# No Rain on This Parade: Data Curation and the Alberta Hail Project Archive

Kelly E. Lau | MAS, MLIS

On the sprawling Canadian prairies, adverse weather conditions such as **hailstorms** can have a significant impact on the agricultural economy by causing millions of dollars in damage to crops.

To study hailstorm physics and find out how to suppress hail, the **Alberta Hail Project** was established in 1956.

Between 1956-1985, a vast amount of **research data** were collected on magnetic tape: precipitation mechanisms, severe storm development, hail suppression, hydrology, and microwave propagation.

CONTEXT

In 1995, the data were transferred to CD-ROM. Data on magnetic tape were degrading and expertise familiar with their collection and calibration were becoming scarce.

CONTEXT

In 2012, CD-ROM is no longer a viable preservation format and this valuable metereological data archive needed to be transferred to a more stable **preservation** and **access** platform.

## DEFINITION

**Data curation** is the active and on-going management of data throughout the **research data lifecycle**: creation, processing, analysis, preservation, access, and re-use.

## OBJECTIVE

Preserve and provide access to the digital radar, aircraft, and surface precipitation data along with supporting calibrations and documentation.

## RATIONALE

Make the Alberta Hail Project Archive available to the **scientific research community**.

QUESTIONS

What data management **services** and **infrastructure** are available or being developed?

Who are the **stewards** for data across the lifecycle and how does data flow among these stewards?

What are the **measures** of successful data stewardship?

What aspects of data management and infrastructure lend themselves most readily to **collaboration**?

MODELS

Open Archival Information System (OAIS)
Research Data Management Lifecycle

FRAMEWORKS

An open-source **digital preservation system** for standards-based, long-term access to research data that uses an integrated suite of tools to process objects from ingest to dissemination.

An open-source **dissemination platform** that researchers can use to upload, share, and access research data.

RELEVANCE

Data curation is an emerging field that brings new **opportunities** and **challenges** for libraries. The growing movement to effectively manage, preserve, retrieve, and re-use research data is one that complements traditional **library missions**.

Data tell **stories** across the sciences, social sciences, and humanities. They help us make sense of the world.

Data help people identify the problems and achievements of the world around us, and evaluate, and reflect on possible **solutions**.

CHALLENGES

The "**long tail**" of data. There are fewer "big data" compared to the thousands of smaller datasets. The myriad of smaller datasets pose their own daunting challenges for management and preservation.

We need skills and tools to extract data's **value**.

If we can't tell the value and **stories** behind our data, we can't make arguments to invest in data management and sharing.

OUTCOMES

Since the end of the Alberta Hail Project, numerous research projects have demonstrated the value of using the Archive: theses, publications, conference proceedings, and scientific reports. Areas of study have included cloud physics, computer science, instrumentation, and meteorology.

**Scientific research** and **collaborations** continue to this day.

OPPORTUNITY

Help researchers create **new knowledge** by making research data that are of interest to scholarly and educational activities available.

OPPORTUNITY

**Make friends**. Form collaborations and **partnerships** while performing your data stewardship responsibilities.

OPPORTUNITY

## BE INSPIRED

**You are engaged in a craft.**

**Your tools are getting better all the time.**

**Your job is challenging and illuminating.**

**Contemporary and future generations are counting on you.**

UNIVERSITY OF ALBERTA