

MOBILISATION DES DONNÉES STRUCTURÉES COMME MOTEUR DE LA DÉCOUVRABILITÉ DU SPECTACLE VIVANT

RAPPORT DE RÉSULTATS
Octobre 2025



Notice bibliographique suggérée :

Ramboanasolo, Gaëlle, Christian Roy, Jean-Robert Bisaillon, Anika Pascale Papillon et Frédéric Julien. 2025. Mobilisation des données structurées comme moteur de la découvrabilité du spectacle vivant : Rapport de résultats. Association canadienne des organismes artistiques (CAPACOA).

Dépôt légal : Bibliothèque et Archives Canada

ISBN : 978-1-0688996-3-8

Photo en page couverture : [Burn Baby, Burn](#), une production de Côté Danse ([Q126952121](#), [K4-155](#)). La représentation de ce spectacle au Burlington Performing Arts Centre a fait partie du groupe B dans les tests A/B. Crédit photo : Aidan Tooth ; interprète : Griffen Grice ; conception des costumes : Yso South.

Ce rapport est publié par CAPACOA, l'Association canadienne des organismes artistiques, dans le cadre du projet Artsdata. Le projet Artsdata bénéficie du soutien du gouvernement du Canada et du Conseil des arts 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 ou du Conseil des arts du Canada.

Canada



Cet ouvrage est mis à disposition sous licence [Creative Commons Attribution 4.0 International](#).

Table des matières

Sommaire	1
Glossaire	3
Contexte et objectif	7
Introduction et hypothèses	7
Méthodologie	9
Prémisse méthodologique	9
Création d'une typologie	9
Algorithme de calcul d'un indice de données structurées	11
Panier de requêtes et de moissonnage des résultats Google Search	12
Étape 1 : État des lieux du référencement Web du spectacle	12
Étape 2 : Développement d'une approche de quantification de l'exposition	13
Études de cas, tests comparatifs A/B et groupe témoin O	14
Préambule	14
Confection des groupes	14
Échantillonnage des événements et paramètres observés	15
Variables cibles : trafic et découvrabilité sur les moteurs de recherche	17
Limites de la méthodologie et décisions quant aux analyses	18
Résultats principaux	20
Résultats de l'indice d'exposition	20
Étape 1 : État des lieux du référencement Web du spectacle	20
Étape 2 : Développement d'une approche de quantification de l'exposition	22
Résultats des test A/B et de l'analyse du trafic Web	24
Analyses descriptives	24
Effet des données structurées sur le trafic des pages d'événements	26
Performance du modèle	26
Résultats observés	28
Certaines propriétés spécifiques ont-elles un impact positif ?	28
Étude de deux événements comparables	29
À propos des événements comparés	29
Résultats observés	30
Résultats satellites	31
Sites de revente de billets	31
Agents conversationnels	31
Activités de communication	34
Recommandations et ressources	35
Conclusion	37
Références	38
ANNEXE – Exemples de données structurées	41
Données structurées du Centre des arts Juliette-Lassonde	41
Données structurées du Sanderson Centre	42
Données structurées proposées par ChatGPT	45

Sommaire

Nos travaux avaient pour objectif d'illustrer et de mesurer l'impact de l'utilisation des données structurées dans la recherche et la recommandation en ligne d'événements des arts de la scène.

En premier lieu, nous avons défini une typologie granulaire des facteurs de rayonnement et de découvrabilité. En second lieu, nous avons établi les méthodologies permettant la création de deux indices de rayonnement permettant le calcul des marqueurs de la découvrabilité : l'un quantifiant la structuration des données; l'autre quantifiant le rayonnement du spectacle vivant sur les pages de résultats des moteurs de recherche (c.-à-d., Google Search). Enfin, nous avons corrélé ces indices avec des données d'achalandage Web tirées de Google Analytics.

Le processus que nous avons eu à cœur d'illustrer consiste à agir sur une série de facteurs (variables d'action, balises Schema) qui ont une incidence sur des marqueurs (variables de conséquences), lesquels permettent de produire des indices mesurés de la découvrabilité :

Facteurs > Marqueurs > Indices.

Cette approche multimodale, nous a permis de tenir compte des facteurs confondants (autres variables pouvant faire fluctuer nos marqueurs et indices, tels le genre, le public cible ou l'achat de campagnes publicitaires) et ainsi d'isoler l'effet des variables de structuration de données.

Les faits saillants de nos travaux sont les suivants.

Même en tenant compte des facteurs confondants, on observe des relations positives entre le niveau de structuration des données des événements et :

- l'exposition sur les pages de résultats de Google ; ainsi que,
- le nombre de pages vues provenant de Google.

La relation entre le niveau de structuration des données et le nombre de pages est relativement linéaire : plus les données sont structurées, plus le nombre de pages vues est important.

Au-delà des données structurées minimales (le nom, la date/heure et le lieu de l'événement), les données décrivant le contenu de l'événement et ses principaux acteurs ont un effet particulièrement marqué. En outre, la *description* de l'événement a un effet statistiquement significatif. De même, les données décrivant l'*artiste sur scène* et l'*organisateur* de l'événement ont aussi un effet non négligeable sur le nombre de pages vues.

Enfin, nous avons pris en considération l'évolution actuelle des habitudes de navigation des internautes, notamment l'usage accru des outils de clavardage assistés par l'intelligence artificielle.

Au terme de ces travaux, tel qu'anticipé par nos hypothèses, nous concluons que l'ajout de données structurées aux pages Web par les diffuseurs d'événements des arts de la scène a des incidences mesurables sur la découvrabilité numérique de ces événements.

L'actuel rapport présente plus granulairement les gains et les incidences de ce travail de documentation des arts de la scène en ligne et en tire quelques recommandations visant à améliorer les bonnes pratiques d'affaires sectorielles en la matière.



[Kuné \(Q125392378, K2-5096\)](#). Le concert de cet ensemble au Sanderson Centre a fait partie du groupe B dans les tests A/B. Crédit photo : Zahra Saleki.

Glossaire

Balisage (on utilise parfois le terme indexation) : action qui consiste à ajouter des données structurées (les balises sémantiques, ex. un triplet [Schema.org](https://schema.org)) à une page Web afin de générer des marqueurs de rayonnement. La présence de ce marqueur constitue un facteur qui incarne un indicateur grâce auquel il est possible de produire un indice calculé.

Consommation (voir marqueur) : valeur quantitative de conversion, de clic, de page vue, d'achat, soit une action posée par l'internaute découlant de son parcours en ligne.

Découvrabilité et rayonnement : ces deux termes peuvent prêter à confusion. Le premier est défini formellement par l'Observatoire de la culture et des communications du Québec et mis en contexte par Michèle Rioux et col. (2019, p.41) et selon la formule consacrée suivante : « La découvrabilité est la capacité d'un contenu culturel à se laisser découvrir aisément par le consommateur qui le recherche et à se le faire proposer au consommateur qui n'en connaissait pas l'existence. »

Le notion de rayonnement s'est cristallisée quant à elle dans le rapport Observatoire de la culture et des communications du Québec (OCCQ) et Danvoye (2021, p.10) et son trait distinctif est d'ajouter la mesure de la consommation effective à celle de la découvrabilité ¹. Par exemple, Rioux et col. (2019) mentionnent le rayonnement par la formule suivante : *favoriser le rayonnement par l'exportation*.

Pour qualifier et quantifier les résultats de mesure de ces phénomènes, le rapport OCCQ et Danvoye utilise le terme « marqueur ».

Données structurées : données et métadonnées conditionnées selon la logique **RDF** (Resource Description Framework) ou « cadre de description des ressources », un modèle de graphe destiné à décrire de façon formelle les ressources Web et leurs métadonnées et à permettre le traitement automatique de telles descriptions (Wikipédia).

Un énoncé de données structuré et formel repose sur un triplet : **sujet, prédicat, objet**.

Il est possible de faire l'analogie du triplet avec la syntaxe habituelle : sujet, verbe, complément, par exemple :

La salle Desjardins est située dans le Centre des arts Juliette-Lassonde, à Saint-Hyacinthe, au Québec.

¹ Voir Figure 1

Les données structurées intégrées aux pages web à l'intention des moteurs de recherche sont le plus souvent représentées selon le vocabulaire [Schema.org](http://schema.org) et encodées dans le format JSON-LD ².

Ainsi représentée, l'affirmation concernant la Salle Desjardins représentée deviendrait le bloc de code suivant :

```
<script type="application/ld+json">
{
  "@context": "http://schema.org",
  "id": "http://kg.artsdata.ca/resource/K2-227",
  "@type": "Place",
  "name": "Salle Desjardins",
  "sameAs": "http://www.wikidata.org/entity/Q111668872",
  "containedInPlace": {
    "@type": "Place",
    "name": "Centre des arts Juliette-Lassonde"
  },
  "address": {
    "@type": "PostalAddress",
    "streetAddress": "1705 rue Saint-Antoine",
    "addressLocality": "Saint-Hyacinthe",
    "addressRegion": "QC",
    "postalCode": "J2S 9E2",
    "addressCountry": "CA"
  }
},
</script>
```

Facteur : action, ajout de balises, de données factuelles, de métadonnées qui permet d'enrichir la description d'un événement avant d'en mesurer les effets sur la consommation.

Identifiant pérenne : code informatique permanent, qu'importe si des informations sont modifiées, et compréhensible ou non par les humains, qui permet aux algorithmes et machines de distinguer une entité du monde réel dans un système d'information. Par exemple, le code ISBN ([978-2376620167](https://www.isbn-international.org/en/978-2376620167)) pour un livre papier vendu en ligne, le code QID ([Q560434](https://www.wikidata.org/wiki/Q560434)) pour toute entité décrite Wikidata ou l'identifiant Artsdata (ad) pour toute entité décrite dans le graphe de connaissances Artsdata ([K2-5867](https://kg.artsdata.ca/resource/K2-5867)).

Indicateur : voir Marqueur.

Indice : pointage calculé adossé à une échelle ou un pourcentage qui *cherche à représenter les marqueurs de rayonnement - c'est-à-dire, dans quelle mesure les meilleures pratiques favorisant le rayonnement ont-elles été appliquées*. Cette notion s'applique également au

² Voir cette documentation Google : <https://developers.google.com/search/docs/appearance/structured-data/intro-structured-data?hl=fr>

concept de découvrabilité. Le projet a mis au point un indice d'exposition sur les pages de résultats de Google et un indice de données structurées mesurant le niveau d'enrichissement pour une page Web donnée utilisant les balises [Schema.org](https://schema.org) appropriées.³

Marqueur : conséquence, variable quantitative de résultat associée au concept de rayonnement et qui témoigne de la présence d'une action posée reposant sur une bonne pratique de balisage : *mesures qui cherchent à représenter explicitement le rayonnement*.

Métadonnée : information documentaire, balise, facteur, associée à un objet numérisé pour permettre à un humain ou à un robot (algorithme, programme informatique, machine) à mieux en comprendre le contenu. *Par exemple*, le **titre** d'un spectacle ou sa **date** de représentation.

Métadonnée enrichie : métadonnée associée principalement aux étapes créatives d'une production en amont de sa mise en marché et de la promotion d'une représentation (essentiellement les œuvres et les expressions - voir la définition OEMI). *Par exemple*, les **contributions** de type écriture, mise en scène, conception des décors ou éclairages. Des informations enrichies liées aux manifestations-représentations peuvent s'ajouter, par exemple, le nom des **artistes interprètes** susceptibles de changer au cours de la vie d'un spectacle.

Microdonnées : format d'encodage permettant d'intégrer des données structurées dans une page Web (documents HTML) afin de permettre son repérage et sa compréhension par les robots de recherche et les intelligences artificielles. Il y a trois principaux formats d'encodage pour les données structurées : microdonnées, RDFa, JSON-LD. Ces trois formats sont supportés par les moteurs de recherche et par Artsdata.

Œuvre / OEMI : abréviation pour référer au modèle de données de la Fédération internationale des associations et institutions de bibliothèques (IFLA) qui inclut les éléments Œuvre / Expression / Manifestation / Item (voir les définitions qui suivent). Ce modèle se prête plus ou moins bien à la description de la chaîne de valeur du spectacle vivant, notamment en raison du caractère complexe de l'œuvre « spectacle » (elle contient ou intègre des expressions de plusieurs autres œuvres) et de l'absence d'une manifestation destinée à être communiquée par le biais d'un support physique ou numérique. En raison de l'ambiguïté du terme « œuvre » tel que défini dans le modèle OEMI, nous lui avons préféré le terme « Spectacle ».

Présence : variable P du modèle LATICCE P/V/R. Dans le présent rapport, la présence est implicite car les tests ont été réalisés à partir d'un panier de référence constitué de l'offre même des diffuseurs et des sites étudiés.

Propriété : dans les données structurées, la propriété permet de décrire des relations entre des objets (c.-à-d. des entités du monde réel) ou des attributs de ceux-ci. Dans le triplet RDF (sujet prédicat objet), la propriété est un prédicat.

³ Voir <https://vitrinelinguistique.oqlf.gouv.qc.ca/fiche-gdt/fiche/2080594/marqueur>

Recommandation : variable R du modèle LATICCE P/V/R, c'est-à-dire le potentiel d'une information ou d'un contenu d'être recommandé dans l'environnement numérique. Il est possible de mesurer la qualité de la recommandation en recourant à de multiples variables secondaires. C'est l'objet même de l'exercice au fondement du présent rapport. Le LATICCE qui s'est employé à mesurer l'accès à la musique en flux numérique sur les plateformes de streaming a identifié les sous variables concordance, pertinence et nouveauté en relation avec le respect ou non du profil de goût d'internautes fictifs répondant à des persona. Dans le présent rapport, la qualité des recommandations est mesurée à l'aide d'un indice d'exposition sur les pages de résultats de recherche de Google.

Représentation : événement au cours duquel un spectacle est exécuté en public. Au moment de publier le rapport, une proposition était à l'étude pour intégrer ce concept à l'ontologie [Schema.org](https://schema.org).⁴

Spectacle : œuvre composée d'un ensemble d'éléments créatifs, incluant des conceptions, destinée à être exécutée en direct pour un public. Autrement dit, le spectacle réfère au contenu qui est présenté sur scène lors d'un événement des arts de la scène. Le terme spectacle est parfois employé par analogie dans l'usage courant pour désigner l'événement lui-même (c'est-à-dire la représentation de spectacle). Au moment de publier le rapport, une proposition était à l'étude pour intégrer ce concept à l'ontologie [Schema.org](https://schema.org).⁵

Tests A/B (incluant groupe témoin O) : procédure de test en vertu de laquelle on compare les résultats de performance de 2 objets à peu près identiques mais se distinguant par une légère variable que l'on contrôle.

Par exemple une représentation de spectacle à 19h ou une représentation du même spectacle à 21h. Laquelle sera la plus fréquentée ?

On peut faire des tests de découvrabilité A/B autour d'une offre pour le même spectacle vendu en ligne qui n'offre aucune donnée structurée (groupe témoin ou O), un deuxième (groupe A) qui présente une structuration minimale des données et un troisième (groupe B) qui inclut des métadonnées enrichies, afin de voir si une différence de résultats s'observe.

