Plan Overview

A Data Management Plan created using DMPonline

Title: A highly automated InSAR detection system for nuclear infrastructure health monitoring

Creator: Stewart Agar

Principal Investigator: James Lawrence, Stewart Agar

Data Manager: Stewart Agar

Project Administrator: James Lawrence

Contributor: Anthony Carpenter

Affiliation: Imperial College London

Template: Imperial College London Generic DMP

ORCID iD: 0000-0002-9781-817X

ORCID iD: 0000-0002-9524-6219

Project abstract:

This project will produce a highly automated system for structural health monitoring (SHM) via analysis of satellite SAR data, specifically developed for EDF's decommissioning of nuclear infrastructure programme. As satellite platforms continue to capture data, key insights regarding the structural health of infrastructure will be distilled to monitoring teams in regular intervals.

Valuable use cases for InSAR technology in Nuclear Decomissioning were demonstrated by Telespazio in 2022 over three reports. This work will provide robust InSAR SHM software and capabilities to EDF and Magnox Ltd., the eventual custodians of decommisioning sites.

The capabilities of such a system will be enhanced through development of novel machine learning approaches to couple SAR data with optical, multispectral, ground-based investigation, and other geospatial data. Machine learning will further be developed to aid user interpretation of observed phenomena.

ID: 123330

Start date: 02-05-2023

End date: 02-05-2025

Last modified: 16-06-2023

Copyright information:

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

A highly automated InSAR detection system for nuclear infrastructure health monitoring

Administrative details

Project Name/Title

A highly automated InSAR detection system for nuclear infrastructure health monitoring

Principle Investigator/Researcher

Stewart Agar

ORCID iD (if applicable)

0000-0002-9524-6219

Project Description

This project will produce a highly automated system for structural health monitoring (SHM) via analysis of satellite SAR data, specifically developed for EDF's decommissioning of nuclear infrastructure programme. As satellite platforms continue to capture data, key insights regarding the structural health of infrastructure will be distilled to monitoring teams in regular intervals.

Data Collection

What data will you create or collect?

SAR images will be processed into a variety of intermediate formats. These generally form 2D or 3D arrays corresponding to the input SAR images or their geographic projection. Examples include: coregistered images, stacks of images ('data cubes'), geocoded coordinated, models for physical processes which modulate the SAR data such as atmospheric delays, ground deformation models.

Raw data will be processed using OpenInSAR software and used to create maps and annotated figures which may be used in reports and publications.

Machine learning models may be trained for purposes such as early identification or interpretation of deformation phenomena and change detection. The trained weights for these models may form valuable data that is computationally expensive to recreate.

Will you be reusing existing data (e.g. data sourced from a 3rd party data provider)?

- Yes (please give details)

The project will make extensive use of European Space Agency satellite data as inputs for deformation monitoring. Data from the Sentinel 1 mission will be used extensively throughout.

Optical data from e.g. Sentinel 2 will also be used to for change detection and feature recognition purposes.

Digital elevation models such as from NASA's SRTM and the Environment Agencies LiDAR campaigns will also be utilised for data processing and visualisation purposes.

What file formats will be used?

SAR data and related products will generally be provided in TIFF (GeoTIFF) format. This is an open format, offering a number of compression options while being flexibile in terms of data dimensionality.

XML or a similar markup language will be used to store metadata. This XML may be embedded within the TIFF files as appropriate.

JPEG2000 is a common format of input optical and multispectral data.
GIS data may also be delivered in Geopackage or Shapefile open formats.

Will you create any software or write any code to process or analyse data?

- Yes (please give details)

A key product of this project is 'OpenInSAR' software to gather and process SAR data into deformation measurements. A further system will be produced to automate OpenInSAR analyses and serves as an adapter and data management system for Cloud computing and HPC platforms. Finally, software will be produced to automate the interpretation and communication of deformation measurements.

Ethics and Legal Responsibilities

Does your research involve human participants?

- No

Will you be processing/collecting personal data?

