Plan Overview

A Data Management Plan created using DMPonline

Title: Dam Safety Early Warning Systems Using Remote Sensing and IoT

Creator: Desmond Evurani

Principal Investigator: Desmond Evurani

Contributor: Dr. Alison Parker, Dr. Francis Hassard

Affiliation: Cranfield University

Template: DCC Template

ORCID iD: 0009-0000-4432-0231

Project abstract:

Dam safety is a critical concern globally, especially in regions with aging infrastructure, climate variability, inadequate maintenance capacity and limited real-time monitoring capabilities. Traditional dam monitoring methods such as periodic inspection and isolated instrumentation are often insufficient to detect early signs of deformation, seepage, or structural instability. Although Remote Sensing (RS) technologies such as Synthetic Aperture Radar (SAR) and optical satellite data have significantly advanced deformation monitoring, their temporal limitations and susceptibility to environmental noise reduce their effectiveness as standalone early warning systems (EWS). Similarly, low-cost Internet of Things (IoT) sensors offer continuous ground measurement but remain limited in spatial coverage. This research proposes an integrated Dam Safety Early Warning System (DSEWS) that combines multi-source RS data with IoT sensor networks through interoperable data fusion and machine learning (ML) based anomaly detection. The study aims to design, develop, and evaluate a cost-effective, scalable, and context appropriate monitoring system tailored for tropical, resource-constrained environments. Using a mixed-methods approach including satellite data analysis, sensor deployment, and predictive modelling, this project seeks to generate an adaptive and actionable early warning framework capable of enhancing dam safety, reducing failure risk, and strengthening climate resilience in vulnerable regions.

ID: 189182

Start date: 28-07-2025

End date: 28-07-2028

Last modified: 18-11-2025

Grant number / URL: https://scholarship.ptdf.gov.ng

The above plan creator(s) have agreed that others may use as much of the text of this plan as they would like in their own plans, and customise it as necessary. You do not need to credit the creator(s) as the source of the language used, but using any of the plan's text does not imply that the creator(s) endorse, or have any relationship to, your project or proposal

Dam Safety Early Warning Systems Using Remote Sensing and IoT

Data Collection

What data will you collect or create?

Datasets:

1. Dam instrumentation Data
2. Remote Sensing Data
3. IoT Sensor Data
4. ML Data
5. Hydrological Data

Dam instrumentation data is data from various on-site instrumentation installed in the dam. Remote sensing data is data collected about the earth's surface and atmosphere using satellites such as Interferometric Synthetic Aperture Radar (InSAR) and Optical data to provide multispectra data for vegetation, moisture and surface water analysis. Internet of Things (IoT) sensors provide on-the-ground data in real- time measurements such as pore pressure, soil moisture levels, reservoir water level, rainfall or structural integrity). Machine Learning (ML) data is annotated datasets used to train, test and validate machine learning models so it can learn to identify patterns and make predictions on water level trends and detect anomalies. Hydrological data is information relating to the movement, distribution and quality of water on Earth.

The Data formats are as follows:

1. Numerical data in csv, tif, or similar formats.
2. Image data in JPEG/PNG/TIFF for visual sensing.
3. Video data in MP4/MOV for dynamic inspections.
4. Structured datasets for model in HDF5, TFRecord and other ML compatible formats.

How will the data be collected or created?

Data will be collected through the following mediums:

1. Field/Site Visits and Investigation: Collect data from on-site instrumentation installed in the dam.
2. Sensors: Collect real-time data using sensors installed and other autonomous platforms. IoT data will be collected in rea-time and stored in a cloud-based or local server environment.
3. Simulations: Generate datasets using software simulation environments.
4. Collaboration: Gather industry/institution provided datasets, utilize open access satellite datasets and publicly available datasets. Collaboration with government agencies, dam managers and operators, European Space Agency, and other satellite monitoring agencies.
5. Data Processing: Preprocess raw data for feature extraction and format standardization. Data preprocessing will be addressing missing values, outliers and noise.