Visibilité : variable V du modèle LATICCE P/V/R. La visibilité est une variable de la découvrabilité qui repose sur la production et l'affichage de contenus éditoriaux par les acteurs humains avant que ne se répercute l'influence des algorithmes de recommandation. Incidemment une grande visibilité aura certainement un impact sur la recommandation automatique.

⁴ Voir la discussion « Proposal (Performing Arts 2) - New sub-class of Event: PerformingArtsEvent ». <https://github.com/schemaorg/schemaorg/issues/4468>

⁵ Voir la discussion "Proposal (Performing Arts 1) - New sub-class of CreativeWork: PerformanceWork". <https://github.com/schemaorg/schemaorg/issues/4467>

Contexte et objectif

Dans le cadre du projet Artsdata, l'Association canadienne des organismes artistiques (CAPACOA) a retenu les services des auteurs afin de développer des indicateurs qui quantifient la découvrabilité numérique des arts de la scène. Le présent rapport résume les travaux effectués lors d'une phase exploratoire s'étant déroulée de janvier à juillet 2024, puis d'une deuxième phase de moissonnage systématique de résultats dans le moteur de recherche Google, réalisée d'août 2024 à août 2025. La méthodologie de ces travaux repose sur la création de modèles qui utilisent diverses variables permettant une mesure reproductible et vérifiable de la découvrabilité en ligne des représentations de spectacles et manifestations des arts de la scène, notamment le recours aux métadonnées structurées du consortium [Schema.org](https://schema.org).

Deux brèves annexes se penchent quant à elles sur deux phénomènes corollaires associés à la problématique étudiée, soit le poids des sites de revente de billets dans les résultats de recherche et le rôle croissant des applications de recommandation alimentées par l'IA.

Le rapport formule enfin des conclusions et des recommandations permettant de stimuler le trafic sur les pages d'événements en arts de la scène.

Introduction et hypothèses

Au mois d'août 2025, la plus récente *Enquête québécoise sur les loisirs culturels et le divertissement* (EQLCD) produite par l'Institut de la statistique du Québec (Observatoire de la culture et des communications du Québec (OCCQ) 2025) et les deux articles publiés par Mario Girard dans La Presse sur le « bon vieux billet en carton »⁶ mettent la table pour un renouvellement de la conversation autour de la fréquentation du spectacle vivant à l'ère numérique.

De façon générale, la fréquentation de spectacles d'artistes locaux tendrait à baisser plus la population est jeune. Cela dit, il s'agit aussi d'un constat pour les autres produits culturels. Quant à Mario Girard, il pointe les environnements d'achat de billets en ligne comme dissuasifs pour la fréquentation des spectacles. Ces observations sommaires font néanmoins porter le regard vers le fait que de possibles perturbations des habitudes culturelles sont en cours.

Nous connaissons l'actuel engouement associé au sujet de la découvrabilité en ligne dans un contexte où l'abondance des sollicitations et produits offerts aux publics explosent. Qu'à cela ne tienne, le secteur des arts de la scène et du spectacle, contrairement à ceux de l'audiovisuel et de l'enregistrement sonore, a jusqu'ici peu fait l'objet de recherches (nonobstant les efforts de notre groupe de chercheurs : A10S et Artsdata 2024; gradient et Artsdata 2024; gradient et dia-log 2023; Estermann et Julien 2019; Julien et Petri 2017) ou de politiques publiques sur cette question. Nos travaux ont donc, et de nouveau, pour but de s'inscrire dans ces

⁶ La Presse, 11 et 19 août 2025 <https://www.lapresse.ca/arts/chroniques/2025-08-19/la-billetterie-et-vous.php>

dynamiques, de sonder si, et comment, les arts de la scène sont eux aussi affectés par le virage numérique des secteurs culturels.

Ainsi nos travaux se fondent sur l'hypothèse de départ inspirée de l'équation de mesure de la découvrabilité du LATICCE à l'effet que plus un événement est exposé (P/V/R : présent, visible, recommandé) (Rioux et al. 2019, 2021, p. 15) , plus il est susceptible d'être consommé par le public, c'est-à-dire d'être d'abord consulté, puis ultimement, de susciter un achat. Nous devons évaluer la mesure en vertu de laquelle les **facteurs de rayonnement** d'une offre culturelle influencent sa consommation.

Les lacunes de nos connaissances sur la découvrabilité des arts de la scène et en matière de recherche sur ce sujet sont de nature diverse. Sur le plan des habitudes des publics, nous devons mieux cerner les pratiques socio-construites associées par exemple, à la migration de l'identification des événements et l'achat de billets sur Internet (Roberge et al. 2020). Comment évoluent les pratiques des publics ? Nous avons tenu compte de ces aspects au meilleur de nos connaissances dans l'élaboration de nos paramètres méthodologiques lorsque vient le moment de lancer nos tests. Dans le but d'essayer de systématiser notre compréhension des paramètres construits, nous avons produit une typologie des variables en présence. Nous y reviendrons dans la section décrivant nos méthodes.

Cela dit, Artsdata a opté, pour le moment, pour une étude des pratiques de mise en marché des organismes artistiques en contexte numérique selon une approche fonctionnaliste et technique.

Les objectifs et questions de recherches initialement formulés pour motiver notre projet se sont fixés pour priorités de constater les résultats d'efforts de stimulation de la découvrabilité et de comparer les pratiques entre diffuseurs et acteurs de la chaîne de valeur des arts vivants. La production d'un état des lieux des manifestations mesurables de la découvrabilité et du rayonnement en constituait l'un des prérequis (CAPACOA et al. 2024).

Nous nous sommes intéressés aux recours aux meilleures pratiques d'indexation des contenus et d'optimisation de la recherche (SEO), notamment à l'utilisation des métadonnées descriptives enrichies, des balises de **données structurées** du référentiel [Schema.org](https://schema.org).

Méthodologie

Prémisse méthodologique

Notre prémisse méthodologique de départ fut de faire reposer nos observations sur Google (Search) comme «baromètre» de l'environnement numérique. Les moteurs de recherche créent des synthèses des données massives publiées en ligne tout en s'alignant le plus possible sur les attentes et intentions des consommateurs.

Nous avons mis en œuvre des travaux de moissonnage systématique et automatique puis des études de cas, le tout reposant sur les résultats de recherche dans le moteur de recherche Google. Les travaux ont reposé tantôt sur un vaste échantillon de requêtes représentatives des parcours de découverte des consommateurs, ensuite sur une sélection de 5 diffuseurs (phase 1), puis de trois diffuseurs lors de tests A/B/O de la phase 2 (soit le Burlington Performance Arts Center (BPAC) (ON), le Sanderson Centre de Brantford (ON) et le Centre des arts Juliette-Lassonde de Saint-Hyacinthe (QC)). Pour chacun des diffuseurs, une liste de représentations de spectacles a été circonscrite ainsi que des périodes temporelles de requêtage des variables identifiées.

Nous avons ultérieurement fait reposer nos analyses sur l'existence ou non d'une structuration des données fondées sur la typologie recommandée. Lors de la phase 1, les événements ont été classés en quintiles et les résultats comparés selon le nombre de pages vues et le volume de données structurées sur leurs pages d'événement. Ensuite, lors de la phase 2, nous avons également testé plusieurs modèles explicatifs, à travers trois scénarios de modélisation A/B/O, pour évaluer précisément l'impact de la structuration des données sur la découvrabilité des événements.

Voici en détail les méthodologies d'observation.

Création d'une typologie

Nos travaux ont reposé sur diverses méthodes quantitatives, celles-ci s'appuyant sur une typologie préalable de divers marqueurs ⁷, entités (éléments, sujet), attributs (propriétés, prédicat), valeurs (objet), indices, indicateurs et facteurs de rayonnement ^{8, 9}.

L'effort de production de notre typologie et des définitions associées s'échelonne sur plusieurs années et a débuté en amont de l'actuel projet de mobilisation des données structurées. Ce

⁷ Voir Figure 1.

⁸ Voir la liste des propriétés associées à la description d'événements proposées dans le modèle de données Artsdata. <https://culturecreates.github.io/artsdata-data-model/classes/event.fr.html>

⁹ Voir notre tableau des typologies  CAPACOA - Typologie des données

travail représente un effort majeur et la typologie produite n'est jamais totalement terminée. Les parties prenantes concernées par l'application des principes d'activation de la découvrabilité doivent notamment adopter une telle typologie de façon consensuelle et en maîtriser les tenants.

Le travail de production de la typologie CAPACOA/Artsdata devait à priori s'aligner sur des normes internationales. Or, contrairement à d'autres secteurs de l'économie de la culture où il existe des consortiums et lieux décisionnels en matière de pratiques normées ¹⁰, un tel espace pour les arts de la scène n'existait pas au démarrage de nos travaux. Par conséquent, nous avons eu à construire une concertation et identifier des lieux propices au débat autour d'une typologie et d'un socle minimal permettant de documenter les entités des arts vivants et les événements, en particulier. C'est ainsi que la communauté internationale regroupée autour du projet Wikidata Performing arts ¹¹ et W3C ¹² ont pu servir de caisse de résonance à nos décisions en matière de définition d'une typologie. CAPACOA a aussi pris la responsabilité de réunir un groupe de travail, le *LODEPA WG 6 Wikidata/Wikimedia*, portant sur les enjeux des métadonnées descriptives pour les arts de la scène. Des acteurs du Canada, du Royaume-Uni, de Suisse, de Belgique et de Croatie participent à ces conversations.

Il n'est pas évident d'intégrer pleinement la nuance entre les concepts associés aux définitions normées du numérique puisqu'ils se superposent avec ceux issus du cadriciel du monde pré-numérique qui parlait plutôt en termes de « stratégies marketing » et de « résultats de vente ». Voir à ce titre la figure 1 qui présente l'ensemble du continuum des marqueurs de rayonnement selon l'Observatoire de la culture et des communications du Québec (2021).

Nos efforts principaux en matière de production d'une typologie se sont donc concentrés sur l'identification des **propriétés** (attributs, prédicats) les plus susceptibles de permettre au spectacle vivant d'émerger dans Google Search. Quelles sont les propriétés qui définissent une représentation donnée ? Nos tests A/B sont notamment intervenus sur ces propriétés.

¹⁰ OCLC pour le livre, DDEX pour la musique enregistrée, EIDR pour l'audiovisuel.

¹¹ https://www.wikidata.org/wiki/Wikidata:WikiProject_Performing_arts

¹² <https://www.w3.org/community/pair-cg/> (groupe désormais inactif)

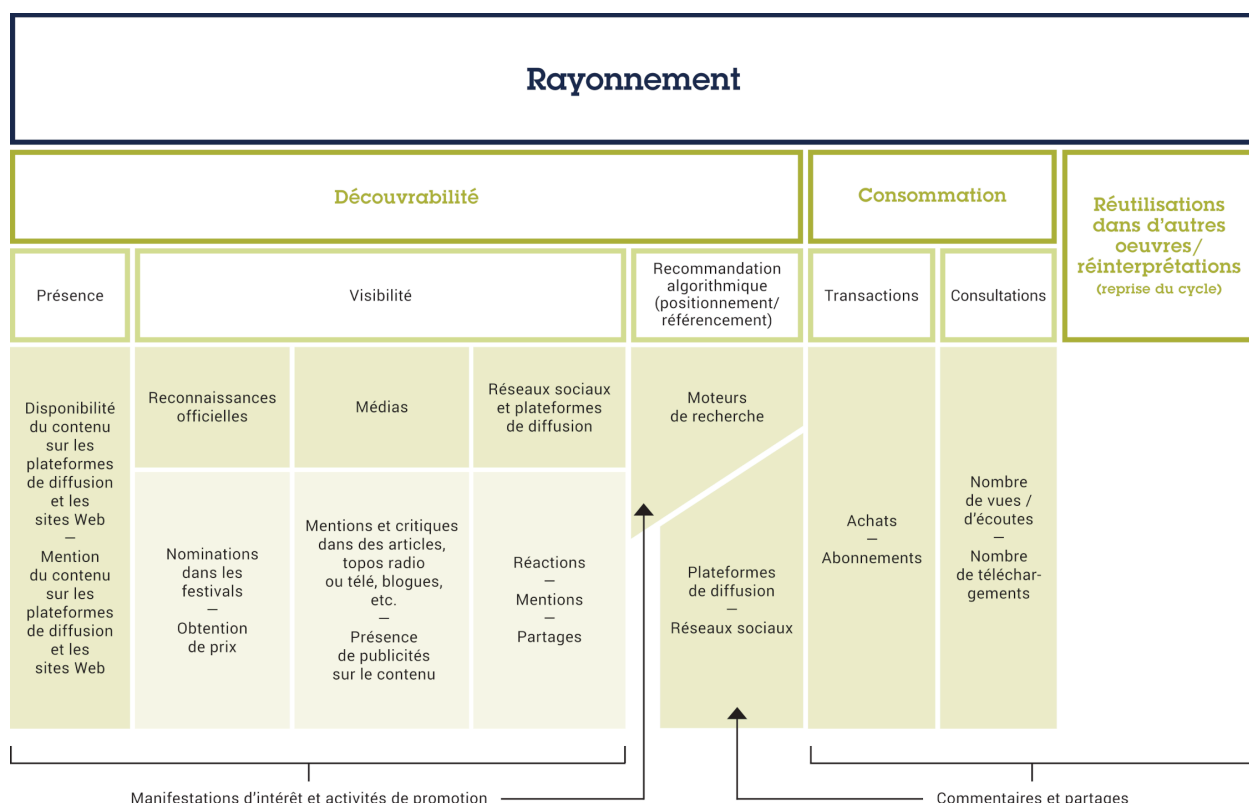


Figure 1. Les marqueurs de rayonnement selon l'Observatoire de la culture et des communications du Québec (OCCQ) et Danvoye (2021).

Algorithme de calcul d'un indice de données structurées

Une fois la typologie établie, il nous est apparu pertinent et réaliste de produire un premier indice fondé sur le seul calcul algorithmique de présence ou d'absence de certaines données structurées sur une page Web visée, tout en pondérant légèrement le poids de cette présence de données structurées selon leur importance pour la désambiguïsation et la recommandation. Cette pondération s'appuyait notamment sur le statut des propriétés associées à la description d'événements tel qu'établi par Artsdata/CAPACOA¹³ selon qu'elles sont jugées essentielles, recommandées ou facultatives. Les propriétés déclarées essentielles sont celles qui sont jugées nécessaires à la désambiguïsation des événements, c'est-à-dire qui permettent de distinguer un événement d'un autre événement qui lui ressemble. Par exemple, la date/heure et le lieu exact sont nécessaires pour détecter des potentiels doublons d'événements. Les propriétés recommandées sont des propriétés demandées par de nombreux consommateurs de données d'Artsdata, comme des calendriers culturels. Par exemple, l'image d'un événement

¹³ Voir la liste des propriétés associées à la classe Event et leur statut dans le modèle de données d'Artsdata. <https://culturecreates.github.io/artsdata-data-model/classes/event.fr.html>

n'est pas requise pour la désambiguïsation d'un événement mais elle est exigée par tous les calendriers culturels (et par les personnes qui les consultent).

