Discovering Data!

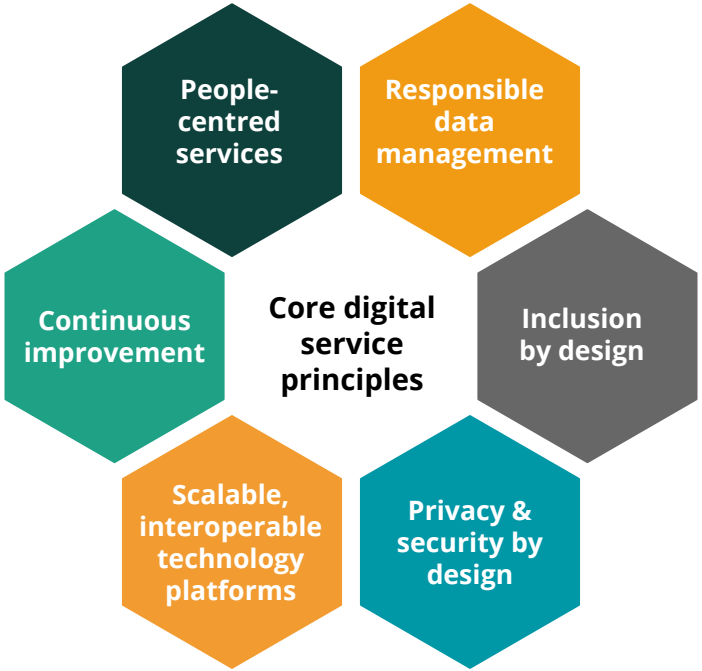
Gov Info Day
2019

Ontario Digital Service and Open Government

Encompassing Digital Services, Open Government and Data Governance the ODS leads and supports government initiatives to harness new technologies and approaches to make government simpler, faster, better.

Simpler, Faster, Better Services Act

- Announced in Budget 2019
- Puts people at the centre in the design, development and delivery of government services
- Mandates public release of non-sensitive government data
- Elevates into legislation principles outlined in the Digital Service Standard
- Formalizes in statute the role of the Chief Digital and Data Officer



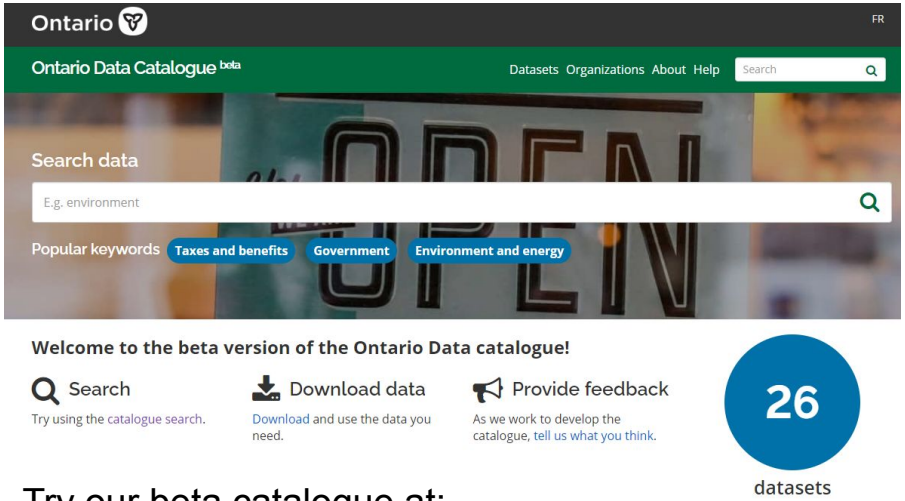
What we do!

The Open Government Team we are focused on bringing our initiatives to life and make information accessible by everyone!

We're best known for our Data Catalogue which we are updating for 2020.

To do this we are:

1) Updating our metadata schema
2) Exploring how we can share government APIs with the public
3) Conduct user testing of our beta catalogue



