ARTSDATA

DRIVING DISCOVERABILITY IN THE PERFORMING ARTS THROUGH STRUCTURED DATA

RESULTS REPORT

October 2025

Gaëlle Ramboanasolo Gradiant inc, Christian Roy and Anika Pascale Papillon A10S, Jean-Robert Bisaillon LATICCE, Frédéric Julien CAPACOA

Suggested bibliographic citation:

Ramboanasolo, Gaëlle, Christian Roy, Jean-Robert Bisaillon, Anika Pascale Papillon and Frédéric Julien. 2025. Driving Discoverability in the Performing Arts through Structured Data: Results Report. Canadian Association for the Performing Arts (CAPACOA).

English translation: Gilbert Bélisle

Legal deposit: Library and Archives Canada

ISBN: 978-1-0688996-4-5

Photo on cover page: *Burn Baby, Burn*, a Côté Danse (Q126952121, K4-155) production. This dance show presented by the Burling Performing Arts Centre Centre was part of the B group in our A/B testing. Photo credit: Aidan Tooth; performer: Griffen Grice; costume design: Yso South.

This report is published by CAPACOA, the Canadian Association for the Performing Arts, as part of the Artsdata project. The Artsdata project is funded by the Government of Canada and the Canada Council for the Arts. The views, opinions, and recommendations expressed in this report are those of the authors and do not necessarily reflect those of the Government of Canada or the Canada Council for the Arts.

Canada

Conseil des arts du Canada | Canada Council for the Arts

Table of Contents

Summary

The objective of this research project was to illustrate and measure the impact of the use of structured data in the online search and recommendation of performing arts events.

We began by defining a granular typology of dissemination and discoverability factors. We developed methodologies for creating two dissemination indices to calculate two discoverability markers: one to quantify structured data and one to quantify the dissemination of live performances in search engine results pages (i.e. Google Search). Finally, we correlated these indices with web traffic data from Google Analytics.

The process we sought to illustrate consists of acting on a series of factors (action variables, Schema tags) that have an impact on markers (consequence variables), and which allow us to produce measured indices of discoverability: **Factors** > **Markers** > **Indices**.

This multimodal approach allowed us to consider confounding factors (i.e. variables such as gender, target audience or advertising campaigns that could cause our markers and indices to fluctuate) in order to isolate the effect of data structuring variables.

The highlights of our work are as follows.

Even when confounding factors are taken into account, positive relationships can be seen between the level of event data structuring and:

- exposure on Google Search results pages; and
- the number of page views coming from Google Search.

The relationship between the level of data structuring and the number of pages is relatively linear: the more structured the data, the greater the number of page views.

Beyond the minimum structured data (the name, date/time, and location of an event), data describing the content of the event and its main participants have a particularly strong effect. In addition, the *description* of an event has a statistically significant effect. Similarly, data describing the *performer* and the event *organizer* also have a significant effect on the number of page views.

Finally, we considered current trends in the browsing habits of Internet users, particularly the increased use of AI-assisted chatbots.

At the end of this work, as anticipated by our hypotheses, we concluded that the addition of structured data to web pages by performing arts event presenters has a measurable impact on the digital discoverability of their events.

This report presents a more detailed account of the benefits and impacts of documenting performing arts events online and proposes some recommendations aimed at improving practices in this area.

Glossary

A/B testing: Testing procedure wherein performance results of two near-identical objects are compared and differences are measured using a control variable (control group O).

Example: Which would have a greater audience, a performance of a show at 7 or at 9 p.m?

To see if search results differ, A/B discoverability tests can be carried out on three live performances that are sold online: one without structured data (control group O), one with minimal structured data (group A), and one with enhanced metadata (group B).

Consumption: Quantitative value of conversion from click to page view to purchase; i.e. an action taken by the Internet user resulting from their online journey (see "Marker").

Discoverability and Dissemination: These two terms can be confusing. The first is formally defined by the Observatoire de la culture et des communications du Québec and contextualized by Michèle Rioux et al. (2019, p.41) according to the following established formula: "Discoverability is the ability of cultural content to be easily discovered by consumers who are looking for it and to be offered to consumers who were not aware of its existence."

The concept of dissemination has been crystallized in a report by the Observatoire de la culture et des communications du Québec (OCCQ) and in Danvoye (2021, p.10). Its distinctive feature is that it adds the measurement of actual consumption to that of discoverability.[1] For example, Rioux et al. (2019) refer to dissemination using the following formula: *promoting dissemination through exports*.

To qualify and quantify the results of measuring these phenomena, the OCCQ and the Danvoye report uses the term "marker."

Enhanced metadata: Metadata associated mainly with the creative stages of a production prior to marketing and promotion (essentially works and expressions–see the definition of "Work/WEMI"). **Contributions**, such as writing, directing, set design, or lighting design, are examples of enhanced metadata. This type of data can also refer to shows. It can be "enhanced" with the names of new **performers**, for example, and with other data that can change during the life of a show.

Factor: Action (addition of tags, factual data, and metadata) that enhances the description of an event before measuring its effects on consumption.

Index: Score based on a scale or percentage; a representation of dissemination *markers—that is, the extent to which best practices promoting dissemination have been applied.* The same notion applies to the concept of discoverability. The project developed an exposure index

[1] See Figure 1.

measuring ranking on search engine results page and a structured data index measuring the enhancement of a given web page using the appropriate Schema.org tags.[2]

Indicator: See “Marker”.

Marker: Consequence (quantitative result variable) associated with the concept of dissemination, which reflects the presence of an action based on good tagging practice, i.e. *measures that seek to explicitly represent dissemination*.

Markup: Addition of structured data (semantic tags such as a Schema.org triple) to a web page to generate dissemination markers. The presence of a dissemination marker is a factor that embodies an indicator, which can be used to produce a calculated index (sometimes referred to as indexing).

Metadata: Documentary information, tag or factor associated with a digitized object to enable a human or robot (algorithm, computer program, machine) to better understand content. A show **title** or a performance **date** are metadata.

Microdata: Encoding format (processing, conversion in machine language) for integrating structured data in a web page (HTML documents) and enabling it to be found and understood by search engines and AI crawlers. There are three main encoding formats for structured data: microdata, RDFa, JSON-LD. All three formats are supported by search engines and Artsdata.

Performance: Event during which a live performance work (i.e. a show) is publicly presented for an audience. Upon publishing this report, a proposal for integrating this concept to the Schema.org vocabulary was under consideration.[3]

Persistent Identifier: Computer code that may or may not be understandable to humans and that allows algorithms and machines to distinguish a real-world entity in an information system. The ISBN code (978-2376620167) of a paperback sold online, the QID code (Q560434) of any entity described in Wikidata, or the Artsdata identifier (ad) of any entity described in the Artsdata knowledge graph (K2-5867), are all persistent identifiers.

Presence: Variable P of the LATICCE P/V/R model. In this report, presence is implicit because the tests were carried out using a reference basket consisting of the offerings of the presenters and websites studied.

Property: In structured data, properties express relationships between objects (i.e. real-world entities) or their attributes. In the RDF triple (subject-predicate-object), the property is a predicate.

Recommendation: Variable R of the LATICCE P/V/R model, i.e. the potential for information or content to be recommended in a digital environment. The quality of a recommendation can be

---

[2] See https://vitrinelinguistique.oqlf.gouv.qc.ca/fiche-gdt/fiche/2080594/marqueur
[3] See the discussion “Proposal (Performing Arts 2) - New sub-class of Event: PerformingArtsEvent”. https://github.com/schemaorg/schemaorg/issues/4468

measured using multiple secondary variables and is what this project has fundamentally undertaken to demonstrate. The LATICCE model, which has been used to measure access to digital music on streaming platforms, has determined that the sub-variables of concordance, relevance, and novelty can determine whether or not the taste profile of fictitious Internet users corresponding to personas is being respected. This report measures the quality of recommendations based on an exposure index assessing the appearance of information on Google search results pages.

Show: A work composed of a set of creative elements, including designs, intended to be performed live for an audience. In other words, a show refers to the content that is presented on stage during a performing arts event. The term “show” is sometimes used in everyday language to refer to the event itself (i.e. the performance of a show). Upon publishing this report, a proposal for integrating this concept into the Schema.org vocabulary was under consideration.[4]

Structured data: Data and metadata conditioned according to **RDF** (Resource Description Framework) logic; a graph model designed to formally describe Web resources and their metadata and enable the automatic processing of such descriptions (Wikipedia). A structured and formal data statement is based on a triple: **subject**-**predicate**-**object**. The triple can be compared to the usual English language syntax subject-verb-object. For example:

Salle Desjardins is located inside Centre des arts Juliette-Lassonde in Saint-Hyacinthe, Québec.

Structure data integrated in web pages for search engines is most often represented in the Schema.org vocabulary and encoded JSON-LD.[5]

Thus, the statement about the Salle Desjardins becomes the following code script:

[4] See the discussion “Proposal (Performing Arts 1) - New sub-class of CreativeWork: PerformanceWork”. https://github.com/schemaorg/schemaorg/issues/4467

[5] See this Google documentation: https://developers.google.com/search/docs/appearance/structured-data/intro-structured-data

```
<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>
```

Visibility: Variable V of the LATICCE P/V/R model. Visibility is a discoverability variable that reflects the production and display of editorial content by human actors prior to any impact from recommendation algorithms. A high visibility will certainly have an impact on automatic recommendations by algorithms.

Work / WEMI: Reference to the IFLA data model, which includes the elements Work / Expression / Manifestation / Item. This model is suitable to a degree for describing the performing arts value chain, particularly due to the complex nature of a “performance work” (it contains or incorporates Expressions from several other works) and the absence of a Manifestation embodied by a physical or virtual Item. Due to the ambiguity of the term "work" as defined in the WEMI model, we have preferred to use the term “show” for the purposes of this research.

Context and Objective

As part of the Artsdata project, the Canadian Association for the Performing Arts (CAPACOA) asked the authors to develop indicators that can quantify the digital discoverability of the performing arts. This report summarizes the work carried out during an exploratory phase that took place between January and July 2024, followed by a second phase of systematic harvesting of results from the Google Search, which took place between August 2024 to August 2025. The methodology used is based on the creation of models that use various variables to enable a reproducible and verifiable measurement of the online discoverability of performing arts shows and events, in particular the use of structured metadata from the Schema.org consortium.

