



**HAL**  
open science

## **Entity / Identity**

Philippe Bootz, Samuel Szoniecky, Abderrahim Bargaoui

► **To cite this version:**

| Philippe Bootz, Samuel Szoniecky, Abderrahim Bargaoui. Entity / Identity. 2009. sic\_00603588

**HAL Id: sic\_00603588**

**[https://archivesic.ccsd.cnrs.fr/sic\\_00603588](https://archivesic.ccsd.cnrs.fr/sic_00603588)**

Preprint submitted on 26 Jun 2011

**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.

# Entity/Identity

*A tool designed to index documents about digital poetry*

Philippe Bootz, Samuel Szoniecky and Abderrahim Bargaoui, Laboratoire Paragraphe  
Translated by Nicolas Cognard

*Symposium e-poetry, Barcelone, 2009*

## 1 Preservation: an answer to obsolescence.

### 1.1 Digital preservation proposals.

One may be surprised by the fact that we introduce a reflection on the identity of digital poetic works while planning to design a tool to index documents. Yet, the idea of such a tool has followed from the reflection on questions of preservation. These questions have led us to ponder on the nature of these works. Thus, the proposed tool will bare the mark of this reflection.

Over the last years, the relation between preservation and indexing has been at the centre of reflections on the preservation of the digital heritage. In that matter, the most important project has been developed by the Guggenheim museum in collaboration with the Langlois foundation. It aims at preserving museum installations that include electronic or digital apparatuses. In this project, as in all the other projects involved in digital preservation, the question of preservation is often treated with regards to the technological obsolescence that threatens every digital media. Then, a technological migration (simulation, recreation, reduplication) is strongly recommended in order to extend these works' lifetime. As a matter of fact, this approach is in keeping with that of variable media, developed by Jon Ippolito (2003,) and which serves as a theoretical basis for general projects of preservation in the technological arts. This approach states that a work is a "unique cultural artifact" (Rinehart, 2003,) in short an "invariant" that would be embodied in a media that is itself variable. The preservation of a work consists in cancelling the variability of the implied media. This cultural artifact is defined within the project of variable media as the state of the artist's intention, merging with the first version exposed by the artist. The work must then be documented by the artist as part of this strategy. The DOCAM project (<http://www.docam.ca/>) falls within this scope by putting forward a wider reflection on the relation between preservation and documentation. It develops case studies, a thesaurus and a good manners guide, among other things.

Other researches lay down some general technical principles on the preservation of the digital heritage by following this same logic, i.e. by assuming the presupposition of an existing "original". Gladney (2004, 2006) acknowledges the difficulty one may have in defining such an original by pointing out the fundamental role of human subjectivity, according to Wittgenstein's *Philosophical investigations* (1953.) Gladney states that: «*Nobody creates an artifact in an indivisible act. What is a version or an original is somebody's subjective choice, or an objective choice guided by subjective social rules*»

(Gladney, 2004, p. 5.) Once this step is performed and in keeping with this approach, the question of the authentication of the copies may be solved by the TDO method (Trustworthy Digital Objects,) which consists in saving metadata and objects at once. This metadata embeds some information on the origin of the copies, on the nature of the reading software and semantic pieces of information (ontological relations) as well as other pieces of information connected with the saved object. The saved object accounts for the fact that “*no document is comprehensible except in the context of other documents.*” (Gladney, 2004, p. 9.) Gladney also insists on the relation between preservation and communication, according to the classical point of view that defines communication as a transmission. For him, preservation consists in making sure that the user gets an exact copy of the original as well as in making him understand the author’s idea. According to Gladney, this can be done thanks to a paratextual documentation based on multiple channels. This problem of understanding questions the interface of the tool that one uses to consult a document<sup>1</sup>. Gladney also brings up the idea that the concept of version is constituent of the digital work. In order to attest the authenticity of the documents, the method resorts to an encoding of the data, relying on the recursive use of certificates.

Thus, this system is more complete than that of variable media, even if the fundamental concepts concerning the nature of these works remain the same.

Lee et al. (2002) puts forward the data re-mediatization (emulation, migration, encapsulation) as a solution to digital preservation. The idea is to use these methods before obsolescence gets the most of the former media, so as to ensure the faithfulness of the transformation. According to him, it is a matter of preserving the digital content and its functionality, as well as the possibility to consult it.

In Europe, the CASPAR project (Giaretta, 2006) [<http://www.casparpreserves.eu/>] brings in the association of preservation and knowledge of the preserved objects as an essential condition for their re-exploitation.

The Digital Preservation Coalition [<http://www.dpconline.org/graphics/index.html>] counts several other international initiatives (Semple & Clifton, 2007). They all give elements of answer to the problem of preservation as part of the use of the semantic Web and of a normalized open archive information system (OAIS).

## **1.2 The question of the state of reference.**

We may sum up these different approaches in the following paradigm: a digital product includes an original version which should be preserved in the actual media, despite the constant change of this latter. This product is a digital object which preservation cannot be performed without a thorough documentation on its original state. This documentation undergoes a description of the author’s state of intention as well as a semantic documentation.

From now on, we have to question the fundamental assumption of these approaches, i.e. the existence of a state of reference which we should be able either to reconstruct or else to preserve. We can imagine that this is perfectly feasible in the case of a museum that deals

---

<sup>1</sup> In the case of a re-mediatization, one can easily notice that the interface of a reading tool is not neutral as far as communication is concerned. Thus, during a conference that I organized on the occasion of “par Monts et par Mots” in the villa Mont-Noir in Saint-Jean Cappel in 2000, very few people were able to distinguish Queneau’s *Hundred thousand billion poems* from Antoine Denize’s CD-ROM *Typewriters* (Gallimard, 1999) and his contribution to Queneau’s work.

with objects: there is indeed an original version of the object, bought by the museum and which maintenance may be complicated by the obsolescence of the technological apparatuses that are used, as well as by the fact that it is impossible to replace them in an identical way. We can also imagine that it is as feasible in the case of digitized analog documents, but how about digital literature?

The ELO approach, related to digital literary works, follows the aforementioned general process. The project advises authors to archive their works and to take into account the question of preservation as soon as they conceive them (Tabbi, 2004.) The essential contribution of this project undoubtedly lies in the community aspect that it displays: preservation is no longer a matter of specialists cut off from the context of creation, nor is this simply the business of professionals. In fact the whole of the community, i.e. all those who maintain a relationship with digital literature in a way or another, must take charge of preservation. Even if ELO explicitly tackles the issue of the definition of the state of reference that is to be preserved, digital literary works are too often treated like objects that would exist on their own and be embodied in a media threatened by multiple causes of obsolescence, identified in the project. Nevertheless, ELO introduces an original solution: the reference is not defined by a single state, but by the totality of the relations that make a document legible: « *From the point of view of long-term digital preservation, however, the entity of interest is not necessarily any discrete object but the working relationship among objects (each of which may mutate) that assures readability. This means that the intact "original work" in its initial instantiation [...] loses its iconic status and becomes just one of many possible manifestations of a preserved work* » (Liu et al., 2005.) Yet, this functional conception seems to fit with the recognition of the technological conditions of execution of the works, for the proposed solutions remain the classical ones of emulation, simulation and documentation. The efficiency of these approaches depends on a given initial state of reference, an “original” that can be described and documented in neutral terms in a new adapted format (called “x-lit” in the project), i.e. a XML-based representation that would be legible both by men and machines. This standard will then have to be able to describe the media, as well as the computing and interactive aspects of the works.

Thus, none of these methods will reach its goal unless the state that is to be preserved is defined. In all these projects, the problem remains unsolved because the preservation issue is tackled from the sole point of view of technological obsolescence. Yet, the definition of the state of reference, if any, can only be performed by taking into account the work’s lifetime that precedes its obsolescence. The aforementioned steps accounting for the situations of production and reception, the search for a state of reference must partake in communication issues revolving around the work.

## **2 From obsolescence to lability: a shift in point of view**

### **2.1 The point of view of unstable media.**

The V2\_ center is an institute based in Rotterdam. It studies unstable media and defines them as opposed to variable media, in the following terms:

« *We make use of the unstable media, that is, all media which make use of electronic waves and frequencies, such as engines, sound, light, video, computers, and so on. Instability is inherent to these media.* » (manifesto for the Unstable Media. 1987, [http://archive.v2.nl/v2\\_archive/projects/capturing/1\\_1\\_inventory.pdf](http://archive.v2.nl/v2_archive/projects/capturing/1_1_inventory.pdf))

[« nous nous servons des médias instables, c'est-à-dire de tous les médias qui utilisent les ondes et fréquences électroniques, tels que les moteurs, le son, la lumière, la vidéo, les ordinateurs, et ainsi de suite. L'instabilité est inhérente à ces médias. ] (quoted and translated by Laforet, 2009, p. 134 note 216)

V2\_ opposes the concept of capture to that of conservation (<http://capturing.projects.v2.nl/>). Capture concerns any instantaneous state of a work's lifetime, without going as far as to presume that this state constitutes an absolute state of reference. The V2\_ project performs its captures in archives and gives pieces of advice to build these up. The project opens the possibility to document various states in the work's life, as well as its environment. In this perspective, it also builds up a thesaurus to document the work.

The conception of unstable media is very close to the one that we present hereafter. The main differences lie in presuppositions, in the area that is to be documented and in the nature of the documentation. Just like the project of variable media, this project relies on the materiality of the work and especially on the concept of media. Thus, even if the work is here totally assumed as a process that interacts with its environment, this project no more comes from the artistic nature of the work than the variable media project does. In fact, it comes from its aesthetic *raison d'être*.

Hence the following questions: what is the aesthetic impact of that variability or instability? What is the aesthetic impact of the media? Is the work a project, or in other words, can we describe one or several states of reference within its process of evolution?

## **2.2 Procedural transformation and lability of the work.**

Communication via digital literary works has already been the subject of analyses that have given way to the formalization of a procedural model (Bootz, 2004). Yet, what fundamentally characterizes communication via digital works in this model is the existence of an "autonomous process", i.e. a "loose" and unpredictable relationship between the author's intentionality that is expressed in the program and the state that is produced as the reader executes this program. Thus, there is a particular transformation, called "procedural transformation" (Bootz, 2003, p. 81) that turns the execution process, observed on the author's machine, into a different execution process observed by the reader. This gives way to an aesthetic divergence between the result experienced by the author and the one that is read by a reader. This divergence is no more due to the existence of a "generativity" than it is caused by interactivity. It is rather the consequence of a fundamental technological propriety related to the creation context. To put it simply, every author is a machine-user and as far as users are concerned, computers do not behave like a Turing machine, for the program of a work contains a great amount of things that are left unsaid. The nature of these things left unsaid is either one of the following two:

- The user cannot master the complete set of instructions displayed during the execution process. He only masters the instructions that he can establish in his author program. This means that he is no more in control of the instructions given by the operating system, than of those given by the software layers (protocols, players...). Yet, the behavior of these layers may produce various results, because of the great variety of technological

contexts. Hence, the author cannot provide readers with identical results, as shown in the case of portable devices, for example.

