

Efficient integration of large amounts of field survey data

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

Victor Cameron

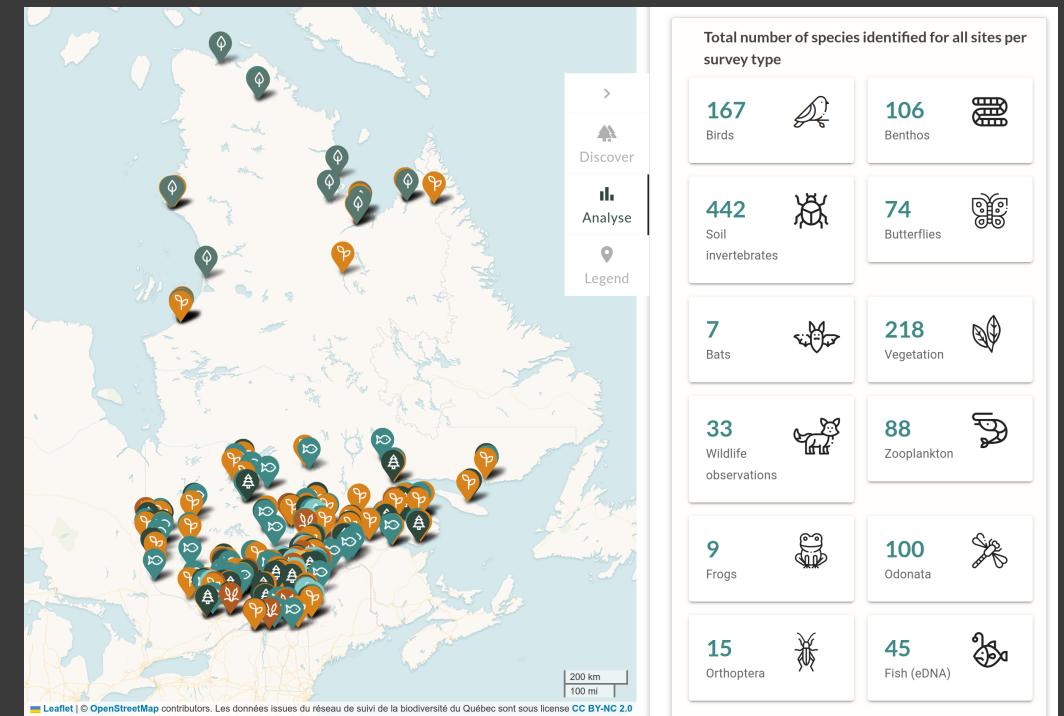
Montréal, 2023



Québec Biodiversity Observation Network

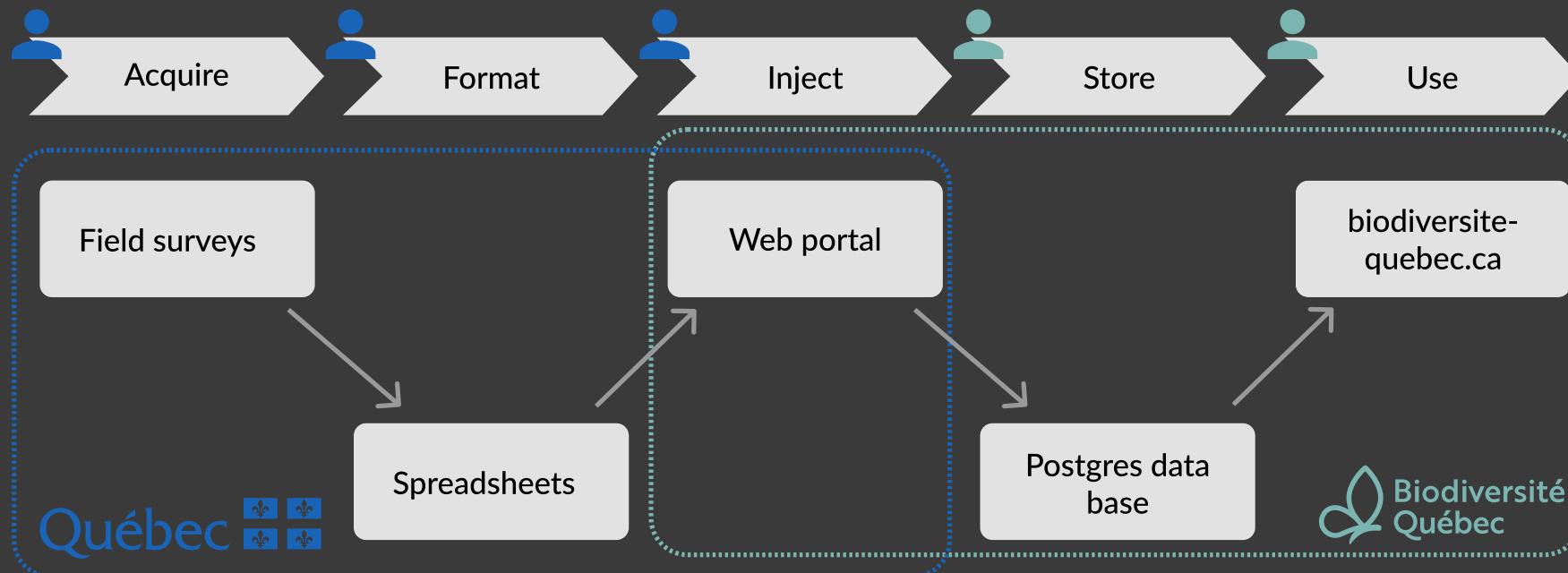
The network is a collection of field surveys to measure ecosystem changes