Two short appendices examine two related phenomena associated with the issue under study: the weight of ticket resale sites in search results and the growing role of recommendation applications powered by AI.

Finally, the report presents conclusions and recommendations for increasing traffic to performing arts event pages.

Introduction and Hypotheses

In August 2025, several publications set the stage for a renewed conversation about live performance attendance in the digital age. The *Enquête québécoise sur les loisirs culturels et le divertissement* (EQLCD), the most recent survey on cultural leisure and entertainment in Québec, was produced by the Institut de la statistique du Québec (Observatoire de la culture et des communications du Québec (OCCQ) 2025). Additionally, two articles were published by Mario Girard in *La Presse* on the "good old paper ticket"[6].

In general, attendance at local artists' shows tends to decrease as the audience age decreases. The same can be said for other cultural products. Mario Girard points to online ticket purchasing environments as a deterrent to attending shows. These summary observations nevertheless draw attention to the fact that possible disruptions to cultural habits are underway.

Online discoverability is currently very popular in a context where there has never been as much choice and as many products. Nevertheless,in contrast to the audiovisual and sound recording industries, there has been little research on the performing arts and the entertainment sector (notwithstanding the efforts of our group of researchers: A10s and Artsdata 2024; gradiant and Artsdata 2024; gradiant and dia-log 2023; Estermann and Julien 2019; Julien and Petri 2017) or public policy on this issue. This project therefore aims, once again, to be part of this dynamic, to explore how performing arts are affected by these cultural digital shifts.

[6] La Presse, August 11 & 19, 2025 https://www.lapresse.ca/arts/chroniques/2025-08-19/la-billetterie-et-vous.php

The initial hypothesis of our work is based on the LATICCE discoverability measurement equation, which states that the more an event is exposed (P/V/R: present, visible, recommended) (Rioux et al. 2019, 2021, p. 15), the likelier it will be consumed by the public, i.e. viewed and ultimately purchased. We set out to assess the extent to which the **visibility factors** of a cultural offering influence its consumption.

There are various gaps in our understanding of the discoverability of the performing arts, due in part to the shortage of research on this topic. In terms of audience habits, we need to better understand the socio-constructed practices associated, for example, with using the Internet to find shows and purchase tickets. (Roberge et al. 2020). How are audience practices evolving? We considered these issues as best we could when we developed the methodological parameters for launching our tests. To systematize our understanding of these constructs, we developed a typology of the variables involved. The methodology section provides a detailed description of the process.

For the time being, Artsdata has chosen to use a functionalist and technical approach to the study of performing arts organizations' marketing practices.

The project's initial research objectives and questions focused on observing the results of efforts to stimulate discoverability by comparing the practices of presenters and players in the performing arts value chain. A prerequisite was the production of an overview of measurable manifestations of discoverability and dissemination (CAPACOA et al. 2024).

The project examined the use of best practices for content indexing and search engine optimization (SEO), and in particular, the use of enhanced descriptive metadata and **structured data** tags from the Schema.org repository.

Kuné (Q125392378, K2-5096). The ensemble's performance at the Sanderson Centre was part of the B group in our A/B testing. Photo credit: Zahra Saleki.

Methodology

Methodological Premise

Our initial methodological premise assumes that the search engine Google Search is a "barometer" of the digital environment. Search engines create summaries of the massive amounts of data published online, while aligning themselves as closely as possible with consumer expectations and intentions.

We carried out systematic and automatic harvesting of the results generated by Google Search and used these results for case studies. The queries used to generate the results were based on a large sample of queries representative of consumers' discovery paths, on a selection of five presenters during phase 1 of the initiative and on three presenters during A/B/O testing in phase 2 (Burlington Performance Arts Center, Sanderson Centre in Brantford, Centre des arts Juliette-Lassonde in Saint-Hyacinthe). Presenter-centric queries were defined for a list of shows and specific time periods.

Analysis involved determining to what degree the data was structured according to the recommended typology. During phase 1, events were classified into quintiles, and the results were compared according to the number of page views and the amount of structured data on the event pages involved. Phase 2 involved the testing of explanatory models based on three A/B/O modeling scenarios, to accurately assess the impact of data structuring on the discoverability of events.

The observation methodologies are detailed in the following pages of this report.

Creation of a Typology

This research is based on various quantitative methods, which are themselves based on a preliminary typology of various markers,[7] entities (elements, subjects), attributes (properties, predicates), values (objects), indices, indicators, and dissemination factors.[8, 9]

The effort to produce a performing arts typology and associated definitions has been ongoing for several years and began prior to the current structured data mobilization project. This research is an integral part of that effort, as the typology produced is never finished. Stakeholders involved in applying the principles of discoverability must adopt such a typology by consensus and understand its implications.

---

[7] See Figure 1.

[8] See the list of properties associated with the event descriptions proposed by Artsdata: https://culturecreates.github.io/artsdata-data-model/classes/event.html

[9] See the CAPACOA Data Typology Table.

The production of the CAPACOA/Artsdata typology was designed to align with international standards. However, unlike other sectors of the cultural economy, where there are consortiums and decision-making bodies for standardized practices,[10] no such space existed for the performing arts when the work began. Consequently, we had to build consensus and identify suitable forums where we could debate a typology, as well as the minimum requirements for documenting aspects of the performing arts, including performances in particular. The international community gathered around WikiProject Performing Arts[11] and W3C[12] served as a sounding board for decisions on defining a typology. CAPACOA also took on the responsibility of convening a working group, LODEPA WG 6 Wikidata/Wikimedia, which addresses data modelling challenges in the performing arts. Stakeholders from Canada, the United Kingdom, Switzerland, Belgium and Croatia are participating in these conversations.

It is not easy to fully integrate the nuances between the concepts associated with standardized digital definitions and those of the pre-digital world, which tended to be viewed as “marketing strategies” and “sales results.” In this regard, Figure 1 represents the spectrum of dissemination markers according to the Observatoire de la culture et des communications du Québec (2021).

Our main efforts in producing a typology have therefore focused on identifying the **properties** (attributes, predicates) most likely to enable live performances to emerge in Google Search. What are the properties that define a given performance? A/B testing focused on these properties.

[10] OCLC for books, DDEX for recorded music, EIDR for audiovisual works.

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

[12] https://www.w3.org/community/pair-cg/ (group now inactive)

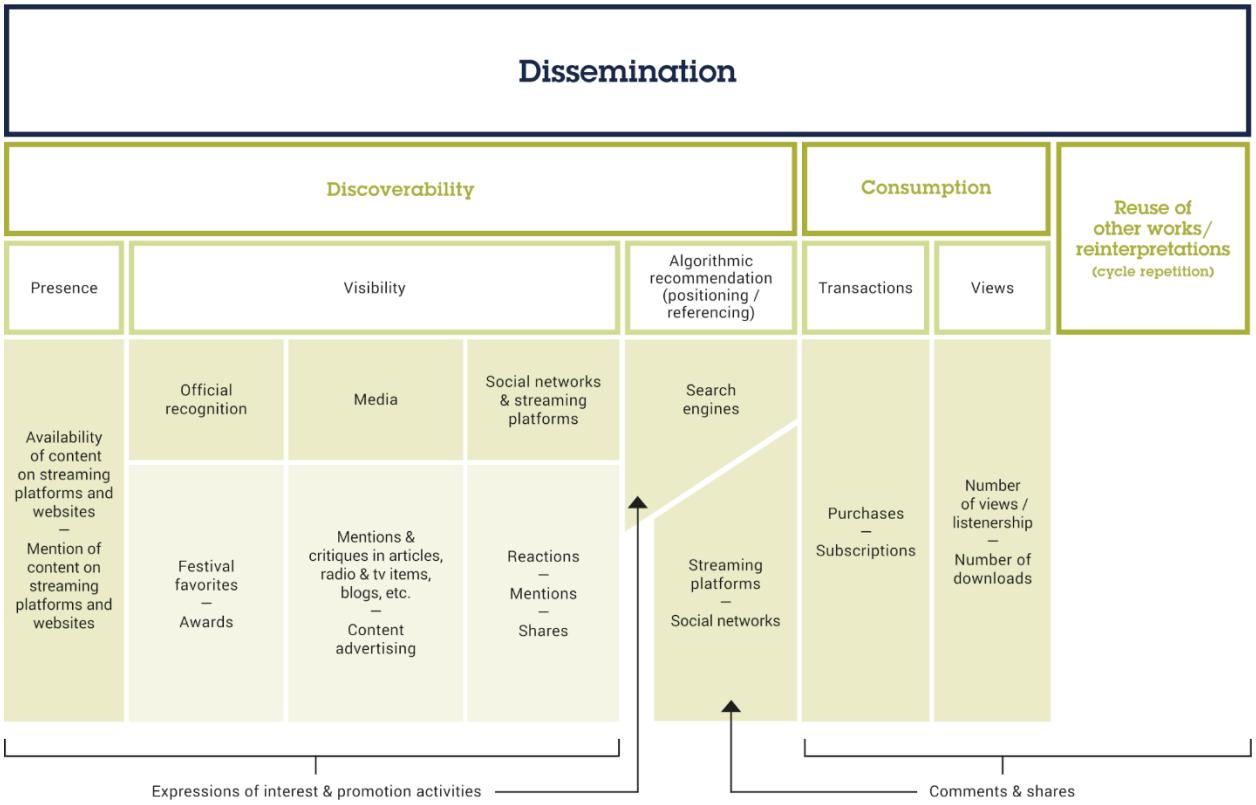


Figure 1. Dissemination markers according to the Observatoire de la culture et des communications du Québec (OCCQ) and Danvoye (2021).

Algorithm for Calculating a Structured Data Index

Once the typology was defined, we felt it was relevant and realistic to produce an initial index based solely on the algorithmic calculation of the presence or absence of certain structured data on a given web page, while slightly weighting the presence of structured data according to its importance for disambiguation and recommendation. This weighting was based on the status of the properties associated with the description of events, as established by Artsdata / CAPACOA:[13] essential, recommended, or optional. Properties declared as essential are those deemed necessary to disambiguate events. These are properties, like the exact date/time and location, that make it possible to distinguish between similar events. Recommended properties are those requested by many Artsdata data consumers and cultural calendars. The image of an event, for example, is one such property. It isn't required for disambiguating an event, but it is required by events listings (and expected by arts goers looking up events).