- The program cannot define the whole of the parameters that are used during the execution. Apart from the parameters handled by the reader, such as the volume and brightness controls, there are lots of parameters that are controlled by the apparatus alone (the running speed of a code line, the data reading speed...) or that depend on the running environment (disk fragmentation, number of active windows, number of tasks processed by the system...)

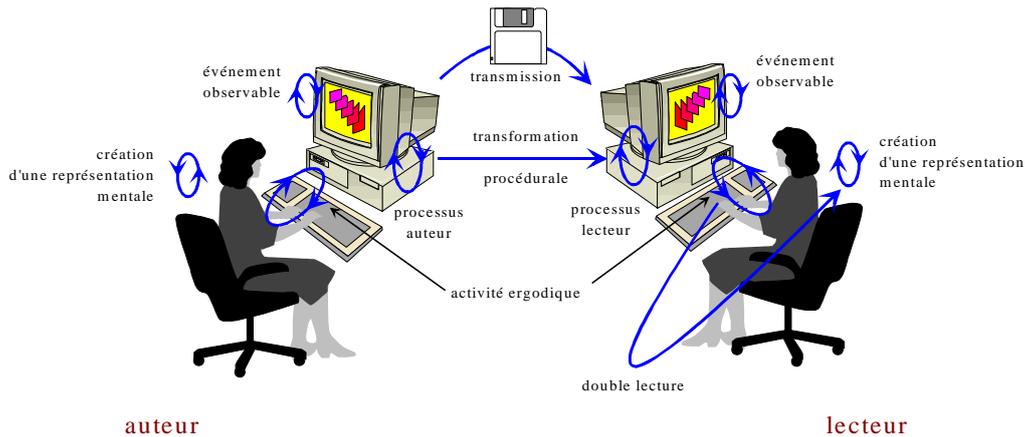


figure 1 : The procedural apparatus and its processes.

Thus, procedural transformation plays a decisive role in the perception of a work and ascribes the implied media with a variable aspect, way before the question of obsolescence is raised. The reader, as any observer in fact, will perceive the work as labile. Therefore, the problem is to establish whether this lability is to be considered as characteristic of the work “on display” or whether it is inherent to the work. Depending on the answer to this question, we may or not define a state of reference.

**a) The state of the author’s intention cannot be used as the state of reference.**

In order to tackle that issue, we have to consider several hypotheses. The first one partakes in the logic of the “variable media” approach and asserts that the state of reference we have to take into account corresponds to one of the author’s states of intention, if any, the initial state. Unfortunately, this state seldom corresponds to the initial state observed by a reader. As a matter of fact, the result observed by the author depends on the whole technological context created by his machine, a state that can neither be thoroughly described nor reproduced identically by the reader’s machine. One might suppose that there is but little difference, but twenty years of experience in publishing the *alire* review testify to the contrary. Thus, some author believed that his work had been changed by the editor as he noticed a slight lag between sound and image on the CD-ROM provided with the review. This lag was due to the fact that he had only tested his work on his hard-drive. The shift from his hard-drive to the CD-ROM had slightly changed the running time of the different reading flows, yet the difference was already significant to him. However, this author never demanded that his work be withdrawn from the review, nor demanded he that the difference between his state of intention and the realized state be mentioned. Very often, authors do not mind broadcasting works that behave differently depending on time and machines. Moreover, the

state of intention that an author designed on his machine may never be published at all. Thus, certain works published in the *alire* review were created on machines that already existed before the creation of the review. When they are run on apparatuses that are different from the original one, such works never produce identical results. The author is aware of that fact and agrees to broadcast his work all the same.

Therefore, we shall not evoke the author's state of intention - nor shall we evoke the state he worked on - as the state of reference for works that are destined to be read in private, while being electronically broadcast in multiple contexts. We call these works "private works". They constitute the vast majority of digital literary works.

We shall not evoke an original state of the work either, a state that would, for instance, correspond to an original broadcast. These works are often broadcast simultaneously and in so many different ways that it is practically impossible to point out the original one, or a reference of perception for that matter. Indeed, which reader could be considered as the point of reference? How could this reference be documented, for inasmuch as it would exist, it would correspond to an event localized in space and time. In order to be reproduced, this event would have to be tangible, which is not the case. There currently is no editorial device that can legitimate an original "display" frame. Thus, we shall no more evoke a state of reception, than a state of intention, in order to define the state of reference of a work we should know how to restore. The conceptions at work in the present methods of preservation infer a pernicious effect, that of "ossifying" the work, of denying its profoundly procedural and contingent nature, so as to try to refocus it on the safety frame of the object. We keep on thinking about the digital work as an object, but that object does not exist.

#### **b) Lability as inherent to the digital work.**

In our opinion, that "variable" nature does not determine the media but it does determine the digital work. It is a characteristic that is inherent to digital works, even if it manifests itself through the technological behavior of the media. Just as well, one could define the lability of these works as a semiotic variability rather than a variability of the media.

Actually, we can take that lability into account and explain it from a semiotic point of view, without referring to the technology of the media. It is not an easy thing to study, because we need to compare observations taken from different technological contexts. This is an unusual protocol. Yet, the study of a prototypical case, i.e. the evolution of the perception of Jean-Marie Dutey's work *mange-texte* (1989) may be helpful. This work is a programmed animated text relying on graphic mutations. At first, the running of the work lasted 20 minutes and the legibility of the texts used to get the most of the graphic treatments that ruled its evolution. This work was then considered as a literary work. By 1994, due to the increasing speed of processors, the running time had been shortened to only a few minutes, which prevented anyone from reading the texts. The status of the work had changed. It was now considered as a graphic work. In 1994, the work was reprogrammed to its current form, in which the program adapts to the machine to some extent, so as to preserve a certain legibility. This solution does not completely correspond with the author's initial intention, because some characteristics have not been preserved. It does not correspond to the preservation of a perceptive state either. It even relies on the fact that it is impossible to reach such preservation. Yet, this programming was performed with the author's full consent.

The change of status between the versions of 1989 and 1994 is directly linked to the intersemiotic behaviour of the work. Let us now examine its semiotic functioning in detail.

The concept of intersemiotic behavior is easier to explain by referring to Klinkenberg's theory of the sign (1996.) He introduces the concept of stimulus in the definition of the sign, besides the three traditional dimensions of the signifier, the signified and the referent. The stimulus is defined as the part of the world that the semiotic decision recognizes as a sign: the ink of the words written in a book, for instance. That notion is particularly helpful as soon as we are confronted with a multi-code system or an intersemiotic system. In such systems, the same physical part can be a stimulus within several semiotic systems. This is what takes place in *mange-texte*.

This work is made of 3 successive stanzas that follow each other according to an intersemiotic process, i.e. the jumping from a semiotic system to another. Each stanza is multi-coded. It obeys a double system, in other words a system that is both linguistic and typographic. Each letter is composed of a void segment and of 4 elementary segments (fig.2,) each of which can be rotated thrice through an angle of 90° so as to lean on each side of a square, materialized in the work for that matter. The set of variants thus obtained constitutes a group of distinctive units of letters, which corresponds to an unusual linguistic level within our language. The pixels that constitute these segments are thus the stimulus of both a linguistic system and a typographic one. The graphic components of this system are completed by two enveloping forms: the square and the circle in which the typographic elements are drawn. The stimulus of these borderlines belongs to the sole graphic system and does no longer partake in the linguistic system. Each letter is made up of 4 elementary enveloping forms. The multi-code system is then made up of two systems of which stimuli partially overlap.



Figure 2 : non-void distinctive units of letters in *mange-texte*

These two systems evolve in different ways throughout the running of the work. The stanzas (linguistic system) follow each other in the form of still screenshots, interrupted by moments when the linguistic system partially disappears. The graphic system, on the other hand, evolves according to the following simple three-step algorithm, recurring from one stanza to another:

- Firstly, the typographic system is subject to a mutation from square to circle: every square is turned into a circle. This is done by cancelling the corner pixel of every square.
- Secondly, every elementary segment is replaced by the segment of the letter that takes its place in the next stanza. The replacement order is random. Every stanza takes on a different color.
- Finally, when the substitution is complete, the transformation order is reversed: every enveloping form gets back from circle to square.

In the first and third steps, the linguistic system is not destroyed, for the letters remain. Simply, they are no longer legible, for the hole in the center of each letter suffices to destroy their legibility with this font. On the contrary, the linguistic system is randomly destroyed in

the second step of the graphic transformation. This happens because the combination of distinct units of different letters no longer forms letters. What takes place is truly an intersemiotic functionality, i.e. the flow of an alternately multi-code system and of a purely graphic one. The mere graphic system remains throughout the whole transformation. This latter does ensure the intersemiotic continuity.

How has the perception of this system evolved over time?

As long as the work was being created, both the typographic and the linguistic systems of the multi-code system were visible as each stanza was displayed and before the stanzas underwent the aforementioned transformation process. The screening time of each stanza lasted several minutes, so that one directed one's attention mainly on deciphering of the letters, the font that was used being hardly legible for unaccustomed readers. During that read-out time, the *texte-à-voir*<sup>1</sup> (literally, text-to-watch) was essentially made up of segments of letters and was fundamentally considered as a text. The transformation phases were slow and the general process that we described above was hardly visible, because of this slowness. The result one could perceive during these phases was truly a graphic system, but not a very interesting one. The typographic aspect of the segments could not be perceived in step 2 and step 3 and the replacement process was hardly perceptible, for it was too slow. As a whole, the work was perceived as a succession of separate stanzas operated by a graphic transition. Its status was the same as that of a literary work. The intersemiotic flow was perceived as the succession of different semiotic systems. On the contrary, in the 1994 version, the read-out time had become much too short to be legible. It was even difficult to realize that texts were being displayed. As for the logic of the transformation process, it had become very visible. The replacement of the segments was operated in a lapse of time that allowed perception and understanding. The intersemiotic system remained the same, but the way one perceived and understood it had changed. Thus, the nature of the work itself had changed. That being so, the work no longer seemed to be a literary one, instead it appeared to be a mere visual product.