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



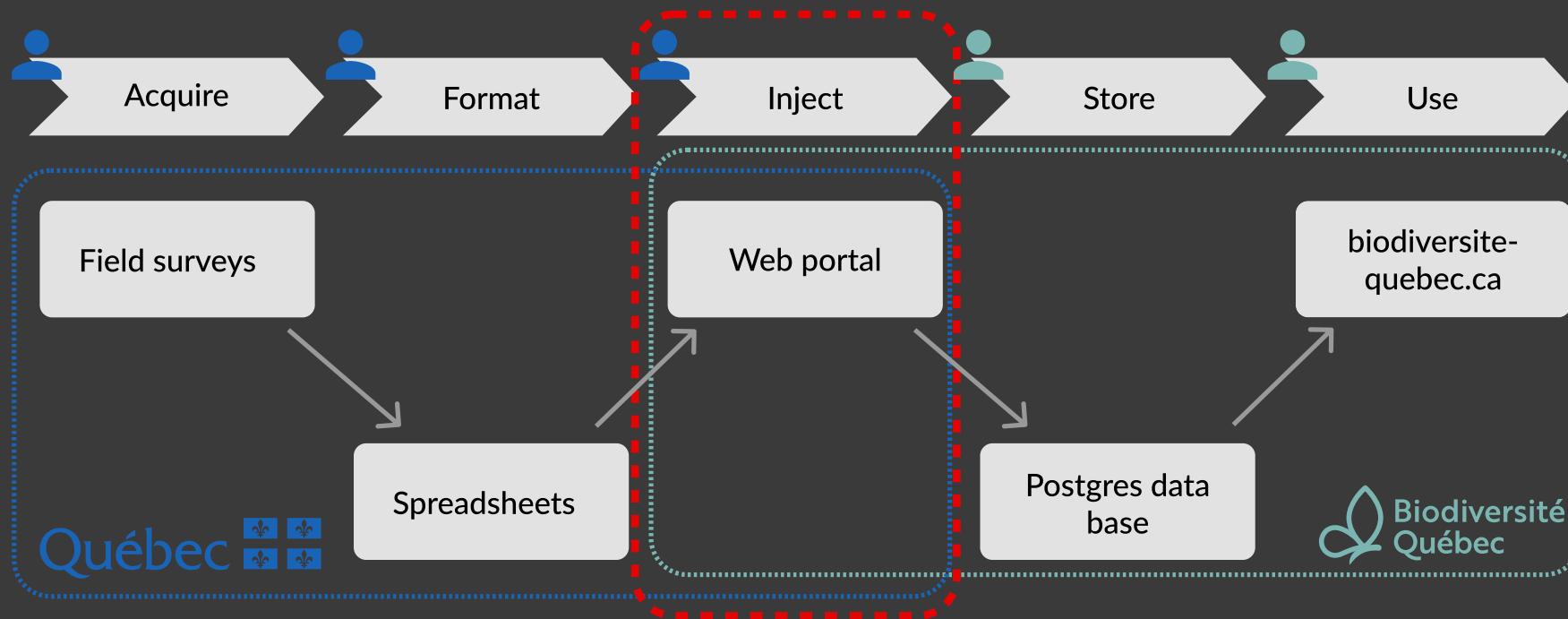
Québec Biodiversity Observation Network

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



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 green plant and blue dragonfly. The top right contains standard navigation icons: a blue pencil for edit, a red trash can for delete, and a magnifying glass for search.

The main content area is titled "Site > Campagne". It is divided into several sections:

- Informations:** Fields for "No. site" (148_101_H01) and "Type d'inventaire" (Papillons), along with date pickers for "Date d'ouverture" (2021-07-05) and "Date de fermeture" (2021-07-05).
- Techniciens:** A list of technicians: "Sylvie Normand" and "Stephanie Cholette", each with a minus sign to remove them. A "+ Ajouter" button is available for adding more.
- Effort:** Fields for "Heure de départ" and "Heure de fin" (both set to hh:mm:ss), and a "Notes" text area.
- Environnement:** Fields for "Vent" (Calm (moins de 1 km/h)) and "Ciel" (Partiellement nuageux (10 à 50 %)), along with temperature input ("Température (C°)" set to 27) and a "Notes" text area.
- Observations:** A table header with columns: Date, Valide, Espèce, Abondance, Média(s), and Notes. Below it, a message says "Aucune données".
- Média:** A section with a search bar labeled "Recherche...", and buttons for "Visualiser" (with a camera icon), "Télécharger" (with a download icon), "Téléverser" (with an upload icon), and a trash can icon.