- No

Will you be processing/collecting special categories of personal data (please select all that apply)?

- None of the above

Are there any IP or copyright restrictions which might influence your use or sharing of the data?

- Yes (please give details)

The Intellectual Propertry (IP) and copyright restrictions outlined in the research project contract will be adhered to. Please see section 4 (below) for 'USE AND EXPLOITATION OF INTELLECTUAL PROPERTY RIGHTS'

4.1 All Background belonging to one Party is and shall remain the exclusive property of the Party owning it (or, where applicable, the third party from whom its right to use the Background has derived).

4.2 Each Party grants the other Parties (to the extent it is able to do so) a royalty-free, non-transferable, non-exclusive, licence to use its Background for the sole purpose of the performance of the Project.
4.3 Results shall vest and be owned as follows:
4.3.1 To the extent that any of the Results are generated or developed by the Company (EDF Energy) alone, without Imperial's intellectual contribution, then they shall vest in and be owned absolutely by the Company;
4.3.2 To the extent that any of the Results are generated or developed by the Company jointly with Imperial, then they shall vest in and be owned jointly by the Company and Imperial, in accordance with each Party's inventive contribution towards such Result(s);
4.3.3 To the extent that any of the Results are generated or developed by Imperial, without the Company's intellectual contribution, then they shall vest in and be owned absolutely by Imperial.
4.4 In the event that it is or may be possible to obtain any registered Intellectual Property Rights in any jointly owned Results, the Company agrees to be responsible for the filing and prosecution on behalf of the Parties and in their joint names of applications for registration, and the maintenance and renewal of any registrations, in such countries as the Parties agree in writing, subject to Imperial co-operating in the provision of all necessary assistance, information and instructions, with respect to the same, provided that:
4.4.1 if Imperial but not the Company wishes to apply for registration in any country or countries, Imperial may do so at its sole cost and expense on behalf of both Parties and in their joint names, and the Company shall provide Imperial with all necessary assistance,

information, and instruction;

4.4.2 neither Party shall amend or abandon any registration in respect of which the Parties are jointly registered as owners unless, in the case of an amendment, the other Party shall have given its prior written consent or, in the case of abandonment, the other Party shall be given the opportunity to maintain the registration at its own cost.

4.4.3 the Party making an application for registration shall consult with the other Party at reasonable intervals concerning the application for and maintenance of such registration.

4.5 Imperial hereby grants to the Company and its Affiliates, a royalty-free, irrevocable, non-transferable, non-exclusive, right and licence to use its Results for the sole purpose of internal research and development.

4.6 The Company hereby grants to Imperial a royalty-free irrevocable, non-transferable, non-exclusive licence to use its Results for their own non-commercial activities such as teaching and scientific or clinical research.

4.7 Imperial hereby grants the Company the following options:

4.7.1 In respect of any Results which are jointly owned, pursuant to Clause 4.3.2, the option to take an exclusive licence to the Intellectual Property Rights of Imperial.

4.7.2 In respect of any Results owned by Imperial, pursuant to Clause 4.3.3, the option to take either a non-exclusive or exclusive licence, which may or may not be restricted by field. The Company may exercise the above options at any time during the three (3) months following the date on which the Company is first notified of the existence of the Results (or by such other date as Imperial and the Company may agree) by serving on Imperial a written notice (an “Option Notice”), in which case such exercise shall be effective on the date such notice is sent by the Company.

4.8 Following Imperial’s receipt of an Option Notice, Imperial and the Company will negotiate in good faith for a period of up to ninety (90) days after Imperial gives written acknowledgement of receipt of the Option Notice (the “Negotiation Period”), the terms of any licence agreement provided for in Clause 4.7 above (the “Licence”). The Licence may be granted by Imperial’s technology transfer partner, Imperial College Innovations Limited, and shall contain all such terms and conditions which are usual and customary in a licence agreement, including but not limited to liability, audit provisions, termination, remuneration and royalties and governing law provisions. The financial terms of any licence will be fair and reasonable in the circumstances and will be negotiated on a case-by-case basis taking into account the scientific and financial contributions of the Parties to the Results being licensed and the subsequent scientific and financial contribution of the Parties that will be necessary to commercially exploit such Results. If the Parties are not able to agree the terms of the Licence within the Negotiation Period, the Company’s rights under clauses 4.7.1 and 4.7.2 will lapse.

