DATA MANAGEMENT PLAN

# 1. Data description and collection or re-use of existing data

## How will new data be collected or produced and/or how will existing data be re-used?

The data generated during the project can be divided into two categories: written by the PI (e.g. notes, spreadsheets, protocols), stored as \*.txt, \*.doc, \*.docx,\*.xlsx, \*.opj files; and data collected during laboratory experiments. Data file format produced during measurements may differ depending on the measurement technique, but mostly it will be numerical (e.g. \*xlsx, \*txt, \*csv) or graphical (e.g. \*jpg, \*png, \*mdt) files. The types of generated data depends strictly on the software used to collect or analyse it. Only new data will be generated during the project. The data collected during the measurements will be stored according to previously established protocol. All the comments and details about synthesis and measurement will be stored in laboratory notebook. Calibration of the measuring equipment will be carried out according to the manufacturer's recommendations. All measurements will be performed in controlled conditions to ensure data uniformity and quality.

Data about sample synthesis: physical form/Microsoft Office format (\*.xlsx or \*.doc/\*docx); volume: max 0.2GB

Information about the procedure for measuring samples: Microsoft Excel and/or Microsoft Word format (\*.xlsx or \*.doc/\*docx); volume: max 0.2GB

spectroscopic data: \*.sp (raw data), \*.csv, \*.dat; volume: max. 0.3GB

X-ray diffraction data: \*.xrdl (raw data) and \*.csv (exported data) ; volume: max. 0.3GB

Scanning electron microscope images: \*.tiff; volume: max. 20GB

X-ray photoelectrone spectroscopy data: \*.vms (raw data), \*.csv (exported data); volume: max. 0.5GB

AFM microscopic data: \*mdt (raw data) and \*.gwy, \*tiff, \*.csv (project and exported data); max. 20GB

Origin spreadsheets: \*.opj; volume: max. 1.5GB

Other data: \*.jpg, \*.tiff, \*.eps, \*.cif

dr inż. Piotr Okoczuk

## What data (for example the types, formats, and volumes) will be collected or produced?

# 2. Documentation and data quality

## What metadata and documentation (for example methodology or data collection and way of organising data) will accompany data?

The data, along with metadata, will be collected immediately after the fabrication of the devices and stored on the hard drive of a computer connected to the measurement system. Each measurement series will be catalogued separately in an individual folder named according to the following format: YYYYMMDD - Name, where YYYY stands for the year, MM for the month, and DD for the day the measurements were performed. “Name” will be a brief description of the experiment to facilitate later identification of the folders. After completing each measurement series, the data and metadata will be copied to the hard drive of the project leader’s computer, where they will be catalogued in the same manner.

## What data quality control measures will be used?

Basic data quality control measures will include the analysis of electrical parameters such as conversion efficiency, fill factor, open-circuit voltage, and short-circuit current, as well as both automated and manual procedures to verify the accuracy and consistency of the data. This will be achieved through regular measurements of the fabricated photovoltaic cells and a reference silicon cell.

# 3. Storage and backup during the research process

## How will data and metadata be stored and backed up during the research process?

Raw measurement data will be copied and stored both on the PI’s computer and an external device, such as a hard drive. Backups will be performed regularly, at a minimum of once per month.

## How will data security and protection of sensitive data be taken care of during the research?

The measurement and scientific equipment, as well as the PI’s computer, are connected to the institution’s secure internal network. Access to the PI’s computer is protected by a password and restricted solely to the PI. Only authorized institutional personnel have access to the measurement devices and associated computers. Data will be routinely backed up and stored on an external storage device.

# 4. Legal requirements, codes of conduct

## If personal data are processed, how will compliance with legislation on personal data and on data security be ensured?

The project will not include the collection or handling of personal data or any other data subject to protection requirements.

## How will other legal issues, such as intelectual property rights and ownership, be managed? What legislation is applicable?

All intellectual property generated throughout the implementation of the project will remain the property of Gdańsk University of Technology and be managed by the university. The data and results will be shared publicly via open access under one of the Creative Commons license.

# 5. Data sharing and long-term preservation

## How and when will data be shared? Are there possible restrictions to data sharing or embargo reasons?

The part of the data will be published in the Institution’s open research data repository – MOST Wiedzy Open Research Data Catalogue. The part of the data will be published in scientific journals which may also require raw data publication or links to the MOST Wiedzy Open Research Data Catalogue be provided with the research articles.

## How will data for preservation be selected, and where will data be preserved long-term (for example a data repository or archive)?

The MOST Wiedzy Open Research Data Catalogue will be the main data repository. Data provided to that repository will be chosen based on its scientific quality and exemplarity. Additionally, the data transferred to the repository will fulfill FAIR requirements and will be categorized and labeled according to the standard file formats. Data published in scientific articles will have priority for long-term keeping. Raw data in open formats (.csv/.dat/.txt/image formats) will be accessible to the general public. The repository holds the CoreTrustSeal certification, confirming its reliability and sustainability. Data deposited there will be automatically categorized for long-term preservation with no expiration date.

## What methods or software tools will be needed to access and use the data?

Depending on the dataset the software used for its processing will be either closed license software or open-source. The data stored in MOST Wiedzy Open Research Data Catalog will be in open formats, e.g.: txt, tiff or csv. The data in raw formats will be provided on direct requests.

## How will the application of a unique and persistent identifier (such us a Digital Object Identifier (DOI)) to each data set be ensured?

The data available in MOST Wiedzy Open Research Data Catalogue platform will be assigned with the DOI number.

# 6. Data management responsibilities and resources

## Who (for example role, position, and institution) will be responsible for data management (i.e the data steward)?

Open Science Competence Center (https://pg.edu.pl/openscience) – established by Gdańsk University of Technology will be responsible for Data Management Plan, data storage and dissemination. The Principal Investigator will be responsible for the procedures assessment and overall data quality.

## What resources (for example financial and time) will be dedicated to data management and ensuring the data will be FAIR (Findable, Accessible, Interoperable, Re-usable)?

No additional resources are required.