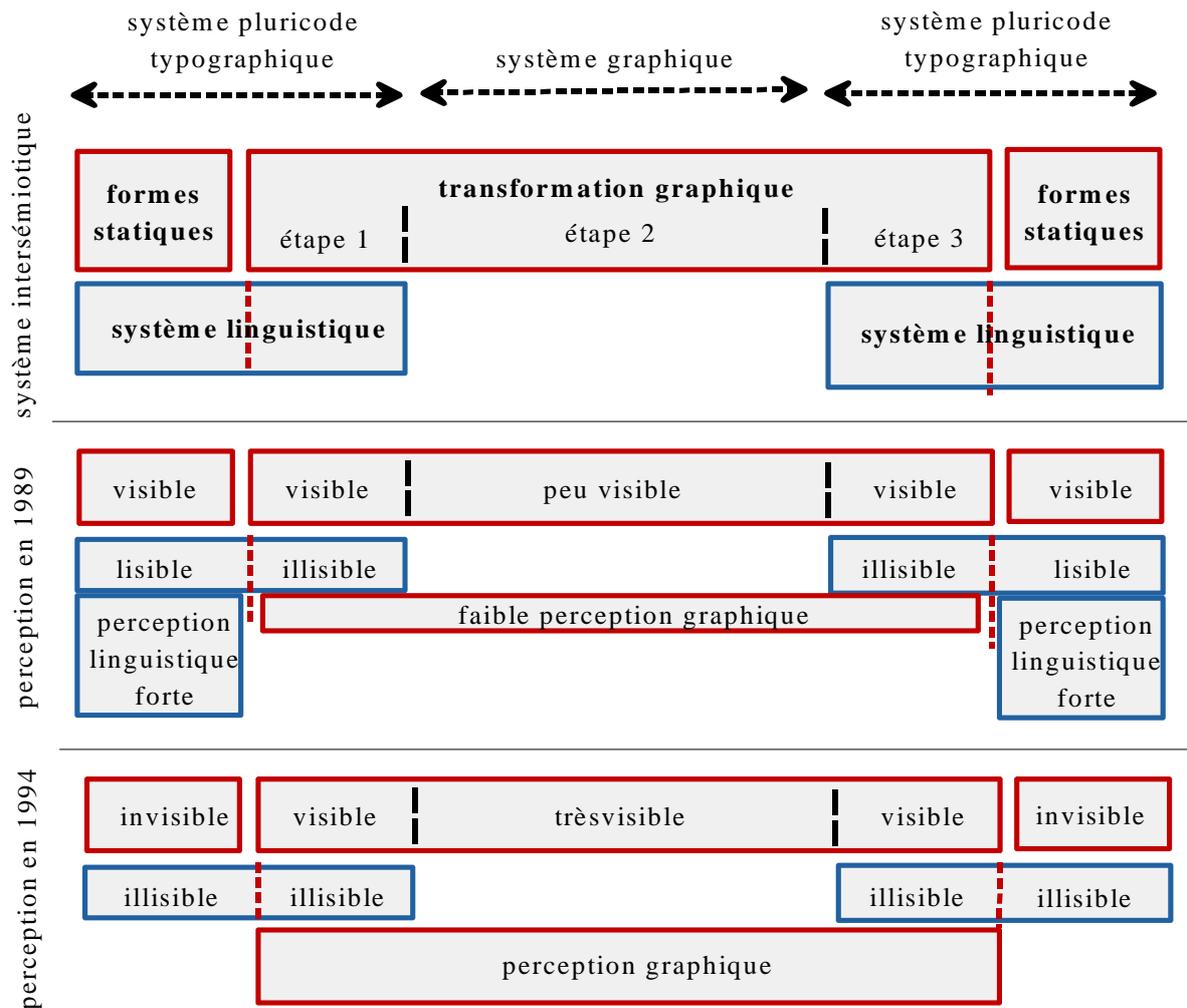


Figure 3 : the intersemiotic functioning of Dutey's *mange-texte* and the evolution of its perception

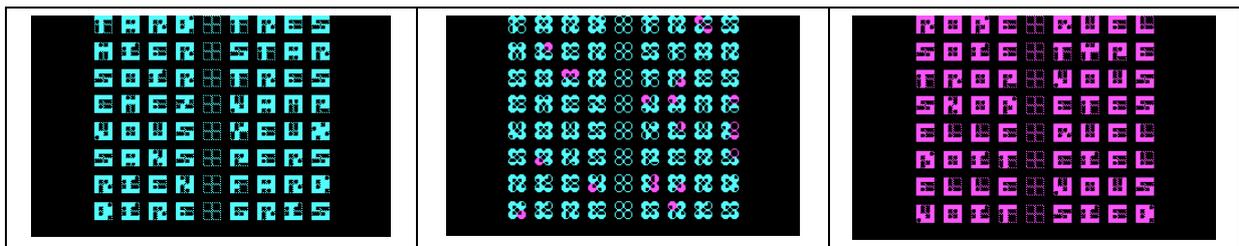


Figure 4 : screenshots of the intersemiotic functioning of Dutey's *mange-texte*

### 3 Preservation conceived as the actualization of a power to act.

#### 3.1 The work as a power to act.

As soon as a private-reading digital work appears labile, one cannot but notice that its ability to mutate constitutes one of its fundamental characteristics, which makes it different from all other kinds of media works. Therefore, it is impossible to define the original version of a work, although any work includes a perennial object that can definitely be archived and described, i.e. its program. Nevertheless, it also contains noticeable states - which reading is variable - and a generating function that one may mistake for the execution process at first

sight. As for the execution process, it implements this labile characteristic. From then on, one cannot freeze the execution process to make it reproduce the same specific state, not even one of those enumerated in the PAD project. This goes against the idea of preservation, because it destroys that lability. Therefore, preservation should not be regarded as a problem of reconstruction of a state. Of course, the problem of preservation remains, for obsolescence constitutes a borderline case of lability. Obsolescence shall be defined in semiotic terms, without reference to any technological evolution. That definition shall be the only one in use. A work is obsolete as soon as its visible components no longer undergo any semiotic process. Defined as such, obsolescence characterizes the “semiotic death” of the work and does not necessarily superimpose on the obsolescence of the technological system.

Rather than defining the work as an object or a state, we propose to characterize the work, above all, as an aesthetic *power to act* within a socio-cultural context. Such an approach is compatible with the point of view of Spinozan ontology.

### **3.2 The power to act in Spinozan ontology.**

The Deleuzian reading of Spinozan ontology (Deleuze, 2001) proposes a valid ontological model for every kind of entities that we call individuals. The individual may of course be a human being, but it may also be an inanimate object such as a document or a work. Deleuze defines the individual thanks to three constitutive dimensions: its infinite and extensive physical and material parts, its singular essence (power to act) and the relationship that links these two first dimensions in the “here and now”. This is in no way a Christian approach. The essence of an individual is not an absolute that would be inherent to its extensive parts. It is rather another point of view on the being. The extensive parts constitute the point of view of the “width”, and the essence that of the “thought”. An individual may mutate by adding external elements to its extensive parts, always according to a particular relationship with its essence. Spinoza suggests that we conceive existence as an experimentation on eternity, or in other words, as an actualization of the potential relationships between extensive parts and essence, in the “here and now”.

The projection of our point of view in a Spinozan ontology amounts to considering works as individuals. As a result, their power to act cannot be assimilated by other individuals as if it were “something that one could possess”. To put it differently, the essence of a work shall never become the extensive part of another individual. It is only knowable. That knowledge may comprise several levels, from an inadequate and purely emotional knowledge to an intuitive knowledge of the essence (such as the author may experience,) through the knowledge of relationships (science and analysis fall within this category.) The conception of preservation, as dealt with in the projects described in the first part of this article, is incompatible with such an ontological conception. These projects tend to establish immortality by freezing relationships. They re-establish extensive parts *ad vitam aeternam*, according to a particular point of view (the author’s point of view in the case of the variable media project,) and to the detriment of all other possibilities.

Yet, in Spinoza’s opinion, every individual tends to preserve his power to act. That power is ceaselessly actualized in the relationships that the extensive parts maintain with the essence. It is the relationship that actualizes the essence. From our perspective, one needs to preserve the possibility of such relationships. Extension may create these relationships; therefore, preserving a work does not amount to freezing (or immortalizing) its extensive parts. It is rather a matter of making the never-ending creation of new extensive parts possible. All of them maintain singular relationships with the essence of the work. Preserving

is not avoiding destruction. Instead, it amounts to assuming this destruction, while making a new *creativity* of the work possible (instead of a new *creation* of the work.) The perception we may have of the present extensive parts of the work only constitutes some specific relationship that we keep with it. We have to accept it, as well as we have to acknowledge the fact that this relationship may disappear. In this regard, preserving does not amount to protecting any extensive part. It is rather a question of providing an actualization of the work’s power to act, whichever form that actualization may take.

Let us now focus on the process of individual extension, so that we shall understand how to preserve the actualization of a work’s power to act. We may consider that the work-related documents - created by an individual - maintain a certain relationship with the essence of that work. These documents constitute the extensive parts of the work depending on that relationship. In other words, the extensive parts of a work are not limited to the objects that we usually use to define a work. They also include each and every document that is related to the work, for these documents actualize a power to act of the work, by acting in the world.

Nevertheless, we can distinguish all the extensive parts of the work from one another, depending on how the digital apparatus is used in the context of a communication between the role of the author and that of the reader. Hence the following procedural model.

