

Towards Performing Arts Information As Linked Data?

Patrick Le Boeuf

► **To cite this version:**

Patrick Le Boeuf. Towards Performing Arts Information As Linked Data?. SIBMAS 2012 Conference : Best Practice! Innovative Techniques for Performing Arts Collections, Libraries and Museums = À la recherche de l'excellence! Approches innovantes dans les collections et bibliothèques des arts du spectacle, Oct 2012, France. <hal-00807942>

HAL Id: hal-00807942

<https://hal-bnf.archives-ouvertes.fr/hal-00807942>

Submitted on 4 Apr 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Towards Performing Arts Information As Linked Data?

Patrick Le Boeuf, Bibliothèque nationale de France, Quai François Mauriac, 75706 Paris cedex 13, France; tel.: 33(0)1 53 79 42 23, email: patrick.le-boeuf@bnf.fr

Abstract.

Semantic Web technologies and Linked Data are of crucial importance for cultural heritage institutions, as they will allow them to share data, both among them and with any kind of other agents who will be thus enabled to reuse that valuable knowledge in a variety of ways. They rely, among other things, on ontologies that provide unified representations of the (often heterogeneous) information produced by such institutions. The FRBR_{OO} model (Functional Requirements for Bibliographic Records, object-oriented) has been under development since 2003 as an extension of the CIDOC CRM (International Committee for Documentation [of ICOM, International Council of Museums], Conceptual Reference Model), the ontology for museum information. The combination of FRBR_{OO} and CIDOC CRM provides a unified ontology for the information produced by both museums and libraries. Information relating to performing arts is explicitly taken into account in this ontology. This paper shows how performing arts, and resources that document them, are modelled in FRBR_{OO}, and how this could help libraries and museums of the performing arts to make the contents of their databases available as Linked Data.

Biographical Statement.

Patrick Le Boeuf is a library curator at the National Library of France (Bibliothèque nationale de France, BnF), Standards and Models Unit, where he has 'specialized' in conceptual modelling of cultural heritage information. He is a member of both the CIDOC CRM SIG (CIDOC CRM Special Interest Group) and the International Working Group on FRBR/CIDOC CRM Harmonization. From 2006 to 2009, he worked in the Performing Arts Department of the same institution, where he was in charge of, among others, the Edward Gordon Craig and the Louis Jouvet collections. He still occasionally writes articles and gives presentations on aspects of Edward Gordon Craig's work and thought.

Twelve years ago, on the occasion of the 23rd SIBMAS (International Association of Libraries and Museums of the Performing Arts) International Congress (Paris, 2000), I gave a presentation¹ on the conceptual model developed by IFLA (International Federation of Library Associations and Institutions) for bibliographic information, and how it could be used in the specific context of the information related to performing arts. That model, known as the FRBR model (Functional Requirements for Bibliographic Records),² was primarily designed to account for descriptions of physical resources, and transient entities such as performances do not fit perfectly in such a framework. I therefore suggested that a working group investigate, as a follow-up to the SIBMAS Congress, what was missing in the FRBR model for it to cover performing arts adequately. Although some email activity was actually carried out for a few weeks after the closing of the congress, it resulted in no substantial outcome. The whole story could have finished there, but, as things happen, the context of information science has undergone dramatic changes since 2000. A few months after the congress, Tim Berners-Lee, James Hendler and Ora Lassila introduced the notion of ‘Semantic Web’ in a seminal article published in 2001.³ In 2006, Tim Berners-Lee introduced another important phrase: ‘Linked Data.’⁴ The difference between ‘Semantic Web’ and ‘Linked Data’ is not always crystal-clear, but it seems that ‘Semantic Web’ refers to a specific technical structure – which can be used internally, within the boundaries of a local information system, not necessarily on the Web only –, while ‘Linked Data’ refers to the actual data that is made available on the Web through that technical structure, and interconnected from site to site. Libraries, archives and museums tend to be more and more interested in transforming their data into ‘Linked Data,’ i.e., in ‘using WWW standards to explore data interoperability not only between institutions but with the Web community and end users as well.’⁵ During the

¹ Patrick Le Boeuf, ‘Le spectacle vivant en tant qu’objet documentaire et le modèle conceptuel de données des FRBR (Functional Requirements for Bibliographic Records)’, in *Arts du spectacle, patrimoine et documentation: XXIII^e congrès international, Paris, 25-30 septembre 2000, Société internationale des bibliothèques et musées des arts du spectacle* (Paris: Bibliothèque nationale de France, SIBMAS, 2002), p. 162-175.

² IFLA Study Group on the functional requirements for bibliographic records, *Functional Requirements for Bibliographic Records: final report* (Munich: K. G. Saur, 1998). Also available online: <<http://www.ifla.org/en/publications/functional-requirements-for-bibliographic-records>> (accessed July 12, 2012).

