Digital art posterity: building a data model for digital art corpora

Louise Fauduet
Bibliothèque nationale de France
France
louise.fauduet@bnf.fr

Céline Thomas
Université Paris 8
France
celine_thomas@gmx.fr

Chu-Yin Chen
Université Paris 8
France
chu-yin.chen@univ-paris8.fr

Adèle Sicre
Université Paris 8
France
adele.sicrevidalou@gmail.com

Nola N'Diaye
École nationale des chartes
France
nolandiaye@gmail.com

Florent Aziosmanoff
Living Art Lab
France
florent.aziosmanoff@gmail.com

ABSTRACT

The research project "Digital art and posterity", February 2015-February 2018, aims to contribute to perpetuating digital artistic heritage by developing a general descriptive system of digital works of art which provides a way to model the technical and artistic characteristics of the work, its significant properties and its formal principles in text form.

This research is performed in collaboration between three institutions: the National Library of France (Bibliothèque nationale de France, BnF), the INRéV laboratory of Paris 8 University and the Living Art Lab. The partners come from different fields — those of conservation, of artistic creation and of academic research — and combine their skills and experiences to devise a conceptual model and a data model true to the specificities of the project's corpora.

This paper presents the main attributes of the proposed information model and ontology, so that all institutions involved in the collection, preservation and diffusion of digital art can provide input regarding its pertinence and interoperability.

CCS CONCEPTS

• Information systems~Document topic models • Information systems~Content analysis and feature selection • Information systems~Ontologies • Applied computing~Performing arts • Applied computing~Media arts • Applied computing~Digital libraries and archives

KEYWORDS

Digital Art, Information Model, Data Model, Interactivity, Transdisciplinarity, Versioning

1 INTRODUCTION

In recent years, there have been an increasing number of papers at iPRES on the topics of emulation solutions and digital art preservation [1]-[6]. Both issues are often intertwined, as computer-generated audiovisual artistic creations that are of a composite nature (sound, picture, video, motion picture, etc.) require preservation processes ranging from emulating the original to recreating the code and/or replacing the input and output devices with technologies of another nature. The obsolescence of hardware and software involved in the presentation of digital art is extremely rapid. And yet we need to ensure that future generations will be able to experience these works while maintaining the possibility to interact with the technical and artistic content, in order to affect the course of its performance.

"*Art numérique et postérité*" ("Digital art and posterity") is a project based on collaboration between a national library, a university department and a private creation and research partner. The project aims to contribute to the preservation of digital art by working on its description.

We are aware that there is a wide range of definitions of digital art among the different communities of artists, researchers and curators. The relevance of the information model we are developing may be biased by the corpora and the professional environment of the project's partners, even though the members come from different backgrounds and the collections are diverse. We will include in this paper a presentation of the partners, their work environment and the art used in the project to give an accurate picture of the information model's immediate applicability.

We will then present the conceptual model designed in the project. The main characteristics of an information model for contemporary digital art are the place given to the artist's intentions for the presentation of the artwork, as reconstituted by the curators or as expressed in his or her own words, and the importance of accounting for successive versions of an artwork.

We will discuss the challenges associated with the modeling process, and the way building a database and an ontology from the conceptual model impacts the overall description of digital art.

2 RESEARCH AS A *MÉNAGE À TROIS*

We will first introduce the actors of the project "Digital Art and Posterity", their contributions and missions. We then will clarify the meaning of "*art numérique*" as we defined it using the corpora of each collaborator as a representative panel of digital art. We will explain the specificities of digital art in order to

highlight the challenges related to its preservation. Finally, we discuss the issue of the mass treatment of digital artworks.

2.1 A versatile project team

The research project "Digital Art and Posterity" is a project shared between three institutions: the Bibliothèque nationale de France (BnF, the National Library of France), the Images Numériques et Réalité Virtuelle research team (INRéV, Digital Images and Virtual Reality), part of the Arts des Images et Art Contemporain laboratory (AIAC E4010, Images Arts and Contemporary Art) of Paris 8 University, and the Living Art Lab, an independent art studio.

This project was made possible thanks to the Labex Arts-H2H, cluster of excellence in arts and human mediations[1], laureate of the program "Investing for the Future" since 2011[2]. All of the Labex's research projects are built around the principle of transdisciplinarity between different fields in the humanities. This experimental dynamic made it possible for the project partners to meet and to collaborate.

The first partner is the BnF, the project leader, and specifically its Audiovisual department. The BnF has a mandate to collect a vast array of publications. Over the years, it has developed best practices to describe documents ranging from manuscripts to video games. The Audiovisual department has built expertise in the field of digitized and born-digital heritage preservation. In the past few decades, the BnF has been gathering a digital artwork collection through legal deposit. It is mostly composed of documents on a physical medium but also entails web archiving and contents distributed online. One of its objectives in our project is to enrich its collection of digital art. It also aims at bettering access to its collections using more standardized description and access processes.

The second partner is the INRéV research team of Paris 8 University. For 30 years, this team has been developing a unique manner of apprehending digital creation. The laboratory's aim was and remains to continuously create, innovate and hybridize art, science and technology. This team's main figures are protagonists of French digital art history. Their artistic practices have given them an insider knowledge of digital techniques and their evolutions. One of the specificities developed by the research team over the years is interactive art with an autonomous and sometimes intelligent behaviour coupled with a reflexion on the creative process. Combined with the undergraduates' and graduates' productions at the Arts et Technologies de l'Image department[3] (ATI, Computer Graphic and Art Department), these artworks constitute a remarkable corpus. INRéV joined the project to start working on its preservation strategy as none is in place at the University.

The third partner is the Living Art Lab, an independent studio which was previously part of Le Cube, a center for digital art creation. Le Cube, through the Living Art Lab, directed and produced interactive art installations, by offering the creators financial resources and IT skills. With experience in fostering and facilitating digital art, the Living Art Lab brings to the project its vision of digital art and its creations. Its goal is to promote and preserve Living Art installations. These form a coherent corpus with a specific approach to interaction which broadens our project's panel.

