What We Heard Report
Challenge Based Solicitation
for
Discoverable and Trusted Data Assets in Support of Decision Making

Purpose of this document:

Under Procurement Process 3.0, Shared Services Canada (SSC) is piloting an improved model of engagement with the private sector. The goal is to collect supplier feedback throughout the procurement process to quickly refine procurement elements if and when necessary. SSC is committed to hear suggestions with an open mind and report back on them to vendors in a transparent way.

This report addresses feedback/comments received via the Slido questionnaire.

Feedback has resulted in a call to modify requirements. Modifications to be made in the final version of the Statement of Challenge (SoC).

Response to Feedback

Scoring method
Comment: While the difference in scoring for 1, 2 and 3 points is clear, the difference between scoring a solution either 4 or 5 seem to be subjective.
SSC Response **Please look at page 36** **Example of scoring:** For the criteria FR1- Storing and publishing capacity: Definitions, business metadata and taxonomy for key business terms and entities. At minimum, the Bidder demonstrates that the Solution contains the following items and ability to save below content to the IBG: - a text box to enter who the Data Steward - a text box or control that shows how the business metadata can be used. - a text box or control that demonstrates how the business metadata is related to technical metadata - a control (tree or drill) that demonstrates taxonomy relationship with other business metadata. The points will be allocated as follow: - All 4 elements demonstrated is equivalent to 3 - meets criterion - More relevant elements = 4 points or 5 points depending of the level of demonstration
SSC Action Clarification provided.

Questions, Responses, and Actions

    1- Problem Statements and Challenges

Part 1: Question 1.3
Comment:

More details on data-related technologies in use would be helpful. We imagine there are a wide variety, but knowing that, for example, "Operational Database Management systems" means Oracle and MS SQL, would be helpful. This is even more important for Data Modelling and ETL tools (some vendors may support very few, or only one of each, of these ...)

SSC Response
Point taken, will modify the package.

SSC Action
Revise SoC.

2- **Technical Requirements**

Part 2: EF1 : Capacité de stockage et de publication

La solution doit stocker et publier les éléments suivants en ce qui concerne les données de SPC :
  o Définitions, métadonnées commerciales et taxonomie pour les principaux termes et entités commerciales.
  o Métadonnées opérationnelles et techniques
  o Possibilité de lier une entrée de catalogue à un terme organisationnel
  o Doit être capable de déterminer si les métadonnées techniques sont un entrepôt de données
  o Publie une procédure pour demander l'accès à un actif de données
  o A la capacité de déterminer si un actif de données est conservé ou éliminé
  o Normes et politiques relatives aux données
  o Prend en charge les normes de métadonnées de l'industrie
  o Renseignements sur les acteurs de la gouvernance des données, y compris les propriétaires et les gestionnaires de données
  o Renseignements sur les données faisant autorité, les données de référence et les données de base

Commented [GC1]: This is in French Please put it in English in the English version

Comment:
The requirements can be misinterpreted due to lack of examples and/or specificity on technical terms. We as vendors do not know how SSC defines terms, i.e. operational, technical and business metadata etc.

Data standards and/or policies are traditionally not published within Integrated Business Glossary (IBG) solutions.

SSC Response:
The industry standards of the types of metadata are established by the Data Management Association, referenced in both the Data Management Book of Knowledge and Dataversity.net.   SSC will define these terms to establish a baseline on what types of metadata we are interested in capturing.

Regarding the comment of what a glossary typically publishes, IBG Solutions are used to automate Data Governance within organizations. Ideally, the IBG is used to capture and socialize the data standards and/or policies to the data governance community.

SSC Action
SSC will provide definitions; including a link to a Dataversity webpage as point of reference for our definitions in the revised Statement of Challenge.

Define types of Metadata

Business Metadata: Provides the meaning of data, by defining terms in everyday language without regard to technical implementation. Focuses largely on the content and condition of the data and includes details related to Data Governance

Technical Metadata: Provides information on the format and structure of the data as needed by computer systems. Examples of technical metadata include physical database tables, access permissions, data models, backup rules, mapping documentation, data lineage

Operational Metadata: This type of metadata "describes details of the processing and accessing of data," according to the DMBOK. Various examples of operational metadata include job execution logs, data sharing rules, error logs, audit results, various version maintenance plans, archive and retention rules, among many others.