³ Tim Berners-Lee, James Hendler, and Ora Lassila, ‘The Semantic Web’, in *Scientific American*, Vol. 5, No. 1 (2001), p. 29-37.

⁴ Tim Berners-Lee, *Linked Data* (2006-07-27, last change: 2009-06-18). Available online: <<http://www.w3.org/DesignIssues/LinkedData.html>> (accessed July 12, 2012).

⁵ Jon Voss, ‘The State of Linked Open Data in Libraries, Archives & Museums’, in *SemTech 2011, June 8, 2011*, slide 2. Available online: <<http://www.slideshare.net/jonvoss/linked-open-data-in-libraries-archives-museums>> (accessed July 19, 2012). For more information about Linked Open Data in the context of cultural heritage

same period, an international working group, consisting of representatives for both the library and museum communities, developed a new definition of the FRBR model as an extension of the conceptual model for museum information, the CIDOC CRM (Conceptual Reference Model),⁶ which was developed under the auspices of the ICOM CIDOC (International Council of Museums, International Committee for Documentation). In order to distinguish that new definition of the FRBR model from the original one, its authors decided to name it FRBR_{OO} (where ‘OO’ refers to the object-oriented formalism which is used in both CIDOC CRM and FRBR_{OO}). CIDOC CRM and FRBR_{OO} are well suited for the Semantic Web, as they express information in the form of triples that can easily be translated into RDF (Resource Description Framework) graphs.⁷ And the good news for SIBMAS is that the FRBR_{OO} model takes performing arts explicitly into account. The way FRBR_{OO} models performances was the topic of two presentations delivered at the CIDOC Annual Conferences of 2008 and 2010.⁸ It is only fair that the same information should now be shared with SIBMAS.

At the Core of the CIDOC CRM

Although it is neither desirable nor possible to give here a complete description of the CIDOC CRM, a brief presentation of its main structures is indispensable for a good understanding of the specific sections of the FRBR_{OO} model that deal with performing arts.

institutions, see the Web site entitled *LOD-LAM* (‘International Linked Open Data in Libraries, Archives, and Museums Summit’) at <<http://lod-lam.net/summit/>> (accessed July 24, 2012). The difference between ‘Linked Open Data’ and just ‘Linked Data’ is that there is no restriction on the way Linked Open Data can be reused by anyone in any kind of context, which can be very different from the context in which the data was originally produced.

⁶ CIDOC CRM Special Interest Group, *Definition of the CIDOC Conceptual Reference Model: version 5.0.4, November 2011* ([Heraklion, Greece]: [ICS-FORTH], 2011). Available online: <http://www.cidoc-crm.org/docs/cidoc_crm_version_5.0.4.pdf>, or: <http://www.cidoc-crm.org/docs/cidoc_crm_version_5.0.4.doc> (accessed July 12, 2012).

⁷ For more information on RDF, see the W3C (World Wide Web Consortium) page devoted to it at: <<http://www.w3.org/RDF/>> (accessed July 12, 2012).

⁸ Martin Doerr, Chryssoula Bekiari, and Patrick Le Boeuf, ‘FRBR_{OO}, a Conceptual Model for Performing Arts’, in *2008 Annual Conference of CIDOC, Athens, September 15-18, 2008*. Available online: <<http://cidoc.mediahost.org/archive/cidoc2008/Documents/papers/drfile.2008-06-42.pdf>> (accessed July 12, 2012).

Ifigenia Dionissiadou, ‘Archives Incorporating Museum Objects: the Case of Performing Arts’, in *2010 Annual Conference of CIDOC, Shanghai, China, November 8-10, 2010*. Available online: <http://cidoc.meta.se/2010/full_papers/dionissiadou.pdf> (text), <http://cidoc.meta.se/2010/presentations/pp_dionissiadou.pdf> (slide presentation) (accessed July 12, 2012).

The CIDOC CRM is said to be event-centred, which means that it focuses on the various events that a given object undergoes during its lifetime: it is produced at a given moment in time in a given place by a given person or group, modified on other occasions by the same person or group or other ones, and eventually possibly destroyed, or reused for the production of another object. In other terms, the event notion is what interconnects museum objects, agents, time-spans, and places. In addition, each individual object, agent, time-span or place is referred to through various kinds of appellations: titles, and identifiers such as inventory numbers, for objects; names of persons or corporate names, for agents; dates, which allow us to situate time-spans of events on a timeline; and place names. Consequently, a generic Appellation class is declared in the CIDOC CRM, along with its various specific subclasses. Another important notion in the CIDOC CRM is the notion of Type, which makes it possible to relate all the aforementioned classes to existing controlled vocabularies such as thesauri and classification schemes.

Although the CIDOC CRM consists of over 80 classes, a simplified schema (see Figure 1) with only 9 classes is sufficient to have a good idea of the overall structure of this sophisticated model.

