

## DATA MANAGEMENT PLAN

DOE policy on data storage, dissemination, sharing, preservation, and protection will be strictly followed by all PI, Co-PIs and collaborating scientists on this project, entitled, “*Consortium on Sensing, Energy-efficient Electronics and Photonics with 2D Materials and Integrated Systems for Training the Next-Generation DOE-NNSA STEM Workforce (SEEP-IT)*”. This project will generate three categories of data including synthesis/fabrication/materials data, spectroscopy/characterization data, and simulation and computational data. The data generated from this proposed project will be managed by relevant individual research groups at the University of North Texas (UNT), the University of Texas at Arlington (UTA) and the University of Arkansas at Pine Bluff (UA-PB), in accordance with this data management plan.

### 1. Data Types and Sources

#### 1.1 Synthesis/Fabrication/Materials Data

The produced data from research projects involving the device fabrication will include photo-resist materials, mask, substrate, exposure, and post exposure development, optimized fabrication conditions (e-beam lithography, holographic lithography and reactive ion\_etching) for sensor and laser device, and photonic crystals in materials of SiN, Gd<sub>2</sub>O<sub>2</sub>S:Tb, Ta<sub>2</sub>O<sub>5</sub>, VO<sub>2</sub>, GaP and Al<sub>0.8</sub>Ga<sub>0.2</sub>P. The produced data from research projects involving synthesis, and doping of materials include precursors used for materials development, optimized growth conditions using vapor based approaches such as chemical vapor deposition (CVD) and its hybrids, pulsed laser deposition (PLD), etc., of transition metal dichalcogenides (TMDs) such as WSe<sub>2</sub>, WS<sub>2</sub>, MoSe<sub>2</sub>, and MoS<sub>2</sub>, van der Waals dielectrics, such as h-BN and two-dimensional hybrid perovskites (CH<sub>3</sub>(CH<sub>2</sub>)<sub>3</sub>NH<sub>3</sub>)<sub>2</sub>(CH<sub>3</sub>NH<sub>3</sub>)<sub>3</sub>Pb<sub>4</sub>I<sub>13</sub>, and three-dimensional perovskites Cs<sub>2</sub>AgBiBr<sub>6</sub>. The structural, electrical, and optical properties of the as-grown materials will be characterized. The data collected will either be stored in ASCII/CSV format or in the form of graphs, drawings/graphics, as appropriate. The software used to draw the diagrams and interpret the data will be noted with each result. The instrumentation specifications and measurement protocols will also be noted as appropriate. Data collected during the experiments will be registered in logbooks that will include the synthesis methods of the materials prepared, and the images collected from the characterization techniques will be registered digitally. The samples prepared will be stored in each PI and co-PI's lab.

All these methods, procedures and observations will be recorded in detail in laboratory notebooks of the students and staff working on this project at the lead and partner sites and all pages in the notebooks will be dated. These data will be stored in their original hard copies and the laboratory notebooks will also be scanned into PDF format periodically for electronic storage. All these data will be stored in a password-protected repository that is accessible to all the collaborators, as needed.

#### 1.2 Spectroscopy/Characterization Data

The synthesized samples and devices will be characterized using structural, spectroscopic, and atomic resolution electron microscopic techniques. The data generated from these characterizations and analysis can be classified into four categories, as follows:

- a) Structural and morphology characterization
- b) Metrology protocols for electrical and optical properties
- c) Lasing and sensing characteristics
- d) Structure-property correlation and recommendations

The bulk structure and morphology characterization methods include x-ray diffraction (XRD), Optical Microscopy (OM), Raman Spectroscopy, and Field-emission (FE-) Scanning Electron Microscopy (SEM). The metrology protocols for atomic structure elucidation include Focused Ion Beam (FIB) milling followed by high-resolution transmission electron microscopy (HR-TEM) and electron energy loss spectroscopy (EELS). The primary optical properties will be characterized via Photoluminescence (PL), UV-VIS optical absorption and transmission spectroscopy, and Raman. The electrical properties will be characterized with the help of physical property measurement facility. Optimized conditions with lasing and sensing devices will be identified and recommended for IP, publication, and archival reports.

### **1.3. Simulation and Computational Data**

The project will generate the following types of computational data:

- a. Structural parameters for materials and devices studied
- b. Sets of property characterization (resonance quality factor, absorption, transmission, e-field intensity)
- c. Computational tools and methods developed to calculate and model specific device properties
- d. Software scripts used to analyze the collected data

The produced data from simulation includes input and output files as well as photonic band structure, reflection and transmission spectra, dielectric functions, and other matrices. The input and output files are in ASCII format while the latter are stored in binary format. Software and analysis scripts are saved either in ASCII text format (bash, python, Matlab, SCM, CTL) or in Jupyter notebooks (for analysis in Python). In cases where such sharing is appropriate (according to standard scientific practice), it is desired that other researchers will re-use programs and produce derivative-software of generic utility to the scientific community. Unpublished data will be shared with other researchers upon request, according to standard scientific practice. The project does not produce data associated with privacy or confidentiality concerns.

