This is an old version; current Working Document: Linked Open Event Data - Vision & Roadmap

Linked Open Event Data - Vision & Roadmap

Beat Estermann, Bern University of Applied Sciences, March 2023

Table of Contents

This is an old version; current Working Document: Linked Open Event Data - Vision & Roadmap

1. Introduction

The **visibility of events in the digital space** is gradually improving. There are various initiatives in Switzerland to pool event data on central data platforms, such as cultural calendars, ticketing platforms and the like. While the integration of this data still involves manual effort in many cases, various players have recognized the potential that lies in automatizing the data transfer.

Yet, like other countries, Switzerland still lacks a comprehensive, shared data infrastructure that allows for the federation and sharing of data from a maximum of event organizers and event platforms. Such a **shared data infrastructure would serve as the basis for a variety of digital services**, such as cultural calendars, recommendation systems, search engines, personal assistants on smartphones, archive and documentation databases, or entries in the free online encyclopedia Wikipedia.

A distributed data infrastructure based on the **FAIR principles (Findable, Accessible, Interoperable, Reusable)** would have the advantage for producers, organizers, and other stakeholders along the value chain that they only have to enter their data once, without having to deal with the lock-in problems resulting from dominant event and ticketing platform providers that rely on a proprietary approach towards data to make their business models work.

Linked data is the technology of choice for linking data from different sources, so that they can be used by a variety of services as if they came from the same data source. Important prerequisites for the publication of data as linked data are a **shared data model** as well as the **use of the same reference databases** (e.g., base registers for persons, organizations, event venues, etc.).

The present document sets out the vision and a roadmap for the establishment of a Linked **Open Event Data** ecosystem in Switzerland. The remainder of the document is structured as follows:

- Section 2 lays out the vision and the architecture of the Linked Open Event Data ecosystem.
- Section 3 contains the proposed roadmap for the roll-out of the Linked Open Event Data ecosystem, with a focus on Switzerland, thereby taking stock of the current state of implementation in the different areas.
- Section 4 contains a set of organizational maturity models which complement the roadmap and reflect the point of view of individual organizations and provide guidance on their implementation of Linked Open Event Data.
- Section 5 contains some indications regarding the process in view of the further development of this document as well as potentially relevant topics that have not been covered yet in the document.

The list of authors and reviewers can be found at the end of the document, in section 6.

2. Linked Open Event Data – Vision and Architecture

2.1. Our Vision

The vision behind Linked Open Event Data is simple and can be summarized in the form of three principles:

1. The different stakeholders that are part of the value networks around cultural events[1] need to enter their data only once; it is then shared with the different players across the sector provided that the people and organizations concerned wish to do so (**once only**)[2].

2. Data and content are freely shared to maximize their accessibility, findability, and re-usability across the value network. As a rule, data is published under the CC-0 Public Domain Dedication[3], and content is published under a free copyright license, such as CC BY[4] or CC BY-SA[5] (**open data by default**). Copyright restrictions, privacy considerations, as well as business secrets apply and need to be taken into account as part of the data publication workflows.

3. To increase its interoperability and re-usability, the data is published according to a shared data model and by linking to the same reference databases (**interoperability by default**).

2.2. Alignment with the Tallinn Principles

In general, the Linked Open Event Data ecosystem is expected to adhere to the principles set forth in the **Tallinn Declaration on eGovernment**[6]. The Tallinn Declaration contains five central principles and several accompanying measures that serve as a common orientation framework for eGovernment development in EU and EFTA states, both nationally and internationally (European Union, 2017, pp. 3ff):

1. In order to increase the efficiency and user orientation (**"user centricity"**), the services of public administrations should be digital, inclusive, accessible and barrier-free by default (**"digital by default"**, **"inclusiveness"**, **"accessibility"**).

2. For central services of the public sector, the principle applies that citizens or businesses need to provide the same information to the public administration only once (**"once only"**); the administration takes care of the internal transfer of the information at the behest of citizens or businesses.

3. Public services must be trustworthy and secure (**"trustworthiness"**, **"security"**), whereby privacy is to be respected and national eID solutions are to be used.

4. In the spirit of openness and transparency, citizens and companies should be enabled to manage data concerning them in base registries themselves and to exercise a certain degree of control

[1] See Estermann B., Julien, F. (2019): A Linked Digital Future for the Performing Arts: Leveraging Synergies along the Value Chain. Canadian Arts Presenting Association (CAPACOA) in cooperation with the Bern University of Applied Sciences, as well as Estermann, B. (2020): Creating a Linked Open Data Ecosystem for the Performing Arts (LODEPA). Arti dello Spettacolo / Performing Arts, Year 6, Special Issue, pp. 31-49, for model of the value network with its different stakeholders in the performing arts domain.
[2] In the Tallinn Declaration on eGovernment, this principle is referred to as the “once only” principle.
[3] https://creativecommons.org/publicdomain/zero/1.0/
[4] https://creativecommons.org/licenses/by/4.0/
[5] https://creativecommons.org/licenses/by-sa/4.0/
[6] https://www.newsd.admin.ch/newsd/message/attachments/49838.pdf

over their use. Furthermore, administrative data that is useful for the economy or society should be made available as open data for free further use **("open by default"**) and linked to reference databases (**"automatic linkages to databases"**). In addition, long-term digital preservation of public sector data and information must be ensured.

5. IT applications of the public sector should be interoperable **("interoperability by default"**). In addition, open source solutions and open standards should be used to increase the reusability of IT applications – both within the public administration and beyond (**"make ICT solutions available for reuse"**). When setting up service infrastructures, duplications should be avoided (**"avoid sectoral duplication of service infrastructures"**).

Several accompanying measures are intended to promote the digital transformation of the public sector. These include the promotion of **"digital skills"** within the public administration at all levels and **"digital leadership skills"** among top civil and public servants. Further accompanying measures are broader use of data by the public administration itself (**"widen and deepen the use of data and analytics"**), faster innovation cycles (**"more and faster innovation"**) and **"agile ways of developing and deploying digital technology"**.

For an overview of the state of implementation of the Tallinn Principles in Switzerland as of 2021, see Marti et al. (2022)[7]. In view of the establishment of the Linked Open Event Data ecosystem, these principles shall apply not only to the infrastructures and services provided by government agencies themselves, but also to infrastructure elements and services that are developed and sustained preponderantly thanks to public funds.

As regards accessibility, both the **accessibility of web content**[8] as well as the provision of **information about the accessibility and inclusiveness of events**[9] themselves are of interest.

These principles are complemented by **ethical considerations** that need to be further specified. The CARE principles for Indigenous Data Governance (Collective Benefit, Authority to Control, Responsibility, Ethics)[10], the Code of Conduct for Trusted Data Spaces[11], as well as the principles of the Digital Public Goods Alliance[12] provide a good starting point when it comes to defining the respective governance criteria.

Furthermore, in order to secure the **economic sustainability** of the emerging Linked Open Event Data ecosystem, it will be important to ensure that those who make important contributions to the shared data infrastructure will also experience its benefits.

