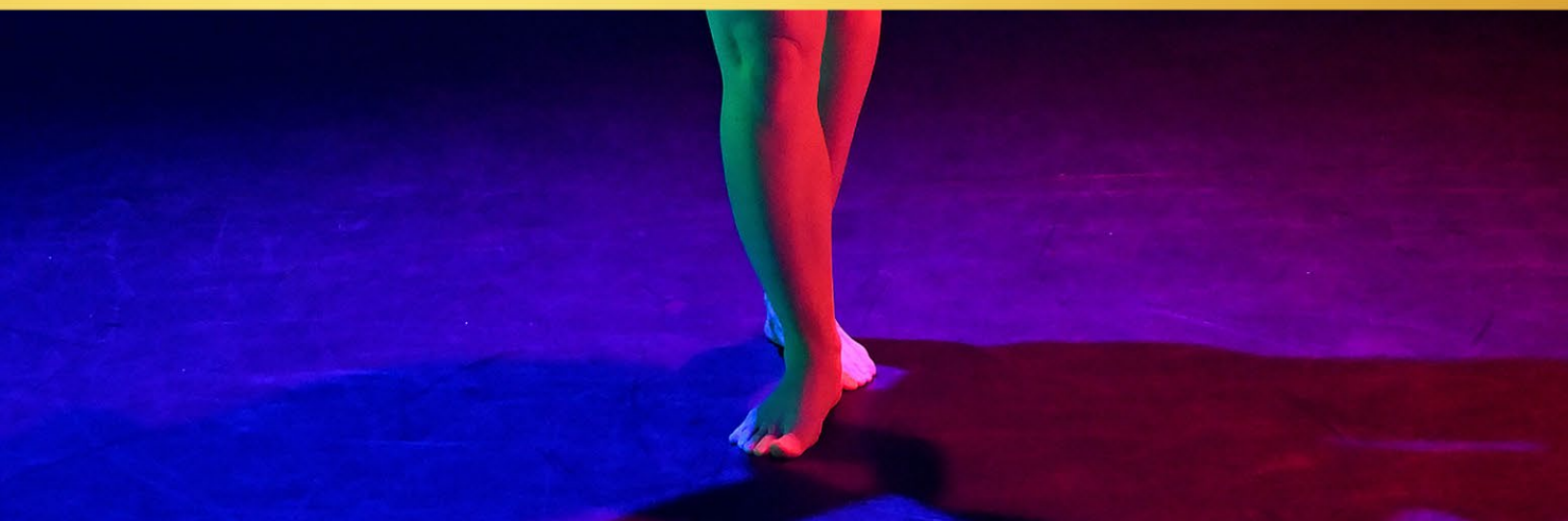




Renforcer et lier les systèmes
d'information des arts de la scène

Rapport de projet





CANADIAN ASSOCIATION
FOR THE PERFORMING ARTS
ASSOCIATION CANADIENNE
DES ORGANISMES ARTISTIQUES

Ce projet a été réalisé par CAPACOA avec l'appui de Synapse C, Anicet Lepetit Ondo et Gregory Saumier-Finch, sous la direction de Frédéric Julien (CAPACOA).

Nous remercions les gestionnaires des plateformes partenaires du projet : RADARTS, l'Association RIDEAU, Organization of Saskatchewan Arts Councils et Ontario Presents. Nous remercions également Véronique Godin et Molly Richardson (RADARTS), Jason Lavigne et Simon Landry (Blaze Studios) pour la mise en œuvre des recommandations du projet dans la plateforme Scènes francophones.

Avertissement : Ce projet a été financé par le Fonds de relance des services communautaires du gouvernement du Canada. Les prises de position, opinions et recommandations de ce rapport sont celles des auteurs du rapport et ne reflètent pas nécessairement celles du gouvernement du Canada.

Canada

Dépôt légal
Bibliothèque et Archives Canada
ISBN : 978-1-9991984-9-7
Publié en juillet 2024.



Cet ouvrage est publié sous la licence [Creative Commons 4.0 International](https://creativecommons.org/licenses/by/4.0/).

Citation de ce rapport :

Ondo, A. L., Saumier-Finch, G., Vézina-Topping, L., (2024). *Renforcer et lier les systèmes d'information des arts de la scène : Rapport de projet*. CAPACOA.

Photo en page couverture : Le spectacle *Glitch* apparaît dans le répertoire de la plateforme [Scène Pro](#) et sa compagnie de production, Bouge de là, s'est vue attribuer l'identifiant [K10-235](#) dans Artsdata. Ce spectacle a tourné dans plusieurs villes au Québec ainsi qu'en Colombie-Britannique et en Ontario. Crédit photo : Suzane O'Neill.

Table des matières

Définitions	4
Contexte	6
CAPACOA	6
Artsdata	6
Le projet : Renforcer et lier les systèmes d'information des arts de la scène.....	6
Enjeux associés aux projets technologiques	7
Enjeux de gestion.....	7
Enjeux organisationnels	7
Enjeux technologiques	7
Partenaires de projet	8
Scène Pro	8
Scènes francophones.....	8
I Want to Showcase	9
Arts Touring Connector	9
Résumé du déroulement du projet	10
Méthodologie	10
Défis rencontrés.....	12
Faits saillants	13
Le processus d'ouverture de données comporte de multiples étapes.....	13
Les données ouvertes liées dans le contexte du graphe de connaissances Artsdata.....	14
L'identifiant pérenne : la clé pour lier les données	16
Méthodes de publication et d'extraction des données ouvertes	17
Observations sur les systèmes d'information analysés	19
Observations sur les ontologies des plateformes analysées.....	21
Observations sur l'équité, la diversité et l'inclusion.....	25
Recommandations	26
Conclusion.....	28

Définitions

Entité : Groupe de choses identifiables qui partagent les mêmes attributs. Il peut s'agir de choses tangibles comme des objets (p.ex : des livres), des personnes (p.ex : des artistes), des lieux (p.ex: des salles de spectacle) ou de choses intangibles (p.ex : des spectacles, des genres musicaux).

Note : Dans les ontologies RDF, le terme « entité » désigne autant une classe, qu'un attribut ou une instance.

Synonymes :

- classe (terme employé dans les ontologies RDF)
- type (terme employé dans l'ontologie Schema.org)

Instance : Occurrence d'une entité dans un jeu de données. Par exemple, la Salle Louis-Fréchette, au Grand Théâtre de Québec, est une occurrence de l'entité « salle de spectacle » dans Scène Pro. Son identifiant Scène Pro est [272](#).

Synonymes :

- individu
- objet
- entité nommée

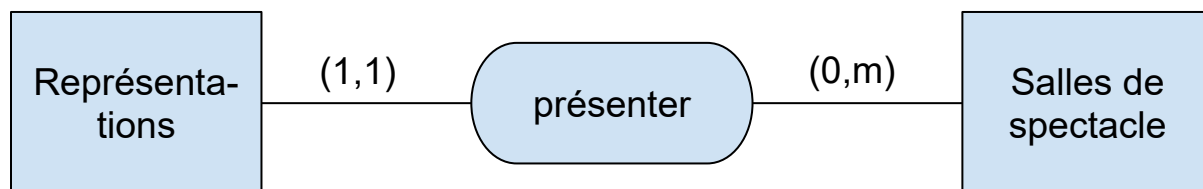
Attribut : Caractéristique ou paramètre d'une entité et, par extension, des individus qui en sont des instances. Par exemple, la capacité est un attribut de l'entité salle de spectacle et sa valeur attendue est un nombre. L'attribut peut aussi décrire une relation entre deux instances. Par exemple, l'instance Salle Louis-Fréchette est associée à l'instance Grand Théâtre de Québec par l'attribut containedInPlace dans Artsdata. L'association d'un attribut à une valeur particulière forme une « pair attribut-valeur ».

Synonymes :

- propriété (terme employé dans les ontologies RDF)
- métadonnée descriptive

Cardinalité : Nombre d'occurrence d'une entité pouvant être liée à une autre entité. Par exemple, une représentation ne peut être liée qu'à une seule salle de spectacle, cette cardinalité s'écrit (1,1). À l'inverse, une salle de spectacle peut être liée à aucune ou plusieurs représentations, cette cardinalité s'écrit (0, m).