In addition to this original trio, we called on other skills to enrich the project by recruiting interns and experts with various backgrounds: contemporary art philosophy, data processing and information science, management of audiovisual archives, computer science, art history, art conservation and digital humanities.

The versatility of the project team makes for a stronger scientific project. We rely on collective intelligence, which leads us to interesting moments of constructive misunderstandings and allows us to move forward to better solutions.

First and foremost, the "*art numérique*" that the partners want to describe and preserve had to be defined and matched with the corpora.

2.2 What is "*art numérique*"?

A conventional definition of "digital art" would be a diverse set of artistic creations based on the specificities of the computer language and the IT system. It characterizes artworks produced by computers and that can be rendered with electronic devices. This art genre began to be produced in the early 1960s and has not ceased developing since then through a wide range of artistic expressions. In Québec *art médiatique* is seen as a whole, and *art numérique* is understood specifically as online art. In the United States and in the UK, cultural institutions manage time-based media art as an ensemble[4], of which software-based art is a subset. In France, *art numérique* or digital art is currently the most widely used category [7]. Its history encompasses the permanent evolution of digital technologies (for instance, virtual reality or augmented reality) and the developments in connected disciplinary fields (such as artificial intelligence, artificial life, cognitive science or robotics).

While we agree with this definition, it isn't precise enough to fit the complexity of our corpora. Among the properties we need to express in our description are the characteristics related to the initial publication and production of the artwork, since the corpora of the partners have different publication histories: published works at the BnF, research results at the ATI-INRéV, exhibited objects at the Living Art Lab.

The first essential element of any description of digital art is that computer programming is at its center. In the project, we

[1] http://www.labex-arts-h2h.fr

[2] http://www.enseignementsup-recherche.gouv.fr/pid24578/investissements-d-avenir.html

[3] http://www.ati-paris8.fr/index.php?lang=en

[4] http://mattersinmediaart.org, https://www.nyu.edu/gsas/dept/fineart/events/time-based-media.htm

incorporate both interactive, intelligent or behavioral artwork and the computer-animated image, since all these art forms are computer-based. The same artistic impulse governed their creation, as is shown in the ATI-INRéV corpus. They have to be preserved together according to specific strategies, and the computer systems used to create such moving images have to be described in the information packages.

Furthermore, building on earlier research on the role of the viewer in creation [8], digital art explores the ability to create an interactive relationship between the program and its external environment and in particular with the spectator [9]. The BnF's corpus offers good examples of elementary multimedia interactive artworks with its art CD-ROMs collection, produced in the 1990s. Although these artworks are described in the BnF's catalog, their bibliographic records are based on a common model for all electronic documents. The existing technical information is enough to render the discs in the reading rooms now, but insufficient to allow long-term preservation. There is, for instance, no record of the basic procedural rules needed to replay the art, although changes in technological context gradually limit our ability to perceive the artwork as intended.

Another expression of the interactive relationship that can be produced by digital art may be defined as "intelligent", "behavioral" or even "autonomous". Artworks in this category, more experimental, have been created within the ATI-INRéV and Living Art laboratories since the mid-1980s. They require captors and actuators, and/or virtual or augmented reality systems; their 3D images are generated in real-time by artificial life or intelligence algorithms. From these artworks emerge autonomous or intelligent behaviors that have been defined as a new form of interactivity. Couchot, Tramus and Bret named it "second interactivity" [10]; Aziosmanoff prefers the term of "Living Art" [11], [12], and Chen, using the concepts of complexity and enaction [13], calls it "enactive digital art" [14]. The description of these artworks has to include their behaviors, their self-actualization processes, and the interaction systems.

Figure 1: *La Funambule virtuelle*, Alain BERTHOZ, Michel BRET, Marie-Hélène TRAMUS, 2000. Interactive installation. Interaction with the spectator through position and rotation captors attached to a belt, with an intelligent 3D computer-animated character developing a rebalancing strategy in real-time [15]. Uses a neural network program (Anyflo)[5].

The most complex description and preservation challenges we have to face come from parts of the corpora heavily reliant on fundamental research on the art-science themes. Research domains such as artificial art/life, artificial art/intelligence, art/enaction, art/emergence have led to very unique and innovative artworks. Examples in our corpora use made-to-measure software ("Anyflo" by Bret[6], "Rodin" by Huitric and Nahas [16]), devices and platforms.

Figure 2: *Morphogénèse*, Chu-Yin CHEN, 1996 (film; 7'13). Generative artwork, artificial life. Software: Anyflo. Result of researches on the creation of a virtual world (biosphere) within which the virtual creatures self-generate according to genetic algorithms [17][18].

While interaction gives the artworks the appearance of unpredictability and instability due to the individualized experience of the spectator, the real instability of the digital objects comes from their successive versions in different software and hardware. In the thirty-odd years of INRéV's existence, the artist-researchers have had to either update their artworks or let them die. Whether at the artist's initiative or at a museum's or gallery's request, repairing artworks for a new exhibition requires different strategies: migration, emulation, recreation. Changes to the original art range from slight modifications to complete distortions. Over time, the successive versions of the artwork can diverge from replicas to form something akin to a series based on the same concepts. Describing versioning of complex objects is one of the main challenges of our project.

Along with this challenge emerges another of our issues: how can we make the distinctive nature of digital artworks work with mass processes?

[5] http://www.anyflo.com/bret/art/2000/fun/fun.htm
[6] http://www.anyflo.com/bret/art/1986/Anyflo.htm

2.3 Dealing with mass

The project partners have very different processes for collecting, describing, preserving and accessing their digital art collections. None of them are currently geared towards presenting digital art in an exhibition context, although it has happened and may happen again.