Definitions come from https://www.dataversity.net/what-is-metadata/

Part 2: EF3 : Fonctionnalité de collaboration

Comment:
Some focus on why this is beneficial would be helpful. Presumably, you want to guide users to the best data for their purpose based on ratings, feedback, certification, etc…

SSC Response:
IBG Solutions are used to automate Data Governance within organization. The benefit of collaboration within the IBG is used to capture and socialize the data standards and/or policies to the data governance community.

SSC Action
Clarification provided.

Part 2: EF4 : Apprentissage machine - Suggestions?

La solution, qui fait appel à un algorithme d'apprentissage machine, recoupe les actifs d'information disponibles afin d'apporter une meilleure compréhension à l'utilisateur.
Le recoupement des métadonnées peut apporter encore plus de renseignements à l'utilisateur. Les références croisées devraient également réduire le temps que les utilisateurs passent à chercher.
** Les actifs d'information sont des éléments qui touchent, contiennent ou décrivent des données

Comment:
Outcome of mandatory requirement is not clearly defined. Expected functionality is not clear. Demo requirements (on page 39 of the CBS document) does not align with mandatory requirement FR4 (on page 61) and specifies different mandatory functional requirement (Collaboration). Presentation evaluation grid on page 39-44, only covers subset of both FR and NFR mandatory requirements. The Crown might down select a solution that would proceed to written response phase without meeting mandatory requirements i.e. FR4-FR9.

SSC Response
This speaks to the difference between a traditional procurement, where sometimes the crown wants to be reassured prior to contract signing that a given solution will meet all performance requirements. This agile procurement leverages a streamline evaluation process. SSC *are purposely choosing to only evaluate the capacity of the solution to meet a subset of the requirements during presentation.*

SSC verified that there were no misalignment between p. 39 and FR4… The presentation for FR4 requirement is noted on p.40.

For this solicitation, the outcomes are not specifically defined.

SSC Action

This corresponds to NF3, not FR4. Clarification provided

Part 2: EF6 : Processus d'établissement de rapports - Suggestions?

La solution de catalogue doit au moins signaler chaque activité dans un processus de travail défini, la date cible d'achèvement de l'activité, l'état de l'entrée de catalogue associée, le nom et l'identification de l'entrée de catalogue, et un différenciateur visuel évident pour les entrées dont la date cible est imminente et dépassée.

Comment:

There are two FR6 requirements in the CBS document (page 62).

SSC Response

Will correct the numbering.

SSC Action

Correct numbering p.61-62

Part 2: EF7 Fonctionnalité de recherche sémantique

La solution doit comprendre une fonctionnalité de recherche sémantique conviviale et robuste qui n'exige pas la connaissance d'un langage d'interrogation.

- La solution doit permettre aux utilisateurs d'ajouter leurs propres balises de recherche à une entrée de catalogue afin d'enrichir les métadonnées.
- La solution doit permettre une recherche graphique et un filtrage des entrées du catalogue.
- La solution doit fournir une représentation graphique (ou visuelle) des relations naturelles entre les éléments de métadonnées.
- La solution doit permettre l'utilisation de la visualisation pour afficher de multiples types de relations entre les actifs de données.
- La solution doit permettre l'utilisation de la technologie d'apprentissage machine pour automatiser le processus de découverte, d'inventaire, de profilage, de marquage et de création de relations sémantiques entre les ressources de données réparties et cloisonnées

Comment:

As it pertains to requirement: "the solution must enable the use of Machine Learning technology to automate the process of discovering, inventorying, profiling, tagging and creating semantic relationships between distributed and siloed data assets".

The expected behavior of machine learning technology is not clearly defined.

SSC Response

Machine Learning technology enhances semantic search functionality. Machine Language technology would attempt to discover natural relationships between information assets. Creating or providing tagging elements to information assets based upon a user's search behavior. As a user interacts with

the system, it will learn on what are the information assets that they are interested in and provide suggestions.

SSC Action

Clarification provided.

Part 2: EF8 Traçabilité des données

La solution permet de capturer et de publier la traçabilité des données à travers les multiples étapes du cycle de vie des données : acquises, partagées, archivées et purgées.

Prise en charge des outils ETC de premier niveau et fourniture de connecteurs natifs et d'API basées sur REST pour rechercher et extraire le contenu de la traçabilité des données.

Comment:

"Tier one ETL tools" are not defined in the document.