[7] Marti, M., Estermann, B., Neuroni, A. (2022): Die Tallinn-Prinzipien als Wegbereiter der digitalen Transformation des öffentlichen Sektors in der Schweiz. In: Mertes, A., Pleger, L. E. (2022): Digitale Transformation der öffentlichen Verwaltung in der Schweiz. Wiesbaden: Springer Gabler.
[8] See for example the WCAG 2.1 standard.
[9] See for example the accessibility information page of Theater Basel or the accessibility app provided by Ginto.
[10] https://www.gida-global.org/care
[11] Under development (OFCOM / DFAE).
[12] https://digitalpublicgoods.net/

2.3. Architecture

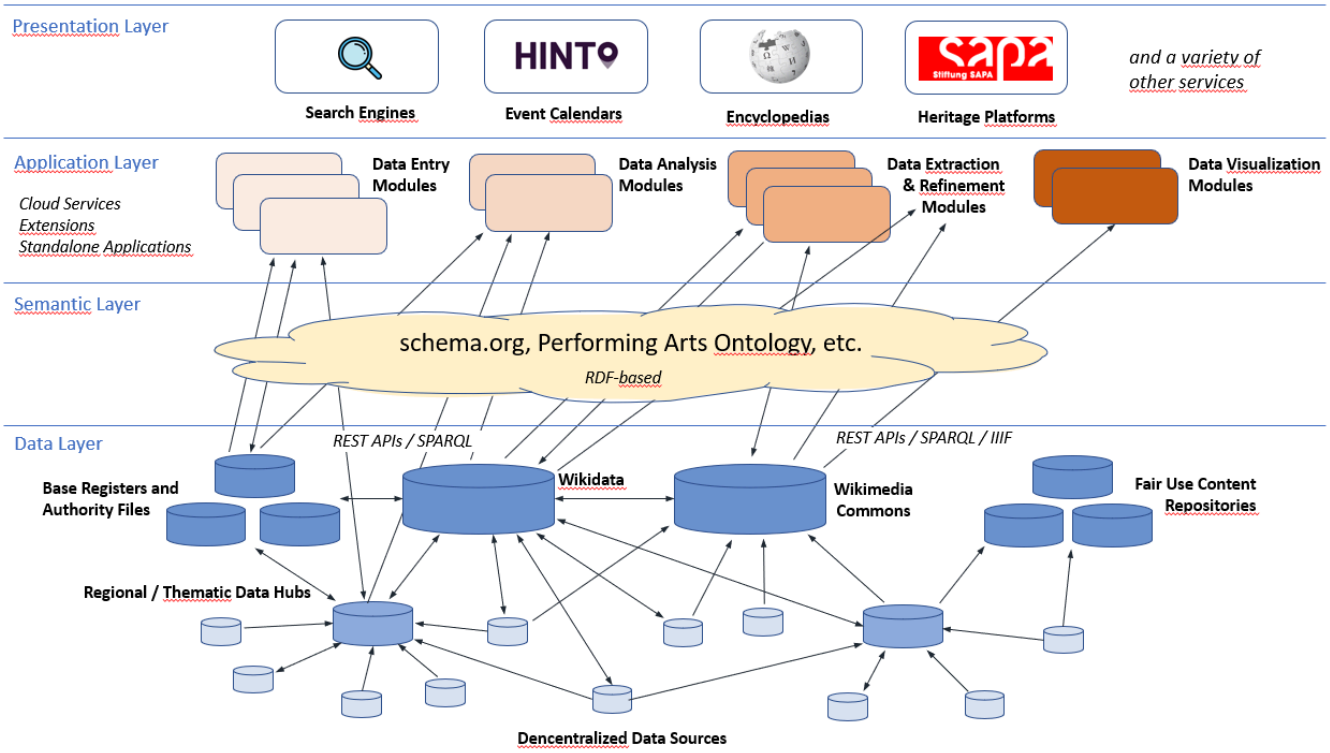

Figure 1: Architectural model of the Linked Open Event Data ecosystem

At the technical level, the vision of a Linked Open Event Data ecosystem consists in a distributed knowledge base for event data, based on linked data technology[13]. Figure 1 gives an overview of the architecture of such a linked open data ecosystem. The architecture is made up of different architectural layers, as can be seen in Figure 1.

Represented at the bottom is the **data layer**, which consists of a distributed database for events data based on linked data* technology, comprising both data platforms through which structured data is made available, and content repositories through which media files (text, image, audio, video, 3D models) are provided. The content repositories are expected to comply with the IIIF* standard, while the structured data is provided through SPARQL* endpoints or openly documented APIs*. The data platforms include base registers* and authority files* (e.g. ISNI*, VIAF*, etc.), which serve as common registers of named entities. While some of these databases take the form of simple triple stores through which data is made available as linked open data, others take the form of community platforms which allow for collaborative data curation, such as Wikidata.

While the structured data is by default made available as open data to facilitate the interlinking of the various databases, a substantial amount of digital content related to the performing arts is covered by

[13] A similar linked open ecosystem, focusing on data from the arts sector, has been described in Estermann B., Julien, F. (2019): A Linked Digital Future for the Performing Arts: Leveraging Synergies along the Value Chain. Canadian Arts Presenting Association (CAPACOA) in cooperation with the Bern University of Applied Sciences and in Estermann, B. (2020): Creating a Linked Open Data Ecosystem for the Performing Arts (LODEPA), Arti dello Spettacolo / Performing Arts, Year 6, Special Issue, pp. 31-49. The text of this section has been sourced from these publications and adapted for the purpose of this document. A glossary of the technical terms marked with an asterisk (*) can be found in Annex 1.

copyright. It is therefore up to the copyright holders and the maintainers of collections to decide whether they want to make the content available under a free copyright license or in some other form of open access regime, or whether they prefer to restrict the access to specific users or user types (e.g. education and research). Here again, an open community platform is available in the form of Wikimedia Commons, the media repository maintained by the Wikimedia Foundation. It is however limited to content that has been published under a free copyright license*.

Displayed right above the data layer is the **semantic layer**. It consists of ontologies, sometimes also referred to as data models, related in some way or the other to events. Along with shared sets of named entities (authority files or base registers), commonly used ontologies represent the shared language that provides the linkages between different datasets within the linked data cloud. Ontologies may be provided as classical standalone RDF* data models, such as FRBRoo (Doerr et al. 2008), or schema.org, and/or they may be implemented and collaboratively maintained within Wikidata.

Next comes the **application layer**, consisting of various cloud services tailored to the needs of some of the “power users” within the linked open data ecosystem. Examples for such cloud services are data extraction tools for data publishers (e.g. helping them to semi-automatically extract metadata from media files), data entry tools that help users enter data into the system, analytical tools for researchers, or data visualization tools for users interested in creating their own tailor-made data visualizations. These services may be provided according to the “software as a service” model and may be integrated into various offers targeted at end users (cf. presentation layer), or they may be provided in form of extensions (e.g. a Wordpress extension for the entry of structured data), or as standalone tools. As outlined above, some of the data platforms and content repositories may take the form of online collaboration and/or crowdsourcing platforms; in this case, they extend into the application layer, providing tools and services that go well beyond providing data and content.

The **presentation layer** heading the schematic structure displayed in Figure 1 finally consists of a myriad of end user services that all draw, at least to some extent, on the elements of the three other layers – by exploiting or aggregating data from various data providers, by using shared ontologies, and/or by making use of cloud services. Thanks to the clear separation of the data layer and the presentation layer, various gateways and services may provide access to the shared data infrastructure. A gateway or a service can be conceived of as a single point of access to a subset of the platform ecosystem, tailored to the needs and expectations of a specific user segment or facilitating a certain task (e.g. searching for upcoming events). In some cases, users simply act as consumers; in others, they take the role of “prosumers,” engaging on crowdsourcing platforms and in collaboration spaces that allow them to make their own contributions to the Linked Open Event Data ecosystem.