The Preservation & Art - Media Archaeology Lab (PAMAL[7]) at the Avignon School of Art has restored digital works of arts in a media archeology perspective, to study the impact of the material context of creation on the artworks' reception [19]. But the partners of our research project have different constraints, and systematic recreation of artworks on software and hardware matching the original is not practical given the volume of collections concerned, and access requests that are relatively unpredictable compared to a planned exhibition project, for instance.

The National Library of France has existing data and metadata repositories, and access systems. Any solution tailored for digital art must be included in these generic systems that combine data on all types of materials, and specific information might lose visibility in the process. There is a catalog and a data repository using semantic web technologies[8]. There are two digital archives, the Scalable Preservation and Archiving Repository (SPAR) [20] and the audiovisual archive. The latter is the current repository for digital art content, but has limited metadata querying capabilities. It is combined with a display system for the audiovisual research reading room, where emulation processes are only very partially automated. Library assistants deal with requests for materials requiring emulation, and have the tools and skills to provide access to relatively standardized content such as CD-ROMs from the 1990s. Any other level of installation and emulation currently requires the intervention of a technician or engineer on a case by case basis.

In contrast, INRéV has no preexisting processes, repositories or tools for preserving its digital art collections, even though it values the works created by its faculty and students highly, both as art and as teaching devices.

Working on an information model for digital art is a mutually beneficial project. It can improve the quality of preservation at the BnF as the information requirements are implemented throughout the digital preservation process with the input of digital art specialists. It can also help the Paris 8 University and the Living Art Lab develop new research on the properties of digital art, and help them evaluate whether the BnF can be a partner in their preservation needs.

The resulting data model has to be both generic enough to describe any digital artwork despite the variety of techniques and approaches in this field, and specific to allow the descriptive and preservation metadata collected to be used for preservation planning and prompt choices of access solutions.

[7] http://pamal.org/wiki/Accueil

[8] catalogue.bnf.fr; data.bnf.fr

3 THE CONCEPTUAL MODEL

The data model we have to build can only be constructed after we have set a conceptual structure. To describe it here, we will introduce our working methodology: we drew up an inventory of the digital artworks from our corpora, identified the information necessary for their preservation, and collected as much data as was judged useful. The structure of the conceptual model that organizes the gathered information was then developed. It was constructed according to the digital artwork's lifecycle and respects its fluctuating nature from genesis to exhibition.

3.1 Awakening sleeping beauties

"Once upon a time, our artworks were forgotten, asleep in the midst of a wild forest. Going through the brambles, we managed to discover marvelous hidden collections; and now the awakening begins..."

The project's team first needed to identify the "sleeping beauties" of our corpora.

The history of creation at the ATI-INRéV has been documented through exhibitions and the memory of this institution has been transmitted by the artist-teachers. However, an inventory of the artworks created by the professors and students has never been made.

The Living Art Lab has always been concerned with documenting the works it fostered. The artworks have been described on the Cube's website at the time of their exhibition[9]. The Living Art Lab's method includes a focus on collecting artists' and IT engineers' testimonies[10]. Nevertheless, there is no digital archive of the collection.

As for the BnF, there was no typology or keywords in their catalog that marked documents as digital art.

We inventoried over 300 works: around 200 from the BnF, 120 from the ATI-INRéV, 70 from the Living Art Lab. We established a limited number of categories: 3D computer-animated films; generative artworks; interactive, intelligent or behavioral installations; virtual reality and augmented reality artworks; gaming creations; interactive multimedia artworks.

In order to evaluate the information and data required for the identification of the artworks, we compared the descriptive records of institutions outside the project, such as the ones created on Navigart and reachable on Videomuseum's website[11], and we interviewed curators. Simultaneously, we had to design a test sample which was technically, historically and aesthetically representative of our corpora, but also narrow enough to allow us to analyze it thoroughly and to specify the information necessary for the preservation of digital artworks

This identification of data performed, we had to awake our beauties; not with a kiss, but with documentation and data

[9] http://lecube.com/fr/coproductions_151

[10] http://lecube.com/fr/living-art-lab_154

[11] http://www.videomuseum.fr

collection. The lack of existing information has led us, as others before, to try to capture information at the source: from the living artists or contributors. We could rely on previous research from the museum community, such as the "The Variable Media Questionnaire" of the Variable Media Network [12] [21], the "Questionnaire for New Media Works" of the DOCAM Research Alliance (Documentation and conservaton of the media arts heritage)[13], or the recommendations of Matters in Media Art[14]. Through several case studies, the Digital Art Conservation project[15], initiated in 2010 by the ZKM (Zentrum für Kunst und Medientechnologie Karlsruhe), was able to apply concrete preservation strategies for a wide range of digital works based on artist interviews [22]. Discussions with researchers and curators in Québec and New York have confirmed this approach to be common practice among heritage institutions.

Technical information about the artwork seemed the aspect the most often absent from records. We began by sending a technical questionnaire on digital artworks to the artists, producers and contributors to gather information on how the work and its system operate (required IT environment, peripheral devices and interfaces, mode of interaction and mode of installation, etc.). We then interviewed them in person, as they were interacting with their artworks; this strategy has been well received so far. We record both the play-through of the emulated artwork on the computer and the artist interview on video. Our partners at the ATI-INRéV have also introduced the method known as the "*entretien d'explicitation*" ("clarifying interview technique") [23] to the project. The *entretien d'explicitation* is an introspective interview technique designed to lead the interviewee to access previously unconscious memories of the creation process and past actions. We aim at retrieving data on both the objective steps of creation and of the implicit knowledge related to the act.

