

## Data Management Plan

# Migration of legacy data to new media formats for long-time storage and maximum visibility: Modern pollen data from the Canadian Arctic (1972/1973)

Harvey Nichols<sup>‡</sup>, Susann Stolze<sup>‡</sup>

<sup>‡</sup> University of Colorado Boulder, Boulder, United States of America

Corresponding author: Susann Stolze ([susann.stolze@gmx.de](mailto:susann.stolze@gmx.de))

Reviewable v1

Received: 23 Aug 2016 | Published: 25 Aug 2016

Citation: Nichols H, Stolze S (2016) Migration of legacy data to new media formats for long-time storage and maximum visibility: Modern pollen data from the Canadian Arctic (1972/1973). Research Ideas and Outcomes 2: e10269. doi: [10.3897/rio.2.e10269](https://doi.org/10.3897/rio.2.e10269)

## Abstract

### Background

This Data Management Plan (DMP) was created using the DMPTool. It describes modern pollen data collected along a 2500 mile (~4000 km) transect in the Canadian Arctic in 1972/73 as part of an NSF funded research project (GB-33497). The project was undertaken at the Institute of Arctic and Alpine Research at the University of Colorado Boulder. This legacy dataset originally stored as a paper copy and on 35-mm film will be migrated to digital formats that allow upload of the dataset to an international open-access library for permanent storage and visibility. This DMP was submitted to the 2015 Best Digital Data Management Plan and Practices - Competition (<https://data.colorado.edu/cudmpguidance>).

### New information

The study provides a baseline pollen dataset for the interpretation of Holocene pollen diagrams from this region and for comparison with modern surface pollen samples,

allowing the assessment of the effects of modern climate change on Arctic and subarctic ecosystems.

## Keywords

data management plan, Canadian Arctic, surface pollen data, modern pollen traps, baseline study, vegetation, climate

## Types of data produced

In 1972/73, Dr. Harvey Nichols, emeritus professor at the University of Colorado Boulder, conducted an NSF funded (GB-33497) baseline study on modern pollen deposition along a 2500 mile transect in the Canadian Arctic. This study was designed to provide palynological data for the interpretation of Holocene pollen diagrams from that area and also to address the effects of modern climate change on Arctic and subarctic ecosystems such as the northward displacement of the Arctic treeline. The annual pollen deposition was monitored at 34 locations in a region of large geographical extent of today's Nunavut territory within the borealtundra transition and arctic tundra ecosystem in the Canadian Arctic for a two-year period, using standard pollen collectors used in aerobiology. The project generated pollen and environmental data for a total of 505 samples. The pollen samples were studied by optical microscopy to determine pollen taxa and abundance for each sample collected. These data were complemented by environmental observations in the field using 35-mm color slide films and Super-8 rolls of film. Original input, processing, and storage of the pollen data in the 1970's were performed with early digital IBM computers using punched cards. The computer programs used were developed by the PI, together with Margaret Nichols and Margaret Eccles at the Institute of Arctic and Alpine Research (INSTAAR). The data have been stored for future reference in physical formats including punched cards, a paper copy, notebooks, color slides, and rolls of film. This extensive dataset of modern pollen deposition represents a unique baseline for future aerobiological, environmental, and climatological studies that seek to reveal the extent and magnitude of climatically induced environmental change linked to climate change in the Canadian Arctic since the 1970's. Given the ongoing scientific debate on the rapid environmental changes happening in this region today, publication of this dataset is of particular interest.

## Data and metadata standards

The legacy data will be migrated to digital data formats common to palynological research. The complete dataset stored on a paper copy will be transcribed into an electronic Excel document (.xls). Color slides will be scanned and saved as digital images in .jpeg format. The video material will be transferred commercially from analog VHS tapes to digital carriers (blue ray disks) as QuickTime (.mov) or Audio Video Interleave (.avi) formats. A

verbal description of the environmental information documented by camera at each site will also be included in the digital research data file (.xls). This file will also store necessary metadata that allow easy use of the dataset. Information on the project aims, environmental settings, sampling strategy, field work, applied methods, and data processing will be stored as a word document (.doc) and in a portable document format (.pdf).

## **Policies for Access and Sharing**

To ensure discoverability and reusability of the research dataset, the data descriptor will be submitted to an international open-access, peer-reviewed journal for the publication of data descriptions. Following a pre-submission enquiry, full submission of the data descriptor to Scientific Data published by the Nature Publishing Group has been strongly encouraged. Upon completion, the dataset will be made publicly available by uploading to the open-access library PANGAEA, approved by the journal as a depository. PANGAEA is a widely recognized and heavily used repository for palynologists and other scientists interested in environmental and climatological research. Because a manuscript for an international peer-reviewed journal is in preparation, a password protection of the dataset will be requested from PANGAEA to ensure a moratorium period. During this period, other researchers interested in the dataset are required to obtain permission for access from the PI and/or Dr. Stolze.

## **Policies for re-use and distribution**

The PI holds the property rights of the data. The data uploaded to PANGAEA can be used according to the terms of the license provided in the included metadata. By submitting the data and metadata to an open-access library and journal, the data will be accessible under a creative commons license, so the users must indicate the original citations of the dataset and metadata.

## **Plans for archiving and preservation**

All legacy storage media and the digital carriers will be securely stored in a dedicated room by the PI. Digital copies of the research dataset will be stored on password protected computers of the PI and Dr. Stolze, with additional back-up copies on securely stored external hard drives to be deposited at INSTAAR. Synchronization of the copies will be performed as necessary. The digital master file including metadata will be stored indefinitely at a legally protected open access library and data management system for Earth and Environmental Science, PANGAEA, which guarantees long-term availability. Assignment of a Digital Object Identifier (DOI) ensures that the data will be archived as a citable data collection. A digital and a newly printed hardcopy of the research data and documentation along with the digital carriers of the photographic and video recordings will

be stored at the INSTAAR library and the Archive Department of the Norlin Library of the University of Colorado Boulder for future reference. We expect that the migration of the legacy data to the new digital formats, publication of the metadata in an open access peer-reviewed journal, and upload to the data library will be completed in 2016.

## **Roles and responsibilities**

Dr. Harvey Nichols designed and executed the baseline study in the 1970's. Due to the technological obsolescence of the legacy source media, Dr. Susann Stolze, an affiliate researcher at INSTAAR, was recently contracted to develop and execute a management plan, with oversight by the PI, to migrate the legacy data to new, digital, media formats for permanent storage, maximum visibility, and easy accessibility. To ensure highest quality of the migrated data, comparison of the data with that stored on the legacy source media will be conducted by S. Stolze and H. Nichols. The PI has taken long-term responsibility for management and sharing of the data. In case of long-term absence of the PI, this responsibility would, with his permission, be transferred to S. Stolze until completion of this data management plan. Upon the archiving of the dataset, the responsibility for data distribution and storage will be assigned to the archival service.