3. Roadmap

The roadmap is set out from the point of view of the **data ecosystem as a whole**. It is complemented by **organizational maturity models** (see section 4) that reflect the point of view of individual organizations (data publishers, data aggregators, event producers and organizers, funding organizations) and provides indications as to how to get from one maturity level to the next.

3.1. Bootstrapping the Linked Open Event Data Ecosystem

The following roadmap is based on an analysis of 28 academic papers and reports on data ecosystems[14] which has identified the conditions required for a data ecosystem to thrive. The measures can be assigned to four dimensions: (1) **data sharing**; (2) **interoperability & shared infrastructures**; (3) **stakeholder involvement**; (4) **economic sustainability** (see figure 2). In the following, the checklists for the different dimensions are presented, along with an overview of what has been achieved so far. Based on this, next steps are suggested for each dimension.

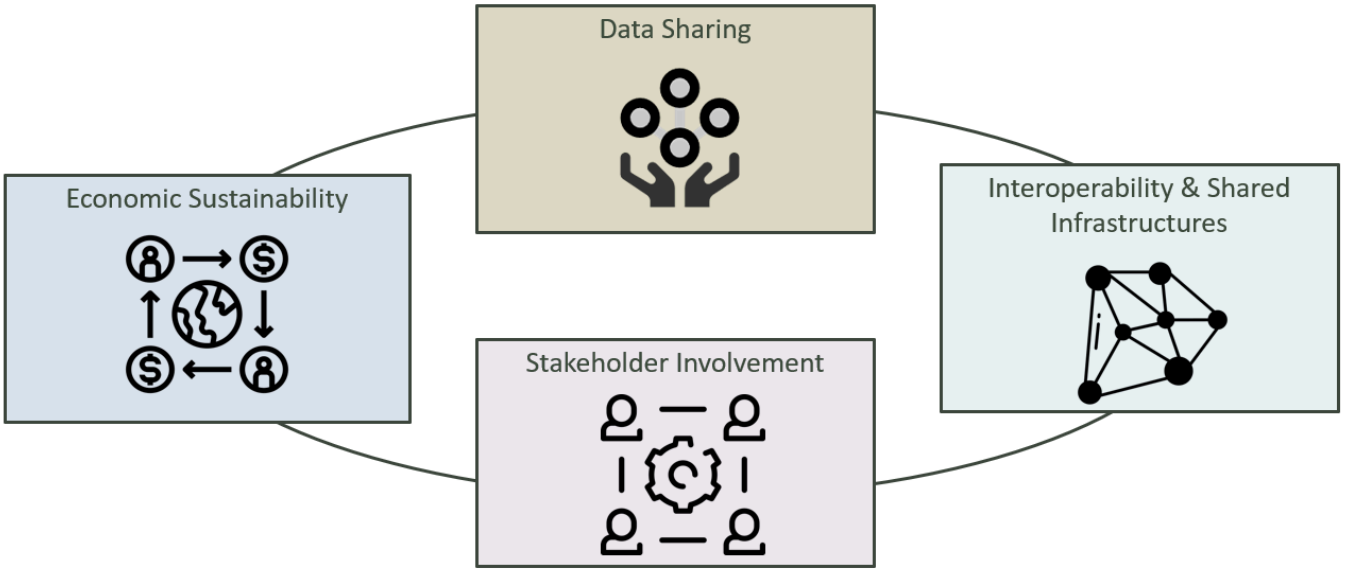

Figure 2: Key aspects of thematic data ecosystems

3.1.1. Data Sharing

The data sharing dimension first and foremost consists in creating a **legal framework and organizational policies** that are in line with the vision of the data ecosystem. In order to establish a Linked Open Event Data ecosystem, the principles set out in the vision need to be endorsed by the participating organizations. Furthermore, **organizational practices** need to be adapted; and for this, a new set of **skills and organizational capabilities** may be required (see figure 3 for a checklist).

[14] Estermann, B. (forthcoming).

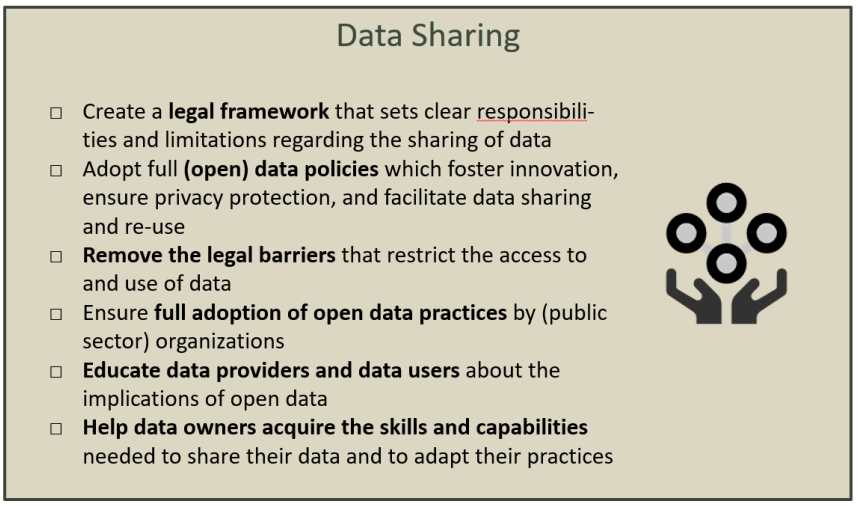

Figure 3: Checklist "Data Sharing"

3.1.1.1. Current State of Implementation

There are various initiatives in Switzerland to **pool event data on central data platforms**, such as cultural calendars, ticketing platforms and the like. While the integration of this data still involves manual effort in many cases, various players have recognized the potential that lies in **automatizing the data transfer**. Several event platforms have started to make their data available to each other. However, this is usually done without making a robust statement regarding the rights of the users in terms of **re-usability of the data**.

The cultural agenda of the canton of Ticino sets a good example, with data that can be obtained as open data via an API[15]. This requires some rights clearance up front, when the data is provided by the event organizers. The Swiss Archive for the Performing Arts (SAPA) takes a similar path by publishing their production database as linked open data.

As regards the **sensitization** of cultural sector organizations and **training of staff members**, we are still at the very beginning. As part of their Linked Digital Future Initiative, the Canadian Association for the Performing Arts (CAPACOA) has made available a series of online resources[16] and has rolled out an extensive training programme for the sector, especially as regards the use of Wikidata[17]. There is a lot of material that can be adapted and re-used in the Swiss context, and there may be opportunities for closer cooperation with CAPACOA in the future.

3.1.1.2. Proposed Next Steps

Activity	Description	Timeframe / Progress
Inventory of data aggregators	Create an inventory of data aggregators documenting their data sharing practices (licenses; technical access)	Q4/2022: initial version created Q1/2023
Sensitization / consultation of data aggregators	Hold a workshop with data aggregators to discuss the vision and roadmap, including the maturity model for organizations	Q1/2023

[15] Die Nutzung der API ist anmeldepflichtig; die Daten werden unter der Creative Commons Zero (CC-0) Lizenz veröffentlicht (siehe weiterführende Informationen zur Agenda culturale sowie zur Banca dati operatori culturali).
[16] https://linkeddigitalfuture.ca/resources/
[17] https://linkeddigitalfuture.ca/wikidata-workshops-season-2/

Activity	Description	Timeframe / Progress
Implementation of data sharing practice on the Hinto platform	The data made available through the Hinto API is released under a CC-0 license; for images, a robust license statement is systematically provided.	Q1/2023

3.1.2. Interoperability and Shared Infrastructures

In the technical domain, ensuring **data harmonization and interoperability of data and systems** is key when it comes to establishing a data ecosystem. The **infrastructures** through which the data is shared should be **stable, scalable, available, and reliable**. Furthermore, **common tools and services** may be provided that facilitate data integration and can serve as showcases for the consumption of the harmonized and interlinked data (see figure 4 for a checklist).