Le schéma de cet exemple pourrait être :



Une telle représentation constitue un **modèle conceptuel de données**. Les éléments dans les rectangles sont des entités, l'élément dans le rectangle arrondi est l'association qui lie les deux entités, et les valeurs entre parenthèses sont les cardinalités.

Ontologie : Représentation formelle d'un ensemble de concepts et des relations entre eux au sein d'un domaine spécifique. Elle sert à structurer et à organiser les connaissances pour en faciliter l'interprétation et le partage.

Alignement d'ontologies : Établissement des correspondances entre les concepts de différentes ontologies, afin de permettre leur interopérabilité et l'intégration de leurs informations.

Synonymes :

- mise en correspondance
- *mapping*



Image 1 : Cette photo est l'image du profil d'artiste de Raphaël Butler sur la plateforme Scènes francophones. Cette image a été réutilisée pour chacune des 14 représentations du spectacle de Raphaël, Tattoo de cœur, présentées par les membres de RADARTS en février et mars 2024. Crédit photo : courtoisie de l'artiste.

Contexte

CAPACOA

La Canadian Association for the Performing Arts/l'Association canadienne des organismes artistiques (CAPACOA) est un organisme national de services aux arts qui soutient la tournée et la diffusion des arts de la scène. Depuis plus d'une décennie, CAPACOA déploie des efforts en soutien à la transformation numérique des arts de la scène, notamment par le biais d'initiatives telles que *Un avenir numérique lié* et Artsdata.

Artsdata

Artsdata est un graphe de connaissances pancanadien pour les arts. Depuis 2018, Artsdata facilite le partage et la réutilisation de données ouvertes entre des sites web, des bases de données et des plateformes de l'industrie du spectacle.

Le projet : Renforcer et lier les systèmes d'information des arts de la scène

Présentation du projet

En lien avec l'initiative Artsdata, CAPACOA et ses partenaires de projet ont analysé la structure de données de systèmes d'information clés du secteur des arts de la scène, afin de permettre une circulation efficace de l'information entre ces systèmes. Le projet en est un de transformation numérique.

Objectifs du projet

L'objectif principal est de cartographier les systèmes d'information clés des arts de la scène pour documenter les disparités dans les ontologies et recommander des améliorations orientées sur l'ouverture des données, l'alignement des ontologies et la réutilisation des données. Lorsque des protocoles ouverts sont adoptés et que des ontologies sont interopérables (c'est-à-dire, des systèmes d'information peuvent se comprendre), cela rend possible la réutilisation des données des arts de la scène à travers le Canada.

Enjeux associés aux projets technologiques

Cette section discute des enjeux généraux qui gravitent autour des projets technologiques et des initiatives de transformation numérique, soit la catégorie de projet dans laquelle s'inscrit le projet discuté dans ce rapport. Il est question d'enjeux sociotechniques, soit des enjeux de gestion, des enjeux organisationnels et des enjeux technologiques. Ces enjeux se manifestent à différentes intensités d'une organisation à l'autre, influençant, dans le cadre de ce projet, le fonctionnement des systèmes d'information analysés.

Enjeux de gestion

Les enjeux de gestion réfèrent aux éléments de planification, d'organisation, de direction et de contrôle inclus dans un processus de gestion de projet. Lors de l'élaboration d'un projet technologique, on se questionne sur le contexte, l'environnement externe, dans lequel l'organisation propriétaire du projet évolue. Est-ce que le projet est cohérent avec le modèle d'affaires dans lequel on tente de le réaliser ? Quels efforts seront requis ? Quels sont les risques ? Est-ce que le projet s'harmonise avec le style de leadership et de gestion présents dans l'organisation ? Est-ce que le projet s'inscrit dans une stratégie claire ? Comment va s'intégrer le projet dans la gouvernance des technologies de l'information ? À quel(s) besoin(s) organisationnel(s) et/ou technologique(s) le projet répond-il ? Quel(s) objectif(s) participe-t-il à atteindre ?

Dans le cadre du projet, ces enjeux se sont manifestés au niveau des stratégies de gestion des partenaires, qui ont à la fois influencé le projet et évolué avec celui-ci.

Enjeux organisationnels

Les enjeux organisationnels mènent à une réflexion sur l'environnement interne de l'organisation propriétaire du projet technologique. Cette réflexion porte sur l'état des connaissances et des compétences du personnel. Sont-elles en adéquation avec les connaissances et les compétences requises pour réaliser le projet ? Est-ce que des habitudes de formation continue sont présentes dans l'organisation ? La réflexion porte aussi sur les effets du projet sur l'organisation et les méthodes de travail du personnel, sur les processus d'affaires et sur les rôles et responsabilités. Comment le personnel va-t-il réagir à ces effets ? Comment sont gérées l'information et la connaissance au sein de l'organisation ? Est-ce qu'un plan de gestion du changement est établi ?

Enjeux technologiques

Les enjeux technologiques concernent les infrastructures matérielles et logicielles, les habitudes d'entretien et de mise à jour, ainsi que les dispositifs administratifs (dont les politiques) encadrant la qualité de l'information et la conservation des données. Il faut se questionner sur l'état des infrastructures, dont leur récence, leur intégration, leur compatibilité, leur extensibilité, leur interopérabilité, la fréquence d'entretien et l'assiduité des mises à jour. Enfin, il faut se questionner sur la nature des données et de l'information : est-ce qu'il y a un potentiel d'atteinte à la vie privée ?

Partenaires de projet

Étant donné l'envergure restreinte du projet, en temps et en ressources, CAPACOA a ciblé quatre plateformes numériques qui centralisent et facilitent des processus d'affaires dans l'industrie du spectacle, au Canada. Ces plateformes avaient en commun qu'elles étaient gérées par des associations du secteur du spectacle avec lesquelles CAPACOA entretenait déjà des relations, ce qui se voulait un facteur de succès pour le projet.

Cette sélection ne constitue cependant pas un inventaire exhaustif des plateformes au Canada, loin de là. Il existe plusieurs autres plateformes à l'intérieur et à l'extérieur du domaine des arts de la scène au Canada. CAPACOA avait répertorié en 2019 plus de 40 différents systèmes d'information se rapportant à 11 cas d'usage¹. De ceux-ci, un bon nombre ne sont plus en opération, ce qui soulevait aussi des préoccupations quant à la pérennité des plateformes retenues.

Scène Pro

<https://scenepro.ca/>

Scène Pro est une plateforme visant à faciliter la coordination du secteur des arts de la scène du Québec autour de la donnée d'offres de spectacles. Scène Pro capitalise sur les inscriptions d'offres de spectacles que font déjà en grand nombre les producteurs (+2000 inscriptions annuelles de RIDEAU, du ROSEQ et de Réseau Scènes) pour générer une base de données unifiée. Scène Pro souhaite agir comme source de données structurées pour l'ensemble du secteur des arts de la scène². Scène Pro est dotée d'une API permettant à des applications tierces d'accéder à ses données.

Référentiel des métadonnées descriptives du spectacle

Le Référentiel du Québec est un modèle de données qui se veut un « outil facilitant le développement de systèmes d'information associés aux spectacles, et leur interopérabilité ».³ La structure de données du Référentiel est alignée sur celle de Scène Pro. Le modèle du Référentiel étant bien documenté, c'est celui qui a servi aux fins de l'équipe de projet.

Scènes francophones

<https://www.scenesfrancophones.ca/>

Chapeautée par l'Alliance Scènes francophones⁴, la plateforme Scènes francophones se veut un « guichet unique numérique qui facilite les rencontres et les échanges amplifiant la diffusion des

¹ Estermann, B. et Julien, F. (2019). *Lier l'avenir numérique des arts de la scène : Comment mobiliser les synergies de la chaîne de valeur* (p. 59, 62). CAPACOA. <https://bac-lac.on.worldcat.org/oclc/1138886653>

² Scène Pro. (2024). À propos. <https://scenepro.ca/a-propos>

³ Référentiel des métadonnées descriptives du spectacle. (2024). Bienvenue. <https://a10s-ca.github.io/referentiel/>

⁴ L'Alliance Scènes francophones est composée de trois réseaux soutenant la diffusion des arts de la scène dans la francophonie canadienne : RADARTS, Réseau Ontario et le Réseau des grands espaces.

arts de la scène francophone au Canada⁵ ». Scènes francophones vise à répondre à divers enjeux et besoins, dont la redondance des formulaires à remplir dans les divers systèmes et la découvrabilité des arts de la scène⁶. La plateforme cherche à servir tous les intervenant.es de l'industrie des arts de la scène : réseaux, diffuseurs, artistes, agents/représentants d'artistes et le grand public.