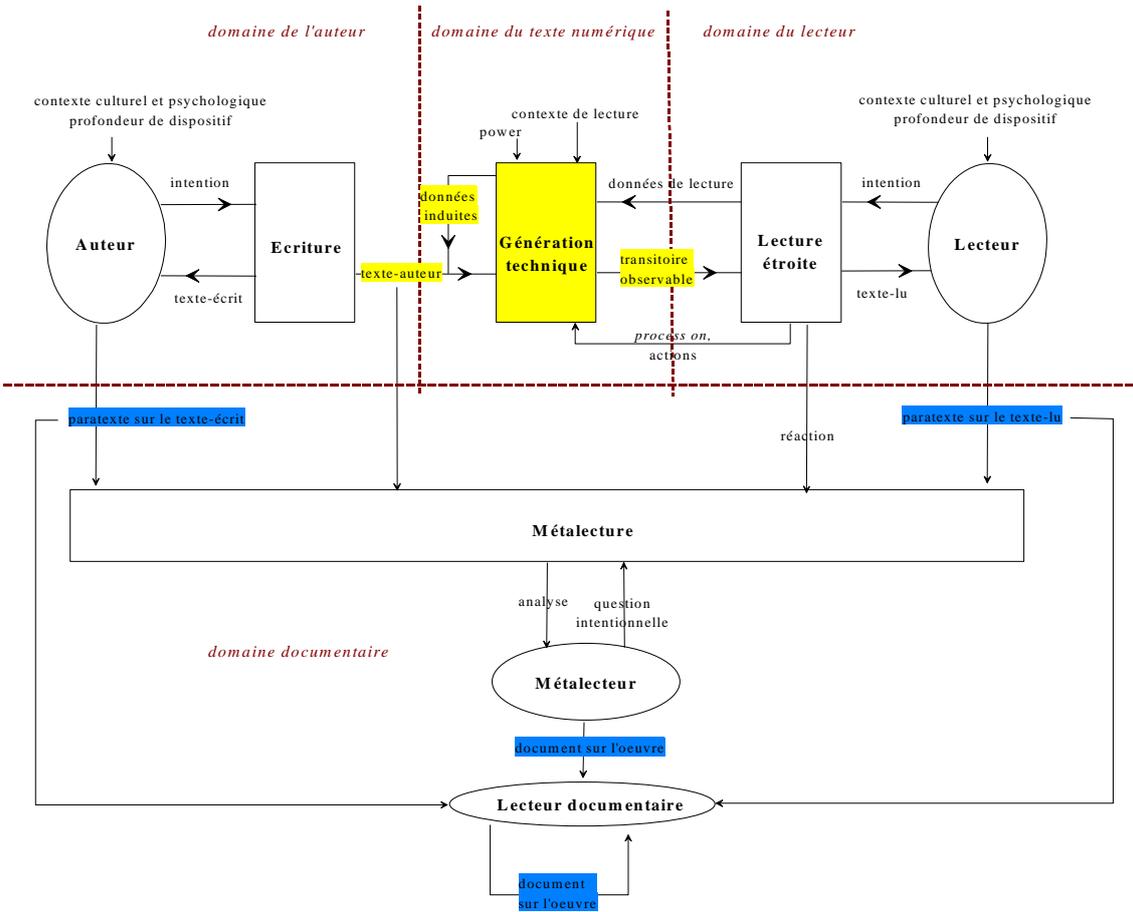


Figure 5 : functional diagram of the extended procedural model (Bootz, 2009)

Rather than using individuals in the communication process, the procedural model uses the “roles” they play. It thus completes the Spinozan ontological model, by giving details on the nature of the shocks that occur between the extensive parts of a work and human

beings' extensive parts. That model defines three roles: the role of the author, the role of the reader and the role of the metareader (notably that of an analyst.) The role of the reader establishes a text-to-watch, based on the noticeable transitory states that are created by the digital and technological apparatus (generating function.) It establishes a mental representation (read text) of a text-to-watch by resorting to interpretation. An individual who is in the position of a reader can re-integrate that representation into physical actions (e.g. the case of a performer) or in documents (e.g. the case of an analysis in close reading.) The role of the author consists in turning the mental representation of the work (which constitutes the true state of intention, i.e. the written text) into a set of documents (author-text) that is likely to be understood by men and to be used as the material granting access to the work. This is done thanks to the concatenation of technological operations, which are to turn them into perceptible states for the reader (observable transient.) The author role can be played by the whole of the creation staff. It shall result in the setting up of documents describing the author-text (graphic guidelines, functional guidelines...) The role of the metareader either consists in sorting out information from the author-text without resorting to the channel of the digital apparatus that turns it into an observable transient, or else in sorting out information on the reading process by analyzing reading situations. The information that is taken from the author-text shall be obtained via paratexts about this author-text (algorithms, presentations of structures or descriptions, listings...) These are usually given by the author himself. The metareading produces documents acknowledging the work. These documents do not have any specific name in this model. The analysis of the diachronic behavior of *mange-texte* presented earlier on is an example of metareading. The information we gathered on the author-text was not the result of a screen analysis (observable transient.) We owe it to Philippe Bootz, who is the programmer of this author-text and therefore perfectly knows the program.

The drawing of a diagram of these roles within the communication process reveals a "field of the work" comprised of specific extensive parts: the author-text, the observable transient, the author's and the reader's machines (software and hardware.) The relations between the extensive parts of the work and the essence of the work have the ability to enter this communication relation that explicitly displays a specific relationship with the digital apparatus, as well as special roles. The other documents do not have this triple propriety. Therefore, we shall identify the extensive parts that constitute the field of the work in the procedural model as the "initial body" of the work. All other documents shall be called "parergon", taking up and widening the Derridan concept of parergon, while keeping its pith and marrow.

### **3.3 Preservation and documentation**

This idea of the nature of the work leads us to consider every document related to a work as expressing one aspect of its power to act. It also leads us to consider the issue of preservation as inseparable from that of documentation. As mentioned earlier on, preserving a work amounts to finding the way to actualize its power to act, notably by documenting its many known aspects. In this perspective, the "initial body" of a work - which traditional methods aim at preserving or at recreating - is only a set of documents among others. Their obsolescence does not mean a loss of the work's power to act as long as its many aspects are documented. Neither does it mean that the "work" should be considered as a material object, of which we should thoroughly describe the extensive parts. What is at stake is its power to act. A work is mortal because its "initial body" shall disappear. Conversely, it is eternal because its power to act remains by expressing itself in relation to other extensive parts, despite the disappearance of its initial body. Indeed, the work's power to act does not only

depend on the perception of its initial body. Maybe more than in any other media, the work may produce side-effects in the digital media. Its power to act can be expressed by secondary discourses that are different from those we use to express the perception of its observable extensive parts. We may consider that every digital work destined to a private reading comprises the dimension of a conceptual work. Not because it *is* a conceptual work, but because the Web, as well as every other digital media, can be considered as a fictional space. Many authors (Lucie de Noutiny, Burgaud, Balpe among others) have underlined this idea in their works. Consequently, the perceptive components of a work shall play the role of a *referent*, while keeping a strong cultural action via secondary discourses. Let us note that this situation is not very far from the one experienced by archeologists. They cannot discover anything indeed, unless they “destroy” the site they work on and document that destruction (thus restoring a power to act.)

The documentation of the extensive parts of a work (initial body and parergon) is operated thanks to the indexing of every document dealing with the work, wherever they come from. Knowledge always depends on the relationship an individual keeps with the object of his knowledge. Therefore, knowledge of the work, which is necessary to preserve the possibility of an actualization of the work’s power to act, shall only be expressed according to different points of view.

As for our research, we shall consider as the fundamental hypothesis that our goal is to preserve the possibility to actualize a work’s power to act in the case of private-reading native digital works. It is an ontological issue rather than a communication one. Let us introduce three hypotheses in relation to this power to act:

- This power to act finds expression in all kinds of documents and not only in recordings or in documents given by the author. As a result, we propose ourselves to deal with all kinds of documents connected with the work, especially scientific analyses and readers’ reflections.
- This power to act cannot be controlled. It shall neither be mastered, nor frozen. A work is an autonomous entity - an “individual” – about which gathered documents shall let us know but very few aspects.
- The relation to the work, which is a prerequisite to its preservation, is based on a knowledge that depends on different points of view. Each point of view may give rise to a specific indexing of documents. That indexing will enable us to re-actualize certain aspects of the work’s power to act, but probably not *all of them at the same time*.

The work is thus perceived as an entity, which aspects we can but partially restore, because all of these need different actualizations. Preserving does not amount to gathering all of these aspects, so as to reconstruct an object. Rather, it is a matter of documenting the work and indexing that documentation according to several points of view, so as to recreate the conditions of a partial actualization. In case we wanted to recreate perceptive relationships, this actualization to come shall take one of the forms enounced in the projects we mentioned earlier on (simulation and recreation, in particular.) It shall also take on other forms (projection, conference, derived work...) that will keep it up-to-date.

In this perspective, the works edited in the first issues of *alire* were preserved thanks to Philippe Bootz, who re-programmed them in 1994. His aim was to build up an electronic

reading show. It was not a question of reproducing a visual state observed before the reading, but a matter of preserving the possibility of certain characteristics of the work to the detriment of possible others. The reading possibility has been preserved to the detriment of specific aesthetic projects possibly appearing on-screen. In *mange-texte* for example, the imperceptibility of the graphic transformation that performs substitutions between the distinctive elements of the letters was not preserved. Now, this was an aesthetic preconception that played with the relationship readers keep with time. In other words, preservation has played with the relationships between extensive parts and essence by modifying them. What is important here is that giving priority to a relationship eventually destroys some others that may prove to be as characteristic (and maybe as productive) of the work over time. For instance, we hardly take into account the fact that the program itself carries aesthetic representations, along with a representation of the world, regardless of the results produced by its running. The re-programming of a work often destroys these aspects, because the major part of the representation inherent to the program is linked with the dialog that takes place between the author and the programming language. The program cannot be reduced to a neutral description of what is displayed on-screen, for it comes from the author as much as it is inherent to the work, even if its display is not operated through the digital reading channel, but through that of the meta-reading. Thus, trying to singly preserve the various aspects of the work is in keeping with the recommended direction. It is precisely the way that Jim Andrews chose to follow as regards BpNichol's work. He reconstructed separate facets of the work thanks to the publishing of the program listing of the work, to an emulation of this work in the form of a current program and to a simulation of the original machine allowing us to run the original program on the most recent machines. Yet, this variety of approaches does not thoroughly account for the initial richness of the work. By sticking to the perceptive elements of the work and by cutting this latter off from its context of reception, these approaches do neither show its innovative characteristic, nor the shift it caused in its present horizon of expectation, for instance. The facets of a work shall not be reduced to its sole perceptive dimension. In effect, the other aspects of a work require the documentation of multiple aspects according to multiple points of view. This is hard to do afterwards, but an adaptation should be welcome, as in the case of a film adaptation of a novel.

The stance we take here on preservation takes up many of the characteristics of the preservation projects we developed in the first part, especially as regards the link they draw between preservation, documentation and indexing. As such, it does not nullify these projects; instead it puts them into perspective, as well as it leads to the extension of their field.