Digital artists are aware of the ephemeral nature of their art. Many have chosen to reassess and recreate their works in order to continue to exhibit or perform them, but nevertheless proceed in different ways. For instance, some of them are forced to slow down technological improvements when they reactivate their works in order to respect the aesthetic aspects of the techniques used at the origin of the creation, while others push forward their works with technological evolution and overwrite previous versions of the artworks, etc. To evaluate artists' different preservation and description requirements according to the beliefs and ideas on which their creative process was based, we organized a study day on December 8 2016[16]. These international artists and digital art pioneers, discussing their expectations and their involvement regarding the future of their artworks, have shown a strong interest in digital art preservation. They have defined the significant properties of their artworks, although they use different phrases for the essential characteristics to maintain through each generation of the piece, instinctively or deliberately.

Whether the artist is the only legitimate source for the description of the works and for preservation strategies choices remains to be determined. Part of our efforts in building an information model for digital art has certainly been dedicated to balancing the artist's word and objective data. While this is not a breakthrough discovery, we aim at translating this line of reasoning into concrete rendering of the data model. It will also be a valuable tool to determine what preservation quality level the BnF is ready to grant to digital artworks, and whether certain types of digital art are more suited to the (relative) mass processes of a library.

3.2 Genesis

The descriptive system we wish to devise includes the concepts and key principles of digital art together with preservation information related to the digital objects. It captures the artwork from its creation to its diffusion, in various possible contexts: exhibitions, classes, reading rooms and so on. We designed our information model around two states of the artwork: the artwork during the creative process, or "artwork-to-be", and the artwork we encounter as a spectator. We present here the "creation process" section of the model, followed by the "reception process" section.

The knowledge and experience of the project partners dealing in artistic creation have given us an opportunity to explore the creative process in detail. Our aim is to preserve and give access to the craftsmanship within the artwork. The first part of our system thus describes the artwork in its creative process: it begins by the "conceptual artwork", where we develop the genesis of the artwork, and then continues to the "realized artwork", where we detail the production process.

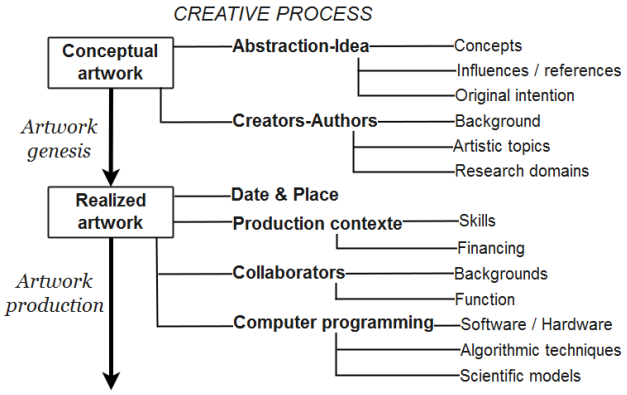

Figure 3: Simplified conceptual model: creative process

[12] http://variablemediaquestionnaire.net
[13] http://www.docam.ca/fr/outils/questionnaire-de-conservation.html
[14] http://mattersinmediaart.org/acquiring-time-based-media-art.html
[15] http://www.digitalartconservation.org/index.php/en.html
[16] http://artnumeriqueposterite.labex-arts-h2h.fr/fr/content/programme

We need to describe the abstract idea of the artwork. This "conceptual artwork" emerges from the individuality of the author(s), their cultural and technical backgrounds, their artistic and scientific references and influences, their intention, etc. In this entity, we highlight the intellectual approach of the artist: the concept that the artworks are intended to convey and the starting point of the artistic project.

This abstract idea is materialized through digital tools and media; it evolves according to the reality it encounters. We named this information entity "realized artwork". For instance, it encompasses the processes involved in the collaboration between the artists and the programmers, technicians, mathematicians, etc. Their contribution has to be understood not as merely technically proficient but also as creative: they too belong to the genesis of the artwork.

Another meaningful element of the creative process is the context of creation: the date, place, event or occasion of execution, the financial model and the production circuit it relied on.

The heart of the piece is the computer program — which is not necessarily written by the artist. We naturally include information entities dedicated to the artwork's internal and external tools, whether already existing or produced specifically for the work, and the way they were used. "Internal system" describes all the software, algorithms and their evolutions. "External system" covers any type of hardware included in the artwork.

Parts of the corpora comprise art generated according to complex theoretical models, which forms a real challenge to defining a general descriptive system. For instance, artificial intelligence is based on algorithmic techniques such as neural networks; and artificial life uses genetic algorithms. Beyond considerations of aesthetics or techniques, the conceptual model needs to accommodate complex relationships between entities. We choose to build a network model where information is ranked from the general to the specific. Our information model has to be open and flexible in order to conform to any digital artwork type.

3.3 Performed, exhibited, rendered: the many lives of the artwork

We designed our information model around two states of the artwork: the artwork during the creative process, or "artwork-to-be", and the artwork we encounter as a spectator. This second part of the model, describing the reception process, further distinguishes between the "performed artwork" and the "experienced artwork".

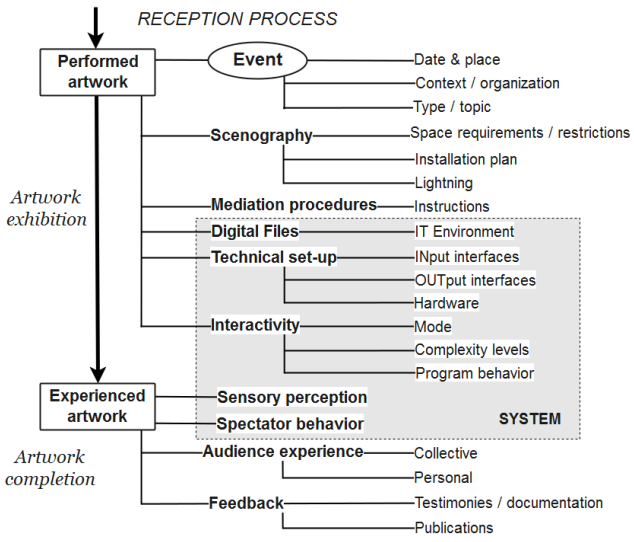

Figure 4: Simplified information model: reception process

Once the artwork has been materialized, it is presented to the public. This is the best-known state of the piece, which we call the "performed artwork". Many of the works in our corpora are based on the concepts of movement, real-time execution and interaction. Their appearance changes continuously and also adapts to the public with whom the work interacts; the works exist under a certain form at a certain time. Thus they are not just artefacts but relationships, connecting the virtual environment and the real world.