[13] See the list of properties associated with the Event class in the Artsdata data model: https://culturecreates.github.io/artsdata-data-model/classes/event.html

Researchers Frédéric Julien and Gregory Saumier-Finch designed the algorithm and submitted it to the team and to members of the Artsdata community of interest for comments and suggestions.

Thus, the resulting structured data index offers a quantitative measure of the exhaustivity of structured data on an event web page. [14]

Defining the Query Basket and Harvesting Google Search Results

Step 1: State of Web Referencing for the Performing Arts

A10s was asked to assess the web referencing of the performing arts[15] to provide a basis for developing approaches to quantify discoverability.

The methodology is based on a set of queries representative of the intentions of performing arts consumers. The team began by generating a large number of queries, which were then filtered to ensure that they were balanced according to several factors. These include various linguistic (French/English) and geographical (major centers/regions) criteria, referencing in Artsdata (entities present or not), sufficient search volume and an absence of ambiguity (search terms that could refer to other contexts, such as a film or a book, were avoided).

In the end, 168 queries were selected and classified into three types deemed representative of consumer intent:

- **Show/artist**: a search for a specific named show or artist (e.g. Jesse Cooke).
- **Recommendations by language (FR, EN, bilingual)**: general searches such as “what to do in Montreal” or “Toronto concerts.”
- **Festival, presenter-producer, venue:** intermediate-level queries where the consumer knows a characteristic (name of festival or venue) without searching for a specific show.

The resulting set of queries was used to examine and compare Google Search results. The set was run on the French and English versions of Google Search Canada. The search results were used to assess the current situation, the main findings of which appear later in this report.

This snapshot of the current state provided guidelines for the development of a discoverability indicator. This indicator should represent the concrete results of SEO efforts rather than the actions themselves, be measurable over time, and remain simple enough that organizations without advanced technical expertise can calculate them. Identified research avenues include the importance of choosing a single version of Google Search adapted to the target market, counting the simple presence of events in enhancements (rather than relying on format

[14] See the README file establishing the parameters of the structured data index calculation algorithm: https://github.com/culturecreates/artsdata-score

[15] See a detailed version of the status report at État des lieux du référencement web des arts de la scène - Rapport

changes), building a basket of queries aligned with the offer and consumer intentions, and finally, weighting the results based on the source (official website of the organizer, aggregators, ticket offices, resellers), in order to reflect the relative value of each type of visibility.

Step 1 is described in detail in the report État des lieux du référencement web des arts de la scène.

Step 2: Developing an Approach to Quantifying Exposure

Based on the findings of the Step 1 assessment, two strategies for quantifying exposure were developed.

The first involves analyzing the Google Search results page of a specific query from the perspective of a presenter. This involves assigning points to each component of the search results page and categorizing the points into two groups: those assigned to the presenter and those assigned to all other players. The index is obtained by dividing the number of points assigned to the presenter by the total number of points. The procedure is documented in detail.

The second strategy extends this approach. It identifies several queries relevant to a presenter and calculates the query's index as previously described, before calculating a weighted average of the different indices. The weighting factors reflect the importance of queries in the discovery process. The procedure for creating the basket of queries and for calculating the weighted index are also documented in detail.

Both strategies were tested by applying them to five presenters at three different times.

The following presenters were used as query targets:

- The Capitol Theatre in Moncton, New Brunswick
- The Théâtre Capitole in Québec City, Québec
- The FirstOntario Performing Arts Centre in St. Catharines, Ontario
- Place des Arts in Montréal, Québec
- The Sanderson Centre for the Performing Arts in Brantford, Ontario

The index was measured three times at seven-week intervals on October 18, 2024, December 6, 2024, and January 24, 2025.

Case Studies, A/B Comparative Testing and Control Group O

Preamble

The following are the main lessons from the exploratory phase based on case studies:

- When measuring discoverability, it is important to distinguish between discoverability "factors" and discoverability "markers." Markers are ways of formulating or indicating that some form of discoverability **has occurred.** They generally relate to an objective, such as a click on a page, exposure on a results page, or even a ticket purchase. Factors relate more to **anything that can contribute** to discoverability, according to the best practices of a proposed discoverability strategy. Most of the discoverability measures we know of are discoverability *factors*.
- In this section, we examine these markers and factors to assess the extent to which best practices—and, in this context, particularly best practices for data structuring—impact the online discoverability of events.

The main overall research question for our hypotheses is thus: What is the impact of data structuring on the online "discoverability" of events?

Comparative Testing Groups

In order to measure the extent to which certain structured data properties in event data influence discoverability, we focused on three (3) scenarios:

- **Group O**: No structured data.
- **Group A**: Minimal data structuring, i.e. only information related to the date (`event:startDate`), location (`event:location`), and name (`event:name`) of the event.
- **Group B**: All elements from Group A (date, location, event name), as well as information about the artists (`event:performer`), including `sameAs` attributes, which allow an artist to be linked to external references. We chose to include the Wikidata link because it is a bridge identifier: entities described on Wikidata often have a multitude of links to persistent identifiers on other platforms. In fact, it appeared to us that although it is frequently overlooked, metadata identifying the main artist on stage (`event:performer`) provides valuable information about the content of a show and could therefore have a positive impact on the discoverability of performances through search engines. We therefore ensured that this data was included in the structured data of group B.

Each event in our sample was assigned to a group and was therefore subject to the associated scenario.

Event Sampling and Parameters Observed

In total, we counted 157 events[16] presented by three venues: Centre des Arts Juliette-Lassonde (CAJL) in St-Hyacinthe, Québec; Burlington Performing Arts Centre (BPAC) in Burlington, Ontario; and Sanderson Centre (Sanderson) in Brantford, Ontario.

These are all multidisciplinary venues. Some even have multiple performance halls. As a result, their respective programming is quite diverse in terms of discipline, target audiences (niche vs. general public), marketing efforts, etc.

Search engine exposure and search engine traffic are influenced by much more than just structured data.

The exploratory phase of this research revealed that **it was necessary to isolate the effects of other "external" variables**. In collaboration with presenters, we therefore identified and annotated certain parameters that were used to assign as accurately as possible a value to each event. The categorization choices were made by presenters. The groups were assigned programmatically to maximize the balance between groups O, A, and B for all parameters.

The following parameters were used:

PARAMETERS	VALUES USED
Type of audience	Niche General public
Venue capacity	Less than 50 50 to 100 101 to 250 251 to 400 401 to 700 700 to 1000 1001 to 1500 1501 to 2000 2000 to 5000 More than 5000
Significant publicity campaign during the observation period	Yes No
Number of days before the event	Whole number Number of days between the date programming is posted online and the date of the event in question

[16] See CAPACOA - Survey on structured data | All events.

Observation period	1-4: 4-week observation period, weeks 1 to 4 after the date programming is posted online 5-8: 4-week observation period, weeks 5 to 8 weeks after the date programming is posted online 9-12: 4-week observation period, weeks 9 to 12 after the date programing is posted online
Discipline	Theatre Music Song Circus Family Comedy Dance

Table 1. Parameters and values used in A/B testing.

Finally, efforts were made to ensure that events were distributed across each O/A/B group, while maintaining a balanced distribution according to the parameters used.

Additional filtering methods

In order to avoid the effect of extreme cases, we performed additional filtering after the initial analysis of the results. A few events were removed because they stood out significantly from the others due to circumstances unique to those events (e.g. exceptional press campaign throughout a tour, an outdoor festival rather than a single indoor performance).

Target Variables: Traffic and Discoverability on Search Engines

Two indexes were produced by A10s and Culture Creates during the project: a search engine exposure index[17] and a data structuring index, respectively.[18]

To assess the impact of data structuring, we examined the relationship between the structured data index and:

- the search engine exposure index (Google Search) for each event; and
- the number of views on each event page originating from a search engine (Google Search).

In addition, we used an explanatory model, which allowed us to isolate the impact of other factors that could influence these two variables.

[17] See previous section, “Defining the Query Basket and Harvesting Google Search Results”, Step 2

[18] See previous section, “Algorithm for Calculating a Structured Data Index”

What is an explanatory model?

A model is a tool that links factors (e.g. visibility, content) to a result (e.g. traffic) and is used to understand influences between factors.

A GLM (generalized linear model) is a type of model that measures the precise effect of one variable while taking other variables into account. The relationship between the data structuring index, website traffic, and discoverability on search engines was analyzed with a GLM model.

This type of model is useful because it clearly shows which factors have a real impact. While an A/B test is limited to a single comparison, a GLM can reveal the individual impact of several variables at once.

Exposure Index

This index was established according to the protocol developed by A10s and annotated by two people at Gradiant. Queries were formulated according to geographical context: event name + "*spectacle*" for events in Quebec, and event name + "show" for those in Ontario. The searches were conducted from the locations concerned (e.g. from Saint-Hyacinthe for the CAJL) to reflect local search conditions.

In addition, other queries were generated for further analysis as necessary, such as queries associated with the presenter. For each of the three presenters, two to three data harvests were carried out.

Presenter	Date 2025-2026 programming posted online	Date of first data harvest	Date of second data harvest	Date of third data harvest
Juliette-Lassonde Arts Center	April 30	May 17	June 7	July 5
Burlington Performing Arts Center	May 29	July 12	August 2	August 30
Sanderson Center	June 10	July 19	August 9	September 6

Table 2. Dates when programming was posted online and data harvested for each presenter.

Number of page views

Traffic was measured using Google Analytics data, isolating sessions whose source was Google Search (*sessionSource*) and whose entry page corresponded directly to the event page (*landingPage*). Only these visits were included. This data was extracted via the Google Analytics API.

Methodology Limitations and Analysis Decisions

The data structuring index was calculated on the basis of all events, but it only figured in the analysis during the actual data harvesting weeks. This is because there were delays between the data being posted online and it being properly structured. In the case of the Sanderson Centre, correctly structured data was only published from the eighth week onwards, after a general indexing error on the site, unrelated to our experiment, was corrected. Thus, only weeks 8 to 12 were considered valid for the Sanderson Centre when we examined data structuring and the number of page views.

