institutions supporting the mobilization, maintenance and growth of marine invertebrate specimen collections and associated databases.*

Strategy: *We leverage the wealth of legacy collections and associated data, the experience of marine biodiversity community members, and the knowledge of the greater digitization community to advance diverse digitization workflows for fluid- and dry-preserved specimens; vastly increase the quantity and quality of accessible marine invertebrate data; and engage an ever-expanding group of researchers, collections professionals, educators, and others in activities that further marine biodiversity research.*

Strategic Planning Process

Initial brainstorming for the collaborative DigIn project began at a University of Florida-hosted workshop in 2019. Many of the initial ideas formulated at the workshop were included in the DigIn proposal to NSF that was funded in 2020 and appear in this Strategic Planning document as well.

The DigIn Steering Group began drafting and editing this document in 2021. The draft was then shared with the full DigIn team for a 30-day comment period, including discussion at a DigIn All Hand’s meeting. Comments from the team were integrated into the document and we have included the names of all who contributed in the Acknowledgements section at the end of the document.

Stakeholders

In a stakeholder analysis we recognized the following groups as valuable to the strategic planning process.

- DigIn Steering Group
- Members of DigIn Working Groups
- Unfunded collaborators including federal institutions such as Smithsonian Institution, and colleagues and associates such as Mary Wicksten

- The full DigIn team of researchers and collections staff, including those still to be hired
- Administrators, colleagues, interns and supporters at DigIn team member's institutions
- Teachers and students at schools near DigIn team member institutions who will become involved in cooperative specimen-based work with DigIn team members
- Staff of iDigBio, the US National Science Foundation's National Resource for Advancing Digitization of Biodiversity Collections
- Other Thematic Collections Networks (TCN) from whom we are learning about project management and organization, digitization strategies, reporting, and other aspects of being a TCN
- The global community of marine biodiversity researchers, taxonomists, and collections managers that have provided invaluable knowledge and expertise
- Biodiversity Information Standards (TDWG), Society for the Preservation of Natural History Collections (SPNHC), and other professional organizations that have created readily-available resources for use in biodiversity and collections-based data mobilization
- Educators, particularly at the undergraduate level, who provide outreach opportunities in formal education settings
- WeDigBio for online, digitization-focused outreach
- Funders, including NSF {and others for certain institutions??}

Environmental Scan

We recognized factors that are, or have the potential to be, helpful (Strengths and Opportunities) or harmful (Weaknesses and Threats). Strengths and Weaknesses are internal to the DigIn project; Opportunities and Threats are external. The factors recognized are current as of the writing of the strategic plan, rather than anticipated with successful implementation of the strategic plan.

Strengths

- Provides an organizational structure and opportunity for community-building among marine biodiversity researchers, institutions, and collections staff
- Develops workflows for the digitization of wet specimens for both field-to-database acquisition of new specimen data and data capture from legacy specimens which include databasing, imaging, georeferencing and/or other elements as required by certain collections and taxa
- Improves nomenclature in collection databases and synchronizes them with WoRMS
- Addresses an identified need to mobilize hundreds of thousands of records of marine invertebrate specimens and make them available via iDigBio
- Improves links between sequence data in GenBank and BOLD and tissue collections in GGBN with collection database records
- Diversity of involved collections and institutions, taxonomic expertise, technical knowledge and general research and digitization skills

Weaknesses

- Lack of a comprehensive pre-existing community of practice for specimen data acquisition, databasing, and propagation among the marine wet collections community
- Lack of standardized specimen information, common databases, or common workflows among institutions and lack of standardized marine localities globally
- Collections are at various points in their digitization workflows, intrinsically require differing approaches across institutions, and have various in-house capacity to advance digitization of their collections

- Differences in institutional support for, e.g., media and promotion, hiring, additional funding, support for collaborations, sustainability of collections and collection databases
- Diverse levels of technical expertise across institutions, in some cases leading to concern about implementing some data processing procedures