We describe the IT environment and system required to run the artwork. We also describe the material environment which supports it: the hardware, components and peripheral devices and their networked system. A significant number of digital art installations also require other non-digital equipment and materials which have to be specified. Documentation such as installation plans have to be integrated into our model. The model allows for description of the scenography: space requirements and restrictions (dimensions and characteristics), lightning (light source, luminous atmosphere…), noise level, etc.

The description of the "performed artwork" also needs, in order for our model to be pertinent, to be completed by information on the system that allows the artwork's relationship to the spectator and participant, and/or the environment. The "system" entity is designed to record the precise role of each interface. The input interfaces characterize the interaction mode: the work can react to a presence, a precise or generic gesture, a position in space, a sound, and so on, and is not restricted to one spectator. The output interfaces describe the perceived reaction of the artwork: the moving image on a display, the sounds from speakers, etc.

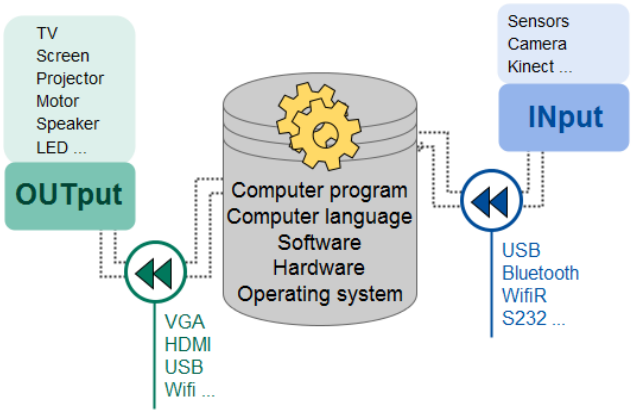

Figure 5: Input-output system diagram (interactive installation)

The interaction protocol has to be precisely described to keep replay and recreation options open. It can be graded on different levels of complexity that go from the simple action-reaction to an actual relationship between two equal objects with their own behavior: the human and the intelligent digital system [24].

Our model links the performed artwork to an entity that represents the performance event; an exhibition, a festival, a conference, and so on. There is a close relationship between the circumstances of the event and the characteristics of the performed artwork: its hardware changes, its backdrop and setting are modified. In addition, it evolves over time, whether is it partly or entirely emulated or recreated to conform to existing contemporary technology.

To complement the description of the conditions of the artwork's performance, we created an entity dedicated to the reception of the piece by the spectator: "experienced artwork". Since the artwork modifies itself anytime it meets an audience, it only exists because someone "read" it, "played" it or with it, "used" it: because someone took part in its full realization. The number of possible experiences is limitless; the implementation of this information entity could include a few representative testimonies.

Textual information on the exhibition materials may be all that documents the artwork if and when it enters the BnF's collections: the library is not well suited to preserving made-for-measure hardware, or even series of working contemporary computers.

By analyzing the existence of the artwork from its conception to its concrete temporal instances, we can grasp the essence of the piece and thus ensure that its description will be relatively stable through time and with different technologies. The information model is built on the description of the different stages of the artwork's life: conceptual, realized, performed and experienced. It is both rhizomic and exponential; it is built to allow for the description of the artwork's consecutive and/or parallel versions. One conceptual artwork can generate an unlimited number of realized artworks, which can generate an unlimited number of performed artworks, which will generate an unlimited number of experienced artworks.

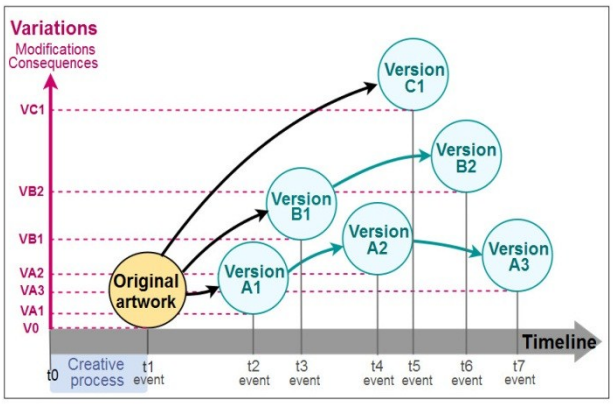

Figure 6: Artwork Lifecycle Diagram.

4 IMPLEMENTING THE INFORMATION MODEL

The conceptual model was our first step towards machine-actionable data models. We discuss here the adaptation of our conceptual model into a database and an ontology, and how the specificities of digital art influenced our choices.

4.1 Data models

Our first aim in building a database model and an ontology was to test the conceptual model through different types of implementation. On one hand, this process helped us consolidate our model and define our concepts; on the other hand, it raised new questions. The relational database and the ontology had to be developed in parallel, and they had to encapsulate all the different scenarios offered by a digital work of art that is based on the unpredictable. Ontologies and databases both have their limits, but the most difficult challenge came with the database: to prevent any recurrence or duplication, and moreover to avoid empty tables and keys in the database.

The digital artwork's lifecycle creates a problem of terminology, in this field of art history where computer technologies are preeminent. It is essential to distinguish the hardware and software used for the creation of the artwork, those used for the initial performance during an exhibition or a festival, those used for later versions of the work, and finally those we may use for purposes of emulation. In the ontology, we get the opportunity to render this distinction by applying the FRBR model. We can link technical metadata to the entities Work, Expression, Manifestation and Item as needed.