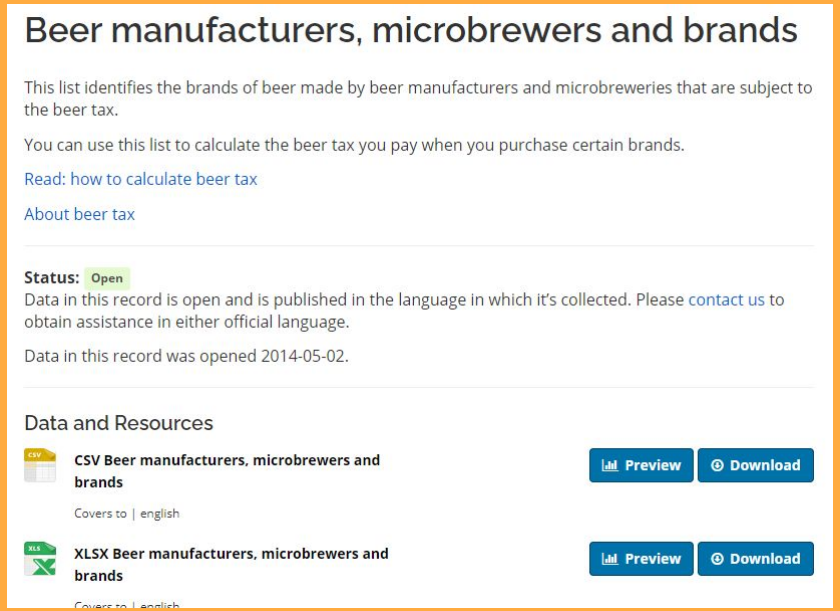
Try our beta catalogue at:
https://stage.data.ontario.ca/

What are we going to tell the public about our data?

Scope

We need a schema for our data catalogue.

Human Readable



Beer manufacturers, microbrewers and brands

This list identifies the brands of beer made by beer manufacturers and microbreweries that are subject to the beer tax.

You can use this list to calculate the beer tax you pay when you purchase certain brands.

Read: how to calculate beer tax

About beer tax

Status: Open

Data in this record is open and is published in the language in which it's collected. Please contact us to obtain assistance in either official language.

Data in this record was opened 2014-05-02.

Data and Resources

CSV Beer manufacturers, microbrewers and brands — Preview | Download

Covers to | english

XLSX Beer manufacturers, microbrewers and brands — Preview | Download

Machine Readable

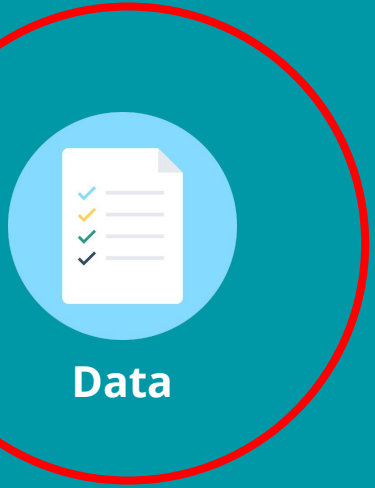

```
{
  "help": "https://data.ontario.ca/api/3/action/help_show?name=package_show",
  "success": true,
  "result": {
    "notes_translated": {
      "fr": "La liste identifie les marques de bière que font les fabricants de bière et les microbrasseri
      utiliser cette liste pour calculer la taxe sur la bière que vous payez lorsque vous achetez certaine
      (http://www.fin.gov.on.ca/fr/tax/bwt/taxincludedpricing.html)\r\n\r\n[Au sujet de la taxe sur la biè
      "en": "This list identifies the brands of beer made by beer manufacturers and microbreweries that ar
      beer tax you pay when you purchase certain brands.\r\n\r\n[Read: how to calculate beer tax](http://w
      tax](http://www.fin.gov.on.ca/en/tax/bwt/index.html)"
    },
    "license_title": "Open Government Licence – Ontario",
    "maintainer": "Open Government",
    "relationships_as_object": [],
    "private": false,
    "maintainer_email": "opendata@ontario.ca",
    "num_tags": 0,
    "maintainer_branch": {
      "fr": "",
      "en": ""
    },
    "keywords": {
      "fr": [
        "Impôts et avantages fiscaux"
      ],
      "en": [
        "Taxes and benefits"
      ]
    },
    "id": "5add9937-4663-47c9-9fd4-996eba649876",
    "metadata_created": "2019-11-07T15:15:45.585462",
    "title_translated": {
```


Metadata - two ways

Records

Apps

Information



Data

Data only

Doesn't need to be complete/perfect, can iterate

Core/MVP + Agile

Dataset
(e.g, Airports)

File (CSV)

File (KML)

File (JSON)

Two things to describe, actually

There are trade-offs

More metadata can be good

Longer and more arduous user experience is bad

Deciding on a schema

A field should:

- **Be useful** (does it answer user needs?)
- **Be clear** (contact information vs author email)
- **Be easy** (is it possible/probable for us to populate?)

