

DataStream got its start in the Mackenzie River Basin 10 years ago and has since established hubs across Canada. Photo credit: Pat Kane.

Dawn of data sharing

Championing water quality data sharing for Great Lakes futures

BY KATHERINE BALPATAKY

WHEN YOU WALK THE SHORELINE of a great lake over days, years, or decades, you can't help pondering the changes you witness. For me, that great lake is Lake Erie. I grew up with Long Point in my backyard and now own a home near Port Colborne, Ontario. In my 48 years, I have seen shorelines transformed by the arrival of new species and the loss of familiar ones, as well as by algal blooms, plastic waste, and climate change. These changes have inspired curiosity and dismay, but also hope. More research and actions are needed to protect these precious lakes. Underlying all this is a need for greater availability of data.

At DataStream, I collaborate with an incredible team bringing passion and expertise to address the freshwater data limitations we face across Canada. While numerous organizations, researchers, and industries collect water data, as a nation, we haven't pulled it together and used it effectively. Practically, we lack the evidence to make important decisions that will shape the health of our Great Lakes—now and into the future. It's a serious challenge that some regard as a national security threat. DataStream is addressing this problem. In the U.S., there is greater integration of publicly available water quality data through the Water Quality Portal. However, due to a recent lapse in government funding, the portal's future is uncertain.



Katherine Balpataky and Crowley, the English Setter, exploring the shores of Lake Erie.

Northern exposure

DataStream is a registered charity supporting a comprehensive data literacy program and an open access platform for sharing western scientific water quality data.

DataStream's story includes hundreds of community groups, governments at all levels, Indigenous water guardians, and freshwater scientists committing their time to data stewardship. Water monitoring is the starting point of a lifecycle that includes working with a specialist to standardize data, prepping metadata, and sharing it for future analysis. As we say at DataStream, "open data is a team sport." The results are inspiring, and we're only beginning to understand its potential.

DataStream exists because 10 years ago, communities of the Mackenzie River Basin in the Northwest Territories were grappling with how to protect freshwater for future generations. They identified data gaps as a barrier. They needed a solution to assess the outcomes of a collaboration between five neighbouring jurisdictions and the Government of the Northwest Territories (GNWT).

To address this gap, the GNWT teamed up with the Gordon Foundation, a philanthropic foundation with a long-standing interest in freshwater protection. Their goal was to build a system to house community-based and government datasets. The resulting 2016 launch of Mackenzie DataStream became the prototype for the current system that houses surface water, sediment, and (soon-to-be) groundwater data, with over 340 organizations engaged, all in the name of freshwater protection.

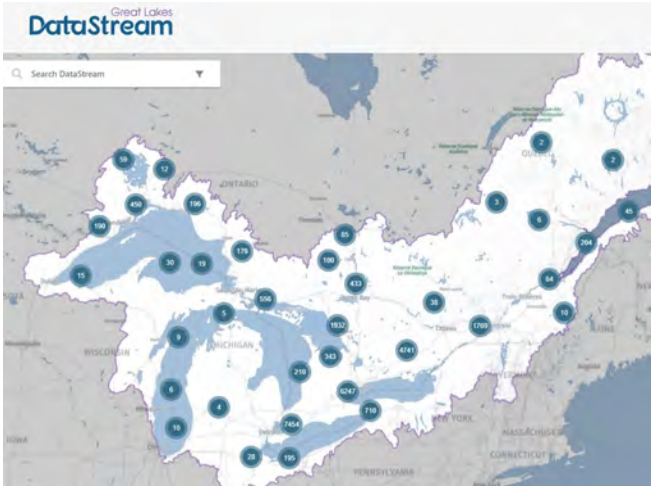
Dizzying days of data sharing

DataStream's cross-sector and cross-country data sharing has been conceptualized as the appropriate type of solution in literature on freshwater data. A 2025 article by Jess Kidd et al. in the journal *Water* synthesized insights from technology sector experts about Canada's freshwater data dilemmas. While our team wasn't mentioned, the results ring true to our experience and approach. The authors pointed to the need for 1) support of an open data culture, 2) enhanced use of data licenses, 3) increased data literacy skills and development, as well as 4) a freshwater data standard guiding collection and management. But there's more to do.

Data sharing is particularly relevant for environmental data and data collected in the public interest, using public funds.

The open data culture the authors speak of is a growing movement worldwide. DataStream talks to groups who are initially unsure about sharing data openly. However, data sharing is particularly relevant for environmental data and data collected in the public interest, using public funds. When we describe transboundary data uses or the time saved (by one researcher's estimation over 60%) by having data standardized in one place, monitoring groups get excited about the potential. Certainly not all data are suitable for open access, such as Indigenous Traditional Knowledge, but most water quality data are fair game.

The lion's share of the effort involved in fostering an open data culture comes from outreach and support. There are



DataStream has regional hubs in the Great Lakes, Lake Winnipeg, the Mackenzie River Basin, the Atlantic region, and the Pacific region.

hundreds, maybe thousands, of community-based monitoring groups aiming to improve freshwater health. They wish to contribute to science, policy, water management, and education efforts, but need help with their data. Although data may be captured by non-scientists, it's highly valuable and sometimes collected with expert advice or guidance in design protocols. It costs money (often taxpayers') to collect it, and it's essential for freshwater science and protection. Therefore, we support them in data formatting and management.

DataStream is growing into a one-stop-shop for making freshwater quality data discoverable, available, machine-readable, and ready for prime time. The platform provides free and open access data, aligns with international open data best practices, and all datasets are published under open data licenses, which provide clarity around ownership, attribution and reuse. Finally, data are shared in a standardized format, based on the WQX (Water Quality Exchange) schema, the most widely used water quality standard in North America.

Here are five inspiring examples of data collaborations from the Great Lakes:

- University of Waterloo researchers crunch big data using AI to assess nutrient loads
- Citizen scientists gather nutrient data on Lake Erie to fight algal blooms
- Volunteers with Ottawa Riverkeeper advance chloride monitoring to advocate for smarter salt management
- DataStream and the Great Lakes Observing System develop an integration to make water data more accessible
- Data Rescue intern curates polyfluoroalkyl substances data from the Université de Montréal

DataStream's efforts in the Great Lakes are supported by the Canada Water Agency.

Katherine Balpataky is the executive director of DataStream where she works with the team to foster strategic partnerships across Canada.