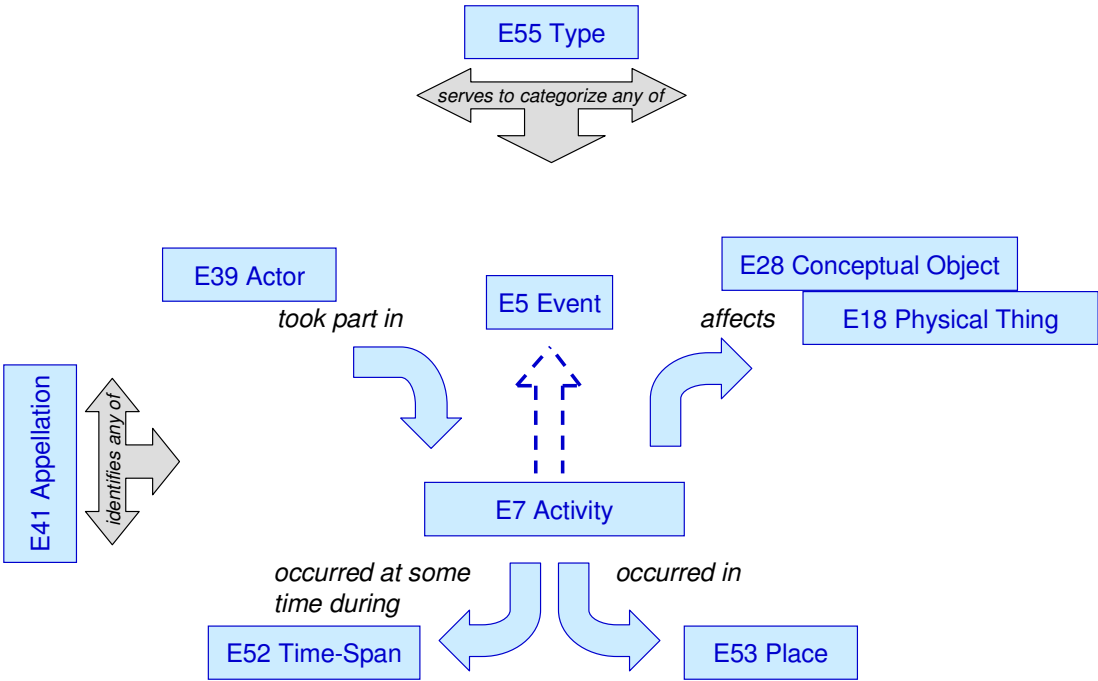


Figure 1

FRBR_{OO} and Performing Arts

The E5 Event⁹ class, and more specifically its subclass E7 Activity, is particularly relevant when it comes to performing arts, as it allows us to treat performances as documentation units. In FRBR_{OO}, a specific subclass of E7 Activity was declared for performances: F31 Performance.¹⁰ It is defined in a sufficiently flexible way to cover individual performances, runs of performances, or complete life cycles of productions, consisting of an original run plus its extensions and tours. In the context of the FRBR_{OO} model, performances are to be understood as activities that consist of behaving in a particular way that follows predetermined specific instructions, and in an audience's presence, be it either direct (live performances) or mediatized (through broadcasting or recording).

The specific instructions followed by performers during performances correspond in the FRBR_{OO} model to the class labelled F25 Performance Plan. An instance of F25 Performance Plan can be minimal; in extreme cases, it can consist of just one word: 'Extemporize!' Those instructions can exist in the performers' minds only, or they can be documented, e.g. in staging books, annotated scores, etc. However, it seems that no documentation of any kind can be said to convey the *complete* set of instructions that performers are supposed to follow while they are performing. An instance of F25 Performance Plan can also exist when the planned performances eventually did not take place, which allows the FRBR_{OO} model to account for documentation relating to both actual performances and unrealized projects.

Beyond such instructions, the FRBR_{OO} model acknowledges the existence of the overall conception that informs and permeates a given production (or musical performance). This notion covers the stage director's conception, or the choreographer's conception, or the conductor's or soloist's conception, etc. It is a set of mere concepts represented in the FRBR_{OO} model as class F20 Performance Work. A given instance of F20 Performance Work can result in any number of instances of F25 Performance Plan, as some specific performing instructions can evolve over time, while the production remains recognizable as a distinct creation. In the FRBR_{OO} model, a production is regarded as a work in its own right, distinct from the text (or score) which is being performed.

⁹ In the CIDOC CRM, each individual class is identified by a code, consisting of letter 'E' and a number, and a label.

¹⁰ In FRBR_{OO}, each individual class is identified by a code, consisting of letter 'F' and a number, and a label.

The relationship between performances and the works that are performed is established at the level of F25 Performance Plan, through a property named *R14 incorporates (is incorporated in)*.¹¹ This means that the instructions which are supposed to be followed during performances include a specific expression of the performed work, e.g., a given translation, or a given version.

Figure 2 illustrates this conceptualization with an example.