SSC Response

Currently SSC only has SSIS and Talend as ETL tools. These tools are not Tier one. Tier One tools are Informatica, IBM Data Stage & MS SQL

SSC Action

SSC to update the solicitation to list the elements in the evaluation criteria to be demonstrated for evaluation purpose. (Attachment 6 – Evaluation criteria). It will also include the list of elements the solution would interact with during the prototype phase.

Part 2: FR9 : Interopérabilité - Suggestions?

La solution assure l'interface et interagit avec l'outil de gestion de la qualité des données.

Comment:

More specificity would be helpful. You could argue that MS Excel is a data quality tool. Suggest "an industry recognized data quality solution" or "a data quality management tool recognized as a leader by industry analysts", or something similar.

Does SSC own one that you would like to work with?

SSC Response

By DQ solution, SSC means commercial products/COTS. While SSC does not have a DQ solution at this time, SSC is considering to acquire one in the near future and want to confirm possibility of connection with eventual tool.

SSC Action

Requirement revised. SSC will post the revised solicitation package on Buy and Sell.

Part 2: ENF1 : Exigence de rendement

Le rendement de la solution doit être conforme à ce qui suit :

Elle doit prendre en charge au minimum 6 000 utilisateurs; 500 utilisateurs simultanés ayant des droits de lecture de base et 100 utilisateurs simultanés ayant un rôle de gestionnaire des données (rôle d'auteur/contributeur); Elle doit prendre en charge 1 000 sources de données contenant jusqu'à 100 000 métadonnées.

Comment:

Please characterize the activities of the 6,000 users - how many of the actual users (as opposed to concurrent) would be authors/contributors?

SSC Response
As per the requirement, minimum 500 concurrent users who have basic read rights and 100 concurrent users who have data stewards roles (author/contributor role).
SSC Action
Clarification Provided

Part 2: ENF2 : Déploiement sur place
La solution proposée doit pouvoir être déployée et exploitée au sein d'un centre de données géré par SPC.
Comment:
This requirement contradicts Treasury Board Secretariat's "Cloud First" policy. Furthermore, all new generation tools have at least certain components that are operated and managed in the cloud, private or public, and in either PaaS, SaaS or IaaS. In order for SSC to leverage best-of breed technology and also align the solution with TBS cloud first policy we strongly recommend to reword this requirement to include hybrid deployments - on-premise metadata collection with cloud enabled interface and metadata repository. TBS "Government of Canada Cloud Adoption Strategy: 2018 update": Cloud services are identified and evaluated as the principal delivery option when initiating IT investments, initiatives, strategies and projects. Using the Government of Canada Right Cloud Selection Guidance to guide their decision, departmental and agency CIOs will select one of the following deployment models, in the following order of priority: a) public cloud b) hybrid cloud c) private cloud d) non-cloud Departmental and agency CIOs will select one of the following cloud service models, in the following order of priority: a) software as a service (SaaS) b) platform as a service (PaaS) c) infrastructure as a service (IaaS)
SSC Response
We are considering feedback and discussing how best to address.
SSC Action
TBD

Part 2: ENF3 Déploiement infonuagique
La solution doit être capable de se connecter et de fonctionner dans une infrastructure basée sur le nuage.
Comment:
This mandatory requirement is in contradiction with mandatory requirement NF2.
SSC Response
We are considering feedback and discussing how best to address.
SSC Action
TBD

Part 2: ENF4 : Plateforme comme service sur le nuage - Suggestions?
La solution doit fournir une plateforme comme service (PaaS); le fournisseur est alors chargé de fournir : • Une plateforme infonuagique afin d'exploiter tous les composants de la solution logicielle

• Tous les services professionnels nécessaires pour gérer la plateforme comme composant de la solution, au moins l'application de la sécurité, les correctifs, la sauvegarde et la récupération, la configuration de la plateforme • La fourniture de la capacité de stockage de la plateforme nécessaire pour gérer toutes les fonctions requises de la solution identifiées dans l'énoncé des défis, 100 Go de capacité pour l'échelonnement des entrées et sorties des fichiers de métadonnées de la solution du GOI
Comment: While these categories are flexible, what is being described is a Managed Service delivered on a Cloud Service Provider's IaaS, rather than an actual PaaS. We recommend to rewrite requirement bullet nr. 2 as "all necessary platform services to manage the platform as a service at the very least application of security, patches, backup & recovery, storage and platform configuration". We recommend removing bullet nr. 3 as storage is implied in the PaaS and is covered with bullet nr. 2
SSC Response We are considering feedback and discussing how best to address.
SSC Action TBD

