



# Points de parole de la Dirigeante principale de l'information

Sommet Big Data & Analytics Canada 2020

Version 1.0  
2020-02-11



Shared Services  
Canada

Services partagés  
Canada

Canada

## **Quelles sont les origines de la stratégie en matière de données**

### **À propos du gouvernement numérique**

- Le gouvernement du Canada (GC) réimagine ses relations avec les citoyens Canadiens
- Afin de garantir que les besoins des Canadiens soient prioritaires, le GC rationalise et sécurise la prestation de services
- Nous voulons garantir une expérience cohérente pour tous les Canadiens en supprimant la nécessité de naviguer entre plusieurs institutions pour effectuer une transaction de service

### **À propos de la feuille de route des données et de l'appel à des stratégies de données ministérielles**

- Notre capacité à réimaginer notre prestation de services dépend de notre capacité à exploiter les données (vieux proverbe : on ne peut pas améliorer ce que l'on ne mesure pas)
- Les ministères fédéraux traitent des quantités croissantes de données et ont la possibilité de tirer parti d'une variété de solutions numériques; les modèles commerciaux évoluent, tout comme les exigences en matière de ressources et d'expertise;
- Avec un paysage de plus en plus complexe et une demande de plus en plus forte des Canadiens pour une prestation de services centrée sur l'utilisateur, les ministères fédéraux doivent passer des données comme résultat des fonctions et des systèmes des programmes aux données comme actif stratégique de l'entreprise
- L'appel lance aux ministères pour qu'ils produisent des stratégies de données est né de ce besoin. Ces stratégies visent à déterminer comment chaque service fera un usage plus stratégique des données tout en protégeant la vie privée des citoyens

## **Comment nous passons des tableurs aux lacs de données**

### **Qu'est-ce qu'un lac de données et en quoi il diffère des autres approches**

- Un lac de données est généralement un magasin unique de toutes les données d'entreprise, y compris des copies brutes des données du système source et des données transformées utilisées pour des tâches telles que la génération de rapports, la visualisation, l'analyse avancée et l'apprentissage automatique. Un lac de données peut inclure des données structurées provenant de bases de données relationnelles (lignes et colonnes), des données semi-structurées (CSV, journaux, XML, JSON), des données non structurées (courriels, documents, PDF) et des données binaires (images, audio, vidéo).
- Le «lac de données» du programme d'analyse opérationnelle d'entreprise (PAOE) de Services partagés Canada (SPC) (également appelé dépôt de données de l'entreprise (DDE) est actuellement une architecture de base de données contenant uniquement des données structurées et semi-structurées utilisant Oracle.
- PAOE a différentes couches au sein de son architecture pour prendre en charge le flux de données, y compris le marécage (copie brute de toutes les données requises de la source), le transfert (nettoyage technique de données), le lac (agrégation et intégration des données) et le lac conservé (centré sur l'utilisateur, affichages de données).

### **Différence entre le DDE et les entrepôts de données traditionnels**

- Nous prenons toutes les données source une fois pour satisfaire les besoins de reportage actuels mais aussi futurs
- Nous ne nous concentrons pas sur la satisfaction d'exigences de reportage spécifiques, mais plutôt sur la fourniture de données d'entreprise comme un atout ou le consommateur peut explorer et produire des rapports selon les besoins
- Toutes les données sont ouvertes par défaut (et seules les données non protégées sont importées dans le DDE)
- Nous ne corrigeons/ne nettoyons pas les données dans le DDE- contrairement aux entrepôts de données, le cadre de qualité des

données du programme de gestion des données traitera les problèmes de qualité des données et procédera à la correction de la qualité dans la source.

## **Quel est notre objectif**

Améliorez la prise de décision axée sur les données à SPC en créant un référentiel central où toutes les données de la SPC d'entreprise sont hébergées pour:

- Diminuer notre fiabilité sur les feuilles de calcul et d'assembler les données à la volée
- Soutenir et améliorer les rapports et analyses reproductibles, cohérents et opportuns

## **Comment allons-nous y arriver**

- Intégrez toutes les données d'entreprise de valeur d'entreprise dans le DDE pour fournir des données comme un atout
- Autoriser l'accès ouvert par défaut à la DDE pour les consommateurs de la SPC afin de soutenir leurs efforts de prise de décision
- Établissez des relations entre les domaines de données pour permettre aux consommateurs de données de 'coudre' facilement deux domaines de données
- Construire des données de référence cohérentes pour permettre une approche cohérente de rapport-parler des pommes avec des pommes.