## **4 ArchiPoeNum: An adapted indexing tool.**

We have already begun to conceive and design a tool (Bootz & Szoniecky, 2008) to index all kinds of documents related to digital poetry (including the initial body of the work.) This tool is based on a functional diagram that is common to both the Spinozan ontology and the procedural model. Here, it takes the form of a Firefox plug-in (<https://addons.mozilla.org/fr/firefox/addon/10935>), in order to allow a free and automatic access to its different evolutions. The whole of the digital literature community will have access to that tool, as well as the developers interested in this Open Source project (<http://code.google.com/p/archipoenum/>). Each indexing will frame the document according to the adopted point of view. Hence, all kinds of documents can be indexed, from an ephemeral performance to a printed article, through the program of an executable work. That tool may be used by connecting to the Web, as well as in an autonomous way. A synchronisation protocol will allow one to easily handle the multiple versions of the indexing

files and to solve their possible conflicts. The data access is ensured by the setting up of an open archive in the form of an OAI storage. This latter will give access to the indexed documents, as well as to the works published on-line.

#### **4.1           Simplicity and standardization for a better access.**

The implementation choices, as regards developments, were caused by the will to give the easiest and most generic tool possible. Our aim is to make the indexing data easily transmissible, either with the ArchiPoeNum tool, or else directly via e-mail. As well, the editing of the indexing processes can be performed with any text-editing software<sup>2</sup>. One requires a browser that is compatible with the basic technology used (i.e. SVG,) in order to check the indexing results. We have chosen the XML language to define and edit the indexing data, for it allows us to gather the following operations within a single independent file:

- all the indexing processes in the form of a procedural code in JavaScript
- the semantic references needed by these processes in the form of a RDF network
- the result of these processes by saving the user's preferences

Thus, we are able to easily build up what we may call ontological agents, i.e. dynamic graphic representations possessing a potential of semantic evolution, which the user may upgrade as he pleases.

In order to make the indexing data work with one another, we need semantic references. These are marked in a RDF network, according to 4 fundamental axes: time, space, concepts, and persons. We will use the following standard formalisms:

- Time

Apart from the Dublin Core metadata specifications (<http://library.princeton.edu/departments/tsd/katmandu/html/dcdate.html>) that we will use to index the bibliometric data in a standard way, we will use the simile timile format ([http://code.google.com/p/simile-widgets/wiki/Timeline\\_EventSources](http://code.google.com/p/simile-widgets/wiki/Timeline_EventSources)) in order to index the more complex temporal data, i.e. the data that is linked with periods of time or readers' reactions, for instance. The usual dating conventions shall be used as far as edited documents are concerned.

- Space

As regards places and in case a document contains a richer data than a simple editing place does, such as in a performance for instance, we will use a KML expression, a standard that is well-known nowadays in the definition of geographical objects in 2D and 3D (<http://code.google.com/apis/kml/documentation/kmlreference.html>)

- Semantics

---

<sup>2</sup> Nevertheless, we advise you to use SVG editing tools such as inkscape (<http://www.inkscape.org/>)

The semantic dimension will be expressed in OWL formalism. Furthermore, other semantic languages may be used, especially via the implementation of already developed ontologies or via the use of a semantic dictionary such as IEMML (<http://www.ieml.org>)

- Persons

A questionnaire will be made available to the person who performs the indexing, in order to record the data related to him/her. No personal questions will be asked, but rather information that will help us distinguish an indexing from another. The data will be in FOAF format, in order to allow the establishment of a social network, as would be the case with a research team for example, or in order to define the users of a multimedia library.

## 4.2 Current state of the tool.

This tool is currently developed by Mundigo, in partnership with the Paragraphe laboratory. Here are the currently implemented functionalities of the tool. At first, these functionalities are presented with UML diagrams that describe the use cases, activity diagrams and transitional state diagrams. Secondly, some commented read-out copies will give details on the interface of the tool and on the access to these functionalities.

### 4.2.1 Use Case

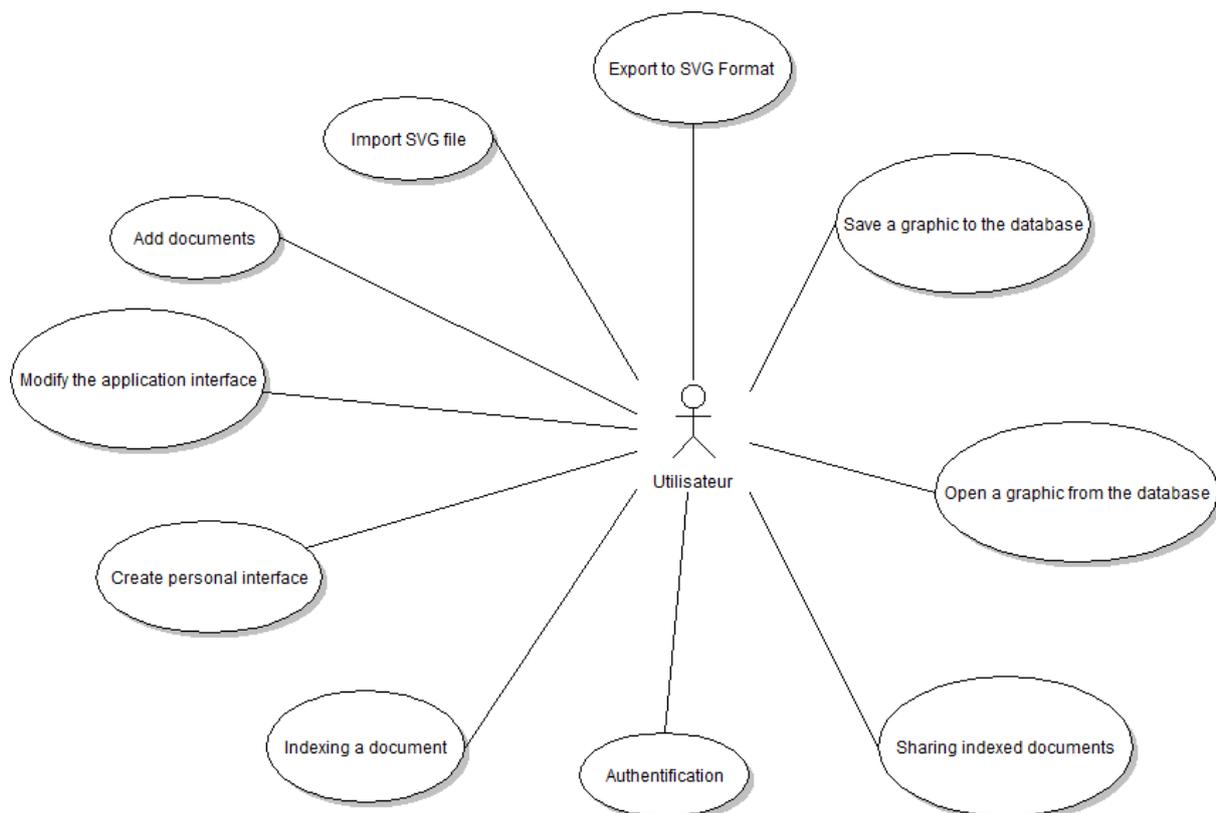


Figure 2 : use case

In our application the user is the only player in the system and can perform several operations:

- A new document Indexer poetry by adding digital information in the application interface.
- Open a document already indexed and stored in the database.
- Save the current document in the database.
- Authenticate by a login and password.

- Export the document as SVG.
- Import an SVG file in the interface.
- Edit and customize the application interface.
- Share documents indexed.

### 4.2.2 Activity diagrams

This diagram describes the different steps to get to export the current document format SVG.

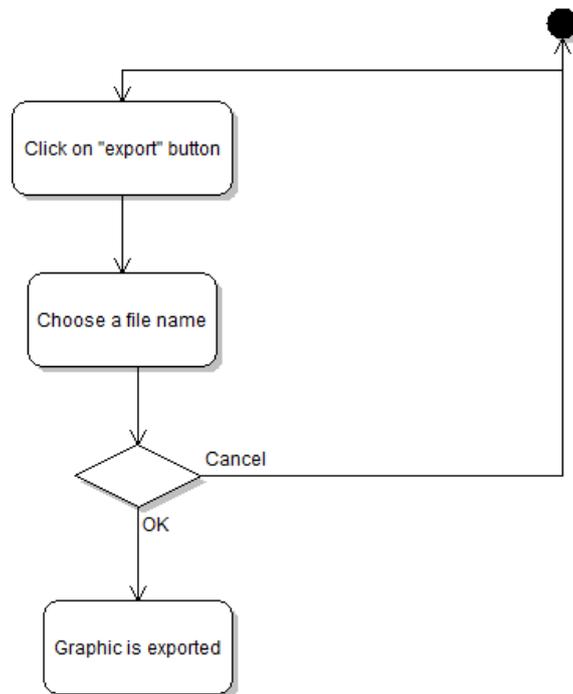


Figure 3 : export SVG

This diagram describes the different steps to import a SVG file in the application interface

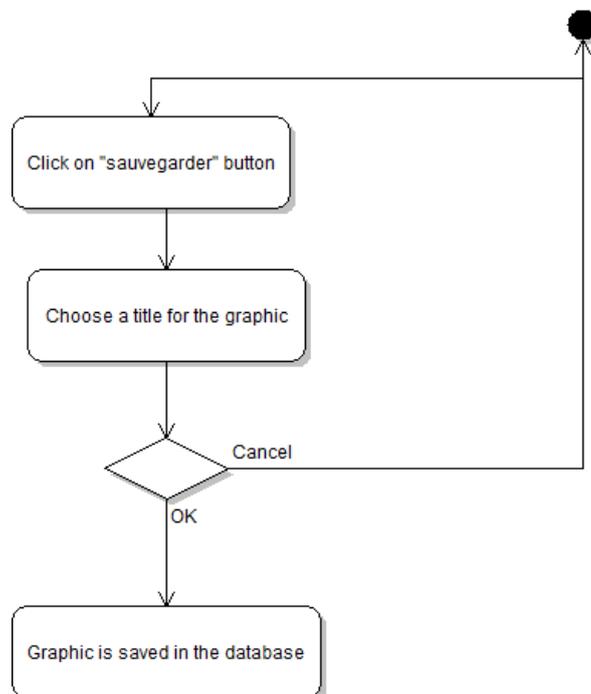


Figure 4 : save SVG

This diagram describes the steps to open a SVG file from the database.