Describe the field! (to not conflate names)

Step 1 - Start with CKAN

URL: https://data.ontario.ca/dataset/beer-manufacturers-microbrewers-and-brands
Title: Beer Manufacturers
Description:
[redacted]
Tags: [redacted]

Ministry: Ministry of Finance
Author/Email: [redacted] / [redacted]@ontario.ca
Maintainer/Email: [redacted] / [redacted]@ontario.ca

Licence: Open Government Licence

Created/Modified: YYYY-MM-DD / YYYY-MM-DD

Step 2 - Dublin Core (International standard)

URL: https://data.ontario.ca/dataset/beer-manufacturers-microbrewers-and-brands
Title: Beer Manufacturers
Description:
[redacted]
Tags: [redacted]

Ministry: Ministry of Finance
Author/Email: [redacted] / [redacted]@ontario.ca
Maintainer/Email: [redacted] / [redacted]@ontario.ca

Licence: Open Government Licence

Created/Modified: YYYY-MM-DD / YYYY-MM-DD

Not included: Subject, coverage, Language, Relation, Source

- Title
- Description
- Publisher
- Contributor/Creator
- Rights
- Created/ Modified

Step 1 - Existing Catalogue

URL: https://data.ontario.ca/dataset/beer-manufacturers-microbrewers-and-brands
Title: Beer Manufacturers
Description:
[redacted]
Tags: [redacted]

Ministry: Ministry of Finance
Author/Email: [redacted] / [redacted]@ontario.ca
Maintainer/Email: [redacted] / [redacted]@ontario.ca

Licence: Open Government Licence

Created/Modified: YYYY-MM-DD / YYYY-MM-DD

Access level
Is the data available?
Exemption
Reason why the data is restricted
Rationale not to Release
Explanation why the data is restricted
Opened Date
The date when the dataset finally has been opened (to the public on the open data catalogue)
Update frequency
How often new data should be added to dataset
Geographic coverage
Text term for the geographic scope of this dataset (e.g., “Ontario” or “Sudbury”)
Temporal coverage
Date range that the dataset describes

Step 4 - Who else?

- GO ITS-46 - **Common Metadata Elements Standard**
- Canada
- Alberta
- Dublin Core Terms
- Etc
- **146 different fields**

Step 5: Other needs

What does the public want/need? (What will we be able to provide?)

E.g.,Last updated

Step 5: Other needs

Do we have the fields we need to assess quality? (not the assessing or values themselves)

Metrics

- **Completeness** - **Primary** - **Timely** - **Accessible** - **Machine Processable and non-proprietary** - **Non-discriminatory and licence free**	- **Reusability** - **Understandability** - **Authenticity** - **Conformance** - **Accessibility** - **Retrievability** - **etc**

Summary - Dataset

URL: https://data.ontario.ca/dataset/beer-manufacturers-microbrewers-and-brands
Title: Beer Manufacturers
Description:

Tags:

Ministry: Ministry of Finance
Author/Email: / @ontario.ca
Maintainer/Email: / @ontario.ca

Licence: Open Government Licence

Created/Modified: YYYY-MM-DD / YYYY-MM-DD
Last time data added: YYYY-MM-DD
Last time anything was touched: YYYY-MM-DD
Current as of: YYYY-MM-DD

Access level: OPEN
Exemption: Privacy
Rationale not to Release

Opened Date: YYYY-MM-DD
Update frequency: monthly
Geographic coverage: Ontario
Geographic breakdown: none
Access instructions:

Summary - Resource

URL: https://data.ontario.ca/dataset/beer-manufacturers-microbrewers-and-brands.xls
Title: Beer Manufacturers - XLX
Description:

Format: XLSX

Version: 1
Temporal Range: YYYY-MM-DD - YYYY-MM-DD
Data birth date: YYYY-MM-DD
Data made public date: YYYY-MM-DD
Type: Data
Language: French
Geographic markers: None

Application Programming Interfaces (APIs) let systems talk to each other and exchange data in structured ways.

APIs are the connective tissue of digital services

Ontario

Building connected digital services: API Guidelines

Introduction

About this guide

This guide is for anyone working to develop digital services for a public service, whether as part of the Ontario Public Service, a government agency, or beyond.

These guidelines are meant to be a roadmap, not a roadblock. They aim to help streamline digital service development for a more consistent and robust product for both developers and end users.

What APIs can do

Application Programming Interfaces (APIs) let applications talk to each other in structured ways. Software developers create APIs to share some functionality or data from an application they've developed with anyone who might want to work with it.

For example, when you look at the weather app on your phone, the information you see is coming from an API built by your local weather service. Their API takes weather prediction data and converts it into a format that other app developers can use. This lets developers focus on building a user-friendly interface (for example, your phone's weather app) without having to worry about the science behind the data.

