

**Data Quality Management – Award of Contract**

<b>Date:</b>	April 30, 2018
<b>To:</b>	Toronto Public Library Board
<b>From:</b>	City Librarian

**SUMMARY**

---

The purpose of this report is to seek Toronto Public Library (TPL) Board approval for an award of contract for implementation of Data Quality Management (DQM) to enable the data management workflow in support of TPL's services.

The Library issued a Request for Proposal (RFP) for the purchase of a DQM application. Ataccama Corporation was the successful proponent.

TPL's 2016-2019 strategic plan calls for transformation to a 21<sup>st</sup> century customer experience. The implementation of DQM will deliver business value by ensuring that all key initiatives and processes are driven with relevant, timely and trustworthy data.

Awards with a value greater than \$500,000 require Board approval.

**RECOMMENDATIONS**

---

**The City Librarian recommends that the Toronto Public Library Board:**

1. awards a contract for a Data Quality Management application to Ataccama Inc. for a total five-year commitment value of up to \$956,520, excluding Harmonized Sales Tax (HST), comprised of:
  - a. \$469,960 of capital costs; and
  - b. \$486,560 in operating costs.

**FINANCIAL IMPACT**

---

The Data Quality Management solution is funded from the Technology Asset Management Program (TAMP) capital budget project of Digital Services & Emerging Technologies (DSET) and the financial impact is described below.

The total amount of the award, \$956,520, is comprised of capital and operating costs over the first five years of the project as detailed in the table below. Awards with a value greater than \$500,000 require Board approval.

### **Data Quality Management System**

<u><b>Year</b></u>	<u><b>Capital</b></u>	<u><b>Operating</b></u>	<u><b>Total</b></u>
2018	469,960		469,960
2019		121,640	121,640
2020		121,640	121,640
2021		121,640	121,640
2022		121,640	121,640
<u><b>Total</b></u>	<u><b>\$ 469,960</b></u>	<u><b>\$ 486,560</b></u>	<u><b>\$ 956,520</b></u>

The capital costs of \$469,960 include product costs (Data Quality Center, Data Quality Issue tracker, Dashboard), professional services and other costs, including onsite support, first-year software maintenance and integration costs.

The operating costs of \$486,560 relate to annual support, maintenance and annual cloud infrastructure operations, and represent a one-time incremental operating budget cost of \$121,640 by 2019.

The Director, Finance and Treasurer has reviewed this financial impact statement and is in agreement with it.

## **ALIGNMENT WITH STRATEGIC PLAN**

TPL's 2016-2019 strategic plan calls for transformation to a 21<sup>st</sup> century customer experience. The implementation of a Data Quality Management Service supports this strategy of advancing the Library's digital platforms by fulfilling data quality dimensions such as completeness, accuracy, consistency and relevancy.

## **ISSUE BACKGROUND**

Data quality is affected by the way data is entered, stored and managed. Poor data quality generates maintenance and repair costs, and also affects customer satisfaction, reputation or even strategic decisions. Data volumes can grow inexorably, and Business Intelligence and analytics applications that depend on clean and relevant data become very crucial to business.

The following is a high-level summary of services aimed to be replaced/optimised:

- manual data entry
- disorganised data
- unintentional overwriting

- no definition of data ownership and data sources
- no process for data governance
- no authority of data source for TPL.