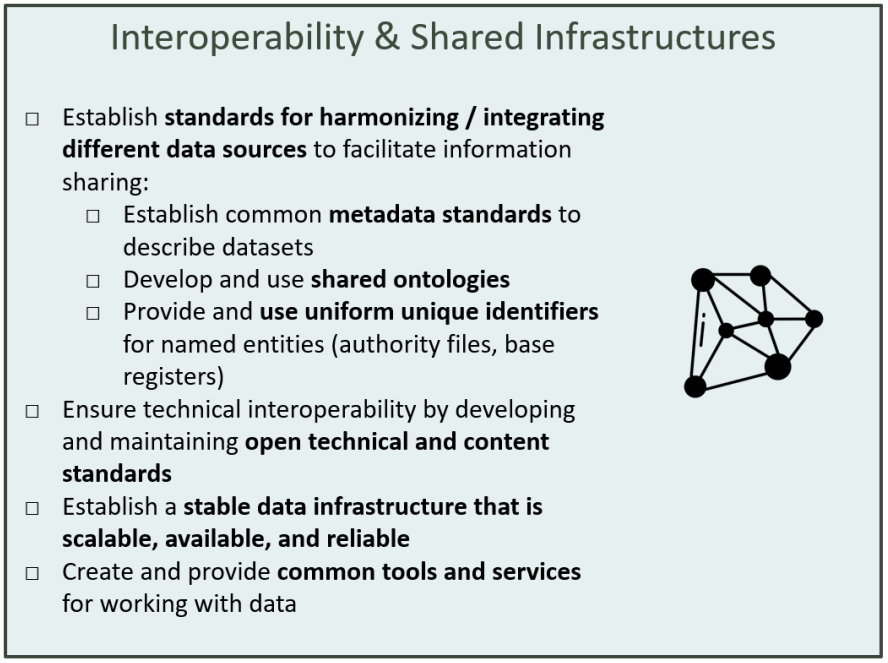

Figure 4: Checklist “Interoperability & Shared Infrastructures”

3.1.2.1. Current State of Implementation

Over the past years, Wikidata has developed into a **data integration hub**, through which **base registers and authority files** are interconnected. This approach has been widely embraced by the heritage sector. The LODEPA community, which aims for the establishment of a Linked Open Data Ecosystem for the Performing Arts[18], has followed suit by ingesting data about organizations, persons, venues, and creative works related to the performing arts as well as data about individual performing arts productions[19]. In this way, **the maintenance of data held across many different organizations can be pooled and tackled in common**, and the respective Wikidata Q-numbers can be used as uniform unique identifiers for the various classes of entities that can easily be linked to other authorities. The data on Wikidata can be

[18] The community consists of about 50 professionals in around 25 countries. It has so far been jointly led by the Bern University of Applied Sciences and CAPACOA. For further information, see: https://linkeddigitalfuture.ca/community/
[19] Data harmonization and integration is mainly coordinated through the WikiProject Performing Arts, the WikiProject Cultural Venues, and related Wikidata projects as well as tri-monthly working group meetings.

accessed through a variety of APIs, including a SPARQL endpoint. Furthermore, the data can be downloaded in form of data dumps[20].

In the form of the Wikidata platform, the Wikimedia Foundation, together with the Wikimedia Movement, provides a **stable, scalable, and reliable infrastructure for collaborative data management**. Thanks to a lively community, many tools have been developed that facilitate the matching, disambiguation, ingest, exploration, monitoring, and visualization of data (e.g. OpenRefine, QuickStatements, Mix'n'Match, Entitree, Integraality, Histropedia, etc.[21]). The Wikidata community has also established processes for data harmonization, following a Bazaar model[22]. Furthermore, there are established policies and processes to deal with potential privacy issues that may arise when managing data about living people[23].

Thus, a large part of the shared infrastructure required for the establishment of a Linked Open Event Data ecosystem already exists. It is ready to be used for the collaborative management of data about people, organizations, venues, creative works, and complex events, such as festivals or performing arts productions. A lot of relevant data has already been ingested into Wikidata, but it will still require a considerable effort in terms of data ingestion and data cleansing before the platform can be used as a reference database for the core classes of items required for the Linked Open Event Data ecosystem.

In addition to Wikidata, several base registers and authority files relevant to the performing arts are available as linked open data or can be queried through an API, including the Virtual Authority File (VIAF)[24] or the International Standard Name Identifier (ISNI)[25]. They play a key role in identifying persons or organizations across several databases. In contrast to Wikidata, VIAF and ISNI are narrower in scope and do not rely on crowdsourcing. Another community project relying on crowdsourcing is MusicBrainz[26]; it can serve as a reference source for music albums, musical works and performers. Further reference sources relevant in the context of the performing arts are the Internet Movie Database (IMDb)[27] for actors, Songkick[28] for concerts, and Discogs[29] as an alternative to MusicBrainz for performers.

While Wikidata is a good starting point for the collaborative management of some aspects of Linked Open Event Data, it has to be kept in mind that Wikidata is not suited for the management of data that cannot be publicly shared. Thus, data that is to be shared only among a number of selected stakeholders cannot be managed in Wikidata. Also, the mechanisms to deal with vandalism are mostly reactive, which means that the platform is not suited as the primary source of business-critical information that is time-sensitive. The platform is also not a good place to store data that is subject to frequent changes or are of no value after a given point in time (e.g. after an event has taken place). For all these reasons, other solutions are needed to store some of the core data about upcoming events, such as time, place, participants, availability of seats, ticketing information, event status, and the like. If such data is to be federated and made available across several organizations and platforms, one or several other solutions are needed.

There are currently two initiatives that have been launched in this area that aim to be in line with the principles of a Linked Open Event Data ecosystem and are relevant to the Swiss market:

---

[20] See https://www.wikidata.org/wiki/Wikidata:Data_access for an overview.
[21] For an overview of available tools, see: https://www.wikidata.org/wiki/Wikidata:Tools
[22] See Eric Raymond's 1997 essay, The Cathedral and the Bazaar.
[23] See: https://www.wikidata.org/wiki/Wikidata:Living_people
[24] http://viaf.org/
[25] https://isni.org/
[26] MusicBrainz, The Open Music Encyclopedia, https://musicbrainz.org/
[27] https://www.imdb.com/
[28] https://www.songkick.com/
[29] Discogs, Music Database and Marketplace, https://www.discogs.com/

- A community of stakeholders under the lead of Theapolis is working towards the establishment of a **common theater production database**[30]; its current geographical scope is the D-A-CH region.

- The Switzerland-based association “Open Event Data” aims to promote open event data by making event data available and sharing them across platforms thanks to open APIs[31]. The event data thus provided is made available on the **event platform** Hinto[32] among others.

Both initiatives are considering using Wikidata as a common data hub where authority data on venues, people, organizations, creative works, and performing arts productions can be managed collaboratively, while some of the event data will be managed outside Wikidata.