API Guidelines alpha

veen multiple related but independent systems. river's licenses could build an API that takes a useful for a car rental company that wants to

https://github.com/ongov/API-Guidelines/

APIs allow you to build reusable components and develop a platform, so that you don't have to reinvent the wheel every time. They are, in essence, the building blocks to a digital ecosystem. Understanding the value of APIs will help you shift

What's in the API Guidelines?

- Business and process considerations
  - API first, a.k.a. "eat your own dogfood"
  - Connections to the Digital Service Standard
  - When to do an API and when not to
- Technical guidelines
  - Data formats
  - Security and privacy
  - Documentation
  - And much more!

Eat your own dogfood


Talk to your users


Use modern (“RESTful”) design

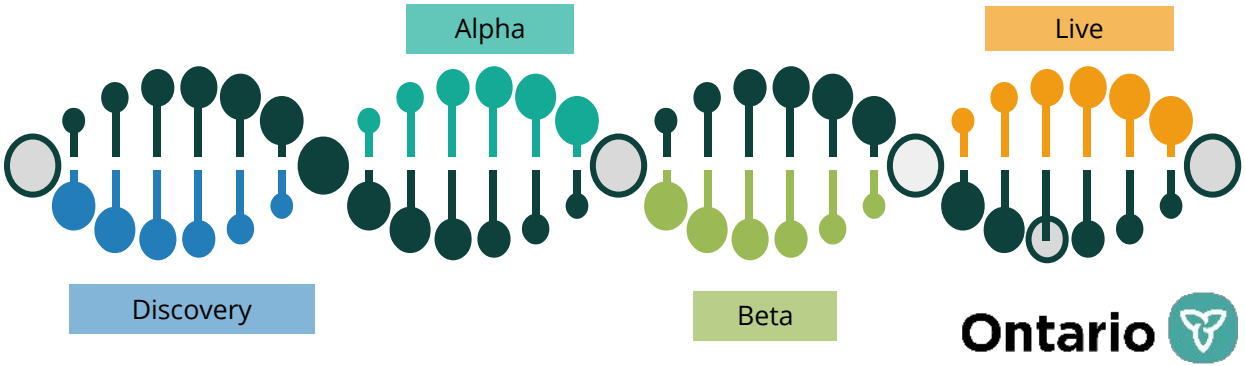
Practice security by design

Even if your API does not expose protected or privileged data, it is still important to consider how to embed secure architecture design into your system

- Use appropriate secure protocols
- Do not send sensitive data in URLs
- Control user access and permissions
- Build internal expertise

What's next

- **Improve the Guidelines** as people use them. Send us your feedback!
- **Create community of practice** of public servants interested in APIs.
- **Research and test** how best to to provide a place to share and discover APIs maintained by the OPS.



Possible vision:

The API Catalogue is an online collection of APIs and learning resources that makes it easy to use APIs that allow access to Government of Ontario data and services

Analytics

Google Analytics:
Current Catalogue Fun Facts

Oct 1, 2019-Oct 31, 2019

- Total sessions: **47,346**
- Total page views: **80,565**
- Average session duration: **00:01:53**
- New vs. Returning Visitor

Page Title	Total Downloads
Registered religious officials	2,628
Private school contact information	1,523
Fuels price survey information	901
Official forms	845
Crown timber charges for forestry companies	669
Ontario public library statistics	557
Well records	502
Municipal boundaries	362
Soil survey	306
Ministry of Natural Resources and Forestry topographic map	297
Top 10 Downloads Total	**21,853**

How can we compare portals?

- Information quality
  - Data sets metrics (Machine processable/readable, Non-discriminatory/restriction free, Non-proprietary)
  - Metadata metrics (Metadata availability, Publisher, Release data & up to date)
- System quality/access
  - Ease of use metrics (Friendliness, Documentation and Tutorial)
  - Performance metrics (API)
  - Search, provision, download metrics (Thematic categories, Open formats, Download, Visualizations and Analytics tools, Language (multiple), Searching & filtering capabilities)
- Service quality
  - Feedback and collaboration metrics (Personalisation, Forum (feedback) & contact form, Request form,Comments)

Customer satisfaction metrics: how are we doing?