In the relational database, we choose to distinguish the hardware and the software as specific tables, and then to separate them from the tables that describe the implementation of hardware and software used by the artwork ("technical set up") or individual performances of artworks ("event"). This simplifies the

description of versions in the database: using junction tables, software and hardware can be ascribed to a specific setup and/or event. By detailing technical metadata at several points of the artwork's lifecycle — conception, performance and emulation —, we create more opportunities for future research in the fields of art history and computer sciences.

However this level of detail, and the tests we performed in populating the database, gave rise to a new issue: defining the core of the digital artwork. All the artworks in our corpora rely on interactivity, and this cannot be expressed only in terms of hardware and software, even if it is strongly related to both. The table "System" aims at filling this void, by describing the intention governing the conception of the program, the cognitive system within which the artist operated, the actions and reactions of the audience, etc. We believe future researchers will value a standardized description of the type of interactivity to search for artworks in the corpora, even as free text describing the experience of the work remains irreplaceable in certain cases.

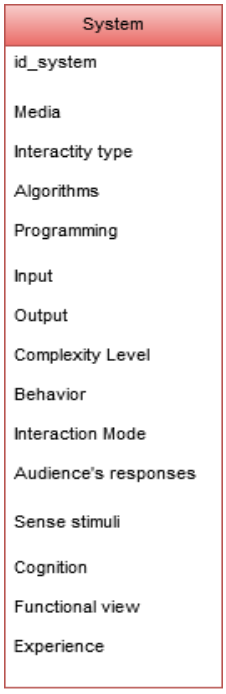

Figure 7: System table.

In our ontology, information relating to interactivity type and conceptual system is linked to the "Item" entity, as it is applicable to the preservation context of a specific instance of the artwork.

Implementing a database has led to retrospective changes in the conceptual model: we now distinguish between conception, as an intellectual activity, and production. Our initial "conception" entity was too vague, encapsulating the intellectual and the technical aspects of the created work. We had to make sure information items such as programs, algorithms, software, hardware etc. were not encapsulated in, and thus tied to, the table describing the genesis of the artwork, when we need to use the same information types to describe technical setups or emulation instances. Besides, there would have simply been too many elements in the one table to make it manageable.

One last example of the way our database and our ontology affected the logic of our information or conceptual model is how is allows us to see its flaws. We have difficulties modeling aesthetic currents and notions: the history of digital art is still being made, and at the moment, it is impossible for us to establish definite styles or movements and to fill fields of the sort in a database or an ontology. This underscores the need of agreement in the wider community on several key definitions in digital art, through a thesaurus for example.

4.2 What choices for our ontology?

We have additional objectives in working on an ontology derived from our conceptual model: making our information model interoperable, making our information machine-actionable, and evaluating the proportion of data that can be standardized. From this evaluation, we wish to assess how this ratio of free text to controlled data affects the quality of preservation.

Achieving interoperability of the data model depends in part on an accurate assessment of the importance given to the free-text description of interaction and behavior of the artwork. We want the ontology to be precise enough to describe the project's corpora while opening it up to the possibility of being used in other contexts. Interaction plays an important part in our description, as it is the common denominator to our three collections. But we include in our model the possibility of artworks without interactivity, with Pierre Morelli's table "Complexity Levels". The first level is indeed: "The object is unresponsive and without any activity. It is totally inert".

Manageability and interoperability also depend on choices relating to integrating existing ontologies. We use a fair number of preexisting classes and properties, some widely used and describing general concepts, some specific to information science and digital preservation. In the first category are SKOS and FOAF[17]. In the second are FRBR-oo [25], Consumer Electronics Ontology and the PREMIS ontology[18]. Used in major research projects such as Biblissima[19] [26] and Doremus[20], FRBR-oo responds to our versioning issue, using the Work and Expression classes. As for the PREMIS ontology, it is designed to address "metadata that a digital archive needs to know for preserving objects" [27], which is at the center of our project. We are particularly interested in "premis:Hardware", describing the IT equipment, "premis:Environment", as the environmental context

[17] https://www.w3.org/TR/skos-reference, http://www.foaf-project.org/
[18] http://new.cidoc-crm.org/frbroo/home-0, http://wiki.goodrelations-vocabulary.org/Vocabularies, http://id.loc.gov/ontologies/premis.html
[19] http://www.biblissima-condorcet.fr/
[20] http://www.doremus.org/

and its documentation, both essentials to make digital objects usable and in "premis:ObjectCharacteristics" to express information about the format, size, fixity or else the creating application of our different files composing the artworks.

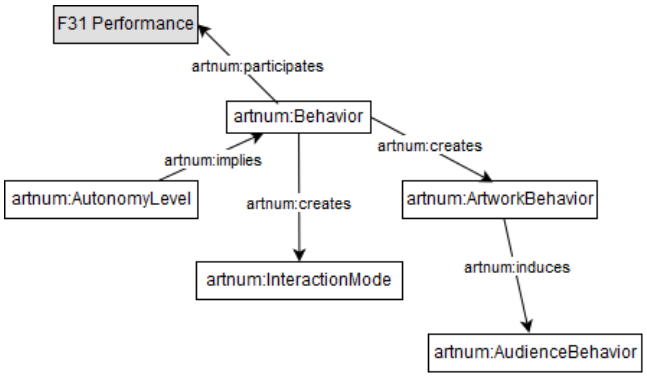

Figure 8: Proposed additional classes and properties.

We choose to limit the creation of classes and properties to those absolutely necessary to express the particularities of digital art with an interactive dimension. The terms "Behavior", "AutonomyLevel" or "InteractionMode" represent these two wishes. They express how an artwork behaves itself and two components of this behavior: its capability of autonomous action (or reaction) and the way it interacts with the audience. In other words, we want to transpose the process and functioning of interaction: how many people it can connect with, what triggers the interaction (breath, heat, movement) etc.

