

DATA MANAGEMENT SCORECARD

OVERALL RATING

Overall Health Check Rating is based on the rolled-up score of all areas defined within Data Management: Data Classification, Lifecycle Management, and Document Management.

IT Infrastructure Health Check: Data Management

May 16<sup>th</sup>, 2017

| Last Year | Data Management   | Health Check Rating |
|-----------|---|---------------------|
| N/A       |  | 1.7                 |

DATA MANAGEMENT

VALUE SCORECARD

DATA MANAGEMENT

| RATING LEVELS | DATA MANAGEMENT      |  MEETS NEEDS |  RELIABILITY |  OBSOLESCENCE |  COST \ EFFORT |
|---------------|----------------------|---|---|--|---|
|               |                      |   |   |  |   |
| 3 HIGH        | Data Classification  |   |   |  |   |
| 2 MEDIUM      | Lifecycle Management |   |   |  |   |
| 1 LOW         | Document Management  |   |   |  |   |

ITEMCURRENT RATINGPREVIOUS RATING

DATA  
MANAGEMENT

1.7

N/A

There are no formal data management policies in place, resulting in insufficient controls and management of the Regions storage, incurring increased data management costs.

Inactive data (that is data that hasn't been accessed in more than a year) accounted for 50% of data residing on a primary storage volume sampled during the assessment - This was the equivalent of 25TB of primary disk.

There are no data classification policies in place. Region staff identified sensitive data (e.g. personally identifiable) exists throughout the various storage systems, however there has not been an assessment to determine the extent of this data.

The Region has procured a document management system (eDocs) and is currently well positioned to deliver a corporate solution.

Please refer to the [Data Management](#) section for the complete review and recommendations

# FINDINGS AND RECOMMENDATIONS

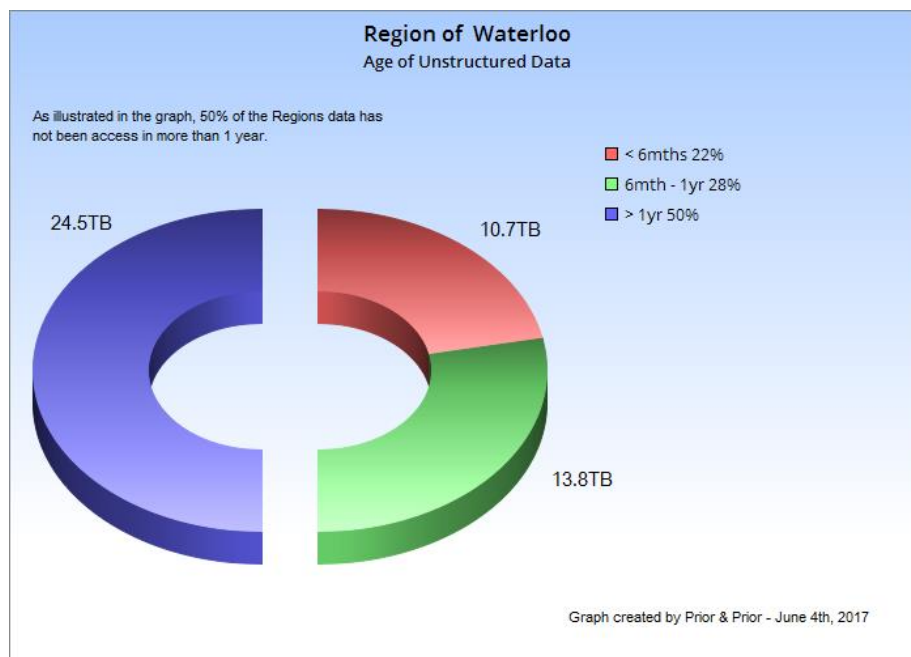
## DATA MANAGEMENT

### DATA LIFECYCLE MANAGEMENT

A data scan was performed on 49TB of unstructured data residing on the Compellent SAN.

The results of the assessment determined that:

1. There are no formal data management policies in place, resulting in insufficient controls and management of the Regions storage, and incurring increased data management costs;
2. Inactive data (that is data that hasn't been accessed in more than a year) accounts for 50% of data residing on a primary storage volume
3. Without formal storage management practices the proliferation of data over the next 3 years will create substantial issues with regard to security, availability, scalability, and cost of data.



## Recommendations

### <sup>10</sup>Data Lifecycle Management (DLM)

Data Lifecycle Management (DLM) is a policy-based approach to managing the flow of an information system's data throughout its lifecycle – from creation and initial storage, to the time it becomes obsolete and is deleted, or is forced to be deleted through legislation.

DLM products attempt to automate processes involved, typically organizing data into separate tiers according to specified policies, and automating data migration from one tier to another based on those criteria. As a rule DLM stores newer data, and data that must be accessed more frequently, on faster, but more expensive storage media. Less critical data is stored on cheaper, but slower media.

### Data Archiving

Retention schedules are the foundation of a successful records management process. These schedules take into account an organization's legal, regulatory and operational requirements while providing guidance on how long records need to be kept and what to do with them when they are no longer needed. It is important to develop a schedule for backing up and archiving all computer records and for keeping current copies of all paper and computer files off-site and accessible.

To do so, it is vital for the Region to determine **what data needs to be backed up and what data needs to be archived**. This can be done with the business users of the data. At the same time, the archival retention periods for the various types of data need to be determined. It is also vital to remember that not all data has the same requirements, while some data may have overlapping requirements.

Backup and archive serve different purposes (recovery vs long-term preservation and retention and there are three (3) key differences between the two solutions:

1. The Data
2. Access
3. Disaster Recovery

---

<sup>10</sup> A proper DLM product provides software for data archiving, retention, regulatory compliance, storage consolidation, backup and recovery optimization, and capacity management across a tiered-storage infrastructure. The software redefines storage management by enabling organizations to profile the criticality of their data within its lifecycle to automate discovery, classification, and placement on the most appropriate storage resource.

## DATA CLASSIFICATION

The Region has no formal data classification processes or policies in place at this time, however Information Management and Archives (IM&A) published a <sup>11</sup>5-year Master Plan (2015 - 2020) that defines the future direction of information management for the Region. This document should be followed with ITS working directly with IM&A to address the requirements for data classification.

It should also be noted that in June, 2017 a project was initiated for the development of PCI DSS compliance within the Region. The first step in protecting credit card and customer data is to know what sensitive data you really store, classify what you have and set up the appropriate security controls. ITS will be required to work with the Finance department to help determine the appropriate technological solutions(s) to address these requirements.

A successful <sup>12</sup>data classification project tells you what data you have, how it's classified, and that it's stored and accessed efficiently. As part of a review of cybersecurity and data loss prevention (DLP) requirements, the Region needs to consider the development of corporate data classification practices.

Please refer to [Appendix A: Data Classification Protocol](#) for a recommended approach

## DOCUMENT MANAGEMENT

The Region has procured software (eDocs) that could be integrated with a data archiving solution as part over an over-arching data management program.

---

<sup>11</sup> Document #1822456

<sup>12</sup> At a high level, **Data Classification** is the process of collecting the business requirements of data and applications, and using those requirements to store, protect and manage data at the appropriate service levels. A data classification project must begin with a definition of what's being classified and what metrics are appropriate for the level of classification desired.