- Overall customer experience
- % of visitors likely to return
- % of visitors likely to recommend
- Ease of getting information or using your service
- Timeliness of getting the information
- Does the information meet and/or exceed expectations?
- Relevance and usefulness of the information

Searching the catalogue

External Search

- Top referring commercial search engines
- Top referring search terms

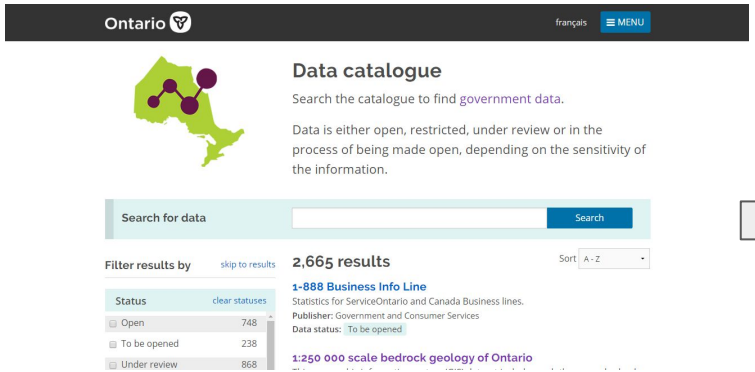
Internal Search

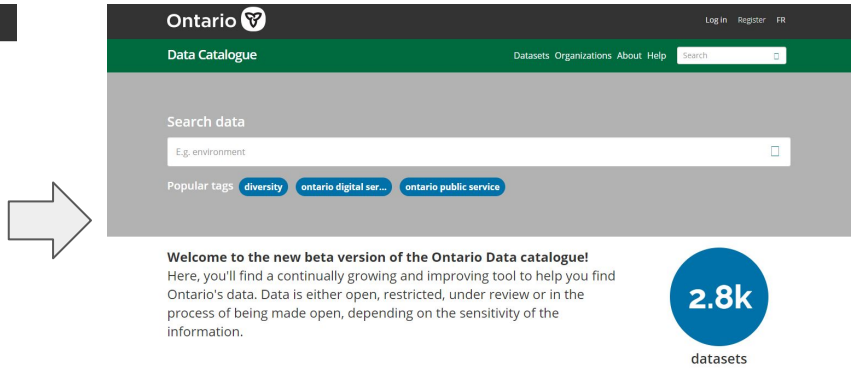
- Top search terms/phrases
- Top “no results” queries
- Percentage of visitors using site search

User Research

New Catalogue, who dis?

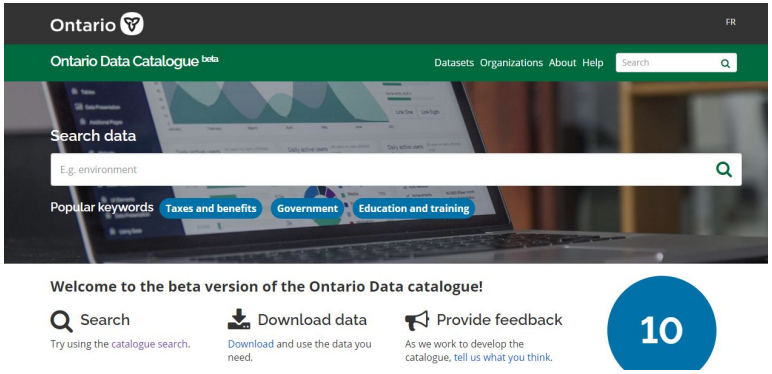
New Catalogue = opportunity to reboot



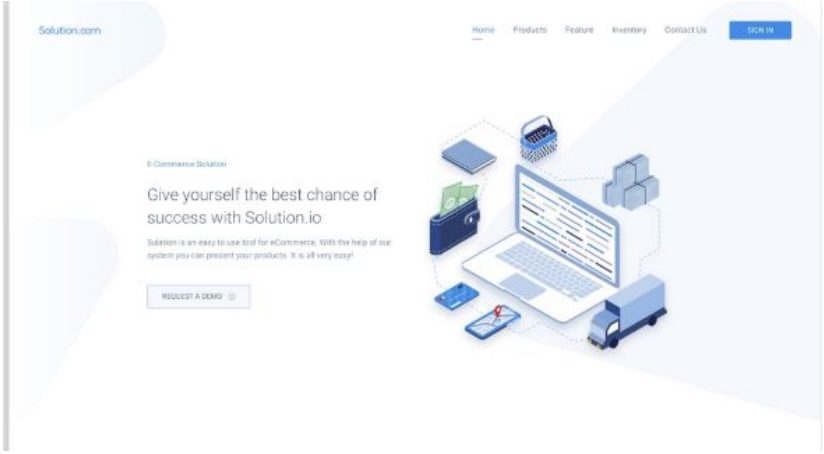


CKAN comes with preset language and layout...

Our UI



How we see our UI



How our Users see our UI

What must a catalogue be able to do well?

Users should be able to...