Les chercheurs Frédéric Julien et Gregory Saumier-Finch ont conçu l'algorithme et l'ont soumis à l'équipe ainsi qu'aux membres de la Communauté d'intérêt Artsdata pour commentaires et ajustements.

Ainsi, l'indice de données structurées fournit une mesure quantitative représentant de l'exhaustivité des données structurées sur une page Web d'événement ¹⁴.

Panier de requêtes et de moissonnage des résultats Google Search

Étape 1 : État des lieux du référencement Web du spectacle

La première étape de la démarche, confiée à A10s, a consisté à faire un état des lieux ¹⁵ du référencement Web du spectacle vivant, avant de développer des approches de quantification de la découvrabilité.

La méthodologie s'appuyait sur un ensemble de requêtes représentatives des intentions des consommateurs des arts de la scène. L'équipe a d'abord généré un grand nombre de requêtes, puis les a filtrées pour s'assurer qu'elles soient équilibrées selon divers critères : linguistiques (français/anglais), géographiques (grands centres/régions), référencement dans Artsdata (entités présentes ou non), volume de recherche suffisant, et absence d'ambiguïté (éviter les termes qui renvoient aussi à d'autres contextes, comme un film ou un livre).

Au final, 168 requêtes ont été retenues et classées en types reflétant l'intention du consommateur. On distingue :

- Spectacle/artiste : recherche précise d'un spectacle ou d'un artiste nommé (ex. : Jesse Cooke).
- Recommandations FR, EN, bilingue : recherches plus générales de type « quoi faire à Montréal » ou « concerts Toronto ».
- Festival, diffuseur-producteur, salle : requêtes intermédiaires où le consommateur connaît une caractéristique (nom d'un festival ou d'une salle) sans chercher un spectacle précis.

Ces requêtes forment le panier d'analyse utilisé pour observer et comparer les résultats dans Google. Elles ont été exécutées sur les versions française et anglaise de Google Canada, puis les résultats ont été analysés pour dresser l'état des lieux, dont les principaux résultats sont présentés plus loin dans le présent rapport.

¹⁴ Voir le fichier README établissant les paramètres de l'algorithme de calcul de l'indice de données structurées : <https://github.com/culturecreates/artsdata-score>

¹⁵ Voir une version détaillée de l'état des lieux au

 État des lieux du référencement web des arts de la scène - Rapport

L'état des lieux a permis d'établir certaines balises pour le développement d'un indicateur de découvrabilité. Celui-ci devrait représenter les résultats concrets des efforts de référencement plutôt que les actions elles-mêmes, être mesurable dans la durée et rester suffisamment simple pour être calculé par les organisations du milieu sans expertise technique avancée. Parmi les pistes identifiées, on note l'importance de choisir une seule version de Google adaptée au marché cible, de compter la simple présence d'événements dans les enrichissements (plutôt que de se baser sur des formats changeants), de bâtir un panier de requêtes aligné à la fois sur l'offre et sur les intentions des consommateurs, et enfin de pondérer les résultats selon la source (site officiel de l'organisateur, agrégateurs, billetteries ou revendeurs), de manière à refléter la valeur relative de chaque type de visibilité.

L'étape 1 est décrite de façon détaillée dans le rapport [État des lieux du référencement Web des arts de la scène](#).

Étape 2 : Développement d'une approche de quantification de l'exposition

À partir des constats de l'état des lieux, nous avons élaboré deux stratégies de quantification de l'exposition.

La première permet d'analyser une page de résultats de recherche (SERP ou *search engine result page*) sur Google suite à une requête en particulier, en prenant la posture d'un diffuseur. Il s'agit d'attribuer des points à chaque composante de la page de résultats de recherche, et de catégoriser les points en deux groupes : ceux attribués au diffuseur, et ceux attribués à tous les autres acteurs. L'indice est obtenu en divisant le total des points attribués au diffuseur par le total de tous les points. La [procédure détaillée est documentée](#).

La deuxième approche s'appuie sur la première. Il s'agit d'identifier un certain nombre de requêtes pertinentes pour un diffuseur, de calculer l'indice obtenu avec la première approche pour chacune de ces requêtes, puis de calculer une moyenne pondérée des différents indices. Les facteurs de pondération tiennent compte de l'importance des requêtes dans le parcours de découverte. La [procédure de création du panier](#) ainsi que le [calcul de l'indice pondéré](#) font également l'objet de documentation détaillée.

Les deux approches ont été testées en les appliquant à cinq diffuseurs, à trois moments différents.

Les diffuseurs ciblés étaient :

- Le Théâtre Capitol à Moncton au Nouveau-Brunswick.
- Le Théâtre Capitole de Québec au Québec.
- Le FirstOntario Performing Arts Centre de St. Catharines en Ontario.
- La Place des Arts de Montréal au Québec.
- Le Sanderson Centre for the Performing Arts de Brantford en Ontario.

L'indice a été mesuré le 18 octobre 2024, le 6 décembre 2024 et le 24 janvier 2025 (donc trois fois, à intervalle de 7 semaines entre chaque mesure).

Études de cas, tests comparatifs A/B et groupe témoin O

Préambule

Voici les principaux apprentissages de notre phase exploratoire reposant sur des études de cas :

- Dans la manière de mesurer la découvrabilité, on distingue les « facteurs » de découvrabilité et les « marqueurs » de découvrabilité. Les marqueurs sont des manières de formuler, d'indiquer **qu'il y a eu** une forme de découvrabilité (davantage lié à un objectif, comme un clic sur une page, une exposition dans une page de résultat, même l'achat d'un billet). Les facteurs sont davantage tout **ce qui peut contribuer** à la découvrabilité selon les bonnes pratiques d'une stratégie de découvrabilité que l'on propose. La majorité des mesures de découvrabilité que nous connaissons sont de l'ordre du *facteur* de découvrabilité.
- Dans cette section, il est question de mettre en relief ces marqueurs et ces facteurs afin d'évaluer concrètement dans quelle mesure ces bonnes pratiques et, dans ce contexte, particulièrement les bonnes pratiques de structuration de données, influencent la découvrabilité des événements en ligne.

Ainsi, la question de recherche principale qui recouvre nos hypothèses est : Quel est l'impact de la structuration de données sur la « découvrabilité » des événements en ligne ?

Confection des groupes

Afin d'évaluer concrètement l'association entre la présence de certaines propriétés de données structurées d'événements et les marqueurs de découvrabilités, nous nous sommes intéressés à trois (3) scénarios :

- Groupe O. Aucune structuration des données.
- Groupe A. Structuration minimale des données, c'est à dire uniquement : les informations sur la date (`event:startDate`), lieu (`event:location`) et le nom (`event:name`) de l'événement.
- Groupe B. Tous les éléments du Groupe A (date, lieu, nom de l'événement) ainsi que les informations sur les artistes (`event:performer`) dont son attribut `sameAs` qui permet de lier l'artiste à des références externes. Nous avons choisi d'y mettre le lien Wikidata car c'est un identifiant passerelle : les entités décrites sur Wikidata y ont très souvent une multitude de liens vers d'autres identifiants pérennes ou vers des identifiants sur d'autres plateformes. En effet, il nous est apparu, et ce malgré le fait que cette

information est fréquemment négligée, que les métadonnées identifiant l'artiste principal sur scène (event : performer) fournissent des informations précieuses à propos du contenu du spectacle présenté et pourraient par conséquent avoir une incidence positive sur la découvrabilité des représentations par le biais des moteurs de recherche. Nous nous sommes ainsi assurés que cette information fasse partie des données structurées du groupe B.

Chacun des événements de notre échantillon a été assigné à un groupe et a donc été soumis au scénario associé.

Échantillonnage des événements et paramètres observés

Au total, nous comptons 157 événements ¹⁶, distribués à travers trois diffuseurs : Centre des Arts Juliette-Lassonde (CAJL) à St-Hyacinthe, QC ; Burlington Performing Arts Centre (BPAC) à Burlington, ON ; et Sanderson Centre (Sanderson) à Brantford, ON.

Ces derniers sont tous des diffuseurs pluridisciplinaires et certains gèrent plusieurs salles. Par conséquent, leurs programmations respectives sont très éclatées à la fois d'un point de vue disciplinaire, mais aussi d'un point de vue du public visé (niche vs. grand public), des efforts marketing qui sont apportés, etc.

L'indice d'exposition sur les moteurs de recherche et le trafic provenant des moteurs de recherche sont influencés par bien plus que la seule structuration de données.

Il est ressorti de la phase exploratoire qu'il **était absolument nécessaire d'isoler les effets d'autres variables « externes »**. En collaboration avec les diffuseurs, nous avons donc identifié et annoté certains paramètres que nous avons utilisés afin de procéder à une attribution aussi juste que possible de tous les événements. Les choix de catégorisation ont été faits par les diffuseurs. Puis, l'assignation des groupes s'est faite programmatiquement afin de maximiser l'équilibre entre les groupes O, A et B, et ce, pour tous les paramètres.

Voici les paramètres utilisés :

PARAMÈTRES	VALEURS PRISES
Type d'audience	Niche Grand public
Capacité de la salle	moins de 50 50 à 100 101 à 250 251 à 400 401 à 700

¹⁶ Voir  CAPACOA - Enquête sur les données structurées | Tous les événements

	700 à 1000 1001 à 1500 1501 à 2000 2000 à 5000 plus de 5000
Campagne considérable pendant la période d'observation	Oui Non
Nombre de jours avant l'événement	Nombre entier Différence en nombre de jours entre la date de la mise en ligne d'une programmation et la date de l'événement concerné
Période d'observation	1-4 : Période d'observation de 4 semaines, de 1 à 4 semaines après la date de mise en ligne de la programmation 5-8 : Période d'observation de 4 semaines, de 5 à 8 semaines après la date de mise en ligne de la programmation 9-12 : Période d'observation de 4 semaines, de 9 à 12 semaines après la date de mise en ligne de la programmation
Discipline	Théâtre Musique Chanson Cirque Famille Humour Danse

Tableau 1. Paramètres et valeurs utilisés par les tests A/B

Finalement, nous nous sommes assurés de distribuer les événements dans chaque groupe O/A/B tout en respectant l'équilibre de cette répartition selon les paramètres utilisés.

Méthodes de filtrage additionnel

Afin d'éviter l'effet de cas extrêmes, nous avons procédé à un filtrage additionnel après les premières analyses de résultats. Quelques événements se distinguaient grandement des autres dû à des circonstances uniques à ces événements (campagne de presse exceptionnelle sur toute une tournée, festival extérieur plutôt que représentation unique en salle) et ont donc été retirés.

Variables cibles : trafic et découvrabilité sur les moteurs de recherche

Dans le cadre des travaux, deux indices ont été produits par A10s et La Culture Créée, respectivement un indice de découvrabilité sur les moteurs de recherche ¹⁷ et un indice de structuration des données ¹⁸.

Afin d'évaluer l'impact de la structuration de données, nous avons étudié la relation entre l'indice de structuration de données et :

- l'indice d'exposition sur les moteurs de recherche (Google) de chaque événement.
- le nombre de vues sur chaque page d'événement provenant des moteurs de recherche (Google).

Pour ce faire, nous avons également utilisé un modèle explicatif, qui permet d'isoler l'impact des autres facteurs pouvant influencer ces deux variables.

Qu'est-ce qu'un modèle explicatif ?

Un modèle, c'est un outil qui met en relation des facteurs (ex. visibilité, contenu) avec un résultat (ex. trafic), pour comprendre ce qui influence quoi.

Un GLM (modèle linéaire généralisé) est un type de modèle qui permet de mesurer l'effet précis d'une variable tout en tenant compte des autres. C'est celui que nous avons utilisé pour analyser la relation entre l'indice de structuration des données et le trafic, et la découvrabilité sur les moteurs de recherche.

Il est utile parce qu'il montre clairement quels facteurs ont un vrai impact. Là où un test A/B se limite à une comparaison unique, un GLM révèle l'impact individuel de plusieurs variables à la fois.

Indice d'exposition

Cet indice a été établi selon le protocole développé par A10s et annoté par deux personnes au sein de Gradiant. Les requêtes ont été formulées en fonction du contexte géographique : nom de l'événement + « spectacle » pour les événements au Québec, et nom de l'événement + « show » pour ceux en Ontario. Les recherches ont été effectuées depuis les localités concernées (par exemple, depuis Saint-Hyacinthe pour le CAJL), afin de refléter les conditions réelles de consultation.

En complément, d'autres requêtes ont été générées à des fins d'analyses ultérieures si nécessaire, telles que des requêtes associées au diffuseur. Pour chacun des trois diffuseurs, deux à trois vagues de moissonnage ont été réalisées.

¹⁷ Voir la section Étape 2 de A10S - Panier de requêtes et de moissonnage des résultats Google Search

¹⁸ Voir la section Algorithme de calcul de l'indice de données structurées

Diffuseur	Date de mise en ligne de la programmation 2025-2026	Date du premier moissonnage	Date du deuxième moissonnage	Date du troisième moissonnage
Centre des Arts Juliette-Lassonde	30 avril	17 mai	7 juin	5 juillet
Burlington Performing Arts Centre	29 mai	12 juillet	2 août	30 août
Sanderson Centre	10 juin	19 juillet	9 août	6 septembre

Tableau 2. Dates de mise en ligne des programmations et dates des moissonnage pour chaque diffuseur

Nombre de pages vues

Le trafic a été mesuré à partir des données de Google Analytics, en isolant les sessions dont la source était Google (sessionSource) et dont la page d'entrée correspondait directement à la page de l'événement (landingPage). Seules ces visites ont été retenues. Ces données ont été extraites via l'API de Google Analytics.

Limites de la méthodologie et décisions quant aux analyses

L'indice de structuration des données a été calculé en tenant compte de l'ensemble des événements, mais n'a été considéré dans l'analyse que sur les semaines effectives. En effet, il y a eu des décalages entre la mise en ligne et la bonne structuration des données. Par exemple, pour le Sanderson Centre, les données n'ont été correctement publiées qu'à partir de la huitième semaine, après la correction d'une erreur générale d'indexation du site, indépendante de notre expérience. Ainsi, seules les semaines 8 à 12 ont été considérées comme valides pour le Sanderson Centre lorsque nous avons mis en relief la structuration de données et le nombre de pages vues.

Pour le CAJL, le principal problème concernait la structuration des données liées aux artistes (donc, les événements du groupe B) : le champ `performer.type` indiquait systématiquement « Person », même lorsqu'il s'agissait d'« Organization » (par exemple une compagnie de

danse). Cette erreur a été corrigée deux semaines après la mise en ligne, juste avant le premier moissonnage du 17 mai, par l'ajout manuel par l'équipe de CAPACOA de balises JSON-LD dans les pages concernées. Un autre enjeu touchait le fait qu'il existe plusieurs salles au CAJL et qu'elles n'étaient pas correctement distinguées : les salles étaient toutes nommées « Centre des arts Juliette-Lassonde de Saint-Hyacinthe » et l'identifiant pérenne qui aurait pu les désambiguïser avait été associé à l'objet « PostalAddress » plutôt qu'à l'objet « Place ». Ce problème a été réglé début juin et nous avons estimé son incidence potentielle relativement mineure, étant donné que le CAJL ne présente pas très souvent des spectacles en simultané dans ses deux salles. Toutes les semaines observées pour le CAJL ont donc été comptabilisées.