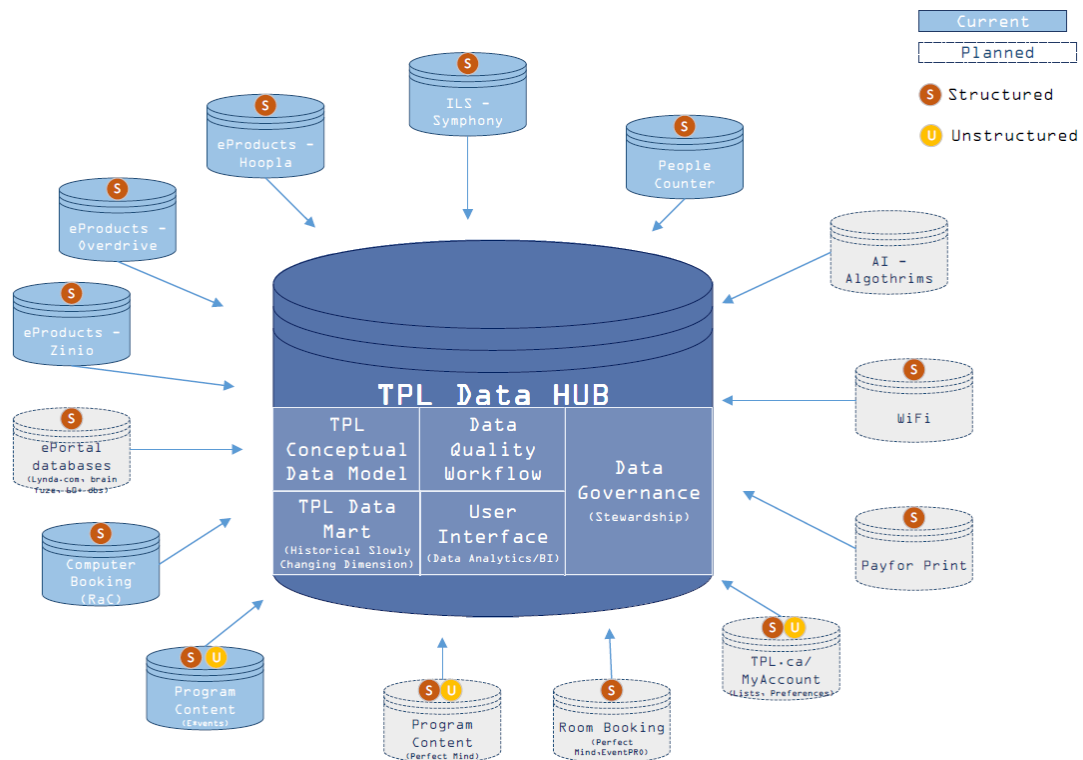
The recognition of similar processes triggered the release of the Request for Proposal for the selection of a software application(s) to serve the Library's business requirements for Data Quality Management.

## COMMENTS

The goal of the Data Quality Management Project are to:

- support general data quality criteria, including Cleansing, Matching, Linking, Parsing, Profiling, etc.
- provide Address Normalization
- provide De-duplication
- provide Data Quality Reporting
- provide Data Governance workflow
- support Data Integrity.

All Data sources defined in the following diagram are key sources for Data Quality Management.



As of 11/14/17

## **Procurement and Evaluation Process**

The Library issued an RFP for the purchase of a DQM Application and received two compliant proposals.

A cross-functional evaluation team comprised of members from different departments reviewed the proposals and assigned scores for proponents' pricing and technical capabilities, including primary services and optional value-added services.

The team reviewed the proposals and considered the information provided that was used to assign the scores for both proponent capabilities and services. The highest-ranked proponent, based on the aggregate scores of the proposals and financial arrangements, was shortlisted. Ataccama was the highest-scoring bidder and is selected to implement the Data Quality solution.

## **CONTACT**

Alex Hocevar; Director, Digital Services & Emerging Technologies; Tel: 416-393-7104;  
Email: [ahocevar@torontopubliclibrary.ca](mailto:ahocevar@torontopubliclibrary.ca)

EY Park; Manager, Information Management & CRM; Tel: 416-393-5907;  
Email: [eypark@torontopubliclibrary.ca](mailto:eypark@torontopubliclibrary.ca)

