Librarians to the Rescue! An LIS-Associated Workflow Model for the Rescue of Research Data at Risk

Louise Patterton
https://orcid.org/0000-0002-8067-8545
University of Pretoria, South Africa
lpatterton@csir.co.za

Theo J. D. Bothma
https://orcid.org/0000-0001-7850-3263
University of Pretoria, South Africa
theo.bothma@up.ac.za

Martie J. van Deventer
https://orcid.org/0000-0002-9776-1177
University of Pretoria, South Africa
mvandeve2017@gmail.com

Abstract

This study presents a data rescue workflow model and explores the potential roles and responsibilities of the library and information services (LIS) sector in rescuing at-risk data. The methodology was multifaceted, consisting of three distinct phases of data collection and analysis. First, content analysis was conducted on 15 purposively selected data rescue publications to identify key activities that are integral to the data rescue process. This phase resulted in the creation of an initial data rescue workflow model. Second, feedback was obtained from six experienced researchers at a South African research institute, chosen from a sample of criterion-based researchers, after they reviewed the initial model. This feedback was used to revise the model. Third, additional feedback on both the initial and revised models was gathered during a mini focus group session with three purposively selected LIS experts at the same research institute. This feedback was used to refine the model and develop a final recommended data rescue workflow model, which is presented in this study. The recommended model outlines the crucial stages, activities, and outputs of data rescue in a chronological sequence. The study's outcomes provide new insights into data rescue and the involvement of LIS professionals, highlighting how various LIS experts can contribute to data rescue projects.

Keywords: data conservation; data rescue; data at risk; research data management; data curation

UNISA
Mousaion
Volume 43 | Number 2 | 2025 | #17232 | 29 pages
https://doi.org/10.25159/0027-2639/17232
ISSN 2663-659X (Online), ISSN 0027-2639 (Print)

Introduction

Data is fundamental to scientific research and discovery, enabling researchers to test hypotheses, gain insights, validate concepts, make new discoveries, and accelerate innovation (Chaudhary 2023). Access to historical data is essential across many research fields: past atmospheric data helps map ozone concentrations in Earth's stratosphere; photographs of glaciers from different time periods provide evidence of climate change; 70-year-old medical records on punch cards reveal how varying cholesterol levels predict future diseases; and stream data from the 1950s aids hydrologists in understanding the impact of non-native tree species on water distribution across landscapes (Griffin 2017). Unfortunately, older data are often at risk of being discarded, damaged, stored in outdated formats, made inaccessible due to obsolete readers, or lost due to poor data management practices. It is, therefore, crucial to rescue and preserve these valuable, often irreplaceable data before they are lost.

Data rescue is described by the World Meteorological Organization (WMO) as the "ongoing process of preserving all data at risk of being lost" (2020). It entails the recovery of "older datasets and metadata from the trappings of their existing media or format" (Research Data Alliance [RDA] 2017b). According to DataFirst (n.d.), "at risk" data are data that are deteriorating, data that are lacking explanatory documentation, or data that are in non-digital formats; data rescue involves identifying such threatened data and capturing the data in digital format for long-term preservation and reuse.

The entire data rescue process typically involves the following steps, in chronological order: cataloguing and appropriately storing paper media to minimise deterioration (WMO 2020); digitising data into a computer-compatible format (WMO 2020); ensuring the availability of necessary metadata, which provides additional context for the data (Griffin 2017); uploading data into international repositories to ensure proper storage, management, and accessibility for sharing, analysis, and reporting (Copernicus Climate Change Service [C3S] 2021); and continuously updating digital records to ensure compatibility with current media and formats (WMO 2020).

Data rescue projects portray diverse features and characteristics; data rescue ventures range from climate sciences to religious studies (Brönnimann et al. 2006; European Space Agency 2018), historic paper data to modern digital data (Antico, Aguiar, and Amsler 2018; Cascone 2022), data at risk from the bottom of the ocean to data at risk from the surface of the moon (Hsu et al. 2013; Modine 2008), with participating parties extending from researchers to library and information services (LIS) professionals to citizen scientists (Carey 2017; Global Surface Air Temperature [GLoSAT] n.d.). The benefits of these data rescue projects include continued access to valuable and unique at-risk data (Griffin 2015; Wippich 2012), historical data being essential in many fields of science (Bradshaw, Rickards, and Aarup 2015), extending the knowledge of the subject (Gallaher et al. 2015), an increase in scientific accuracy (Eveleth 2014), and cost-effectiveness, as data rescue is less expensive resource-wise than regenerating the data (Diviacco et al. 2015).

Data rescue activities share significant overlap with typical research data management tasks, such as repository uploads, metadata quality control, digitisation efforts, and advising on data storage and preservation. Globally, many academic and research institutions have established data management divisions, often integrated within the institution's library and information science department (Perrier, Blondal, and MacDonald 2018; Sheikh, Malik, and Adnan 2023). Given that the LIS sector already performs key data management tasks found in data rescue projects, investigating its current and future involvement in these activities is a logical next step. Therefore, this article examines the global involvement of the LIS sector in data rescue projects. Additionally, the study aims to gather information on published data rescue projects and guidelines, leading to the development of a data rescue workflow model. Based on these insights, the article also outlines the potential roles and responsibilities of the LIS sector during a data rescue project.

Literature Review

Published data rescue literature reveals that the LIS sector is involved in a range of rescue activities (Eke 2017; Kijas 2018; Hoffman et al. 2020); these activities are summarised in Table 1. Data rescue entails identifying, digitising, preserving, and sharing valuable data and associated metadata at risk of loss, and the table below reports on locating data, assessing data, uploading to repositories, adding metadata to datasets, and ensuring the long-term accessibility of data.

Table 1: LIS sector involvement in data rescue

Data rescue activity	Reference
Different levels of LIS involvement in data rescue: • Surveying researchers to determine data needs, and raising data rescue awareness • Deep web archiving • Rescuing data needed, valued, and prioritised by the research community • Organising and managing a data rescue event	Eke 2017
Major involvement in data rescue projects	Kijas 2018; Thompson 2017;
Locate data	Thompson 2017
Long-term storage, curation, and preservation of data	Thompson, Davenport Robertson, and Greenberg 2014; WMO 2016
Use of libraries when creating data inventories	WMO 2016
Assess data and determine potential data users	Palmer, Weber, and Cragin 2011; Hoffman et al. 2020
Upload data to repositories; share data	Hoffman et al. 2020; Kijas 2018; Thompson 2017
Creation of data rescue material and publications	Hoffman et al. 2020; Thompson 2017
Membership of data rescue groups	Hoffman et al. 2020; Thompson 2017
Involvement in environmental scanning of data at risk/data rescue projects	Kijas 2018; Thompson, Davenport Robertson, and Greenberg 2014
Provide data rescue training	Thompson, Davenport Robertson, and Greenberg 2014
Host data rescue events	Eke 2017; Kijas 2018; Raughley 2017

Published outputs provide evidence of the LIS sector's current involvement in data rescue globally, but in the absence of a data rescue workflow model or data rescue activity flowchart it is not clear how LIS professionals are involved throughout the entire rescue project, and how the sector plays a part during the data rescue workflow. An evolving next step is to present to the LIS sector a recommended and more formalised data rescue model indicating the various stages, activities, decisions, and outputs forming part of the process. Access to a model will provide libraries and LIS professionals with a better understanding of data rescue and enable the identification of potential and realistic roles and responsibilities linked to their skills, experience, and staff complement.

