

# CIBERTEXTUALIDADES

Publicação do CECLICO - Centro de Estudos Culturais, da Linguagem e do Comportamento



TEMA DE CIBERTEXTUALIDADES 05

## ELECTRONIC PUBLISHING MODELS FOR EXPERIMENTAL LITERATURE

Organização de **Rui Torres** e **Manuel Portela**

# ficha técnica

## DIRECTOR

**Rui Torres**

## DIRECTOR-ADJUNTO

**Pedro Reis**

## CONSELHO DE REDACÇÃO

**Jorge Luiz Antonio** - Investigador Independente

**Sérgio Bairon** - Universidade de São Paulo, Brasil

**Pedro Barbosa** - Investigador Independente (Professor Aposentado,  
Escola Superior de Música e Artes do Espectáculo, Portugal)

**Luis Carlos Petry** - Pontifícia Universidade Católica de São Paulo, Brasil

**Manuel Portela** - Universidade de Coimbra, Portugal

**Pedro Reis** - Universidade Fernando Pessoa, Porto

**Fátima Silva** - Universidade Fernando Pessoa, Porto

**Rui Torres** - Universidade Fernando Pessoa, Porto

## COMISSÃO DE HONRA

**Maria Augusta Babo** - Universidade Nova de Lisboa, Portugal

**Jean-Pierre Balpe** - Université de Paris VIII, França

**Jay David Bolter** - Georgia Tech, Atlanta, E.U.A.

**Philippe Bootz** - Université de Paris VIII, França

**Claus Clüver** - Indiana University, Bloomington, E.U.A.

**José Augusto Mourão** (in memoriam)

**Winfried Nöth** - Universität Kassel, Alemanha

**Lúcia Santaella** - PUC-São Paulo, Brasil

**Alckmar Luiz dos Santos** - Universidade Federal de Santa Catarina, Brasil

**Alain Vuillemin** - Université d'Artois, França

## TÍTULO

**Revista Cibertextualidades 05 (anual) - 2013**

© Universidade Fernando Pessoa

## EDIÇÃO

edições UNIVERSIDADE FERNANDO PESSOA

Praça 9 de Abril, 349 | 4249-004 Porto

edicoes@ufp.pt | www.ufp.pt

## DESIGN

Oficina Gráfica

da Universidade Fernando Pessoa

## DEPÓSITO LEGAL

241 161/06

## ISSN

1646-4435



## Vers une ontologie du domaine de la poésie numérique

Philippe Bootz<sup>1</sup> & Samuel Szoniecky<sup>2</sup>

**RÉSUMÉ:** L'essai se propose d'établir le modèle théorique d'un outil d'indexation fondé sur l'ontologie spinoziste appliqué au domaine de la poésie numérique. Pour ce faire, il commence par proposer une modélisation de l'individu spinoziste dans le formalisme du modèle entité/relation utilisé pour décrire des bases de données relationnelles fondée sur les trois dimensions que sont l'essence, les parties extensives et les relations entre essence et parties extensives. Il démontre que cette schématisation est compatible avec le schéma de communication du modèle procédural utilisé pour analyser les oeuvres de poésie numérique. Il propose alors un protocole d'indexation fondé sur la mise en oeuvre de ces deux modèles. L'intérêt principal de cet outil est de prendre nativement en compte la possibilité de divers points de vue et la focalisation de l'indexation sur des parties précises de l'oeuvre.

**MOTS-CLÉ:** indexation, ontologie, spinoza, poésie numérique

**RESUMO:** Este artigo tem como objectivo propor um modelo teórico para a criação de uma ferramenta de indexação baseada na ontologia de Spinoza aplicada ao campo da poesia digital. Para tal, começa por propor uma modelagem do indivíduo em Spinoza no contexto do formalismo do modelo entidade/relação utilizado para descrever bancos de dados relacionais com base nas três dimensões que são a essência, as partes extensivas e as relações entre essência e as partes extensivas. Pretende-se mostrar que esse mapeamento é compatível com o esquema de comunicação do modelo processual utilizado para analisar as obras de poesia electrónica. Em seguida, propõe um protocolo de indexação baseado na implementação destes dois modelos. A principal vantagem desta ferramenta é ter em conta a possibilidade de diferentes pontos de vista e a focalização da indexação sobre partes específicas da obra.

