The Swiss Archive of the Performing Arts is a private foundation and the consolidation of several cultural heritage institutions from the field of dance and theater. One result of this merger is the necessity to join various legacy databases of the previous institutions into one common database that connects the manifold holdings and data collections. The currently developed solution will be a graph database and thus is part of an ongoing paradigm shift in regard to the representation of knowledge. Besides of cleansing and enriching the existing data and building the new technical infrastructure, the development of a data model that can cope with the heterogeneity of the data is a major task. Data models for performing arts have evolved over the recent years and this article tries to give an overview over central problems herein and how they have been addressed.

Three problems and the tools to solve them

As an essentially intangible form of culture, performing arts have special requirements when it comes to data modeling, which also can be conceived as three central problems of representation.

1. What is the actual work of art in performing arts? Is it the production as the context in which decisions about the selection and realization of a play are made, in which stage scenery, costumes, and light are designed? Or is it the actual performance on stage? If yes, then how do we have to understand differences between individual performances? We could raise further questions here in what I want to call the problem of the artwork.

2. At its core, a database is a structured form of description and descriptions are accounts of specific entities. In the case of performing arts central parts of what we consider to be the subject of interest are immaterial and unlike other cultural artifacts are no longer at hand. That’s why archives and museums collect objects and documentations as proxies. This difficulty to seize the work of art directly constitutes the problem of description.

3. There is not only an aggregation of works and objects but also of persons and institutions. The creation of performing arts works is not an individual affair but a complex network of interdependent authorships. Texts by people, who died a long time ago, are used; the border between creative decision and technical services blurs in the course of collaboration; responsibilities are often concealed by institutional structures; functions as much as specific roles are not simply fulfilled but constantly reinterpreted. This is what I want to call the problem of agency.

In order to formalize descriptions stored in databases and to make them machine-readable and linkable, formal database models or ontologies are required, that define what can be said meaningfully about a work, its authorship, history, and context. As there is no comprehensive and generally accepted ontology specifically for performing arts, more generic ontologies from the fields of archives, museums and libraries have to be adopted and combined.
1. **CIDOC CRM** is an event-centered reference model for museums that is widely accepted and versatile in its application. However, the advantage of being flexible enough to allow individual interpretations here, can easily turn into a disadvantage and permit also widely incompatible applications. At the same time this seems to be the best option to link databases from different cultural fields, which requires communities that negotiate shared implementations. (CIDOC CRM Special Interest Group 2015)

2. **FRBRoo** is an extension of CIDOC CRM based on the FRBR-standard for libraries, where apparently it is hardly used. But for us it is valuable because of its coverage of literary sources and its compatibility with CIDOC. (International Working Group on FRBR and CIDOC CRM Harmonisation 2017)

3. **Records in Context (RiC)** might turn into a future standard for archives and allow to transcend the provenance limitation of current standards. At the moment, there is no definitive version of RiC but as the provenance of archival holdings is not relevant for this article, we also do not need it. (ICA. EGAD - Expert Group on Archival Description 2016)

**Solving the artwork problem**

While it is reasonable that for performing arts works there cannot be an immediate representation, this does not mean that the situation for other art forms is simpler. The philosopher Nicholas Wolterstorff (1975) has shown that all kinds of art forms deploy their own intricate ontological structures. May it be “performance-works” or “object-works”, they both feature an ontological distinction between the work as such and its manifestation. It might only be that the object character of paintings, sculptures, and photographic prints misleads us to simplify their representations when we take the object that can be measured and described for the work itself.

We encounter such an erroneous simplification when we look at library catalogs that list different editions of works as unrelated to each other and thus leave it to us to reconstruct their provenance and context. The FRBRoo ontology provides a possible solution here when it distinguishes between an abstract **Work**, its concrete **Expression**, a specific **Manifestation**, and finally its physical **Items**. An author conceives a **Work**, puts it into words as one of possibly many **Expressions**, each published edition is then a **Manifestation**, and a single copy of such an edition would be an **Item**. This structure can be branched out to connect multiple entities of each of the classes, resulting in a tree-like structure that not only makes each entity traceable in relation to the original work but also relates different editions or translation of a work to each other.