I Want to Showcase

<https://iwanttoshowcase.ca/>

La plateforme I Want to Showcase offre une façon simple pour des artistes de poser leur candidature à plusieurs vitrines⁷ sur les arts de la scène à travers l'Amérique du Nord, le tout à un seul et même endroit en ligne. Ce système sert donc les organisateurs d'événements de type vitrine, ainsi que les artistes et leurs représentants. I Want to Showcase est une collaboration entre 10 conférences partenaires au Canada et aux États-Unis⁸.

Arts Touring Connector

<https://arts-tc.ca/>

Arts touring connector (ATC) est une plateforme qui vise à faciliter la planification des tournées de spectacles. Ce processus appelé « programmation en groupe » est mené par les diffuseurs et les réseaux qui les soutiennent. Il a pour but de rendre les tournées plus efficaces et d'accroître leur impact pour les artistes, les diffuseurs et les communautés. Les artistes et les spectacles sont ajoutés à la plateforme quand un diffuseur en fait la demande. De plus, une proposition de tournée est affichée lorsqu'au minimum trois diffuseurs ont exprimé leur intérêt pour un spectacle. À l'heure actuelle, quatre réseaux de diffuseurs utilisent ATC et deux autres sont en cours d'intégration. Chaque réseau approche la coordination de tournée différemment, ce qui implique que la manière dont ils se servent de la plateforme peut varier, entre autres au niveau de la quantité d'informations divulguées auprès des agents, des gérants et des artistes.

⁵ Scènes francophones (2024). À propos : [scenesfrancophones.ca](https://www.scenesfrancophones.ca/).

<https://www.scenesfrancophones.ca/%C3%A0-propos/scenesfrancophonesca>

⁶ Scènes francophones. (2024). À propos.

<https://www.scenesfrancophones.ca/%C3%A0-propos/scenesfrancophones>

⁷ Une vitrine est une occasion pour un artiste de présenter un extrait de spectacle (ou une intégrale) à un auditoire composé d'organismes diffuseurs, lesquels peuvent ensuite acheter les droits pour présenter une ou plusieurs représentations du spectacle.

⁸ I Want To Showcase. (2024). Our Partner Conferences.

<https://iwanttoshowcase.ca/partner-conferences.php>

Résumé du déroulement du projet

Cette section présente la méthodologie mise en place pour réaliser le projet, suivie d'un résumé des défis rencontrés à la réalisation.

Méthodologie

1. Modélisation conceptuelle des systèmes d'information

La première grande étape de l'analyse est une modélisation conceptuelle de la structure de données des systèmes d'information analysés. Pour réaliser cette modélisation, l'équipe a effectué au minimum deux entretiens avec les équipes opérant les systèmes d'information analysés.

Un premier entretien visait à comprendre l'origine des plateformes, le contexte qui a mené à l'idéation et le contexte dans lequel les plateformes évoluent maintenant. Les analystes tentaient d'obtenir des réponses à des questions telles que

- D'où est venue l'idée ?
- Quel(s) besoin(s) sont derrière la conception des plateformes ?
- Quelles étaient les intentions initiales au moment d'entamer la conception ?
- Où en est le développement de la plateforme par rapport à ces intentions initiales ?

Ce premier entretien était suivi d'un second avec les personnes responsables de la conception technique des plateformes, dans le but de comprendre les prémices technologiques et techniques qui ont orienté le développement des plateformes.

De là, les analystes se sont attelés à la modélisation des systèmes d'information : un dictionnaire entité-attribut a été créé pour chacune, puis, lorsque possible, un graphe de dépendance et un modèle conceptuel de données ont aussi été produits. L'objectif était de faire ressortir les types d'entités (aussi appelés des « classes ») utilisées par le système d'information analysé, leurs attributs, les valeurs possibles et les relations (incluant les cardinalités) entre ces entités. Les documents générés par cette modélisation ont ensuite servi de base pour l'analyse des alignements ontologiques.

2. Analyse des alignements ontologiques

Aligner les ontologies vise à minimiser les problèmes d'hétérogénéité qui surviennent lorsque différentes sources de données tentent d'échanger leurs données. Dans ce travail, plusieurs niveaux d'hétérogénéité ont été abordés, à savoir : le niveau syntaxique, terminologique, conceptuel et sémiotique.

Au niveau syntaxico-terminologique, l'hétérogénéité se reflète dans le fait d'attribuer des noms aux entités (classes, propriétés, relations, etc.) qui constituent une ontologie. Dans cette catégorie, le travail vise à :

- Gérer la synonymie lorsque différentes désignations sont employées pour faire référence à une même entité.
- Gérer la polysémie (un même nom pouvant désigner plusieurs entités distinctes).

- Gérer des concepts provenant de différentes langues (Français, Anglais) utilisés pour désigner une même entité.

Le niveau conceptuel : Les divergences conceptuelles que nous pouvons observer dans les sources de données sont généralement basées sur :

- La granularité : Certaines sources de données décrivent les mêmes concepts avec des niveaux de détail différents.
- La couverture : La différence entre deux ontologies peut résider dans l'étendue de la couverture du domaine décrit. Elles peuvent couvrir des parties différentes ou des parties qui se chevauchent.
- La perspective : Même si deux ontologies représentent exactement le même domaine, elles peuvent l'aborder de points de vue différents.

Le niveau sémiotique : ce type d'hétérogénéité survient lorsqu'il y a une différence d'interprétation de la même ontologie par diverses personnes ou communautés.

Pour atténuer ces disparités entre les sources de données, nous analysons d'abord les sources existantes, puis établissons des correspondances (basées sur des relations d'équivalence, de subsomption, etc.) entre des entités telles que les classes et les propriétés, afin de réconcilier les points de vue sémantiques et les opinions de plusieurs experts.

Après avoir analysé la structure, la nomenclature et le niveau de granularité des données, deux types de correspondances peuvent être identifiés. D'une part, les correspondances simples. Ce type de correspondance est de type (1:1), où chaque entité d'une ontologie est associée à une seule entité de l'autre ontologie. D'autre part, les correspondances complexes. Elles peuvent être soit de type (1:m), où une entité d'une ontologie est associée à plusieurs entités de l'autre ontologie, soit de type (n:m), où plusieurs entités d'une ontologie sont associées à plusieurs entités de l'autre ontologie.

Pour ce travail, nous nous sommes appuyés sur des méthodes terminologiques (sémantiques et syntaxiques) d'alignement, ainsi que sur celles basées sur la structure interne et externe des ontologies, englobant les entités, leurs relations et les propriétés qu'elles partagent. Nous avons commencé par implémenter des ontologies indépendantes issues de ces différentes sources de données. Ensuite, nous avons entrepris une phase de fusion de ces sources, chacune conservant son contexte (IRI). Enfin, nous avons aligné les entités existantes.

3. Mise en application

Au terme de la modélisation et de l'analyse de l'alignement de chaque plateforme, nous avons remis à l'organisme gestionnaire un rapport privé avec des observations et des recommandations. Puis nous avons offert du soutien aux gestionnaires de plateforme dans des processus d'ouverture de données.

Dans le cas de Scènes francophones, nous avons accompagné les gestionnaires dans la mise en œuvre de la plupart des recommandations, tel que détaillé plus bas. Nous avons aussi procédé à l'extraction et au chargement de toutes leurs données ouvertes dans Artsdata. Ces données ont

par la suite été réutilisées par le gouvernement du Nouveau-Brunswick dans sa nouvelle plateforme Arts Culture NB (en développement).

Dans le cas de Scène Pro, nous avons procédé à l'extraction et au chargement de certaines de leurs données - diffuseurs et salles de spectacles - dans Artsdata.

Dans le cas de I Want To Showcase, les partenaires de la plateforme ont accepté les recommandations et ont entamé un processus d'ouverture de données. Ce processus était en démarrage au moment de la conclusion du projet.