Enfin, une erreur d'assignation a empêché d'intégrer correctement les événements du BPAC, qui ont conservé leur niveau de structuration initial. Nous n'avons donc pas eu de groupe témoin pour le BPAC. Par conséquent, lorsque l'analyse portait sur les regroupements (A/B/O), seuls les événements du CAJL et du Sanderson Centre ont été retenus. Cela ne nous a toutefois pas empêché d'étudier les liens entre l'indice de structuration de données et la fréquentation du site Web. Les analyses portant sur l'indice de structuration de données, et non les groupes, incluent ainsi les trois diffuseurs.

Résultats principaux

Des observations préliminaires nous amènent à penser que les requêtes associées à des spectacles sont sujettes à une importante volatilité dans les résultats. En effet, certains facteurs qu'il est difficile d'inclure dans le protocole de sélection des requêtes ont un grand impact. Par exemple, l'un de ces facteurs est l'existence ou non de campagnes publicitaires classiques de plus ou moins grande envergure.

De façon générale, on peut émettre l'hypothèse que des tendances de fond plus fortes s'exprimeraient avec un plus grand nombre d'échantillons. Nos tests auraient ainsi intérêt à se dérouler de façon longitudinale, un mandat qui échoit en principe à un observatoire au mandat et au financement pérenne. Nous verrons aussi à la section « Résultats satellites » que l'émergence récente des robots conversationnels ajoute une nouvelle couche de complexité.

Nos travaux ont néanmoins permis de définir une typologie granulaire des facteurs de rayonnement qui devrait constituer un socle sur lequel faire reposer les recherches futures sur la découvrabilité des arts de la scène. Voir à ce titre notre section « Méthodologie ».

De façon générale, nos constats indiquent des corrélations positives entre la qualité et la quantité de données structurées présentes dans les pages Web des acteurs des arts de la scène et les statistiques de fréquentation en ligne et en salle.

Résultats de l'indice d'exposition

Étape 1 : État des lieux du référencement Web du spectacle

L'état des lieux confié à la firme A10S a permis de faire plusieurs observations sur la composition des pages de résultats de recherche en réponse à des requêtes associées à la découvrabilité du spectacle.

En effet, les résultats organiques issus des requêtes reflètent une diversité d'acteurs. Globalement, un tiers provient directement de la chaîne professionnelle des arts de la scène (lieux, producteurs, diffuseurs), un autre tiers de médias et d'agrégateurs (souvent liés au tourisme), et le reste de diverses plateformes. Toutefois, lorsqu'on regarde uniquement les premiers résultats, la chaîne professionnelle domine largement : plus des deux tiers des positions de tête leur appartiennent, signe que leurs sites sont bien structurés et crédibles aux yeux de Google. De plus, la répartition des résultats varie peu entre les différentes versions de Google (français/anglais, mobile/ordinateur).

En revanche, la composition des résultats change fortement selon le type de requête. Les réseaux sociaux apparaissent massivement pour les recherches liées à des artistes, diffuseurs ou festivals, alors qu'ils sont quasi absents pour les requêtes de recommandation, dominées

par les médias et les agrégateurs. On peut observer ces impacts graphiquement dans le graphique suivant :

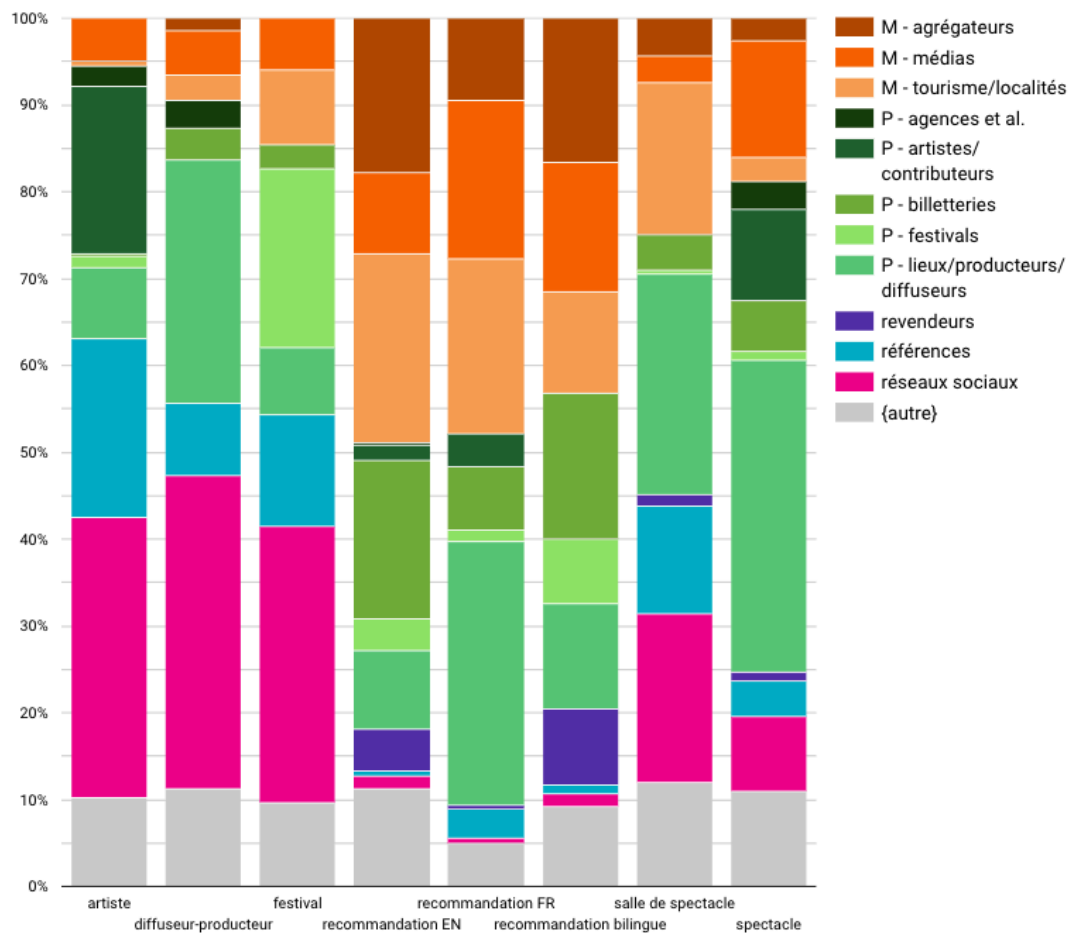


Figure 2. Comparaison de la répartition des catégories de sites sur les résultats organiques, pour toutes les requêtes, au moissonnage le plus récent, en fonction du type de requête.

Les différences linguistiques sont aussi notables : les billetteries et les revendeurs occupent davantage d'espace dans les requêtes en anglais ou bilingues, alors que les sites de diffuseurs/producteurs sont plus visibles en français. Enfin, parmi les domaines les plus présents, on retrouve surtout de grands acteurs internationaux comme Eventbrite, Wikipédia et Facebook, avec peu d'acteurs locaux, à l'exception de la Place des Arts et de Mtl.org, ce dernier se démarquant grâce à un référencement organique particulièrement optimisé.

Les résultats détaillés et des observations complémentaires sur les résultats enrichis sont présentés dans le rapport [État des lieux du référencement Web des arts de la scène](#).

Étape 2 : Développement d'une approche de quantification de l'exposition

Les résultats et l'analyse sont présentés sous forme visuelle [dans ce rapport](#). À titre d'exemple, le rapport permet de constater que l'indice varie dans le temps de façon différente pour chacun des diffuseurs :

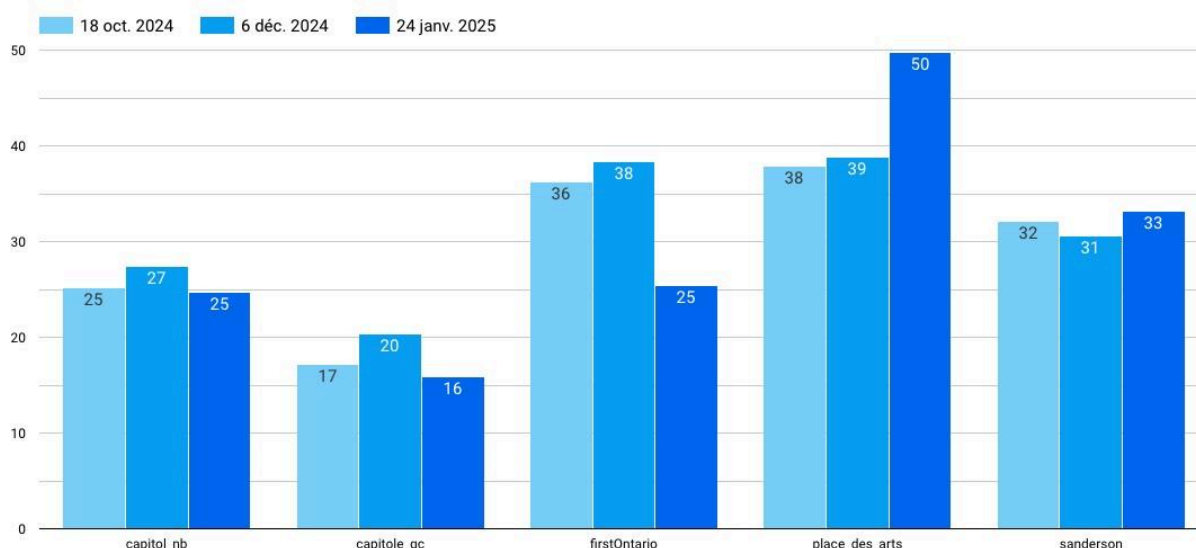


Figure 3. Indice d'exposition par diffuseur en vertu des chacune des dates de moissonnage.

L'analyse des résultats présente diverses difficultés. Certaines variations significatives de l'indice pour un diffuseur ne semblent pas nécessairement s'expliquer par un changement sur le site Web du diffuseur en question.

Des observations préliminaires nous laissent penser que les requêtes associées à des spectacles amènent de la volatilité dans les résultats. En effet, certains facteurs que l'on peut difficilement prendre en considération dans le protocole de sélection des requêtes ont un grand impact. À titre d'exemples, les bons résultats de la Place des Arts au dernier moissonnage s'expliquent par la sélection de deux spectacles (parmi trois) qui sont des exclusivités de ce diffuseur : la « concurrence » dans l'univers numérique était absente pour ces requêtes, l'indice global est tiré vers le haut. Pour cette raison, une version améliorée de l'indice pourrait éventuellement implanter un ou plusieurs des changements suivants :

- avoir des critères différents pour sélectionner les requêtes de spectacle ;
- augmenter le nombre de requêtes associées à des spectacles, de façon à réduire l'impact de spectacles individuels sur l'indice ;
- isoler les requêtes associées à des spectacles et faire un sous-indice distinct pour ce volet seulement (qui serait forcément plus volatile), puis un sous-indice sans tenir compte des spectacles (sous l'hypothèse que celui-ci serait moins volatile).

Par ailleurs, il faut souligner que nous disposons de données moissonnées à seulement trois reprises. On peut émettre l'hypothèse que des tendances de fond plus fortes s'exprimeraient avec un plus grand nombre d'échantillons.

Néanmoins, nous jugeons que la démarche démontre qu'il est possible de développer un indice d'exposition quantifié du point de vue d'un diffuseur de spectacles. L'indice peut servir à plusieurs fins.

Par exemple, une observation dans le temps de l'indice, pour un diffuseur donné, pourrait permettre de constater le résultat d'efforts délibérés, ou de changements imprévus dans l'environnement. En effet, lors de la refonte d'un site Web, l'indice peut servir d'outil pour s'assurer que l'exposition n'est pas affectée négativement par les changements, ou au contraire, pour confirmer que le nouveau site amène davantage d'exposition. Cette utilisation potentielle a été observée durant nos travaux, lorsque le Théâtre Capitol (de Moncton) a lancé un nouveau site Web.

L'indice pourrait également être utilisé pour observer l'évolution du rayonnement d'une cohorte donnée de diffuseurs. La méthode de pointage d'une capture, utilisée pour le calcul de l'indice, pourrait également être utilisée indépendamment pour réaliser des analyses (ex : tests A/B sur des efforts de découvrabilité sur un lot de spectacles).

Des résultats plus détaillés sont présentés dans le rapport [Élaboration d'un indicateur de rayonnement du spectacle sur les moteurs de recherche](#).

Résultats des test A/B et de l'analyse du trafic Web

Analyses descriptives

Les tests comparatifs A/B confiés à la firme Gradient nous ont permis d'observer le trafic provenant des moteurs de recherche sur plusieurs périodes. Un cycle d'observation quotidien laissait trop de place à des pics de trafic. Un cycle d'observation hebdomadaire ou sur des périodes de plusieurs semaines a permis de lisser les effets soudains pouvant survenir sur une base quotidienne. Afin d'aligner autant que possible nos résultats sur l'effet de la structuration de données sur le trafic avec ceux que nous tirerions de l'effet sur l'indice d'exposition, nous avons fait le choix d'observer des périodes de quatre semaines.

Découvrabilité sur les moteurs de recherche

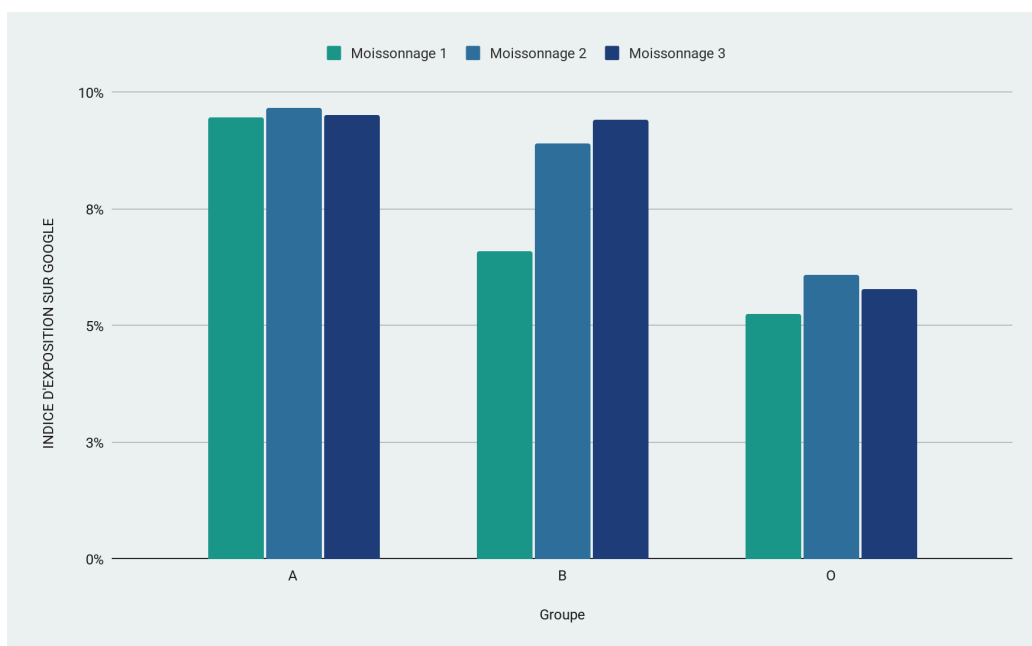


Figure 4. Indice d'exposition moyen par groupe (A - Métadonnées enrichies; B - Ajout des interprètes; O - Groupe témoin - Aucun balisage) pour le diffuseur Centre des Arts Juliette-Lassonde par date de moissonnage.