As to the development of **shared data models** in the area of cultural events, various initiatives exist, the most notable ones being schema.org and the activities of the LODEPA community:

- Schema.org[33] is a reference website publishing documentation and guidelines for using structured data mark-up on web-pages (called microdata). Its main objective is to standardize HTML tags to be used by webmasters for creating rich results (displayed as visual data or infographic tables on search engine result pages). It is also used as an ontology to publish classical linked data, using RDF and triplestores.

- The LODEPA community’s efforts towards data harmonization in the area of the performing arts, is notably taking place through the WikiProjects Performing Arts[34], Cultural Venues[35], and Cultural Events[36], as well as through the W3C PAIR Community Group[37].

There is currently no harmonized practice regarding the description of simple events among data aggregators in Switzerland.

As regards the **common tools and services** required for the establishment of the Linked Open Event Data ecosystem, many of the tools used in the context of Wikidata (see above) are domain-agnostic and can be used by data publishers or aggregators willing to use Wikidata as a reference database. There are however two types of services that would be particularly useful when it comes to bootstrapping the Linked Open Event Data ecosystem:

- Platforms for data collection that provide users the possibility to directly link their data entries to Wikidata or other commonly used base registers or authority files. Many platforms currently let users enter their data in the form of strings. Some platforms have some form of internal reconciliation. Reconciliation against external reference databases is still the exception.

- Services which federate and display event data made available as linked open data and make the federated data available for reuse. Such services are important showcases that can demonstrate the effectiveness of data harmonization efforts and can help facilitate the use of

[30] See: https://theapolis-support.de/projekt-inszenierungsdatenbank/
[31] See: https://www.openeventdata.ch/de/mitmachen-43.html
[32] https://www.hinto.ch/
[33] See: https://schema.org/Event
[34] https://www.wikidata.org/wiki/Wikidata:WikiProject_Performing_arts
[35] https://www.wikidata.org/wiki/Wikidata:WikiProject_Cultural_venues
[36] https://www.wikidata.org/wiki/Wikidata:WikiProject_Cultural_events
[37] https://www.w3.org/community/pair-cg/

federated data. There are currently several applications under development that may eventually serve this purpose, among them the Hinto platform[38] and the Culture in Time app[39].

An overview of relevant software tools and components can be found in Annex 2.

3.1.2.2. Proposed Next Steps

Activity	Description	Timeframe / Progress
Start using Wikidata as reference database for event venues	Link all the venues in the Hinto database and in the SAPA database to corresponding items in Wikidata. Resolve any remaining conflation issues in Wikidata (e.g. describing a place and an organization in the same Wikidata item).	Q4/2022: Pilot ingest SAPA venues initiated Q1/2023
Overview of data models currently employed	Provide an overview of the data models that are currently employed by data publishers and data aggregators who are interested in aligning their data models in view of the establishment of a Linked Open Event Data ecosystem.	Q1/2023
Overview of current data exchange practices	Provide an overview of the data exchange practices (APIs, etc.) currently employed among data publishers and data aggregators who are interested in harmonizing their practices in view of the establishment of a Linked Open Event Data ecosystem.	Q1/2023
Showcase Hinto	Promote the Hinto platform as a showcase for the progress of data sharing and harmonization practices in Switzerland.	Q2/2023
Showcase Culture in Time	Promote the Culture in Time app as a showcase for the progress of data sharing and harmonization practices in the performing arts domain across the globe.	Q2/2023
Start using Wikidata as reference database for organizations	Link all the organizations in the Hinto database and in the SAPA database to corresponding items in Wikidata. Resolve any remaining conflation issues in Wikidata (e.g. describing a place and an organization in the same Wipkidata item).	Q4/2022: Pilot ingest SAPA organizations initiated Q2/2023

3.1.3. Stakeholder Involvement

The aspect “stakeholder involvement” comprises three important dimensions (see figure 5):

1. ensuring diversity by involving all relevant stakeholder groups in the ecosystem;
2. defining and appropriately filling network roles and setting up adequate organizational structures;
3. empowering (potential) members of the ecosystem to actively participate by helping them acquire the necessary knowhow.

To achieve this, network leadership needs to be provided, which also means that a shared vision needs to be developed in order to rally the various stakeholders behind the same goals.

[38] https://www.hinto.ch/
[39] https://culture-intime-production.herokuapp.com/

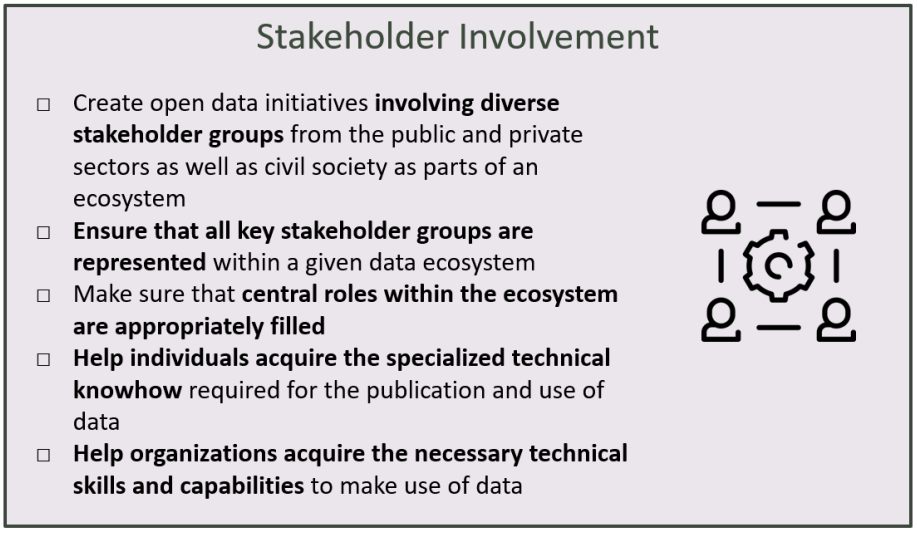

Figure 5: Checklist "Stakeholder involvement"

3.1.3.1. Current State of Implementation

The LODEPA network operates as a loose network without any formal organizational structure and without central funding. Together with their Canadian counterparts, Swiss partners have played a leading role from the outset – both in terms of network coordination and when it comes to deploying and using linked data technology. The network comprises a broad variety of stakeholders from about 25 countries. Network cohesion, effectiveness, and growth is however hampered by the lack of funding for network coordination tasks, real life meetings, and conferences. Still, the network offers a good setting to coordinate and harmonize data management practices in the area of performing arts at an international level.

As regards the stakeholder involvement at the level of events data in Switzerland, the situation should be improved. There have been many initiatives to cooperate at a regional level[40], and the larger event data platforms have been able to gather a considerable customer base. However, overall coordination and a commonly agreed course of action among the key stakeholders is lacking.

3.1.3.2. Proposed Next Steps

Activity	Description	Timeframe
Stakeholder analysis	Create an overview of key stakeholder groups and identify possible representatives	Q1/2023
Foster a shared vision among key stakeholders	Hold a workshop with representatives of key stakeholder groups to agree on a common vision and roadmap	Q2/2023
Designate or create a supporting organization	Designate one of the existing organizations as a supporting organization for the implementation of Linked Open Event Data in Switzerland, or alternatively, create a new organization for this purpose (association, cooperative). Formulate a mandate, clarify forms of membership, etc.	Q3-Q4/2023

[40] See for example: Verein Open Event Data, kulturzueri.ch, the cooperation between the event platforms Saiten and thurgaukultur.ch, etc.