Défis rencontrés

Envergure du projet

Le défi de réaliser l'analyse de quatre plateformes sur une année, l'une à la suite de l'autre, était un programme ambitieux dès le départ. La réalisation du travail auprès de la première plateforme a pris plus de temps que prévu, mais a permis de dégager des méthodes qui ont pu être répétées par la suite. L'analyse de la première plateforme a en quelque sorte servi de projet-pilote. Une part du travail d'analyse a dû être réalisé sur plusieurs plateformes en même temps afin de s'adapter aux échéanciers des partenaires : il n'était pas toujours simple de trouver des chronologies communes.

Mobilisation des partenaires

Certaines étapes qui n'étaient pas prévues ont été ajoutées au niveau de la communication avec les partenaires et leurs sous-traitants, occasionnant des retards dans l'obtention des données. Les modes de gouvernance variés des différentes plateformes et la sous-traitance à des tiers des aspects techniques de la gestion de ces plateformes ont constitué des défis à l'obtention des données. D'emblée, la gouvernance des données (et des plateformes) n'étaient pas inclus dans le périmètre du projet, à l'exception des éléments de gouvernance requis pour rendre les données ouvertes.

Difficultés techniques au niveau de l'obtention des données

Outre les enjeux liés à la mobilisation des partenaires, des difficultés techniques ont été rencontrées au niveau de l'obtention des données. Par exemple, certaines plateformes n'utilisent pas de système de gestion de base de données relationnelle (RDBMS), système qui facilite grandement la modélisation de la structure des données. Néanmoins, les plateformes organisées sur la base d'un système de gestion de contenu (CMS) sans base de données relationnelle ont été modélisées en fonction de l'information disponible. Également, une des analyses de plateforme n'a pas pu être réalisée en raison d'enjeux de propriété intellectuelle.

Faits saillants

Le processus d'ouverture de données comporte de multiples étapes

Le processus d'ouverture de données comporte de multiples étapes

L'intérêt des données ouvertes dans le contexte du secteur du spectacle se trouve dans les nombreux avantages individuels et collectifs qui peuvent découler de leur réutilisation⁹, notamment sous forme de gains d'efficacité. Dans le cadre de ce projet, l'ouverture des données des plateformes présente aussi un intérêt dans une perspective de découvrabilité des artistes et de leurs spectacles, ce que préconise la plateforme Scènes francophones.

Pour être « ouverte » et donner lieu à ces avantages, une donnée doit répondre à certaines conditions.

Les données doivent être mises à disposition selon une licence permettant à quiconque de les copier, utiliser et republier, ce qui inclut notamment la possibilité de les combiner à d'autres jeux de données ouvertes, pour n'importe quelles fins, incluant pour des fins commerciales¹⁰. C'est ce qu'implique la licence Creative Commons CC0.

Les données doivent aussi être faciles à trouver. Elles doivent être accessibles gratuitement, dans un format pratique (lisible et traitable par machine) et dans leur intégralité. Elles doivent enfin se conformer à des normes afin d'être interopérables, ainsi que le préconisent les principes FAIR¹¹.

Avant d'être rendues ouvertes, les données doivent être préparées par les organisations qui les détiennent afin d'assurer leur qualité. La qualité des données peut être évaluée selon plusieurs critères : exactitude/fiabilité, complétude, actualité, cohérence/comparabilité et intelligibilité¹².

Les efforts entourant la qualité des données doivent être soutenus par un processus de gestion de la donnée incluant les étapes de nettoyage (repérer et supprimer les doublons, retirer tous les champs inexacts ou incomplets, etc.), de standardisation et, parfois, d'anonymisation des données (dans le cas de renseignements personnels). Cela requiert notamment la responsabilisation d'un ou des employé(s) en matière de qualité des jeux de données. C'est l'approche retenue par Scène Pro, qui est dotée d'une technicienne en contrôle de qualité.

Le processus d'ouverture des données ne se termine pas à la publication de celles-ci. Une fois que les données sont publiées, l'organisation qui les a rendues disponibles doit s'assurer qu'elles

⁹ Wylie, M., Sculthorp, M., Gagnon-Turcotte, S. et Chatwin, M. (2021). *Une voie prometteuse pour promouvoir la gouvernance des données dans le secteur des arts de la scène : analyser les chartes et les principes pour la gouvernance des données*. Nord Ouvert et Association canadienne des organismes artistiques (CAPACOA), 24-26. <https://bac-lac.on.worldcat.org/oclc/1264796075>

¹⁰ Open Knowledge Foundation. (s. d.). *What is Open Data? Open Data Handbook*. <https://opendatahandbook.org/guide/en/what-is-open-data/>

¹¹ Les principes FAIR résument en quatre points universels les propriétés commerciales et techniques à considérer pour générer et/ou rendre des données Faciles à trouver, Accessibles, Interopérables et Réutilisables (Findable, Accessible, Interoperable et Reusable en anglais), aussi bien pour les machines et les algorithmes que pour les humains.

Go FAIR. (s. d.). *Fair Principles*. <https://www.go-fair.org/fair-principles/>

¹² Projet dia-log. (2023). Cahier des recommandations. <https://dia-log.ca/r%C3%A9sultats-et-bilan>

demeurent accessibles dans le temps. C'est-à-dire qu'elles doivent sécuriser un budget stable qui permettra d'allouer les ressources nécessaires à la mise à jour continue de la plateforme et de données elles-mêmes. Dans le cas de la plateforme I Want To Showcase, chaque partenaire verse une contribution annuelle pour l'entretien de la plateforme.

Considérations de consentement en contexte de données ouvertes

Bien sûr, avant d'ouvrir des données, il est important de s'assurer qu'aucun renseignement personnel ni aucune donnée sensible ne seront divulgués sans le consentement éclairé des personnes concernées.

Certaines données peuvent cependant être considérées comme plus ou moins sensibles tout dépendamment du contexte. Dans les plateformes analysées, certaines données qui pourraient autrement être considérées des renseignements personnels prennent une dimension publique dans la mesure où elles décrivent des artistes dans le cadre de leur pratique professionnelle. Par exemple, lorsque l'artiste utilise son prénom et son nom en guise de nom de scène, cela devient une information de nature publique. De la même façon, la région de provenance de l'artiste peut également être ouverte dans l'objectif de revendiquer une appartenance à un territoire. Ces données sont considérées comme moins sensibles que, par exemple, les données d'auto-déclaration d'appartenance à un groupe en quête d'équité. Il faut donc se rappeler que la sensibilité de certaines informations est subjective et peut varier selon les personnes et les contextes.

Une politique de confidentialité claire est utile afin d'explicitier lesquelles des données les concernant pourraient être ouvertes et lesquelles demeureront privées. Une politique de données ouvertes peut également venir compléter ces deux documents-cadre. C'est l'approche qui a été recommandée et mise en œuvre par Scènes francophones lors du projet. La *Politique de gestion de données* de Scènes francophones précise quelles données seront utilisées de quelles façons et est assortie d'une politique de données ouvertes¹³.

Toujours dans le cadre du projet, CAPACOA a procédé à une bonification et à une traduction de la publication [Lier ses données – Trousse de sensibilisation et directive sur le consentement pour les associations artistiques](#)¹⁴. Cette ressource détaille différentes stratégies pour obtenir un consentement éclairé et cite des exemples pour chacune.

Les données ouvertes liées dans le contexte du graphe de connaissances Artsdata

Artsdata est une base de données distribuée, c'est-à-dire une base de connaissances liant entre elles des bases de données locales du secteur des arts, de même que d'autres bases de connaissances dans le web des données ouvertes et liées. Chaque fournisseur de données a son

¹³ <https://scenesfrancophones.ca/legal/politique-de-donn%C3%A9es-ouvertes>

¹⁴ MacIntosh, B. et Julien, F. (2024). Édition 1.1. <https://linkeddigitalfuture.ca/fr/2022/10/25/louverture-des-donnees-passe-par-les-organismes-de-services-aux-arts/>

propre graphe à l'intérieur d'Artsdata. Par exemple, les données de Scène Pro sont chargées dans le graphe suivant :

<http://kg.artsdata.ca/culture-creates/artsdata-planet-scenepro/scenepro-orgs>

Lorsque l'équipe d'Artsdata entreprend d'appuyer un potentiel contributeur dans un processus d'ouverture de données, elle procède d'abord à une analyse sommaire des données du contributeur, conformément aux étapes décrites plus haut. Au besoin, elle peut aussi soutenir le fournisseur dans l'amélioration de la qualité des données et la mise en œuvre de principes FAIR, notamment dans la mise à disposition des données selon un protocole ouvert et lisible par machine.