On the left side of the interface is a sidebar with the following menu items:

- Cartographie
- Cellules
- Sites
- Campagnes
- Taxons de référence
- Victor Cameron
- Déconnexion

At the bottom center is the "Fonds vert" logo with the text "© 2016 - développé par iEILab".

Partners need an interactive interface

Our web portal

✓ Data availability

✓ Data injection

✓ Data quality control

✓ Data privacy

✓ Well-designed interface

⚠ User experience is more complex than anticipated

⚠ Limited flexibility to accomodate new survey types

⚠ Vulnerable to technological obsolescence

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



jxf@mastodon.social
@jxxf

Optimist: The glass is $\frac{1}{2}$ full.
Pessimist: The glass is $\frac{1}{2}$ empty.
Excel: The glass is January 2nd.

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 with a dark theme. The top menu bar includes 'Fichier', 'Accueil', 'Insertion', 'Dessin', 'Mise en page', 'Formules', 'Données', 'Partager', and '...'. The ribbon tabs are 'Accueil' (selected), 'Insertion', 'Dessin', 'Mise en page', 'Formules', 'Données', and 'Partager'. The formula bar shows 'A10' and 'Texte'. The data starts at row 1 with columns A, B, C, D, and E labeled 'campaigns_type', 'sites_site_code', 'landmarks_lat', 'landmarks_lon', and 'campaigns'. Row 10 is highlighted in green, while rows 11 through 35 are pink. Cell A10 contains the value 'Camapigns_type'. A tooltip appears over cell A10 stating: "'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 téléchargement' (selected), and '7. Cha...'. Row numbers 1 through 35 are visible on the left.

	A	B	C	D	E
1	campaigns_type	sites_site_code	landmarks_lat	landmarks_lon	campaigns
10	Camapigns_type				
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

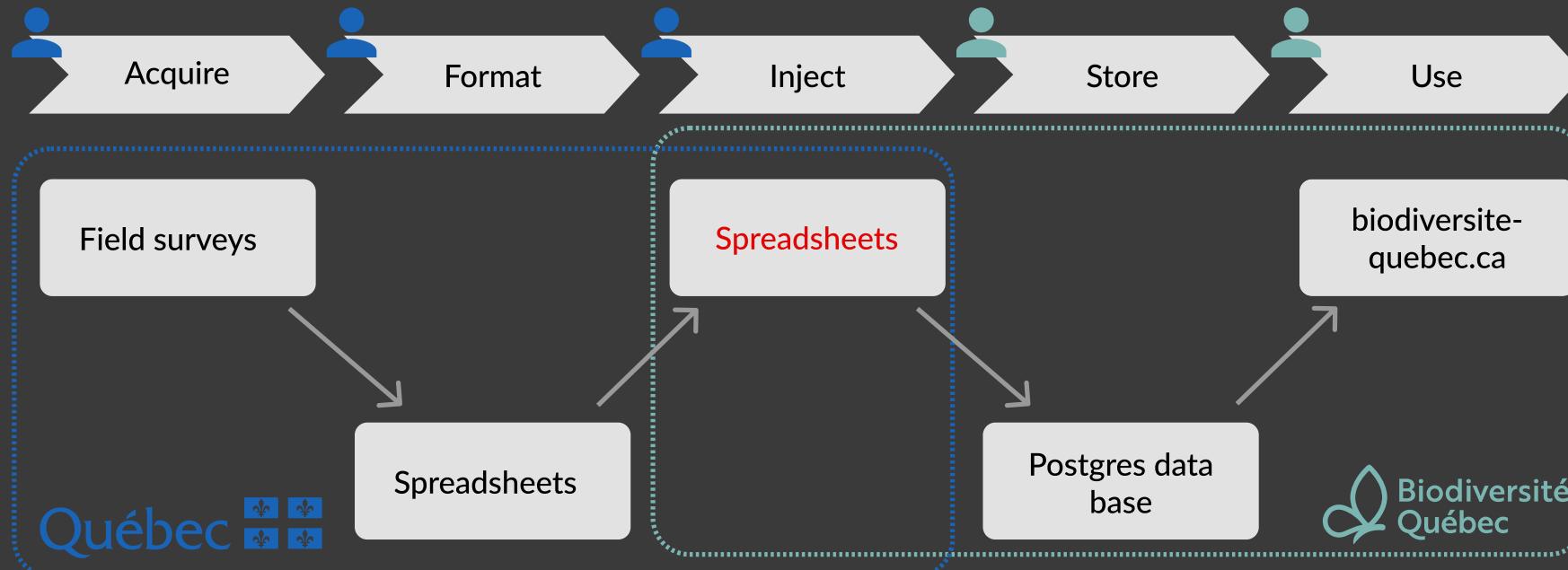
Template = data + metadata + formatting guide + validation rules

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

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

Reach out!

Victor Cameron

Eco-informatician, lecturer

victor.cameron@usherbrooke.ca

www.biodiversite-quebec.ca



Laboratoire
d'écologie
intégrative



Université de
Sherbrooke