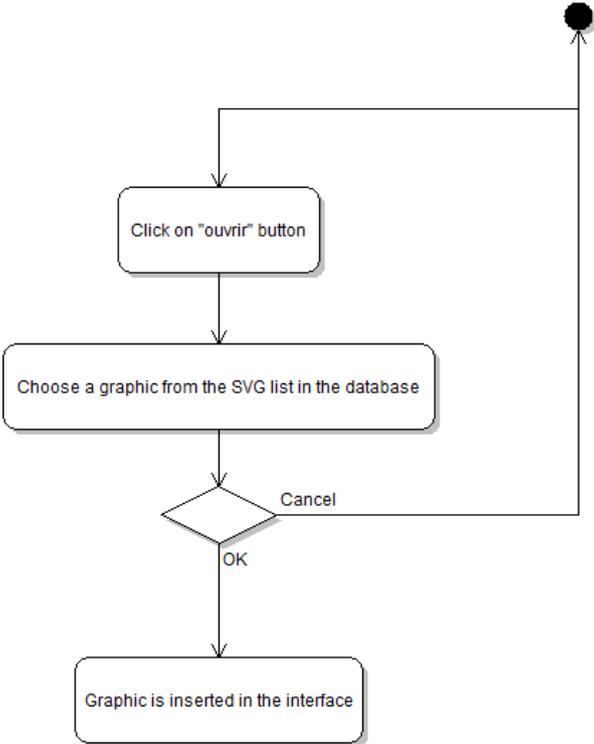


Figure 5 : open SVG

This diagram describes the authentication of a user in the application.

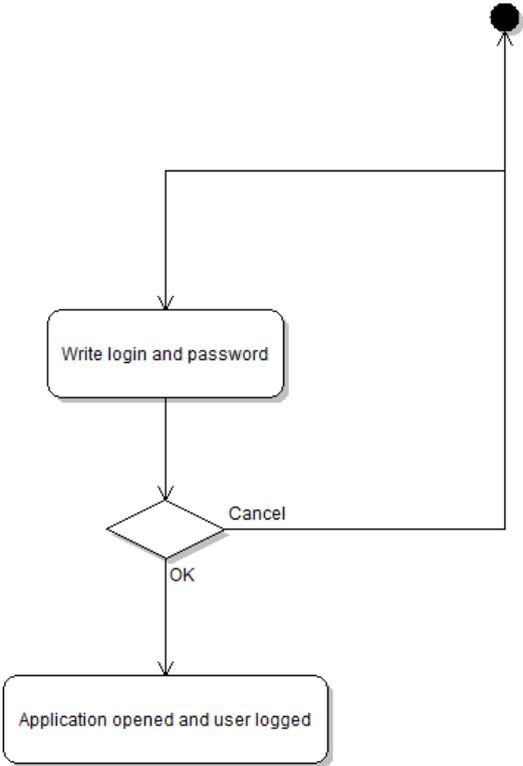


Figure 6 : login

This diagram describes how users can customize its interface.

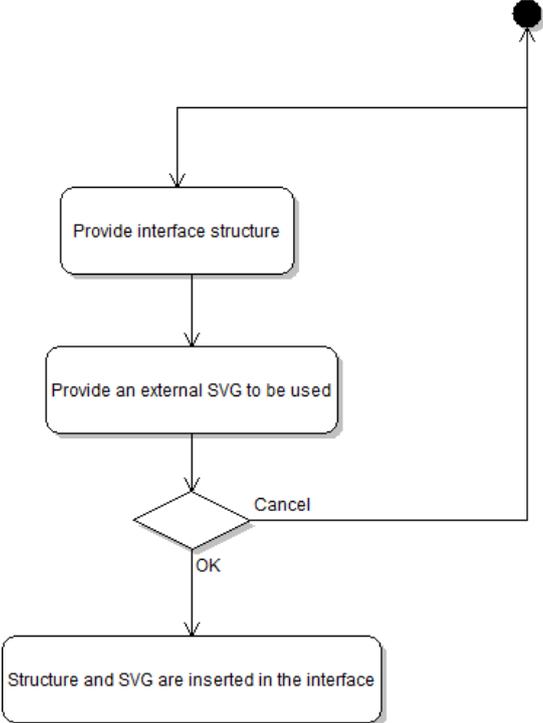


Figure 6 : customize interface

This diagram describes the different stages for the indexing of a document in the application.

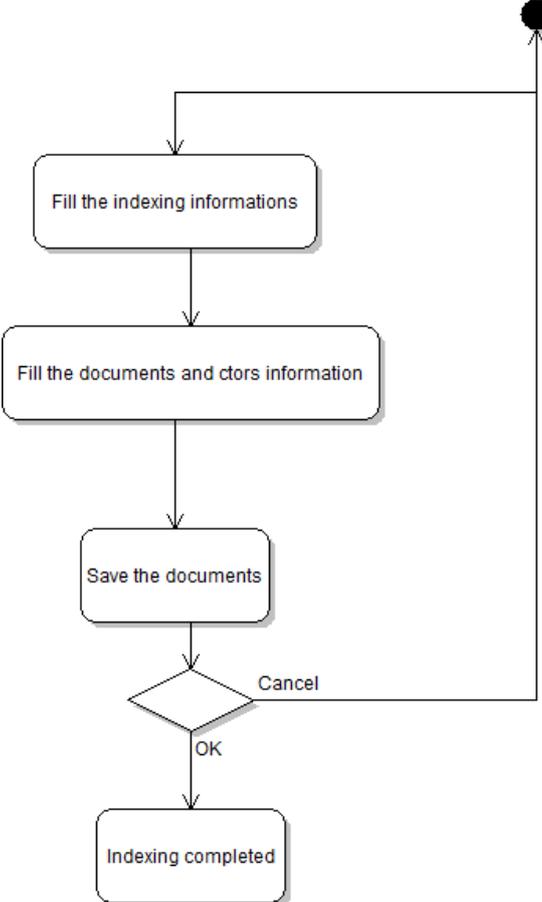


Figure 7 : indexing document

This diagram describes the different steps to share a document with other users.

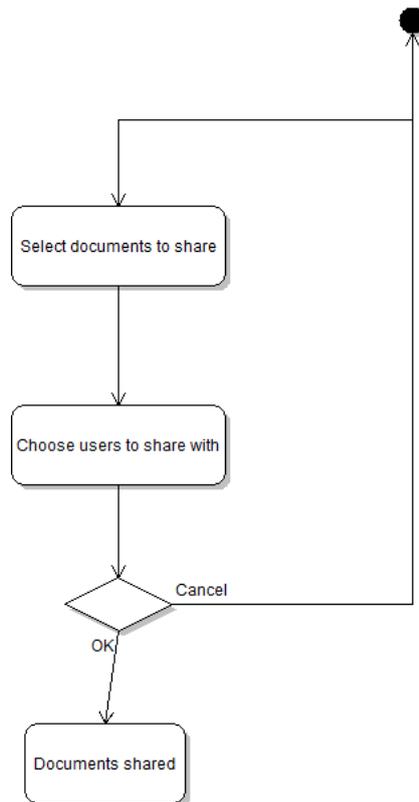


Figure 8 : share document

#### 4. 2. 3 State Chart Diagram

This diagram describes the state of the system during the export.

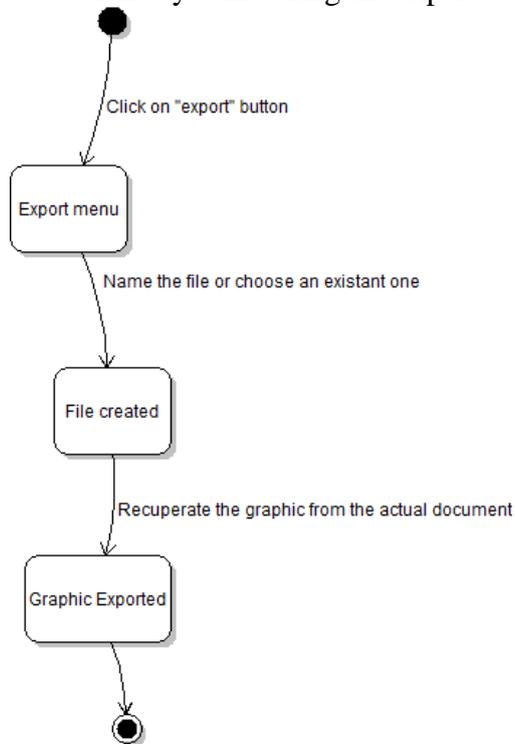


Figure 9 : export

This diagram describes the state of the system during the import.

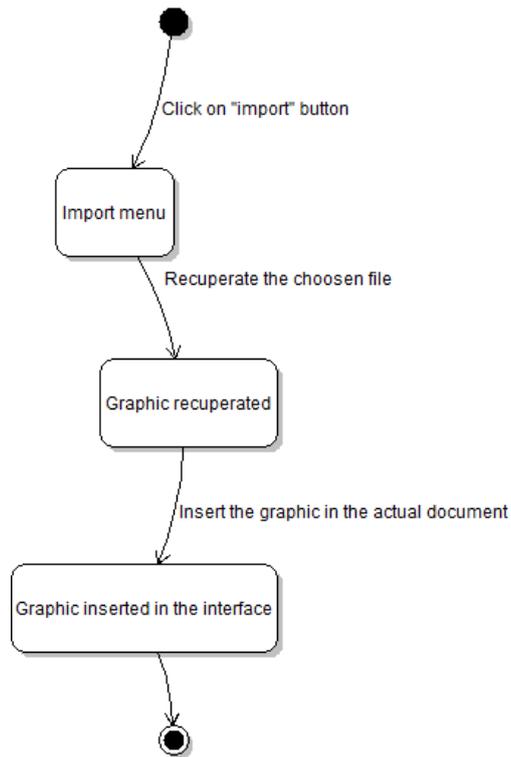


Figure 10 : import

This diagram describes the state of the system during the backup file in the database.