Les deux analyses tendent à confirmer nos hypothèses, mais de façon nuancée. Pour l'indice d'exposition, les groupes A et B affichent en moyenne de meilleurs résultats que le groupe témoin à tous les moissonnages. Dans le cas du groupe B, la hausse observée à partir du deuxième moissonnage correspond vraisemblablement au moment où les microdonnées du CAJL ont été correctement déployées. Voir à ce titre notre section « Limites de la méthodologie ».

Trafic depuis les moteurs de recherche

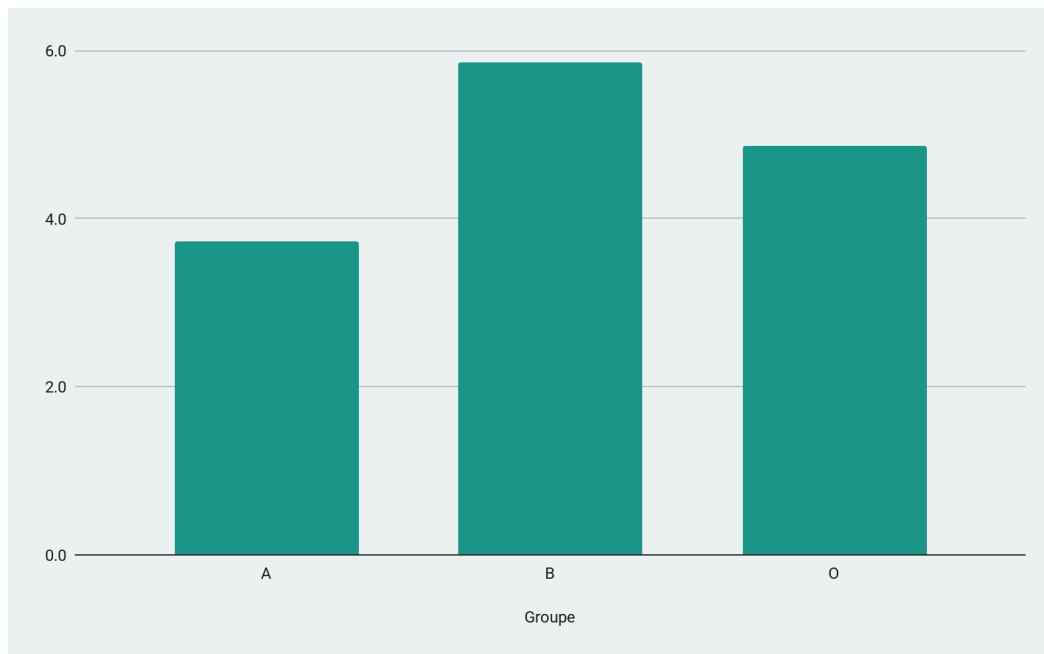


Figure 5. Vues moyennes sur quatre semaines effectives pour les diffuseurs Centre des arts Juliette-Lassonde et Sanderson Centre pour les groupes A - Métadonnées enrichies; B - Ajout des interprètes; O - Groupe témoin - Aucun balisage

Les données de fréquentation montrent toutefois une dynamique différente : si le groupe B enregistre un nombre moyen de pages vues supérieur à celui du groupe témoin, les événements du groupe A obtiennent des résultats inférieurs. Cette situation pourrait être liée à un biais d'échantillonnage dans la composition des groupes, mais elle rappelle aussi que **l'exposition mesurée ne se traduit pas automatiquement en trafic**.

Ces écarts s'expliquent en partie par les limites méthodologiques de notre indicateur. L'indice d'exposition ne repose que sur une requête par événement et correspond à un ratio sensible au volume total de résultats présents dans une page. Ainsi, un même nombre de résultats liés au diffuseur peut générer un indice plus faible si la page contient davantage de liens. De plus, le trafic observé peut provenir d'autres requêtes que celles retenues, ce qui introduit un décalage entre exposition mesurée et trafic réel.

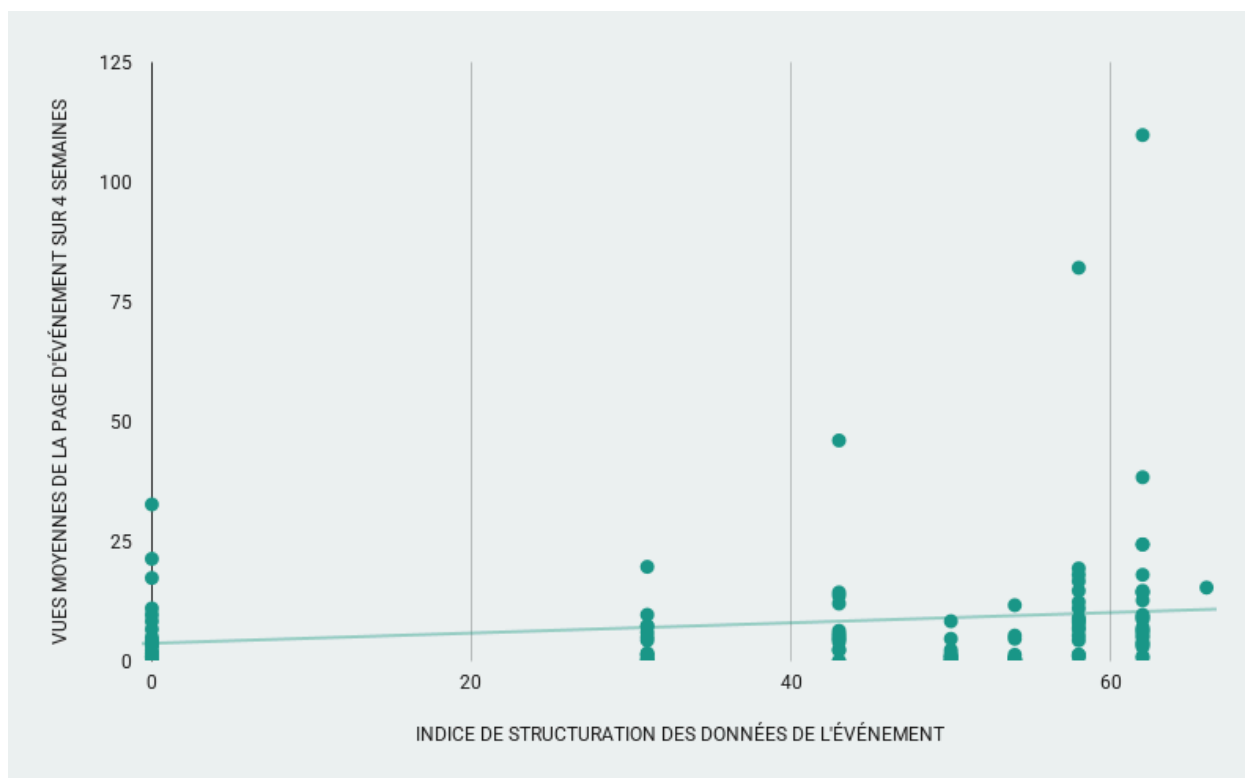


Figure 6. Relation entre l'indice de structuration de données et les vues moyennes sur 4 semaines. Note de lecture : Chaque point correspond à un événement observé sur une période de 4 semaines. Par exemple, le point situé à l'abscisse 58 et à l'ordonnée 82 signifie que l'événement observé a un indice de structuration de données de 58 et a reçu en moyenne 82 vues sur une période de 4 semaines.

Effet des données structurées sur le trafic des pages d'événements

Performance du modèle

Le modèle n'a pas pour objectif de prédire de façon exacte le nombre de vues d'une page Web. Il s'agit plutôt d'un modèle explicatif, qui nous permet de comprendre les tendances générales. Sur une période de quatre semaines, l'erreur moyenne d'estimation est d'environ ± 5 vues par page. Il est important de préciser que la fréquentation des pages d'événements en provenance de Google reste globalement faible. En effet, les moteurs de recherche déploient des efforts considérables pour fournir des réponses à même la page de résultats et cette tendance vers les réponses « zéro clic » a transpara dans nos données.

L'intérêt principal du modèle est donc de mettre en évidence les relations entre certaines données structurées (par exemple, les données décrivant le contenu du spectacle) et la fréquentation issue des moteurs de recherche. Les résultats confirment une corrélation positive et une tendance encourageante, ce qui va dans le sens des hypothèses de notre recherche.

Un autre point important est que la fréquentation des pages Web ne suit pas une distribution normale (voir Figure 7). Elle s'apparente plutôt à une distribution en « longue traîne », où quelques pages attirent beaucoup de visites tandis que la majorité en reçoivent très peu. Pour tenir compte de cette particularité, nous avons utilisé une fonction de distribution adaptée : la loi de Tweedie avec un paramètre log de 1,5.

Enfin, après plusieurs étapes de sélection, les variables explicatives retenues dans le modèle sont les suivantes :

- Discipline ;
- Type d'audience ;
- Campagne considérable dans la fenêtre temporelle observée ;
- Capacité de la salle ;
- Nombre de jours avant l'événement ;
- Groupe ;
- Indice de données structurées
- Diffuseur ;
- Période (période de 4 semaines, pouvant être 1-4, 5-8 ou 9-12).

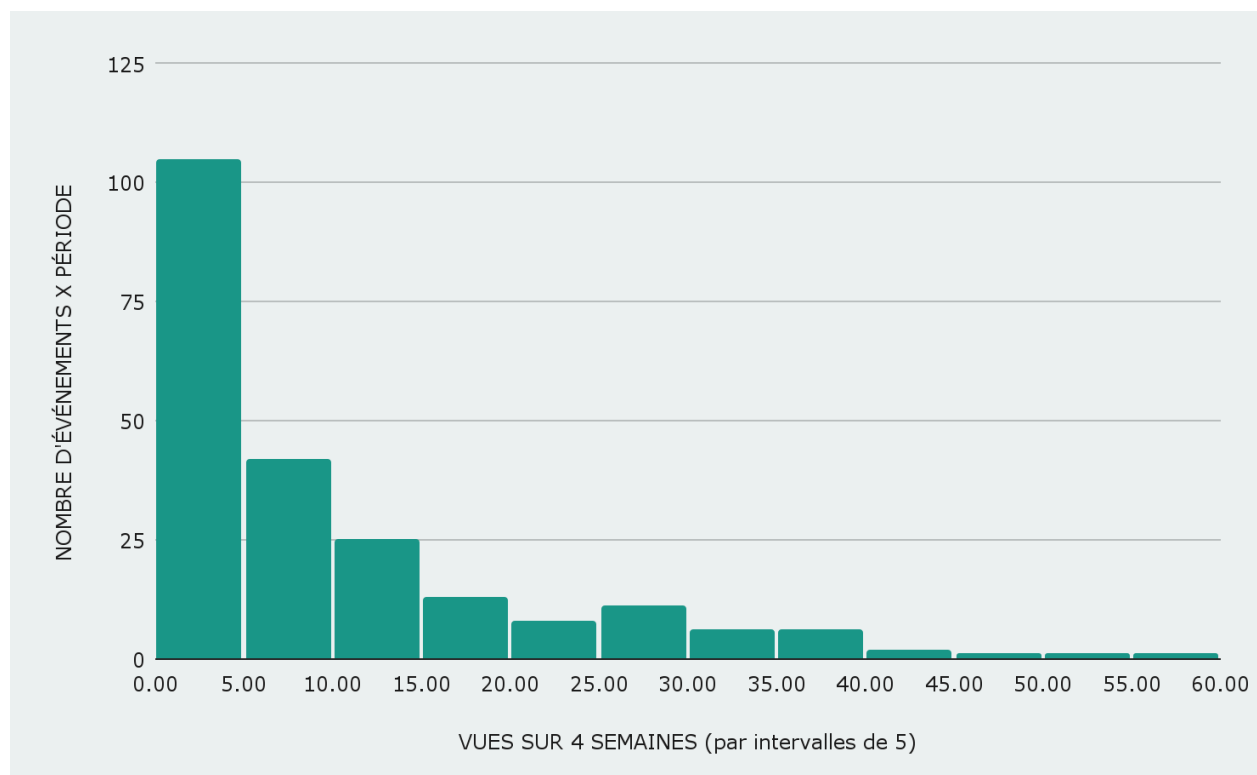


Figure 7. Histogramme du nombre de pages vues par période de 4 semaines (hors valeurs extrêmes). Note de lecture : Par exemple, la première barre indique que 105 événements ont attiré peu de visites (entre 0 et 5) pendant une période de 4 semaines. À l'opposé, un événement a reçu plus de 55 vues.

Résultats observés

Notre analyse confirme que la qualité de la structuration des données (l'indice de données structurées) a un effet direct et mesurable sur la visibilité des événements. En moyenne, **chaque amélioration de 10 points de l'indice de structuration de données entraîne une hausse d'environ +3 % des pages vues en provenance de Google sur quatre semaines** (voir la figure 6). Cet effet est encore plus marqué dans certaines situations : il peut quasiment doubler entre la 9^e et la 12^e semaine suivant la mise en ligne (+6,3 %), et reste particulièrement favorable pour la musique de niche (+2,6 %) ou les concerts en salles moyennes de 500 à 1000 places (+2,3 %). Les campagnes promotionnelles, lorsqu'elles sont soutenues par une bonne structuration des données, amplifient également cet effet (+2,5 à +2,9 %).

En guise de référence et d'exemples, les données importantes pour la désambiguïsation des lieux (`location.address.postalcode` et `location.sameAs`) ont une valeur de 4 points sur l'indice de structuration de données. Les données recommandées (ex., `description`, `URL de l'image`, `performer`, `offers.url`) ont une valeur de 2 points. Enfin, les données optionnelles (ex., `duration`, `endDate`, `eventAttendanceMode`) ont une valeur de 1 point. Ainsi, un événement n'ayant que les données minimales obtiendrait un pointage de 28, alors qu'un événement ayant aussi les propriétés de désambiguïsation et les propriétés recommandées aurait un pointage de 62.¹⁹

Cependant, l'impact varie selon le contexte. L'effet de l'indice est moins marqué, voire légèrement négatif, dans certains cas : lorsque la page d'un événement est mise en ligne 70 jours ou plus mois avant sa tenue (-3 %), pour les représentations de théâtre (-2 % à -3 %) dans certaines périodes ou salles, pour la danse (-1 % environ) ou encore pour les événements familiaux dans de petites salles.

Certaines propriétés spécifiques ont-elles un impact positif ?