For the CAJL, the main problem concerned the structuring of data related to artists (i.e. events in group B): the `performer.type` field systematically indicated `Person`, even when it was an `Organization` (such as a dance company). This error was corrected two weeks after going live, just before the first data harvest on May 17, by the CAPACOA team manually adding JSON-LD tags to the relevant pages. Another issue concerned the fact that there are several venues at the CAJL which were not correctly distinguished: the venues were all named “Centre des arts Juliette-Lassonde de Saint-Hyacinthe” and the persistent identifier that could have disambiguated them had been associated with the `PostalAddress` object rather than the `Place` object. This issue was resolved in early June, and we considered its potential impact to be relatively minor, given that the CAJL does not often present shows simultaneously in both of its venues. Therefore, data from all weeks observed for the CAJL were analysed.

Finally, an assignment error prevented BPAC events from being correctly assigned. There was therefore no control group for the BPAC. Consequently, when the analysis focused on groupings (A/B/O), only events at the CAJL and Sanderson Centre were included in the analysis. This limitation did, however, not prevent us from studying the correlation between the structured data index and website traffic. The analyses based on the structured data index included data from all three presenting organizations.

Main Findings

Preliminary observation leads us to believe that the results of queries for shows are subject to significant volatility due to certain factors that are difficult to include in the query selection protocol and that have a major impact on the results. The presence or absence of traditional advertising campaigns of any scale is one such factor.

In general, we can hypothesize that a larger sample size would reveal stronger underlying trends. Testing would benefit from being conducted longitudinally, a task that usually falls to an institute with a long-term mandate and sustainable funding. The recent emergence of chatbots adds a new layer of complexity, as can be seen in the “Satellite Results” section.

Nevertheless, this research has made it possible to define a granular typology of dissemination factors that should serve as a foundation for future research on the discoverability of the performing arts. See the “Methodology” section for more information.

In general, our findings indicate positive correlations between both the quality and quantity of structured data present on performing arts events’ web pages, online visits to these web pages, as well as their venue attendance statistics.

Exposure Index Results

Step 1: State of Web Referencing for Performing Arts Shows

The assessment conducted by A10s consultants made several observations about the composition of search results pages of queries related to the discoverability of shows.

The organic results from the queries reveal a diversity of players. Overall, one-third are directly from the professional performing arts value chain (venues, producers, presenters), one third are from media and aggregators (often linked to tourism), and the rest originate from various platforms. However, when only the top search results are considered, the professional sector dominates with more than two-thirds of the top results, a sign that their sites are well-structured and deemed credible by Google Search. Furthermore, the distribution of results varies little between the different versions of Google Search (French/English, mobile/computer).

The composition of search results varies considerably with the type of query. Social networks appear heavily in searches related to artists, presenters, or festivals, while they are practically absent in recommendation queries, which are dominated by media outlets and aggregators. The following chart illustrates the composition of the search results of different types of queries by website category.

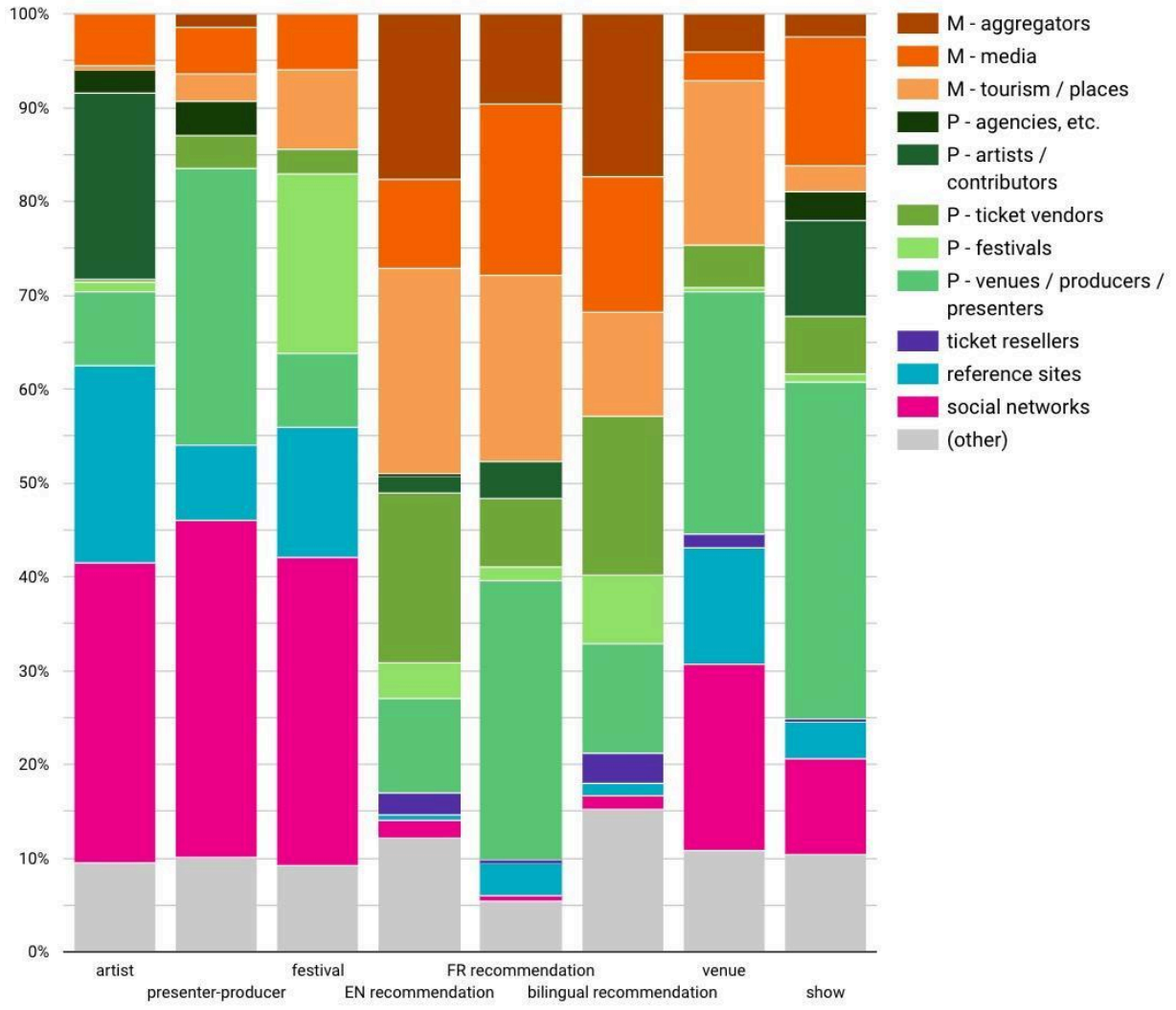


Figure 2. Comparison of the distribution of website categories in organic results for all queries, based on the most recent data harvest, by query type.

There are also notable linguistic differences. There are more ticket sellers and resellers in the results of English or bilingual queries, while presenter/producer sites are more visible in French queries. Except for a few local players, major international players, such as Eventbrite, Wikipedia, and Facebook, provide most of the prominent domains in search results. The Place des Arts and Mtl.org stand out, thanks to particularly optimized organic SEO.

Detailed results and additional observations on the enhanced results are presented in the report État des lieux du référencement web des arts de la scène.

Step 2: Development of a Strategy to Quantify Exposure

The results and analysis are presented visually in this report. The following example from that report shows that the exposure index of each presenter varies differently over time.

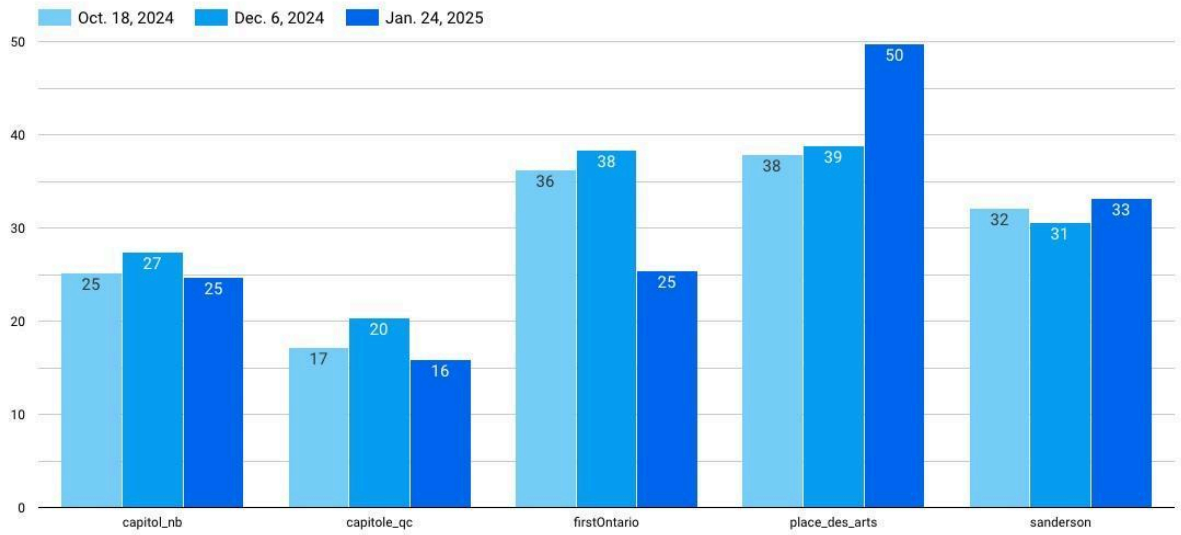


Figure 3. Exposure index by presenter on each data harvest date.

Analyzing search results this way presents several challenges. Some significant variations in the index for a given presenter cannot be easily explained by a change on that organization's website.