This study also investigated literature that would lead to the creation of a data rescue workflow model. These publications are summarised and discussed in the next section, as the 15 data rescue publications selected underwent a process of content analysis

during this study to determine mandatory data rescue stages, crucial data rescue activities, and the chronological flow of vital components of a rescue project.

Methods and Results

Three different data collection methods were used in this study, and each contributed in a sequential manner towards the creation of a data rescue workflow model. The methods (see Figure 1) comprise:

- selection and content analysis of 15 data rescue publications leading to the creation of an initial data rescue workflow model (see Figure 2),
- feedback, suggestions, and critique provided to the researchers via email after the initial model was shared with researchers; the feedback was used to revise the initial model, and
- feedback, suggestions, and critique on the revised model collected from three LIS experts during a mini focus group session; the feedback was used to create the final recommended data rescue workflow model (see Figures 3, 4, 5, 6, and 7).

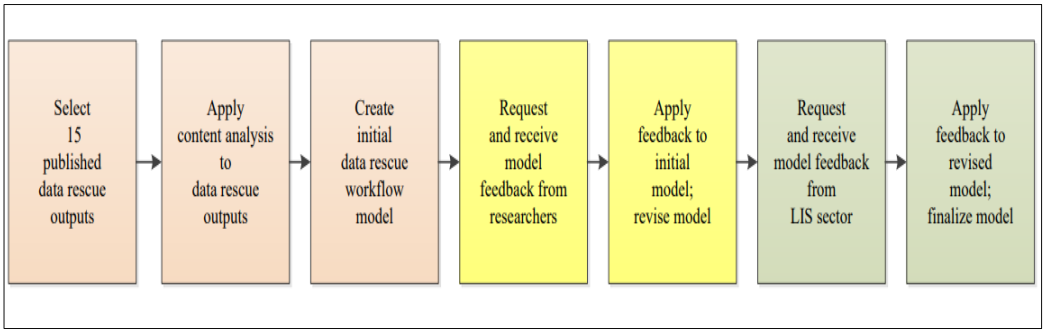

Figure 1: Summarised research process of this study

Content Analysis and Derived Results

Fifteen data rescue publications were purposively selected based on the completeness of their descriptions of data rescue steps. These publications were identified through Google Scholar and chosen for the availability of detailed descriptions of data rescue steps, stages, and guidelines. The date range was incidental to the selection process. The data that were rescued were historical in nature, and detailed information about the original data collection methods, conditions, and circumstances (i.e., metadata) was not necessarily included in each data rescue publication.

Content analysis, a methodology applicable to various types of texts such as written speeches, newspaper articles, documents, reports, digital media, pictures, and audio-visual content (Costa 2022), was applied to this sample of publications to identify the

crucial components of a proposed data rescue workflow model. The components to be identified included activities, stages, outputs, and inventories associated with data rescue. This method of analysis provides researchers with valuable insights into prevailing themes, patterns, and trends (Ashikuzzaman 2018) and was considered the most suitable approach for determining data rescue activities, stages, role players, stakeholders, outputs, skills, and other key features, ultimately leading to the development of a data rescue model.

Content Analysis Sample

Table 2 illustrates the main details of the 15 data rescue publications selected for content analysis:

Table 2: Data rescue publications used for content analysis

Author and date	Rescue topic	Publication	Data format	Discipline
Brönnimann et al. 2006.	A guide for digitising manuscript climate data	Article	Paper	Climate sciences
Griffin 2006	Rescuing and recovering lost or endangered data	Article	"Non-digital" data	Discipline-agnostic
Antuña et al. 2008	Solar radiation data rescue at Camagüey, Cuba	Article	Paper	Solar radiation
Güntsch et al. 2012	Efficient rescue of threatened biodiversity data using reBiND workflows	Article	Paper and early digital data	Biodiversity
International Environmental Data Rescue Organization (IEDRO) 2014	Description of rescue steps used by the organisation	Web page	Mostly paper	Climate sciences
WMO 2014	Guidelines for hydrological data rescue	Online manual	Mostly paper	Hydrology

Diviacco et al. 2015	Data rescue to extend the value of vintage seismic data: The OGS-SNAP experience	Article	Paper data, magnetic tape data	Seismology
Wiebe and Allison 2015	Rescue of dark zooplankton data	Article	Paper data, old digital formats	Zooplankton
Mavraki et al. 2016	Rescuing biogeographic legacy data: The Thor Expedition	Article	Published historical data	Biogeographic
DataFirst n.d.; Woolfrey 2016	The Rescue of "at risk" data on forced resettlement in South Africa	Workshop presentation, also website	Paper data	Sociological data
WMO 2016	Guidelines on best practices for climate data rescue	Online manual	Mostly paper	Climate sciences
Downs and Chen 2017	Curation of scientific data at risk of loss	Book chapter	Modern digital data	Environmental and socio-economic data
Kennedy 2017	Guidelines for marine species occurrence data rescue	Online manual	Early digital and digital formats	Marine species
Ryan et al. 2017	Integrating data rescue into the classroom	Article, also poster	Transcription of electronic sheets	Precipitation
Research Data Alliance (RDA) 2017a	Data rescue guidelines of the RDA DATA Rescue Interest Group	Website	Mostly paper	Discipline-agnostic

Content Analysis Steps

The selected publications were scrutinised and a flexible coding framework, which allowed for the addition of categories or concepts throughout the analysis process, was developed. Framework categories were the following:

- data rescue steps common to the majority of rescue projects,
- data rescue steps being the exception,
- data rescue steps identified as mandatory activities,
- outstanding/pivotal data rescue features detected,

- pre-rescue activities,
- post-rescue activities,
- role of data formats,
- role of research discipline,
- data rescue role-players, participants, and stakeholders,
- the concept of data rescue teams,
- the role of specific skills/expertise, and
- outputs/deliverables created during rescue.

The text was coded and coded segments were added to a group of similar segments. Categories were merged, and the final narrative, comprising stages, activities, and outputs crucial for inclusion in the initial data rescue workflow model, was drafted.

Initial Data Rescue Workflow Model

Nine main rescue activities, forming the workflow model's main rescue stages, formed the backbone of the initial model. The stages were:

1. a project initiation phase, including data assessment, appointment of a data rescue project team, and creation of a data rescue project plan as well as a data management team,
2. storage and preservation of data at risk,
3. creation of data inventories,
4. imaging of paper-based data at risk,
5. digitisation of paper-based data at risk,
6. creation of metadata and data documentation,
7. sharing of digitised data to repositories,
8. long-term preservation of rescued (digitised) data, and
9. project closure.

A diagrammatical representation of the initial model, based on content analysis findings, was drafted and is shown in Figure 2.

