OVERVIEW

The national effort to digitize information in biocollections as outlined in the NIBA Strategic Plan has been successfully catalyzed by funding from NSF and activities of iDigBio. Digitization priorities have been defined, best practices have been integrated into workflows, cyberinfrastructure resources including a search portal have been provided, working groups have focused on major challenges, and collaborations with data providers/users have been built. These community-driven activities have led to improved digitization practices, increased involvement in digitization and training, and adoption of informatics tools that improve the efficiency and scalability of digitization in all types of collections. In its 4th year, iDigBio is working with staff in 439 collections in 268 institutions. Communication among stakeholders to increase access to collections data has been established through workshops, webinars, and social media. iDigBio has sponsored attendance of 1,245 unique participants from 511 institutions to 63 workshops that targeted digitization. iDigBio has ingested 448 recordsets containing 28 million records and 5 million images. All data ingested are indexed so that queries and other types of index-based access are supported. Searches can be done through a Web-based graphical interface or through APIs. Search and analytical tools enable users to mine diverse data such as taxonomy, location, images, traits, and vocalizations. With funding for another 5 years, iDigBio will continue its successful strategies with an increasing emphasis on data improvement and use in research and outreach. The national resource created by ADBC will be increasingly recognized as essential for information on biodiversity and digitization. As the scientific and societal benefits of validated collections data are realized, digitization will become a sustained practice in biocollections.

INTELLECTUAL MERIT

Activities in Phase 2 will focus on many of the same strategies as in Phase 1, but with more emphasis on data quality and use, outreach, and sustainability. iDigBio will host workshops on digitization processes, targeting preparation types that are difficult to database or image, on data quality improvement, and on data use. Engaging staff at TCN institutions, as well as at collections where digitization is not taking place or where digitization is occurring but from which records are not being harvested by iDigBio, will continue through personal contacts, workshops, and social media. iDigBio will work to enable cutting-edge research through improved access to data, easier access to research tools and software, and improved data visualization. Research workflows will be developed to illustrate how data can be used in research applications, probing the availability of current tools and pushing current limits of tools and resources to handle biodiversity data. New user-friendly interfaces will be developed to expose new capabilities of iDigBio APIs through new visualizations in the specimen portal, new appliances, or integration with third-party applications software. iDigBio also will address issues such as maximizing data use and value, ensuring appropriate attribution, and extending data schemas to include additional information to meet expectations of users.

Broader Impacts

Key objectives of iDigBio are to foster project awareness in the scientific community, identify stakeholders, develop educational resources related to digitization, and broaden participation of underrepresented groups. Outreach has been accomplished through courses at UF and FSU on digitization-related topics, informational videos available online, and workshops on a variety of topics including digitization workflows, georeferencing, data carpentry, project management, collections management, citizen science, education & outreach, and broadening participation. E&O efforts in Phase 2 will emphasize expanding and coordinating activities across TCNs and others in the collections community through programs that emphasize digitization, biodiversity informatics, and the value of collections data for understanding biodiversity. Training of students will build upon successful elements in Phase 1, including cyber-enabled courses that connect students across institutions. Lifelong learners will be targeted with the goal of promoting participation directly in the national digitization effort through close integration of Citizen Science with digitization activities. iDigBio will continue its efforts to broaden participation in biology by sponsoring career workshops targeting underrepresented groups.