## **Comment augmentons-nous notre connaissance des données**

- L'objectif est de démystifier les données: les rendre réelles, comprendre leur importance
- Nous concevons et mettons en œuvre une stratégie de gestion du changement organisationnel (GCO) comprenant des programmes de sensibilisation et d'éducation liés aux données

## **Sur quelle technologie repose notre DDE**

Ensembles de données de:

- Système de suivi de l'intégration opérationnelle (SSIO) – système SPC
- Bureau de contrôle d'entreprise (BCE) – système SPC

- Gestion des services de technologie de l'information – Bureau de contrôle d'entreprise (GSTI-BCE) – Autres extraits ministériels de données BCE
- Technologie de l'information du gouvernement du Canada (TIGC) Plan stratégique – extrait SCT
- La gestion du portefeuille d'application de données (GPAD) – extrait SCT

Technologie pour soutenir l'intégration et la livraison des données BCE:

- Base de données Oracle
- Transfer de fichiers gérer et sécuriser (TFGS) pour transporter en toute sécuriser des ensembles de données depuis l'extérieur de SPC vers l'infrastructure BCE
- Services d'intégration des serveurs SQL (SISS) pour extraire, transformer et charger (ETC) les données à travers le BCE
- SAS pour façonner, rapporter et analyser les données BCE

## **Quel est notre état final et quel est notre chemin pour y arriver**

- L'état final est basé sur les recommandations de Gartner
- Centre d'excellence en analyse de données (partenariat entre les entreprises et les technologies de l'information / gestion de l'information) - avec un modèle en étoile
- L'analyse en tant que service
- Les données en tant que service
- L'objectif au cours de la troisième année et au-delà est de passer de l'analyse «de rapport» à l'analyse «prédictive» et «normative» afin de permettre des décisions d'affaires. La voie pour y parvenir est de continuer à construire le dépôt de données de l'entreprise, à continuer de favoriser et développer l'expertise en matière de culture et d'analyse des données, d'assurer le financement après la troisième année, de définir les rôles de dirigeant principale de l'analyse et de dirigeant principale des données, de continuer à renforcer les capacités technologiques (y compris l'apprentissage automatique, Intelligence artificielle); et continuer à définir et à reproduire la culture des données, la littératie des données et l'ouverture par défaut (tout en assurant la confidentialité et la sécurité des données)

## **Notre approche pour les intendants de données**

### **Qu'entend-on par «intendants de données»**

- Les intendants de données sont les personnes d'une organisation responsable de la sécurité, de la qualité et de l'accès aux données de leur secteur de programme
- Ils établissent, mettent en œuvre et appliquent les politiques, les processus, les procédures et les normes en matière de données.
- Ils gèrent également les problèmes de qualité des données tout en protégeant contre la perte de données.
- Ils sont essentiels à la fourniture d'actifs de données fiables qui permettent aux dirigeants de prendre des décisions éclairées.

### **Qui fait un bon intendant de données**

- Familiarité avec les concepts de base de données
- Une bonne maîtrise de l'intégration et de la gestion des données
- Connaissance spécifique de la plate-forme (s) en place
- Une bonne compréhension des pratiques exemplaires en matière d'entrée de données/mettre à jour les meilleures pratiques
- Forte orientation du service à la clientèle.
- Capacités éprouvées d'analyse et de résolution de problèmes.
- Bonnes compétences en communication écrite, orale et interpersonnelle.
- Capacité de mener des recherches sur les questions de données et, au besoin.
- Capacité de présenter des idées dans une langue convivial et favorable aux affaires

### **Quelles sont nos attentes envers eux**

- Ils appliquent une saine gestion des données, conformément à leur rôle et à leurs responsabilités
- Ils participent activement à la communauté de pratique des intendants de données de SPC
- Partager et promouvoir des normes, des modèles et des protocoles de données communs pour faciliter la bonne gestion et l'utilisation des données à travers SPC et avec des organisations externes (par exemple, les normes API, les normes de données ouvertes, les modèles d'échange de données)

- Guide du changement de culture dans l'ensemble du GC en sensibilisant les gens aux outils, aux conseils et aux normes existants et en plaidant pour leur utilisation
- Identifier les opportunités d'échange de données en fonction de la valeur commerciale du Ministère et du GC
- Collaborer avec d'autres groupes et communautés du GC (p. Ex. La communauté de pratique des données d'entreprise du GC, l' Académie du numérique et les services numériques canadiens) pour assurer l'harmonisation des efforts visant à tirer parti des données en tant qu'actifs commerciaux et stratégiques
- Aligner sur la norme numérique GC «Soyez un bon intendant de données»

## Gouvernement Ouvert

Le gouvernement ouvert (GO) est un moyen de sauvegarder notre démocratie et d'alimenter les solutions à des problèmes complexes.

