This data management plan (DMP) is intended to outline the methods for managing the data generated as a part of the research in this proposal. It will explicitly describe the types of data generated, the dissemination and sharing of this data, and overall management of the data over time.

**Types of data**

This project will generate physical samples, digital images, measurements about cotton fibers, and software used to produce these data. The physical samples are small amounts of fibers collected from Gossypium species and Upland cotton diversity panels. The digital images are high-resolution images of cotton fibers captured on digital light microscopes in formats of bmp, jpeg, and tif. The measurement data are quantitative numbers regarding fiber size and shape obtained from the system to be developed in this project and other textile testing instruments. All measurement data will be organized and analyzed in Excel, and saved in .xlsx files. The data may be also analyzed in SPSS and Matlab, and the output will be presented using Microsoft Word and PowerPoint. The software to be developed in this project will be written in the Visual C++ and Python, and stored in .cpp, .py and other standard forms. The data generated by the software will be in text files to be imported into Excel.

**Data and metadata standards**

The project team will choose file formats for the digital data and information that will enable the most effective and most secure management over time. All the data formats, such as Excel, Word, jpeg and cpp, are the most open formats and widely used.

Metadata records will be created to describe each of the project's physical samples and digital resources. Metadata will be used for 1) description of the samples and resources, and 2) management and curation of the resource over time. These can be generally achieved through the use of a descriptive metadata standard such as Dublin Core (http://dublincore.org/). This aligns with the standard metadata approach by UNT Libraries in its digital library (http://digital.library.unt.edu/), which will serve as one dissemination venue for publicly available project results, outputs, data, and information products.

**Policies for access and sharing, and privacy**

The project team will work with the UNT Libraries in appraising the information objects and derived data to determine potential benefits to future researchers, and accordingly will make those available publicly through the UNT Libraries' Digital Library or long-term preservation. The team will work to protect privacy and confidentiality concerns, as well as to respect any proprietary or intellectual property rights. For data sources that are embargoed for some period of time, the metadata records will be available to allow discovery of the resources.

All technical reports, presentations, and publications of the project will be available via the UNT's open access Scholarly Works Repository (http://digital.library.unt.edu/explore/collections/UNTSW/), hosted in UNT's Digital Libraries.

All physical samples will be stored securely in the Principal Investigator (PI)'s and co-PIs' research lab. They may be available upon request to the PI or co-PIs.

The software to be developed in the project will be stored in the password-protected cloud space (UNT's OneDrive) during the project time. After the project, the final version of the software will be made available on GitHub.

**Policies for re-use, re-distribution, derivatives**
To the extent possible, data and derived datasets collected and/or created by the project will be available to other researchers for reuse. In consultation with all stakeholders, decisions about open or limited availability will result from the appraisal discussed above. For data and datasets determined by the project team to require limited access, requests can be made to the PI and co-PIs. Metadata describing all data and derived datasets will include information about requesting access and use of the data. Others using the project's collected, created, and/or derived data products will be requested to provide a standard citation (proper attribution provided by the project) to credit both the USDA NIFA and this project for access and use of the data.

**Plans for archiving and preservation**
All data and images will be stored in the password-protected cloud space (UNT's OneDrive), which is routinely backed up and serves as a first layer for curating and archiving of data during the project's duration.

All processed data will be preserved for at least 3 years after the end of the project in facilities provided by the UNT Libraries' Digital Library (http://digital.library.unt.edu), which is committed to long-term access and stewardship of publicly available research outputs. All project-related materials, e.g. technical reports, presentations, and publications will be stored with long-term accessibility in the UNT Scholarly Works open access repository (http://digital.library.unt.edu/scholarlyworks), which is supported through the UNT Digital Library.  The software to be developed in the project will be retained on GitHub for seven years after the end of the project.

**Roles, Responsibilities and Reporting**
The PI and co-PIs will implement the proposed DMP in his/her research space when creating and managing all data, images, samples and software described above.  Specifically, the PI (Xu) will be responsible for generating the images and microscopic measurements of cotton samples, software codes. The co-PI (Hinze) will be responsible for producing two panels of cotton samples and the metadata of each sample.  The co-PI (Kothari) will be responsible for generating fiber measurement data with traditional fiber-testing equipment. All the PIs will be responsible for sharing reports, publications and datasets.  No additional fund is needed for the DMP implementation.

In our annual and final reports to NIFA, we will include progress in data sharing (publications, database, software, curriculum, outreach materials, etc.). The final report will describe the data produced during the award period and the components to be stored and preserved with the expected duration after the award ends.