

## Computing and Data Management Plan for American Family Funding Initiative Projects<sup>1</sup>

Data generated, processed and stored throughout the course of a research project is a valuable asset involving significant resources. Thus, it is important that we understand your computing and data management needs. All Phase II applicants for the American Family Funding Initiative awards will need to submit a document describing their plan to manage the data involved in the proposed project, as well as the computing resources that will be required to process that data. This plan is not part of the initial application for funding.

The prompts provided below will help you create a brief narrative document outlining your computing and data management needs. Please limit this document to no more than 2 pages. A complete plan will describe your data process and articulate a clear vision of your technical approach and your computational resource requirements. Including specifics about all the tools you will use is not essential. This document includes resources that can help you access available computing tools and understand costs.

**NOTE:** If you expect your project will use proprietary (American Family) data or materials, then we ask you to communicate directly with [AmFam Research](#) about computing and data arrangements.

Components of the plan:

### 1) Your Research Data Inventory

You will need to create an inventory of the data you plan to use or collect.

- What type of data is involved (e.g., text, images or algorithms)?
- What is the source of the data?
  - Are the data and/or other deliverables considered [standard or proprietary](#)?
  - Will it be generated by this project or accessed from another source?
  - Does the data need to be regulated in any way (e.g., proprietary or HIPAA or IRB regulated)?
- Where will the data be stored during the project and after the project?
- What software and/or platforms will be required to work with the data?
- What is the estimated amount of data to be collected or accessed?
- Will you share the data outside your team? With whom?
- What documentation will you create to ensure the data is usable in the future?
- Who are the team member(s) responsible for managing and documenting your research data? Please provide name(s), title(s) and email address(es).

***Do you have questions about your data management needs?***

[Research Data Services can help!](#)

***Additionally, this website provides guidance on how to write a data management plan:***

<https://researchdata.wisc.edu/how-to-create-a-dmp>

### 2) Your Computing and Storage Resource Requirements

Create a list of computing resource requirements for your project.

- Do you have a set of tools you are comfortable using and prefer?
  - Examples: R/RStudio, Python, Jupyter Notebook/Lab/Hub, PyCharm, Other IDEs, Excel....

---

<sup>1</sup> Adapted from Research Data Resources (<https://researchdata.wisc.edu/how-to-create-a-dmp/>)

- What types of computing and data processing are involved in your project? Please share with us your current understanding of your computing needs. What support will you need should your project be selected for funding?
  - How will you process the data?
  - Will you process data in stages?
  - How will the data be accessed for computing?
  - What is the frequency of processing and/or data access?
  - How much overall compute time do you need? Please provide an estimate in hours.
  - Will you need high throughput, high performance or GPU computing (for example, to analyze data or train models)?
- Are there security considerations beyond those indicated above in your data inventory?
- Will you share data with additional collaborators, especially outside of the UW?
- What team member(s) will be responsible for managing and documenting your computing requirements?
- What are the costs associated with the required computing or storage resources? If you aren't sure of a specific dollar amount, consider using our "[T-shirt sizes](#)" to help you estimate costs.

Do you have questions or need more information about your computing and storage resource requirements or costs?

[Research Cyberinfrastructure](#) can help! [rci@g-groups.wisc.edu](mailto:rci@g-groups.wisc.edu)

The resources listed below are broadly available to faculty and staff on campus:

- <https://researchci.it.wisc.edu/>: Research Cyberinfrastructure can advise, consult and refer you to UW resources.
- <https://researchdata.wisc.edu/>: Research Data Services can advise and consult on data lifecycle management and data management plans.
- <https://storage.researchdata.wisc.edu/>: The Data Storage Finder Tool is an interactive website to match storage offerings to your specific needs.
- <https://it.wisc.edu/services/#research>: This site lists IT resources useful to researchers. Your department may have additional tools or collaboratives.
- [UW has contracts](#) with cloud computing vendors [AWS](#), [Azure](#) and [GCP](#). Each provides campus with services at or below their published rates ([AWS](#), [Azure](#) and [GCP](#)). The calculators can be challenging to use if you are not familiar with them, so feel free to reach out to Research Cyberinfrastructure (contact information below).
- <https://chtc.cs.wisc.edu/>: The Center for High Throughput Computing can partner with researchers who have more extensive computing needs.

We recommend you contact your local IT staff for guidance on resource availability and usage within your unit or department. Your department may have access to additional, more domain-targeted resources. If your unit or department does not currently use any of the services listed above or is unsure how to proceed with your use case, please feel free to contact Research Cyberinfrastructure for help ([Research Cyberinfrastructure](#) or [rci@g-groups.wisc.edu](mailto:rci@g-groups.wisc.edu)).

The Data Science Institute has GPU resources you can use for your project. Contact [research@datascience.wisc.edu](mailto:research@datascience.wisc.edu) for more information.

Campus and local services may not be appropriate for projects that involve proprietary or restricted data. In these situations, should your project be selected for funding, we ask you to communicate directly with [AmFam Research](#) about computing and data arrangements, if this is the case.