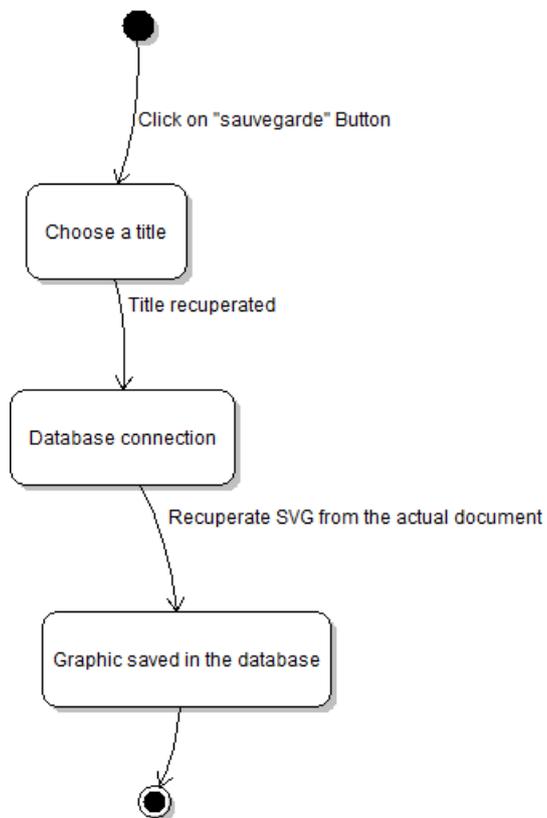


Figure 11 : backup

This diagram describes the state of the system during the opening phase of a document from the database.

