

2015 iDigBio Annual Evaluation Year 4 Impact Evaluation

Summary

The Year 4 evaluation focused on the impacts of the two areas of greatest activity during the first years of the project—digitization workforce training and cyberinfrastructure. While adequate progress has also been made in the areas of education, outreach and research uses of data, these areas depend on a robust specimen data portal to be impactful. To date, work in these areas has largely focused on raising awareness of the existence and potential of the national digitization effort. While over 90% of the community gives iDigBio high marks for raising awareness, it is premature to measure impacts in these areas.

Digitization and workforce training have directly impacted more than 1200 individuals and 500 institutions with an undetermined number of additional individuals indirectly impacted (via train-the-trainer activities, etc.). Post-training evaluations consistently reveal increases in awareness, skill, or knowledge of workshop topics often among 100% of respondents. The vast majority of those who have expressed an opinion believe that the workshops and webinars have moved the digitization forward beyond what would be expected without them, contributed to community building, increased access to expertise, led to new collaborations, and improved the digitization of collections.

With over 17,000 users, the iDigBio website is increasingly viewed as the place to go for digitization resources and to learn about upcoming events and developments related to digitization in the collections community. Two-thirds of those who responded to a community survey report they visit the website three to four times a year, most often to access information about workshops, other digitization resources, and to learn about upcoming events. Further evidence of the website's impact is that over 400 Adobe Connect recordings of workshops and meetings have been viewed nearly 6500 times in total. The community views the website as a valuable resource, noting that much of the information available on the site simply does not exist elsewhere.

As of late May 2015, the iDigBio search portal has ingested 448 recordsets containing a total of 28 million records for 84 million specimens and 5 million images. More than three quarters (78%) of respondents on the 2015 Community Survey reported visiting the iDigBio specimen portal, and the number of survey respondents who have contributed data to the portal has doubled in the past year. (This does not necessarily include those who submit their data through another initiative).

The impact of iDigBio can also be measured via “anticipated” and “unanticipated” outcomes. The most commonly experienced “anticipated outcomes” reported by respondents were (1) increased digitization of collections, (2) increased ability to share data, and (3) increased collaboration among collections. The three most commonly experienced “unanticipated outcomes” were (1) improved collection management, (2) increased quality of specimen data, and (3) increased involvement and numbers of undergraduates working in collections.

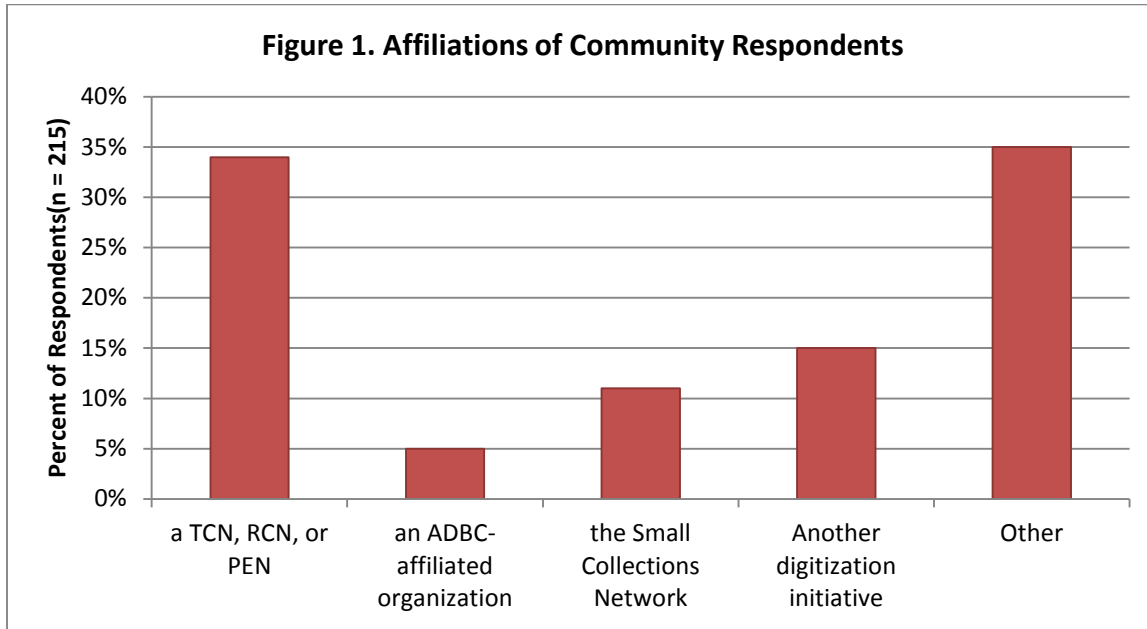
Approach

The goal of the Year 4 Impact Evaluation was to document the impact of iDigBio via a multi-method approach including participant observation, surveys, interviews, and analysis of project records. This evaluation targets the areas of greatest impact to date, those in which the Project Evaluator has been most closely involved, and areas for which impact measures are currently available.

Invitations to participate in the 2015 Annual Community Survey were emailed to over 1300 individuals who have attended iDigBio events, subscribe to the newsletter, are affiliated with a TCN, or are collaborating/partnering with iDigBio in some way. To encourage broad representation, anonymous links to the survey were also provided via the newsletter, Facebook, and Twitter. Nearly 250 individuals responded (including iDigBio team members).

One-third (34%) of respondents (excluding members of iDigBio) identified themselves as current members of a TCN, RCN, or PEN. Respondents who identified themselves as “other” included representatives of NSF Bio Centers, SPNHC, Specify, VertNet, Symbiota/SEINet, Paleobiology Database, National Park Service, wildlife agencies, and