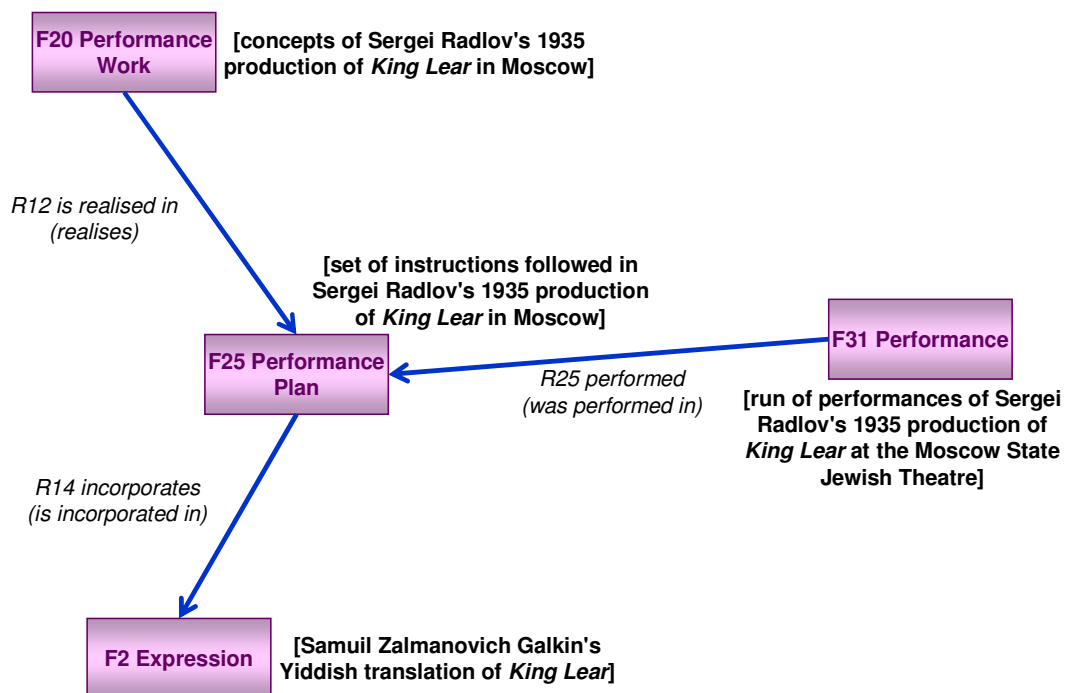


Figure 2

The way performing arts are modelled in FRBR₀₀ is also relevant for some currents of contemporary art, such as performance art, live art, and happenings, as an instance of F25 Performance Plan does not necessarily have to 'incorporate' a pre-existing expression of another work.

The physical items held by libraries and museums of the performing arts are related to the aforementioned entities in a number of ways. Incidental music is a distinct work, an expression of which is incorporated in the performance plan followed during the

¹¹ In FRBR₀₀, each individual property is identified by a code, consisting of letter 'R' and a number, and two labels, one from domain to range, the other, in brackets, from range to domain.

performances (see Figure 3). Set and costume designs are instances of the CIDOC CRM class E29 Design or Procedure that are used for the activities of building the sets and making the costumes, and associated with the performance plan (see Figure 4). Audio and video recordings, and photographs too, are linked to an instance of F31 Performance through an instance of F29 Recording Event, which results in an instance of F26 Recording (see Figure 5). Press clippings are carriers of works that are about an instance of F31 Performance (see Figure 6). Advertising materials such as posters and programmes are physical items that were produced specifically to serve as such in the activity of performance, a notion modelled in the CIDOC CRM through property *P19 was intended use of (was made for)* (see Figure 7).

