Documenting Marine Biodiversity through Digitization of Invertebrate Collections

DigIn
Core rationale

More than double the number of digitized specimen lots of non-molluscan marine invertebrates in North America

Reconstruct communities sampled in the past

Biogeographic distributional patterns in time and space

Biodiversity synthesis linking extended specimen data

Nomenclatural standardization, facilitating taxonomic resolution

Unifying the collections community with best practices and training
DigIn participating institutions
Participating Institutions

Academy of Natural Sciences, Philadelphia, PA
American Museum of Natural History, New York, NY
Museum of Comparative Zoology, Harvard University, Cambridge, MA
Virginia Institute of Marine Science, College of William & Mary, Gloucester Point, VA
Yale University Peabody Museum of Natural History, Boston, MA [subaward]

Auburn University, Auburn, AL
Fish and Wildlife Research Institute, St. Petersburg, FL
Harbor Branch Oceanographic Institute, Florida Atlantic University, Fort Pierce, FL
North Carolina Museum of Natural Sciences, Raleigh, NC
Rosenstiel School of Marine & Atmospheric Science, University of Miami, Miami, FL
University of Alabama, Tuscaloosa, AL
University of Florida, Gainesville, FL

Field Museum of Natural History, Chicago, IL [subaward]
University of Colorado, Boulder, CO

Arizona State University, Tempe, AZ [subaward]
Bishop Museum, Honolulu, HI
California Academy of Sciences, San Francisco, CA
Natural History Museum of Los Angeles County, Los Angeles, CA [lead]
Q-Quatics, Laguna, Philippines [subaward]
Santa Barbara Museum of Natural History, Santa Barbara, CA
Scripps Institution of Oceanography, University of California San Diego, CA
Institutional holdings
Key challenge for DigIn: Workflow diversity

Marshalling existing digitized records — diversity of sources and platforms

Expedition station data — cross-collection, cross-institutional

*Shared centralized resource for all collections*

Georeferencing — extension and adaptation of terrestrial approaches

Specimen data

*Data capture from paper/cards — transcription and OCR*

*Linking existing specimen images (minimal new specimen imaging)*

*Linking existing genetic/tissue data*

*Nomenclatural reconciliation (based primarily on WoRMS)*
Multiple workflows to capture data from specimen containers
DigIn broader impacts

STEM educators and curatorial sustainability

*Cal State University Dominguez Hills — pilot with local teachers*

*Develop museum/college partnerships nationwide*

Virtual field experiences — co-created by science leads and educators

*Combining geospatial data, underwater imagery, and specimen data*

*Focus on Channel Islands, Strait of Juan de Fuca, and Florida Keys*

Public engagement in data transcription

*Local and remote participation in specimen data acquisition*
DigIn goals and objectives summary

Digitize and disseminate **835,000 lots**, representing **7,500,000** specimens.
Mobilize and disseminate an additional **210,000 lots** that are digitized.
Georeference **175,000 station records**.
Mobilize or create **464,000 images** of specimens and type specimens.
Reconcile nomenclature across collections.
Link extended specimen data to GenBank, BOLD, and GGBN.
Disseminate through diverse aggregators, including iDigBio.
Develop and disseminate best practices for field-to-digitization workflows.
Co-create lessons with K-16 educators that build from and add to our digital products.
Integrate undergraduate training, public engagement, and online participation.
Unite the marine collections community with a sustainable digitization ecosystem.