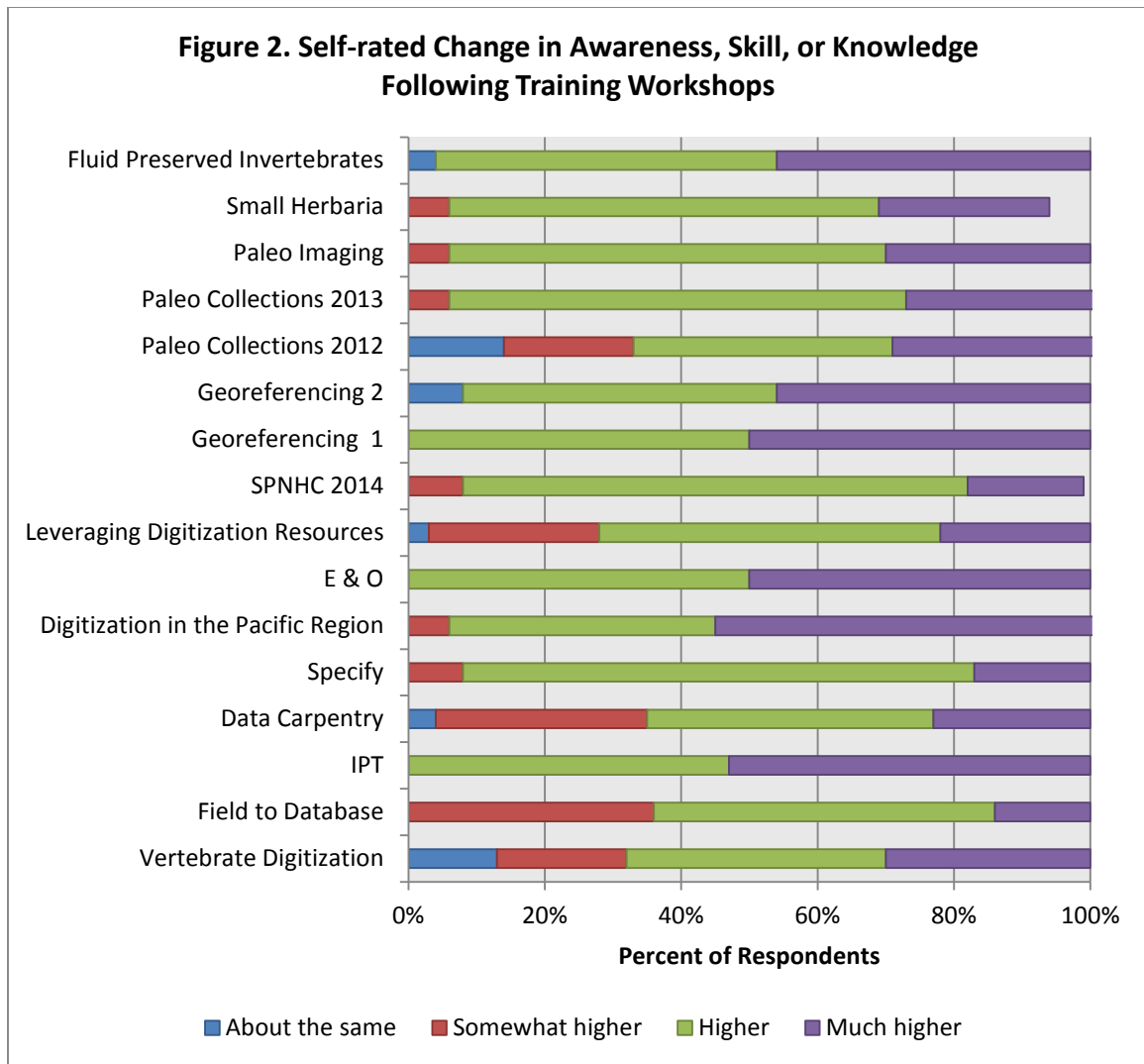
unaffiliated museums and herbaria including several outside of the U.S., as well as university scientists, a high school AP Biology teacher, and members of former TCNs.



Digitization and Workforce Training (Inreach)

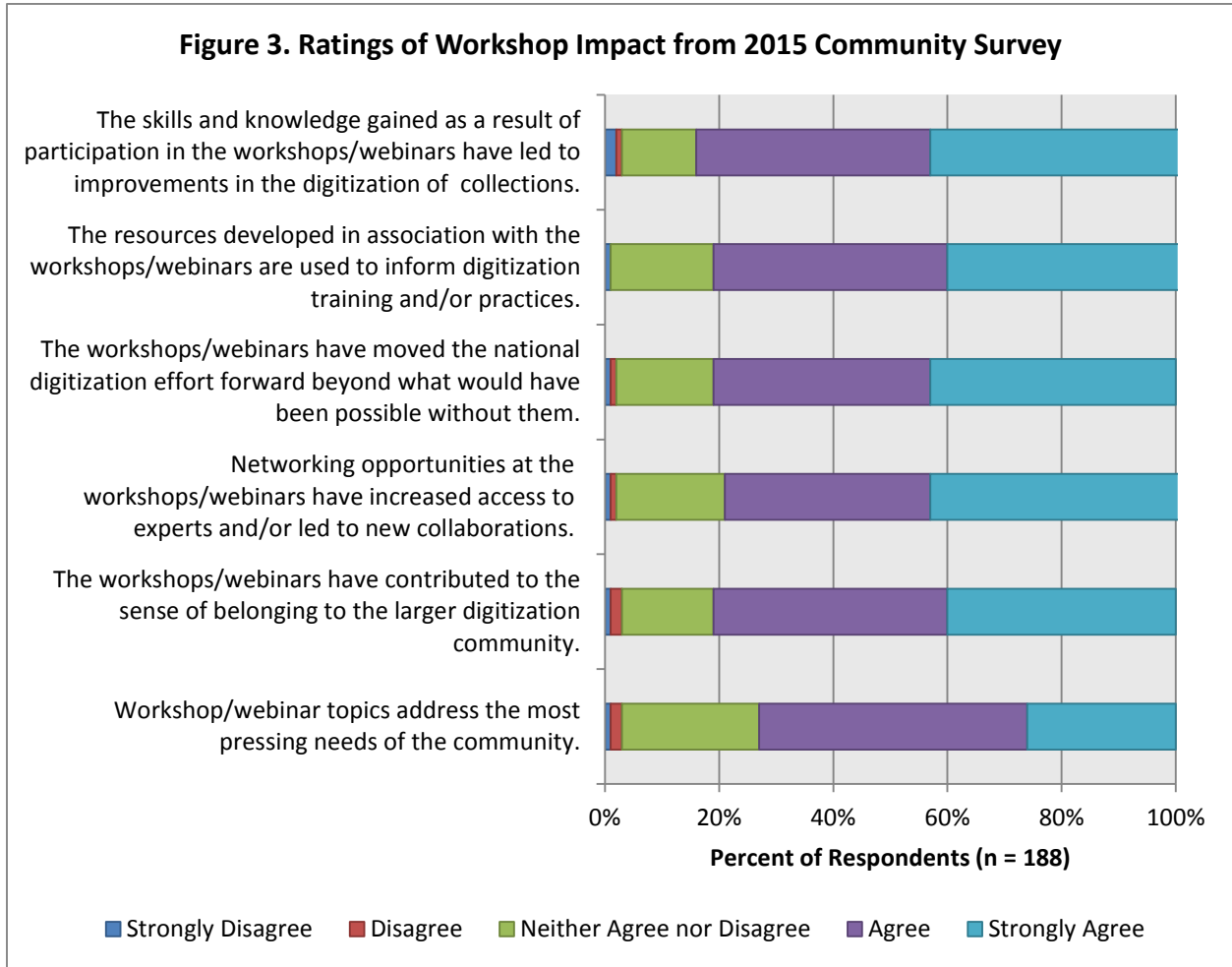
Since its inception, iDigBio has conducted over 60 workshops, symposia, and meetings involving more than 1200 individuals and over 500 institutions (iDigBio refers to educational efforts involving the collections community “inreach.”). The purposes of these events varied: Some were designed primarily to raise awareness about the national collections digitization effort, others involved working groups focused on developing a product (e.g., policies and standards, publications or workshops), some were hackathons, while others provided training for the collections community. It is the last category that is the focus of attention here because (1) post-workshop surveys provide tangible evidence of impact and (2) results of annual surveys indicate that the community (broadly defined) rates the training and outreach efforts of iDigBio as one of its primary strengths and a major contribution to the digitization effort.