Vient ensuite le processus appelé « extraction-transformation-chargement » :

- Extraire les données avec l'option technologique la plus appropriée en fonction par exemple de l'état et de la structure du site Web ou de la disponibilité d'une API ;
- Transformer ces données au format RDF, ce qui implique un travail de mise en correspondance du modèle de données du contributeur avec le modèle de données Artsdata (lequel s'appuie sur l'ontologie Schema.org) et les autres modèles chargés dans Artsdata ;
- Charger les données ainsi extraites et transformées dans une instance de base de données graphe dédiée aux données du contributeur ;
- Générer et publier des métadonnées administratives de traçabilité indiquant le contributeur comme la source des données contribuées ;
- Mettre les données du contributeur à disposition pour les consommateurs de données par le biais du terminal SPARQL et des APIs d'Artsdata sous la licence CC0 1.0 Universal et/ou les autres licences copyleft employées par le Contributeur, conformément aux principes FAIR.

À la suite de ce processus, l'équipe d'Artsdata procède à des travaux d'enrichissement des données :

- Réconcilier les instances du jeu de données du contributeur avec la bibliothèque d'instances déjà présentes dans Artsdata (lorsque la qualité des données le permet) ;
- Récupérer les identifiants pérennes associées à ces entités dans Wikidata, Musicbrainz et d'autres bases de connaissances ouvertes (lorsque la qualité des données le permet) ;
- Assigner des identifiants pérennes Artsdata à ces entités (lorsque les données répondent aux exigences minimales pour l'assignation d'identifiants) ;
- Inférer des nouvelles données à partir des données existantes dans Artsdata et de leurs liens avec des entités réconciliées dans d'autres bases de connaissance (lorsque la qualité des données le permet).

Dans le cadre du projet, l'extraction-transformation-chargement des données de Scènes francophones a grandement été facilité du fait que la plateforme encodait déjà ses données au format JSON-LD, un format d'encodage courant pour les données ouvertes et liées.

L'identifiant pérenne : la clé pour lier les données

Dans son processus de réconciliation de données et dans l'objectif de partager des données bien désambiguïsées, Artsdata s'appuie sur les identifiants pérennes. Cela ne constitue toutefois pas une exigence propre à Artsdata, mais plutôt une pratique essentielle pour tout processus de partage de données :

L'attribution d'identifiants uniques et pérennes aux entités que l'on décrit, quelle que soit leur nature, est absolument nécessaire pour garantir la bonne gestion, l'accessibilité et la réutilisabilité des données et des métadonnées que l'on produit¹⁵.

Un identifiant pérenne est une chaîne de caractères alphanumériques qui désigne une entité ou une instance (par exemple, un artiste) de façon permanente et globale, *indépendamment de la base de données où elle est référencée*. Lorsqu'un même identifiant pérenne est référencé dans plusieurs bases de données, on l'appelle parfois un « identifiant passerelle » car il facilite la réconciliation et la circulation des données entre les bases de données¹⁶.

Les identifiants pérennes sont fréquemment exprimés sous forme d'**identifiants uniformes de ressources** (URI). Une URI permet à la fois d'identifier une ressource et d'accéder à ses métadonnées. Une URI est donc actionnable : elle se résout à une localisation sur le Web (tout comme l'URL) où les métadonnées sont mises à disposition dans un format lisible par les humains (par ex., un document HTML) et/ou par les machines (par ex., une application JSON-LD ou un document RDF).

En plus d'identifier et de localiser une entité ou une instance d'entité, l'URI permet aussi de *lier* ces ressources entre elles, de la même manière qu'un hyperlien lie des pages web.

C'est pour l'ensemble de ces raisons que les principes FAIR affirment ceci :

Le principe F1 [les (méta)données se voient assigner des identifiants pérennes globalement uniques] est incontestablement le plus important car il sera ardu de mettre en œuvre les autres aspects de FAIR sans identifiants pérennes globalement uniques¹⁷.

Un identifiant sous forme d'URI (Uniform Resource Identifier) est donc dans le web des données ouvertes et liées ce qu'une clé primaire est dans une base de données relationnelle.

¹⁵ France (2015). *Identifiants pérennes pour les ressources culturelles: Vade-mecum pour les producteurs de données*. <https://www.culture.gouv.fr/Espace-documentation/Publications-revues/Identifiants-perennes-pour-les-ressources-numeriques>

¹⁶ Le modèle de données Artsdata formule des recommandations relativement aux identifiants passerelles, accessibles à <https://culturecreates.github.io/artsdata-data-model/index.fr.html#identifiants-passerelles>

¹⁷ Go FAIR (s.d.). *FAIR Principles*. <https://www.go-fair.org/fair-principles/f1-meta-data-assigned-globally-unique-persistent-identifiers/>

Pour faciliter la circulation de données entre une base de données relationnelle et le web des données ouvertes liées, il est possible de définir un plan de nommage permettant de convertir la clé primaire d'une instance décrite dans la base de données (par exemple, un organisme diffuseur ou un artiste) en une URI désignant la même entité à travers le web des données.

Dans le cadre du projet, le plan de nommage retenu par la plateforme Scène francophone, selon les recommandations de l'équipe de projet, a été le suivant :

`https://scenesfrancophones.ca/id/<clé primaire de l'instance>`

Ainsi, l'URI <https://scenesfrancophones.ca/id/39> désigne le Centre culturel de Caraquet, dont la clé primaire est 39 dans la base de données de Scènes francophones. Lorsque cet URI est appelé dans un navigateur, il se résout à l'URL <https://scenesfrancophones.ca/diffuseurs/centre-culturel-de-caraquet>, qui correspond à la page web du Centre culturel dans le répertoire de diffuseurs de Scènes francophones.

Enfin, dans d'autres circonstances, il aurait aussi été possible de nommer un URI en ajoutant à la suite de l'URL d'une page web existante le symbole # suivi d'une chaîne de caractères faisant office d'identifiant unique. Cette chaîne de caractères s'appelle un « identificateur de fragment ». Cette méthode est expliquée dans la documentation d'Artsdata¹⁸.

Il ne faut pas confondre l'URI d'une chose et l'URL de la page web où cette chose est décrite. L'URL désigne un *support numérique* qui contient des contenus (textuels et audiovisuels) à propos de choses du monde réel (personnes, organismes, événements, etc.). L'URI désigne plutôt une chose décrite dans les contenus de la page web.

Méthodes de publication et d'extraction des données ouvertes

Pour favoriser la réutilisation des données, celles-ci doivent être rendues accessibles selon des protocoles ouverts et lisibles par machine. Dans le cadre du projet, quatre méthodes de publication et d'extraction des données ont été envisagées pour permettre à Artsdata (ou à d'autres parties) d'accéder aux données des plateformes.

Publication et extraction frontale page par page

Cette méthode implique d'utiliser un « robot » pour moissonner les données structurées, notamment les données suivant l'ontologie Schema¹⁹, et les données non structurées dans le corps de la page web. Les avantages de cette méthode sont :

- la simplicité de mise en place, tant pour le fournisseur de données que pour Artsdata;
- la sécurité des données, parce qu'elles sont accessibles en lecture seulement;
- l'automatisation des processus;
- l'incidence positive des données structurées Schema sur les robots d'indexation des moteurs de recherche.

¹⁸ <https://culturecreates.github.io/artsdata-data-model/gabarits-jsonld/README.fr.html#fn:2>

¹⁹ <https://schema.org/docs/schemas.html>

Le désavantage de cette méthode est que peu de plateformes sont en mesure d'extraire des données de cette manière. Par ailleurs, la configuration d'un robot de moissonnage étant spécifique à la structure du site, tout changement à celle-ci peut perturber le processus.

Cette approche a été initialement considérée pour l'extraction des données de Scènes francophones. Toutefois, comme aucune section du site public n'exposait les données sur les salles de spectacles, l'approche suivante a dû être employée.

Publication et extraction frontale par type de média MIME

Dans ce mode d'extraction, un jeu de données est exposé séquentiellement sur une seule URL dans un type de média MIME approprié (par exemple, `application/json`, `application/ld+json`, `application/rdf+xml`, `text/xml`, `text/turtle`, etc.). Parfois, une négociation de contenu²⁰ peut être employée afin de présenter le jeu de données selon différents types MIME (incluant ou non une page HTML lisible par des humains). Les avantages de cette méthode sont :

- une simplicité relative (en comparaison à la mise en place d'une API);
- la sécurité, parce que les données sont accessibles en lecture seulement;
- l'automatisation des processus;
- le potentiel de contribuer à la découvrabilité des entités exposées selon cette méthode;
- la négociation de contenu.

