

Efficient integration of large amounts of field survey data

Shifting from a state-of-the-art Web Portal to Spreadsheets

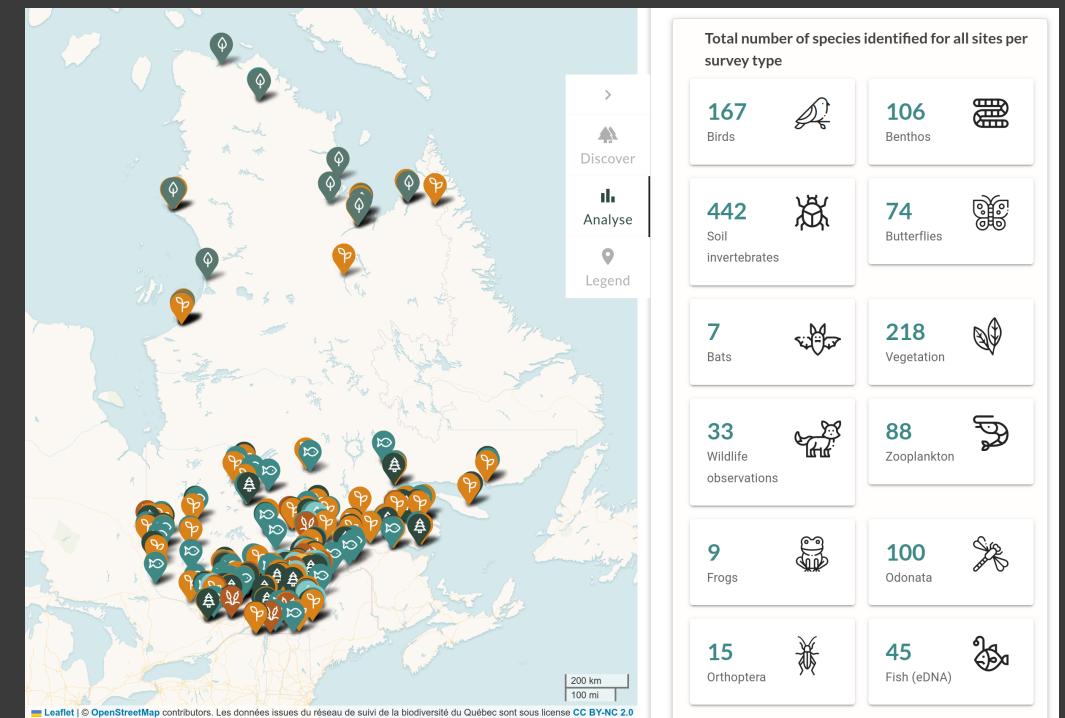
Victor Cameron
2023



Québec Biodiversity Observation Network

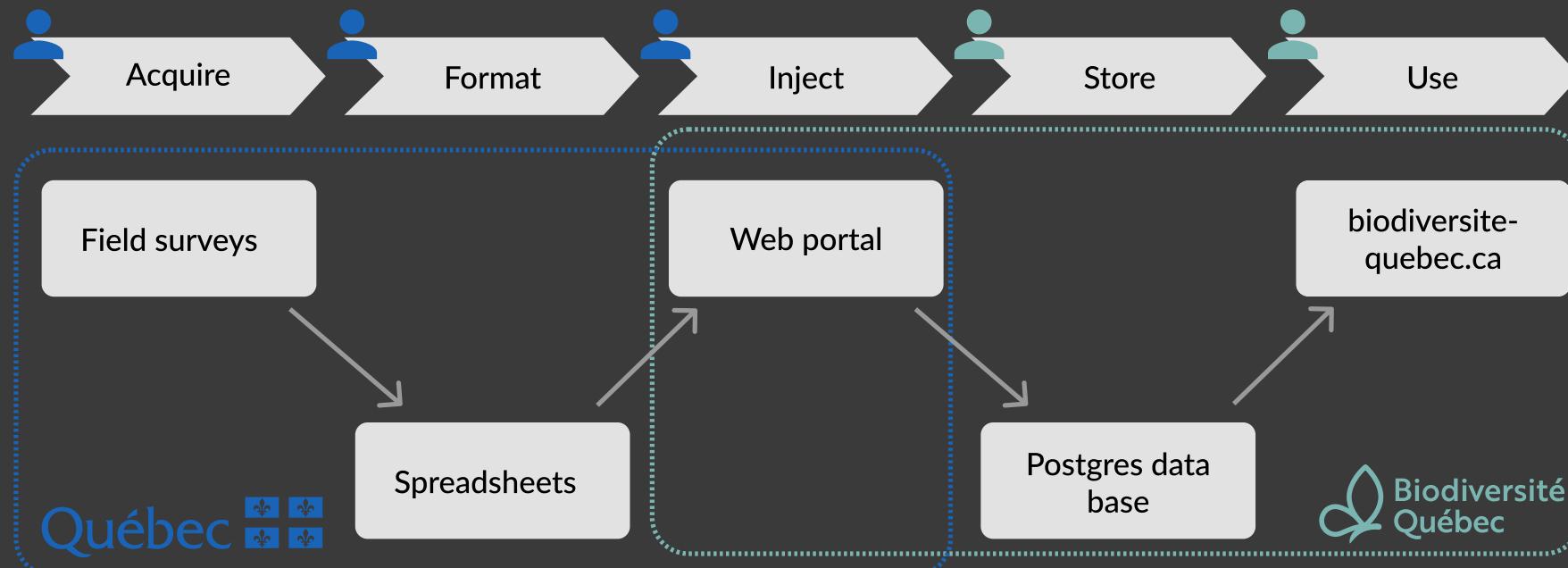
The network is a collection of field surveys to measure ecosystem changes

- 355 sites
- Monitoring since 2016
- 15+ survey types
- ~1,000,000 observations per year