Preliminary observations suggest that queries associated with shows generate volatile results. Factors that are difficult to account for in the query selection protocol can have a major impact. For example, the strong showing in Place des Arts results in the last harvest can be explained by the selection of two shows (out of three) that were exclusive to this presenter. As there was no "competition" in the digital universe for these queries, the Place des Arts overall index shot up. For these reasons, an improved version of the index could potentially implement one or more of the following changes:

- Choose different criteria for selecting show queries.
- Increase the number of queries associated with shows to reduce the impact of individual shows on the index.
- Isolate queries associated with shows and create a subindex for this component (which would necessarily be volatile), along with a second subindex without the show criteria (the assumption being that it would be less volatile).

Furthermore, it should be noted that data was collected on only three occasions. A larger sample size would presumably reveal stronger underlying trends.

We nevertheless believe that this approach demonstrates that it is possible to develop a quantified exposure index from the perspective of a show presenter. The index can be used for several purposes.

Observing the index over time for a given presenter could reveal the results of deliberate efforts or unexpected changes in the environment. When redesigning a website, for example, the index can be used as a tool to ensure that exposure is not negatively affected by the changes, or conversely, to confirm that the new site is generating more exposure. This potential use was observed during work surrounding the launch of a new website for Capitol Theatre in Moncton, New Brunswick.

The index could also be used to observe changes in the visibility of a given cohort of presenters. The capture scoring method used to calculate the index could also be used independently to perform analyses (e.g. A/B testing on discoverability efforts for a batch of shows).

More detailed results are presented in the report Élaboration d'un indicateur de rayonnement du spectacle sur les moteurs de recherche.

Second chances, by RUBBERBANDance Group (Q16640268, K10-122). This dance show presented at Centre des arts Juliette-Lassonde was part of the B group in our A/B testing. Photo credit: David Wong.

Results of A/B Testing and Web Traffic Analysis

Descriptive analyses

A/B comparative tests conducted by Gradiant were used to observe search engine traffic over several periods. A daily observation cycle left too much room for traffic spikes. Observation cycles covering several weeks smoothed out sudden effects that can occur daily. We chose to harvest data over four-week periods so that the search results on the impact of data structuring on traffic would align closely with the results on the impact of the exposure index.

Discoverability on search engines

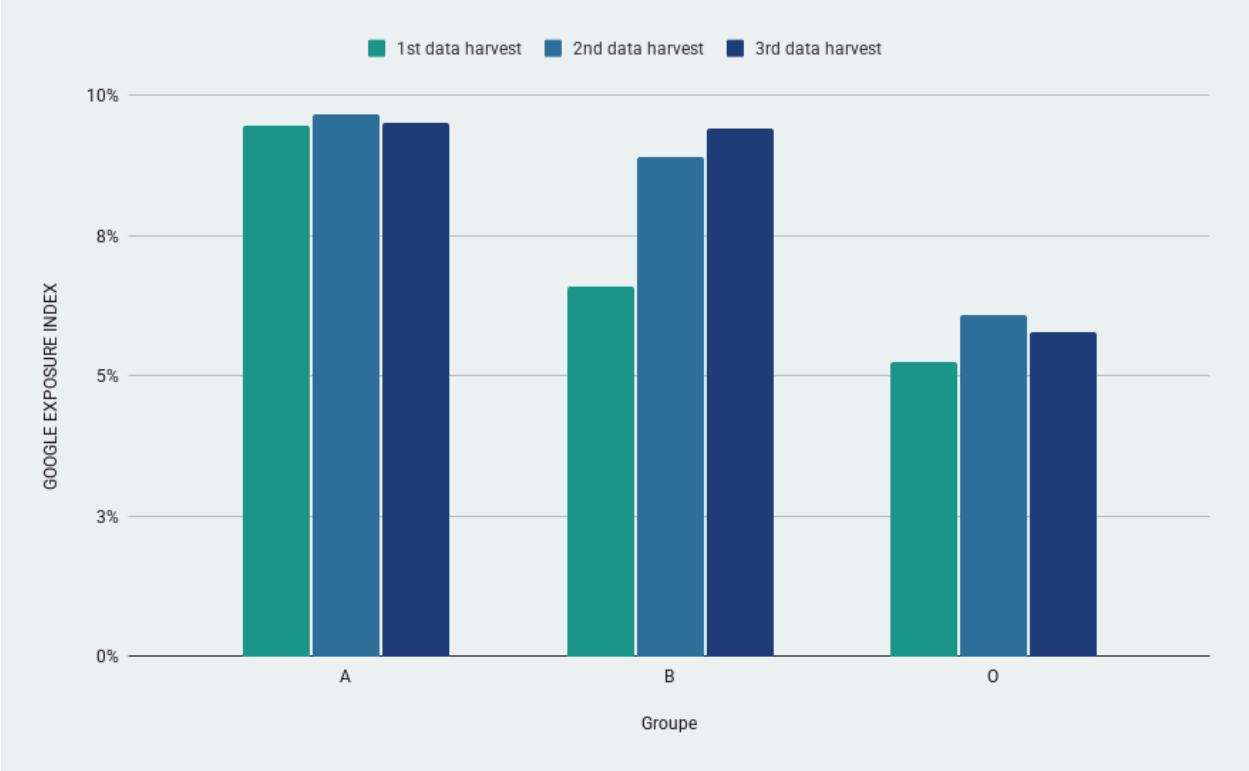


Figure 4. Average exposure index by group (A - enhanced metadata; B - addition of performers; O - control group - no markup) for the presenter Centre des Arts Juliette-Lassonde, by harvest date.

Both analyses tend to confirm our hypotheses, but in a nuanced way. For the exposure index, groups A and B show better results on average than the control group for all data harvests. In the case of group B, the increase observed in the second likely corresponds to the moment when microdata was correctly deployed on CAJL's website. See the "Methodology Limitations" section for more information.

Search Engine Traffic

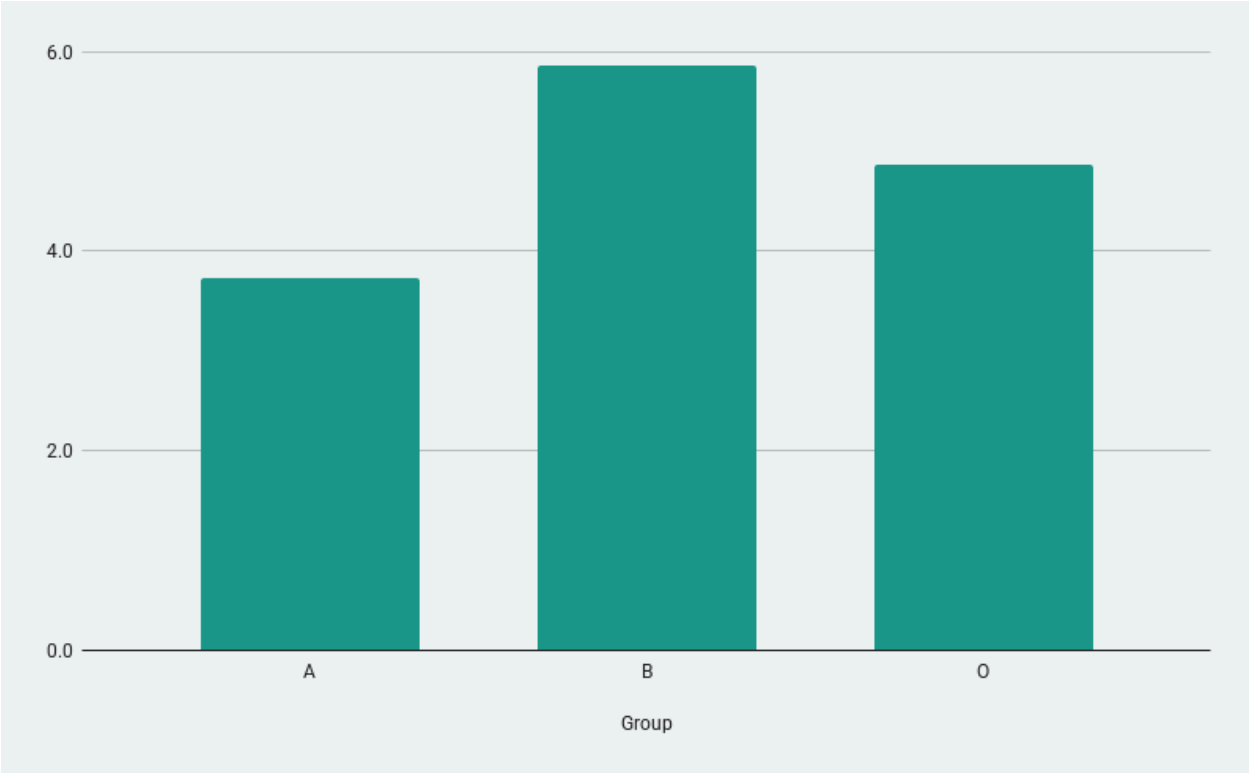


Figure 5. Average views over four weeks for presenters Juliette-Lassonde Arts Center and Sanderson Center for groups A - enhanced metadata; B - addition of performers; O - control group - no markup.

However, the traffic data show a different dynamic: while group B has a higher average number of page views than the control group, group A's events have lower results. This situation could be related to a sampling bias in the composition of the groups, but it also reminds us that **measured exposure does not automatically translate into website traffic**.

These discrepancies can be explained in part by the methodological limitations of our indicator. The exposure index is based on only one query per event and corresponds to a ratio that is sensitive to the total number of results on a page. Thus, the same number of results related to the presenter will generate a lower index when the page contains more links. In addition, the traffic observed may come from queries other than those selected, which introduces a discrepancy between measured exposure and actual traffic.