Opportunities

- Large amount of legacy data
- Possible sharing of expeditionary and collection data among institutions
- Creating a strong community builds sustainability into digitization efforts within institutions and for global efforts
- Ability to work with other TCNs, institutions and global experts to advance digitization of marine invertebrates
- Advanced stage of prior TCNs provides rich expertise that can be accessed
- Contribute to global knowledgebase of marine invertebrate information
- Increase public awareness of marine research and research collections
- Interchange with IZIG or other organizations to gain and share workflows and other resources

Threats

- (Current) Limited ability for onsite work in collections
- (Current) delays in hiring staff
- Lack of existing protocols and workflows for wet specimen digitization
- Some specimen lots may have already lost their collection data or the data is so minimal that it is no longer interpretable
- Some institutions are not certain of continued ability to persist beyond the program timeframe, possibly leading to “orphaned” collections

Sustainability

Here we address potentially high-impact factors in the list in the previous section.

Maintaining Strengths

- Provide regular opportunities for DigIn members to collaborate, socialize and share knowledge with each other and global marine biodiversity researchers
- Provide continued support for the technological and infrastructural tools, e.g., Zoom, Slack, Google Drive, to maintain ongoing, transparent and efficient communication and productivity
- Support DigIn researchers in publishing results of our nomenclatural, technical and scientific work in a diversity of outlets ranging from blog posts and white papers to peer-reviewed publications.

Addressing Weaknesses

- Continue to expand and establish new opportunities to work from home, work onsite with limited people, and work asynchronously
- Establish workflows and protocols that are modular and can be customized and evolving to meet the needs of individual collections and collecting events, taxa, and institutional needs
- Establish templates and examples by which institutions can share project highlights over social media, popular press, and scientific outlets in light of limited person-power to be able to do this within certain institutions

- Design and deliver training that is appropriate to augment technical skills where needed

Capitalizing on Opportunities

- Host workshops, symposia, and discussions at virtual and in-person conferences such as SPNHC, TDWG, {Western Society of Naturalists, other?} and at our various institutions to share results of DigIn activities, provide opportunities to learn from experts outside of the DigIn circle, and broaden the DigIn network
- Develop community science activities, e.g., Zooniverse-based transcription projects, that engage participants worldwide
- Through organized outreach workshops, create relationships between participating institutions and their neighboring schools that will persist as a nexus for collections-based work assistance and public engagement
- Establish an Expedition database to pool and share cruise data with DigIn members and other global marine biodiversity researchers

Mitigating Threats

- Support efforts to develop digitization workflows and protocols in the DigIn project and make these available to the global community
- Produce content, e.g., annual newsletters and blog posts, describing project activities for broad distribution
- Merge community with existing national and global organizations having strongly overlapping interests (e.g. IZIG [Invertebrate Zoology Interest Group])
- Engage with additional communities, e.g., community science, other (non marine invert) wet collections, that are currently on the fringe of DigIn's network

Goals and Objectives

We recognized the following high-level goals with supporting objectives following prioritization of elements in the preceding Sustainability section. We aspired to make the objectives SMART—specific, measurable, attainable, realistic, and time-bound.

1. **Digitize specimens** of marine invertebrates that are housed in museum and university collections around the country
 - a. Database and make available online 835K lots, representing 7.5 million specimens
 - i. Establish a Google spreadsheet for reporting progress made, to be used by each collection for quarterly reporting. *2021*
 - ii. Establish transparent tracking tools that update with project-wide progress. *2021*
 - iii. Establish workflows to aid institutions in accomplishing digitization goals. *2021*
 - iv. Support institutional responsibilities to maintain collections and/or databases including regularly pushing data to iDigBio and GBIF
 - b. Mobilize an additional 210K lots that are databased but not yet available or readily discoverable online
 - i. Identify the collections and data to be put online and do any final data cleaning. *2021*
 - ii. Work with Cat Chapman and/or other iDigBio staff to get data ingested to the iDigBio portal and GBIF. *2021*