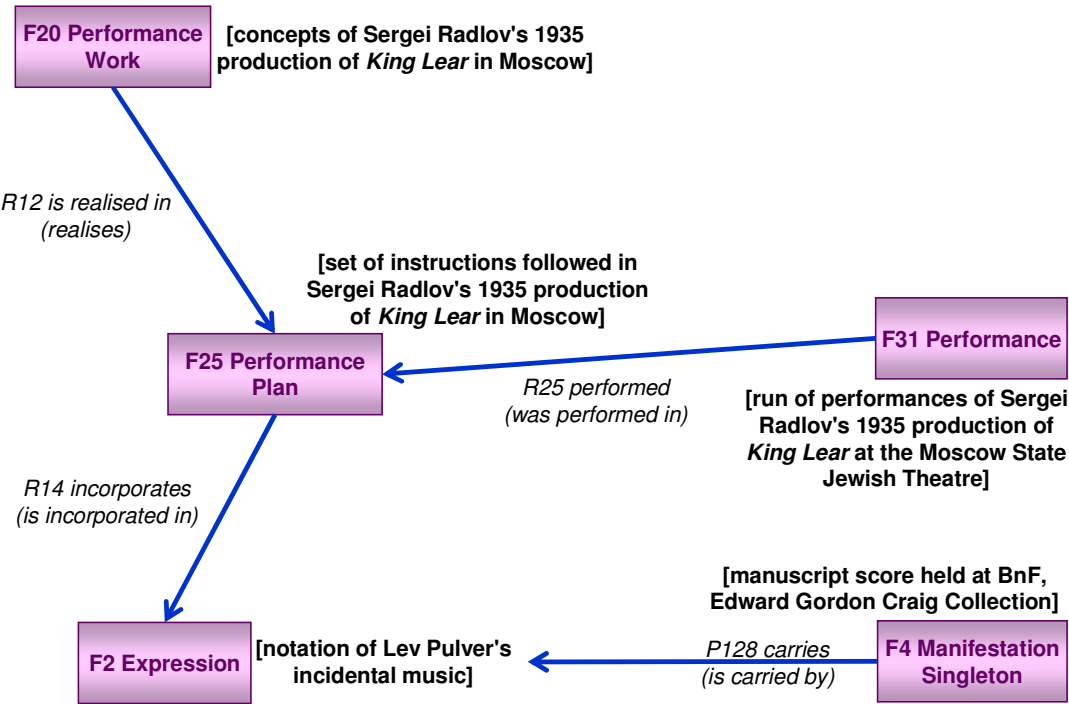


Figure 3

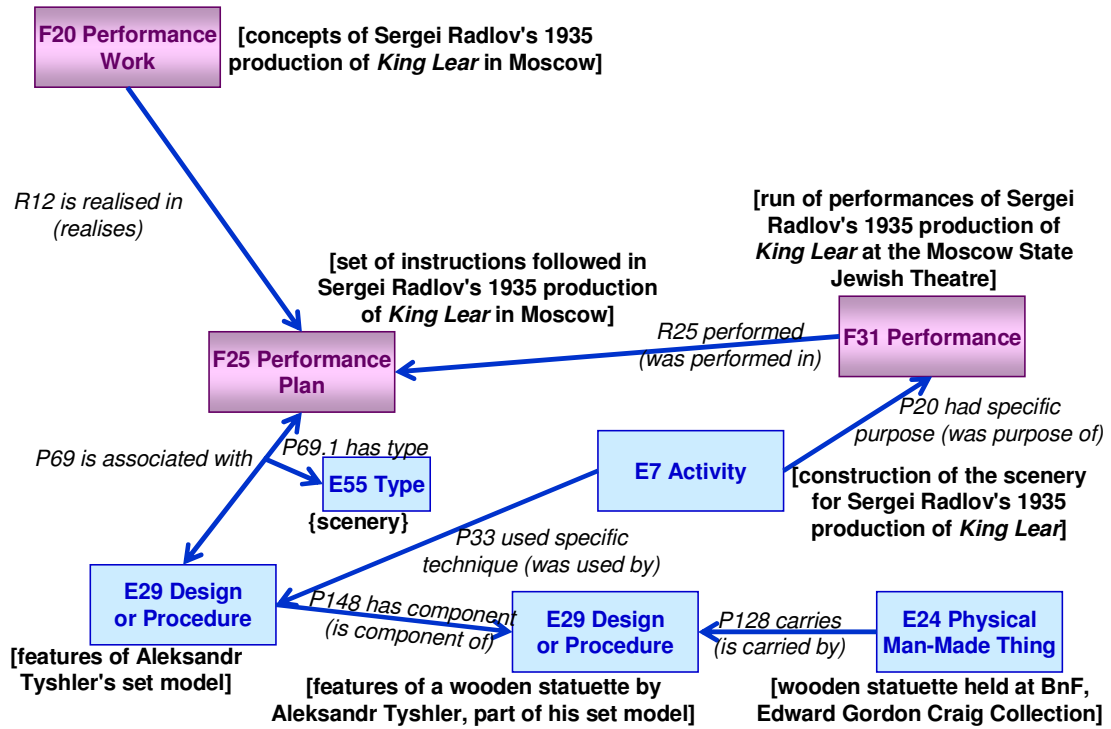


Figure 4

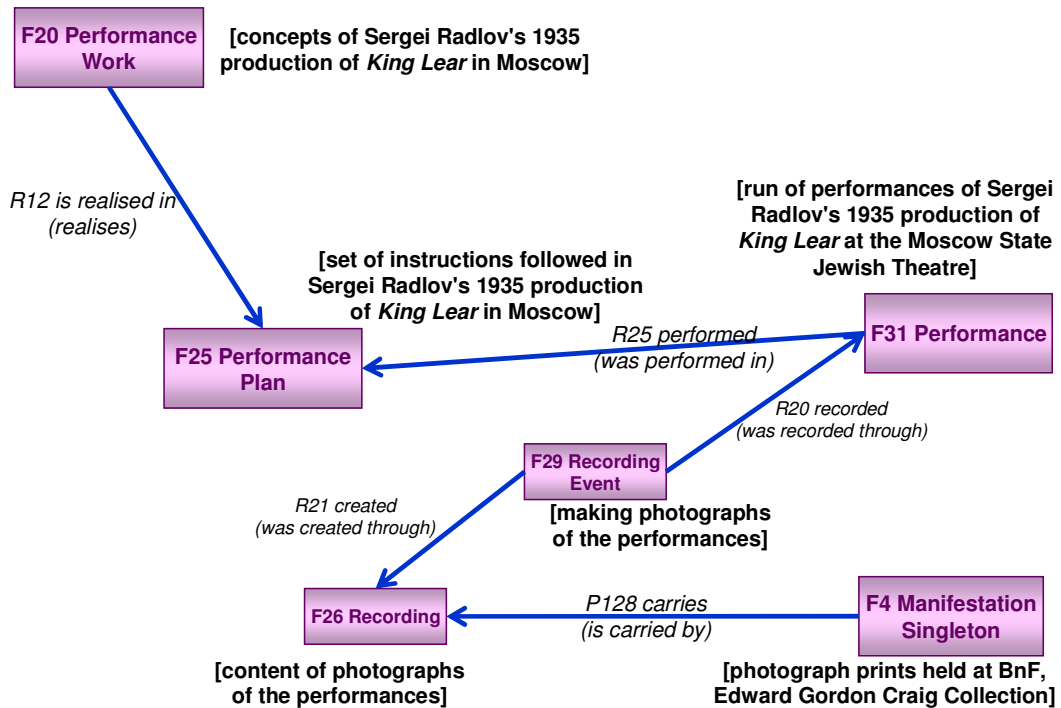


Figure 5

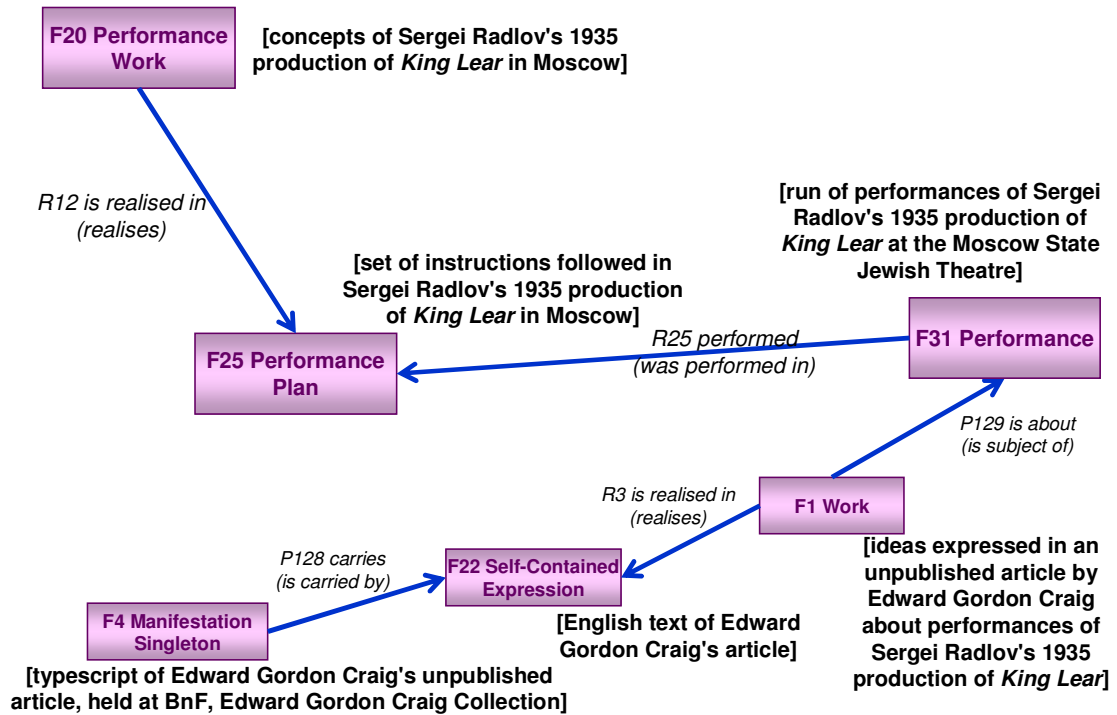


Figure 6

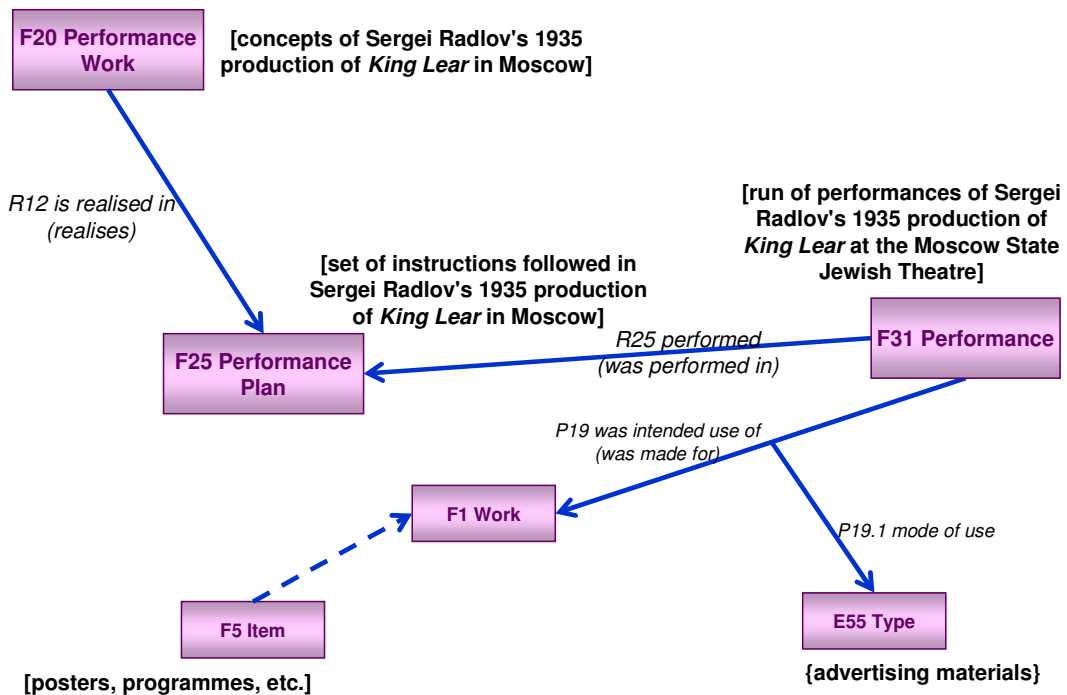


Figure 7

Performing Arts Information and Linked Open Data

At the heart of Linked Open Data lies the principle that the information produced by any institution should be made available for everyone else (even in very different domains of interest or activity) to reuse and recombine with their own information (or anyone else's information) in order to produce even richer information. The fact that FRBR_{OO}, if used together with CIDOC CRM, is an ontology that accounts for both library and museum data, is of particular interest for SIBMAS, which, as its very name indicates, has both libraries and museums as members. Thanks to such an ontology, it would be possible for a library to reuse data produced by a museum about a given production, if that library happens to hold physical items that document the same production, or vice versa. The data produced by various members of SIBMAS could be aggregated in thematic portals devoted to given aspects of performing arts (such as: theatre activity in a given country, or on a given period, or of a given artist who has an international career, etc.). Cultural heritage institutions