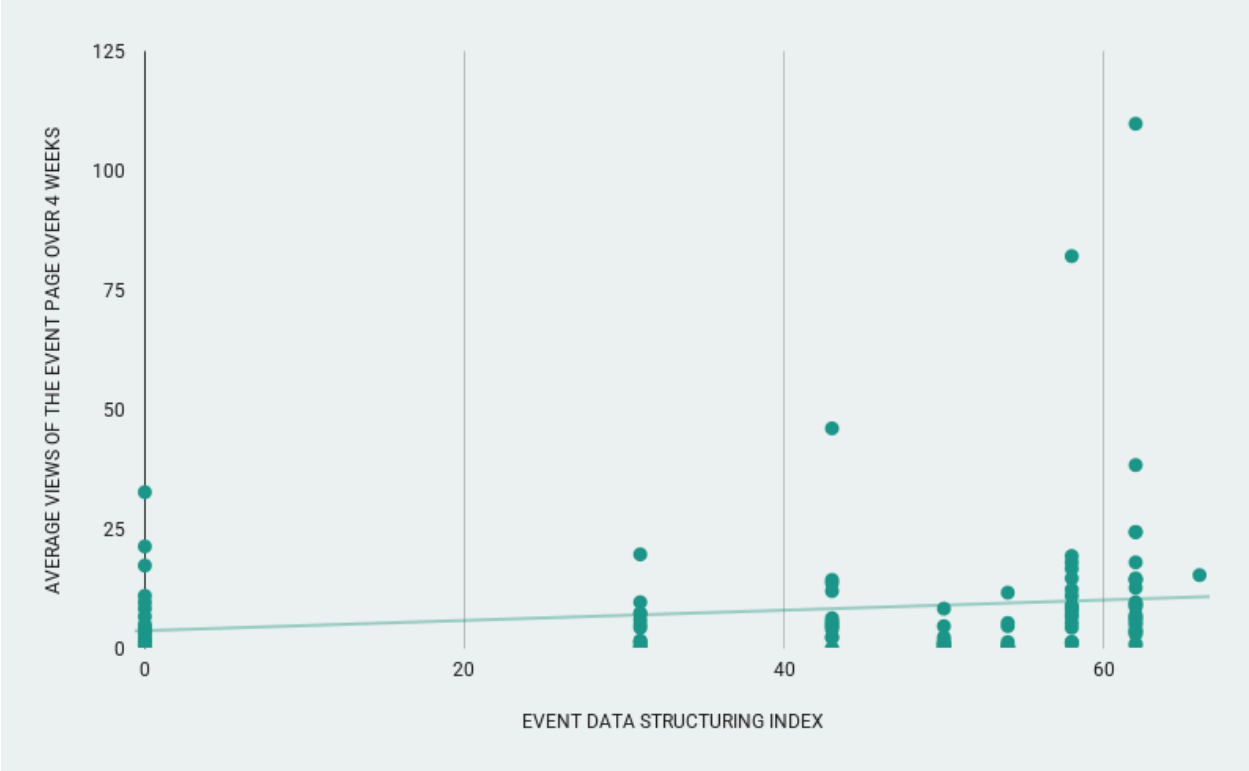


Figure 6. Relationship between the data structuring index and average views over 4 weeks. Reading note: Each point corresponds to an event observed over a four-week period. For example, the point located at 58 on the horizontal axis and at 82 on the vertical axis means that the observed event has a data structuring index of 58 and received an average of 82 views over a four-week period.

Effect of Structured Data on Event Page Traffic

Model Performance

The model is not designed to predict with any accuracy the number of views of a web page. It is rather an explanatory model that allows us to understand general trends. The average estimation error is approximately ±5 views per page over a four-week period. It is important to note that traffic to event pages from Google Search remains low overall. Search engines now expend considerable efforts to provide answers directly on the results page. Our data reflects this trend toward “zero-click” answers.

The model's main value lies in its ability to highlight the relationship between certain structured data (for example, data describing the contents of the show) and traffic from search engines. The model confirms a positive correlation and an encouraging trend, which is consistent with this research’s hypotheses.

Another important point is that web page traffic does not follow a normal distribution (see Figure 7). Rather, it resembles dissemination with a “long tail”, where a few pages attract a lot

of visits while most pages receive little traffic. To take this particularity into account, we used an adapted distribution function, the Tweedie law with a log parameter of 1.5.

Finally, after several selection steps, we arrived at the following explanatory variables in the model:

- Discipline
- Audience type
- Significant campaign in the observed time period
- Venue capacity
- Number of days before the event
- Group
- Structured data index
- Presenter
- Period (4-week period, which can be 1-4, 5-8, or 9-12)

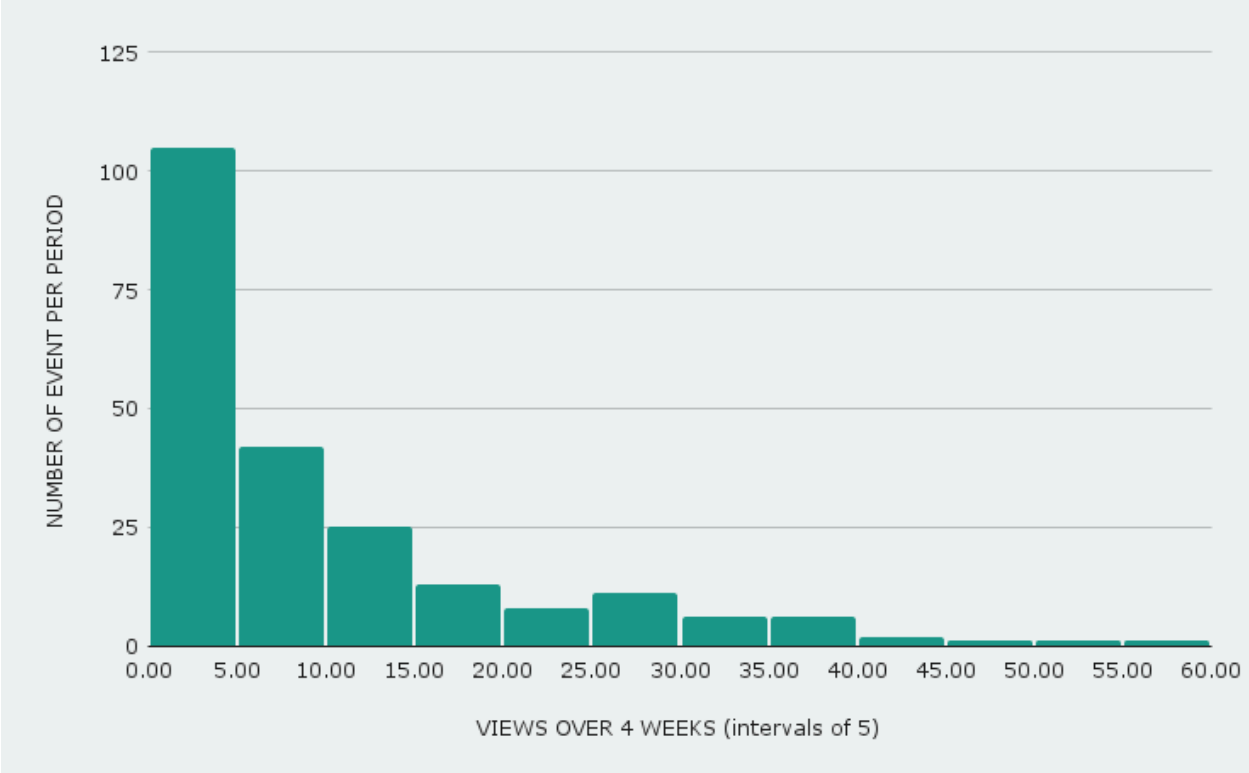


Figure 7. Histogram of page views by 4-week period (excluding extreme values). Reading note: For example, the first bar indicates that 105 events received few views (between 0 and 5) during a 4-week period. In contrast, 1 event received more than 55 views.

Results Observed

Our analysis confirms that the quality data structuring (the structured data index) has a direct and measurable impact on event visibility. On average, **each 10-point improvement in the data structuring index leads to an increase of approximately 3% in page views from Google Search over four weeks** (see Figure 6). This effect is even more pronounced in certain situations: it can almost double between the 9th and 12th week after the event is posted online (+6.3%) and remains particularly favorable for niche music (+2.6%) or concerts in medium-sized venues with 500 to 1,000 seats (+2.3%). Promotional campaigns, when supported by good data structuring, also amplify this effect (+2.5 to 2.9%).

For reference, data that is important for disambiguating locations (`location.address.postalcode` and `location.sameAs`) has a value of 4 points on the data structuring index. Recommended data (like `description`, `image`, `performer`, and `offers.url`) has a value of 2 points. Optional data (like `duration`, `endDate`, `eventAttendanceMode`) is worth 1 point. Thus, an event with only the minimum data would score 28, while an event with data for both disambiguation and recommended properties would score 62.[19]

However, impact varies with the context. The effect of the index is less pronounced or even slightly negative in certain cases, such as when an event is posted online 70 days or more in advance (-3%). Negative impacts are also observed for theater performances (-2% to -3%) in certain periods or venues, and for dance (approximately -1%), or family events in small venues.

Do specific properties have a positive impact?

An important question was whether or not specific properties of events could have a positive impact on page views. These properties are not directly included in the main model. We proceeded to examine them separately, making comparisons by group and using data from the CAJL and the Sanderson Center.

As mentioned above, the number of page views has a "long tail" distribution rather than a normal distribution. Given this fact, a simple comparison of averages is not meaningful. We therefore used a Wilcoxon rank sum test, which compares medians rather than averages and is better suited for this type of data.

The results indicate a generally positive difference for events with structured data. In some cases, the observed effect is marginally significant. However, one striking point did emerge: the presence of a description has a statistically significant effect on traffic, confirming the central role of event descriptions in the attractiveness of pages to Google Search.

---

[19] For full details of structured data properties and their scores, see the README file for the structured data index: https://github.com/culturecreates/artsdata-score

With minimum data properties	Average visits over 4 weeks	Median visits over 4 weeks
No	5	3
Yes	9	5

With the `organizer` property	Average visits over 4 weeks	Median visits over 4 weeks
No	4	2
Yes	12	7

With the `description` property	Average visits over 4 weeks	Median visits over 4 weeks
No	4	2
Yes	12	7

With the `Performer` property and `sameAs` link	Average visits over 4 weeks	Median visits over 4 weeks
No	7	4
Yes	11	6

Study of Two Comparable Events

About the Events

To illustrate our overall results, we conducted a more targeted analysis of certain events that we deemed comparable. Given that the results showed that the difference was particularly noticeable for music performances (concerts), we chose to compare two music events. Both events took place at the Sanderson Centre just one month apart (November 4 and December 4), and featured Canadian country artists, James Barker Band and Dallas Smith. These two artists enjoy a similar level of fame, with approximately 400,000 and 300,000 subscribers respectively on Spotify, as of the publication date of the report.