The two first FRBRoo classes – **Work** and **Expression** – have subclasses for the domain of performing arts, which differ mainly in their denominations but feature the same relational logic: An abstract **Performance Work** is realized in a **Performance Plan**. In both cases, there may be more than one **Expression** or **Performance Plan** as multitudes of realizations. Where modeling of performing arts differs is that instead of the **Manifestation** there is a **Performance** (a subclass of the CIDOC CRM Activity) that realizes the **Performance Plan**, and thus replaces both, **Manifestation**, and **Item**.

Doerr, Le Boeuf, and Bekiari (2008) where the first who demonstrated the application of FRBRoo for performing arts and its advantage to be extensible with other CIDOC CRM classes.

The process from a **Work** to either an **Item** or a **Performance** builds one axis of FRBRoo. The other is created through the relationships between the expressions of different works. So, in the case of a Hamlet production in Geneva, the **Performance Plan** incorporates a French translation of the original English text and FRBRoo allows to connect these three works – the original play, the translation, and the staging. This way productions can be classified according the version or translation of a play they use. The same way a video documentation of the production extends the axis of references. The solution to what I have called the artwork problem thus lies in the possibility to draw networks of different steps of realization and adaption. I will regard this as sufficient for the moment even if incorporation is a very broad concept and FRBRoo does not foresee to qualify such relations any further.
Solving the description problem

Description in the domain of performing arts is happening on two levels: addressing the work (or a network of works, as we have just seen) directly and indirectly by describing artifacts that are related to them. The description of a present object like a poster or a sketch for a stage design is a much more straightforward task than making statements about an absent performance. Archives and museums of the performing arts collect and preserve these objects as a form of documentation. Being an ontology for museums, the description of physical items is a core feature of CIDOC CRM. So it is not a problem to describe the objects in performing arts collections with it. What is less prevalent is to describe their relationships to a Performance Work or Plan. In a 2012 paper, Patrick Le Boeuf has given the example of a wooden statuette held at the French National Library that was a part of the stage model for Sergei Radlov’s 1935 production of King Lear in Moscow. It is a Physical Man-Made Thing carrying a Design or Procedure that is associated with a Performance Plan and was involved in an Activity that aimed at a Performance of the play. Le Boeuf’s diagram (Le Boeuf 2012: Fig. 4) demonstrates two ways how the necessary connections can be made: It is part of the general design process of the stage, which can be either seen as associated with the Performance Plan or it can more directly be linked to the Performance itself as a technique that is used to construct the actual stage. (Le Boeuf, likely, does not see these two options as mutually exclusive but as complementary.) However, I think we are still lacking experience to agree upon how this is done best. But we can see how CIDOC CRM provides options to connect the description of related works to the performing arts works themselves.

Solving the agency problem

The two previous approaches to documentation elide that performing arts works are the results of team efforts and involve more than one author. They support descriptions of relationships between works, objects and concepts but they do not address the specifics of authorship and agency in performing arts. This was a central issue in the development of the data model for the Swiss Performing Arts Platform. Together with the Bern University of Applied Sciences we developed a data model that was published as a draft two years ago (Estermann and Schneeberger 2017).
At its core it is built on CIDOC CRM, FRBRoo and the draft of Records in Context available at the time. Figure 1 shows its core structure with the FRBRoo axis with Work, Expression, Manifestation, and Item. Next to it are the subclasses for performing arts with a slight variation in that the FRBRoo Performance appears twice, once as Performing Arts Production and then as the actual Performance. In the lower left we have the provenance documentation of an item according to RiC. 

Where the model actually extends its reference ontologies, is in the rendering of agencies as shown in a concrete example (Fig. 2). It assigns the individual functions involved in performing arts to the different steps of realization: Performance Work, Performance Plan, and Performing Arts Production. It also complements the relationship between the Performance Plan, and the play it is based on beyond a simple incorporation as it renders the roles defined in the play and their casting in the production.