Une question importante était de savoir si certaines propriétés particulières des événements pouvaient avoir un impact positif sur la fréquentation des pages. Ces propriétés n'ont pas été intégrées directement dans notre modèle principal. Toutefois, nous les avons examinées séparément, en procédant à des comparaisons par groupe, en utilisant notamment les données du CAJL et du Sanderson Centre.

Comme rappelé plus haut, le nombre de pages vues suit une distribution en « longue traîne » et non une distribution normale. Dans ce contexte, une simple comparaison de moyennes n'est pas appropriée. Nous avons donc utilisé un test statistique de Wilcoxon, qui compare les médianes plutôt que les moyennes, et qui est mieux adapté à ce type de données.

¹⁹ Pour le détail complet des propriétés de données structurées et de leur pointage, consultez le fichier README de l'indice de données structurées : <https://github.com/culturecreates/artsdata-score>

Les résultats montrent que la différence est généralement positive pour les événements dont les données sont structurées. Dans certains cas, l'effet observé reste marginalement significatif. En revanche, un point marquant est ressorti : la présence d'une description a un effet statistiquement significatif sur la fréquentation, confirmant son rôle central dans l'attractivité des pages pour Google Search.

Présence des propriétés minimales	Moyenne de la fréquentation sur 4 semaines	Médiane de la fréquentation sur 4 semaines
<i>Non</i>	5	3
<i>Oui</i>	9	5

Présence de la propriété organizer	Moyenne de la fréquentation sur 4 semaines	Médiane de la fréquentation sur 4 semaines
<i>Non</i>	4	2
<i>Oui</i>	12	7

Présence de la propriété description	Moyenne de la fréquentation sur 4 semaines	Médiane de la fréquentation sur 4 semaines
<i>Non</i>	4	2
<i>Oui</i>	12	7

Présence de la propriété performer avec lien sameAs	Moyenne de la fréquentation sur 4 semaines	Médiane de la fréquentation sur 4 semaines
<i>Non</i>	7	4
<i>Oui</i>	11	6

Étude de deux événements comparables

À propos des événements comparés

Afin d'exemplifier nos résultats globaux, nous avons procédé à une analyse plus ciblée pour certains événements que nous avons jugé comparables. Compte tenu des résultats obtenus qui montraient que l'écart était particulièrement notable pour les spectacles de musique

(concerts), le choix s'est porté sur la comparaison de deux événements de musique. Les deux événements ont lieu au Sanderson Centre, avec 1 mois seulement d'écart (4 novembre vs 4 décembre) mettant en scène tous deux des artistes country canadiens, James Barker Band et Dallas Smith, qui se situent dans une zone de notoriété comparable, avec respectivement environ 400 000 et 300 000 abonnés sur Spotify en date de la publication du rapport.

Leur proximité musicale et leur audience similaire en font de bons candidats pour une comparaison contrôlée : James Barker Band sert de représentant du groupe témoin (O), tandis que Dallas Smith représente le groupe A.

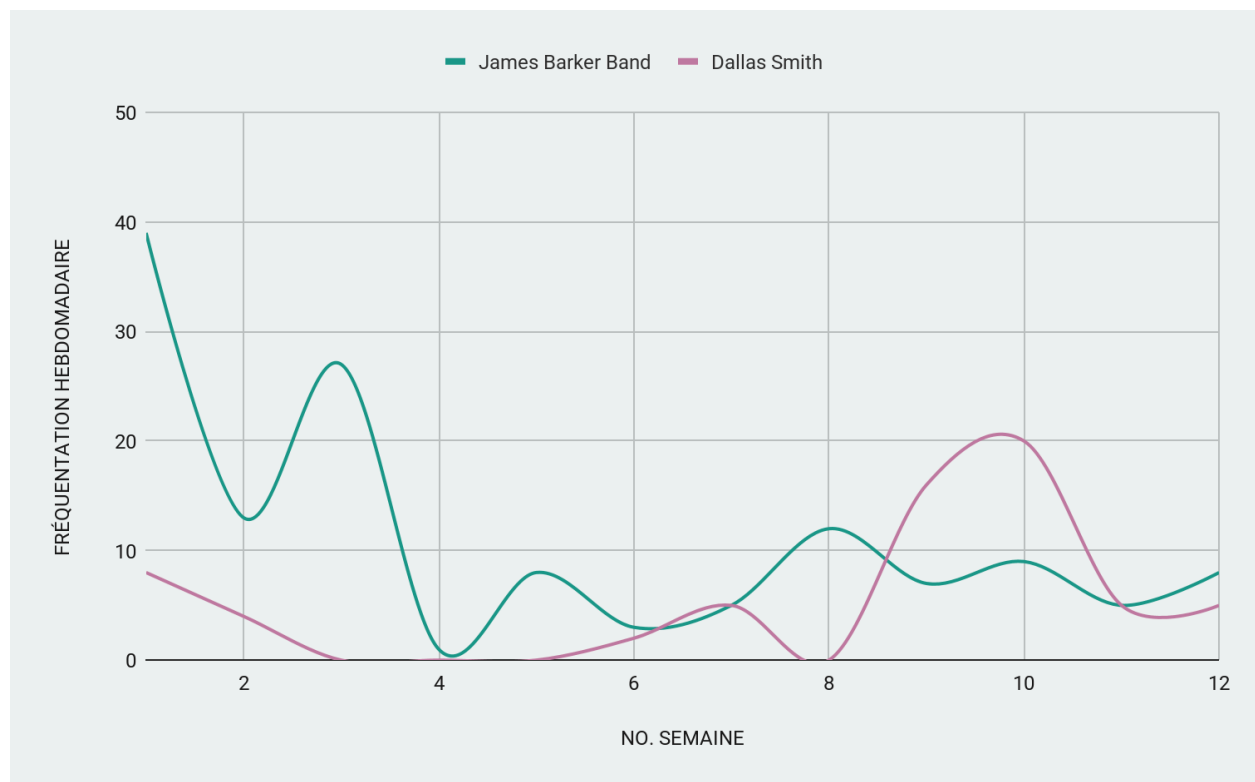


Figure 8. Diagramme du nombre de pages vues par semaine pour deux événements : l'un avec des données structurées pendant toute la période de 12 semaines (James Barker Band), l'autre avec des données structurées à partir de la 8^{ème} semaine (Dallas Smith).

Résultats observés

Ce qui est particulièrement intéressant avec le Sanderson Centre, c'est que nous pouvons observer les fluctuations du nombre de pages vues avant et après la mise en ligne des données structurées. En effet, l'un des deux événements avait des données structurées dès le première semaine, alors que, pour l'autre, les données structurées n'ont été effectives qu'à partir de la huitième semaine.

Pour James Barker Band, malgré un grand nombre de pages vues dès la mise en ligne, celui-ci s'estompe et l'on constate une forte baisse entre les périodes 1-4 et 5-8, suivie d'une stabilisation en période 9-12. Globalement, la tendance est à la diminution, sans signe de reprise significative.

La dynamique est différente pour Dallas Smith : après des niveaux bas, un bond survient en période 9-12, atteignant 46 pages vues.

La période 9-12 est la première où les données structurées ont été effectives, ce qui permet une lecture plus fiable des écarts observés. Le fait que le pic de Dallas Smith survienne précisément à ce moment donne du poids à notre hypothèse.

Résultats satellites

Sites de revente de billets

L'état des lieux du référencement Web des arts de la scène a permis de constater l'espace occupé par les sites de revente de billets. De façon générale, sur l'ensemble des résultats analysés pour toutes les requêtes, ces sites représentent en moyenne 1,2% des résultats.

On observe toutefois d'assez grandes variations en fonction du type et de la langue des requêtes. En effet, pour les requêtes associées à la recommandation (donc, qui ne mentionnent pas un artiste, un spectacle ou une salle en particulier) et qui sont en anglais ou avec des termes bilingues, les sites de revente occupent 2,6% de l'espace, soit plus du double de la moyenne. Contrairement à ce qui précède, le site [Billets.ca](#), qui représente de 50 à 70% de l'espace occupé par des revendeurs, semble suivre une logique différente en ce sens qu'il performe nettement mieux que les autres dans les recherches axées sur des spectacles précis.

Nous n'avons pas analysé ce qui explique les différences basées sur la langue. Nous pouvons toutefois émettre quelques hypothèses :

- Le marché de la revente est occupé par des acteurs étrangers qui sont mieux référencés en anglais qu'en français.
- Le marché local francophone est moins intéressant, à cause de sa taille, pour les revendeurs de billets, qui y font en conséquent moins d'efforts.
- Les diffuseurs québécois et leurs plateformes de billetterie utilisent davantage de données structurées, ou d'autres stratégies de référencement, et offrent donc une meilleure « résistance » à l'occupation de l'espace par les revendeurs.

Agents conversationnels

Tel que nous le soulignons à la section « Méthodologie », notre prémisse de départ fut de faire reposer nos observations sur Google (Search) comme « baromètre » de l’environnement numérique. Il en va de même quant à nos hypothèses initiales : les *moteurs de recherches créent des synthèses des données massives publiées en ligne tout en s’alignant le plus possible sur les attentes et intentions des consommateurs*.

Or, les robots et agents conversationnels créés sur les avancées de l’intelligence artificielle, les grands modèles de langage, l’apprentissage profond et machine reposent sur les mêmes postulats : produire une synthèse fondée sur l’analyse de données massives afin de répondre à une question.

En fin de parcours de recherche, il nous est donc apparu nécessaire de fouiller l’incidence possible de ces nouveaux outils sur nos hypothèses et questions de départ, de réfléchir à l’impact de l’IA sur les habitudes de recherche d’événements par les internautes.

Déjà des études tendent à démontrer que l’introduction de l’IA générative dans les pages de résultats a un impact notable sur les comportements en matière de recherche, entre autres en réduisant le volume des clics extrants vers les sites des organisations culturelles.

Les sommaires générés par IA apparaissent en haut, à la droite des premiers résultats de recherche. Ils puisent leurs réponses auprès de multiples sources et répondent instantanément à la requête de l’internaute. Ils évitent souvent aux gens de devoir naviguer vers des sites tiers (...) ²⁰

Nous avons effectué trois tests sommaires dans quatre applications IA – Mistral, Gemini, Leo-Brave et ChatGPT – et noté nos observations dans le document complémentaire suivant :

 Robots conversationnels et Scénarios utilisateurs | Tests sommaires 2025 .

Nos questions ont cherché à évaluer la qualité des réponses à trois requêtes :

1. SVP Indique-moi les spectacles à Montréal entre les 6 et 13 juin 2025.
2. Puis-je voir une pièce de théâtre avec Geneviève Brouillette cet automne au Québec ?
(Nous cherchions à nous faire recommander la pièce *Peut contenir des traces d’égo* au Centre des arts Juliette-Lassonde, le 25 octobre 2025, un événement faisant partie de notre panier de requêtes.)
3. L’IA utilise-t-elle les données structurées Schema.org ?

Nos premiers résultats montrent que les performances de l’IA sont très différentes selon que l’application réalise des recherches en ligne en temps réel ou se fonde seulement sur des données massives d’entraînement initiales.

²⁰ Voir Downs et Unitt 2025, une enquête portant sur 100 sites culturels du Royaume-Uni. Traduction des auteurs.

Les premières, ChatGPT et Gemini sont très performantes, les secondes, Mistral et Brave beaucoup moins. Les secondes ne font pas usage des balises Schema alors que les premières y auraient systématiquement recours et de façon documentée.

Un autre constat est fort intéressant, les applications effectuant des recherches en ligne en temps réel sont désormais en mesure de nous partager leurs sources.

À ce titre, les sites officiels des diffuseurs, les sites régionaux d'information culturelles et touristiques, les médias généralistes et le site de recommandation de sorties La Vitrine sont tous mis à contribution dans les résultats. Les sites des billetteries semblent jouer un rôle marginal. La répartition du poids quant aux sources moissonnées pourrait faire l'objet d'efforts et de plus amples recherches.

Lorsque nous avons posé la question « l'IA utilise-t-elle les données structurées Schema.org ? », la réponse de Google AI Pro a confirmé les intuitions de départ et les résultats de l'actuel projet :

Oui, les systèmes d'IA de Google, y compris les modèles de langage comme Gemini (auquel Google AI Pro se réfère), utilisent les données structurées Schema.org. Ces données jouent un rôle crucial dans la manière dont les modèles d'IA comprennent et traitent l'information du web.

Oui, les systèmes d'IA de Google utilisent la propriété `performer` de Schema.org pour identifier les artistes participant à un événement. Cette balise fait partie des données structurées de type `Event`, qui sont essentielles pour aider Google à comprendre le contenu de votre page Web et à le présenter de manière enrichie dans ses résultats de recherche. ²¹

ChatGPT nous offre même un exemple de code de balisage à utiliser pour un artiste participant à un événement (voir Annexe).

De façon générale, les tests permettent de penser que la recherche d'événement en ligne pourrait rapidement muter vers l'usage des agents conversationnels mais aussi que les données enrichies et structurées y sont déjà mises à profit :

1. Optimise your content for AI Overviews (...) Structured data: Use schema markup, a technical way to describe the content of your website to Google. For cultural organisations, schema for events, exhibitions, artworks, and historical information is especially relevant. (Downs et Unitt 2025, p.14)

Néanmoins, deux questions importantes demeurent : Devons-nous accepter les réponses des robots à nos requêtes sans effectuer de recherche indépendante ? Et les réponses à ces

²¹ Requêtes exécutées le 8 septembre 2025. Voir le document complémentaire

 Robots conversationnels et Scénarios utilisateurs | Tests sommaires 2025

requêtes peuvent-elles contenir des biais que seule une recherche indépendante pourrait repérer ?

Activités de communication

Nos conclusions générales et résultats sont consignés dans un diaporama synthétique qui a fait l'objet d'une présentation lors de la Matinée Numérique du 18 septembre 2025^{22, 23}.

Une deuxième activité de communication s'est tenue le 15 octobre à l'initiative de CAPACOA²⁴.



Second chances, du Groupe RUBBERBANDance (Q16640268, K10-122). La représentation de ce spectacle au Centre des arts Juliette-Lassonde faisait partie du groupe B dans les tests A/B. Crédit photo : David Wong.

²² Voir https://passerelles.encommun.io/c/les-matinees-numeriques/m?m=14842#m_14842

²³ Voir le diaporama [Mesure de la découvrabilité | Matinées numériques 20250918](#)

²⁴ Voir <https://www.artsdata.ca/fr/evenements/les-cles-pour-dynamiser-la-decouvrabilite-des-evenements-15-oct>

Recommandations et ressources

À la lumière des constats de cette étude, dans l'optique de favoriser la découvrabilité des événements des arts de la scène, les parties prenantes qui présentent des représentations de spectacles ou qui y contribuent en tant qu'artiste principal ou organismes de production auraient intérêt à :

- Baliser leurs pages d'événements en recourant aux données structurées Schema.org et s'assurer d'en donner l'accès aux agents d'indexation des moteurs de recherche et des agents conversationnels ;
- Veiller à ce que ces données contiennent minimalement toutes les données permettant de distinguer des événements similaires – le nom (name), la date / heure (startDate) et le lieu précis de l'événement (location, en renseignant l'adresse complète et un identifiant pérenne sameAs ²⁵) ;
- Accorder une attention particulière aux métadonnées décrivant le contenu de l'événement – en particulier la description (description) ;
- Bien décrire et identifier, à l'aide d'identifiants pérennes (sameAs), les principaux contributeurs de l'événement – le diffuseur (organizer) et l'artiste sur scène (performer).