### Avantage

- Répondre à la demande croissante d'informations gouvernementales
- Besoin croissant à la fois d'une fonction publique engagée et sensible aux besoins du public et de la participation des citoyens
- Développement rapide des opportunités numériques transformant la vie quotidienne des citoyens, des secteurs public et privé
- Reconnaissance croissante du potentiel du gouvernement ouvert pour réduire les coûts, créer des gains d'efficacité et augmenter la productivité à l'intérieur et à l'extérieur du gouvernement

### Exemples

- Guide de natation et guide aquatique à but non lucratif de Toronto
  - Ils intègrent des données provenant des administrations municipales, provinciales et fédérales pour maintenir à jour des informations sur les niveaux de sécurité des plages et des sources d'eau de l'Ontario.
- Rappeler des applications
  - Applications pour fournir des informations de rappel à la fois sur un réfrigérateur intelligent connecté par Samsung et des assistants vocaux utilisant Google Home et Amazon Alexa.
  - API du portail du gouvernement ouvert sur les rappels et les alertes de sécurité fournie par Santé Canada. La réutilisation d'informations

provenant de plusieurs sources, y compris l'Agence canadienne d'inspection des aliments et le site Web Canadiens en santé, permettrait aux utilisateurs de rechercher des rappels, des avertissements et des alertes pour des aliments, des produits de consommation et des produits médicaux.

- Népal 2015 tremblement de terre - données entraînées secours en cas de catastrophe
  - D'énormes quantités de données publiées sur le Humanitarian Data Exchange coupant à travers les questions de compétence liées à la mobilisation internationale et aux organisations multilatérales impliquées.
  - Les données indiquent les sources d'eau, les hôpitaux et la nécessité d'une aide immédiate
  - A aidé à diriger les efforts là où ils étaient le plus nécessaires

## **Libre-service avec d'autres départements**

- Grâce à la solution cloud que nous testons actuellement, SSC peut fournir des données et des outils d'analyse non classifiés à d'autres ministères pour utiliser nos données pour leurs propres analyses. Permet la co-création de tableaux de bord et de visualisations, afin de permettre à SSC et à ses partenaires d'obtenir des informations sur nos données.
- En raison de la multitude d'outils dans notre solution cloud, permettra à tout utilisateur, quel que soit son niveau de compétence, de consommer nos données à des fins d'analyse. Les ensembles d'outils incluent des outils pour les utilisateurs non experts ayant une expérience semblable à Google (Thoughtspot), des outils pour les ressources de visualisation (SAS) et des outils pour les scientifiques de données avancés (R, Python).
- En plus des aspects techniques, une feuille de route et une stratégie de promotion sont en cours d'élaboration afin d'annoncer la valeur commerciale.

## **Capacités d'analyse**

- la mise en place d'un référentiel de données d'entreprise axé sur le chargement d'ensembles de données clés dans le BCE, afin d'effectuer des rapports et des analyses de nos services
- Définition d'une source unique de vérité et de données fiables pour l'analyse
- Élaboration de rapports standard à bouton-poussoir pouvant être utilisés pour plusieurs lignes de services



## Que cherchons-nous

Utiliser des données d'entreprise clés dans notre entrepôt de données, rechercher des réponses et des informations sur nos services, telles que: comment nous nous débrouillons en tant que fournisseur de services, quels sont nos domaines de faiblesse, où sont nos goulots d'étranglement, combien de temps faut-il pour fournir un service , temps pour résoudre l'incident, etc.)

## Qu'avons-nous appris d'un point de vue technique / analytique

- Vous devez d'abord être axé sur les données - commencez par les bonnes données
- Besoin d'inclure la gestion des données (qualité des données, gouvernance des données, données de référence, programme d'alphabétisation des données)
- Un outil ne fera pas tout. Exigence de diversité de l'ensemble d'outils d'analyse afin de fournir l'éventail complet des capacités au sein du département (différents outils basés sur différents rôles, et différents besoins). Cependant, ne veulent pas d'un gratuit pour tous, donc besoin de déterminer les profils d'utilisateurs, et d'attribuer des outils en conséquence
- La conception des capacités et de l'infrastructure analytiques est complexe, le temps de se déployer longtemps dans une solution sur place. Cloud plus rapide à déployer, quelle solution est nécessaire pour déterminer quelles données peuvent être dans le cloud, problèmes de connectivité avec le déplacement de grands ensembles de données à partir de systèmes sur site vers le cloud.
- Cloud excellent pour la preuve de concepts de capacités / outils d'analyse.

## Ce qu'il faut savoir du point de vue du DPI

Ressources difficiles à trouver - en particulier des équivalent temps plein (ETP). Il est important d'avoir une stratégie de ressources humaines innovante qui se concentre désormais uniquement sur l'acquisition de ressources, mais également sur le renforcement de l'expertise nécessaire avec les ressources existantes.