- c. Georeference 175K existing and newly digitized locality records
 - i. Establish georeferencing working group. *2021*
 - ii. Establish procedures for preparing geographic data, submitting data to the Symbiota/CoGeo portal for georeferencing, and repatriating coordinates back to institutional databases
 - iii. Identify cases where georeferencing can be automated, and identify records for which georeferencing will be outsourced. *2021*
 - iv. Georeference existing records. *2022–2023*
 - v. Georeference new records. *2022-2024*
 - vi. ...
 - d. Mobilize or create 464K images of specimens and types
 - i. Create an imaging workflow that is modular and customizable to different collections. *2021*
 - ii. Image specimens. *2022-2024*
 - iii. Mobilize already existing images and associate them with specimen records when applicable. *2022*
2. **Improve DigIn's databases** and our connections to appropriate databases and aggregators
- a. Improve nomenclature in collection databases and synchronize them with WoRMS
 - i. Establish nomenclature working group. *2021*
 - ii. Include nomenclature updates to digitization workflows. *2021*
 - iii. Automate or provide opportunities for regular improvements to nomenclature in DigIn collections. *2022*
 - iv. Synchronize nomenclature with WoRMS by working with WoRMS editors. *2023*
 - b. Improve links between collection database records and sequence data in GenBank and BOLD, as well as molecular-grade tissue records in GGBN
 - i. Identify DigIn collections and records with sequence data. *2022*
 - ii. Improve digitization workflows to include links to GenBank, BOLD, and GGBN where appropriate. *2022*
 - iii. Ensure sequence data from DigIn collections are in GenBank and BOLD. *2023*
 - iv. Ensure linkage exists between collections records and GGBN tissue records, where appropriate. *2023*
3. **Unite the marine collections community** to develop diverse resources on marine biodiversity based on these data
- a. Establish and propagate best practices for field-to-digitization workflows
 - i. Host workshop to develop above specific to marine wet collections. *2023*
 - ii. Host workshop materials on DigIn website and on iDigBio. *2023*
 - b. Host conference symposium highlighting collaborative work of DigIn and broader marine collections community to mobilize data
 - i. Identify conference. *2023*
 - ii. Publish paper or collection of papers based on symposium that reflects advances in marine collections data mobilization. *2023*

- c. Establish an Expedition Database to aggregate marine expedition station data housed in DigIn institutions across the country
 - i. Establish expedition data working group *2021*
 - ii. Identify DigIn institutions with primary expedition/cruise station data. *2021*
 - iii. Work with XX to create a searchable online database to house the station data. *2022*
 - d. Meld US community efforts and international ones
4. Develop tools and resources for **education and outreach**
- a. Establish outreach working group *2021*
 - b. Collaborate with undergraduate educators to develop open and modifiable lessons that build from and add to our digital products
 - i. Work with Cal State University Dominguez Hills (CSUDH) team to deliver a week-long pilot workshop for teachers to develop institutional relationship *2021*
 - ii. Work with CSUDH team to develop pilot workshop. *2021*
 - iii. Begin linking participating institutions with their local schools to identify potential workshop participants
 - iv. Implement pilot workshop with CSUDH. *2022*
 - v. Bring in participating institution professionals and their local teachers for workshops to create sustainable work/engagement relationships at their home locations. *2023–2024*
 - c. Engage the online public via community science projects, e.g., iNaturalist, Zooniverse, and provide in-person opportunities that leverage these platforms, e.g., City Nature Challenge, WeDigBio onsite events
 - i. Identify DigIn institutions capable of and interested in hosting outreach events. *2021*
 - ii. Work with institutions to develop a cross-institution iNaturalist project focused on documenting marine inverts with photos, while also learning about DigIn collections. *2022*
 - iii. Work with institutions to develop Zooniverse projects focused on digitizing specimen data; may include transcription of specimen labels, transcription of cruise logs and/ or annotating images. *2022*
 - iv. Host WeDigBio events based on Zooniverse projects. *2022*

Evaluation

The Strategic Plan will be evaluated twice a year, in January and July, in conjunction with those quarterly reports. At that time we will evaluate successes and shortcomings, update this document, and will consider patterns in how tasks are completed (or not).

Acknowledgements