Part 2: ENF8 : Sécurité La solution proposée doit assurer la transmission sécurisée et confidentielle de l'information avec une protection adéquate contre les attaques malveillantes et l'exposition accidentelle (p. ex. contamination croisée) en utilisant l'infrastructure de sécurité au sein de SPC.
Comment: More description, or a pointer documentation for "the security infrastructure within SSC" would be helpful, given that this is a mandatory requirement.
SSC Response SSC is reviewing this requirement.
SSC Action TBD

Part 2: ENF10 : Outil d'ingestion de métadonnées - Suggestions? La solution doit permettre d'acquérir des métadonnées à partir des actifs d'information, qui ne font pas partie de la structure typique des dépôts. La classification des actifs d'information va comme suit : • les sources de données qui sont classées comme des systèmes de gestion de bases de données opérationnelles (SGBDO) (infonuagique, sur place, hybride) • les systèmes de modèles de données • les systèmes de renseignement opérationnel • les systèmes d'extraction, de transformation et de chargement • les sources de données de fichiers non hiérarchiques, c'est-à-dire au moins les fichiers .text, .CSV, .xls
Comment: More specifics would be helpful. There may be vendors that only support one (i.e. their) solution in

these spaces. It is not clear what "metadata that falls outside of the typical repositories structure" means. We suggest removing this requirement as it seems duplicate unless further clarification is provided.

SSC Response

SSC will update the requirement.

SSC Action

Content to be updated.

Part 2: ENF14 : Environnements pris en charge et compatibilité des logiciels

Fournir une documentation indiquant quels sont les environnements pris en charge pour le logiciel. De même, fournir toute compatibilité logicielle qui démontre que les composants logiciels ou les systèmes peuvent fonctionner de manière satisfaisante ensemble sur le même ordinateur ou sur des ordinateurs différents qui sont reliés par un réseau informatique.

Comment:

Requirement is not clear. Consider providing a use case describing what is required functionality.

SSC Response

Use Case for Supported Environments.

Some software versions may not be suitable or supported in certain environments. For instance, one can no longer connect SAS EG 8.1 to MS SQL Server 2012. As MS SQL Server 2012 is not a supported SQL Server instance from Microsoft.

SSC Action

Clarification provided.

3- Evaluation Criteria

Part 3: EF2 Capacité de connexion - Suggestions?

Le soumissionnaire devrait démontrer comment la solution proposée peut se connecter à des sources de données qui sont classées comme des systèmes de gestion de bases de données opérationnelles (SGBDO) et des technologies infonuagiques.

Comment:

While there is a definition for required OPDMBS connectivity and metadata collection (on page 38 of CBS document), it is not clearly defined what are the tools used to manage data for Business Intelligence, Data Modeling, Extract Transform & Load, and Enterprise Service Business.

If these are mandatory ("at a minimum), wouldn't the score be 5 or 0?

SSC Response

SSC is considering the feedback received and determining how to address.

SSC Action

TBD

Part 3: EF2 Capacité de connexion

Le soumissionnaire devrait démontrer comment la solution proposée peut se connecter à des sources de données qui sont classées comme des systèmes de gestion de bases de données opérationnelles (SGBDO) et des technologies infonuagiques.

Au minimum, le soumissionnaire démontre que la solution est capable de se connecter aux types de sources de données suivants :

Plateformes infonuagiques (au minimum, elles doivent prendre en charge)
1. AWS
2. Azure

SGBDO (au minimum, devrait prendre en charge)
1. Oracle
2. SQL Server de Microsoft
3. DB2
4. Netezza
5. SAP (Hana)
6. MySQL
7. MariaDB
8. MongoDB

Des points supplémentaires seront attribués pour une solution qui prend en charge la connexion hybride, c'est-à-dire qui pourrait se connecter aux deux types de sources de données sans connecteur tiers supplémentaire.

Comment: If these are mandatory ("at a minimum), wouldn't the score be 5 or 0? Beyond the "minimum" why would there be any points for support of a solution you don't use?

SSC Response
As this requirement may be revised, response to this comment is pending.

SSC Action
Clarification to come.

Part 3: EF4 Collaboration

Le soumissionnaire devrait démontrer comment la solution proposée permet une collaboration entre plusieurs utilisateurs.