**PALAVRAS-CHAVE:** Indexação, ontologias, Spinoza, poesia electrónica

<sup>1</sup> Docteur en physique. Professeur agrégé de physique. Docteur en sciences de l'Information et de la Communication. Maître de Conférences à l'Université Paris 8, France. Co-responsable de l'équipe Ecritures et Hypermediations Numériques du Laboratoire Paragraphe. Mail: philippe.bootz@univ-paris8.fr

<sup>2</sup> Docteur en science de l'information et de la communication. DEA Histoire de l'art. Ingénieur informatique pour des grands comptes.

## 1. Introduction

Cet essai se propose d'établir le modèle théorique d'une indexation du domaine de la poésie numérique fondée sur l'ontologie de Spinoza commentée par Deleuze<sup>3</sup>. Son objectif est de concevoir un outil graphique simple, permettant une indexation semi-automatique des documents et une visualisation graphique des résultats.

Spinoza propose un modèle ontologique valable pour tout type d'entités individualisées dénommées individus. L'individu peut donc être un être vivant, bien sûr, mais également un objet inanimé comme un document ou une œuvre. Spinoza oppose éternité et immortalité en considérant qu'un individu peut se transformer et mourir. Il définit ainsi l'individu à l'aide de trois dimensions constitutives : ses parties extensives physiques et matérielles, son essence singulière et le rapport qui les unit dans l'ici et maintenant. Un individu peut se modifier en ajoutant à ses parties extensives des éléments extérieurs, mais toujours selon un rapport particulier à son essence. Les relations entre individus se réalisent au moyen de contacts entre parties extensives nommés «chocs». Un choc fournit, pour une entité vivante, une première expérience affective et émotionnelle, ce que Spinoza nomme une «connaissance inadéquate». Les «connaissances adéquates» reposent, en plus des chocs, sur des relations entre rapports ou entre essences. Ces connaissances correspondent à des connaissances scientifiques ou intuitives. La mort d'un individu se traduit par la perte du rapport entre les parties extensives et l'essence mais cette dernière est éternelle et constitutive du «grand tout» universel des essences.

La transformation et la mort des individus décrite par le modèle spinoziste apporte une flexibilité qui permet facilement de traduire les modifications et versions diverses des œuvres comme des reconfigurations de rapports. Il devient ainsi possible de distinguer dans l'ontologie les différentes configurations d'une œuvre et de tenir compte de son évolution chronologique. Par ailleurs, les parties extensibles s'identifient aisément aux composantes de l'œuvre observables et manipulables par une personne (lecteur, auteur...) en contact avec l'œuvre.

En ce qui concerne le document, le modèle fournit ipso facto une séparation entre le plan de l'expression (parties extensives) et celui du contenu, donc les concepts qu'il contient (essence). Les concepts développés dans ces documents fournissent une connaissance sur l'essence de l'œuvre mais ne constituent nullement cette essence. En effet, dans ce modèle, les documents sont produits dans un rapport spécifique à l'œuvre. On en déduit donc que l'essence de l'œuvre

<sup>3</sup> Deleuze, Spinoza, éternité et immortalité, cours audio, Gallimard.

est inaccessible dans l'absolu mais qu'en revanche une connaissance sur cette essence se construit par ces documents. Cette connaissance peut être infinie et s'organise nécessairement en points de vue qui correspondent aux rapports que le créateur du document établit avec l'œuvre. Ce créateur peut tenir un rôle quelconque : auteur, lecteur, analyste, performeur, technicien... Notons encore que le modèle permet de retenir toutes sortes de «connaissances», depuis les simples sentiments (c'est bon, j'aime...) jusqu'aux prises de position suggestives en passant par les résultats argumentés et étayés par une description scientifique. Le modèle permet donc de traiter tout type d'approche de l'œuvre. Les rapports qui s'établissent entre le document et l'œuvre en vue d'obtenir cette connaissance reposent sur la communication entre le créateur du document et l'œuvre. Ils traduisent donc des points de vue portés sur l'œuvre.

Philippe Bootz développe depuis plusieurs années un modèle de communication qui permet de classer les rapports qu'un individu entretient avec l'œuvre sur la base des relations qu'il installe avec les parties extensives de l'œuvre. Cette classification fait alors ressortir des rôles dénommés acteurs. Les acteurs sont caractérisés par un objectif commun (par exemple analyse sémiotique, interprétation par la lecture...). On peut ainsi regrouper les concepts développés par les approches dans un rôle donné en un faisceau de relations conceptuelles qui constitue une ontologie d'acteur établie de façon classique par un faisceau de relations entre concepts. De telles ontologies restent encore largement à construire sur la base des recherches actuelles menées sur la poésie numérique dans toutes nos disciplines. Retenons ici que le modèle d'indexation proposé est ainsi extensible à l'infini par la multiplication des ontologies d'acteurs (approche historique, approche technique, approche littéraire, approche sémiotique, approche de l'auteur, approche de lecteur...)