Given their similarities in musical genre and audience, they made good candidates for a controlled comparison. James Barker Band serves as the control group (O), and Dallas Smith represents group A.

Results

The Sanderson Centre's data is of particular interest as it allowed us to assess the page view fluctuation before and after structured data was posted online, as one of the two events had structured data from the start and the other event had structured data added as of Week 8 of the data harvest.

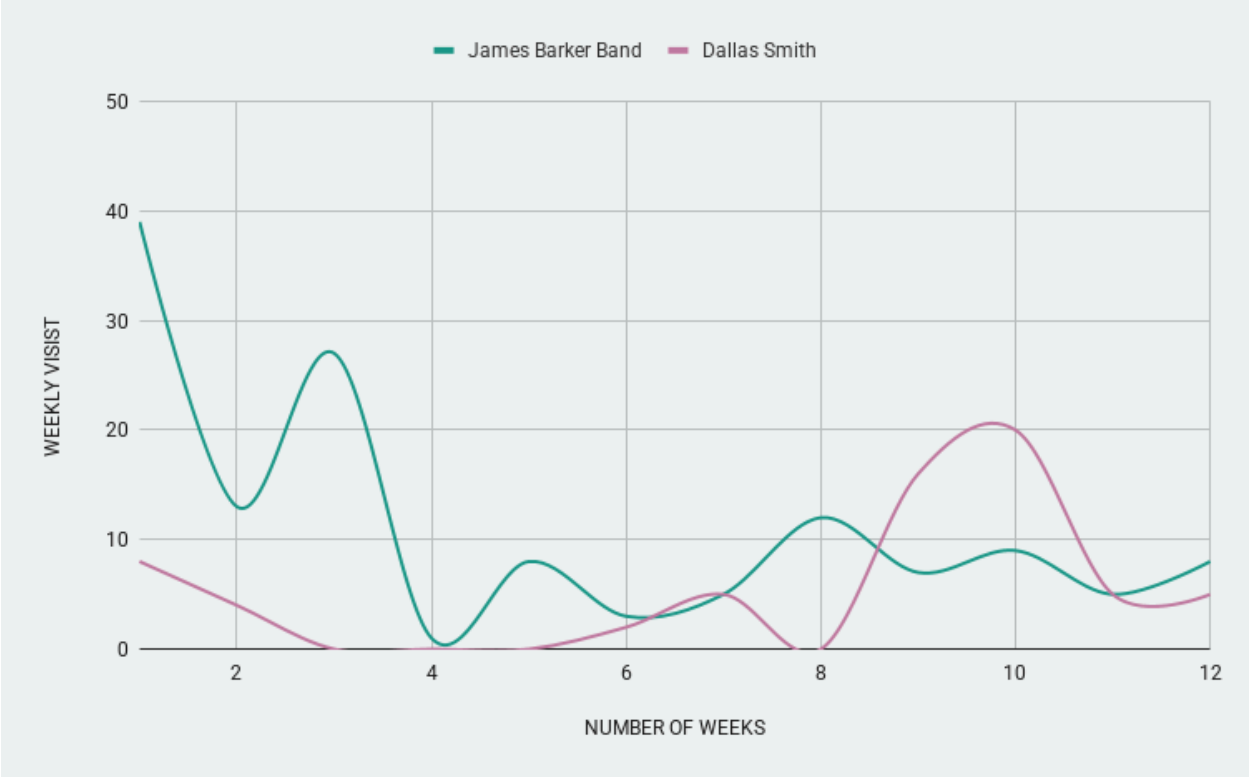


Figure 8. Diagram of the number of page views per week for two shows. The James Barker Band curve has structure data for the entire 12 weeks. The Dallas Smith curve has structure data only as of the 8th week.

For the James Barker Band, the large number of page views when the data was first posted gave way to a sharp decline between Periods 1-4 and 5-8, before smoothing out in Period 9-12. Overall, the trend is downward, with no sign of a significant recovery.

The dynamic is different for Dallas Smith, where after low levels, there was a jump in Period 9-12, reaching 46 page views.

Period 9-12, the first period for which Dallas Smith has structured data, provides a reliable glimpse at the difference structuring data can make. The fact that the Dallas Smith peak occurs precisely at this point is an illustration of this.

Satellite Results

Ticket Resale Sites

The overview of web referencing for the performing arts revealed the space occupied by ticket resale sites. Overall, across all the results analyzed for all queries, these sites represent an average of 1.2% of the results.

However, there are large variations depending on the type and language of the queries. For queries associated with recommendations (i.e. those that do not mention a specific artist, show, or venue) and that are in English or use bilingual terms, ticket resale sites constitute 2.6% of search results, more than double their usual average. In contrast to these sites, the Billets.ca site, which accounts for 50 to 70% of the space occupied by resellers, seems to follow a different logic, in that it performs significantly better than others in searches focused on specific shows.

We have not analyzed the differences based on language, although we can make a few assumptions in this regard:

- The resale market is dominated by foreign players who are better referenced in English than in French.
- The local French-speaking market is less attractive to ticket resellers because of its size, so they put less effort into it.
- Quebec presenters and their ticketing platforms use more structured data or other referencing strategies, and are therefore more resistant to being pushed back in search results by resellers.

Chatbots

As we pointed out in the methodology section, our starting premise was to base our observations on Google Search as a “barometer” of the digital environment. We also stated our initial assumption that *search engines create syntheses of the massive amounts of data published online while aligning themselves as closely as possible with consumer expectations and intentions.*

However, searches with robots and chatbots created using advances in artificial intelligence, such as large language models, deep learning, and machine learning, are all based on the same assumption: answering a question by synthesizing the analysis of massive amounts of data.

At the end of our research, we therefore felt it necessary to explore the possible impact of these new tools on our initial hypotheses and questions, and to consider the impact of AI on internet users' event search habits.

Studies already suggest that the introduction of generative AI in search results pages has a significant impact on search behavior, including reducing the volume of clicks to cultural organization websites.

> *AI-generated summaries sit right at the top of results, pulling together information from multiple sources and answering questions instantly. They often give people what they need without ever needing to visit a website (...).*[20]

We conducted three summary tests in four AI applications – Mistral, Gemini, Leo-Brave, and ChatGPT – and noted our observations in the following supplementary document: Robots conversationnels et Scénarios utilisateurs | Tests sommaires 2025.

Our questions sought to evaluate the quality of responses to three queries:

1. *SVP Indique-moi les spectacles à Montréal entre les 6 et 13 juin 2025.* (Please tell me about shows in Montréal between June 6 and June 13, 2025)
2. *Puis-je voir une pièce de théâtre avec Geneviève Brouillette cet automne au Québec ?* (Can I see a play starring Geneviève Brouillette this fall in Quebec?
   Note: we were looking for a recommendation for the play *Peut contenir des traces d'ego* at the Centre des arts Juliette-Lassonde on October 25, 2025, an event that was part of our basket of queries.)
3. *L'IA utilise-t-elle les données structurées Schema.org ?* (Does AI use Schema.org structured data?)

Our initial results show that AI performance varies greatly depending on whether the application performs real-time online searches or relies solely on initial massive training data.

The former, ChatGPT and Gemini, perform very well, while the latter, Mistral and Brave, perform much less well. The latter do not use Schema tags, while the former systematically use them in a documented manner.

Another interesting finding is that applications performing real-time online searches now share their sources with us.

As such, the official websites of presenters, regional cultural and tourist information sites, general media outlets, and the event recommendation site La Vitrine can all appear in chatbot results. Ticket sales sites seem to play a marginal role. The weighting of the sources harvested could be the subject of further effort and research.

When we asked chatbots "Does AI use Schema.org structured data?", Google AI Pro's answers confirmed our starting assumptions and the results of the current project:

> *Yes, Google's AI systems, including language models such as Gemini (which Google AI Pro refers to), use Schema.org structured data. This data plays a crucial role in how AI models understand and process information from the web.*

[20] See Downs and Unitt 2025, a survey of 100 cultural sites in the United Kingdom.

> *Yes, Google's AI systems use Schema.org's **`performer`** property to identify the artists participating in an event. This tag is part of the **`Event`** structured data, which is essential in helping Google understand the content of your web page and present it in an enriched way in its search results*.[21]

ChatGPT even provided us with sample markup code to use for an artist participating in an event (see Appendix).

In general, the tests suggest that online event searches could quickly shift to chatbots and that enhanced and structured data is already being put to good use:

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

Nevertheless, important questions remain: Should we accept robots' responses to our queries without conducting independent research, and could the responses to these queries contain biases that only independent research could detect?

Communication Activities

Our general conclusions and results are recorded in a summary slideshow that was presented at the Matinée numérique event on September 18, 2025.[22,23]

A second communication activity was held on October 15, hosted by CAPACOA.[24]

[21] Prompts made on September 8, 2025. For more information, see Robots conversationnels et Scénarios utilisateurs | Tests sommaires 2025

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

[23] See the slideshow Mesure de la découvrabilité | Matinées numériques 20250918

[24] See https://www.artsdata.ca/en/events/structured-data-a-powerhouse-for-event-discoverability-oct-15

Recommendations and Resources

Given the findings of this research and with a view to promoting the discoverability of performing arts events, stakeholders who present performances or who contribute to them as lead artists or production organizations would reap discoverability benefits by adopting the following effective data management practices:

- Tag event pages using Schema.org structured data to ensure that search engine crawlers and chatbots have access to it.
- Ensure that structured data contains at a minimum all information needed to distinguish between similar events—the show name (`name`), venue (`location`), and date and time (`startDate`) of the event, by providing all address details and a persistent identifier (`sameAs`) for the specific hall in a building with more than one performance space.[25]
- Pay close attention to metadata describing the content of the event, specifically the `description` property.
- Use persistent identifiers (`sameAs`) to describe and identify the main contributors of the event – the presenter (`organizer`) and the performer (`performer`).

Moreover, integrating enhanced structured data into a website does not only meet the requirements of discoverability via search engines and chatbots. High-quality structured data can be harvested and reused by other websites, such as open-data powered cultural calendars.[26] This can lighten the data entry burden for cultural managers, who would otherwise have to publish the same information on numerous sites. Structured data can also be harvested and preserved for archival purposes, a service provided by the Artsdata knowledge graph. Without such archiving, information about performances often disappears along with the web pages where it was displayed, depriving large language models of valuable sources of information for developing knowledge about the performing arts.