## **SIGNATURE**

---

Vickery Bowles  
City Librarian



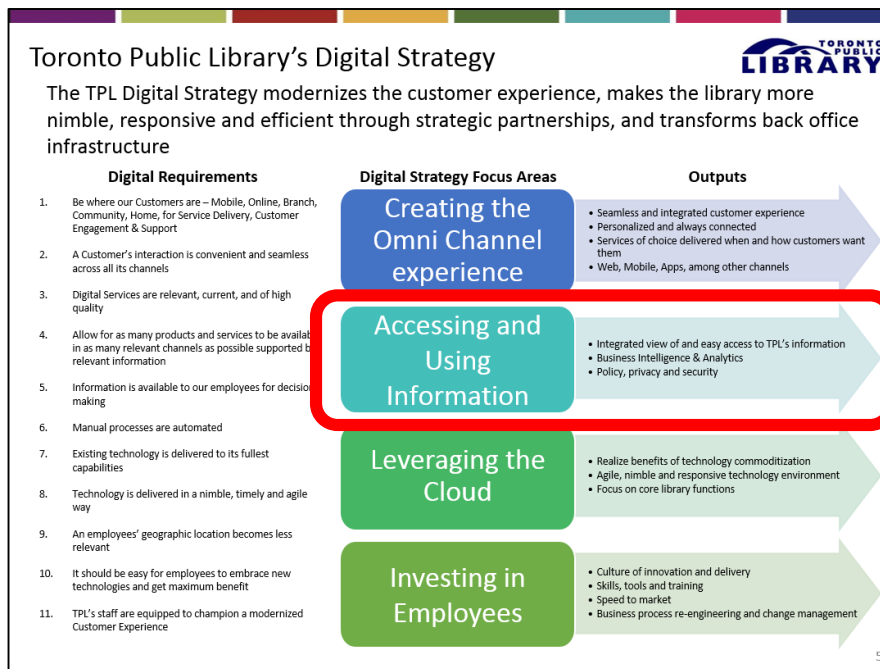
# Digital Strategy – Data Quality Management

Alex Hocevar & EY Park

April 2018

# TPL's Digital Strategy

TPL has identified the access and use of information as a focus area within its Digital Strategy



The ability for TPL to execute in this area is rapidly emerging

# Current State Assessment

What is the current state of TPL's Data Governance and Data Quality capability?

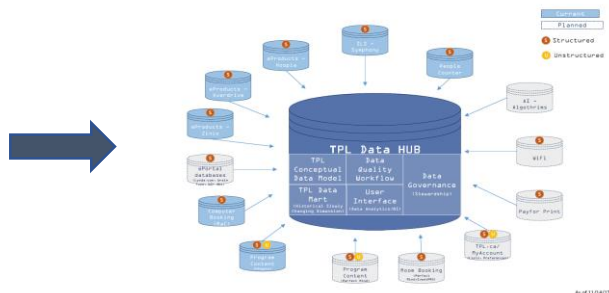
1. TPL has **multiple originations** of data.
2. Our data is very **disparate**.
3. Our technology and the systems that support data are **undisciplined**.
4. All data captured at TPL is **accessible**.
5. We **track** customers and their transactions to varying degrees.

To be effective,

**Data must be accurate, accessible and repeatable.**

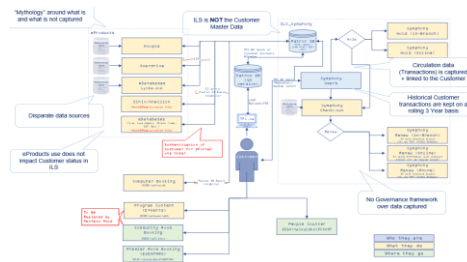
TPL has begun work in a number of areas to enhance it's Data capabilities:

1. One version of the “Truth”



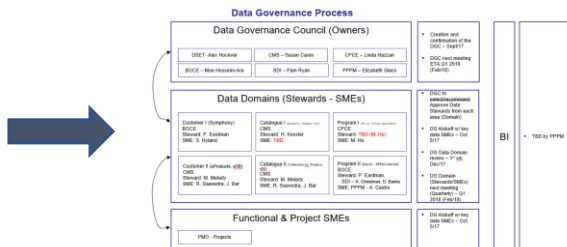
- TPL Data Hub
- Consolidated view of Data across TPL

