Plan Overview

A Data Management Plan created using DMPTuuli

Title: High-Performance Bioinspired Conductive Cellulose Films for Thin Film Solar Cells (BIOSOL)

Creator: Mahboubeh Hadadian

Principal Investigator: Riikka Suhonen, Kati Miettunen

Contributor: Marja Valimäki, Mahboubeh Hadadian

Affiliation: University of Turku

Funder: The Research Council of Finland (former The Academy of Finland)

Template: Research Council of Finland

ORCID iD: 0000-0002-6564-6262

Project abstract:

The main goal of BIOSOL is to develop novel high performance conductive cellulose films for optoelectronics and thus open new lines of products for forestry industry. Cellulose allows detaching from fossil resources and utilizing renewable material streams that does not interfere global food production. Cellulose films can deliver very high transparency and can withstand high temperatures as well as enable functionalization. We ambitiously demonstrate printing of perovskite solar cells on these novel flexible substrates as well as pilot production of the novel conductive cellulose substrates. BIOSOL bridges highly ambitious scientific goals and creation of significant societal impact - this work requires multidisciplinary team with different types of institutes with complementary skillsets.

ID: 29948

Start date: 01-01-2026

End date: 31-12-2029

Last modified: 19-12-2025

Grant number / URL: 373667

High-Performance Bioinspired Conductive Cellulose Films for Thin Film Solar Cells (BIOSOL)

1. General description of data

1.1 What kinds of data is your research based on? What data will be collected, produced or reused? What file formats will the data be in? Additionally, give a rough estimate of the size of the data produced/collected.

The data of the BIOSOL are based on developing conductive cellulose films and their application in emerging photovoltaics. Table 1 shows the data that will be produced during the project.

Table 1. Overview of the datasets, including data owner, file type, approximate size, and required software.

Data	Format	Software	Estimated size	Owner
Cellulose photographs	.tif	Image viewer	2 GB	UTU
Cellulose micrographs	.tif	Image viewer	1 GB	UTU
Cellulose SEM images	.tif	Image viewer	1 GB	UTU
Optical transmittance data	.csv	Origin	500 MB	UTU/VTT
Angle dependent transmittance data	.csv	Matlab	500 MB	UTU/VTT
Surface roughness data	.csv and .tif	Origin and Gwyddion	1 GB	UTU/VTT
Contact angle data	.csv and .tif	Origin and Image	500 MB	UTU/VTT
Mechanical test data	.csv	Origin	500 MB	UTU
Oxygen and water vapor permeation data	.csv	Origin	500 MB	UTU/VTT
Cellulose accelerated aging data	.csv and .tif	Origin and Python	3 GB	UTU/VTT
Leaf skeletons database	.xml	Endnote or Mendeley	100 MB	UTU
Leaves and leaf skeletons photographs	.tif	Image viewer	1 GB	UTU
PSC photographs	.tif	Image viewer	2 GB	UTU/VTT
PSCs IV and EIS characterization	.csv	Origin	500 MB	UTU/VTT
PSCs accelerated aging and stability data	.csv and .tif	Origin and Python	3 GB	UTU/VTT
PSC modelling data	.mat	Matlab and Comsol	100 MB	UTU
Plots	.opju and .tif	Origin and Image viewer	500 MB	UTU/VTT

Documentation: Notes, methodological descriptions, and analysis scripts (e.g., Python, MATLAB, Origin).

1.2 How will the consistency and quality of data be controlled?

Organization's network drives. These drives are managed by each organization's IT department and include regular backups and inherent permissions control to ensure only authorized researchers have access to the raw data. All data produced during the project will also be saved on the different organization's local network drives and, when applicable, on the CSC Allas service. The project ensures data consistency and quality through the implementation of standardized protocols and rigorous validation measures. All experiments and processes will adhere to predefined methods, documented

in a centralized repository to guarantee reproducibility. Instruments such as spectrometers, solar simulators, and scanning electron microscopes will be regularly calibrated using standard reference materials. Automated scripts will validate raw data formats to detect anomalies, and data will be directly logged from instruments to minimize manual entry errors.