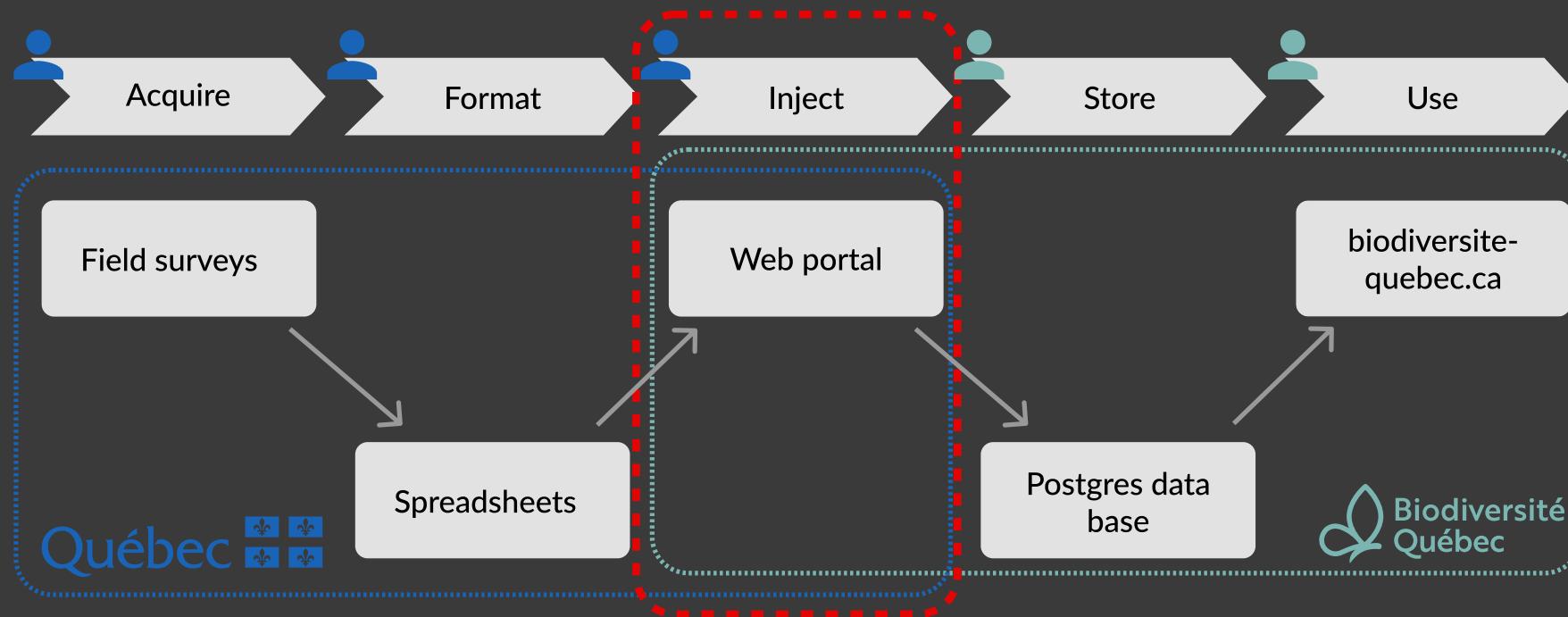
Québec Biodiversity Observation Network

Our mandate: create and manage the data workflow



Québec Biodiversity Observation Network