## **2. Content and Format**

Digital metadata such as images, simulation results, and software will be stored electronically. Electronic repositories for data will be backed up regularly. Standards and protocols regarding data storage have been established in the collaborating laboratories. Data and protocols for publication will be generated using software such as Microsoft Office, Latex, Proteus (thermal analysis), Adobe Photoshop, MATLAB, COMSOL, SolidWorks etc. Summary/report, experimental data, papers, and lecture notes will be stored in \*.pdf format. CAD models of the design devices will be stored in \*.dxf format; Raman and photoluminescence (PL) data, lasing emission data, and

electronic structure data are in \*.txt format; AFM data for layered sample thicknesses, SEM images of the fabricated samples, photos of the experimental setup, and photos of the educational activities will be stored in \*.jpg format. Other data, not published or otherwise available in the theses of students participating in the project, will be made available as picture format files (\*.jpg, or \*.pdf files), as appropriate.

The main file formats that will be used for the data are .csv, .xml, .doc, .pdf, .eps, .jpeg, .tif, .txt, and .stl. These file formats and software are widely used, and the formats of data generated are acceptable, adjustable and exchangeable by most publishers, industry, academia and government.

### **3. Sharing and Preservation**

Data generated under the project will be freely distributed to the research community according to DOE guidelines on data sharing. According to standard scientific practice, unpublished data will be shared with other researchers upon request. Every effort will be made to ensure that public release of data occurs at the earliest reasonable time. Data generated from this project will be disseminated to the broader scientific community through: a) presentations at regional, national and international research conferences and universities, b) manuscripts published in refereed journal, and c) student theses and dissertations. We will make timely manuscript submissions to high-impact scientific/technical journals, and the data will be available immediately following the acceptance of the paper, unless such data is explicitly “embargoed” by the journal. If allowed, we will also include the data as supplementary information to the published article. The published article will have a Data Availability Statement: Data will be available upon request.

When possible, theses, dissertations, and papers published as a result of the proposed project will be made available on public websites (e.g., faculty members’ websites, library website, or institutional repositories). This will make many of our results available in portable document format (\*.pdf).

For electronic deliverables of this project, the bulk of data generated by the project will be stored by the university data storage system that are meant for long-term storage at UNT, UTA and UA-PB. For hardware deliverables of this project, we plan to archive them in each PI and co-PI’s lab. This will enable easy management and future re-use of the hardware for continuation of proposed research directions.

The bulk of data generated by the project will also be stored by the super computer (Talon) data storage system at UNT. Talon is made up of 224 separate computer servers and has 200 terabytes of storage, which is enough to house 10 copies of all of the books in the Library of Congress or 40,000 DVDs, and operates at a speed of 20 Teraflops.

Our minimum planned data retention period is three years following the conclusion of DOE support for this project or three years following public release, whichever is later. In practice, we frequently retain data for a much longer period of time (e.g. decades). An extended period of data retention will be implemented for students working on the project prior to their graduation or thesis/dissertation completion. Research data that support patents will be retained for the entire term of the patent.

**Data generated through collaborations among UNT, UTA, UA-PB and NNSA/DOE Laboratories: ANL and SNL:** The data will be made accessible as soon as possible to the public, in accordance with the principles stated in the Office of Science Statement on Digital Data

Management: <https://science.osti.gov/Funding-Opportunities/Digital-Data-Management>. Peer-reviewed publications produced by this project will be made available via the Department of Energy's "Public Access Gateway for Energy and Science" (DOE PAGES), which provides free public access to peer-reviewed publications twelve months after the date of publication.

**Intellectual Property Rights:** As publications result from the proposed research, the PIs will acknowledge the support from DOE with the associated grant numbers. It is not expected that any of the data should create ethical or privacy issues. The IP on the significant results developed in the project will be safeguarded by means of appropriate registration, if necessary, with United States Patents and Trademarks Office, and other results will be exploited with the help of the corresponding University Technology Transfer Office, which has previously supported the filing of patents of each research group. The ownership of the invention will be distributed based on the work done, according to the US public service invention law. DOE Principles and Guidelines will be strictly followed to pursue intellectual property and patents relating to the proposed project and will ensure the availability to the research community.

**Policies and Provisions for Re-Use and Re-Distribution:** It is anticipated that there will be considerable interest in the developed computational tools, models, designs, and methods within the nuclear security, nanotechnology and photonics community. The data generated from this project will not be copyrighted, with no license issues, and there will be no permission restrictions placed on the data.

#### **4. Protection**

Undergraduate and graduate students, and post-doctoral fellows will participate in the proposed research. Their personal privacy, and personally identifiable information will be protected during the annual evaluation of the project.

#### **5. Rationale**

The proposed data sharing and preservation will enable the validation of research results in Photonics, Electrical Engineering, Nanotechnology Engineering, Microelectronics, Nuclear Engineering, Nuclear Security-Radiation Detection Systems, and speed up the research in those areas. The proposed DMP will also make publications resulting from the proposed research open, machine-readable, reusable, and digitally accessible to the public at the time of publication. This project does not involve a DOE facility beyond what is conventionally made available to approved users.