Au minimum, la solution devrait permettre aux utilisateurs de :
- partager
- collaborer
- avoir un processus de type messagerie associé à la fonctionnalité de collaboration

Comment: Please define "share" and "collaborate"

SSC Response
IBG Solutions are used to automate Data Governance within organization. The benefit of collaboration within the IBG is used to capture and socialize the data standards and/or policies to the data governance community. This being the case, we're referring to collaboration on definitions, business rules, and data standards. Sharing refers to the sharing of information drawn from the content of the IBG. Example: Sharing of definitions of business terminology; sharing of the names of data stewards.

SSC Action
Clarification provided.

Part 3: ENF4 Capacité du soumissionnaire à faire évoluer la solution

Le Canada aimerait sélectionner un soumissionnaire qui a la capacité de maintenir sa solution à la pointe de la technologie.

Le soumissionnaire a la possibilité de déterminer les éléments de la feuille de route de l'innovation. Pour guider sa préparation, les éléments sont des exemples de ce qui pourrait être inclus dans une feuille de route pour l'innovation.

Le soumissionnaire devrait présenter sa feuille de route pour l'innovation, sous la forme d'une représentation visuelle autrement, contenant les éléments suivants :

- Calendrier : Indiquer le délai nécessaire pour concrétiser l'innovation
- Étapes du processus : Définissez vos étapes
- Flux ICR : Montrez ce que vous mesurerez dans le cadre du projet.
- Il s'agit de montrer les sujets généraux que le projet d'innovation explorera.
- Flux de menaces : Il s'agit des éléments qui doivent être gérés et atténués
- Ressources : Quelles ressources mobiliser pour concrétiser l'innovation?

Comment: It's not clear how this requirement will be scored during presentation. Furthermore, it is not an usual practice for vendors to share their product roadmaps.

SSC Response
This requirement enables us to choose a solution and a vendor that remains on top of technology trend (i.e. has not gone stale). It may not be usual practice. One could say this level of vendor consultation is not usual practice. These roadmaps would be shared with the Crown only.

SSC Action
Clarification provided.

Part 3: ENF5 Capacité du soumissionnaire à développer la solution

Le soumissionnaire devrait démontrer l'extensibilité de la solution qu'il propose.

Des points seront accordés pour une solution qui, avec un minimum d'efforts, peut :
- augmenter la capacité de débit de transactions;
- intégrer de nouveaux utilisateurs;
- être déployé dans l'environnement d'un autre service.

Pour cette exigence :
« effort minimal » signifie ne nécessitant aucun changement à l'architecture ni aux intégrations telles que déployées actuellement.

Comment: The Bidder should demonstrate how the proposed Solution is scalable.

SSC Response

With a SaaS service, the vendor typically takes on the responsibility to ensure that the solution has the necessary capacity to handle SSC needs.

This being the case, this requirement is not needed.

SSC Action

This requirement will be removed from the solicitation

4- Price Proposal

Section 4 – Financial Bid form

How clear is your understanding of how the price is to be submitted?

Comment: In the chapter "2.1 Licensed Software and maintenance support" it is specified that
"License is for perpetual for unlimited number of users". In our opinion this licensing approach is not aligned with modern SaaS/PaaS cloud licensing models. Allowing for a different licensing approach, one with usage and consumption in mind, would allow the Crown for easier solution scalability, addition of new data sources or users.

Furthermore, usage and/or consumption licensing will have less CapEx financial impact on the earlier phases of the project and in the future. Additionally, we believe there are missing steps to preclude successful deployment which include end-user roll-out, data governance process setup, piloting the solution with the business community etc. More clarification will need to be provided. Many aspects go into pricing including what type of users or connections you have you. More information will help us be more accurate in terms of final pricing.

The $25 k limitation on the POC component is a potential challenge. Will need to understand in more detail the deliverables in detail.

SSC Response

The current pricing model is based on solution that would be on-prem. SSC will include a different pricing model for a Could solution. This model will be discussed during the upcoming wave of Invitation to Refine.

The details of the Prototype work segment are provided on Section 8 of the SoC. To clarify:

Scope of prototype:
- Three months/ 60 days business days in duration
- Solution connected to
  - SSC's Enterprise Data Repository (max. 10 data sources)
  - SSC's Cloud Data Zone (max 10 datasets)
- Support limited to:
  - Training max 10 people (trainers and admin/support)
  - Technical support

Additional elements critical to a successful deployment (end user roll out, data governance, etc.) are led by SSC. The scope of the contract does not included professional services for these activities.

SSC Action

The pricing model will be further discussed with vendors