Les désavantages de cette méthode sont :

- un potentiel de latence dans la mise à jour des données qui dépend de la fréquence de rafraîchissement de la mémoire cache;
- le recours à des compétences techniques particulières pour la mise en place et pour la transformation des données vers le format de média désiré;
- la nécessité d'entretenir les scripts d'extraction et de transformation en fonction des changements appliqués à la base de données.

Cette stratégie a été proposée et retenue par Scènes francophones dans le cadre du projet. Tant le déploiement par les développeurs de Scènes francophones que l'extraction par Artsdata se sont avérés simples et efficaces.

▷ [Terminaux JSON-LD de Scènes francophones](#)

Extraction par API

Une API (Application Programming Interface) permet à deux logiciels ou plus de communiquer. Alors qu'une interface utilisateur fait le pont entre un humain et un logiciel, une API fait le pont entre différents logiciels. Dans le cadre de ce projet, une API serait une interface permettant d'interroger directement une base de données pour en extraire les données d'intérêt. Les avantages de cette méthode sont :

²⁰ https://fr.wikipedia.org/wiki/N%C3%A9gociation_de_contenu

- la possibilité d'être appelée par un grand nombre d'application pour réaliser un grand nombre d'opérations, outre la lecture de données;
- l'automatisation.

Les désavantages sont :

- la difficulté de mise en place (en comparaison avec les autres méthodes proposées dans ce document);
- la nécessité d'entretenir l'API.

Cette stratégie a été mise en place par Scène Pro lors d'un projet antérieur réalisé avec l'appui de l'initiative *Un avenir numérique lié* de CAPACOA.

▷ [Documentation de l'API de Scène Pro](#)

Exportation des données au format CSV

La manière la plus simple d'extraire des données est de les exporter de leur source en un fichier CSV. Ce type de fichier peut aisément être lu par un logiciel tel qu'Excel, ce qui facilite l'étape de transformation. Le désavantage de cette méthode est qu'elle nécessite une intervention manuelle (à moins de mettre en place un script qui génère lui-même le fichier à fréquence régulière). Ainsi, ce mode d'extraction convient bien pour une extraction ponctuelle, mais moins à un processus que l'on souhaite récurrent.

Observations sur les systèmes d'information analysés

Le but premier et commun aux plateformes analysées est de centraliser et faciliter des processus d'affaires de l'industrie du spectacle, notamment les demandes et la sélection des vitrines aux différents événements de réseautage du milieu. Typiquement, ces événements de réseautage permettent aux artistes de faire valoir leur spectacle dans l'espoir qu'il soit acheté par un ou des diffuseurs. Ainsi, le processus de dépôt de candidature permet aux artistes de faire une demande pour obtenir une vitrine à ces événements de réseautage, c'est-à-dire un moment sur scène pendant lequel ils présentent un extrait ou l'intégralité du spectacle qu'ils souhaitent vendre aux diffuseurs.

Certaines plateformes gèrent également le processus d'inscription aux événements de réseautage qui permet d'obtenir son accréditation pour participer à l'événement. Enfin, certaines plateformes gèrent le processus de paiement, d'autres non.

Certaines plateformes utilisent l'information qu'ils collectent à propos de la programmation des diffuseurs pour bâtir un calendrier de l'offre culturelle que le public peut consulter, ce qui participe à la découvrabilité des artistes et des spectacles.

Tableau 1 : Aperçu des plateformes et des processus d'affaires soutenus par chacune

Processus d'affaires	Scène Pro	Scènes francophones	I Want To Showcase	Arts Touring Connector
Candidature de vitrines	✓	✓	✓	
Répertoire d'offres de spectacles	✓	✓		
Répertoires de diffuseurs et de lieux	✓	✓		✓
Programmation en groupe				✓
Calendrier d'événements		✓		

Enfin, certaines plateformes couvrent des régions spécifiques du Canada ou des communautés linguistiques particulières.

Étant donné la visée pancanadienne d'Artsdata, la présence de ces plateformes constitue une opportunité d'accès à un large éventail de données sur l'écosystème des arts de la scène du Canada.

Cependant, l'analyse a révélé un haut niveau d'hétérogénéité dans la nature et la structure des données collectées par celles-ci à propos des artistes, des spectacles, des diffuseurs et des lieux de diffusion. Tant les entités principales que leurs attributs diffèrent considérablement d'une plateforme à une autre. Cela s'explique en partie par la variation et l'étendue des processus d'affaires intégrés aux plateformes (candidature, inscription, paiement, calendrier), ainsi que par les contextes régionaux ou linguistiques dans lesquels les gestionnaires et utilisateurs des plateformes opèrent.

Cette hétérogénéité fait obstacle à la mise en place d'un processus uniforme et standard d'ouverture des données. Artsdata ne peut appliquer la même recette d'une plateforme à une autre, ce qui restreint les gains possibles en termes d'effort et de temps.

Observations sur les ontologies des plateformes analysées

Dans le cadre de notre analyse, nous avons identifié des correspondances fondées à la fois sur les entités (ou classes) et leurs attributs (ou propriétés).

Nous avons identifié quatre principaux groupes d'entités présents dans la plupart des sources de données :

- Des entités désignant des acteurs, soit des personnes physiques, des groupes ou des organismes;
- Des entités désignant des spectacles, dans le sens d'œuvres destinées à être représentées sur scène;
- Des entités désignant des événements et plus spécifiquement des représentations de spectacles;
- Des entités désignant des lieux, notamment des salles de spectacles.

Observations :

- Malgré de nombreuses classes équivalentes entre les ontologies, il a été constaté que ces classes partagent peu de propriétés en commun. Lorsque les classes équivalentes partagent peu de propriétés, les ontologies alignées peuvent être moins réutilisables dans d'autres contextes. L'absence de propriétés communes limite l'interopérabilité entre les systèmes utilisant ces ontologies. Les échanges de données peuvent être incomplets entre les systèmes. Une des méthodes pour remédier à cela serait d'enrichir davantage ces ontologies en incluant les propriétés manquantes ou en les rendant plus compatibles. Cela peut impliquer l'ajout de nouvelles propriétés ou l'adaptation des définitions existantes.
- Des disparités basées sur la synonymie et les concepts issus de différentes langues ont été observées. Par exemple, dans I Want To Showcase, le terme « Performer » est utilisé pour décrire un acteur du domaine des arts de la scène qui propose un extrait de spectacle pour un événement vitrine. Arts Touring Connector et Scènes francophones emploient quant à eux le terme « Artist » / « Artiste » (respectivement) pour exprimer un concept similaire en lien avec des représentations de spectacles. Dans les trois plateformes, le terme peut désigner autant un artiste individuel, qu'un groupe, un collectif ou une compagnie.
- Des disparités basées sur la cardinalité ont été observées. ATC a établi une cardinalité plusieurs à plusieurs (0,m:1,m) pour la relation entre la classe « Artist » et la classe « Show » : un artiste peut être lié à aucun, un ou plusieurs spectacles ; un spectacle doit être lié à au moins un ou à plusieurs artistes. Cette cardinalité permet par exemple de représenter deux compagnies de théâtre qui s'unissent pour coproduire un spectacle ou encore deux musiciens qui forment un duo le temps d'une tournée. Scènes francophones avait pour sa part conçu son interface de sorte que les utilisateurs s'y voyaient contraints de se conformer à une cardinalité un à un : chaque artiste ne pouvait avoir qu'un seul spectacle, et chaque spectacle n'était associé qu'à un seul artiste. On voyait par

conséquent apparaître dans le jeu de données des doublons (deux instances pour le même artiste) ou encore une instance d'artiste combinant deux compagnies de théâtre. À la lumière de ces enjeux, Scènes francophones a apporté des correctifs en cours de projet et a mis en œuvre une cardinalité plusieurs à plusieurs, à l'instar de ATC.