In addition, the project will employ metadata standards aligned with FAIR principles to make datasets findable, accessible, interoperable, and reusable. Version control systems will track changes in datasets and analysis scripts, ensuring transparency and traceability. Weekly team reviews will identify and address inconsistencies early in the process, while error logs will document any issues encountered during data collection or processing. By combining these measures with rigorous oversight and documentation, BIOSOL will maintain the highest standards of data quality and integrity, supporting reproducibility and impactful research outcomes.

2. Ethical and legal compliance

2.1 What legal issues are related to your data management? (For example, GDPR and other legislation affecting data processing.)

BIOSOL project will comply with all relevant legal and ethical regulations concerning data management, including the General Data Protection Regulation (GDPR) and Finnish data protection legislation. Although the project does not involve personal or sensitive data, proper measures will be in place to ensure that all collected and produced data adhere to applicable laws governing intellectual property rights (IPR), data sharing, and research ethics. Ownership and usage rights of data will be clearly defined in agreements between collaborators and partners to prevent disputes.

2.2 How will you manage the rights of the data you use, produce and share?

The rights to the data used, produced, and shared in the *BIOSOL* project will be managed through clearly defined agreements among the project partners. These agreements will outline the ownership, access rights, and responsibilities related to the data, ensuring fair and transparent use by all parties. Data ownership will typically reside with the organization or researcher responsible for its creation, while shared datasets will be governed by established intellectual property frameworks and licensing terms. For data intended for public dissemination, open-access principles will be followed whenever possible. Data will be shared under suitable licenses, such as Creative Commons (CC BY) for publications and datasets or other open data licenses that encourage reuse while safeguarding attribution rights.

3. Documentation and metadata

3. How will you document your data in order to make the data findable, accessible, interoperable and re-usable for you and others? What kind of metadata standards, README files or other documentation will you use to help others to understand and use your data?

To ensure that the data produced in BIOSOL is findable, accessible, interoperable, and reusable, all datasets will be documented using metadata aligned with best practices in the research field. Each dataset will include a comprehensive README file detailing the experimental procedures, data collection methodologies, calibration details, and any relevant information. All data and metadata will be systematically named by date, time, and a brief description, organized within a coherent folder structure. Relevant data will be archived in trusted repositories or university archives, ensuring availability with clear licensing terms to facilitate reuse and support further research.

4. Storage and backup during the research project

4.1 Where will your data be stored, and how will the data be backed up?

The data produced in the project will be stored in secure, reliable, and accessible location, named Taltio network drive, which is maintaied by IT department which include proper data backup and data security. Dataset can be stored in CSC's Allas platform for possibility of sharing the data with the collaborators, which has its own backup system. Backup procedures will be automated and carried out on a regular basis to minimize the risk of data loss. The data produced in collaboration institutions will also be copied and moved to the Taltio NetWok drive according to agreements. This multi-layered approach ensures that the data remains safe, accessible, and retrievable throughout the project lifecycle and beyond.

4.2 Who will be responsible for controlling access to your data, and how will secured access be controlled?

Access to the data produced in the *BIOSOL* project will be controlled by the project’s PI and data manager responsible for Material Engineering department, Teemo Hynnä. They will oversee data storage, access, and data security.

5. Opening, publishing and archiving the data after the research project

5.1 What part of the data can be made openly available or published? Where and when will the data, or its metadata, be made available?

BIOSOL aims to publish and make openly available all research data that can be shared without restrictions. Data generated in the project, such as general research findings, metadata, and datasets related to cellulose substrates and perovskite solar cells (PSCs), will be made openly available. During the project, access to data is restricted to the researchers involved. Scientific articles will be published throughout the project, and the data and metadata supporting the published results will be made available together with the articles. Where possible, data will be stored by the publisher; otherwise, it will be deposited in recognized open repositories such as UTU's networks drives, Taltio and Zenodo, and CSC digital preservation service. Research software, scripts, and data processing workflows developed in the project will be shared via **GitHub**, with stable releases archived in Zenodo to ensure long-term accessibility and citation. At the end of the project, all shareable data and metadata

produced during the project will be archived in appropriate repositories (e.g., Zenodo, Fairdata/Taltio, or NOMAD) and made publicly available.