Figure 2 shows the responses to one “impact” question asked on the post-event survey following 16 training workshops. (Sample questions: *How does your knowledge of imaging techniques for paleo specimens now compare to that prior to the workshop? Please rate your level of knowledge on how to launch and maintain a collections digitization program following the workshop.*). As shown in Figure 1, the vast majority of participants in iDigBio training workshops report an increase in awareness, skill, or knowledge post-workshop.

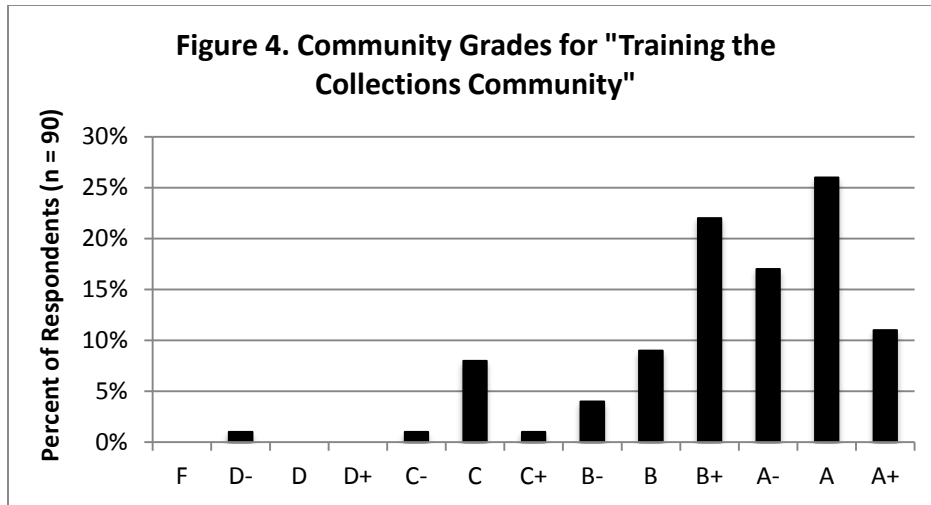


There is not comparable data for every workshop because the post-workshop evaluations of the earliest workshops were primarily focused on assessing needs—that is, topics to be addressed in future workshops—and soliciting feedback on ways to improve the workshop organization, format, delivery, and materials. The results of those surveys (in conjunction with informal feedback) did shape future workshop planning as broadly focused workshops were followed by a series targeting specific collection types, remote access to workshops was instituted, and increasing attention (and time) was devoted to facilitating collaboration and networking, among other changes.