- Des disparités conceptuelles basées sur le niveau de granularité ont été observées. Par exemple, dans l'ontologie du Référentiel du Québec, le terme « Contributeur » est beaucoup plus générique et possède divers sens. En revanche, l'ontologie de Scènes francophones offre des détails plus précis sur ce concept générique, en utilisant des termes spécifiques comme « Artiste », « Diffuseur » et « Réseau ». Lorsque des concepts génériques dans une ontologie doivent être alignés avec des concepts plus spécifiques dans une autre, il devient difficile de définir des correspondances précises et significatives. Cela nécessite des règles de mapping plus complexes et précises pour gérer les différences de granularité. En considérant l'exemple mentionné sur le concept de « Contributeur », les correspondances qui existeront dans ce cas seront complexes et de nature (1:m), où **1** représente une seule entité d'une source de données (IWTS, ATC, Scènes francophones) et **m**, son équivalence sous forme d'une formule logique dans le sens du terme « Contributeur » dans le Référentiel du Québec. En prenant l'exemple du concept générique « Contributeur » dans le Référentiel, et en considérant les concepts « Artiste », « Diffuseur » et « Réseau » dans Scènes francophones, nous pouvons établir des correspondances de ces trois concepts avec celui du Référentiel de la manière suivante :

Tableau 2 : Exemples de correspondance complexe

Concept générique (Référentiel du Québec)	Concept spécifique (Scènes francophones)	Équivalence en logique des propositions	Description de l'équivalence
Contributeur	Artiste	$\forall x(\text{Artiste}(x) \Leftrightarrow (\text{Contributeur}(x) \wedge \text{typeContributeur}(x, \text{"Artiste"})))$	"Artiste" est équivalent à Contributeur où typeContributeur a pour valeur "Artiste".
Contributeur	Diffuseur	$\forall x(\text{Diffuseur}(x) \Leftrightarrow (\text{Contributeur}(x) \wedge \text{typeContributeur}(x, \text{"Diffuseur"})))$	"Diffuseur" est équivalent à Contributeur où typeContributeur a pour valeur "Diffuseur".
Contributeur	Réseau	$\forall x(\text{Réseau}(x) \Leftrightarrow (\text{Contributeur}(x) \wedge \text{typeContributeur}(x, \text{"Réseau"})))$	"Réseau" est équivalent à Contributeur où typeContributeur a pour valeur "Réseau".

- Comme cela a déjà été mentionné dans les observations sur les systèmes d'information analysés, des disparités existent dans la couverture des ontologies, ainsi qu'une polysémie des termes employés dans différentes sources. Automatiser le processus d'alignement devient plus complexe en présence de ces différences dans les définitions et l'utilisation des mêmes termes, ce qui peut conduire à des incohérences lorsqu'on tente de fusionner ou d'interopérer les données des différentes plateformes. Cela nécessite alors une intervention humaine pour interpréter ces divergences et, surtout, des experts du domaine pour déterminer les correspondances entre les entités.

Le tableau 3 qui suit présente l'alignement des classes. Pour des fins de concision, les propriétés ont été exclues du tableau. Pour plus de détails, veuillez vous référer au chiffrier :

▷ [Alignement des ontologies des Arts de la scènes](#)

Interprétation des données du tableau : Les correspondances parfaites représentent les classes ayant une signification identique ou très similaire (ici on parlera de correspondance équivalente). Lorsqu'une classe est plus générique qu'une autre, la relation entre ces deux classes devrait être celle de superclasse et sous-classe. Nous dirons donc que cette superclasse subsume l'autre, et cette relation devient une relation de subsomption.

Préfixes des ontologies référencée dans le tableau :

- **adr:** désigne PREFIX adr: <http://kg.artsdata.ca/resource/>
- **schema:** désigne PREFIX schema: <http://schema.org/>

Tableau 3 : Correspondances identifiées lors de l'analyse des ontologies

Référentiel du Québec (RQ)		Scènes Francophones (SF)		Artsdata		IWTS		ATC	
Entités	Nature correspondance	Entités	Nature correspondance	Entités	Nature correspondance	Entités	Nature correspondance	Entités	Nature correspondance
Contributeur and typeContributeur value Artiste or groupe	Contributeur Subsume des entités de (SF, ATC,IWTS)	Artiste	match parfait des entités- (ATC,IWTS,SF) -SubclassOf Contributeur (RQ)	PerformingArtGroup or schema:Person (and profession value performer).	<i>PerformingArtGroup est une sous classes de Artiste(SF) Performer (IWTS) et Artist (ATC)</i>	Performer	match parfait des entités- (ATC,IWTS,SF) - SubclassOf Contributeur (RQ)	Artist	match parfait des entités- (ATC,IWTS,SF) - SubclassOf Contributeur (RQ)
Lieu	Sous classe de 'Place' Dans AD	/	/	schema: Place	<i>Subsume l'entité "lieu" de RQ</i>		/	/	/
Salle	"Match Parfait (RQ,ATC,SF)" subclassOf Lieu(RQ)	SallesDeSpectacles	"Match Parfait (RQ,ATC,SF)" subclassOf Lieu(RQ)	/	/	/	/	Performance_ space	"Match Parfait (RQ,ATC,SF)" subclassOf Lieu(RQ)
Spectacle	"Match Parfait (RQ,ATC,SF)" and SubclassOf Oeuvre (RQ)	OffreSpectacle (Spectacle)	"Match Parfait (RQ,ATC,SF)" and SubclassOf Oeuvre (RQ)	schema:CreativeWork	Pour CreativeWork, cette entité n'est pas clairement spécifiée, ce qui complique l'établissement des correspondances.	/	/	Show	"Match Parfait (RQ,ATC,SF)" and SubclassOf Oeuvre (RQ)
Offre				schema:Offer	<i>Subsume l'entité "Offre" de RQ</i>				
Contributeur and typeContributeur value Diffuseur	Contributeur Subsume des entités de (SF,AD,ATC)	Diffuseur	"Match Parfait entités de (ATC,SF,AD)" SubclassOf Contributeur (RQ) and adr:ArtsOrganization(AD)	adr: PresentingOrganization	"Match Parfait entités de (ATC,SF,AD)" SubclassOf Contributeur (RQ) and adr:ArtsOrganization(AD)	/	/	Member	"Match Parfait entités de (ATC,SF,AD)" SubclassOf Contributeur (RQ) and adr:ArtsOrganization(AD)
/	/	EvenementContact	"Match Parfait(SF,AD,IWTS)" SubclassOf schema:event(AD)	adr:Conference	"Match Parfait(SF,AD,IWTS)" SubclassOf schema:event(AD)	Conference	"Match Parfait(SF,AD,IWTS)" SubclassOf schema:event(AD)	/	/
/	/	RepresentantArtiste	"Match Parfait (ATC,SF,AD)" SubclassOf adr:ArtsOrganisation (AD)	adr: ArtistRepresentative	"Match Parfait (ATC,SF,AD)" SubclassOf adr:ArtsOrganisation (AD)	/	/	Representative	"Match Parfait (ATC,SF,AD)" SubclassOf adr: ArtsOrganisation (AD)
Représentation	"Match Parfait(RQ,SF,AD)" SubclassOf schema:event(AD)	Évènement	"Match Parfait(RQ,SF,AD)" SubclassOf schema:event(AD)	adr: PerformingArtsEvent	"Match Parfait(RQ,SF,AD)" SubclassOf schema:event(AD)	/	/	/	/
Série	" SuperClasse de adr:EventSerie(AD)			adr: EventSeries	"Match Parfait(AD,ATC)" SubclassOf SÉRIE(RQ) and schema:event(AD)	/	/	Tour	"Match Parfait(AD,ATC)" SubclassOf SÉRIE(RQ) and schema:event(AD)
Contributeur and typeContributeur value Réseau de diffusion	Contributeur Subsume des entités de (SF,AD,ATC)	Réseau	"Match Parfait des entités de (ATC,SF,AD)" SubclassOf Contributeur (RQ) and ArtsOrganization(AD)	adr: ArtsServiceOrganization	"Match Parfait des entités de (ATC,SF,AD)" SubclassOf Contributeur (RQ) and adr:ArtsOrganization(AD)	/	/	Organization	"Match Parfait des entités de (ATC,SF,AD)" SubclassOf Contributeur (RQ) and adr:ArtsOrganization(AD)

Observations sur l'équité, la diversité et l'inclusion

De manière générale, l'information sur les publics cibles ne précise pas si les spectacles proposés peuvent être appréciés par un public sourd, malentendant, ou non-voyant.