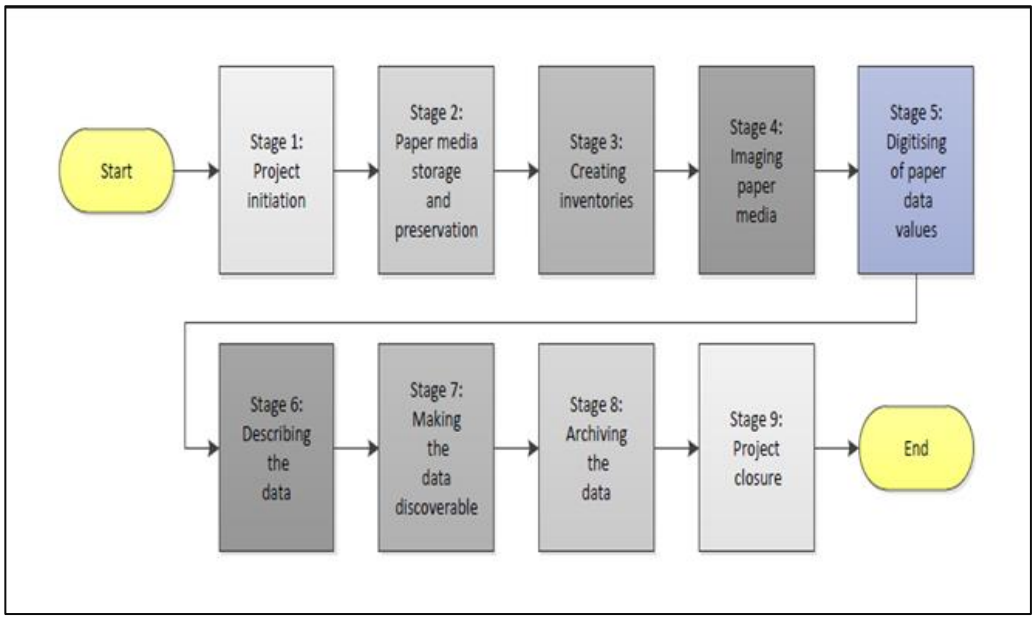

Figure 2: Summary of initial data rescue workflow model

This initial version of the data rescue workflow model is regarded as an antecedent to the eventual recommended data rescue workflow model illustrated in Figures 3, 4, 5, 6, and 7.

Research Sector Feedback and Derived Results

Feedback on the initial model was requested from a sample of eight experienced researchers based at a leading South African research institute. The sample was purposively selected, with all researchers in the position of "Research Group Leader" at the institute being invited to participate in a study on data at risk and data rescue. The request for feedback followed the completion of virtual one-on-one interviews with eight respondents, where they discussed their data rescue experiences and challenges. These interviews, along with the request for feedback, were part of a larger study on data at risk and data rescue (Patterton 2023). The interviewed researchers were provided with a copy of the model and asked to provide feedback via email on its usability for rescuing data at risk within their research groups. Feedback was received from six of the eight sample members.

Feedback

Researcher feedback indicated that the issues listed below were areas of concern.

- The model was described as complex.
- Researchers wanted the LIS sector to be involved in the activities shown in the model.

- Researchers recommended that the model include more formats than paper-based data only.
- Researchers recommended that the model should provide guidelines accompanying the various activities.
- Researchers stated that it was unclear whether all or only certain stages/activities were to be performed during the data rescue process.
- Researchers were unsure whether the model would also address the rescue of confidential data.

Based on the feedback received from researchers, the initial model was revised. The revised model is described in the next section.

Feedback and Model Revision

Besides expert researcher feedback, the authors also realised that the following issues would need to be addressed when revising the initial model:

- a data rescue model would need to cater for full rescue as well as partial rescue, as some rescue projects would be unable to implement all rescue stages and activities;
- the model required a descriptive summary;
- the model required a data assessment activity;
- the model required a data destruction activity; and
- the use of assorted colours and shapes ought to be considered.

Incorporating feedback and the above-listed realisations into the rescue model resulted in the creation of a revised model (see Figure 3). Despite various model updates, the feedback received from the sample of experienced researchers was limited and did not provide sufficient critical input to extensively revise the model. While the summarised revised model would be retained as a single-page view of the recommended model, it became clear that additional stage-specific information needed to be added to the model. It was therefore decided to approach the LIS-sector to obtain additional and novel feedback regarding the components of a usable data rescue workflow model.

LIS Sector Feedback and Derived Results

Three LIS experts, based at the same research institute and having a combined LIS-related experience exceeding five decades, participated in a mini focus group session where the initial and revised models were discussed. This methodological stage focused on the LIS sector as it was anticipated that this institutional division would provide a novel and unique perspective on data rescue and a data rescue model. Criterion sampling was used to select LIS experts; the criterion was research library experience, and in particular experience regarding collaborative research library activities that could form part of a data rescue process and workflow. According to Dudovskiy (2019), this purposive sampling method may prove to be effective when only limited numbers of

people can serve as primary data sources due to the nature of the research design and aims and objectives.

Feedback

Data rescue model feedback and critique obtained from the LIS sector via the focus group session was plentiful and evident of in-depth scrutiny. The table below contains the issues flagged by LIS experts as areas of concern.

Table 3: Feedback from LIS experts

Model category	Feedback
Compact summary	• A single-page summary is crucial to newcomers. • Verbs instead of nouns should be used to indicate stage names (i.e., "Describe the data" instead of "Data documentation"). This comment is applicable to the entire model. • Stage names should ideally be self-explanatory. • The number of activity blocks within the stage should be limited. • A glossary of applicable terms and abbreviations should be included. • The model should include the treatment of confidential data. • Guidance regarding data inventories should be included.
Extended summary	The importance of data assessment should be stated.
Stage 1	• Each image should be accompanied by a summary. • The importance of metadata when assessing data should be stated. • A checklist to accompany data assessment should be included. • Data assessment should involve discipline-specific researchers. • The storage location of pre-rescued data should be stated. • Inventory responsibility should be stated.
Stage 2	Guidelines regarding team selection should be included.
Stage 6	• Ensure that guidance is generic and not institute-specific. • LIS indexers should probably be the main party during this stage. • Consider steps to ensure FAIR data practices.
Stage 7	• Moving data to a preservation format can also be performed earlier. • Regular monitoring and annual testing of data access and download are vital but are archival tasks, not a data rescue task. • The guidelines should contain information on zipping of data.
Stage 8	• The promotion of data depends on the type of data and data ownership; guidelines are crucial. • The stage should be simplified as it currently contains duplicate activities.

LIS sector feedback was used to revise the model. The revisions resulted in a workflow model henceforth referred to as the recommended data rescue workflow model.

Recommended Data Rescue Workflow Model

The main characteristics of the recommended data rescue workflow model are listed below.

- The data rescue model is multi-layered as this approach enables users of the model to view a summarised single-page overview of the model (shown in Figure 3) or a more detailed yet still compact summary of the various rescue stages (shown in Figures 4, 5, 6, and 7).
- The summarised revised data rescue model (see Section "Feedback and Model Revision") was retained as a single-page portrayal of the model.
- The model is a generic data rescue model and is neither focused on nor excluding any research discipline.
- Diagrams and flowcharts are used to illustrate the recommended model.
- The recommended workflow consists of eight main stages, namely:

1. preparing for data rescue,
2. planning for data rescue,
3. storing and preserving the data,
4. digitising or converting the data,
5. describing the data,
6. sharing the data,
7. preserving the data (in the long term), and
8. project closure

- The model differentiates between full data rescue and partial data rescue. The terms "full data rescue" and "partial data rescue" were constructed during this study to differentiate between a rescue project containing all proposed data rescue stages and a rescue project where select stages of the data rescue workflow are implemented.
- The model contains steps for four different data formats: paper data, early digital data, modern digital data, and samples.

Data Rescue Summary Diagram

This single-page diagram is the first layer of the data rescue workflow model and provides a synopsis of the main data rescue stages and activities of the recommended data rescue workflow model. As an outline, its main objective is the single-page portrayal of data rescue to rescue novices.

The diagram includes the main stages and crucial activities but excludes aspects such as roles and responsibilities, references to guidance documents, references to all outputs created during the various stages, and data rescue decisions relevant to specific stages.