5.2 Where will data with long-term value be archived, and for how long?

During the project, data will be stored on the Taltio network drive, which is maintained by the IT department at UTU. This system includes inherent permission controls to restrict access to authorized researchers only. Raw data will remain unaltered and protected with daily and monthly automatic backups. Data with long-term value from the *BIOSOL* will be archived in trusted and secure repositories that comply with international standards for data preservation. In UTU's networks drives, Taltio and Zenodo. In addition, the data will also be saved on UTU's local network drives and CSC's Allas platform, which facilitates sharing among project collaborators. Each dataset will have comprehensive metadata documentation, according to best practices in the research field. This metadata will include detailed procedural notes, calibration details, and experimental parameters. The data and metadata will be systematically named by date, time, and a brief description, ensuring an organized and intuitive data structure. All metadata files will be stored in the same folder as the datasets themselves. The data will be stored for long (minimum of 10 years), mid, or short-term (a few years), ensuring that the data remains accessible and usable for future research, verification, and reuse. This long-term preservation will include proper metadata, documentation, and versioning to ensure data accessibility and usability over time.

6. Data management responsibilities and resources

6.1 Who (for example role, position, and institution) will be responsible for data management?

The responsibility for data management within the *BIOSOL* project will be assigned to the project's Data Manager, who will be a senior researcher or research coordinator at the leading institution of the project (University of Turku). The Data Manager will oversee the entire data lifecycle, including data collection, quality control, documentation, storage, sharing, and archiving. Additionally, individual researchers and postdoctoral researchers involved in specific aspects of the project will be responsible for managing data within their areas of expertise, ensuring adherence to best practices in data handling. They will work closely with the Data Manager to ensure that the data is well-documented, securely stored, and made accessible according to the project's data management plan.

6.2 What resources will be required for your data management procedures to ensure that the data can be opened and preserved according to FAIR principles (Findable, Accessible, Interoperable, Re-usable)?

To ensure that the data can be opened and preserved according to the FAIR principles, the following resources will be required:

1. **Data Management Tools and Software**:
    - A secure data management platform (e.g., institutional repository, cloud-based solution) for storing and sharing data.
    - Tools for creating and maintaining metadata (e.g., data documentation tools, metadata

standards such as Dublin Core or DataCite).
    - Software for version control and data tracking to ensure data integrity over time.
2. **Training and Expertise**:
    - Researchers and project staff will receive training in FAIR principles and best practices for data management to ensure compliance.
    - The project team will include individuals with expertise in data curation and metadata standards to support the documentation and preparation of datasets for long-term access.
3. **Data Sharing and Archiving Resources**:
    - Collaboration with trusted data repositories and archives (e.g., Zenodo, Figshare, or a discipline-specific repository) that comply with open-access and long-term preservation standards.
    - Dedicated funding for the storage and ongoing maintenance of the data and its associated metadata for the required duration.
4. **Hardware and Infrastructure**:
    - Sufficient computational resources for data processing and backup, including cloud storage services or institutional servers with regular backup procedures.
    - Infrastructure for secure access control and user authentication to ensure that only authorized personnel can modify or access sensitive data.
5. **Compliance with Legal and Ethical Requirements** :
    - Legal advisors or data privacy experts to ensure compliance with GDPR, intellectual property laws, and ethical data handling practices, especially for sensitive or proprietary data.

These resources will be used to ensure that the data collected, produced, and shared throughout the *BIOSOL* project is Findable, Accessible, Interoperable, and Re-usable by researchers, policymakers, and the broader community.