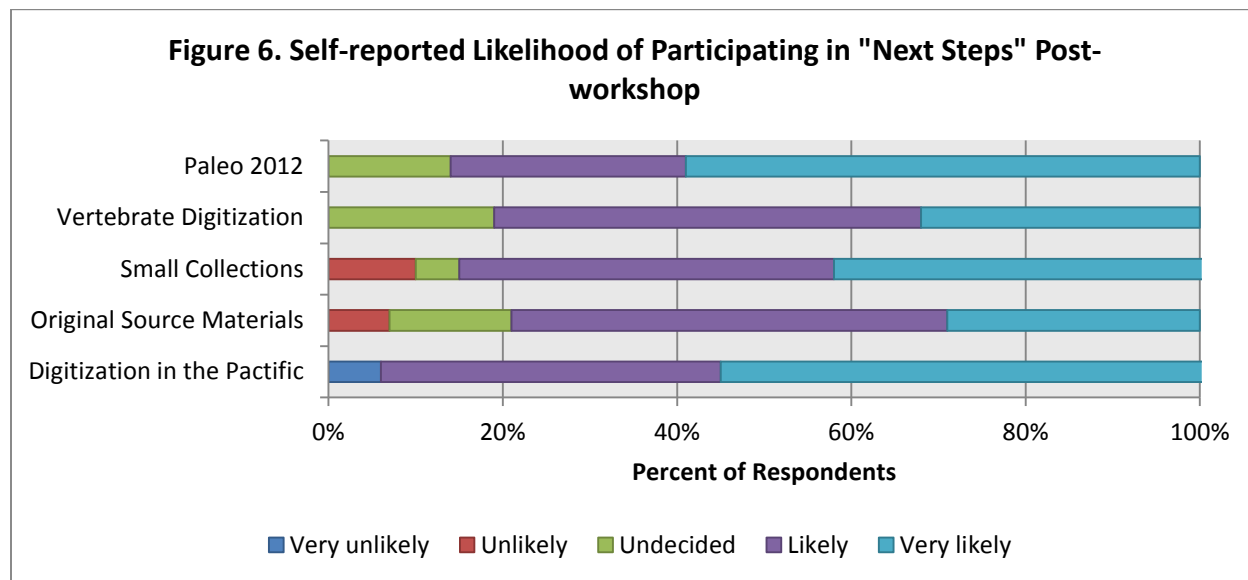
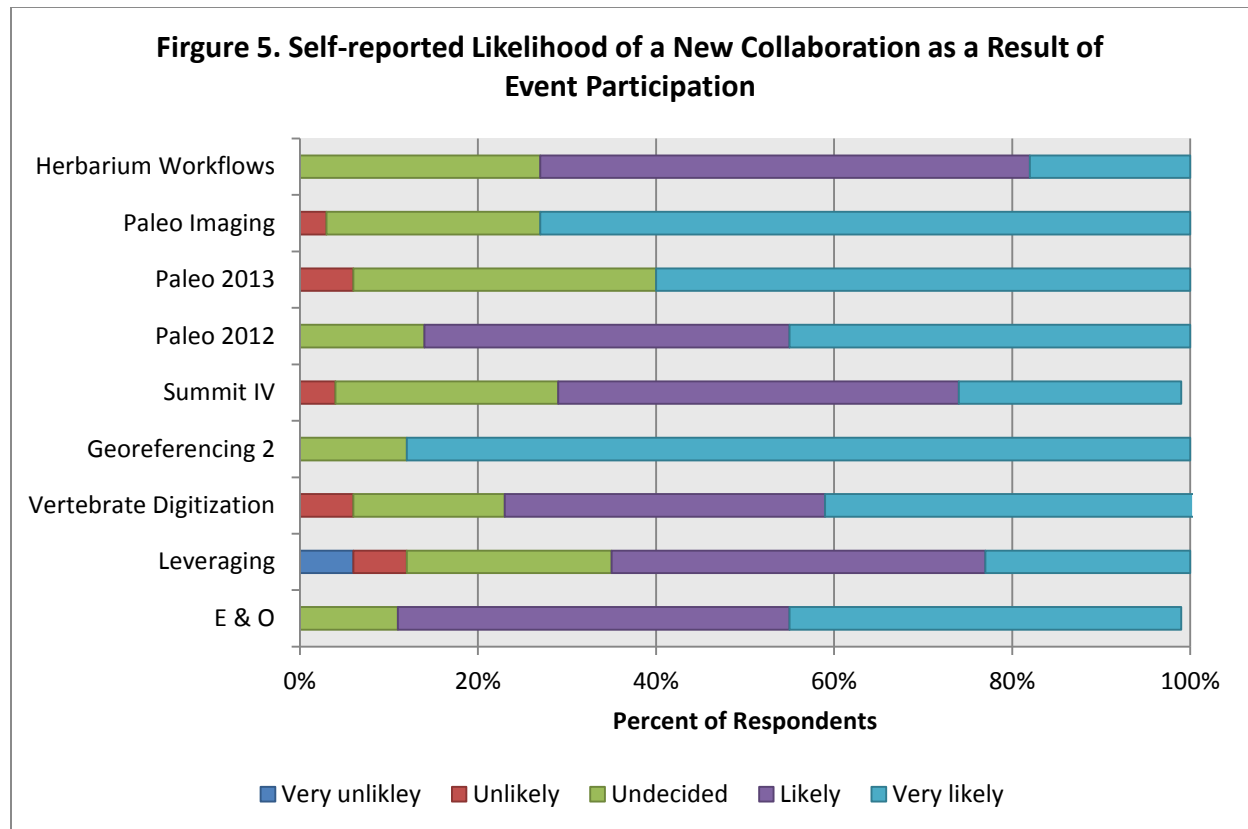
The purpose of providing digitization training for the collections community is, of course, to increase the digitization of collections. Nearly three-quarters of 2015 Community Survey respondents have “personally experienced or observed” an increase in the digitization of collections, while 80% “agree” or “strongly agree” that the workshops and webinars have moved the national digitization effort forward beyond what would have been possible without them (see Figure 3). As one respondent observed, *“I fully support what you have done so far, and appreciate the training that has been made available (staff from my institution have attended a number of workshops and share the info). I think the effort is driving my institution to digitize the collections at a faster rate than would have been done without the federal support. That being said, we are still a couple of years behind in being ready to upload our collections data (e.g. still cleaning up and converting old databases to KE, still digitizing specimen data especially for invertebrate collections, etc.)”*



When asked to grade iDigBio’s effort at training the collections community, 53% of respondents awarded a grade in the “A” range, and 36% in the “B” range. Some of the lower grades may be artifacts of the question structure as several respondents who gave grades lower than “B” indicated that they had “no basis to judge” so simply awarded an average grade of “C” rather than skipping the question (see Figure 4).



A second perhaps unanticipated outcome of the workshops as well as other iDigBio-sponsored events (e.g., hackathons, Summits) was the increase in collaboration and communication reported among participants. As shown in Figures 5 and 6, a majority of respondents to post-workshop surveys report that it is “likely” or “very likely” that they would begin a new collaboration or research network as a result of participating in the workshop and/or participate in any planned “next steps” (e.g., contribute to a publication or wiki, organize a workshop or symposia). A limitation of this data, of course, is that it is based on only those who chose to respond to post-workshop surveys (which averages around 60%) and reflects participants’ best intentions. With the maturation of ADBC and iDigBio, future evaluation efforts should include additional documentation of collaboration such as joint journal articles, conference symposia, workshops, webinars and the like for which information is in the public domain; to date, efforts to obtain follow-up data from workshop participants themselves have not been fruitful. That said, despite the limitations of the post-workshop data, there is converging evidence from the annual community surveys. For example, 80% of respondents to the 2015 community survey “agree” or “strongly agree” that networking opportunities at workshops have increased access to experts and/or led to new collaborations (see Figure 3 above).