De manière générale, il n'est pas précisé si les lieux présentant les spectacles peuvent accueillir un public handicapé en évitant que celui-ci soit en situation de handicap²¹.

Les plateformes analysées comportent un ou des champ(s) qui permettent aux individus d'autodéclarer leur appartenance culturelle ou leur appartenance à la communauté LGBTQIA2+. Les options disponibles pour autodéclarer son appartenance culturelle ne sont pas standards ou uniformes d'une plateforme à l'autre.

Toutes les plateformes ne comptent pas un champ permettant d'autodéclarer son identité de genre.

Toutes les plateformes ne permettent pas de distinguer aisément l'artiste solo du groupe ou du collectif, ce qui rend plus difficile pour les individus d'autodéclarer leur appartenance culturelle et leur identité de genre.

Davantage de ressources en EDI sont disponibles sur le site web de Mass Culture [ici](#)²².

²¹ « La personne exposée à [un handicap] est désignée par le terme [personne handicapée](#). On dit que cette personne se trouve en [situation de handicap](#) lorsque son autonomie est limitée parce qu'elle n'est pas en mesure d'exercer pleinement une activité courante ou un rôle social en raison de l'absence de [facilitateurs](#), destinés à lever les obstacles dans son environnement. Par exemple, si une personne ayant une incapacité liée à la mobilité doit monter à un étage supérieur, elle ne sera pas en situation de handicap si elle se déplace en fauteuil roulant et a accès à un ascenseur. Au contraire, elle sera en situation de handicap s'il lui est impossible de prendre un ascenseur. »

<https://vitrinelinguistique.oqlf.gouv.qc.ca/fiche-gdt/fiche/8385452/handicap>

²² Bernicky, S. (2024). *La spirale de l'équité dans les arts publics*. <https://massculture.ca/recherche-en-residence/la-spirale-de-lequite-dans-les-arts-publics/#>

Recommandations

Recommandation générale sur la conception d'un système d'information et son alignement ontologique

La modélisation conceptuelle devrait faire l'objet d'une attention particulière lors de la conception et de l'entretien d'un système d'information. La modélisation conceptuelle des données permet d'identifier, de définir et de visualiser tous les éléments d'information qui doivent être représentés dans le système d'information, de même que leurs relations, indépendamment des choix technologiques qui seront faits. Cette modélisation devrait être réalisée tôt dans le processus de conception d'un système d'information, telle qu'une plateforme web. Elle doit s'appuyer sur une expertise du domaine davantage que sur une expertise technique. Elle doit enfin être documentée – notamment dans un dictionnaire entités-attributs et dans des schémas – et mise à jour à mesure que la compréhension du fonctionnement du milieu évolue.

Au moment de réaliser la modélisation, les différents référentiels ontologiques du milieu des arts devraient être consultés et considérés afin d'harmoniser, autant que faire se peut, la structure de données aux modèles ontologiques existants. Aligner ontologiquement la modélisation conceptuelle sur les référentiels ontologiques existants, lorsque possible, favorisera la pérennité des plateformes numériques ainsi créées. L'alignement ontologique implique de revoir la manière dont les classes sont définies de sorte qu'elles s'accordent mieux avec les définitions des référentiels existants. L'alignement ontologique peut aussi impliquer l'ajout d'attributs employés pour décrire les classes. Enfin, il peut être nécessaire de réviser les noms donnés aux classes et aux attributs afin d'éviter les problèmes de synonymie, d'homonymie et de polysémie.

La modélisation conceptuelle n'est pas seulement une bonne pratique de design. Elle peut également fournir divers documents qui peuvent être demandés, en cas d'incident de sécurité ou confidentialité, par les autorités en charge de la protection de la vie privée et des renseignements personnels au Canada.

Il est important de souligner que les changements apportés au système d'information peuvent déclencher le besoin de réviser le(s) processus d'affaires exprimé(s) par le système d'information. Alors, l'envergure de l'initiative sera toute autre et il est recommandé de revoir la définition du projet dans lequel s'inscrivent les changements au système d'information afin de valider les échéanciers, les budgets, les contraintes, les attentes et la disponibilité des ressources appropriées.

Recommandation sur la granularité du concept artiste

Dans la collecte de données réalisée par les plateformes, il serait judicieux d'associer des types d'entités plus précis aux instances d'artistes :

- Artiste solo (c'est-à-dire une personne physique);
- Groupe ou collectif (par exemple, un duo ou un groupe de musique);
- Organisme (par exemple, une compagnie de théâtre ou de danse).

Le fait d'augmenter la granularité, soit le niveau de précision, de l'entité artiste présente plusieurs avantages :

- désambiguïser l'artiste principal et les interprètes qui l'accompagnent;
- permettre une utilisation adéquate des identifiants pérennes;
- permettre l'autodéclaration de l'appartenance culturelle des groupes et collectifs indépendamment de l'autodéclaration de l'appartenance culturelle des individus.

En définitive, ce genre de granularité aura une incidence positive sur le niveau d'interopérabilité des différents jeux de données, favorisant leur ouverture et leur mise en commun par le biais d'un système tiers tel que Artsdata.

Recommandations de l'ordre de l'EDI

Les données concernant l'identité de genre et l'appartenance culturelle ne devraient être rendues ouvertes que si, et seulement si, les plateformes sont en mesure de garantir leur anonymisation. La teneur de ces données pouvant porter préjudice aux personnes concernées par les données, l'anonymisation est essentielle.

À long terme, il pourrait être utile aux plateformes de mettre en place des modules d'accessibilité dans leur interface.



Image 2 : Le spectacle Êtres de bois, de la compagnie Fleuve Espace Danse, apparaît dans le répertoire d'offres de spectacles de Scène Pro. Interprète : Victoria Coté. Crédit photo : Emie-Liza Caron St-Pierre.

Conclusion

Au sortir de la pandémie, il est clair que des améliorations et une consolidation des systèmes d'information des arts de la scène au Canada seraient utiles pour accroître la résilience du secteur. Devant ce constat, CAPACOA, stratégiquement placé au cœur du secteur canadien des arts de la scène, a soumis un projet au Fonds de relance des services communautaires dans l'objectif de cartographier les systèmes d'information clés des arts de la scène, documenter les disparités dans les ontologies et recommander des améliorations orientées sur l'ouverture et la réutilisation des données.

CAPACOA a mobilisé quatre de ses partenaires, chacun propriétaire d'une plateforme numérique recueillant, entreposant et utilisant de la donnée sur les arts de la scène, dans diverses régions du Canada. Les systèmes d'information que chacune de ces plateformes sous-tendent ont été modélisés et analysés, ce qui a mis en lumière l'hétérogénéité des modèles ontologiques utilisés par ces plateformes. Il s'agit là du principal résultat de ce projet puisqu'il donne une mesure du travail à accomplir et qu'il a permis, déjà, de mener des expérimentations visant à amoindrir cette hétérogénéité et à entamer des processus d'ouverture de données.

L'autre résultat principal de ce projet est le constat de l'impact des modes de gouvernance des plateformes sur leur fonctionnement et, dans le contexte du projet, sur les processus d'ouverture et de partage de données. Loin d'être une critique sur les choix de gouvernance des organisations propriétaires des plateformes numériques analysées, il s'agit plutôt de reconnaître l'importance et la validité de ces différents modes de gouvernance afin d'en tenir compte dans les étapes à venir ou dans tout autre projet similaire.

Enfin, il importe de catégoriser l'initiative en tant que projet de transformation numérique puisqu'en l'essence, le projet vise à ouvrir les données des arts de la scène et à favoriser leur réutilisation. En tant que projet de transformation numérique, celui-ci doit s'inscrire dans une planification stratégique allant au-delà de la décision d'acquérir ou non la « dernière application à la mode »²³ et toujours prendre en compte les différents enjeux sociotechniques inhérents à cette catégorie de projet.

²³ Pelletier, C. et Cloutier, L. M. (2019). *Challenges of digital transformation in SMEs : Exploration of IT-related perceptions in a service ecosystem*. Dans Proceedings of 52nd Hawaii International Conference on System Sciences (HICSS), Grand Wailea, HI.
<https://scholarspace.manoa.hawaii.edu/bitstream/10125/59934/0494.pdf>