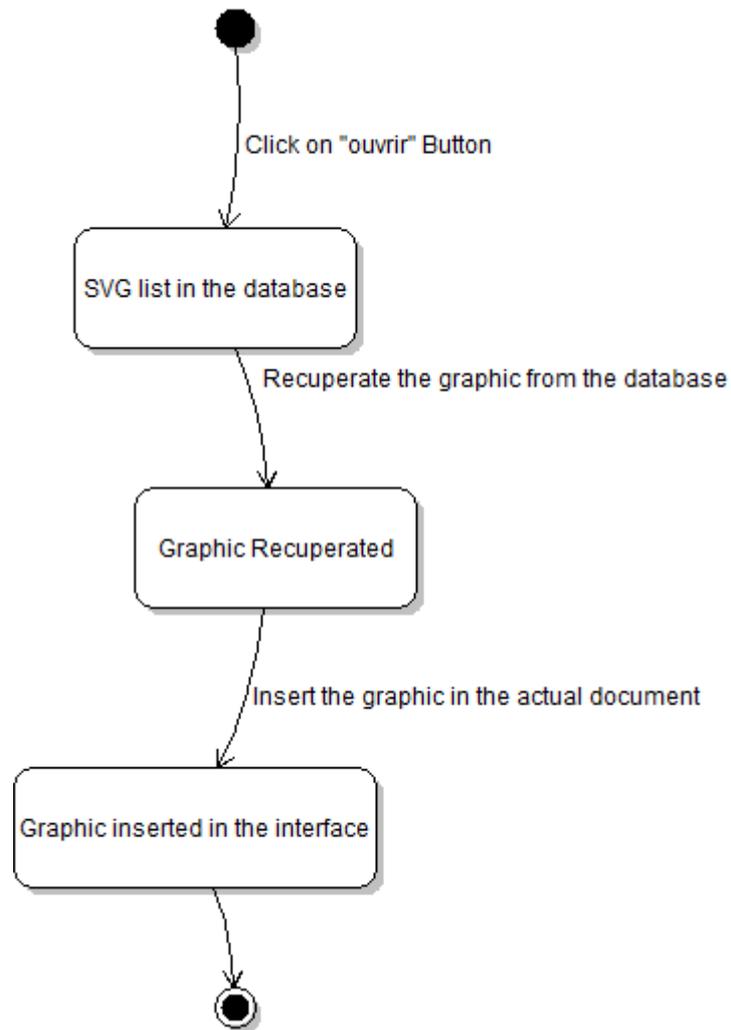


Figure 12 : open document

#### 4. 2. 4 Screenshot

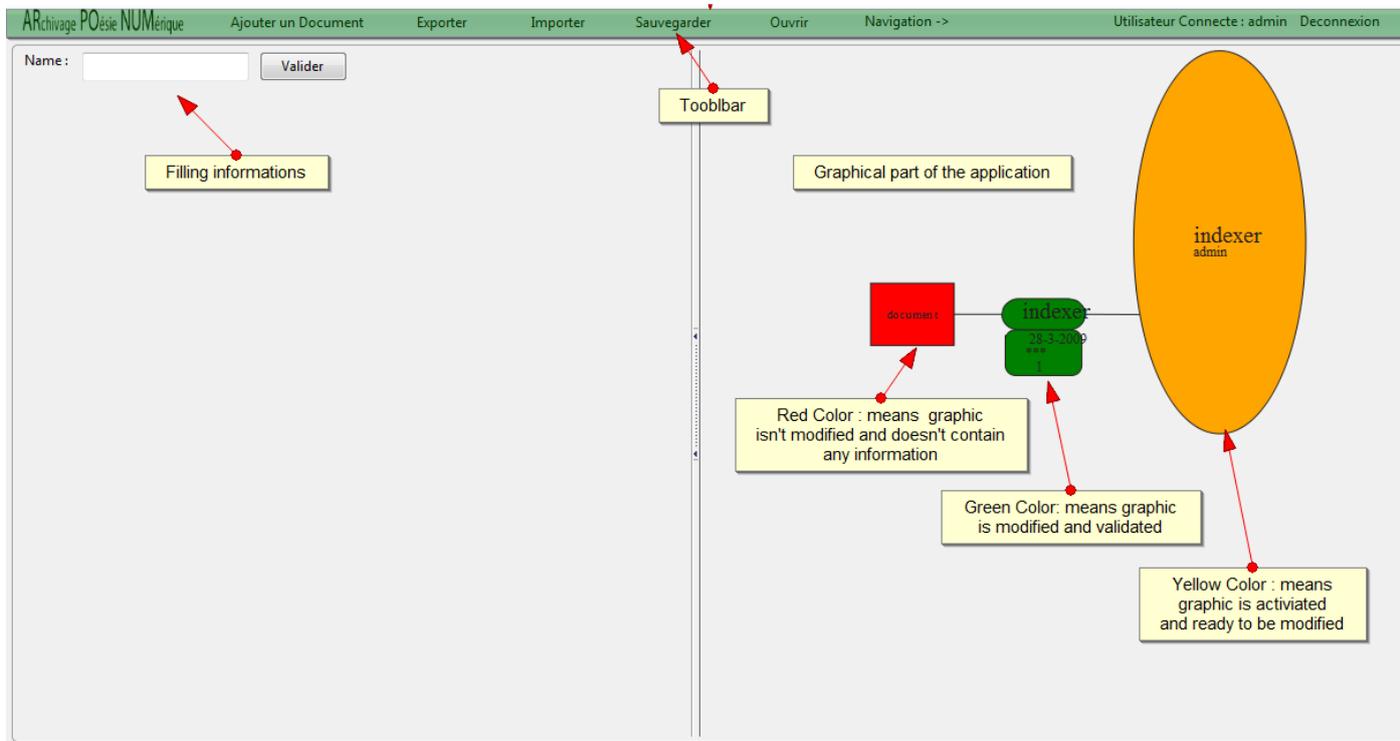


Figure 13 : Interface color signification

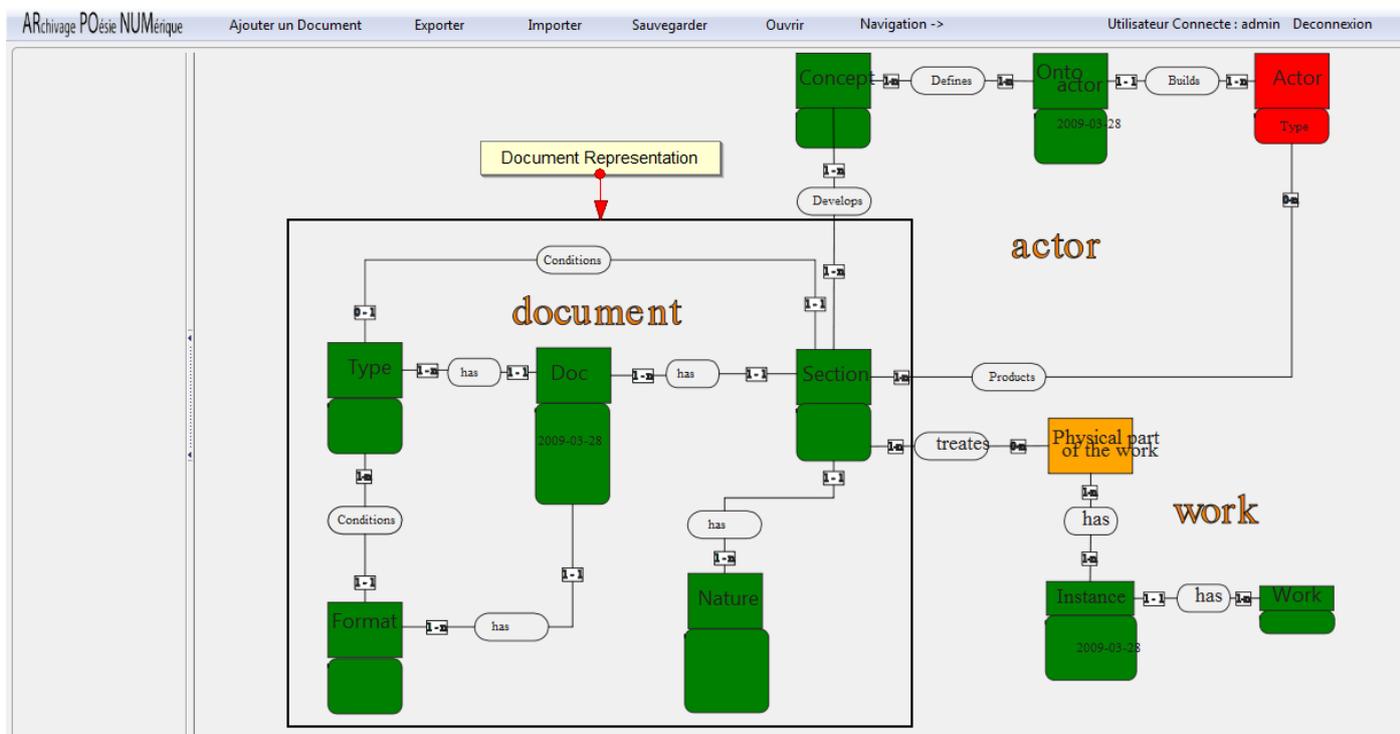


Figure 14 : Document representation

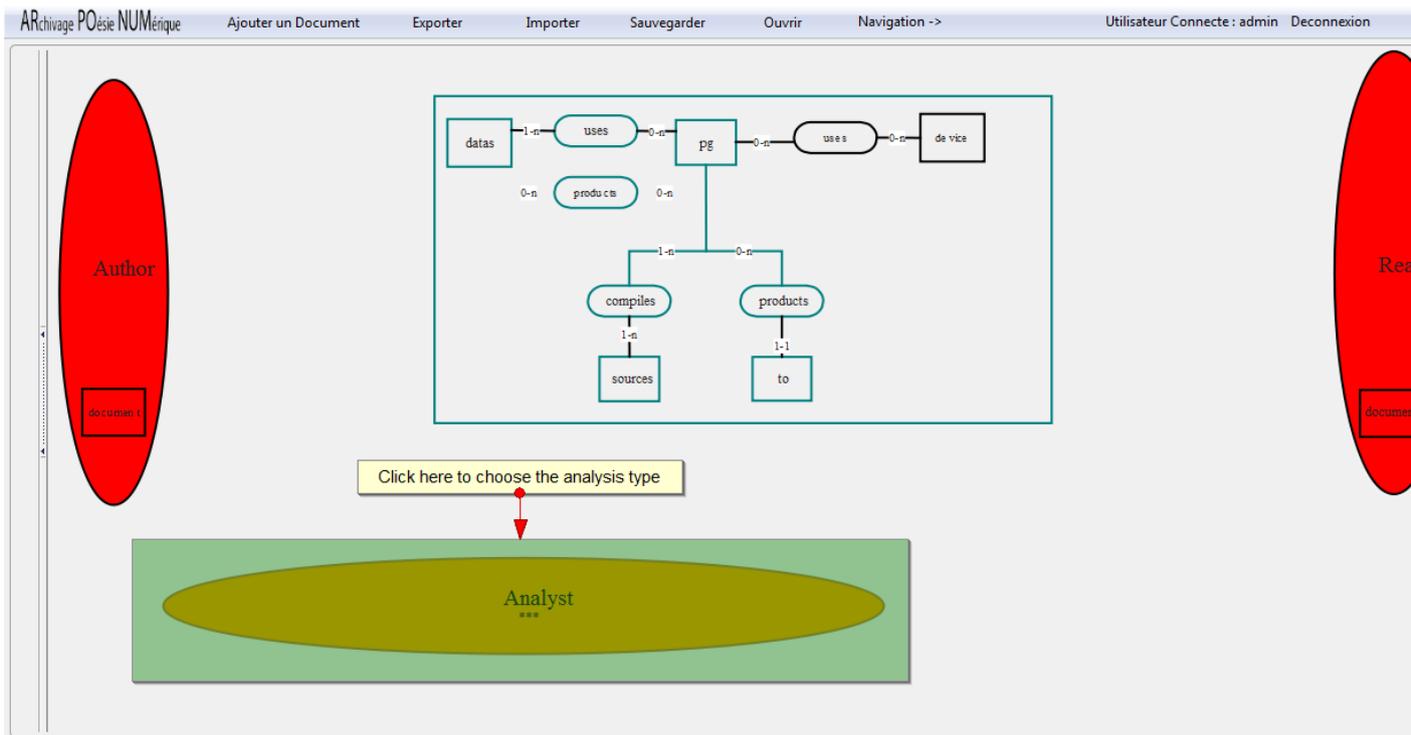


Figure 15 : Actor representation

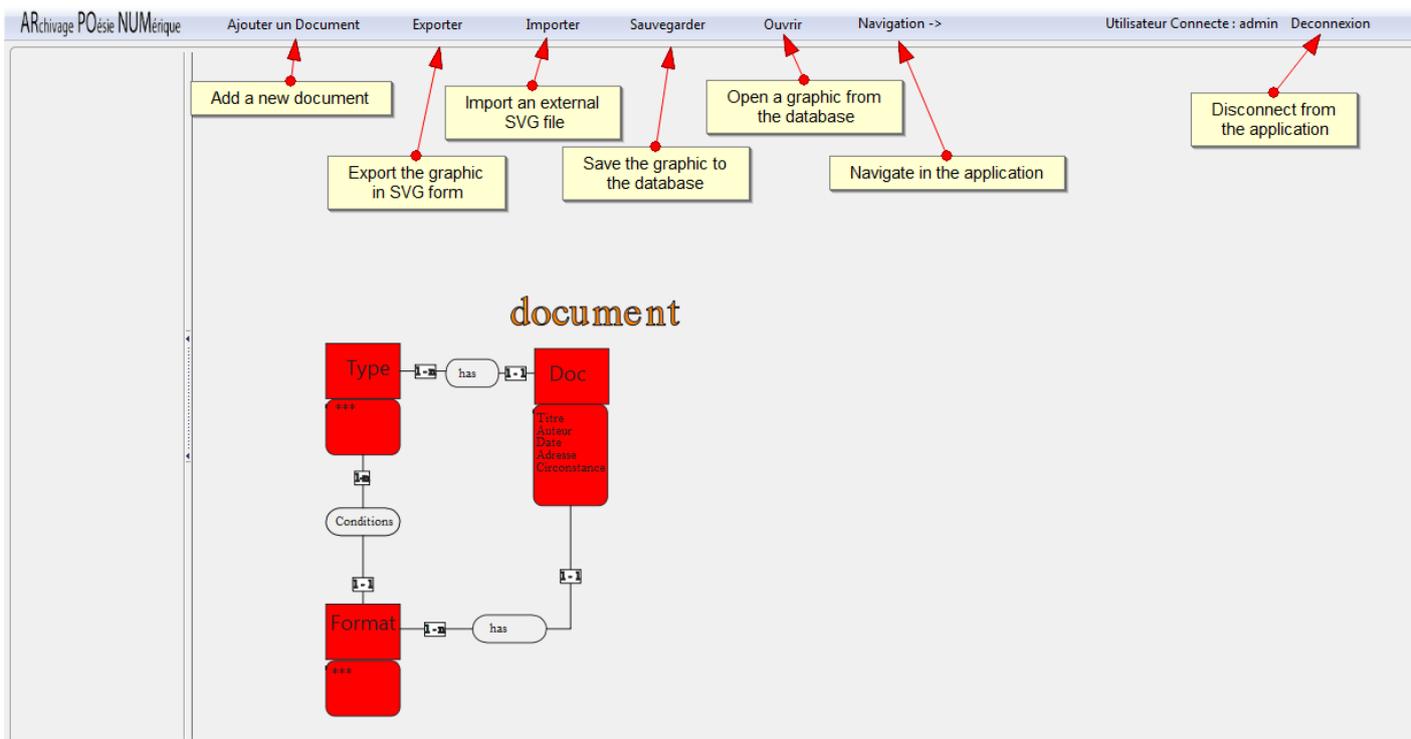


Figure 16 : Document representation

## 5 Bibliographie

Bootz, Ph. (2003). « Construction du sens par les processus dans le modèle procédural », *Hypertextes, hypermedias*, Hermes, H2PTM'03, 2003, pp. 79-88.

Bootz, Ph. (2004). “der/die leser ; reader/readers”, article bilingue allemand/anglais, in F. Block, C. heibach, K. Wenz (éds), *p0es1s. Asthetik digitaler Poesie, The Aesthetics of Digital Poetry*, Hatje Cantz Verlag, Ostfildern, Deutschland, pp. 93 – 121, <http://www.p0es1s.net/poetics/symposion2001/bootz.pdf>

Bootz, Ph. (2010). “Quelques hypothèses sur la lecture en poésie numérique », Lecture numérique et usages du Web, actes des journées d’études 14 et 15 avril 2009, Tunis : Institut Supérieur de Documentation, pp. 125-151..

Bootz, Ph. & Szoniecky, S. (2009). « toward an ontology of the field of difgital poetry », colloque electronic Literature in Europe, Bergen 11-13 septembre 2008, <http://elitineurope.net/node/11>

Deleuze, G. (2001). *Spinoza : immortalité et éternité*, cédrom/K7 audio. Paris : Gallimard

Derrida, J. (1978). *La Vérité en peinture*, Paris : Flammarion, pp. 64 et 83

Giaretta et al. (2006). Caspar and a European Infrastructure for Digital Preservation. *European Research Consortium for Informatics and Mathematics*, News 66. [http://www.ercim.org/publication/Ercim\\_News/enw66/giaretta.html](http://www.ercim.org/publication/Ercim_News/enw66/giaretta.html)

Gladney, H. M. (2004). “Principles for Digital Preservation, preprint”, Electronic Resource Preservation and Access Network, [http://eprints.erpanet.org/70/01/Principles\\_preprint.pdf](http://eprints.erpanet.org/70/01/Principles_preprint.pdf)

Gladney, H. M. (2006). Principles for Digital Preservation. *Communications of the ACM*, 49(2), pp. 111-116

Ippolito, J., Depocas ,A., Jones, C .(2003). *L’approche des médias variables : la permanence par le changement* ; New York et Montréal : The Solomon R. Guggenheim Foundation, et la fondation Daniel Langlois pour l’art, la science et la technologie.

Klinkenberg, J.M. (1996). Précis de sémiotique générale. Louvain-la-Neuve : De Boeck & Lancier.

Laforet, A. (2009). *La conservation du net art au muse. Les strategies à l’oeuvre*. Thèse de doctorat, Université d’Avignon et des Pays de Vaucluse. <http://www.sakasama.net/conservationnetart/these/theseannelaforet0309.pdf>

Lee, K.-H., Slattery, O., Lu, R., Tang, X. & McCrary V. (2002). The State of the Art and Practice in Digital Preservation. *Journal of Research of the National Institute of Standards and Technology*, 107(1), pp. 93 - 106. <http://nvl.nist.gov/pub/nistpubs/jres/107/1/j71lee.pdf>

Liu, A., Durand, D., Montfort, N., Proffit, M., Quin, L., Réty, J-H., and Wardrip-Fruin, N. (2005). *Born-again Bits. A Framework for Migrating Electronic Literature*. ELO, <http://eliterature.org/pad/bab.html>

Semple,N. & Clifton, G. (2007), *DPC/PADI What’s new in digital preservation, n° 16 (april-august 2007)*, <http://www.dpconline.org/docs/whatsnew/whatsnew16.pdf>

Tabbi, J. (2004). Acid-Free Bits and the ELO PAD Project, <http://eliterature.org/pad/afb.html#sec0>

Wittgenstein, L. (1953). *Investigations philosophiques*. Trad. Fr. 1961, Paris : Gallimard

### **Oeuvre analysée :**

Dutey, J. M (1989). Le mange-texte, *alire1*.

---

<sup>1</sup> In other words, the part of the read-out that is to be interpreted.