A potential negative impact of iDigBio and the associated workshops (or at least drawing attention to the workshops via a community survey) is that a subset of the collections community reports feeling isolated; survey respondents express concern that if they are not part of a TCN, RCN, or PEN, they do not benefit from iDigBio, while others note that lack of funding prevents them from attending professional conferences where they might interact with iDigBio personnel. It is worth noting that the cost of participating in iDigBio workshops has been covered in the past (and remote participation is free); while preference for acceptance to workshops may have

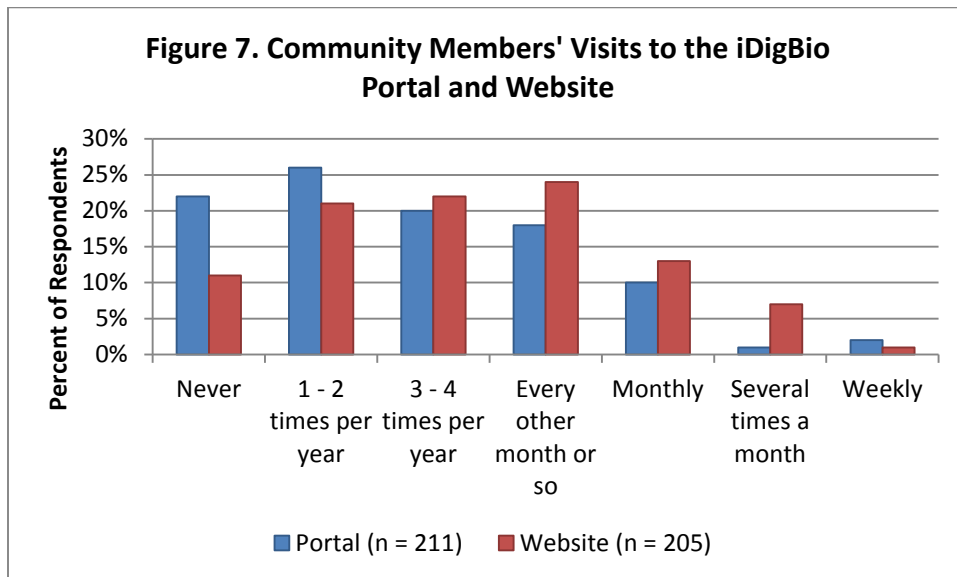
been given to members of TCNs, no survey respondent has commented that they have been denied access to workshops.

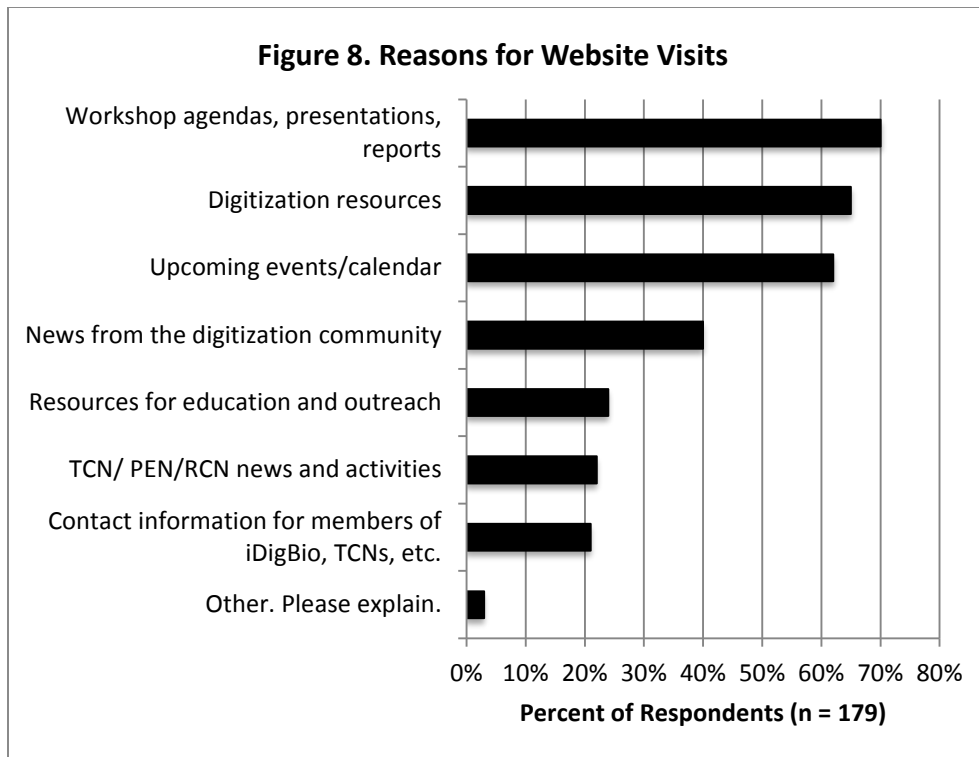
iDigBio has made concentrated efforts to reach out to the entire collections community, particularly those in small collections. A research survey on small collections undertaken by Gil Nelson (iDigBio/FSU) and Anna Monfils (Central Michigan University) in early 2015 led to over 120 individuals to request to be added to the Small Collections Network (SCNet) listserv. In partnership with SCNet, iDigBio has offered 16 no-cost webinars aimed at improving the capacity of those working in small collections.

Website and Portal

One of iDigBio's most impactful accomplishments is the creation of the iDigBio website. With approximately 17,000 users, the website serves as a centralized location for digitization resources and a place to learn about upcoming events and developments related to digitization in the collections community. Technical resources available on the website include data ingestion guidelines, workflows, GUID Guide, reviews of biological collections databases, tool, and data publishing portals among others. The website also provides links to the working groups associated with iDigBio and provides access to meeting minutes, products, and recordings. In addition, the website provides information about both past and upcoming events of interest to the community and is not limited to those sponsored by iDigBio. Those events that are sponsored by iDigBio have associated wikis, which include agendas, presentations, products and documents as well as recordings. As will be apparent below, the website is indeed perceived by many in the community as a highly valuable and unique resource. That said, our annual community surveys consistently reveal that a good percentage of users find the website difficult to navigate. This is especially true for those who visit the website infrequently and who have relatively little technical knowledge. Hopefully, iDigBio can reorganize its website content in the coming years in response to community requests to make it easier to find information.

Two-thirds of community respondents visit the iDigBio at least three to four times per year (see Figure 7). The most popular reasons for visiting the website are to access workshop agendas, presentation, and reports (70%) other digitization resources (65%), and to learn about upcoming events (62%) (see Figure 8). Further evidence of website popularity can be found by examining the number of times workshop and meeting recordings have been viewed. To date, there are over 400 Adobe Connect recordings of events that have been viewed nearly 6500 times in total (range = 1 to 604). Nearly 30% have been viewed 20 or more times.