3.1.4. Economic Sustainability

The fourth dimension is concerned with the economic sustainability of the data ecosystem: First of all, it is important that those stakeholders who are called upon to make an extra effort to contribute to the ecosystem will also be able to benefit from it. Market participants who embrace data sharing practices may need to adjust their business models. For that, new business models that do not rely on the appropriation of data or on creating customer lock-in may need to be devised. In some cases, it may be useful to have the public sector or charities step in to make some initial investments.

But most importantly, on an ecosystem level, it is important that data sharing practices be developed from a demand-driven perspective and have a clear value-added business case. The data provided needs to satisfy customer needs or serve business interests. If not, the ecosystem as a whole will not yield sufficient benefits to ensure its sustainability over time. This is why it is important to involve not only the data providers, but also the potential users and beneficiaries of the data ecosystem (see figure 6).

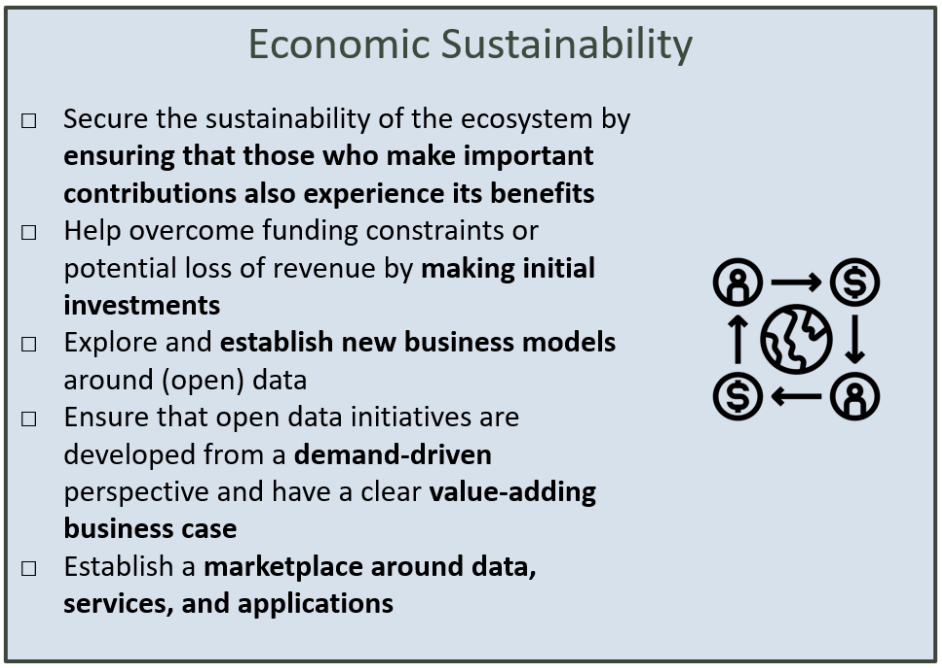

Figure 6: Checklist “Economic Sustainability”

3.1.5.1. Current State of Implementation

As has been noted above, a large part of the shared infrastructure required for the establishment of a Linked Open Event Data ecosystem already exists in the form of Wikidata, its community and the tools that have been developed to interact with it. The platform is provided by the Wikimedia Foundation, and its development is mainly ensured by Wikimedia Germany. The provision and further development of the platform is funded by donations. Thanks to the wide reach of the Wikimedia movement’s flagship project Wikipedia, the Wikimedia organizations are regularly able to raise a sufficient amount of money to ensure the long-term sustainability of its platform ecosystem.

The Wikidata community itself functions as an open collaborative community that engages in commons-based peer production[41]. Contributions are made on a voluntary basis without direct compensation, either by individual volunteers, by public sector and/or non-profit organizations, or by

[41] See: “Commons-based peer production” (Wikipedia)

companies that see a benefit in maintaining a shared data infrastructure. The data curated on the Wikidata platform is made available for free without restrictions regarding its reuse.

When it comes to implementing a Linked Open Event Data ecosystem in Switzerland, there are several aspects that will require dedicated funding, ideally at the network level:

- The development of common tools and services, especially as regards showcase implementations that use federated data.
- Network coordination to provide leadership for the agreement on a shared vision and a common roadmap, for the coordination of harmonization and standardization work, as well as for the coordination of outreach and training activities.
- An outreach, sensitization and training campaign to allow for the active involvement of a variety of stakeholders.

There have been some first tentative efforts in these areas, mostly promoted by individual organizations or regional groupings[42]; apart from the project aiming at creating a shared theater production database, coordinated by Theapolis and funded by the Government of Germany, these activities have not had a truly national or international scope[43].

At the level of individual organizations, it is important that they see a business case in harmonizing the data model and in pooling base registers and authority files across the sector. Furthermore, they will need to adapt their tools and business practices to be able to productively interact with the open collaborative community on Wikidata. In some cases, business models will need to be adapted in order to avoid reliance on proprietary data policies and reliance on customer lock-in in a situation characterized by important network effects.

The situation of every organization is different. While sharing event data explicitly as open data is not common practice yet among data aggregators in Switzerland, there have been various initiatives to facilitate and automate the data transfer between platforms, and there have been efforts to improve the permeability in the dissemination of event data across platform boundaries. However, the majority of intermediaries are still holding back when it comes to making the aggregated data available as open data for free use by third parties. The Canton of Ticino's cultural agenda sets a good example here, as its data can be obtained as open data via an API[44].

In the area of theater production data, the Swiss Archive of the Performing Arts (SAPA) is breaking new ground by making its production history database available as linked open data and intends to maintain a large part of its data in the collaborative online database Wikidata in future, where it can be used and supplemented by third parties. But even in the case of SAPA, the acquisition of data from producers and organizers of theater events is characterized by “media breaks” (data is manually copied and entered into the database). Furthermore, there is untapped potential for using this data for the promotion and diffusion of theater events, but this would require the data cleansing processes to be located further upstream in the data value chain.

---

[42] E.g. the Hinto Platform, or the cooperation project between Saiten and thurgaukultur.ch, the shared theater production database promoted by Theapolis.
[43] We are focusing here only on activities with a geographical reach covering Switzerland or parts thereof. There have been other initiatives with a national scope, e.g. in Canada.
[44] The use of the API requires registration; the data are published under the Creative Commons Zero (CC-0) License (see further information about the Agenda culturale and the Banca dati operatori culturali).

This is an old version; current Working Document: Linked Open Event Data - Vision & Roadmap

3.1.5.2. Proposed Next Steps

Activity	Description	Timeframe
Identify and coordinate among “First Movers”	Identify the organizations that are most likely to join the Linked Open Event Data bandwagon by developing their tools and business practices accordingly. Start coordinating their activities in order to maximize synergies (and in turn the benefits of the First Movers) at an early stage.	Q1/2023
Fundraising	Find funding solutions for network level activities after the end of the Innocheque project.	Q1-Q2/2023

4. Organizational Maturity Models

This section contains the organizational maturity models which complement the roadmap and reflect the point of view of individual organizations and provide indications as to how to get from one maturity level to the next as regards their implementation of Linked Open Event Data.

The proposed maturity models are tentative and serve as a basis for further discussion.

4.1. Maturity Model for Data Publishers and Data Aggregators

The table below contains a set of guiding questions for the maturity model. The response modalities are developed in a separate document (spreadsheet) in direct cooperation with representatives of data publishers / data aggregators. The aspects highlighted in green are deemed to be the most important at this stage *(to be further discussed)* and have thus been included in the initial version of the maturity model.