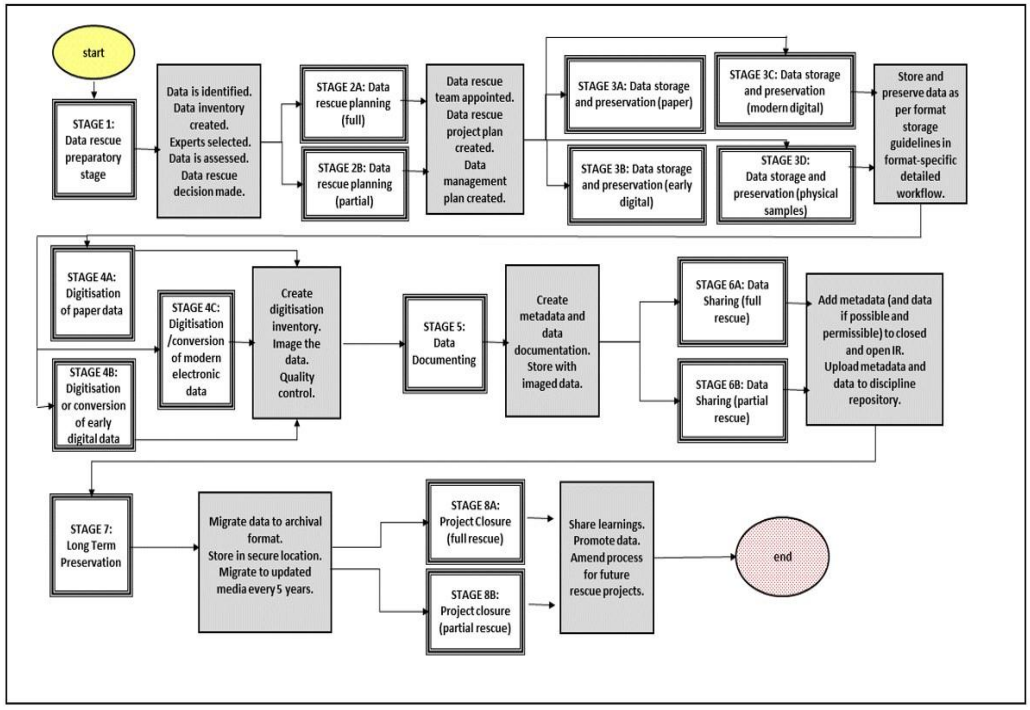

Figure 3: Summary of recommended data rescue workflow model

Data Rescue Compact Workflow (Extended Summary)

The data rescue compact workflow portrays an extended summary of the data rescue workflow model. The workflow comprises eight main stages; the stages are identical to the eight main stages listed in the single-page summary diagram. This layer provides more information than the previous section's summary diagram; additional details include the data formats involved, whether the stage involves full rescue or partial rescue (i.e., Stream A or B), and important decisions forming part of specific rescue stages.