When asked where they might locate the resources currently available on the website if it no longer existed, only 17% of respondents could identify specific sources and, in some cases, it appears those individuals were thinking about specimen records rather than digitization resources per se. Eleven percent reported they would search online (“random web search”), while 16% would rely on listservs, “word of mouth,” personal contacts, and contacting museum professionals directly. Based on what we have learned about the apparent isolation of many in the collections community, having to reach out to others for assistance may prove a large obstacle for many. Indeed, the positive ratings of the networking opportunities offered at workshops no doubt reflect that participants make personal connections with individuals that they can contact later if needed.

Nineteen percent of respondents reported that they did not think they could find the resources now available on the iDigBio website anywhere else. As one individual observed, “I have no idea. I don’t think I could. One of the most valuable aspects of the website is all the wikis and being able to see past workshop presentations and other such documentation that may never be published but are essential resources.” Another noted that some of the resources would not even exist without iDigBio: “There’s no one place where this information would be available, and without iDigBio none of the workshop and working group resources would exist.”

Fourteen percent of respondents suggested various strategies for finding the information, but cited the amount of additional effort that would be required. Representative comments:

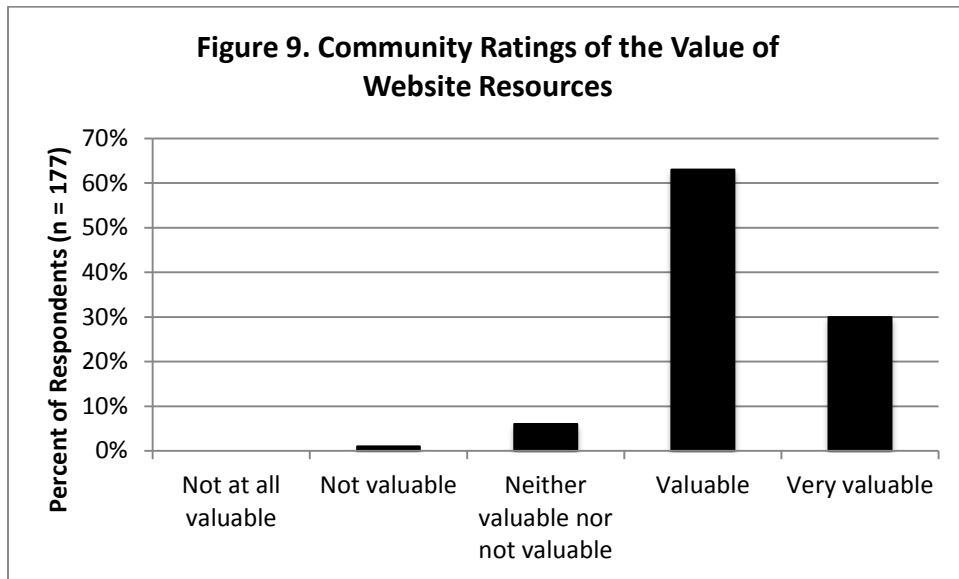
“I think that I would contact other museums to learn how they are doing the digitization process, image, workflows etc., but it would not be that easy, accurate or fast as we can do it using IDigBio website.”

“I would have to search on the web for various digitization topics which would be time consuming and not always find me the results that I need. Also, I would have to contact other institutions to find out how they have implemented their digitization projects. For me, the Documentation/Data Ingestion page on the iDigBio website is very useful.”

“I the case of digitization information, protocols, etc .I probably would NOT find it anywhere, except perhaps through a Google search. No place else has this compiled in this comprehensive way. THIS part has been absolutely irreplaceable as I began digitizing this spring.”

“Some specimen records would come from searching GBIF, but iDigBio is a great help with resources for imaging and data management protocols as well as specimen records. I would probably end up reinventing the wheel if I didn't have iDigBio to connect me with people who may have already done what I need.”

When asked to evaluate the value of the resources on the iDigBio website, over 90% of community respondents rate the resources available on the iDigBio website as either “valuable” or “very valuable” (see Figure 9.) Representative comments:

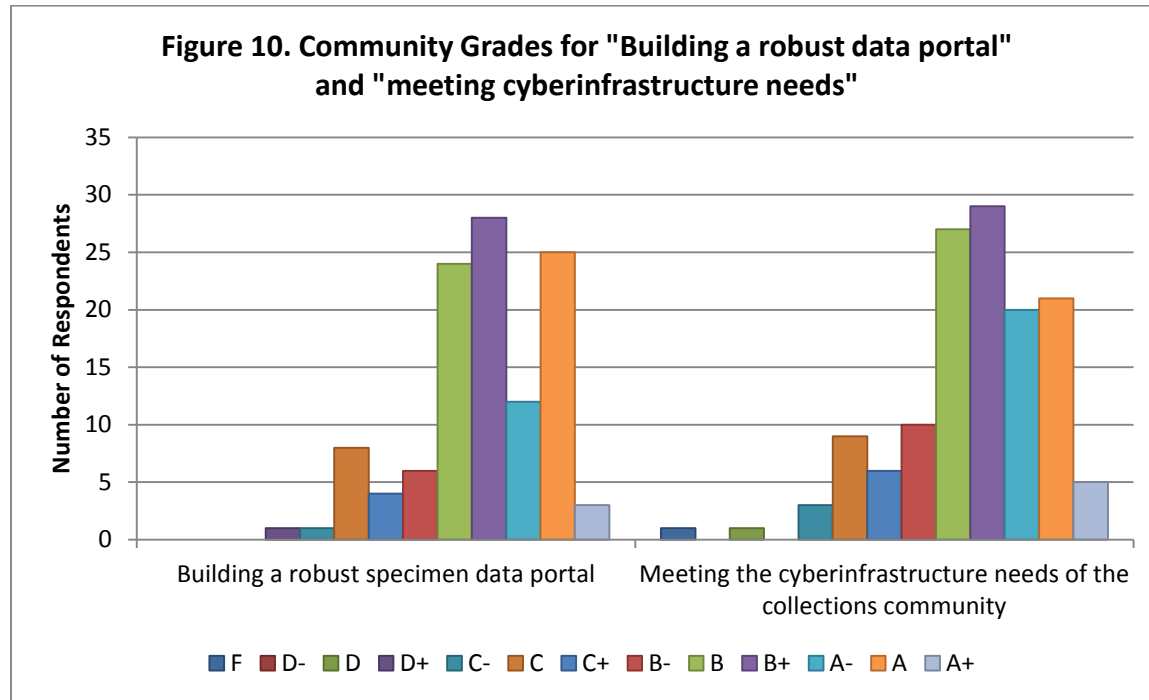


“It is an outstanding resource! It is packed with information, and has frequently saved me from wasting time figuring out solutions to problems already solved by somebody else.” (Curator/university faculty affiliated with another digitization initiative)