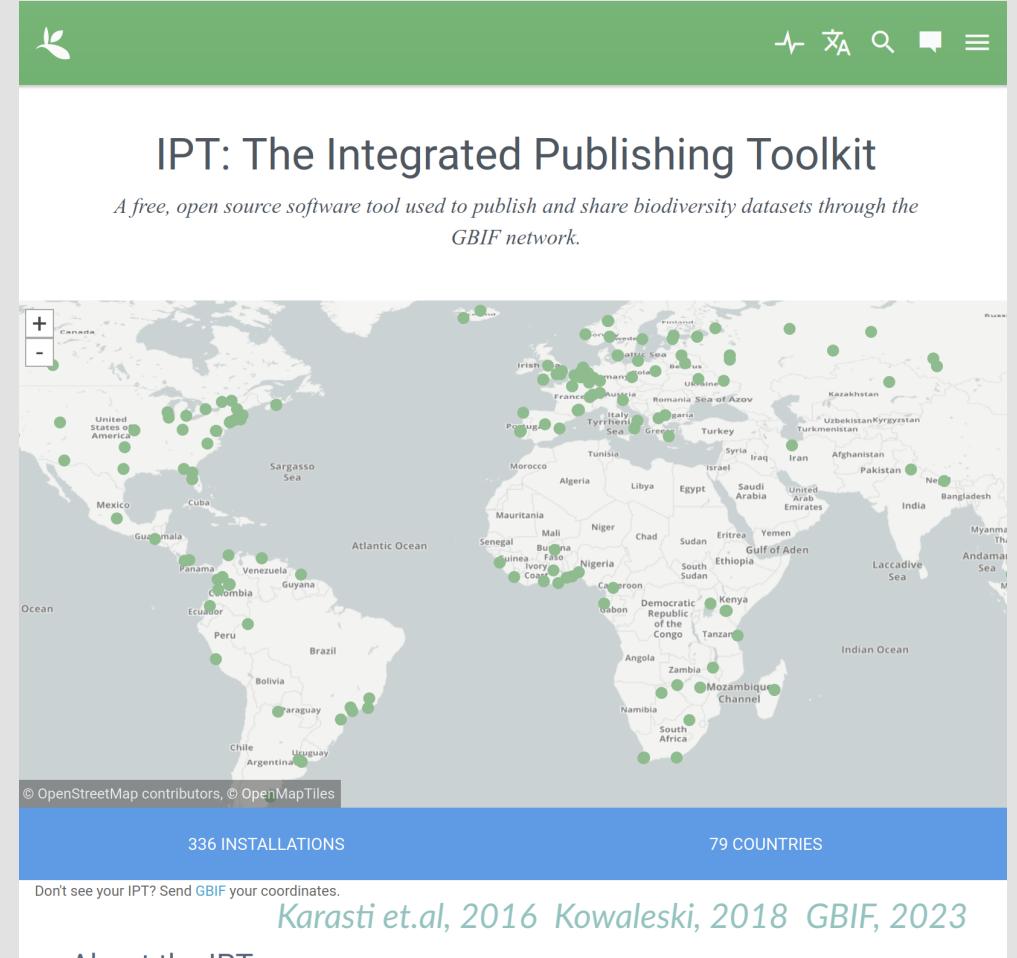
Data formatting, management, and **software skill sets** are *barriers* to a project's data workflow