4.9 In the event that the Company does not exercise its option pursuant to Clause 4.7.1, Imperial and the Company may themselves separately commercially exploit or non-exclusively sub-license jointly owned Results to third parties with the other joint owner’s consent, such consent not to be unreasonably withheld or delayed.

Data Storage and Security

How much data do you expect to generate?

- > 10 TB

How will you store and back-up your data during the project?

Data will be stored using the Imperial College Research Data Store which offers RAID redundancy and options for long-term data archiving where required.

Long-term back ups of deliverables documents will be stored via Microsoft Teams cloud storage.

Software will be backed up on cloud services such as GitHub.

The project team will liase with EDF data management team where it is advantageous for EDF to host data locally.

How will you manage access and security?

Any personal or highly sensitive data stored on College servers will be encrypted and only accessed by authorised members of the project.

Data Documentation and Metadata

How will the data be documented to ensure it can be understood?

All OpenInSAR products will be produced with attached meta data. The software will inherently track and update this metadata when loading in new data and performing individual data processing steps, including software version information. As such it is intended that each OpenInSAR product will provide a complete data lineage for re-use and reproducibility.

In accordance with schedule 4 of the research project contract, the following good data management practices will be adhered to:

- Research data will be accurately recorded in accordance with good scientific practices by the people conducting the research.
- Data trails will be kept to allow demonstration and reconstruction of key decisions made during the research, presentations, and conclusions.

In addition:

- All datasets will be thoroughly annotated with metadata to provide users with all necessary details on the origin or manipulation of the data in order to favour and facilitate re-use and reproducibility.
- Published data and other deliverables will be tracked in a Register of Deliverables, which takes the form of a '.xlsx' spreadsheet which will be versioned and saved via Microsoft Teams cloud storage.

Will you be using any domain specific or widely used metadata standards to describe your data?

- Yes (please give details below)

Metadata for OpenInSAR products will adhere to ISO standard 19115-1:2014: Geographic information - Metadata, where relevant.

Data Preservation and Sharing

What are your plans for long-term preservation and data sharing?

OpenInSAR source code will be stored and distributed via Github. Major software versions will be issued a Digital Object Identifier (DOI) via online services such as Zenodo.

The project team will share key data and outcomes directly with EDF and, where possible, via publication in open access journals.

Data requiring long-term preservation will be identified during meetings with the project team and its sponsors. Online public research repositories such as Zenodo can be used to publish data while Imperial College offers an archiving service for data sets that are large or where public access is not desirable.

Will there be any restrictions on accessing the data?

- Yes (please give details)

Data that cannot be shared publicly will remain secured upon College or EDF servers as appropriate. In general, no restrictions have been identified with OpenInSAR products as these are derived from existing, publicly available information. EDF may however provide protectively marker information to the project team, which will be handled securely and according to EDF policies.

How will potential users find out about your data?

The primary user of the data is EDF Energy. They will be made aware of the data through the project deliverables:

- Monthly progress meetings.
- Quarterly results presentations/workshops.
- Written reports.

All published papers will include a data access statement providing details of where the dataset can be found and under what conditions it can be accessed.

Software and published data will be promoted via stakeholder email networks, conferences, project web pages, and social media.

Responsibilities and Resources

Who is responsible for implementing this plan?

Stewart Agar will have responsibility for study-wide data management, metadata creation, data security and quality assurance of data.

Will you require any additional resources to deliver this plan?

Resources for data hosting via College systems have been costed into the project. Imperial College Research Data Storage provides flexible capabilities to meet this projects needs.