Aspect	Guiding Question
Strategy	To what extent is the organizational strategy aligned with the vision of a Linked Open Event Data ecosystem?
Business Model	To what extent is the organization's business model compatible with the key principles of the Linked Open Event Data ecosystem?
Leadership	To what extent does the leadership culture empower staff members to engage in collaborative data management projects, such as Wikidata?
Staff Skills	To what extent do staff members have the necessary skill set and/or access to relevant training opportunities that empower them to actively participate in collaborative data management?
Staff Resources	To what extent does the organization have the necessary staff resources to support the implementation of Linked Open Event Data?
Financial Resources	To what extent does the organization have the necessary financial resources to support the implementation of Linked Open Event Data?
Data Reuse	To what extent does the way the data is managed by the organization facilitate the re-use of the data in different contexts?
Once Only	To what extent is the data management of the organization organized in such a way that users (private individuals or companies) have to enter the same data only once?
Mydata	To what extent is the personal data of private individuals provided in such a way that they can view and maintain their data electronically by themselves?
Open Data	To what extent does the organization make event related data available as open data?
Open Content	To what extent does the organization make event related content (texts, photographs, trailers, video recordings) available as open content?
Data Exchange with Data Aggregators	How does the data exchange with other data aggregators / data publishers take place?
Data Exchange with Clients	How does the data exchange with clients (businesses, individuals) take place?
Use of Harmonized Data Models	To what extent are the data models used in the specialized applications of the organization harmonized with those of other organizations holding similar data?

Base Registers / Authority Files	To what extent are the data holdings of the organization linked to central reference databases (base registers, authority files)?
Technology	To what extent does the IT architecture of the tools used to manage event data support their integration with the emerging Linked Open Event Data ecosystem?
Tools	To what extent do the software tools used to manage event data support the users' interaction with and contribution to the Linked Open Event Data ecosystem?
Authentication	To what extent does the organization have access to a widely accepted electronic authentication solution for transactions both with internal and external counterparts?
Synergies	When procuring or developing IT systems and specialized applications, to what extent is attention paid to exploiting synergies with like-minded organizations?
Open Source Software	To what extent is preference given to open source solutions when procuring or developing software?

4.2. Maturity Model for Producers and Organizers of Cultural Events

The table below contains a set of guiding questions for the maturity model. The response modalities are developed in a separate document (spreadsheet) in direct cooperation with representatives of producers and organizers of cultural events. The aspects highlighted in green are deemed to be the most important at this stage *(to be further discussed)* and have thus been included in the initial version of the maturity model.

Aspect	Guiding Question
Strategy	To what extent is the organizational strategy aligned with the vision of a Linked Open Event Data ecosystem?
Business Model	To what extent is the organization's business model compatible with the key principles of the Linked Open Event Data ecosystem?
Leadership	To what extent does the leadership culture empower staff members to engage in collaborative data management projects, such as Wikidata?
Staff Skills	To what extent do staff members have the necessary skill set and/or access to relevant training opportunities that empower them to actively participate in collaborative data management?
Staff Resources	To what extent does the organization have the necessary staff resources to support the implementation of Linked Open Event Data?
Financial Resources	To what extent does the organization have the necessary financial resources to support the implementation of Linked Open Event Data?
Data Reuse	To what extent does the way the data is managed by the organization facilitate the re-use of the data in different contexts?
Once Only	To what extent is the data management of the organization organized in such a way that users have to enter the same data only once?
Mydata	To what extent is the personal data of private individuals provided in such a way that they can view and maintain their data electronically by themselves?
Open Data	To what extent does the organization make event related data available as open data?
Open Content	To what extent does the organization make event related content (texts, photographs, trailers, video recordings) available as open content?

This is an old version; current Working Document: Linked Open Event Data - Vision & Roadmap

Data Exchange with Data Aggregators	How does the data exchange with other data aggregators / data publishers take place?
Data Exchange with Clients	How does the data exchange with clients (businesses, individuals) take place?
Use of Harmonized Data Models	To what extent are the data models used in the specialized applications of the organization harmonized with those of other organizations holding similar data?
Base Registers / Authority Files	To what extent are the data holdings of the organization linked to central reference databases (base registers, authority files)?
Technology	To what extent does the IT architecture of the tools used to manage event data support their integration with the emerging Linked Open Event Data ecosystem?
Tools	To what extent do the software tools used to manage event data support the users' interaction with and contribution to the Linked Open Event Data ecosystem?
Authentication	To what extent does the organization have access to a widely accepted electronic authentication solution for transactions both with internal and external counterparts?
Synergies	When procuring or developing IT systems and specialized applications, to what extent is attention paid to exploiting synergies with like-minded organizations?
Open Source Software	To what extent is preference given to open source solutions when procuring or developing software?

4.3. Maturity Model for Funding Organizations

[to be developed at a later stage]

5. Concluding Remarks and Outlook

- Further process / next steps regarding the further elaboration of this document: outreach to potential reviewers of the document / potential first movers when it comes to implementing Linked Open Event Data in Switzerland.

- Aspects that have hitherto not been sufficiently covered in the document:

    - Weitere Anknüpfungspunkte, die im Dokument bisher nicht berücksichtigt sind (siehe Whitepaperkapitel zu Datenraum für Kultur- und Kulturerbedaten):

        (1) Umsetzung des Once-Only-Prinzips bei der Kulturförderung;

        (2) Mitberücksichtigung der Perspektive Digitales Gedächtnis im Rahmen der staatlichen Förderpolitik (das muss sich nicht auf kulturelle Events beschränken).

    - Aus Sicht von fugu ist es wichtig, auch Sportevents und gewerbliche Events mit in den Blick zu nehmen. Der Veranstaltungskalender sollte sich nicht auf Kulturevents allein beschränken.

    - Importance of a tiered approach: not all events need to be promoted across platforms; not all cultural events need to be documented in view of long-term preservation / digital memory; this should be reflected in the document, for many organizations and individual event organizers, the adoption of the proposed practices will be purely optional.

    - When including funding organizations, e.g. by promoting the Once Only Principle in the area of cultural promotion, it is important to clearly differentiate between the aspects pertaining to the technical implementation of the infrastructure and the political aspects related to funding policies; the former must be strictly neutral with regard to the latter (Corodis).

    - At what point do we need to offer trainings (e.g. in form of webinars) for performing arts professionals and performing arts students?

    - Positioning of the Hinto Platform and the Data Ecosystem as a Global Digital Public Good.

6. Acknowledgements

The first draft version of this document has been established thanks to an Innocheque project[45] on behalf of fugu GmbH.

Author(s):

- Beat Estermann, Bern University of Applied Sciences / Bern Academy of the Arts

Reviewer:

- Stefan Grünig, fugu GmbH

[45] https://www.innosuisse.ch/inno/en/home/promotion-of-national-projects/innovation-cheque.html

Annex 1: Glossary

The following glossary of technical terms has been sourced from Estermann (2020) and expanded for the purpose of this document.