Partners need an interactive interface

GBIF's IPT

- ✓ Facilitates biodiversity information sharing
- ✓ Manages and speeds requests to GBIF's server
- ✓ Encourages data-publication by increasing *usability*
- ✓ Regroups data management services



Karasti et.al, 2016 Kowaleski, 2018 GBIF, 2023

Partners need an interactive interface

GBIF's IPT

- ✓ Facilitates biodiversity information sharing
- ✓ Manages and speeds requests to GBIF's server
- ✓ Encourages data-publication by increasing *usability*
- ✓ Regroups data management services

- ⚠ Intensive development and maintenance
- ⚠ Iterative process
- ⚠ Training and support
- ⚠ Adapt to the needs of the network

Partners need an interactive interface

Our web portal

- ✓ Data availability
- ✓ Data injection
- ✓ Data quality control
- ✓ Data privacy
- ✓ Well-designed interface

The screenshot shows a web-based application for biodiversity monitoring. At the top left is the logo "Suivi Biodiversité Québec" featuring a stylized frog, flower, and dragonfly. The top right has a blue header bar with icons for edit, delete, and search. The main content area is titled "Site > Campagne". It includes several sections: "Informations" (Site number: 148_101_H01, Inventory type: Butterflies, Dates: July 5-7, 2021, Technicians: Sylvie Normand, Stephanie Cholette), "Effort" (Departure time: hh:mm:ss, End time: hh:mm:ss, Notes), "Environnement" (Wind: Calm, Sky: Partly cloudy, Temperature: 27°C, Notes), and "Observations" (Search bar, Date, Validity, Species, Abundance, Media, Notes). A sidebar on the left lists navigation items: Cartographie, Cellules, Sites (selected), Campagnes, Taxons de référence, Victor Cameron, and Déconnexion. At the bottom is a footer with the Fonds vert logo and copyright information: "© 2016 - développé par iELab".

Partners need an interactive interface

Our web portal

- ✓ Data availability
- ✓ Data injection
- ✓ Data quality control
- ✓ Data privacy
- ✓ Well-designed interface

- ⚠ Vulnerable to technological obsolescence
- ⚠ Limited flexibility to accomodate new data schemas
- ⚠ User experience is more complex than anticipated

Lessons learned

- 💡 Schema of acquired data evolves with the needs of field technicians and protocols
- 💡 Skill level and comfort with software tools vary widely between partners, participants, and technicians who manipulate the data
- 💡 Data management and usage require quality control and interactive interfaces

Spreadsheet-based template system



Spreadsheet-based template system

Template = data + metadata + formatting guide + validation rules

Spreadsheet templates

- ✓ Builds on partners' existing skill sets and workflows
- ✓ Easier to maintain
- ✓ Fast adoption and important reduction of errors
- ✓ Fast to develop and to adapt

The screenshot shows a Microsoft Excel spreadsheet titled "Template de téléversement". The table has columns labeled "campaigns_type", "sites_site_code", "landmarks_lat", "landmarks_lon", and "campaigns". Row 10 is highlighted in green, and row 11 is highlighted in red. A tooltip in cell A11 states: "'Camapigns_type' devrait être sélectionné parmi les 'Valeurs acceptées'". The status bar at the bottom shows tabs for "4. Valeurs acceptées", "5. Glossaire", "6. Template de téléversement" (which is selected), and "7. Cha".

