**Defense Advanced Research Projects Agency (DARPA)**

**Defense Sciences Office (DSO)**

**Administrative and National Policy Requirements Document Template**

[No page limit has been specified for the Administrative and National Policy Requirements document. A Data Management Plan is required as part of the submission for this proposal.

Instructional text and guidance are provided within brackets. These texts should be deleted from your final submission.

To ensure the reproducibility of results and the accessibility of program accomplishments to future users, we require proposers document the necessary and sufficient scope of data that may be applicable to these goals. Performers will be expected to document both the proprietary and non-proprietary products of the program (including raw unprocessed data, rarified data sets, test data, experimental designs, software source code and executables, build scripts, process sequence, programmatic communication and other collaboration activities, as well as other data) to ensure the retention and potential reusability of this information.

When possible, DARPA may also share some or all of the program-generated data with the broader research community as open data (with permission to access, reuse, and redistribute under appropriate licensing terms) to the extent permitted by applicable law and regulations (e.g., privacy, security, rights in data, and export control). The complete scope of program-generated data described above may go considerably beyond the scope of data to be made public. Hence, it is expected that as part of a DMP proposers delineate their specific data products that are suitable for public release and how they intend to capture and represent this information.

This DMP should include enough detail to ensure that the data products delivered to DARPA (or made public) are adequate for use by an independent third party in recreation and verification of the scientific results.

Proposers are only asked to explicitly document program data, how much there will be and how they intend to manage it as they execute the program. As this is effort that is required to execute the program, DARPA does not expect the existence of a DMP to produce additional cost burden on proposers for data management requirements during or after the period of performance.

Fundamental Research: <https://www.darpa.mil/work-with-us/for-universities/fundamental-research>

<https://iresearch-cms.tau.ac.il/sites/resauth.tau.ac.il/files/DARPA%20Topological%20Excitations%20in%20Electronics.pdf>

# Data Management Plan

## I. Data Inventory

### [The types of data, software, models, and other materials to be produced in the course of the project.]

[Guidance: Describe what data and other research products you will generate in the course of your project. Include an estimate of the size or amount of data produced (KB, MB, GB, TB), and the formats. Examples of research products include Data sets, including test, experimental, and measurement data; Narratives, including observational logs, journals, and collaborations; Analyses; Decisions, including alternatives, exploration branches, and determinations; Design of experiments and simulations, including setup, ingest, and outputs; Codes (with build scripts, development history and versions), Software (executables with source), Algorithms, Data consumed or produced by software; Models or simulations (computational or mathematical); Bibliographies and citations used by your research; Recordings of various physical phenomena (including images, videos, sensor data, etc.); and any other products created but not included in this list.

Consider the following:

* What data will be generated in the research?
* What data types will you be creating or capturing?
* If you will be using existing data, state that fact and include your sources and permissions.
* What is the relationship between the data you are collecting and the existing data?
* How much data will be produced?
* What formats will your data be in?]

[Enter content here]

**II. Data Capture**

**[The plans for capturing the data,** **including the extent and specific mechanisms to be used during the period of performance for the program.]**

[Guidance: Describe how you intend to capture the data during your project. Include the specific mechanisms such as software packages (include version and operating system), specialized equipment, active storage (local, cloud), data backups, procedures, methodologies, processes.

Consider the following:

* How will you capture or create the data?
* What software will you use to capture, analyze, store, process, visualize or manipulate the data?
* How will you store the data?
* What instruments will you use for this project?
* What specialized equipment will you use?]

[Enter content here]

## III. Data management and metadata standards

### [The data management standards to be used for data and research products, and metadata format and content.]

[Guidance: Describe the format of your data, and think about what metadata that would be needed to be able to use these files. Metadata may entail descriptions of research details such as: experiments, apparatuses, computational codes, software, instruments, etc.

Consider the following:

* What form will the metadata describing/documenting your data take?
* How will you create or capture these details?
* Which metadata standards will you use and why have you chosen them? (E.g. accepted domain-local standards, widespread usage, and funder requirements).
* What contextual details (metadata) are needed to make the data you capture or collect meaningful?

Resources:

DOD Public Access Plan: <http://www.dtic.mil/dtic/pdf/dod_public_access_plan_feb2015.pdf>

Project Open Data: <https://project-open-data.cio.gov>

Project Open Data Metadata Schema v.1.1: <https://project-open-data.cio.gov/v1.1/schema/>

Defense Technical Information Center: <https://project-open-data.cio.gov/v1.1/schema/> ]

[Enter content here]

## IV. Data access and sharing

### [Conditions for access and sharing including protection of privacy, confidentiality, security, intellectual property, reuse, redistribution, the creation of derivative works, or other rights or requirements. Identify which data sets and research products constitute the deliverables part of the program execution plan. Explain how the DMP enhances validation and validity of the results, and how it may support future scientific discoveries and engineering innovations.]

[Guidance: Describe how and when the data will become available, including specific timeframes for each data resource or research product. Identify which data sets and research products will be the deliverables for the project. If any data or resources will not be sharable, note them and explain why. If there is an embargo period for sharing any of the data or resources, make sure you provide details explaining the delay, such as publisher, IP, or collaboration reasons. Describe methods for addressing and protecting sensitive data, including participant anonymity, privacy or data redaction. If the data is of a sensitive nature, but can be shared using a restricted access method, describe the mechanism that will be used. If there are any anticipated or possible future data quality issues, explain how they will be resolved. Identify any restrictions on sharing any of the data resources, how they will be implemented, and who can access the data. Sharable data and research products should be stored in publicly accessible databases, where appropriate and available. Consider using institutional repositories, discipline-specific repositories, government-supported repositories or a central environment such as the cloud, or any combination of these. Explain how the DMP enhances validation and reproducibility of results, and how it may support future scientific discoveries and engineering innovation.

Consider the following:

* Which data sets and research products will constitute the deliverables for the project?
* How will you make the data available?
* What resources are needed to access or use the data? Examples are software or equipment.
* When will you make the data available?
* What is the process for gaining access to the data?
* How long will the original data and research products be retained before making them available for wider distribution?
* Are there any sensitive data resources that need to be anonymized or restricted?
* Are there any anticipated data quality issues that need to be addressed?
* Are there any embargo periods for political, commercial, or patent reasons? If so, give details.
* Are there ethical and privacy issues? If so, how will these be resolved?
* Who will hold the intellectual property rights to the data and how might this affect data access?
* Will there be any restrictions for sharing specific data sets or research products?
* How does the DMP support future research and innovation?
* How does the DMP enhance validation and reproducibility of results?]

[Enter content here]

## V. Plans for archiving and preservation

### [Plans for archiving datasets, research products and other digitally formatted data resources, and for preservation of access.]

[Guidance: Describe physical and cyber resources and facilities that will be used to effectively preserve and store research data long term. These can include third-party facilities and repositories. Describe the transformations needed to make the data sets and resources accessible long term.

Consider the following:

* What is the long-term strategy for maintaining, curating, and archiving the data?
* Which archives, repositories, or databases have you identified as locations to deposit data and research products?
* What procedures do your intended long-term data storage facilities have in place for preservation and backup?
* How long will or should data be kept beyond the life of the project?
* What data will be preserved for the long-term?
* On what basis will data be selected for long-term preservation?
* What metadata and supplemental documentation will be submitted alongside the data or created on deposit in order to make the data reusable?
* What transformations will be required in order to make the data reusable?
* How will adherence to this data management plan be checked or demonstrated?
* Will any permission restrictions need to be placed on the data? If so, what methods do you intend to employ?]

 [Enter content here]