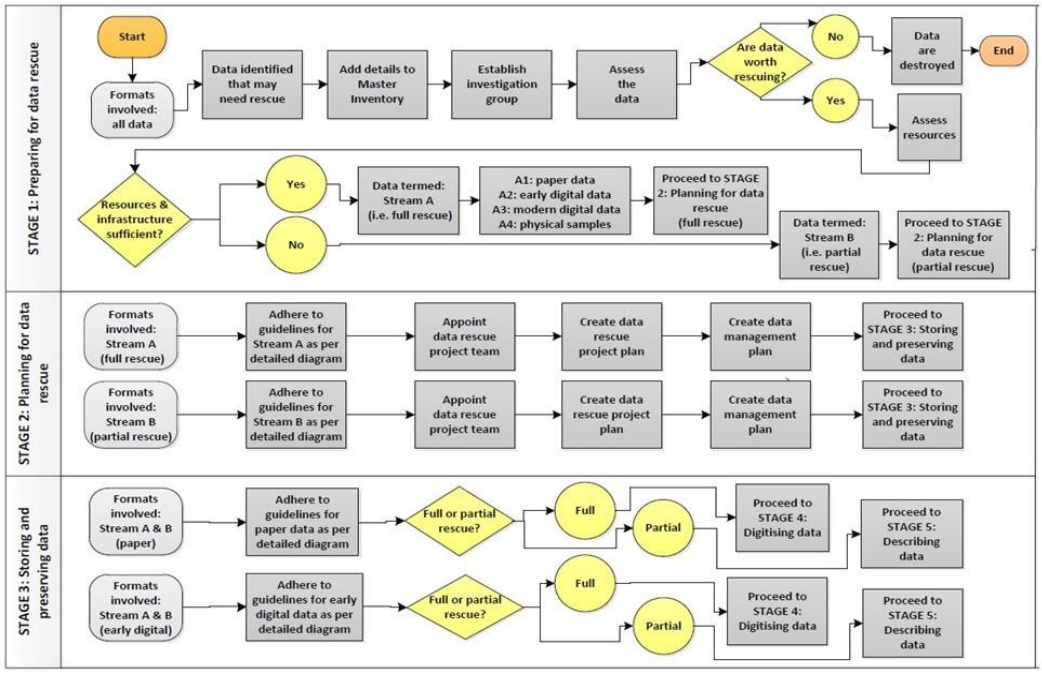

Figure 4: Data rescue stages 1–3

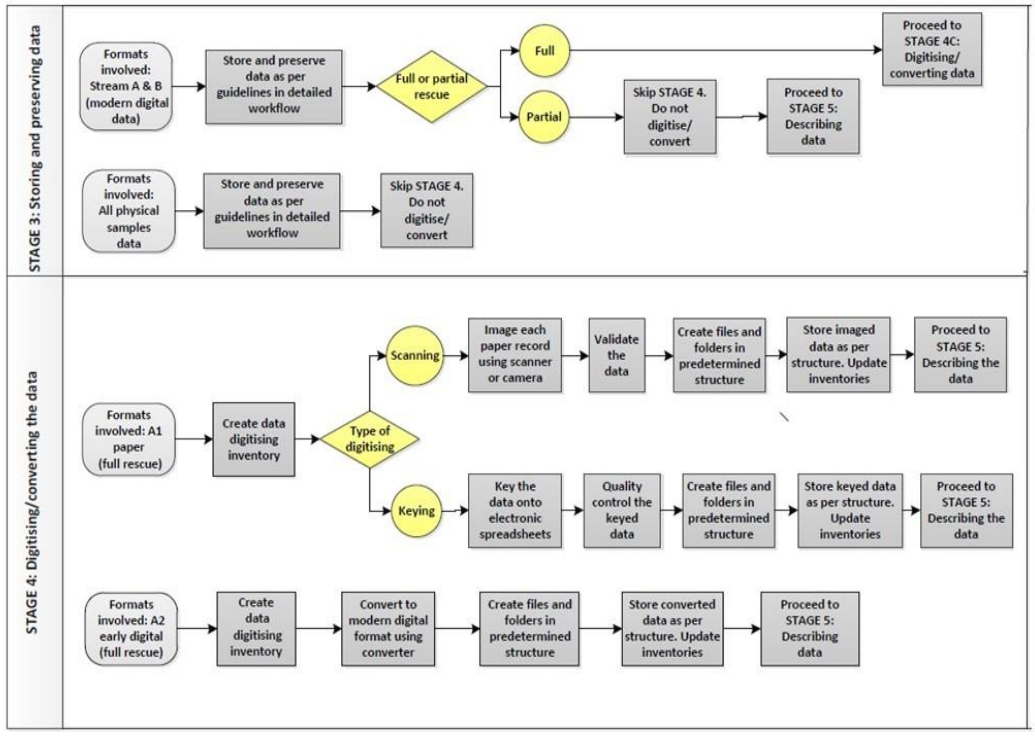

Figure 5: Data rescue stages 3–4

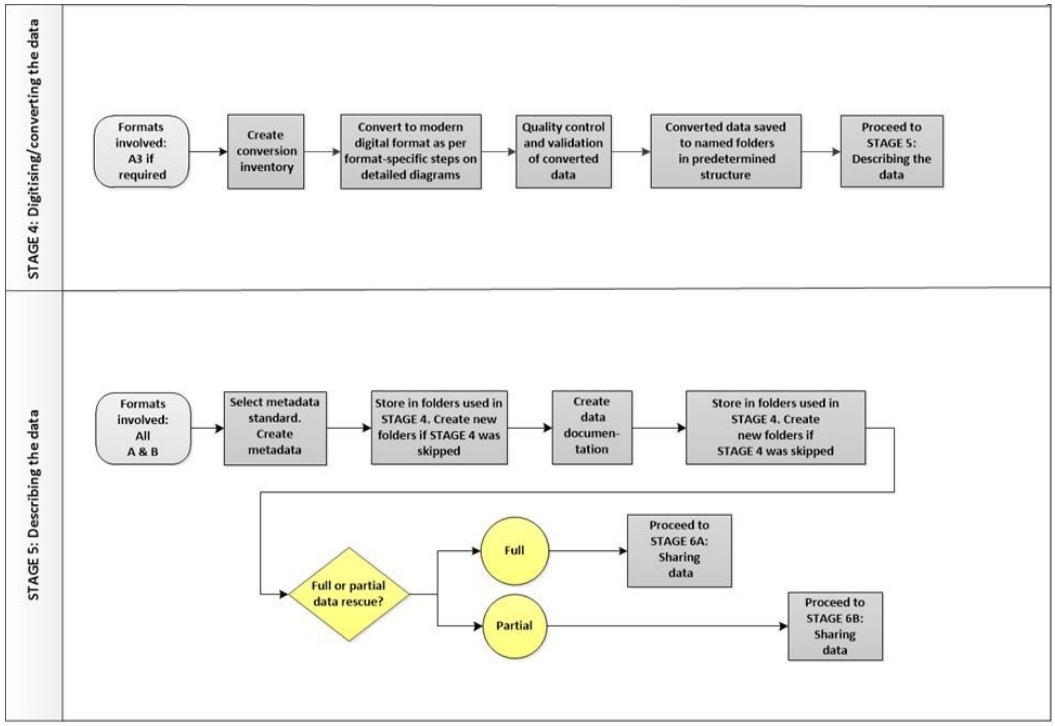

Figure 6: Data rescue stages 4–5

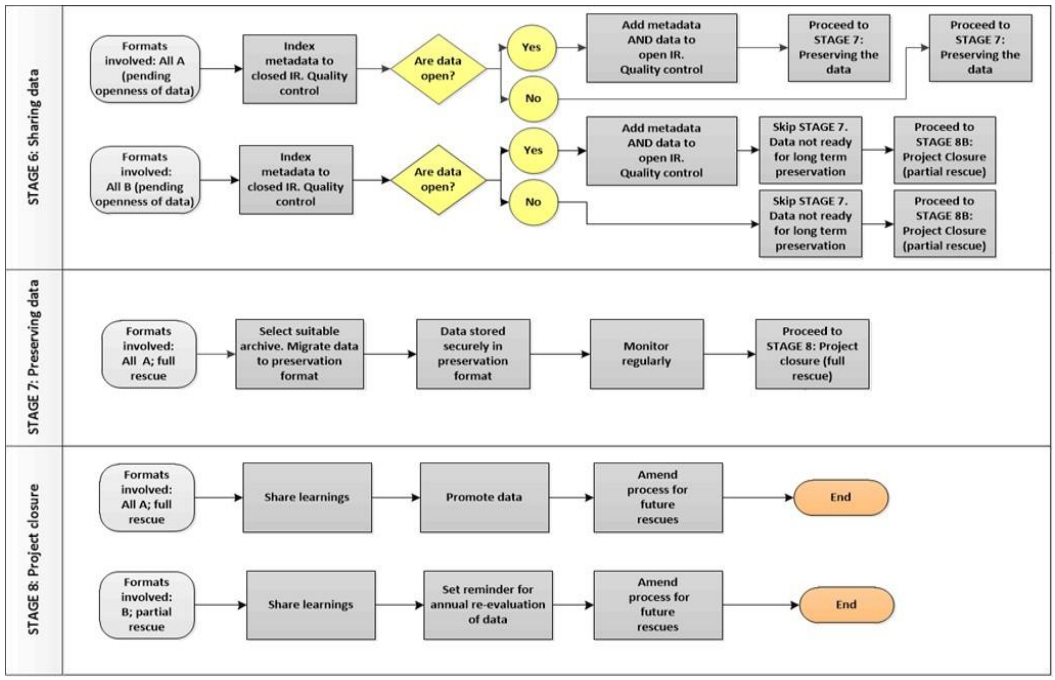

Figure 7: Data rescue stages 6–8

Discussion

While the information contained within the diagrams is mostly self-explanatory, it is not immediately evident how the LIS sector can be involved with the model, or with data rescue. This section details the potential involvement of the LIS sector and its various professional and semi-professional positions during a generic data rescue project.

Roles and responsibilities of the LIS sector when implementing the data rescue model are indicated in Table 4, and it is evident that the sector can potentially be involved in all stages of the model. “Roles” refer to the LIS professional’s position on a data rescue team, while “responsibilities” refer to the tasks and duties of their role during a rescue project. The LIS positions included in the table are not fixed in stone and are based on the LIS positions of the research library where the first author is based. Similarly, the roles and responsibilities described emanate from the current duties performed by the LIS sector at the research library where the first author is based. The extent to which data rescue projects will involve the LIS sector depends on various factors: the availability of LIS professionals, the research discipline, and terminology used when referring to LIS positions. Finally, the LIS sector can also be involved in a supervisory position and as the creator of data rescue guidelines throughout the entire project and during each separate stage.

Table 4: Potential data rescue roles and responsibilities of the LIS sector

Data Rescue Stage	Role	Responsibility
Stage 1: Preparing for data rescue	**Trainer** *LIS data specialist*	• Trains participants • Drafts training material
	Data assessor *Information specialist* *LIS data specialist*	• Assists in assessing suitability of data for rescue, and value of data • Assessment done in collaboration with assessment team members
	Venue organiser *Archives technician*	• Ensures secure and sufficient space for pre-assessed data • Communicates with data custodian regarding delivery of data
	Documentation creator *LIS data specialist*	• Creates stage documentation • Adds relevant details to inventory
	Resources assessor *LIS data specialist*	• Assists in assessing the feasibility of the institute's data rescue resources
	Data destruction coordinator *LIS data specialist* *Records manager*	• Consults with ICT or Waste Services regarding data destruction
	Project manager *LIS data specialist* *Records manager*	• Plans activities and communicates with role-players • Sets deadlines • Oversees activities of this stage • Updates project report