	A	B	C	D	E
1	campaigns_type	sites_site_code	landmarks_lat	landmarks_lon	campaigns
10					
11		"Camapigns_type" devrait être sélectionné parmi les "Valeurs acceptées"			
12					
13					
14					
15					
16					
17					
18					
19					
20					
21					
22					
23					
24					
25					
26					
27					
28					
29					
30					
31					
32					
33					
34					
35					

Spreadsheet-based template system

Template = data + metadata + formatting guide + validation rules

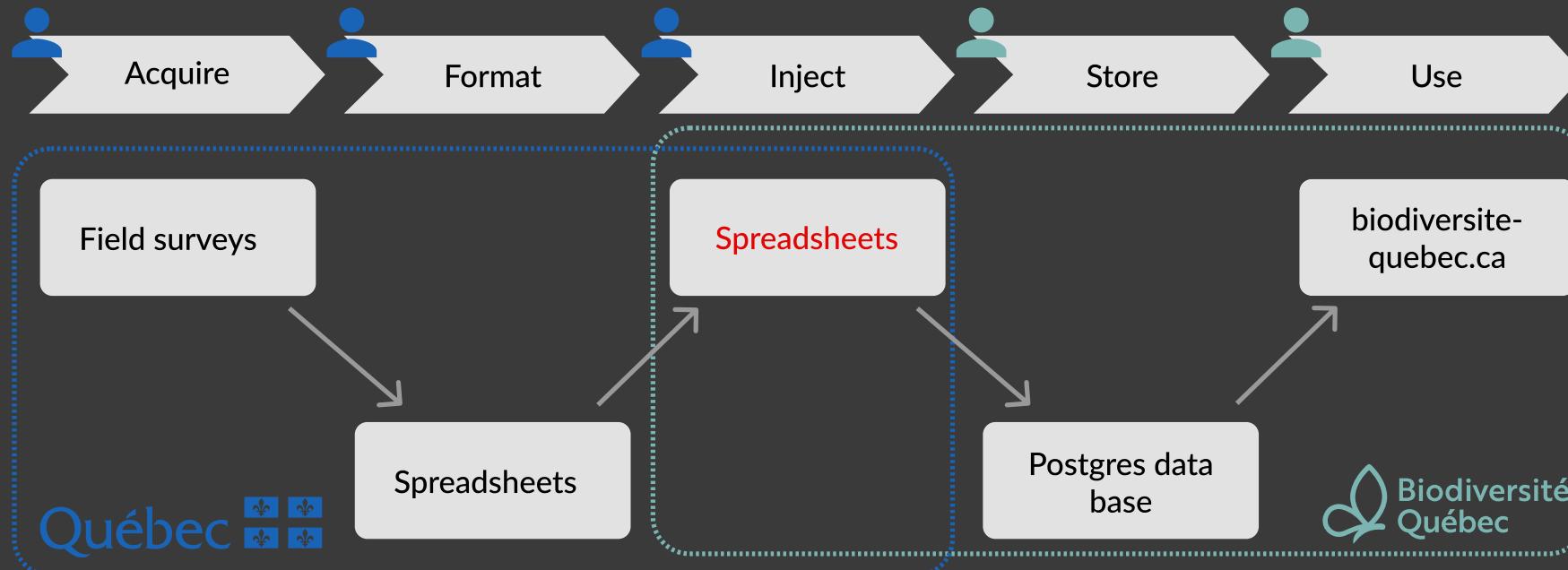
Spreadsheet templates

- ✓ Builds on partners' existing skill sets and workflows
 - ✓ Easier to maintain
 - ✓ Fast adoption and important reduction of errors
 - ✓ Fast to develop and to adapt
- ⚠ Proprietary software
 - ⚠ Difficult to make reproducible

However, web-based visualization tools have proven effective in communicating.

Spreadsheet-based template system

The revised workflow better meets the project's needs



Take home messages

Opt for tools that fit the team skill set and processes

Evolving needs and varied level of skill of partners present large challenges

Budget and plan for highly qualified human resources for maintenance and improvements