“Although I personally have not used the websites and workshops extensively, my staff and graduate students have used them extensively. We have made data from tens of thousands of specimens available and thousands of images. Participation in this program has also led us to adopt Specify as our database, switching from Biota--this has been great!” (University faculty affiliated with a TCN)

As of late May 2015, the iDigBio search portal has ingested 448 recordsets containing a total of 28 million records for 84 million specimens and 5 million images. More than three quarters (78%) of respondents on the 2015 Community Survey reported visiting the iDigBio specimen portal (see Figure 7 above). The number of respondents who reported submitting data to the portal has approximately doubled since 2014 (there were not questions about the portal on the 2013 Community Survey), Nearly half (49%) of those with data reported submitting it to the portal. Of these, only 20% required more than minimal assistance to successfully submit data. Most respondents (72%) who required assistance were “satisfied” or “very satisfied” with the help received. As shown in Figure 10, levels of community satisfaction with the portal and efforts to meet the cyberinfrastructure needs of the community are high. Thirty-six percent of respondents give the portal building effort a grade in the “A” range while 43% give it a grade in the “B” range. Grades for meeting the cyberinfrastructure needs are slightly higher, with 35% of respondents granting a grade in the “A” range and 50% in the “B” range. As noted earlier, the lower grades

should be interpreted with caution as some individuals with little familiarity opted to give average grades of “C” instead of skipping the question.