For each study site a folder will be created with the naming convention:
StudyDate[YYYY-MM-DD]_StudySite e.g. 2025-11-15_Cranfield

Within these folders, all files pertaining to a particular study site must use the following naming convention

StudySite_DataDescription_Version e.g. Cranfield_LocationData_V1
Cranfield_DamSafetyModelScript_V1

When similar data is collected on different dates files will be created with the following naming convention

StudySite_DataDescription_DateCollected[YYYY-MM-DD]_Version e.g. Cranfield_WaterLevel_2025-11-16_V1

When a file needs to be preserved (e.g. for review) a new version of the file must be created with an incremental version number before any further modification are made e.g.

Cranfield_LocationData_V1 should be copied to Cranfield_LocationData_V2 and only the "_V2" file modified from that point on preserving the "_V1" file

Documentation and Metadata

What documentation and metadata will accompany the data?

The data will be all stored in a repository. I will capture all information in a database with links to each item in a readme text file and file header. Metadata will be added to all datasets to ensure each dataset is understandable and reusable before it is submitted to a repository. All documents will comply with Cranfield's Management of Research Data Policy.

Documentation will include:

- Data collection methodology, date and location.
- Measurement instruments and units used.
- Data processing and analysis methodology.
- Folder and file naming convention.
- Data processing/analysis script annotation.
- Explanations of acronyms, abbreviations and variables including full description of any table headings.
- Any software (including versions) required to view data/documents.
- Link to papers/thesis produced from data

Ethics and Legal Compliance

How will you manage any ethical issues?

No personal or sensitive data will be collected throughout this project. I will manage all ethical issues if any with the laid down standards it required in line with the Cranfield University's ethics framework, however, the research does not involve the collection, use or processing of personal or sensitive data. One Drive/Sharepoint systems, backed up regularly, and accessible only to authorised project team members. There are no legal issues with the research project. The work focuses solely on academic

analysis of dam safety monitoring. All tools and materials used are open source for academic use.

How will you manage copyright and Intellectual Property Rights (IPR) issues?

No data sharing or copyright and intellectual property rights issues are anticipated. The data collected, results and analysis for this project would belong to Cranfield University and I.

Storage and Backup

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

While adhering to Cranfield University's policies on data storage and backup, the primary data storage location for this project will be on secure Cranfield University provided One Drive. Secondary storage location will be on Laptop and Desktop, also secure physical storage for any sensitive data will be implemented using external hard dives in locked cabinets.

Automated backups will be implemented at regular intervals, example bi-weekly depending on the data's criticality. this will ensure at least three copies of the data namely: Primary copy (active storage), Backup (external or cloud-based storage), and Offsite copy (to protect against local disasters).

How will you manage access and security?

No additional security is required in this project. Nevertheless, the network drive stores data under encryption, and all copies require password.

Selection and Preservation

Which data are of long-term value and should be retained, shared, and/or preserved?

All relevant datasets will be preserved in full, and will be placed in the Cranfield University repository, CORD, with a retention period of at least 10 years.

Dataset relevance will be determined prior to thesis submission by myself, my supervisor and project partners based on:

- data required to validate the research documented in the thesis.
- possible reuse of the data for further research or training

What is the long-term preservation plan for the dataset?

This involves maintaining the integrity, accessibility, and usability of the dataset for external periods, perhaps beyond the duration of this project. The long-term preservation plan for the dataset for this project incorporates aspects such as Data storage and Backup, Metadata and Documentation, Long-Term Access and Distribution and Regular Data Review and Updates.

Data will be preserved long-term in CORD; there will be no costs and sufficient storage space is available.

Data Sharing

How will you share the data?

Through Cranfield University's research repository such as CORD.

Cord's figshare repository system and assignment of a DOI for the data will make it findable/citable.

Are any restrictions on data sharing required?

No restriction on data sharing are anticipated.

Responsibilities and Resources

Who will be responsible for data management?

Desmond Evurani

What resources will you require to deliver your plan?

Special support is available from the Cranfield University's Research Data Management regarding the DMP and use of CORD. CORD is free to use.