



Canada Revenue
Agency

Agence du revenu
du Canada

Meeting Client Expectations for Timely and Accurate Tax Data

2019 Statistics of Income Consultants Panel Meeting –
Washington D.C., September 2019

About Us

- The Canada Revenue Agency (CRA) regularly provides tax and benefits data about individuals, corporations, trusts, excise, etc. to Canadian government departments and agencies at the federal and provincial/territorial levels.
- Data and statistics are also provided on demand and data sets published on the web.
- Commitment to Open Government – increasingly open by default.

Publications

- Government of Canada - Open Data
- CRA Corporate Reports

Ad Hoc Data Requests

- Media
- Ministerial correspondence
- Academics
- Think tanks & professional associations
- Industries and Municipalities
- Parliamentary Budget Officer
- General public

Let's go back...

- In 2015-16 the CRA received about 370 ad hoc requests for data/statistics and produced nearly 390 regularly scheduled products.
- Big data was becoming a “thing”.
- But some critical infrastructure and legacy source systems were reaching their end life and were not built with statistical reporting in mind.
- We had an opportunity!

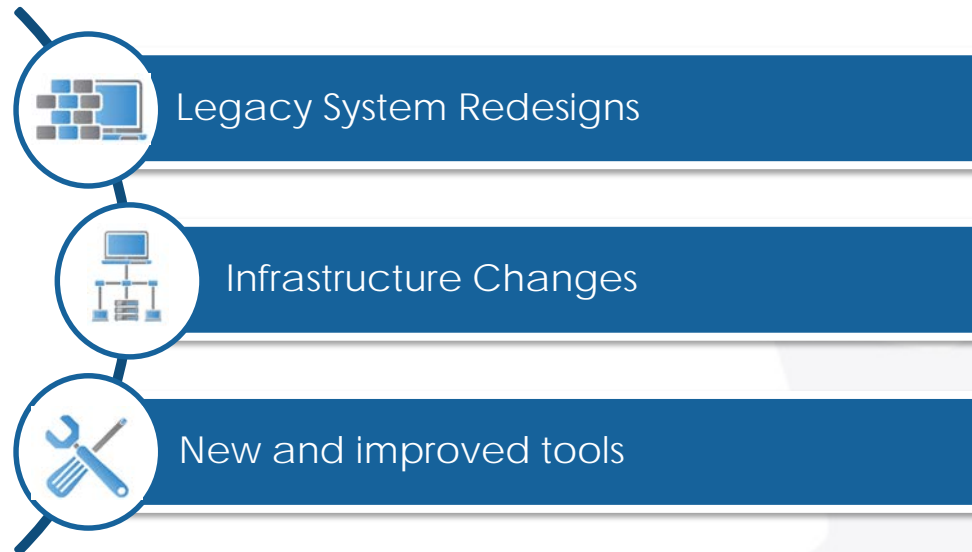
Fast forward four years

“The world’s most valuable resource is no longer oil, but data”
– *The Economist* May 6, 2017

- Computing power has increased
- The cost of storage has gone down
- The value of data is better appreciated for:
 - Decision Making
 - Predictive modeling
 - Behavioural insights
 - Measuring productivity
 - Service & compliance
- Demand for data and analytical services has increased and become more complex
- A.I., machine learning, and other emerging technologies are reshaping the way we look at data
- There is a seemingly endless supply of data
- We are responding

In the last year

- By 2018-19 the CRA received over 550 ad hoc requests and produced roughly 460 regular products...and we do it with about the same number of people.
- We established a Chief Data Officer (CDO) to provide horizontal leadership, optimize data business value, ensure high quality data, enhance our capacity to conduct research, promote/share expertise around BI.
- The CRA has instituted a growth fund for ongoing enhancements to the BI environment and Agency Data Lake.
- We got better and we got faster:



We do a lot to ensure data accuracy

- In May 2018, the Agency Data Clearing House was established to ensure tax and benefit data released by the CRA are: accurate; consistent; complete; and, subject to appropriate disclosure controls.
- Provide technical training and advice to support employees.
- Examine economic factors and legislative policy.
- Implementing a Data Quality Framework to:
 - Ensure adequate data quality assurance checkpoints are in place;
 - Define the roles and responsibilities and governance with respect to data quality;
 - Establish data quality standards and practices;
 - Develop and define data quality indicators;
 - Implement a data quality issues resolution process;
 - Ensure robust metadata practices are applied.

We do a lot to ensure data accuracy (cont'd)

- Our data quality assurance (DQA) processes:
 - Measure central tendencies (e.g., mean, median, mode, standard deviation);
 - Detect outliers and dominance;
 - Examine data over time;
 - Compare similar sets of data; and,
 - Serve as a feedback loop to functional programs/data owners

...but there are still challenges

- Clients seek access to data more frequently and some prefer to do their own data quality review.
- Increased transparency – rising expectation to make more data available publicly.
- Legacy systems are still undergoing redesign.
- Finite resources...can't just throw money at it.



Factors influencing timeliness

- Completeness (e.g., Preliminary vs. Final Data)
 - When are data ready and reliable for public consumption?
 - At what point is it reliable enough for policy research and analysis?
- Competing priorities – we are first and foremost a tax administration
- Complexity of data – how would you like that sliced?
- Data accessibility – it's captured where?
- Low tolerance for error – Data quality assurance takes time

Timeliness Challenges

- We need to be more efficient, but:
 - How much Quality Assurance is the right amount?
 - Data are always changing;
 - Manual processes still exist;
 - Tools and infrastructure are not yet optimized; and,
 - Did I mention resources?

How do we meet the needs of our clients?

- Communicate and engage
- (Re)design forms and systems that consider the needs of data and statistical reporting right from the start
- Continue to invest in infrastructure
- Employee (re)training

If we don't find the balance what happens?



Too Fast

- Inaccurate/missing information
- Misinterpretation
- Risk of disclosure
- Public image



Too Slow

- Data not available to researchers & policy makers or outdated
- Delayed decision-making
- Public image

Advice please...



From your perspective what can we do to achieve the right balance?

As demand for data and statistics continues to increase...



...we need to find the balance between timeliness and accuracy.