- Understand what a data catalogue is
- Find a dataset
- Understand whether data is available
- Understand when a dataset was last updated
- Understand when a dataset will next be updated
- Understand which file to click on

Can they?

The tasks: “Try to see if you can find a dataset that might answer your question.”

Task 1

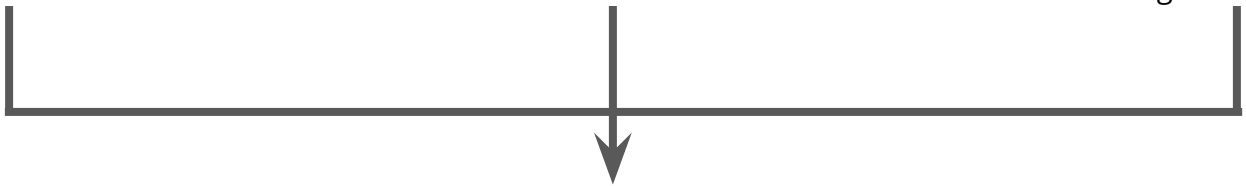
You’re thinking about placing your children in private school and you want to see what is available in your area.

Task 2

You want to find out if there’s a simple list of tax credits available in the data catalogue.

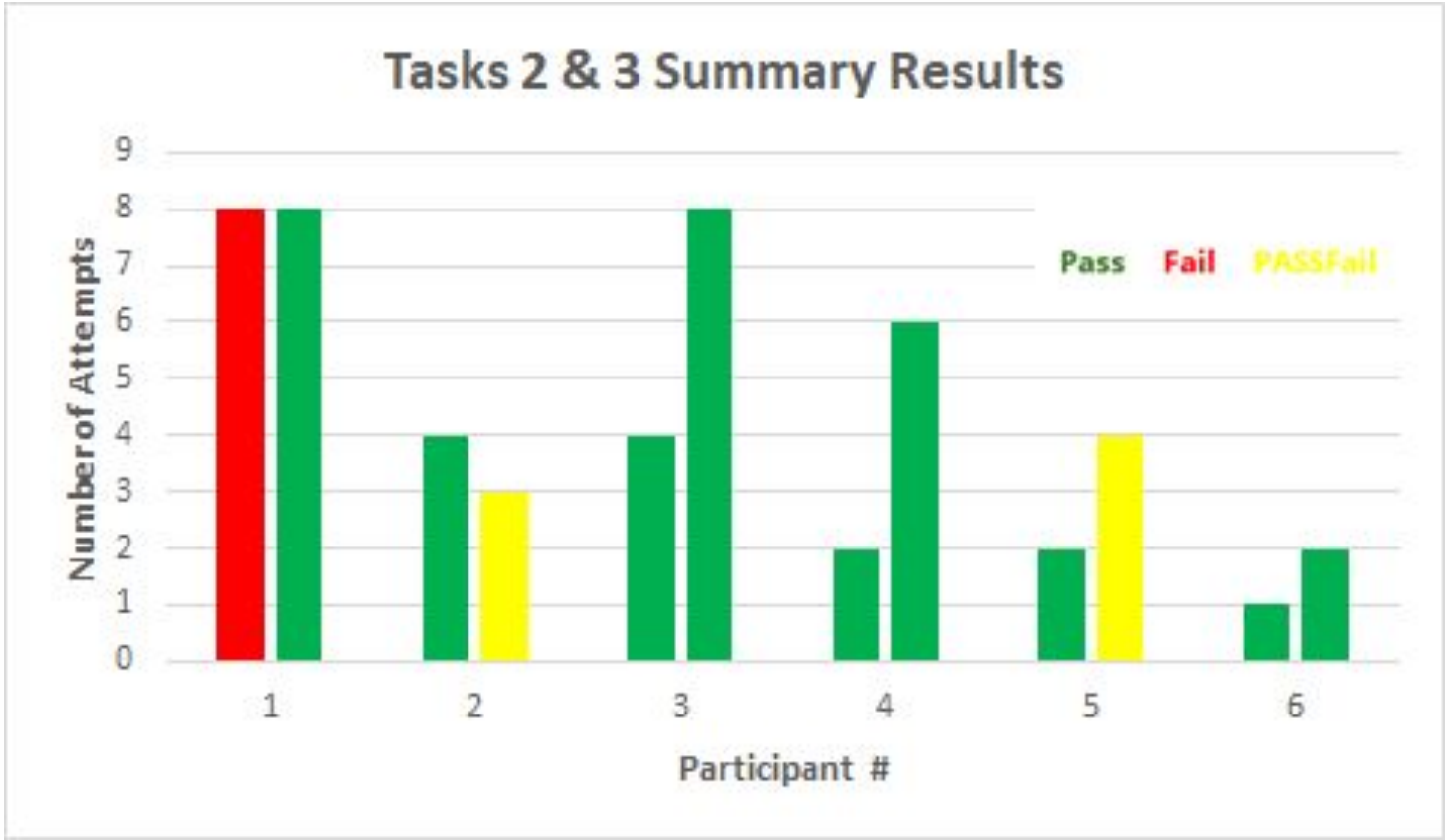
Task 3

You want to find out what the wait times are for calling the business info line. You want to check whether the data catalogue has this data.