However, some of the data experts we are working with are reevaluating the priority given to class and property reuse. They contend that, except for RDF vocabs widely used (such as Dublin Core, SKOS, FOAF or even PREMIS), the data needs would be better served by dedicated classes and properties, rather than using other vocabularies in a "Frankenmonster ontology", as a matter of mastering our data. Indeed, it would be almost safer to manage our own classes and properties rather than using small ontologies that, even if they already exist in the linked data world, are not maintained on a regular basis. Furthermore, we are not aware of any institution that has used the FRBR-oo or CIDOC-CRM ontologies in its processes.

Implementing the conceptual model as a database and an ontology raises the question of the coexistence of controlled vocabularies and full text fields: what degree of standardization is efficient? This depends on several factors: at the art community level, are concepts sufficiently stable to create a precise thesaurus? At the preservation community level, are existing data models adapted to the digital art field? Is there a stable source for normalized information on formats, software and hardware? At the institutional level, are there enough staff to implement the description in a variety of controlled fields? What fields will be useful to index and propose to the users as a request entry point?

The technical and artistic fields in our model are designed with the aim of populating them with controlled vocabularies, but some will have to be full text. Indeed, it does not seem feasible or advisable to express the creation process, or the user's experience, as a set vocabulary, even if concepts have begun to emerge. We will also have to add documentation (images, videos, testimonies, etc.) to the information package in order to satisfy the digital art historians. A better understanding of the proportion of full text to documents to controlled data in our descriptions will emerge in the third year of the project, as we scale up our tests on the corpora. The right proportion will be the one which satisfies the artist's, the researcher's and the archivist's needs.

4.3 Our to-do list

We are aware that our three-year research project will not lead us to a complete data model implementation at the end of 2017. Sorting the partners' corpora, gathering data and preparing our test corpus has been more time consuming than anticipated. However, we do expect to revise the information model and the ontology within this year, using the input we get from additional artists' testimonies as well as recommendations from fellow art historians, librarians, and experts from the digital preservation community.

We will try to experiment a mapping from the conceptual model to the BnF's catalog format, INTERMARC. Description formats and tools are currently being revised to facilitate the adoption of practices compatible with FRBR models, which is an opportunity for our project, as our model has a work / expression / manifestation / item structure. We will also scale up the tests of the data model on the project's corpora, using the test set which has been defined in the second year of our research.

The closing conference for "*Art numérique et postérité*" in December will give us a chance to share our results with the art and information science communities. We intend to show artworks from the corpora through virtualization and emulation, and discuss the delta between the extent of our standardized metadata and the aspects of the work that can only be rendered by free text descriptions.

5 CONCLUSIONS

We are aware that the mix of artists, art historians and librarians in the project is both an asset and a liability. Therefore, we are tracking another type of result, which we hope to foster: the partners have experienced vast acculturation to one another's preoccupations. While the members of INRéV had no knowledge of international best practices in digital preservation at the beginning of the project, they are now discussed in the curriculum of the ATI Department. The importance of gathering information on the artwork as it is being made is presented to the masters students who are studying the technologies of digital art creation.

Conversely, the library staff had little knowledge of the significant properties of interactive digital art, which made defining the information packages and evaluating the quality of emulation and virtualization difficult. A first training session with the library assistants in charge of facilitating access to the documents in the audiovisual research reading room has allowed us to consolidate the procedure for giving access to the emulated

artists' CD-ROMs collection. Others will follow as more corpora are ingested in the BnF's archive.

In the future, we will have to monitor our information and data models to ensure that they remain applicable as digital art evolves.

We have built the models to encompass a large array of digital art creations, and have tried to articulate interaction systems, hardware and software in a way that allows for fast-evolving technologies. Yet we are keenly aware that preservation is a continuous process and that we have to balance the requirements for a stable model with the multifaceted and ephemeral nature of digital art.

ACKNOWLEDGMENTS

We would like to thank the team who made this project possible: Élodie Bertrand, Marie Saladin, Jean-Philippe Humblot, Bertrand Caron, Sébastien Peyrard, Pierre Choffé, Françoise Leresche at the BnF; Marie-Hélène Tramus, Vincent Meyrueis, Edmond Couchot and Michel Bret at Paris 8 University.

Our interns brought new skills to our research: Catherine Champenois, Juliette Fattal, Tigran Ghaplanyan, Catherine Helmer, Yu-Ci Huang, Gaétan Lemaitre, Victor Martin, Julie Milhiet, Nola N'Diaye, Laetitia Perez.

We received invaluable help from the Arts-H2H Labex team and our external collaborators: Pauline Cellard, Peter Stirling, Mehdi Bourgeois, Jean-Luc Soret.

Finally we would like to thank the artists and researchers, who are the very heart of our project.

REFERENCES

[1] Dragan Espenschied, Oleg Stobbe, Thomas Liebetraut and Klaus Rechert. 2016. Exhibiting Digital Art via Emulation - Boot-to-Emulator with the EMiL Kiosk System. In *Proceedings of the 13th International Conference on Digital Preservation* (iPRES 2016). Swiss National Library, Bern, 37-40.

[2] Klaus Rechert, Patricia Falcao and Tom Ensom. 2016. Towards a Risk Model for Emulation-based Preservation Strategies: A Case Study from the Software-based Art Domain. In *Proceedings of the 13th International Conference on Digital Preservation* (iPRES 2016). Swiss National Library, Bern, 139-148.

[3] Frank Padberg, Philipp Tögel, Daniel Irrgang and Martin Häberle. 2016. A Case Study on Emulation-based Preservation in the Museum: Flusser Hypertext. In *Proceedings of the 13th International Conference on Digital Preservation* (iPRES 2016). Swiss National Library, Bern, 149-158.