Authority file	An authority file (or base register) contains a list of named entities which many data publishers link against. The term “authority file” is widely used within the heritage sector.
API	An “Application Programming Interface” is an interface between an application software and an application platform, through which all services are provided.
Base register	A base register (or authority file) contains a list of named entities which many data publishers link against. The term “base register” is widely used in the government sector and often refers to a register that is maintained by a public authority under a specific legal mandate.
Free copyright licenses	In most jurisdictions, any creative work is protected by copyright by default, which means that it can only be replicated with the consent of its creator. Copyright protection is typically granted for the lifetime of the creator plus 70 years. A copyright license gives a licensee authorization to use a work from the copyright owner, often in exchange for payment or subject to some restrictions. In order to facilitate the sharing and re-use of content, several organizations have established standard copyright licenses. Some of them allow anyone to freely share and re-use the content for any purpose, including commercial use, and are therefore referred to as “free copyright licenses”. Examples include the Creative Commons Attribution (CC BY) and the Creative Commons Attribution Share Alike (CC BY-SA) licenses.
GND	The “Gemeinsame Normdatei” (Integrated Authority File) is an international authority file for personal names, subject headings and corporate bodies. It is used for documentation in libraries and increasingly also by other types of heritage institutions. The GND is managed by the German National Library in cooperation with various regional library networks in German-speaking Europe and other partners.
HTTP	“Hypertext Transfer Protocol”, an application protocol for distributed, collaborative, and hypermedia information systems; it provides the foundation of data communication for the World Wide Web.

IIIF	The International Image Interoperability Framework (IIIF) defines application programming interfaces that provide a standardized method for describing and delivering images over the web. Similarly to the linked data approach, this standard supports a decentralized platform architecture where content is made available by various servers from where it can be dynamically integrated into various online services provided by clients. While IIIF programming interfaces for image content have existed for several years, the IIIF standard for audiovisual content is currently under development.
ISNI	“International Standard Name Identifier” refers to a unique, global, cross-domain, standard, persistent identifier for a person or an organization involved in the production and exploitation of creative content. The development of the ISNI register is mainly driven by libraries, rights management societies, stakeholders of the book supply chain, aggregators and service suppliers.
Linked Data	“Linked Data” is structured data which is interlinked with other data so it becomes more useful through semantic queries. It builds upon standard Web technologies such as HTTP, RDF and URIs. The term has been coined by Tim Berners-Lee, director of the World Wide Web Consortium (W3C) in a 2006 design note about the Semantic Web project. Linked data may also be open data, in which case it is usually described as “Linked Open Data” (LOD).
Ontology	An ontology (sometimes also referred to as a data model or conceptual model) is a shared language that is used to represent aspects of the world in form of data. An ontology typically consist of classes and properties for which definitions are provided as well as rules how the classes and properties may be combined among each other.
Open Data	“Open Data” is data that is freely available to everyone to use and republish as they wish, without restrictions from copyright, patents or other mechanisms of control.
RDF	“Resource Description Framework”, a family of W3C specifications used for the formal representation and exchange of linked data.
Reification	Reification in knowledge representation is the process of turning a statement into an object. Reification involves the representation of factual assertions that are referred to by other assertions. It is required to further qualify individual statements, e.g. to indicate their source or the period of their validity.

Semantic Web	The Semantic Web is an extension of the World Wide Web through standards set by the World Wide Web Consortium (W3C). The goal of the Semantic Web is to make Internet data machine-readable. To enable the encoding of semantics with the data, technologies such as Resource Description Framework (RDF) and Web Ontology Language (OWL) are used. The Semantic Web provides a common framework that allows data to be shared and reused across application, enterprise, and community boundaries.
SPARQL	“Simple Protocol and RDF Query Language” is the language used to define queries to a datastore of linked data. A standard API that can be queried using SPARQL is called a “SPARQL endpoint”.
URI	A “Uniform Resource Identifier” is a string of characters used to identify names or resources on the Internet.
VIAF	The “Virtual International Authority File” is an international authority file resulting from a joint project of several national libraries to interlink their national authority files. Since 2012, the service is operated by the Online Computer Library Center (OCLC).
W3C	The “World Wide Web Consortium” is the main international standards organization for the World Wide Web.
Wikidata	Wikidata is a collaboratively edited knowledge base hosted by the Wikimedia Foundation and maintained by an online community which anybody can join. It runs on the MediaWiki software and uses the Wikibase extension to store structured data. The software provides the necessary online collaboration features, such as user management, version history, watchlists, discussion pages, etc. The data is provided free of copyright and can be queried through a SPARQL endpoint. One of the main purposes of Wikidata is to serve as a repository for the structured data used in the context of the free online encyclopaedia Wikipedia.

Annex 2: Existing Software Tools & Components

The table below contains a list of existing software tools and components that are useful in view of the establishment of a Linked Open Event Data ecosystem.

Name	Description / Link	State of Development
Artsdata.ca	Canadian Knowledge Graph for the performing arts, which is open, accessible and community sourced. Developed and maintained by Culture Creates on behalf of the Canadian Association for the Performing Arts (CAPACOA) http://kg.artsdata.ca/ See also the presentation (2023) at: https://www.youtube.com/watch?v=wE5EkoV5NIU The Artsdata.ca platform is a set of RESTful APIs to read and write data in the Artsdata.ca knowledge graph: - Reconciliation API - for named entity recognition https://reconciliation-api.github.io/testbench/#/client/https%3A%2F%2Fapi.artsdata.ca%2Frecon - Artsdata Databus API - to upload data to Artsdata.ca. An account is required. - ID Minting API - To mint new URIs in Artsdata. An account is required. - Search API - under development http://artsdata-satellite-org-finder.s3-website.ca-central-1.amazonaws.com/ - Knowledge Card API (with ranking system) - under development	Actively collecting data from 30+ data sources including presenters' websites and arts service organizations' APIs. Early stage roll-out of APIs to a handful of data clients in 2023.
Footlight	A suite of tools that help arts organizations manage their event data.. Footlight is designed to work well with Artsdata.ca and other LOD sources for named entity reconciliation and data enrichment.	Footlight Console private beta released in 2018. Footlight CMS planned release in Q2 2023. Footlight CMS currently runs on MongoDB; in the future, it may be set up to run also directly on Triple Stores.

	Footlight reuses data in Artsdata and publishes data back to Artsdata.ca for re-use by others. Footlight currently comes with the following tools: **1. Footlight CMS** For arts administrators to create and manage event data for their public facing web site. https://cms.footlight.io **2. Footlight Console** For arts administrators to convert events into linked open data by crawling websites. Footlight Console can publish structured data about events using a JSON-LD code snippet placed on the client's website, and also publish structured data in Artsdata.ca. Footlight Console Developed and maintained by Culture Creates.	
Culture in Time	A tool to facilitate the large scale use of linked open data across several data sources. https://culture-intime-production.herokuapp.com/en https://culture-intime-staging.herokuapp.com/en Developed and maintained by Culture Creates in cooperation with Estermann Digital Transformation Consulting, currently with the support from SAPA (Swiss Archive of the Performing Arts).	Under development
Airtable Plugin	Plugin for Airtable that supports data reconciliation with Artsdata.ca. Developed and maintained by Culture Creates.	
Hinto **(powered by Fuman)**	Event aggregation platform. Developed and maintained by fugu GmbH. Based on fugu's own open source CMS Framework Fuman, which is geared towards efficient data management and multilingual support. https://www.hinto.ch https://www.fuman.ch	Hinto is currently under development. Fuman has been in use for almost 2 decades (development started in 2002).

This is an old version; current Working Document: Linked Open Event Data - Vision & Roadmap

OLE Interface	“Open Linked Event Data Interface” - An API for the open exchange of event data. https://www.hinto.ch/en/arbeiten-mit-hinto/ole-widget/ole-in-detail-70.html	
OLE Plugin	Wordpress Plugin facilitating the publication of event data according to the OLE standard. https://www.hinto.ch/en/arbeiten-mit-hinto/ole-widget/ole-in-detail-70.html	