

## Chat Question Answered from the Nov CCTS TIERS presentation: 2023 and Me: An overview of the new NIH data management and sharing policy

Link to recording of webinar and slides: [https://mediaspace.uab.edu/media/t/1\\_w1vynzx5](https://mediaspace.uab.edu/media/t/1_w1vynzx5)

Note: The text of some questions has been modified for clarification purposes.

### General questions about the policy

1. Is the DMSP still limited to two pages?
  - a. Yes. the DMSP covering all 6 elements is limited to 2 pages.
2. Does the DMSP impact scoring. Is this aspect subject to change throughout FY23?
  - a. The DMSP does not impact scoring and there is no indication that this aspect of the policy will change during FY23 or beyond.
3. How will this be policed? checking that all this data is available
  - a. Compliance with the DMSP should be documented in the annual RPPR and NIH staff will review this section of the RPPR. According to the [NIH FAQ](#) on the DMS policy: Noncompliance with Plans may result in the NIH ICO adding special Terms and Conditions of Award or terminating the award. If award recipients are not compliant with Plans at the end of the award, noncompliance may be factored into future funding decisions.
4. Do we need to develop a plan for an existing grant (turned in with the next RPPR)
  - a. No, the policy is not retroactive and only applies to awards with submission dates on or after January 25, 2023.
5. Many journals restrict number of references; it will be interesting to see how this helps citations of your work/data
  - a. When reusing data, the citation of the primary datasets is as important as citing the publication it's associated with.
6. It is not our job to make it so that it is "easy" for people to make use of the data plus the software. Is that not correct?
  - a. It is the job of a PI to carry out the DMSP. The DMSP should convey how data will be shared in a FAIR way – findable, accessible, interoperable, and reusable.

### Questions regarding which data to share

1. I collect unique data on nonhuman primates, which encompasses literally terabyte of physiologic telemetry data. We are the only lab in the world that does this, and we have developed fitting and summarization routines over many years using both NIH and UAB resources. Do I have to share all the data, and the software tools we use to filter and summarize them? I should say that it will be exceedingly complex to make our software tools work in a way that can be tied to shared data downloaded by anyone, since it populates our custom database.
  - a. This is a complex issue that would be best addressed with a 1-on-1 consultation with the Research Data Management librarian.

2. Do they want all our flow cytometry data?
  - a. The policy calls for PIs to maximize sharing. For flow cytometry, this could mean sharing data from individual replicates or aggregate data from multiple experiments. The data should be of sufficient quality and quantity to replicate the findings in final figures and tables. To decide which level of data will be shared, follow the best practices for the field and exercise good judgement.
3. What about not sharing because of potential for discovery/patent?
  - a. It may depend on the award mechanism. Under the Small Business Innovation Research (SBIR) and Small Business Technology Transfer (STTR) Program Policy Directive, awardees may withhold applicable data for 20 years after the award date. For other awards, the policy is less clear on embargos on data sharing.
4. If the consent forms allow for sharing of de-identified data from human subjects, does the policy then REQUIRE that data to be shared openly?
  - a. No, the policy does not require that data to be shared openly. The PI controls what level of data is share publicly (if any) and what data will be shared in a controlled fashion, such as after signing a data use agreement.
5. What about [the sharing of] clinical data (dates, demographics, surgical outcomes) - raw database of data.
  - a. The NIH DMS policy applies to clinical data generated or used in NIH-funded research studies. Data must be properly de-identified prior to sharing.
6. Is human sequence data considered PHI?
  - a. Genetic information is health information protected by the Privacy Rule. When sharing human sequence data consider what level of data needs to be shared, for example, full sequence files or just key variants. See [This article](#) for additional information.
7. What about Core Facilities funded through an NIH Center grant? Some generated data is covered by the assisted PI's grant DMP. But what if the PI being assisted is not funded through NIH?
  - a. The core facility will have its own DMSP to follow for their center grant. Speak to the core facility about what data is being shared and how.

#### Questions about when and where to share data

1. I guess my question is where is all this data meant to be shared? It's fine for sequence data, where there are government-run databases, but where do I share raw data for bacterial growth curves or enzyme assays or photographs of test tubes? Is there guidance from the NIH on this?
  - a. The policy encourages use of established data repositories. The PI is free to choose a [data repository](#) based off the NIH guidelines for repository selection. Subject specific repositories are preferable. If there is no appropriate subject repository, a [generalist repository](#) will accept data regardless of data type, format, content or disciplinary focus. Often the submission cost is minimal or zero.
2. Choosing an appropriate repository seems like a key step. Will we be able to set up a meeting with you or someone else knowledgeable to help us figure out the right one to go with for our particular data types?
  - a. Yes. The [Office of Scholarly Communication](#) offers consultations to aid in the selection of an appropriate data repository. Email [mihertz@uab.edu](mailto:mihertz@uab.edu) to set up a consultation.