[4] Catherine Jones, Brian Matthews and Ian Gent. 2015. Software Reuse, Repurposing and Reproducibility. In *Proceedings of the 12th International Conference on Digital Preservation* (iPRES 2015). School of Information and Library Science, University of North Carolina at Chapel Hill. http://phaidra.univie.ac.at/o:429590

[5] Dragan Espenschied, Isgandar Valizada, Oleg Stobbe, Thomas Liebetraut and Klaus Rechert. 2015. (Re-)publication of Preserved, Interactive Content – Theresa Duncan CD-ROMs: Visionary Videogames for Girls. In *Proceedings of the 12th International Conference on Digital Preservation* (iPRES 2015). School of Information and Library Science, University of North Carolina at Chapel Hill. http://phaidra.univie.ac.at/o:429586

[6] Klaus Rechert, Thomas Liebetraut, Oleg Stobbe, Isgandar Valizada and Tobias Steinke. 2015. Characterization of CD-ROMs for Emulation-Based Access. In *Proceedings of the 12th International Conference on Digital Preservation* (iPRES 2015). School of Information and Library Science, University of North Carolina at Chapel Hill. http://phaidra.univie.ac.at/o:429556

[7] Jean-Pierre Balpe. 2000. *Contexte de l'art numérique*. Hermes Science, Paris.

[8] Franck Popper. 1985. *Art, Action et Participation, l'artiste et la créativité aujourd'hui* (2nd. ed.). Klincksieck, Paris.

[9] Jean-Louis Boissier. 2009. *La relation comme forme : l'interactivité en art* (new augmented ed.). Musée d'art moderne et contemporain, Genève ; Les presses du réel, Dijon.

[10] Edmond Couchot and Norbert Hillaire. 2003. *L'art numérique : comment la technologie vient au monde de l'art*. Flammarion, Paris.

[11] Florent Aziosmanoff. 2009. *Living art, l'art numérique*. CNRS, Paris.

[12] Florent Aziosmanoff. 2015. *Living art, fondations : au Coeur de la nouvelle économie*. CNRS, Paris.

[13] Francisco Varela. 1996. *Invitation aux sciences cognitives*, translated by Pierre Lavoie. Seuil, Paris.

[14] Chu-Yin Chen. 2017. Un monde pleinement énacté. In *Action, énaction : l'émergence de l'œuvre d'art*, Xavier Lambert (Ed.). L'Harmattan, Paris, 53-76.

[15] Alain Berthoz, Michel Bret and Marie-Hélène Tramus. 2015. Création n° 2 : La funambule virtuelle. In *Hybrid*, 02 (2015). http://www.hybrid.univ-paris8.fr/lodel/index.php?id=533

[16] Hervé Huitric and Monique Nahas. 1984. Realistic effects with Rodin. In *Frontiers in Computer Graphics* (Computer Graphics Tokyo'84), Tosiyasu L. Kunii (Ed.). Springer Verlag, Tokyo, Berlin, Heidelberg, NewYork, 159-168.

[17] Chu-Yin Chen. 2005. Digital Art with Artificial Life: When the Artwork Becomes Autonomous. In *VRIC, Laval virtual 2005 proceedings: 7th Virtual reality international conference*. Laval Virtual, Laval, 217-222.

[18] Chu-Yin Chen and Jean-Claude Hoyami. 2007. Autonomous Systems for Interactive Digital Art. In *X Generative Art International conference* (GA2007). Politecnico di Milano University, Milan, 63-72.

[19] Emmanuel Guez and ESAA (École Supérieure d'Art d'Avignon, Eds.). 2014. Archéologie des média. *mcd* (magazine des cultures digitales) 75 (Sept.-Nov. 2014).

[20] Bertrand Caron, Thomas Ledoux, Stéphane Reecht and Jean-Philippe Tramoni. 2015. Experiment, Document and Decide: a collaborative approach to Preservation planning at the BnF. In *Proceedings of the 12th International Conference on Digital Preservation* (iPRES 2015). School of Information and Library Science, University of North Carolina at Chapel Hill. http://phaidra.univie.ac.at/o:429538

[21] John Ippolito. 2003. Accommodating the Unpredictable: The Variable Media Questionnaire. In *Permanence Through Change: The Variable Media Approach*, Alain Depocas, Jon Ippolito, and Caitlin Jones (Eds.). Guggenheim Museum Publications, New York, Fondation Daniel Langlois pour l'art, la science et la technologie, Montréal, 47-53.

[22] Bernhard Serexhe (Ed.). 2013. Preservation of Digital Art: Theory and Practice. The digital art conservation project. ZKM Centre for Art and Media, Karlsruhe.

[23] Pierre Vermersch. 2010. *L'entretien d'explicitation* (6th ed.). ESF, Issy-les-Moulineaux.

[24] Pierre Morelli. 2000. *Multimédia et création, contribution des artistes au développement d'une écriture multimédia*. PhD thesis. Université de Metz, France.

[25] Chryssoula Bekiari, Martin Doerr, Patrick Le Bœuf and Pat Riva (Eds.). 2016. *Definition of FRBRoo: A Conceptual Model for Bibliographic Information in Object-Oriented Formalism*. IFLA, Den Haag. https://www.ifla.org/files/assets/cataloguing/FRBRoo/frbroo_v_2.4.pdf

[26] Matthieu Bonicel. 2013. Biblissima : observatoire du patrimoine écrit du Moyen Âge et de la Renaissance. *Bulletin des bibliothèques de France*, 2013, 5, 23-26. http://bbf.enssib.fr/consulter/bbf-2013-05-0023-005

[27] Angela Di Iorio and Bertrand Caron. 2016. PREMIS 3.0 Ontology: Improving Semantic Interoperability of Preservation Metadata. In *Proceedings of the 13th International Conference on Digital Preservation* (iPRES 2016). Swiss National Library, Bern, 32-36.