Au demeurant, l'intégration de données structurées enrichies à un site Web ne répond pas qu'aux seuls impératifs de la découvrabilité via les moteurs de recherches ou les agents conversationnels. Des données structurées de qualité peuvent notamment être moissonnées et réutilisées sur d'autres sites, comme les calendriers culturels alimentés par des données ouvertes ²⁶. Cela peut alléger le fardeau de saisie de données pour les gestionnaires culturels, qui doivent autrement publier les mêmes informations sur de nombreux sites. Les données structurées peuvent par ailleurs être moissonnées, puis préservées pour des fins d'archives – un service qui est offert par le graphe de connaissances Artsdata. Faute d'un tel archivage, les informations à propos des représentations de spectacles disparaissent souvent avec les pages Web où elles sont affichées, privant ainsi les grands modèles de langage de sources précieuses d'informations pour développer une connaissance du domaine des arts de la scène.

Bien que cette étude n'ait porté que sur les données événementielles, dans la mesure où les données décrivant le contenu et les artistes ont une incidence sur la découvrabilité des événements, les artistes et les compagnies artistiques auraient sans doute avantage à déployer leurs propres données structurées – les décrivant et décrivant leurs œuvres – sur leur site Web. De cette façon, ils affirment leur autorité sur les données les concernant et exercent un contrôle

²⁵ Pour les bonnes pratiques concernant la propriété location, consultez <https://culturecreates.github.io/artsdata-data-model/location.fr.html>

²⁶ À titre d'exemple, le calendrier culturel du site ArtsCultureNB est alimenté par des données ouvertes provenant d'Artsdata. <https://artsculturenb.ca/fr/arts-de-la-scene/>

sur celles-ci. Cela pourrait de plus aider les robots d'indexation à mettre en relation les données événementielles avec celles sur les créateurs et leurs œuvres.

Enfin, les tests A/B de cette étude permettent de croire que les identifiants pérennes de Wikidata pour le lieu (`location.sameAs`) et l'artiste sur scène (`performer.sameAs`) contribuent à la découverte du spectacle vivant. Cependant, l'édition d'éléments Wikidata requiert des connaissances pointues qui sont encore moins faciles d'accès que les connaissances à propos des données structurées Schema. Ce défi peut être surmonté de deux façons : en recourant aux services de consultants spécialisés dans les données ouvertes ou en adhérant à une association qui offre à ses membres la possibilité de publier leurs informations dans Wikidata. De fait, un grand nombre d'associations au Canada ont déjà réalisé des versements ponctuels de données dans Wikidata et Artsdata ; certaines ont même établi des processus de publication en continu.²⁷

Afin d'aider les organismes artistiques dans la mise en œuvre de ces recommandations, l'équipe d'Artsdata a publié plusieurs **ressources** pour aider les gestionnaires culturels à intégrer du balisage de données structurées dans leurs page Web, notamment :

- des gabarits de données structurées incluant toutes les métadonnées d'enrichissement identifiées dans ce rapport (et plus encore);
- des pages de documentation détaillant comment bien renseigner chaque propriété ;
- des tutoriels vidéo illustrant comment récupérer des identifiants pérennes ;
- une banque d'identifiants pérennes à propos de plus de 10 000 personnes, organismes et lieux des arts de la scène au Canada²⁸;
- des services de découvrabilité numériques sur demande.

Les lecteurs souhaitant en apprendre davantage peuvent se référer à la page d'[introduction aux données structurées d'événements](#)²⁹ ou à celle d'[introduction aux identifiants pérennes](#)³⁰.

²⁷ Voir *Artsdata 2023-2024 Annual Report*. Accessible en ligne à <https://www.artsdata.ca/fr/nouvelles/une-constellation-de-donnees>.

²⁸ Voir à titre d'exemple, les artistes et leurs identifiants pérennes présents dans le graphe Artsdata <https://s.zazuko.com/36BcdkR>

²⁹ Voir <https://www.artsdata.ca/fr/ressources/donnees-structurees>

³⁰ Voir <https://www.artsdata.ca/fr/ressources/bien-identifie>

Conclusion

Selon les résultats de cette étude, il s'avère que certaines données structurées, telles que représentées sur les pages de résultats de Google, sont en phase avec les attentes des internautes lorsque vient le temps d'effectuer des recherches d'événements des arts de la scène à fréquenter. En ce sens, elles seraient donc motrices de découvrabilité.

Soulignons de nouveau l'intérêt à accorder à la description de l'événement ou encore à l'identification de l'organisateur et des artistes sur scène.

Ces recommandations valent pour la découvrabilité des événements via Google Search en date de 2025. Sur la base des connaissances actuelles à propos du référencement génératif (GEO ou Generative Engine Optimization : techniques d'optimisation de contenus pour les moteurs de recherche génératifs tels que Copilot et Gemini) et de nos tests avec des agents conversationnels, il apparaît par ailleurs que ces recommandations sont aussi pertinentes pour les grands modèles de langage.

De façon générale, nos tests permettent de penser que la recherche d'événement en ligne pourrait rapidement muter vers l'usage d'agents conversationnels, mais aussi que les données enrichies et structurées y sont déjà mises à profit tout en offrant une garantie à notre portée quant à la qualité des réponses générées par l'IA.

La grande volatilité des résultats de nos recherches tend aussi à démontrer que cette recherche indépendante doit être pérenne et longitudinale.

L'effort de balisage des actifs en ligne à l'aide de données structurées, par les diffuseurs et parties prenantes de la chaîne économique des arts de la scène, devrait aussi être soutenu et pérenne. Évidemment il s'agit d'un défi qui s'ajoute à la charge de travail déjà importante de ces acteurs et qui interpelle l'éventuelle nécessité de créer de nouvelles compétences au sein des équipes de travail.

À court terme, la mobilisation des données structurées au bénéfice de la promotion des arts de la scène est un problème sectoriel certes, mais devient aussi un enjeu de politique publique. Sa résolution relève probablement d'un bureau, d'un conseil, d'un groupe de réflexion et/ou d'un institut statistique au financement constant. Cela constitue désormais un enjeu de diversité et de souveraineté culturelle.

Références

A10s, et Artsdata. 2024. État des lieux du référencement web des arts de la scène. CAPACOA.
<https://www.google.com/url?q=https://docs.google.com/document/d/19-LOObj8egplVcNDeExzg2y1vF5D10uFpBhpOSJFo3M/edit?tab%3Dt.0&sa=D&source=docs&ust=1756999315420469&usg=AOvVaw3o7-Ttrwo4DstfahZrBQk>

A10s, Christian Roy, Anika Pascale Papillon, et Marc Simonnot. 2025. Élaboration d'un indicateur de rayonnement du spectacle sur les moteurs de recherche. CAPACOA.

CAPACOA, Anicet Lepetit Ondo, Lucie Vézina-Topping, et Gregory Saumier-Finch. 2024. Renforcer et lier les systèmes d'information des arts de la scène. CAPACOA.
<https://capacoa.ca/fr/2024/07/un-point-de-bascule-vers-les-donnees-ouvertes/>

Desmarais, Éric, et Gemini 2.0. 2025. ArtIA - Rapport intérimaire. Sporobole, Société des arts technologiques (SAT), Projet collectif.
<https://praxis.encommun.io/b/artia/n/SMHwIKCGwbdOKHjwAqe0j6rNrBg/>

Downs, Beth, et Chris Unitt. 2025. The Impact of AI Overviews in the Cultural Sector. Royaume-Uni: One Further.
<https://www.onefurther.com/blog/ai-overviews-cultural-sector-report>

Estermann, Beat. 2020. « Creating a Linked Open Data Ecosystem for the Performing Arts (LODEPA) ». <https://doi.org/10.24451/ARBOR.11958>

Estermann, Beat, et Frédéric Julien. 2019. Lier l'avenir numérique des arts de la scène : Comment mobiliser les synergies de la chaîne de valeur. Association canadienne des organismes artistiques (CAPACOA) en coopération avec la Haute École spécialisée bernoise.
https://capacoa.ca/documents/research/ANL_Rapport_2019.pdf

gradient, et Artsdata. 2024. Vers un score de découvrabilité : modélisation explicative multifactorielle. CAPACOA.

gradient, et dia-log. 2023. Mesurer la structuration de données au service de la découvrabilité en ligne - Résultats du projet dia-log. Montréal: Gradient.ai. Synthèse.
https://img1.wsimg.com/blobby/go/83839bab-4939-4901-9f83-b10db51c07da/downloads/Annexe10_BilanGradient_2023-08-24_VF.pdf

Guillon, Vincent, et Emmanuel Vergès. 2025. « La découvrabilité : Un défi culturel à l'ère des plateformes ». L'Observatoire - La revue des politiques culturelles (64): 96.
<https://www.pug.fr/produit/2175/9782706157790/la-decouvrabilite-un-defi-culturel-a-l-ere-des-plateformes>

Jonquet, François. 2025. « SoTicket, une réponse coopérative à des enjeux de souveraineté numérique et de lien au public ». L'Observatoire - La revue des politiques culturelles (64): 59-62. www.observatoire-culture.net

Julien, Frédéric, et Inga Petri. 2017. La numérisation des arts du spectacle. CAPACOA et Strategic Moves. <http://bit.ly/numartsfr>

Le Gall, Anne, et Judith Afopka. 2025. L'interopérabilité des systèmes d'information dans le spectacle vivant : un état des lieux inédit et des pistes pour le secteur. PMT Consultants, TMNlab.

<https://www.tmnlab.com/2025/09/22/linteroperabilite-des-systemes-dinformation-dans-le-spectacle-vivant-un-etat-des-lieux-inedit-et-des-pistes-pour-le-secteur/>

Observatoire de la culture et des communications du Québec (OCCQ). 2020. « Projet d'indicateur(s) de rayonnement de l'offre culturelle québécoise dans les réseaux numériques - Document de référence pour les invitations à soumissionner relativement à l'étude des données pouvant alimenter les indicateurs ».

<https://statistique.quebec.ca/fr/fichier/conception-indicateurs-rayonnement-offre-culturelle-quebecoise-reseaux-numeriques-faisabilite.pdf>

Observatoire de la culture et des communications du Québec (OCCQ), et Marik Danvoye. 2021. Conception d'indicateurs de rayonnement de l'offre culturelle québécoise dans les réseaux numériques. Québec: Institut de la statistique du Québec.

<https://statistique.quebec.ca/fr/fichier/conception-indicateurs-rayonnement-offre-culturelle-quebecoise-reseaux-numeriques-faisabilite.pdf>

Observatoire de la culture et des communications du Québec (OCCQ), Institut de la statistique du Québec. 2025. « Faits saillants sur la fréquentation des arts de la scène au Québec de 2004 à 2023 ». Institut de la statistique du Québec.

<https://statistique.quebec.ca/fr/produit/publication/faits-saillants-frequentation-arts-scene-quebec-donnees-annuelles-2023> (20 août 2025).

Rioux, Michèle, Jean-Robert Bisailon, Guy-Philippe Wells, Martin Tétu, et Sophie Dubois-Paradis. 2021. Être ou ne pas être découvrable ? Présence, visibilité et recommandation des propositions culturelles en ligne : La musique et l'audiovisuel.

UQAM-CEIM-LATICCE. https://ieim.uqam.ca/spip.php?page=article-ceim&id_article=13145

Rioux, Michèle, Jean-Robert Bisailon, Guy-Philippe Wells, et Danielle Desjardins. 2019. « Définir la découvrabilité ». In Découvrabilité des produits culturels en ligne. Un objectif prioritaire pour la coopération bilatérale France-Québec. Rapport réalisé pour le Ministère de la culture et des communications du Québec et le Ministère des Relations internationales et de la Francophonie du Québec, Montréal: UQAM-LATICCE-CEIM, 40-43.

https://www.ceim.uqam.ca/IMG/pdf/ceim_laticce_de_couvrabilite_des_produits_culturels_en_ligne_un_objectif_prioritaire_pour_la_cooperation_bilaterale_france-quebec_2019.pdf

Roberge, Jonathan, Georges Azzaria, Guy Bellavance, et Christian Poirier. 2020. « Chantier sur l'adaptation des droits d'auteur à l'ère numérique. Synthèse de la validation. » : 55.

https://www.mcc.gouv.qc.ca/fileadmin/documents/ForumDroitAuteur2016/INRS_Chantier_sur_adaptation_droits_auteur_rapport_FINAL_16-04-20.pdf

Routhier, Christine, Marik Danvoye, Dominique Jutras, Clément Laberge, Jean-Robert Bisaillon, Danielle Desjardins, et Philippe Michon. 2017. État des lieux sur les métadonnées relatives aux contenus culturels. Québec: (OCCQ) Institut de la statistique du Québec, Observatoire de la culture et des communications du Québec.

<http://www.stat.gouv.qc.ca/statistiques/culture/etat-lieux-metadonnees.pdf>

Saumier-Finch, Gregory. Scène Pro en tant que RDF : Rapport technique sur la faisabilité d'une passerelle vers le Web des données. RIDEAU-CAPACOA.

ANNEXE – Exemples de données structurées

Consultez la page de ressources de Artsdata concernant l'usage de données structurées : <https://www.artsdata.ca/fr/ressources/donnees-structurees>

Voici deux exemples de balisage recommandé assurant une meilleure optimisation des données structurées pour une page Web visée :

Données structurées du Centre des arts Juliette-Lassonde

Exemple de code de balisage du Centre des arts Juliette- Lassonde. On remarque que les **données minimales** et les **données d'enrichissement** identifiées dans ce rapport ont toutes été renseignées. On remarque de surcroît que l'objet de type "Place" définit autant la salle de spectacle que le bâtiment dans lequel elle est située.

```
<script type="application/ld+json">
{
  "@context": "http://schema.org",
  "id": "http://kg.artsdata.ca/resource/K43-1503",
  "type": "Event",
  "name": "Second Chances",
  "url": "https://www.centredesarts.ca/spectacle/rubberband-second-chances",
  "startDate": "2026-04-16T20:00:00-04:00",
  "endDate": "2026-04-16T22:00:00-04:00",
  "location": {
    "@type": "Place",
    "name": "Salle Desjardins",
    "sameAs": "http://www.wikidata.org/entity/Q111668872",
    "containedInPlace": {
      "@type": "Place",
      "name": "Centre des arts Juliette-Lassonde"
    },
  },
  "address": {
    "@type": "PostalAddress",
    "streetAddress": "1705 rue Saint-Antoine",
    "addressLocality": "Saint-Hyacinthe",
    "addressRegion": "QC",
    "postalCode": "J2S 9E2",
    "addressCountry": "CA"
  },
  "description": "Second Chances, le nouveau spectacle de répertoire de RUBBERBAND comprend dans un premier temps Commissions Suite ...",
  "image": "https://www.centredesarts.ca/img-activites/rubberband-second-chances-1280x720.webp",
}
```

```

"performer": {
  "type": "Organization",
  "name": "Rubberband",
  "sameAs": "http://www.wikidata.org/entity/Q16640268",
},
"organizer": {
  "@type": "Organization",
  "name": "Centre des arts Juliette-Lassonde de Saint-Hyacinthe",
  "url": "https://www.centredesarts.ca/"
},
"offers": {
  "@type": "Offer",
  "availability": "http://schema.org/InStock",
  "validFrom": "2025-04-30",
  "validThrough": "2026-04-16T20:00:00",
  "price": "42.00",
  "priceCurrency": "CAD",
  "url": "https://centredesarts.tuxedobillet.com/main/rubberband-260416"
},
"eventStatus": "http://schema.org/EventScheduled",
"eventAttendanceMode": "http://schema.org/OfflineEventAttendanceMode"
}
</script>

```

Note : Ce code diffère quelque peu du code de balisage employé [sur le site du Centre des arts Juliette-Lassonde](#). En outre, le Centre des arts a opté pour le format d'encodage « microdonnées » plutôt que le format « JSON-LD », tel qu'utilisé plus haut. Le format JSON-LD est plus couramment utilisé et il est plus facile à lire pour des humains.