3. What is the timeline—do the data have to become publicly available within a certain period after generation? after publication? Where is the raw data to be deposited—at NIH managed databases or UAB managed?
  - a. Data must be shared at the time of publication or at the end of the award period. UAB does not have an institutional data repository. Select an appropriate externally-managed data repository for raw data. Here are lists of NIH-supported and generalist [repositories](#). Note, it is not required that the repository be NIH-supported. Before selecting an outside data repository, verify it meets the [desirable characteristics](#) of a data repository.
4. Will IRB need to approve where data is deposited or dictate which repository we use?
  - a. The PI should work with the IRB to determine which data can be deposited and if a data repository meets the security standards necessary to safeguard sensitive data. The IRB can recommend data repositories but cannot dictate which one is used. Ultimately these decisions are up to the PI of the grant.
5. Who is officially responsible for long-term storage of very large quantities of data? Say a research group collects >500 Terabytes of data and then either leaves academia or goes to another institution that simply does not have storage capabilities. Is UAB then responsible for long-term storage of this data, since it was collected at UAB? It can cost huge amounts of money to store large datasets like this, where there are no public databases available for storage (like microscopy). To what extent should indirect costs and institution be responsible vs. the PI and the direct costs of the grant?
  - a. This is a policy/business decision for UAB and the funding sponsor.
6. Data repositories have size limits that do not come close to the amount of space needed for some labs. What is the plan for these situations? (See my earlier question about whether the institution is responsible).
  - a. If you have large datasets, reach out to data repositories during the planning stage to develop the data sharing pipeline and determine costs up front.
7. Would be nice if UAB developed their own core repository we could all use and they could handle requests/access to the data
  - a. This is under evaluation. If a sponsor requires data to be submitted to a particular repository, then that is the appropriate course of action. Otherwise, the PI can choose from any number of data repositories prioritizing subject specific repositories over generalist repositories.
8. The policy about making data available at the end of the performance period is really problematic if the study isn't yet published. Perhaps data could be deposited but placed under an embargo for up to two years if the study was ongoing. Who is best person at the NIH to contact if we want to express concerns?
  - a. This type of concern should be brought up with your program officer.

### **Questions about resources to help prepare a DMSP**

1. Will there be resources that investigators can access during grant preparation to prepare this plan?
  - a. Yes. Resources including sample plans and templates can be found in this [Guide](#) which will be updated as more guidance becomes available.
2. Will UAB have examples available prior to Jan 25?

- a. Yes. We currently have one example available in this [Guide](#). Other Sample DMSPs from UAB researchers are being developed and will be available before Jan 25. For additional examples not necessarily in the NIH format, check out [this catalog](#) of over 150 data management plans.
3. This question was addressed to Research Computing: I think was looking into better solutions for storage and sharing of research data. Do you have any updates on how this ties in with this rule change?
  - a. UAB Research Computing has invested heavily to provide storage resources for the conduct of research. Detailed information about the storage resources is available at [https://docs.rc.uab.edu/data\\_management/storage/](https://docs.rc.uab.edu/data_management/storage/). These storage resources are designed to support the collect, process, and analyze phases of a research data life cycle.

### Questions regarding budgeting for data management and sharing

1. After the NIH grant has ended, how are the repository maintenance-related cost to be paid?
  - a. Many data repositories will accept and maintain data at no cost to the depositor. If the deposition/preservation of your data does incur costs those can be budgeted into the grant and paid up-front before the award period ends.
2. Following up on requesting funds for data management and sharing (at least for the duration of the grant), I assume this is something that we as PIs would be responsible for budgeting into direct costs (and not a separate bucket of funds)? If having to build into direct cost budget, and then the budget is cut, is NIH thinking about how PIs would still be able to still do the data sharing given the amount of resources it would take to do so?
  - a. This is a pressing concern for researchers, however, there is no official statement from NIH on if/how the budget will increase in the future to account for the added responsibilities of sharing data.
3. I am in need of a data source (rebase.org), and wonder whom to contact to subscribe this data source? It is like subscribing to an electronic journal?
  - a. The funds needed to purchase datasets can be included in the budget but should NOT be included in the data management and sharing section of the budget. Costs associated with collecting or otherwise gaining access to research data (e.g., data access fees) are considered costs of doing research.
4. Will study section reviewers ding your grant if not enough money is allocated to depositing data?
  - a. Probably not, but there is no clear guidance on this question. It is likely that this type of issue would be resolved during the Just-in-time phase.
5. This level of data management takes a lot of time and effort, and it seems likely that every lab (even small labs) will need a data scientist on staff to support this activity. Is it acceptable to budget a data scientist purely for purposes of compliance with this new policy?
  - a. Yes, it is acceptable to include a %FTE for a data scientist in the budget.