Stage 2: Planning for data rescue	**Trainer** *LIS data specialist*	• Trains participants • Drafts training material
	Documentation creator *LIS data specialist*	• Creates stage documentation
	Data management plan (DMP) creator tasks *LIS data specialist*	• Assists with DMP creation • DMP editing, sharing, and storage
	Team member *Archives technician*	• Suitable tasks as indicated by project leader
	Project manager *LIS data specialist* *Records manager*	• Drafts the data rescue project plan • Manages the data rescue project team • Assigns team member roles and responsibilities • Takes responsibility for the project folder • Distributes timelines and meeting dates
Stage 3: Storing and preserving paper data	**Project manager** *LIS data specialist* *Records manager*	• Monitors rescue progress • Updates rescue progress report
	Trainer *LIS data specialist*	• Trains participants • Drafts training material
	Documentation creator *LIS data specialist*	• Creates stage documentation
	Archive cleaner/data cleaner *LIS data specialist* *Archive technician* *Intern*	• Removes dust from paper data, boxes, and shelves • Oversees dusting and vacuuming activities
	Data curator *Archive technician* *LIS data specialist* *Intern*	• Labels boxes and shelves • Stores paper data in boxes; stores other data formats securely
	Archivist *Records manager*	• Ensures archive meets environmental requirements • Ensures archive meets space requirements • Ensures fire-suppressant system is functional and serviced regularly

Stage 3: Storing and preserving early digital data	**Project manager** *LIS data specialist* *Records manager*	• Monitors progress • Updates the progress report
	Trainer *LIS data specialist*	• Trains participants
	Documentation creator *LIS data specialist*	• Creates stage documentation
	Format expert *Not yet determined*	• Inspects the data and data reader • Assists with the appropriate storage of data and reader • Locates metadata
	Data storage expert *Designated LIS professional or researcher*	• Stores pre-digitised data as per best practices
Stage 3: Storing and preserving modern data	**Trainer** *Format expert* *LIS data specialist*	• Trains participants
	Documentation creator *LIS data specialist* *Format expert*	• Creates stage documentation
	Project manager *LIS data specialist* *Records manager*	• Monitors progress • Updates progress report
	Data storage expert *Designated LIS professional or researcher*	• Stores early digital data as per best practices
Stage 3: Storing and preserving physical samples and specimens	**Trainer** *Format expert* *LIS data specialist*	• Trains participants • Drafts training material
	Documentation creator *LIS data specialist* *Format expert*	• Creates stage documentation

	Project manager *LIS data specialist* *Records manager*	• Monitors progress • Updates progress report
	Data storage expert *Designated research library or researcher*	• Stores samples and specimens as per best practices
Stage 4: Digitising/converting data	**Trainer** *Format expert* *LIS data specialist*	• Trains participants • Drafts training material
	Documentation creator *LIS data specialist* *Format expert*	• Creates stage documentation
	Digitiser *Digitisation clerk* *Archives technician* *LIS data specialist* *Intern*	• Creates digitisation inventory • Scans images or keys data • Converts early digital data to common, open, modern format
	Quality controller *LIS data specialist*	• Validates that all files have been scanned/keyed, converted • Quality controls the resultant data
	Project manager *LIS data specialist*	• Monitors progress • Updates progress report
Stage 5: Describing the data	**Trainer** *Format expert* *LIS data specialist*	• Trains participants • Drafts training material
	Documentation creator *LIS data specialist* *Format expert*	• Creates stage documentation
	Creator of metadata and related tasks *LIS data specialist* *Information specialist*	• Determines metadata standard required by repository • Advises on metadata creation • Advises on data documentation creation • Assists with metadata creation • Reviews metadata created by research base • Responsible for linked metadata and data documentation

	Project manager *LIS data specialist*	• Monitors progress • Updates progress report
Stage 6: Sharing the data	**Trainer** *LIS data specialist*	• Trains participants
	Documentation creator *LIS data specialist*	• Creates stage documentation • Drafts training material
	Repository professional *LIS data specialist* *Indexers*	• Uploads data to selected repositories • Uploads metadata and data documentation to selected repositories • Shares DOI with relevant parties
	Project manager *LIS data specialist*	• Monitors/supervises progress • Updates progress report
Stage 7: Preserving the data (for the long term)	**Trainer** *LIS data specialist* *ICT division*	• Trains participants • Drafts training material
	Documentation creator *LIS data specialist* *ICT division*	• Creates stage documentation
	Archive selector *LIS data specialist*	• Provides guidance regarding the requirements for secure and stable preservation location • Assists in selecting appropriate location
	ICT collaborator *Records manager* *LIS data specialist*	• Collaborates with ICT or entity selected for long-term preservation • Relays needs and requirements
	Project manager *LIS data specialist* *Records manager*	• Monitors progress • Updates progress report
	Data curator *LIS data specialist* *Records manager*	• Gives guidance about suitable preservation formats • Regularly monitors accessibility of preserved data
Stage 8: Project closure	**Trainer** *LIS data specialist*	• Trains team members • Drafts training material
	Documentation creator *LIS data specialist*	• Creates stage documentation
	Supervisor *LIS data specialist*	• Oversees stage

	Marketing *LIS data specialist*	• Liaises with the institute's communications department • Promotes rescued data • Shares data rescue learnings
	Author/presenter *LIS data specialist*	• Publishes article on rescued data/rescue project • Presents paper on rescued data/rescue project
	Miscellaneous closure tasks *LIS data specialist*	• Monitors usage statistics of digitised data • Provides feedback regarding workability of rescue model • Annual reassessment of data not rescued
	Evaluator of rescue model *LIS data specialist*	• Re-evaluates rescue model • Updates rescue model as needed
	Data assessor *LIS data specialist*	• Annual reassessment of data not rescued
	Project manager *LIS data specialist*	• Thanks involved parties • Marketing activities • Final report drafting and distribution • Post-rescue meeting chair; collates feedback
Data rescue event hosting	**Event organiser/host** *Designated research library professional*	• Communications (pre, during, post event) • Ensures spaces for data and rescuers are suitable • Arranges access to premises, parking, and related matters • Establishes security measures • Ensures equipment access • Evaluates equipment and tools • Organises catering • Arranges and tests Wi-Fi access • Ensures training area meets environmental standards; ensures safety, health, environment, and quality (SHEQ) regulations are adhered to

Recommendations and Next Steps

Based on the results of this study and the potential future use of the model, several recommendations regarding the way forward are made.

Create Awareness in the LIS Sector Regarding Data at Risk and Data Rescue

Data at risk: Awareness creation around data at risk should include, at a minimum, the topics of correct handling of older data, correct data-storing procedures, the crucial role

of a data management plan for every dataset, the significance of metadata and data documentation, ideal data preservation formats, and the importance of data migration. Awareness and implementation of mentioned factors will lessen the odds of data ending up at risk and in need of rescue.

Data rescue: While the correct handling of data will lessen the need for data rescue, it is vital that an institute's LIS sector realises that the institute potentially houses historic, valuable (even unique) data at risk that cannot be recollected. To ensure future access to the data and to minimise decay, such data would need to be rescued. To ensure the meaningful participation of the LIS sector in rescue projects, it is crucial that LIS professionals (listed in Table 4) be aware of what data rescue entails, be exposed to the recommended data rescue workflow model, be aware of data rescue interest groups or communities of practices, and be cognisant of the role of collaboration, sharing of expertise, and sharing of scarce equipment and tools.