- ## 2. Common Data Model across TPL



- Data Flows, definitions, triggers, reporting standards, data exchange

- ### 3. Data Governance Framework



- Data Governance workflow
- Data Quality Management Application



# TPL's Data Quality – Examples of Challenges

## Before

### Customer

Name: Doe, John **Doe Doe**  
Profile: **ALL**  
Gender: **<Blank>**  
Language Pref: **<Blank>**  
Date of Birth: **January 1, 1900**  
Address: 123 Yonge St.  
City: **New York**  
Country: Canada**a**  
Postal Code: **H0H H0H**

Correct Duplicates

Trigger call to  
action to verify  
required fields

Standardize  
addresses

Flag incorrect /  
default data entry  
& process errors

## After

### Customer

Name: **Doe, John**  
Profile: **Adult**  
Gender: **Male**  
Language Pref: **English**  
Date of Birth: **February 11, 1956**  
Address: **123 Yonge St.**  
City: **Toronto**  
Country: **Canada**  
Postal Code: **1L3 3A6**

### Item/ Catalogue

Title: Thirteen ABC  
ISBN: 123456, **Paperback**  
Author: **1998**; Smith, Mary  
Publisher: **New Y**, YY Books, **1998**  
Description: **Pbk.**

Business rule  
driven parsing

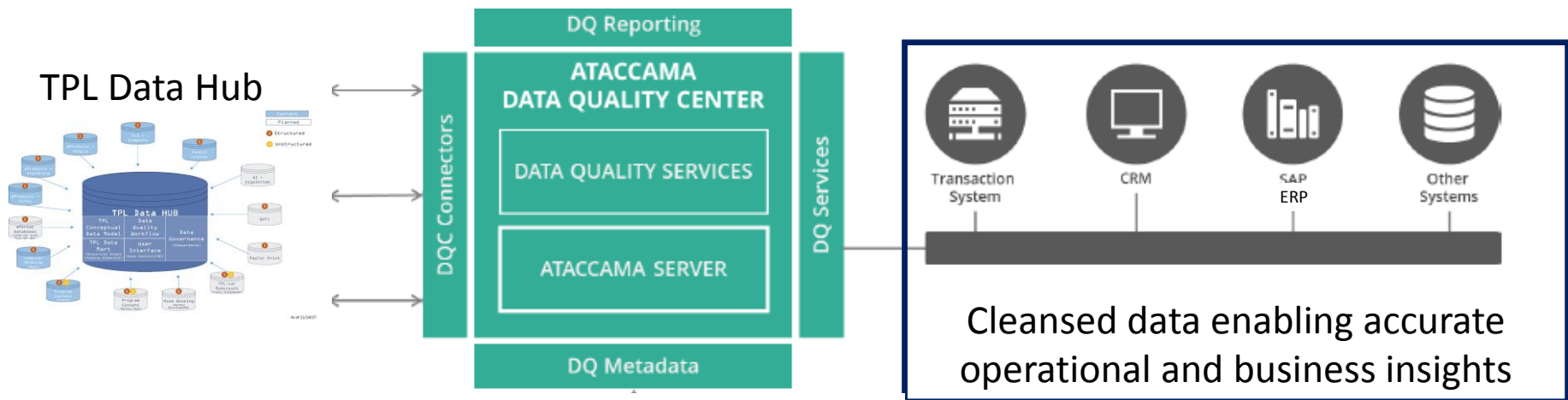
TPL owned data  
dictionary to  
replace system/  
inconsistent  
data/values

### Item/ Catalogue

Title: **Thirteen ABC**  
ISBN: **123456**  
Author: **Smith, Mary**  
Publisher: **YY Books**  
Description: **Paperback**

# Data Quality Management Application – What is it?

The TPL Data quality Management application is a new technology solution which will assist in managing Data Governance workflows and will automate the standardization and cleansing of Data from across the organization.



## Assist with:

1. Manual data entry issues
2. Organizing data
3. Reducing unintentional overwriting
4. Definition of data ownership and data sources
5. Establishment of process for data governance
6. Creation of authoritative data source for TPL

## Budget & Next Steps

### Budget

- Capital \$486,560
- Annual Operating Costs \$121,640

### Next Steps

1. Award of RFP
2. Develop project plan with a 6 to 8 month implementation timeline
3. Continue to enhance TPL Information Management capabilities