could benefit from the data produced by other types of institutions – or the other way round. Indeed, cultural heritage institutions must take the risk that the information they would publish as Linked Open Data might end up in any kind of context which they would never have imagined; as Jon Voss puts it, metadata you publish as Linked Open Data is 'mostly out of your control. Yet it can be a part of something much greater than any of the component parts.'¹² Some projects aimed at exposing performing arts information as Linked (Open) Data are already in existence, and FRBR_{OO} has the potential to federate them. Here is a brief survey of four such or similar projects.

In 2010, the DEN Foundation (Digital Erfgoed Nederland / Digital Heritage Netherlands), the Library of the University of Amsterdam, and the Theatre Institute of the Netherlands launched the 'Dutch Culture Link' project, which consists of publishing 'performance metadata and related digital material as linked data.'¹³ This project uses heterogeneous metadata from the Library of the University of Amsterdam and the Theatre Institute of the Netherlands, which it integrates through its specially developed 'theatre play

¹² Jon Voss, 'The State of Linked Open Data...', slide 16.

¹³ Lukas Koster, 'Dutch Culture Link', in *CommonPlace.net*, posted on October 7th, 2010. Available online: <http://commonplace.net/2010/10/dutch-culture-link/> (accessed July 18, 2012).

data model.’¹⁴ Although FRBR_{OO} would not suffice to solve all the problems encountered by the project team,¹⁵ it would make the data resulting from such a project more easily available for similar projects, and progressively the community would benefit from an ever-increasing dataset on performing arts.

Another project aimed at publishing performing arts information as Linked Data was the SPACE/Travelogue project, led by VTi (Vlaams Theater Instituut) in Brussels. SPACE (Supporting Performing Arts Circulation in Europe) was a three-year project (2008-2011) the purpose of which was ‘to enhance data exchange and to develop research about international mobility in the performing arts production’ through the development of Travelogue, ‘an online prototype meant to monitor the international mobility of the performing arts using Semantic Web technology.’¹⁶ No library or museum was involved in that project,¹⁷ but clearly both cultural heritage institutions and the type of institutions involved in the SPACE/Travelogue project would benefit from reusing each other’s information through Linked Data mechanisms. A specific ontology had to be developed for that project.¹⁸ Mappings between FRBR_{OO} and the SPACE/Travelogue ontology would enable SIBMAS and SPACE members to share their information on the Semantic Web.