...can you tell me whether you can actually access the data in this dataset?
...you want data from the most recent year possible. Can you tell me what years are available to you?
...can you tell me when you think this dataset will next be updated?
...can you tell me when you think this dataset was last updated?
...can you tell me what regions this data might include or describes?
...can you tell me in what languages the data is available?

Why usability testing?



And yet...

The next time you think you might need to use the data catalogue to complete a task, would you feel positive and comfortable before starting your task or negative and uncomfortable?

Positive	Positive	Positive	Positive	Positive	Positive

People will tell you their experience was fine, even if it wasn’t.

Summary of the results

Content/ Language	1. People are skimming by “status” but still seeking the information. 2. Users are confused by the term "maintainer email" 3. Users don't understand why non-open data is in the data catalogue
Navigation/ Organization	1. It isn't clear how to switch between english and french for resource data content. 2. Not clear to users that they are seeing all datasets or how to see all datasets
Visualization	1. Users try to map data without geographic information 2. Users find the spreadsheet visualization small and difficult to read 3. Graphing interface difficult to use
Search	1. Getting the search to yield the results they’re thinking of
Comprehension	1. Recognizing the metadata can answer their questions 2. Interpret and manipulate the dataset to answer their questions

Not just the issues..

Sometime people were **hesitant**. Are there opportunities to make them feel confident / make them feel like they’re not guessing?

Other people were successful but **frustrated**. Can we reduce friction?

People can give the right answer with the **wrong method**s

People can give the right answer **unconfidently** (or give the wrong answer confidently!)

People won't tell you what's wrong

1. **Collect observations**
   "Can I enter in latitude and longitude here?"

2. **What's the issue here?**
   Users don't know why there's a "map" option for some datasets

3. **Consider solutions**
   Could hide the map option for some datasets
   Could have more explainer text somewhere on what datasets are mappable

4. **Test (does this make it better?)**
   TBD

Comprehension

When you’re online shopping (or whatever), are you fully processing and reading, or is your brain taking advantage of UX shortcuts?

People are:

- Skimming
- Scrolling past

Despite that, how can we assure comprehension?

Examples:

can you tell me whether you can actually access the data in this dataset?

Access Level / Status - Ways of determining

1

Status: Open

Data in this record is open and is published in the language in which it's collected. Please contact us to obtain assistance in either official language.

Data in this record was opened 2014-03-03.

