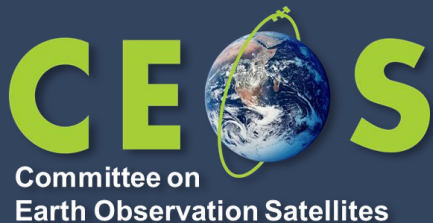# Software preservation White Paper

I.Maggio, Starion for ESA
Agenda Item 5.3
WGISS-58
16-17 October 2024
Sioux Falls, South Dakota, USA

CEOS

Committee on
Earth Observation Satellites

Number	Title	Status	Creation year	Completion date
D DATA-24-01	White Paper on EO Data collections management and governance	open	2024	2025 Q1
D DATA-24-02	White paper on Software preservation	open	2024	2025 Q3

View Deliverable
← Back

D DATA-24-02: White paper on Software preservation

✉ Share
🔗 Permalink

Number:	DATA-24-02
Title:	White paper on Software preservation
Status:	Open
Creation Year:	2024
Completion Date:	2025 Q3 (361 days remaining)
Description:	

> This white paper would address and recommend techniques to ensure preservation and reusability of software tools related to EO missions (e.g. processors, exploitation and visualization tools).

Link to GEO Work Programme:	
External Reference:	
Responsible Users:	Mirko Albani
Responsible CEOS Entities:	WGISS
Contributing Agencies:	WGISS Member Agencies
Relevant Sustainable Development Goal (SDG):	
Progress Reports:	
Comments:	
Actions:	
Created:	2024-02-26 13:47:40 UTC
Last Updated:	2024-02-26 13:47:57 UTC

This presentation provides a starting point for the discussion and for a dedicated session to be organised at WGISS#59

The white paper should be intended to *assist data/software managers* in the field of Earth observation (EO) with the task of ensuring the long-term preservation of EO missions and data related software, thus improving data accessibility and usability for current and potential future users.

The intended audience should comprise data and software providers, decision makers and scientists, and data managers/stewards for data centres and repositories.

It will introduce the concept of software preservation, outlining its importance and justifying the importance of expending effort to ensure effective preservation of EO software.

Details of the main principles of software preservation will be provided, as well as brief descriptions of the primary strategies that may be implemented by data managers, together with challenges.

WHY TO PRESERVE SOFTWARE

- There are several reasons that justify expending the effort required to preserve software, foremost of which is the simple fact that *since data is being preserved, any associated software should be preserved alongside it to maintain the data's maximum value.*

- Having access to the original software is crucial for any potential reanalysis/reproduction of earlier work, ensuring that *researchers can understand how data was processed in the past*, thus enabling meaningful comparisons with modern data analysis.

- Data processing software provides valuable historical context for scientific data as it reflects the analysis methods and technologies used over time.

Collect
Software, documents, knowledge

Preserve
Robust, complete, redundancy

Share
Accessible, open, persistant

These three principles, which correspond to the different phases of the software preservation process, help to guarantee the robust preservation of software while maximising benefit to, and usage by, potential future users of the software.
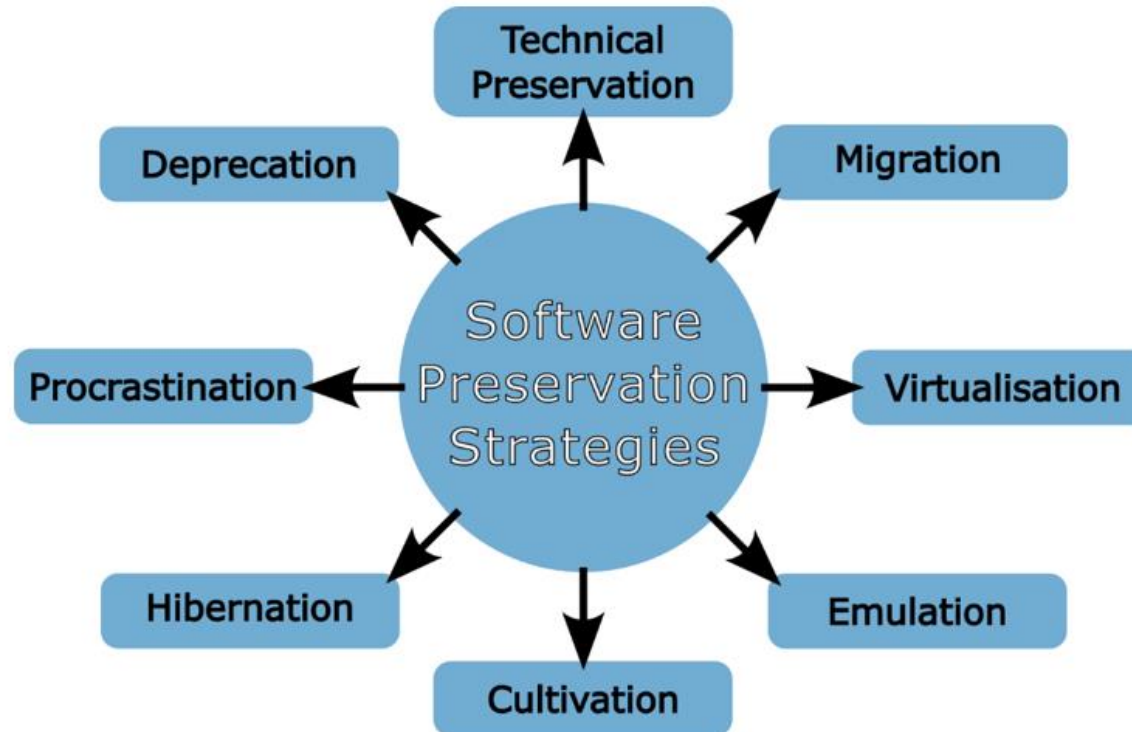
A wide array of *challenges face software preservation*, from technical complications to legal issues and resource availability.

A well planned and proactive approach to software preservation may help to avoid some of the most encountered challenges, particularly for larger, well-resourced organisations.

- lack of documentation and expert knowledge represent a challenge for the preservation process *(mitigated by early adoption of a preservation strategy that ensures an adequate documentation and expert knowledge collection)*.
- preserving the full functionality of the software may be difficult as data managers/preservationists may not be familiar enough with the entire software to accurately assess whether all aspects are functioning as designed.
- risk of partial or full obsolescence prior to beginning the preservation process poses a critical treat that is particularly pronounced for software that depends on any specialised hardware *(mitigated by a good preservation planning)*.

The need to avert these risks further reinforces the importance of planning the preservation activities as soon as possible so that software and hardware may be preserved while they are still available.

- intellectual property (IP) rights which may prevent the distribution of the software without first receiving explicit permission from the rights holder.

- software preservation often requires the manipulation of the software in a manner which may not be permitted by the licence agreements.

- resource availability, both in terms of personnel and finances.

- software preservation activities can involve expenses, particular if employing a strategy of technical preservation for software that relies on specialist hardware that must be purchased and preserved.

- even if all the required expertise are available, preservation is still a time-consuming process that requires significant effort and commitment from the personnel involved.

WAY FORWARD

A dedicated session will be organized at WGISS#59 to share experiences and technical information to be used to draft the relevant CEOS White Paper