GloPAD (Global Performing Arts Database) is a ‘Web accessible database combining digital objects (...) with authoritative and detailed descriptive information about each item.’¹⁹ It aggregates heterogeneous metadata from a number of institutions by mapping them to a common metadata structure built around the event of performance. This approach is close to

¹⁴ Lukas Koster, ‘Linking Library and Theatre Data’, in *International Group of Ex Libris Users 2011 IGeLU Conference, University of Haifa, September 11-13, 2011*, slide 16. Available online: <<http://igelu.org/wp-content/uploads/2011/09/Linking-library-to-theatre-Koster.pdf>> (accessed July 18, 2012).

¹⁵ ‘If both the library catalogue and the theatre production database were FRBRised, we could in theory link on the Work level and cover all individual versions. But we would still need a matching mechanism on that Work level of course. (...) If the published datasets have a flat internal structure (...), then they just don’t provide sufficient added value for third party use. In order to make your published open data useful for others, you have to make implicit relations explicit. And this (...) requires a lot of processing.’ Lukas Koster, ‘Missing Links’, in *CommonPlace.net*, posted on March 28, 2011. Available online: <<http://commonplace.net/2011/03/missing-links/>> (accessed July 18th, 2012).

¹⁶ Vlaams Theater Instituut, *SPACE: General Background*. Available online: <<http://vti.be/nl/over-vti/projects/space-general-background>> (accessed July 11, 2012).

¹⁷ ‘The members of SPACE share a position in between politics and the artistic field in their own countries, work as centres of information, promote the (performing) arts on a national and international level and are experienced in supporting and running European cultural projects.’ *Ibid.*

¹⁸ Koen Van der Auwera and Bob Van Landuyt, *Travelogue Data Model Documentation* (Brussels: SPACE, October 2011). Available online: <http://www.arts-mobility.info/documents/Travelogue_web_DEEL2.pdf> (accessed July 18, 2012).

¹⁹ Joshua Young, *On Metadata: Performing Arts Materials in Our Digital World* (Ithaca, NY: GloPAC (Global Performing Arts Consortium), [2006]). Available online: <http://www.glopac.org/pdf/On_Metadata_v3.pdf> (accessed July 18, 2012). Although the documentation available on GloPAD does not explicitly state that GloPAD makes use of Linked Data technology, it is likely to be a development in the near future.

what is done in CIDOC CRM/FRBR_{OO}, and there are great similarities in the ways both GloPAD and FRBR_{OO} model theatrical productions.

ECLAP (European Collected Library of Artistic Performance) is an ‘online archive for all the performing arts in Europe,’²⁰ a performing arts digital library that aggregates both digitized items and their metadata from various institutions. Its target audience consists mainly of teachers and students, and the techniques employed focus on e-learning.²¹ The ECLAP documentation mentions both the CIDOC CRM and the FRBR_{OO} models.²² The information aggregated by ECLAP is aggregated in turn in Europeana through a mapping to EDM (Europeana Data Model). As things happen, a taskforce has just been formed (in July 2012) with the objective of harmonizing EDM and FRBR_{OO}, which in the future should result in Europeana serving performing arts better.

Conclusion

As Lukas Koster puts it, ‘people calling on institutions to publish their data as linked open data tend to say it’s easy once you know how to do it.’²³ I am certainly not claiming that it is easy. Our legacy datasets often lack precision, which makes it difficult to map them to FRBR_{OO}. But it is worth a try, as the FRBR_{OO} ontology has an important role to play in the publication of performing arts information as Linked Data. The idea is not that the aforementioned projects in that field should change their own metadata structure (which they would not be willing to do, anyway) and adopt FRBR_{OO}. The idea, on the contrary, is that each individual existing database can keep its own ontological commitment, while the FRBR_{OO} model, used in combination with the CIDOC CRM, can provide a common ontology for all Linked Data projects that focus on performing arts, and relate them to more general Linked Data projects that have to do with library and museum information.

²⁰ ECLAP Web site, <<http://www.eclap.eu>> (accessed July 18, 2012).

²¹ Pierfrancesco Bellini, Paolo Nesi, Michela Paolucci, and Marco Serena, ‘Models and Tools for Aggregating and Annotating Content on ECLAP’, in *2011 IEEE International Symposium on Multimedia* (Washington, DC: IEEE Computer Society Press, 2011), p. 210-215.

²² European Collected Library of Artistic Performance, *DE4.1 Metadata Descriptors Identification and Definition* ([Florence, Italy]: ECLAP, 2011). Available online: <<http://pro.europeana.eu/documents/858753/0/ECLAP+DE4.1+Metadata+descriptors+Identification+and+Definition.pdf>> (accessed July 18, 2012).

²³ Lukas Koster, ‘Missing Links’...