

About You and your organization

- Please provide your name, title, and your role(s) at your organization .
 - What is your area of expertise?
- Could you please describe your organization in your own words?

Data Life Cycle

- Does the Data Life Cycle figure capture all the steps that data goes through in your organization?
- Are any critical stages missing? If so, please describe.
- What technical functions or tasks are performed at each DLC stage?
- Do you conduct a requirements analysis for any stages of the DLC when designing, building, updating, and maintaining your CI? If so, would you be willing to share any documentation or details about how you conduct the analysis?
- Can you describe how the data flows between stages in your DLC?
- Which DLC stages or cross-cutting elements (e.g., disaster recovery, identity management) do you experience challenges with and might appreciate external expertise/help with?

Your CI

- Please describe the hardware, software, middleware, services, etc. that you are using throughout the DLC.
- For the different components of your CI, are they built in-house, purchased from a vendor, adapted from an open-source tool, or by some other means?
- Please describe any computing resources you use that are offered by governmental and/or academic computing providers (e.g., local campus cluster, XSEDE, Open Science Grid (OSG), National Laboratories)?
- Please describe any cloud services/resources your facility makes use of for compute, storage, web hosting, etc.? What providers do you use (e.g., AWS SWS, AWS Lambda, Google, Microsoft)?

Data

- What types of data does your organization manage (e.g., photos, measurements, physical samples) ?
- What file formats (e.g., csv, jpg, NETCDF) does your organization manage?
- On average, how much data do you manage each day/week/month/year?
- What changes in type, size, and volume do you anticipate over the course of your organization 's lifetime?
- Please describe any data, and accompanying metadata, that you receive from outside sources. From what sources do you receive these data?
- How do you disseminate your data (e.g., APIs/web services, shipped hard drives, data portal, 3rd party repository, other collaboration)?
- What kinds of metadata accompanies the data?
- How do you publish any accompanying metadata? Do you use any ontologies or Google/schema.org to make the data more discoverable?
- To date, how much data do you estimate you have published?

User Community

- Tell us about your user community?
- How many users access your data over a day/week/month/etc.? If you are able to provide numbers for viewing vs. downloading, we'd be interested to know that.
- What are your policies for access and use (e.g., internal users have access to all data we produce/acquire, while external users can only access an aggregate of the data)?

CI CoE

- The CI CoE pilot is proposing to offer the following services and assistance to large facilities:
 - Connecting large facilities with common interests and projects
 - Consulting/advising on a variety of cyberinfrastructure topics (e.g., data management, identity management, disaster recovery)
 - Technology reviews and assessments
 - A knowledgebase of best practices and lessons learned
 - Assistance with workforce development issues
- Would these be useful to your organization and if so, how?

- Are there any other topics or services that you think a Center of Excellence should provide expertise in or offer to large facilities?
- Are there any additional topics that were not included in these questions that you think a Center of Excellence should explore?