Address the LIS Tertiary Curriculum

To enable LIS students to understand the concept of data at risk and the practices around data rescue, it is vital that the tertiary LIS curriculum considers the inclusion of data rescue in its teachings. With the South African LIS curriculum already including modules on digitisation, data management, and research practices, the added exposure to data rescue prior to stepping into the LIS profession ensures the following:

- prospective LIS professionals will possess knowledge of data at risk and data rescue,
- new LIS professionals will be able to take an active role in data at risk/data rescue skills transfer,
- newly appointed LIS professionals will be of more value to their respective institutes, as they will be able to play a vital part in ensuring an institute's research history is preserved and shared, and
- a new generation of LIS professionals will understand the importance of collaboration with an institute's research sector.

It is recommended that the topics of data at risk and data rescue best be included in a postgraduate module on data management, and only after students have been exposed to data management concepts including data management plans, data repositories, data sharing, and open data. Alternatively, the exposure to data at risk and data rescue may form part of a module on research practices, and ideally, after the aforementioned data management modules have been completed. Even including this discipline in earlier modules should be considered, as data is important across all levels of study.

Promote the Model by Releasing It under a Creative Commons Licence

An important step entails assigning a Creative Commons ShareAlike License to the recommended data rescue workflow model, enabling the use, sharing, and adaptation of the model. The designation of a licence could potentially result in its use at different institutes and involve a range of disciplines and a variety of data formats. Addressing current limitations of the model is anticipated and welcomed.

Launch Data Rescue Project Involving LIS Role Players

A major next step, following the previous recommendations, entails the launch of a data rescue project at the same South African research institute. The recommended data rescue workflow model will be used as guidance and will be preceded by a data rescue awareness session tailor-made for the institute's LIS sector, a data rescue awareness session focusing on the institute's research sector and an institute-wide call for locating data at risk.

Adaptations to the model emanating from this project and envisaged external projects are encouraged and should ideally be published and shared with the wider LIS community.

Additional Steps

An institutional survey of vintage and older data readers is bound to reveal the locations and availability of equipment, thereby enabling the view and reuse of data in a historic format.

Investigating funding opportunities for the rescue of unique historic data, such as the financial assistance provided by IEDRO (2014), could enable the implementation of rescue projects that would otherwise not have been viable.

Conclusion

This article's main contribution lies in the presentation of a recommended workflow model illustrating the ways in which the LIS sector can be involved in data rescue. Refining of the model and the drafting of adaptations are anticipated. For most institutes, various preparatory activities will have to be completed before embarking on a data rescue project, and these steps have been touched on in this article.

It is hoped that parties implementing the recommended data rescue workflow model, or parties adding changes and value to the recommended model, will share their experiences with the wider LIS sector. The sharing and distribution of obtained insight and innovation is bound to enhance the prevalence of LIS involvement in data rescue.

References

Antico, A., R. O. Aguiar, and M. L. Amsler. 2018. “Hydrometric Data Rescue in the Paraná River Basin.” *Water Resources Research* 54 (2): 1368–1381. https://doi.org/10.1002/2017WR020897.

Antuña, J. C., A. Fonte, R. Estevan, B. Barja, R. Acea, and J. C. Antuña Jr. 2008. “Solar Radiation Data Rescue at Camagüey, Cuba.” *Bulletin of the American Meteorological Society* 89 (10): 1507–1512. https://doi.org/10.1175/2008BAMS2368.1.

Ashikuzzaman, M. D. 2018. “What Is Content Analysis?” LIS Education Network, published November 29, 2018, updated December 24, 2023. Accessed February 9, 2025. https://www.lisedunetwork.com/content-analysis/.

Bradshaw, E., L. Rickards, and T. Aarup. 2015. “Sea Level Data Archaeology and the Global Sea Level Observing System (GLOSS).” *GeoResJ* 6: 9–16. https://doi.org/10.1016/j.grj.2015.02.005.

Brönnimann, S., J. Annis, W. Dann, T. Ewen, A. N. Grant, T. Griesser, S. Krähenmann, C. Mohr, M. Scherer, and C. Vogler. 2006. “A Guide for Digitizing Manuscript Climate Data.” *Climate of the Past* 2 (2): 137–144. https://doi.org/10.5194/cp-2-137-2006.

Carey, M. 2017. “University Hosts Federal Climate Data Rescue Event.” *The Hoya*, February 23. Accessed July 17, 2024. https://thehoya.com/business/university-hosts-federal-climate-data-rescue-event/.

Cascone, S. 2022. “How Tech Experts in the West Are Rushing to Save the Digital Archives of Ukraine’s Museums.” *Artnet*, March 14. Accessed July 17, 2024. https://news.artnet.com/art-world/saving-ukrainian-cultural-heritage-online-2084036.

Chaudhary. B. 2023. “Importance of Data.” Term paper, Tribhuvan University. Accessed February 9, 2025. https://www.researchgate.net/publication/374544786_Importance_of_Data_A_Term_Paper.

Costa, K. 2022. “Systematic Guide to Qualitative Data Analysis within the C.O.S.T.A Postgraduate Research Model.” Accessed February 9, 2025. https://www.coursehero.com/file/134892723/SYSTEMATIC-GUIDE-TO-THEMATIC-QUALITATIVE-DATA-ANALYSISdocx/.

C3S (Copernicus Climate Change Service). 2021. “Data Rescue Projects.” Accessed July 17, 2024. https://datarescue.climate.copernicus.eu/.

DataFirst. n.d. “DataFirst and the Rescue of Data at Risk: Partnering to enable Research Access to Historical African Data.” Accessed July 17, 2024. https://www.datafirst.uct.ac.za/services/data-rescue?highlight=WyJyZXNjdWVkIl0=.

Diviacco, P., N. Wardell, E. Forlin, C. Sauli, M. Burca, A. Busato, J. Centonze, and C. Pelos. 2015. “Data Rescue to Extend the Value of Vintage Seismic Data: The OGS-SNAP Experience.” *GeoResJ* 6: 44–52. https://doi.org/10.1016/j.grj.2015.01.006.

Downs, R. R., and R. S. Chen. 2017. “Curation of Scientific Data at Risk of Loss: Data Rescue and Dissemination.” In *Curating Research Data: Practical Strategies for Your Digital Repository*, Vol. 1, edited by L. R. Johnston, 263–277. Chicago: American Library Association. Accessed July 17, 2024. https://academiccommons.columbia.edu/doi/10.7916/D8W09BMQ.

Dudovskiy, J. 2019. “Purposive Sampling.” Business Research Methodology. Accessed July 7, 2025. https://research-methodology.net/sampling-in-primary-data-collection/purposive-sampling/.

Eke, K. 2017. “How Libraries Can Help Rescue Data.” PowerPoint slides. SlideShare. Accessed July 17, 2024. https://www.slideshare.net/kimeke/how-libraries-can-help-levels-of-data-rescue?qid=460c012e-76e2-4e1f-b142-7292775eb3b4&v=&b=&from_search=26/16/engaging-i.

ESA (European Space Agency). 2018. “ESA and Vatican Work to Preserve Heritage Data.” European Space Agency, May 4. Accessed July 17, 2024. https://www.esa.int/Our_Activities/Observing_the_Earth/ESA_and_Vatican_work_to_preserve_heritage_data.

Eveleth, R. 2014. “The Quest to Scan Millions of Weather Records.” *The Atlantic*, August 25. Accessed February 9, 2025. https://www.theatlantic.com/technology/archive/2014/08/the-quest-to-scan-millionsof-weather-records/378962/.

Gallaher, D., G.G. Campbell, W. Meier, J. Moses, and D. Wingo. 2015. “The Process of Bringing Dark Data to Light: The Rescue of the Early Nimbus Satellite Data.” *GeoResJ* 6: 124–134. https://doi.org/10.1016/j.grj.2015.02.013.