Although this study focused only on event data, and since data describing content and artists affects the discoverability of events, artists and arts companies would likely benefit from the use of structured data on their own websites to describe themselves and their works. By doing this, they would assert their authority over the data concerning them and exercise control over it. It would allow web crawlers to link event data to creators and works.

Finally, A/B testing suggests that the presence of Wikidata persistent identifiers for the venue (`location.sameAs`) and the performer (`performer.sameAs`) contributes to the discoverability of live performances. However, editing Wikidata items doesn't come easy and requires specialized knowledge far more advanced than for Schema structured data. Obstacles to adoption can be overcome in two ways: with the help of consultants specializing in open data

[25] For best practices regarding the location property, see https://culturecreates.github.io/artsdata-data-model/location.html

[26] As an example, ArtsCultureNB's events listings are powered by open data pulled from Artsdata. https://artsculturenb.ca/en/performing-arts/

or by joining an association that offers its members the opportunity to publish their information in Wikidata. Many associations in Canada have already made one-time data drops in Wikidata and Artsdata; some even have continuous publication processes.[27]

To help arts organizations implement these recommendations, the Artsdata team published several useful **resources** to help cultural managers add structured data markups into web pages:

- Structured data templates including all the enhanced metadata identified in this report.
- Documentation pages detailing how to correctly fill in each property.
- Video tutorials illustrating how to retrieve persistent identifiers.
- Database of persistent identifiers for more than 10,000 individuals, organizations, and performing arts venues in Canada.[28]
- On-demand digital discoverability services.

To get started with data structuring, we recommend reading Artsdata's **introduction to structured event data**[29] and **introduction to persistent identifiers**.[30]

[27] CAPACOA, *Artsdata Annual Report 2023-2024.* Available at https://www.artsdata.ca/en/news/constellation-of-data

[28] See, for example, this query of artists and their persistent identifiers in Artsdata graph: https://s.zazuko.com/36BcdkR

[29] See https://www.artsdata.ca/en/resources/structured-data

[30] See https://www.artsdata.ca/en/resources/identifiable-and-findable

Conclusion

According to the results of this study, it appears that certain structured data, as represented on Google's results pages, are in line with Internet users' expectations when it comes to searching for performing arts events to attend. In this sense, they would therefore be drivers of discoverability.

Most notably, a good event description is paramount, and so is the identification of the performers and the organizer.

These recommendations apply to the discoverability of events via Google Search in 2025. They are also relevant for large language models, considering what we currently know about generative search optimization (optimization techniques used for generative engines such as Copilot and Gemini), and our own tests with chatbots.

Generally speaking, our findings suggest that online event search could quickly be taken over by chatbots. The good news is that enhanced structured data appears to be already being put to good use by some chatbots. This type of data is a good way of guaranteeing the quality of AI-generated responses.

The high volatility of our research results highlights the importance of conducting this independent research continuously over time.

The effort to tag online assets using structured data by presenters and stakeholders in the performing arts value chain should be supported and sustained. Obviously, this is a challenge that adds to the already heavy workload of those directly concerned and raises the question of the possible need to create new skills within work teams.

In the short term, the use of structured data to promote performing arts events is certainly a sectoral issue, but it is also becoming a public policy issue. Its resolution will likely require an office, a council, a think tank, and/or a statistical institute with consistent funding. It is now a matter of cultural diversity and sovereignty.

References

A10s and 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-LOObj8egpIVcNDeExzg2y1vF5D10uFpBhpOSJFo3M/edit?tab%3Dt.0&sa=D&source=docs&ust=1756999315420469&usg=AOvVaw3o7-TtrwoI4DstfahZrBQk

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

CAPACOA, Anicet Lepetit Ondo, Lucie Vézina-Topping, and 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, and Gemini 2.0. 2025. ArtIA - *Rapport intérimaire. Sporobole*, Société des arts technologiques (SAT), Collective project. https://praxis.encommun.io/b/artia/n/SMHwIKCGwbdOKHjwAqe0j6rNrBg/

Downs, Beth, and Chris Unitt. 2025. *The Impact of AI Overviews in the Cultural Sector*. United Kingdom: 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, and Frédéric Julien. 2019. *Lier l'avenir numérique des arts de la scène : Comment mobiliser les synergies de la chaîne de valeur*. Canadian Association for the Performing Arts (CAPACOA) in cooperation with Bern University of Applied Sciences. https://capacoa.ca/documents/research/ANL_Rapport_2019.pdf

gradiant, and Artsdata. 2024. *Vers un score de découvrabilité : modélisation explicative multifactorielle.* CAPACOA.

gradiant, and 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_BilanGradiant_2023-08-24_VF.pdf

Guillon, Vincent, and 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, and Inga Petri. 2017*. La numérisation des arts du spectacle.* CAPACOA et Strategic Moves. http://bit.ly/numartsfr

Le Gall, Anne, and 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), and 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 Bisaillon, Guy-Philippe Wells, Martin Tétu, and 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 Bisaillon, Guy-Philippe Wells, and Danielle Desjardins. 2019. "Définir la découvrabilité." In *Découvrabilité des produits culturels en ligne. Un objectif prioritaire pour la coopération bilatérale France-Québec.* Report prepared for the Ministère de la culture et des communications du Québec and the 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_coope_ration_bilate_rale_france-que_bec_2019.pdf

Roberge, Jonathan, Georges Azzaria, Guy Bellavance, and 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, and Philippe Michon. 2017. *État des lieux sur les métadonnées relatives aux contenus culturels.* Québec: 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.

APPENDIX – Examples of structured data

See the Artsdata resource page on the use of structured data:
https://www.artsdata.ca/en/resources/structured-data

Here are two examples of recommended markup that optimizes structured data on a target web page:

Structured Data for Centre des arts Juliette-Lassonde

Sample markup code for Centre des arts Juliette-Lassonde. In this example, all of the **minimum data** and enhanced data discussed in this report has been entered. Of particular note, “Place” appears as `@type` for both the venue and the building in which it is located.

```
<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-1280x72
0.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: This code differs slightly from the markup code that appears on the website of Centre des arts Juliette-Lassonde. In addition, the Centre uses a microdata encoding format different from the JSON-LD format used above. The JSON-LD format has the advantage of being more commonly used and easier for humans to read.

Structured Data for the Sanderson Centre

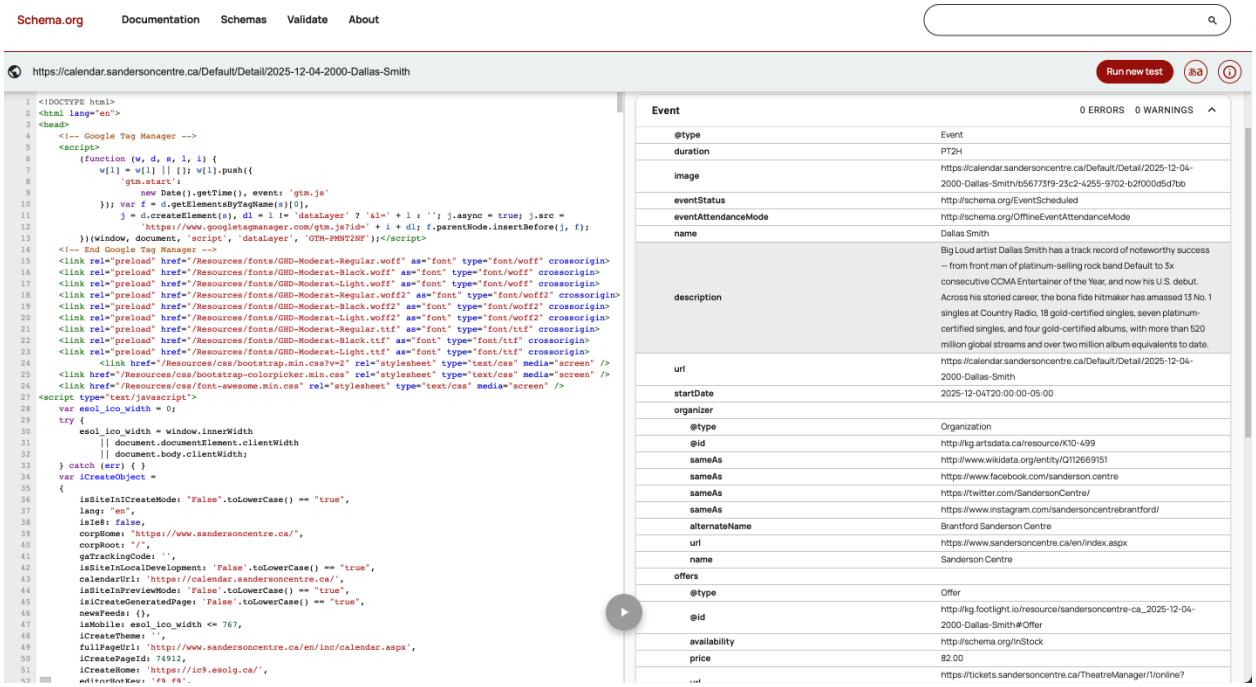
Here is an example of structured data markup for the Sanderson Centre, for the performance of Dallas Smith (cited in the “Study of Two Comparable Events” section). Of note, the `location` and the `organizer` entities are very clearly identified.

```
<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-Dal
las-Smith#WebPage",
      "@type":"WebPage",
      "inLanguage":"en",
"url":"https://calendar.sandersoncentre.ca/Default/Detail/2025-12-04-2000-Dal
las-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-D
allas-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-Dal
las-Smith",
   "startDate":"2025-12-04T20:00:00-05:00",
   "endDate": null,
   "workPerformed": null,
   "performer": null
 }
</script>
```


Here is a screenshot from the same structured data, as viewed in the Schema.org Validator:



Structured Data Proposed by ChatGPT

This example of **markup code** generated by **ChatGPT** for an artist participating in an event. As can be seen from the highlighting, it appears to validate most of our research findings, although ChatGPT neglected to integrate persistent identifiers with the `sameAs` property.


```
<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"
    }
  },
  "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"
```