

## Workshop Introduction 7-12-2012

1. Title slide – begin the workshop by providing some context about the development of a community strategic plan for all collections digitization.
2. After the Office of Science and Technology Policy identified scientific collections as an area of concern and set up a federal interagency working group, an internal group at NSF identified several problems with biodiversity collections and the data that was available within them. We argued that if we did not know what was known, we could not know what was unknown, and that some attention needed to be paid to this so-called ‘dark data’.
3. NSF began an assessment of the collections it had supported with improvement funds over 20 years. This assessment and the assessment of federal collections were proceeding in tandem, but NSF needed to query nonfederal entities, which involved a longer process approval by the Office of Management and Budget.
4. The results of the survey were summarized in a short document made publicly available,
5. The longer and complete results of the survey were also made accessible and the overall report from the Interagency Working Group was published. Several issues were identified in both surveys.
6. Additional issues identified – emphasis on the accessibility and exchange of data.
7. NSF realized that the biodiversity collection community needed to develop a plan for accomplishing this – so some funds were devoted to two focus group meetings and a wiki to involve the broad community.
8. The community did develop a strategic plan for an Online Resource  
  
VISION STATEMENT: This program will develop an inclusive, vibrant, alliance of U.S. biological collections that collectively document the nation’s biodiversity resources and create a dynamic electronic resource that will serve the country’s needs in answering critical questions about the environment, human health, biosecurity, commerce, and the biological sciences.
9. NSF then examined its programs to see how they aligned with this strategic plan.
10. From the two missing components of the plan, NSF set forth a new program under the Division of Emerging Frontiers – the division which incubates innovative approaches across bio.
11. The main goals of digitization were defined and the components that should be included were outlined. Digitization can mean different things to different fields. Thus NSF has defined what we expect from these types of projects as above. We envision the online accessibility to include everything one can find by walking into a collection and picking up the specimen and the associated data gleaned from it.

12. At the moment, the program criteria are defined as shown. The emphasis is still on innovation as the program undergoes development and the research themes are critical to link to the 3 research divisions within BIO and the training and resources are important for linking to the infrastructure activities in BIO.

13. There are certainly new cyberinfrastructure needs which are identified, and we are hopeful that other programs across NSF and elsewhere can be leveraged to provide for many of these activities.

14. Well, there are other challenges as well given the diversity of collections – and I hope you discuss many of these over the course of this workshop as well as gain an appreciation for what a tremendous resource might be created from collections information being made available to the global community.