The SPA data model by purpose is agnostic toward a specific technical implementation and – as comprehensive as it may be – is not complete. It is a stable basis, which has evolved into two parallel projects with different technical basis and distinct objectives. One is the Wikidata project for performing arts that was initiated by Beat Estermann. The other is its implementation for the future database of the Swiss Archive of the Performing Arts. The SPA specifications are our implementation of the SPA data model in RDF. It is a work-in-progress documentation, which iteratively grows with the migration of our legacy databases. So far, it covers places, persons, groups, venues, and performing arts works while still lacking some crucial aspects of these. Besides of its commitment to RDF, as the technical framework for our data, the specification makes three changes in comparison to the original model:

- It abstains from introducing own classes and properties but instead proposes the usage of the original classes and relations as defined by CIDOC CRM, FRBRoo and other ontologies.
- This goes along with the introduction of controlled vocabularies that provide domain-specific knowledge not covered by the used ontologies.
- And finally, it suggests the reification of agency in event-based classes instead of qualified relationships. This reification of agency becomes necessary to

Fig. 2 Theater production (modeling example), from: Estermann and Schneeberger (2017)
allow make further statements on the individual contributions.

Going back to the composition of performing arts works in FRBRoo, it is important to note that the Performance Work and Plan are subclasses of CIDOC CRM's Conceptual Object while the Performance itself is an Activity. There are two different options for introducing further Activities as renderings of agencies:

- We can connect Performance Work and Performance Plan by means of an Expression Creation entity.
- We can split Activities like the Performance or the Expression Creation into subordinate Activities.

This way the Expression Creation contains all Activities that are production-specific but are not happening on the stage. Split into sub-activities they can be assigned to specific Actors – like the production is carried out by the theater as a Group, direction is done by a Person and so on (Fig. 3).

The Performance entity, which realizes the Performance Plan, here turns into an abstract event because most statements that can be made about it apply to any performance of the work. Subordinate Performance entities then allow to make statements about specific performances. This can be the case if an actor was replaced on one occasion or if there is a recording of one specific performance. Therefore, we use the abstract Performance entity to describe the cast of a production and create at least one specific Performance entity for the premiere, which then holds the latter’s date.

Each of these Activities is not only connected with a higher-level Activity and an Actor who carried it out. It also gets a type from the controlled vocabulary (like “choreography” or “costume design”) as much as an optional label that supplements the type. An arbitrary comment can also be added (Fig. 4). Just think of a performance that does not take place in a theatrical environment and where the person who arranged the space holds a credit for “installation”. In our vocabulary we have “stage design” as an Activity, which comes close but is not the same as “installation”. Thus, in this case the activity would be typed as “stage design” but labeled as “installation”. Activities can also be further contextualized: They can fall within broader Activities like a long-term programming of a theater director or the general concept of theater seasons shared by all houses.

Open issues

The presented model for describing activities remains incomplete. Two of its most crucial shortcomings are that it omits orders and roles.

- The question of how to represent orders is a problem that is not addressed by the RDF standard and for which there is no broadly accepted workaround. What makes this even more challenging is that we not only need an option to express orders in RDF, we also need a software to edit them. (And no model that does not facilitate the editing of its content can be a useful model.) So we are in a catch-22 situation:
As I could hopefully show there is a slow but steady progress in modeling performing arts but further discussion is needed to come to a more comprehensive and viable modeling in this domain.

• Roles are problematic for other reasons and to call them problematic here might be surprising as the SPA data model already includes roles. A weak excuse for the omission of roles in our current specifications is that we do not have them in our existing data. We have records of some 65,000 Swiss productions in our database but these records only list the actors (in the correct order) but without their roles. And it would be great to easily get a list with all actors who have played a specific role. But this requires connecting the role played on stage with the role defined by the play. We likely can expect that there should be a stable understanding of the identity of Hamlet. But what to do with a production that reduces the number of fairies in Shakespeare’s *Midsummer Night’s Dream* from four to three? Did they just dispose fairy number four? Did they merge it with one of the others? To clarify such questions is no longer feasible when you are dealing with a large number of productions.

Fig. 4 Rendering single activities
Bibliography


Notes

1 https://www.wikidata.org/wiki/Wikidata:WikiProject_Performing_arts
2 https://sapa.github.io/spa-specifications