2

Data and Resources

Private school contact information

Covers 2019-10-04 to 2019-10-04 | english

Preview Download

VS

Data and Resources

This dataset has no data

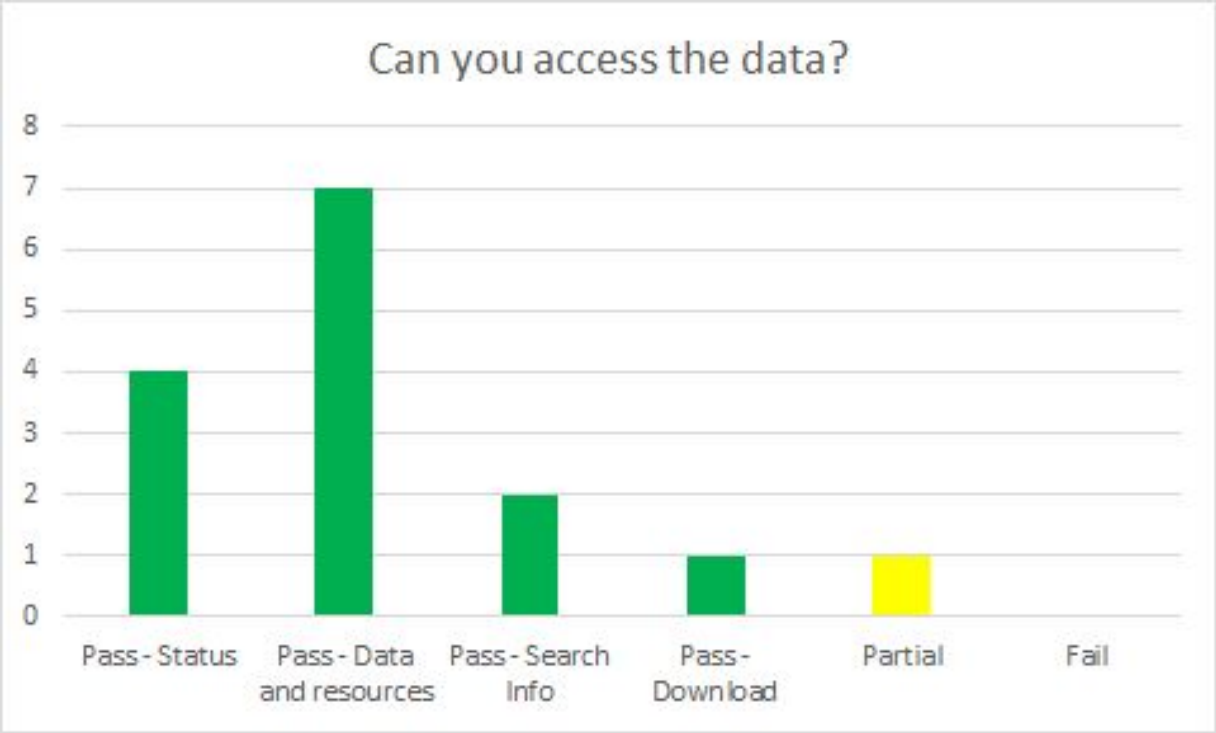
3

Average report card marks

Average marks data is no longer collected as of 2010. Median marks data is collected 2010 onwards for elementary schools for two reporting periods. Includes data for schools.

This dataset has no data

Access Level / Status - Success/Failure



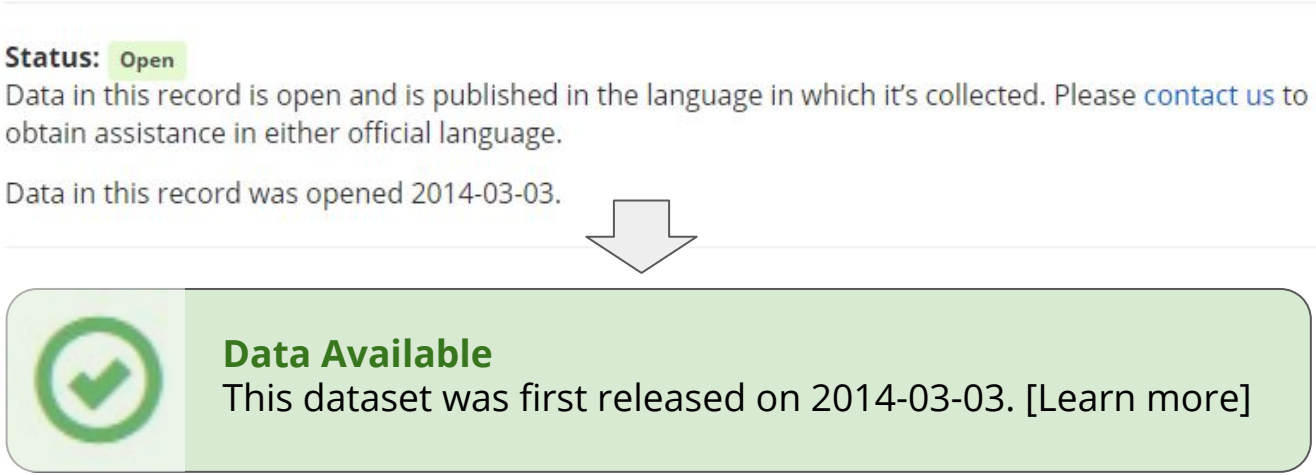
Access Level / Status - Issues

Status panel is not catching attention. Status panel language is not clear.

- 4 people continued to look for non-open datasets elsewhere
- 4 people expressed confusion about the definition of a status (even after reading the description out loud).
- 2 people didn't understand the point of having non-open data on the catalogue.
- While ultimate success was high, users **took a long time** to confirm non-open datasets weren't accessible and expressed frustration

Access Level / Status - Next steps

- Move status panel
- Simplify language
- Add graphical element



Test this in December!

User analytics and research to enhance data discovery

Thank you.

opendata@ontario.ca

https://www.ontario.ca/page/open-government

christine.hagyard@ontario.ca
bianca.sayan@ontario.ca
paul.vet@ontario.ca
cara.scarfone@ontario.ca
razieh.faraji@ontario.ca
allen.kwan@ontario.ca