Données structurées du Sanderson Centre

Voici un exemple de balisage de données structurées du Sanderson Centre pour la représentation de Dallas Smith (cité dans la section « Étude de deux événements comparables »). On peut y voir que le lieu et l'organisateur sont très clairement identifiés.

```

<script type="application/ld+json">
{"@context":
  {"@vocab": "http://schema.org/",
   "footlight": "http://kg.footlight.io/resource/"
  },
  "@type": "Event",
  "organizer": {
    "@id": "http://kg.artsdata.ca/resource/K10-499",
    "@type": "Organization",
    "sameAs": [
      "http://www.wikidata.org/entity/Q112669151",
      "https://www.facebook.com/sanderson.centre",

```

```

    "https://twitter.com/SandersonCentre/",
    "https://www.instagram.com/sandersoncentrebrantford/"
  ],
  "alternateName": [
    {"@language": "en", "@value": "Brantford Sanderson Centre"},
    "Brantford Sanderson Centre"
  ],
  "url": {"@id": "https://www.sandersoncentre.ca/en/index.aspx"},
  "name": {"@language": "en", "@value": "Sanderson Centre"}
},
{
  "duration": "PT7200S",
  "offers": {
    "@id": "footlight:sandersoncentre-ca_2025-12-04-2000-Dallas-Smith#Offer",
    "@type": "Offer",
    "availability": "InStock",
    "price": "82.00",
    "url": "https://tickets.sandersoncentre.ca/TheatreManager/1/online?performance=1885"
  },
  "mainEntityOfPage": {
    "@id": "https://calendar.sandersoncentre.ca/Default/Detail/2025-12-04-2000-Dallas-Smith#WebPage",
    "@type": "WebPage",
    "inLanguage": "en",
    "url": "https://calendar.sandersoncentre.ca/Default/Detail/2025-12-04-2000-Dallas-Smith",
    "lastReviewed": "2025-10-06T16:17:12.690-04:00"
  },
  "location": {
    "@id": "http://kg.artsdata.ca/resource/K11-192",
    "@type": ["Place", "PerformingArtsTheater"],
    "sameAs": [
      "http://www.wikidata.org/entity/Q38386517",
      "https://en.wikipedia.org/wiki/The_Sanderson_Centre"
    ],
    "address": {
      "@id": "http://kg.artsdata.ca/resource/K11-192#PostalAddress",
      "@type": "PostalAddress",
      "addressCountry": "CA",
      "addressLocality": "Brantford",
      "addressRegion": "ON",
      "postalCode": "N3T 2J2",
      "streetAddress": "88 Dalhousie Street"
    }
  }
}

```

```

    },
    "url": {"@id": "https://www.sandersoncentre.ca/en/index.aspx"},
    "name": {"@language": "en", "@value": "Sanderson Centre"}
  },
  "image": "https://calendar.sandersoncentre.ca/Default/Detail/2025-12-04-2000-Dallas-Smith/b56773f9-23c2-4255-9702-b2f000d5d7bb",
  "eventStatus": "http://schema.org/EventScheduled",
  "eventAttendanceMode": "http://schema.org/OfflineEventAttendanceMode",
  "name": {"@language": "en", "@value": "Dallas Smith"},
  "description": {"@language": "en", "@value": "Big Loud artist Dallas Smith has a track record of noteworthy success – from front man of platinum-selling rock band Default to 3x consecutive CCMA Entertainer of the Year, and now his U.S. debut. Across his storied career, the bona fide hitmaker has amassed 13 No. 1 singles at Country Radio, 18 gold-certified singles, seven platinum-certified singles, and four gold-certified albums, with more than 520 million global streams and over two million album equivalents to date."},
  "url": "https://calendar.sandersoncentre.ca/Default/Detail/2025-12-04-2000-Dallas-Smith",
  "startDate": "2025-12-04T20:00:00-05:00",
  "endDate": null,
  "workPerformed": null,
  "performer": null
}
</script>

```

Voici une capture d'écran des mêmes données, telles que représentées dans le Validateur de [Schema.org](https://schema.org/) :

Schema.orgDocumentationSchemasValidateAbout

https://calendar.sandersoncentre.ca/Default/Detail/2025-12-04-2000-Dallas-SmithRun new test

<!-- Google Tag Manager -->

<script>

function(w,d,s,l,i){

w[l]=w[l]||[];w[l].push({

'gtm.start':

new Date().getTime(),event:'gtm.js'

});var f=d.getElementsByTagName(s)[0],

j=d.createElement(s),dl=l!='dataLayer'?'&l='+l:'';j.async=true;j.src=

https://www.googletagmanager.com/gtm.js?id='+i+dl;f.parentNode.insertBefore(f,j);

})(window,document,'script','dataLayer','GTM-PN0T2N0');

</script>

<!-- End Google Tag Manager -->

<link rel="preload" href="/Resources/fonts/GMD-Moderat-Regular.woff" as="font" type="font/woff" crossorigin>

<link rel="preload" href="/Resources/fonts/GMD-Moderat-Black.woff" as="font" type="font/woff" crossorigin>

<link rel="preload" href="/Resources/fonts/GMD-Moderat-Light.woff" as="font" type="font/woff" crossorigin>

<link rel="preload" href="/Resources/fonts/GMD-Moderat-Regular.woff2" as="font" type="font/woff2" crossorigin>

<link rel="preload" href="/Resources/fonts/GMD-Moderat-Black.woff2" as="font" type="font/woff2" crossorigin>

<link rel="preload" href="/Resources/fonts/GMD-Moderat-Light.woff2" as="font" type="font/woff2" crossorigin>

<link rel="preload" href="/Resources/fonts/GMD-Moderat-Regular.ttf" as="font" type="font/ttf" crossorigin>

<link rel="preload" href="/Resources/fonts/GMD-Moderat-Black.ttf" as="font" type="font/ttf" crossorigin>

<link rel="preload" href="/Resources/fonts/GMD-Moderat-Light.ttf" as="font" type="font/ttf" crossorigin>

<link href="/Resources/css/bootstrap.min.css?v=2" rel="stylesheet" type="text/css" media="screen" />

<link href="/Resources/css/bootstrap-colorpicker.min.css" rel="stylesheet" type="text/css" media="screen" />

<link href="/Resources/css/font-awesome.min.css" rel="stylesheet" type="text/css" media="screen" />

<script type="text/javascript">

var esol_ico_width = 0;

try {

esol_ico_width = window.innerWidth

|| document.documentElement.clientWidth

|| document.body.clientWidth;

} catch (err) { }

var iCreateObject =

{

isSiteInCreateMode: 'false'.toLowerCase() == "true",

lang: 'en',

isLe8: false,

corpHome: 'https://www.sandersoncentre.ca/',

corpRoot: '/',

getTrackingCode: '',

isSiteInLocalDevelopment: 'false'.toLowerCase() == "true",

calendarUrl: 'https://calendar.sandersoncentre.ca/',

isSiteInPreviewMode: 'false'.toLowerCase() == "true",

isCreateGeneratedPage: 'false'.toLowerCase() == "true",

newFeeder: {},

isMobile: esol_ico_width <= 767,

iCreateTheme: '',

fullPageTitle: 'http://www.sandersoncentre.ca/en/inc/calendar.aspx',

iCreatePageId: 78912,

iCreateHome: 'https://ic9.esolig.ca/',

editFrontView: 'Fe Fe'

}

}

Event

0 ERRORS 0 WARNINGS

| | | | | | | | | | | | | | | | | | | | |
|---------------------|--|-------|--------------|-----|---|--------------|---|--------|--|--------|--|--------|---|---------------|----------------------------|-----|--|------|------------------|
| @type | Event | | | | | | | | | | | | | | | | | | |
| duration | PT2H | | | | | | | | | | | | | | | | | | |
| image | https://calendar.sandersoncentre.ca/Default/Detail/2025-12-04-2000-Dallas-Smith/b66775f9-23c2-4255-9702-b2f000d5d7bb | | | | | | | | | | | | | | | | | | |
| eventStatus | http://schema.org/EventScheduled | | | | | | | | | | | | | | | | | | |
| eventAttendanceMode | http://schema.org/OfflineEventAttendanceMode | | | | | | | | | | | | | | | | | | |
| name | Dallas Smith | | | | | | | | | | | | | | | | | | |
| description | Big Loud artist Dallas Smith has a track record of noteworthy success – from front man of platinum-selling rock band Default to 3x consecutive CMAA Entertainer of the Year, and now his U.S. debut. Across his storied career, the bona fide hitmaker has amassed 13 No. 1 singles at Country Radio, 18 gold-certified singles, seven platinum-certified singles, and four gold-certified albums, with more than 520 million global streams and over two million album equivalents to date. | | | | | | | | | | | | | | | | | | |
| url | https://calendar.sandersoncentre.ca/Default/Detail/2025-12-04-2000-Dallas-Smith | | | | | | | | | | | | | | | | | | |
| startDate | 2025-12-04T20:00:00-05:00 | | | | | | | | | | | | | | | | | | |
| organizer | <table><tr><td>@type</td><td>Organization</td></tr><tr><td>@id</td><td>http://kg.artsdata.ca/resource/K10-499</td></tr><tr><td>sameAs</td><td>http://www.wikidata.org/entity/Q112669151</td></tr><tr><td>sameAs</td><td>https://www.facebook.com/sandersoncentre</td></tr><tr><td>sameAs</td><td>https://twitter.com/SandersonCentre/</td></tr><tr><td>sameAs</td><td>https://www.instagram.com/sandersoncentrebrentford/</td></tr><tr><td>alternateName</td><td>Brantford Sanderson Centre</td></tr><tr><td>url</td><td>https://www.sandersoncentre.ca/en/index.aspx</td></tr><tr><td>name</td><td>Sanderson Centre</td></tr></table> | @type | Organization | @id | http://kg.artsdata.ca/resource/K10-499 | sameAs | http://www.wikidata.org/entity/Q112669151 | sameAs | https://www.facebook.com/sandersoncentre | sameAs | https://twitter.com/SandersonCentre/ | sameAs | https://www.instagram.com/sandersoncentrebrentford/ | alternateName | Brantford Sanderson Centre | url | https://www.sandersoncentre.ca/en/index.aspx | name | Sanderson Centre |
| @type | Organization | | | | | | | | | | | | | | | | | | |
| @id | http://kg.artsdata.ca/resource/K10-499 | | | | | | | | | | | | | | | | | | |
| sameAs | http://www.wikidata.org/entity/Q112669151 | | | | | | | | | | | | | | | | | | |
| sameAs | https://www.facebook.com/sandersoncentre | | | | | | | | | | | | | | | | | | |
| sameAs | https://twitter.com/SandersonCentre/ | | | | | | | | | | | | | | | | | | |
| sameAs | https://www.instagram.com/sandersoncentrebrentford/ | | | | | | | | | | | | | | | | | | |
| alternateName | Brantford Sanderson Centre | | | | | | | | | | | | | | | | | | |
| url | https://www.sandersoncentre.ca/en/index.aspx | | | | | | | | | | | | | | | | | | |
| name | Sanderson Centre | | | | | | | | | | | | | | | | | | |
| offers | <table><tr><td>@type</td><td>Offer</td></tr><tr><td>@id</td><td>http://kg.footlight.io/resource/sandersoncentre-ca_2025-12-04-2000-Dallas-Smith#Offer</td></tr><tr><td>availability</td><td>http://schema.org/InStock</td></tr><tr><td>price</td><td>82.00</td></tr><tr><td>url</td><td>https://tickets.sandersoncentre.ca/TheatreManager/!online?</td></tr></table> | @type | Offer | @id | http://kg.footlight.io/resource/sandersoncentre-ca_2025-12-04-2000-Dallas-Smith#Offer | availability | http://schema.org/InStock | price | 82.00 | url | https://tickets.sandersoncentre.ca/TheatreManager/!online? | | | | | | | | |
| @type | Offer | | | | | | | | | | | | | | | | | | |
| @id | http://kg.footlight.io/resource/sandersoncentre-ca_2025-12-04-2000-Dallas-Smith#Offer | | | | | | | | | | | | | | | | | | |
| availability | http://schema.org/InStock | | | | | | | | | | | | | | | | | | |
| price | 82.00 | | | | | | | | | | | | | | | | | | |
| url | https://tickets.sandersoncentre.ca/TheatreManager/!online? | | | | | | | | | | | | | | | | | | |

Données structurées proposées par ChatGPT

ChatGPT offre un exemple de **code de balisage** à utiliser pour un artiste participant à un événement. Soulignons que celui-ci valide la plupart de nos conclusions, même si ChatGPT a négligé l'intégration des identifiants pérennes avec la propriété "sameAs".

```
<script type="application/ld+json">
{
  "@context": "https://schema.org",
  "@type": "Event",
  "name": "Concert de Jazz - Trio Lumière",
  "startDate": "2025-09-15T20:00",
  "endDate": "2025-09-15T22:00",
  "eventStatus": "https://schema.org/EventScheduled",
  "eventAttendanceMode": "https://schema.org/OfflineEventAttendanceMode",
  "location": {
    "@type": "Place",
    "name": "Théâtre de la Ville",
    "address": {
      "@type": "PostalAddress",
      "streetAddress": "12 Rue des Arts",
      "addressLocality": "Paris",
      "postalCode": "75004",
      "addressCountry": "FR"
    }
  }
},
```

45 | 49

```
"image": [  
  "https://exemple.com/images/concert.jpg"  
],  
"description": "Une soirée exceptionnelle de jazz moderne avec le Trio  
Lumière.",  
"performer": {  
  "@type": "MusicGroup",  
  "name": "Trio Lumière"  
},  
"offers": {  
  "@type": "Offer",  
  "url": "https://exemple.com/billets",  
  "price": "35",  
  "priceCurrency": "EUR",  
  "availability": "https://schema.org/InStock",  
  "validFrom": "2025-07-01T12:00"  
},  
"organizer": {  
  "@type": "Organization",  
  "name": "Jazz Events Paris",  
  "url": "https://jazzeventsparis.com"  
}
```