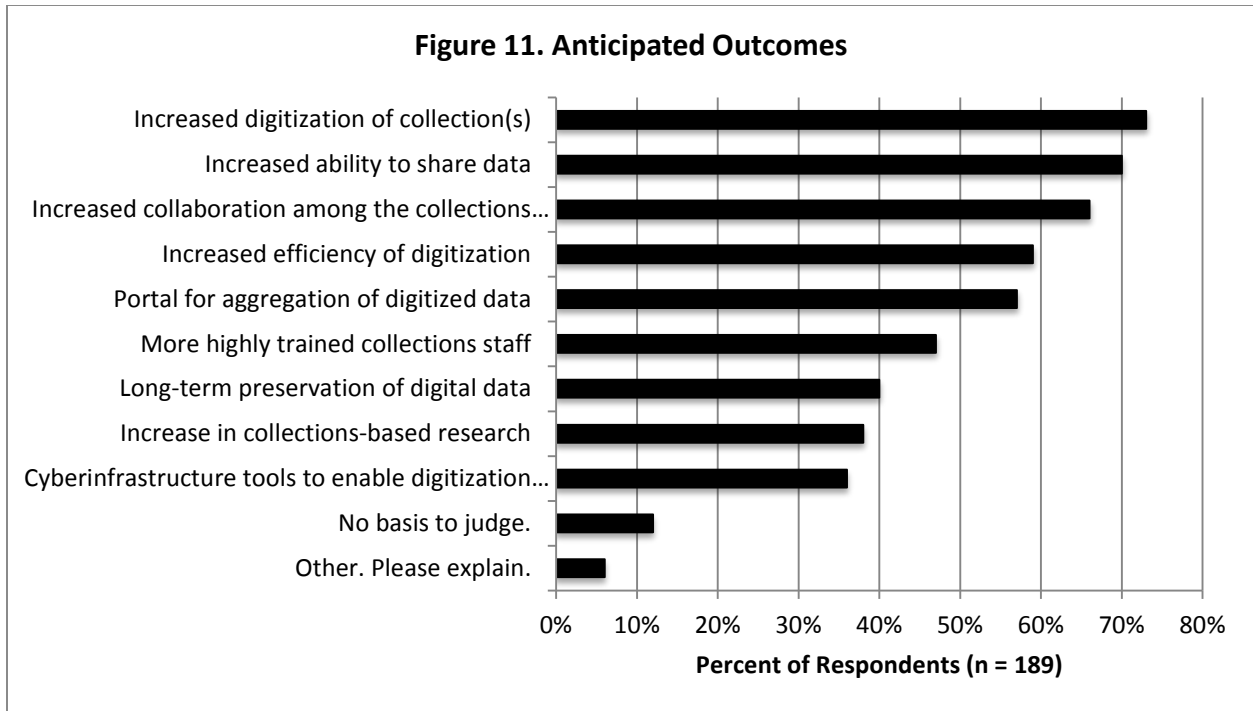


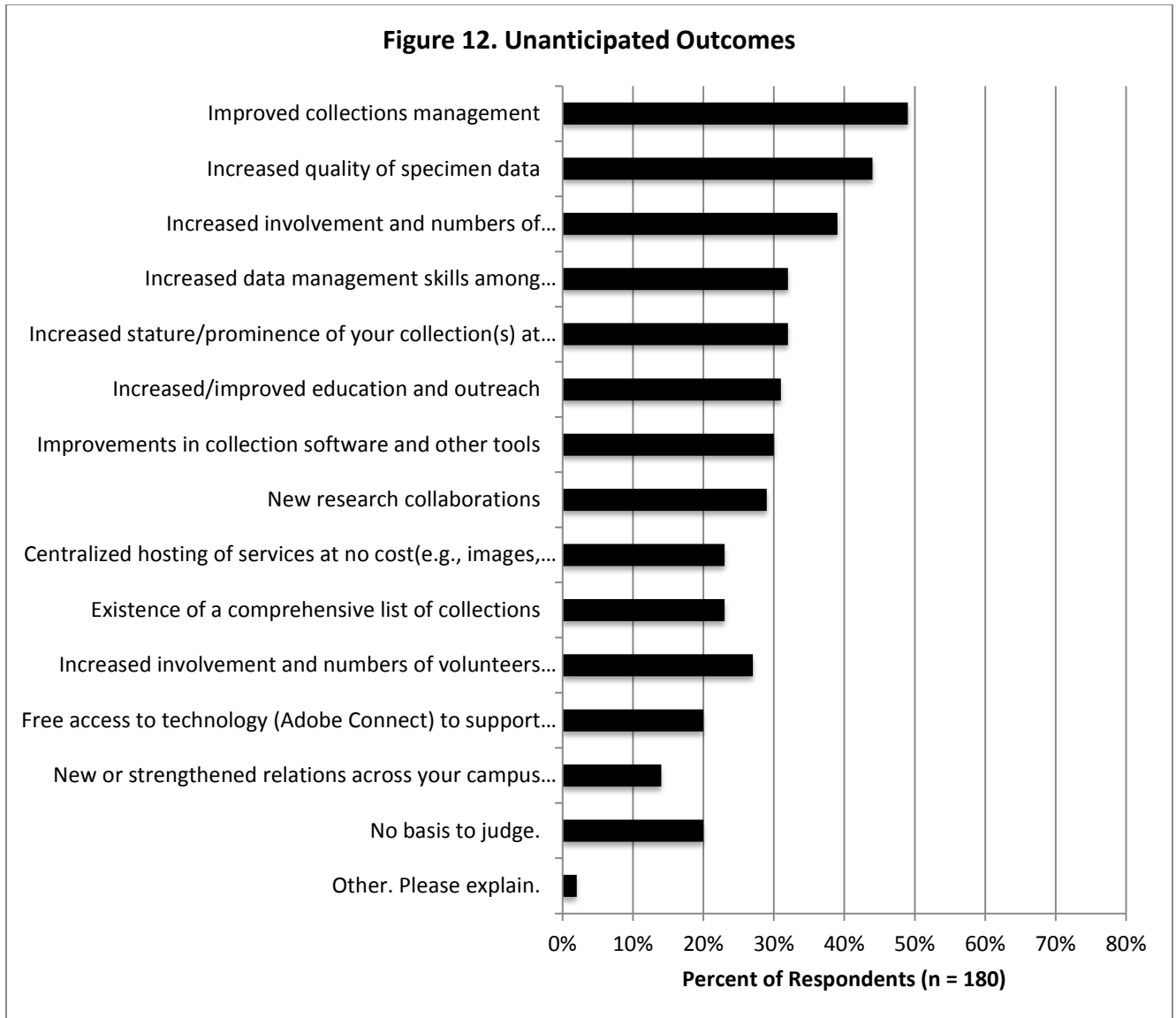
Anticipated and Unanticipated Outcomes

To further evaluate iDigBio impacts, we included questions on the 2015 Community Survey that asked respondents to reflect on outcomes they have personally experienced or observed as a consequence of iDigBio. We divided outcomes into those that should have been “anticipated”—that is, they reflect the goals of ADBC and the national digitization effort and metrics for our own implementation plan—and potential “unanticipated outcomes” based on observations previously shared by the iDigBio team and members of TCNs.

The most commonly experienced “anticipated outcomes” reported by respondents were (1) increased digitization of collections, (2) increased ability to share data, and (3) increased collaboration among collections. Three-quarters or more of those respondents who felt they were in a position to judge reported these outcomes (see Figure 11—note that the figure includes the percentage who responded “no basis to judge.” We included that percentage to provide a broader context for all of the survey results reported here).

The three most commonly experienced “unanticipated outcomes” were (1) improved collection management, (2) increased quality of specimen data, and (3) increased involvement and numbers of undergraduates working in collections. When those who answered “no basis to judge” are eliminated from the analysis, 50% or more reported these three outcomes (see Figure 12). Respondents were encouraged to describe additional outcomes as well.





Of course, not all of the outcomes are perceived as positive. Negative outcomes cited by individual respondents include a decrease in collaboration, especially among larger and smaller collections; the development of mediocre products as the result of developing multiple databasing tools (spreading the funding around); duplication of effort in developing a comprehensive list of collections; and concern about data quality.

Another respondent acknowledged the outcomes, but suggested they are now just beginning to emerge:

“I was being tough regarding the list [anticipated] above. iDigBio has accomplished all of these goals to some extent, but I think we have a ways to go regarding increasing collections based research, ability to share data (and reuse for research), and training of collections staff.”

Additional outcomes, largely unanticipated, are described below:

“Our TCN is starting a list of collection benefits associated with digitization that we did not anticipate: finding types, pre-curation increases curation level, evaluation of overall specimen quality and preservation. Specimen

digitization prompts us to search of old field notes and digitize those as well, reconnect field notes with parts of the collections.”

“I have witnessed first-hand an increase in quality of specimen data, as I sent a couple of corrections to collections managers. I only see that improving with the implementation of tools like Filtered Push.”

“Increase in loan requests and questions about holdings since our data became available on Symbiota. Also able to map species distributions more clearly, leading to better questions about species ecology.”

“My boss has finally gotten the message that through the process of digitization, the condition of the specimens and their organization has been greatly improved. As a result, we have been able to get a bit more institutional funds shunted toward digitization that were originally slated just for curation.”

“Increased stature/prominence (and awareness of) our collection and institution in the community as a result of sharing the digital items through social media and websites.”

“Physical infrastructure for curation also increased as a result of great visibility and activity.”

“I think iDigBio has done an excellent job bringing to the table smaller collections that would never have otherwise participated in the national digitization effort. The resources and training offered by iDigBio are invaluable for such smaller collections, and would not have come from anywhere else.”

“The iDigBio efforts have given me leverage to use internally in my organization to allocate more resources (hardware & salary) toward digitization and online deployment of data.”

“We are developing innovative uses of plant collections to investigate the effects of climate change on plant leafing out times and fruiting times. This would have been more difficult prior to digitization.”

“I think the public portals have influenced the administration here to begin upgrading data systems.”

“It's nice to be able to augment incomplete specimen data with data from what are clearly "dupes" at other institutions.”

“They are learning: a. better / easier management of data inside spreadsheets b. how to use collaborative tools like Google Docs, Google Forms c. how to use web conferencing software like Adobe Connect and Google Hangouts All of these increase both the opportunity for collaboration, but also make it easier to collaborate. And many of these folks self-report they now use these tools (or similar) as they plan conferences / symposia, workshops.”

Overall grades

Finally, as an overall measure of the impact iDigBio is having on the national collections effort, respondents were asked to grade the efforts across a range of goals. Grades for “training the collections community,” “building a robust specimen data portal,” and “meeting the cyberinfrastructure needs of the collections community” were presented in Figures 4 and 10 above. For these goals, the percentage of grade of B- or better were 89%, 79%, and 85%, respectively. The same finding holds for all the goals listed in Figure 13, with at least 82% of respondents awarded a grade of B- or better to each.

Figure 13. Community Grades for iDigBio Efforts