GLoSAT (Global Surface Air Temperature). n.d. “Data Rescue.” Accessed July 17, 2024. https://www.glosat.org/data-rescue.

Griffin, E. 2006. “Rescuing and Recovering Lost or Endangered Data.” *Data Science Journal* 4: 21–26. Accessed July 17, 2024. https://doi.org/10.2481/dsj.4.21.

Griffin, E. 2017. “Rescue Old Data Before It’s Too Late.” *Nature* 545: 267. https://doi.org/10.1038/545267a.

Griffin, R. E. 2015. “When Are Old Data New Data?” *GeoResJ* 6: 92–97. https://doi.org/10.1016/j.grj.2015.02.004.

Güntsch, A. D., D. Fichtmüller, A. Kirchhoff, and G. Berendsohn. 2012. “Efficient Rescue of Threatened Biodiversity Data Using reBiND Workflows.” *Plant Biosystems—An International Journal Dealing with All Aspects of Plant Biology* 146 (4): 752–755. https://doi.org/10.1080/11263504.2012.740086.

Hoffman, K. M., C. T. Clarke, H. S. Y. Shiue, P. Nicholas, M. Shaw, and K. Fenlon. 2020. *Data Rescue: An Assessment Framework for Legacy Research Collections*. University of Maryland, University Libraries, Digital Repository at the University of Maryland (DRUM). Accessed July 17, 2024. https://drum.lib.umd.edu/bitstream/handle/1903/26475/DataRescue_whitepaper_final_NALDC.pdf?sequence=1&isAllowed=y.

Hsu, L., K. Lehnert, S. Carbotte, V. Ferrini, J. Delano, J. B. Gill, and M. Tivey. 2013. "Rescue of Long-Tail Data from the Ocean Bottom to the Moon." Paper presented at the AGU Fall Meeting, San Francisco, United States, December 9–13, 2013. Accessed July 17, 2024. https://www.slideshare.net/hsuleslie/hsu-agu2013?qid=e164d577-cec1-45ae-9dc4-ef38c3c47652&v=&b=&from_search=22.

IEDRO (International Environmental Data Rescue Organization). 2014. "The Process." Accessed July 17, 2024. https://iedro.org/about-us/how-this-all-works/.

Kennedy, M. 2017. *Guidelines for Marine Species Occurrence Data Rescue—The OBIS Canada Cookbook*. Accessed July 17, 2024. https://www.coinatlantic.ca/_files/ugd/cf2ff9_82e7008749294b83b31c4a8a9ecd99cf.pdf.

Kijas, A. E. 2018. "Engaging in Small Data Rescue." *Humanities Commons*, February 9. Accessed July 17, 2024. https://works.hcommons.org/records/fxsz3-jgc69.

Mavraki, D., L. Fanini, M. Tsompanou, V. Gerovasileiou, S. Nikolopoulou, E. Chatzinikolaou, W. Plaitis, and S. Faulwetter. 2016. "Rescuing Biogeographic Legacy Data: The 'Thor' Expedition, a Historical Oceanographic Expedition to the Mediterranean Sea." *Biodiversity Data Journal* 4: e11054. https://doi.org/10.3897/BDJ.4.e11054.

Modine, A. 2008. "Vintage IBM Tape Drive in Apollo Moon Dust Rescue." *The Register*, November 11. Accessed July 17, 2024. https://www.theregister.com/2008/11/11/vintage_ibm_tape_drive_moon_dust_data/.

Palmer, C. L., N. M. Weber, and M. H. Cragin. 2011. "The Analytic Potential of Scientific Data: Understanding Re-use Value." *Proceedings of the American Society for Information Science and Technology* 48 (1): 1–10. https://doi.org/10.1002/meet.2011.14504801174.

Patterton, L. H. 2023. "Data Rescue: Defining a Comprehensive Workflow That Includes the Roles and Responsibilities of the Research Library." PhD diss., University of Pretoria. https://repository.up.ac.za/handle/2263/89829.

Perrier, L., E. Blondal, and H. MacDonald. 2018. "Exploring the Experiences of Academic Libraries with Research Data Management: A Meta-Ethnographic Analysis of Qualitative Studies." *Library and Information Science Research* 40 (3–4): 173–183. https://doi.org/10.5281/zenodo.1324412.

Raughley, L. 2017. “Library Participates in Effort to Preserve Government Data.” University of Michigan Library, January 30. Accessed December 4, 2023. https://record.umich.edu/articles/library-participates-effort-preserve-government-data/.

RDA (Research Data Alliance). 2017a. *Guidelines for Data Rescue*. Accessed July 17, 2024. https://www.rd-alliance.org/groups/data-rescue/forum/topic/guidelines-for-data-rescue/.

RDA (Research Data Alliance). 2017b. “IG Data Rescue RDA 9th Plenary Meeting.” January 11, 2017. Accessed December 4, 2023. https://www.rd-alliance.org/ig-data-rescue-rda-9th-plenary-meeting.

Ryan, C., C. Broderick, M. Curley, C. Daly, C. Duffy, P. Thorne, M. Treanor, S. Walsh, and C. Murphy. 2017. “Integrating Data Rescue into the Classroom.” Accessed July 17, 2024. https://www.maynoothuniversity.ie/sites/default/files/assets/document/Data%20Rescue%20poster_0.pdf.

Sheikh, A., A. Malik, and R. Adnan. 2023. “Evolution of Research Data Management in Academic Libraries: A Review of the Literature.” *Information Development* 41 (2): 305–319. https://doi.org/10.1177/02666669231157405.

Thompson, C. A., W. Davenport Robertson, and J. Greenberg. 2014. “Where Have All the Scientific Data Gone? LIS Perspective on the Data-At-Risk Predicament.” *College and Research Libraries* 75 (6): 842–861. https://doi.org/10.5860/crl.75.6.842.

Thompson, K. 2017. “Documentation as Data Rescue: Restoring a Collection of Canadian Health Survey Files.” *Against the Grain* 29 (6): 33–35. https://doi.org/10.7771/2380-176X.7876.

Wiebe, P. H., and M. D. Allison. 2015. “Bringing Dark Data into the Light: A Case Study of the Recovery of Northwestern Atlantic Zooplankton Data Collected in the 1970s and 1980s.” *GeoResJ* 6: 195–201. https://doi.org/10.1016/j.grj.2015.03.001.

Wippich, C. 2012. “Preserving Science for the Ages—USGS Data Rescue.” Fact Sheet 2012–3078. Reston: United States Geological Survey. https://doi.org/10.3133/fs20123078.

WMO (World Meteorological Organization). 2014. *Guidelines for Hydrological Data Rescue*. Geneva: WMO. Accessed July 17, 2024. https://library.wmo.int/index.php?lvl=notice_display&id=16792#.YXJuBxpBxAU.

WHO (World Meteorological Organization). 2016. *Guidelines on Best Practices for Climate Data Rescue*. Geneva: WMO. Accessed July 17, 2024. https://repository.oceanbestpractices.org/bitstream/handle/11329/366/1182_en.pdf?sequence=1&isAllowed=y.

WMO (World Meteorological Organization). 2020. “Data Rescue.” Accessed February 9, 2025. https://community.wmo.int/data-rescue.

Woolfrey, L. 2016. “The Rescue of ‘at Risk’ Data on Forced Resettlement in South Africa.” Paper presented at the RDA/CODATA Workshop on the Rescue of Data at Risk, National Center for Atmospheric Research, Boulder, Colorado, United States, September 8–9, 2016.