La relation qu'une personne donnée entretient avec l'œuvre est également changeante. Elle se traduit par une reconfiguration des rapports que la personne entretient avec l'œuvre, et ce au cours même de la création du document. Autrement dit, dans un même document, on peut repérer les rôles successifs endossés par la personne, ses différents points de vue, même si l'objectif de l'approche est unique. On pourra donc repérer des sections dans les documents, chaque section étant relative à un point de vue. Ce sont elles, en définitive, qui devront être indexées. La définition des sections permet ainsi de régler avec la finesse voulue l'indexation des documents : depuis une indexation très approximative dans laquelle le document dans son ensemble est considéré comme une seule section jusqu'à une indexation pointilliste dans laquelle on décompose le document en sections très petites. Par exemple, on peut indexer le document en repérant les sections où l'auteur du document se positionne comme lecteur, celles où il établit une analyse littéraire, d'autres où il oriente son discours dans une direction sémiotique, autant de sections qui utilisent des ontologies d'acteurs différentes. Si on indexe le document complet comme une section, alors l'indexation fera référence à l'ontologie d'acteur la plus prégnante.

Ainsi donc, l'association de l'approche ontologique spinoziste, d'ontologies d'acteurs établies par des spécialistes de diverses disciplines et d'un modèle de communication décrivant les parties extensives de l'œuvre et les contacts entre les acteurs et ces parties permet-elle théoriquement de réaliser une indexation souple et flexible, à condition que tous ces modèles puissent être décrits dans un même formalisme.

Nous allons montrer qu'un formalisme graphique utilisant un modèle ensembliste et un modèle entité/association permet de décrire le modèle spinoziste. Nous établirons ensuite le modèle ontologique de l'œuvre en appliquant ce formalisme au modèle procédural. Nous déduisons également de cette formalisation la définition de plusieurs rôles repérés comme des acteurs dans le modèle procédural. L'analyse des relations entre le concepteur de document et l'œuvre telle qu'elle se dévoile dans le modèle procédural nous permettra enfin de proposer une ontologie du document et de détailler la description de ses parties extensives. L'adjonction d'ontologies d'acteurs du domaine (ou de folksonomies) permettra alors de proposer une base de concepts en vue de l'indexation de ces documents.

Outre sa cohérence théorique, le modèle proposé permet de concevoir un outil d'indexation graphique simple et puissant pouvant traiter tout type de document. En effet, le formalisme ensembliste du modèle ontologique est parfaitement approprié à la construction d'une interface graphique vectorielle SVG. La visualisation des connaissances sur une œuvre ou un ensemble d'œuvres peut également se faire graphiquement à l'aide d'ensembles de Venn. Enfin la description des relations sous forme de modèles entités/associations fournit une visualisation vectorielle pouvant servir d'interface graphique pour la détermination à la souris des points de vue adoptés. Certains schémas théoriques seront donc transformés en formalisme SVG et exploités par l'interface.

Nous nous sommes attachés à développer le modèle formel de façon à dégager les modalités de construction des schémas d'interface. L'exemple proposé est juste destiné à montrer comment un tel outil pourrait être utilisé pour indexer un document et comment il pourrait permettre une visualisation des informations. L'objectif, à terme, n'est pas de réaliser une indexation automatique mais plutôt de construire une base de données d'indexations autonome permettant de visualiser des points de vue différents.

## 2. Schématisation du modèle de Spinoza et Deleuze

### 2.1. Schématisation de l'individu

Spinoza définit un individu selon trois dimensions comme «une infinité de parties extensives extérieures les unes aux autres (1° dimension) qui appartiennent à l'individu sous des rapports caractéristiques (2° dimension) qui le constituent et expriment une essence singulière (3° dimension) qui le constitue», cette essence étant un degré de puissance.

On peut aisément représenter cette conception (Fig. 1) par un modèle Entité/Association (modèle E/A) dans lequel essence et parties extensives sont des entités et les rapports des relations. Ce schéma est également un schéma ensembliste constitué de 3 sous-ensembles différents (parties extensives, rapport, essence) de l'ensemble «individu» schématisé par une «patate» ensembliste. C'est pour rappeler cette double représentation mathématique dans laquelle les entités du modèle E/A sont des sous-ensembles typés que je schématise dans ce modèle en leur affectant un symbole graphique différent :



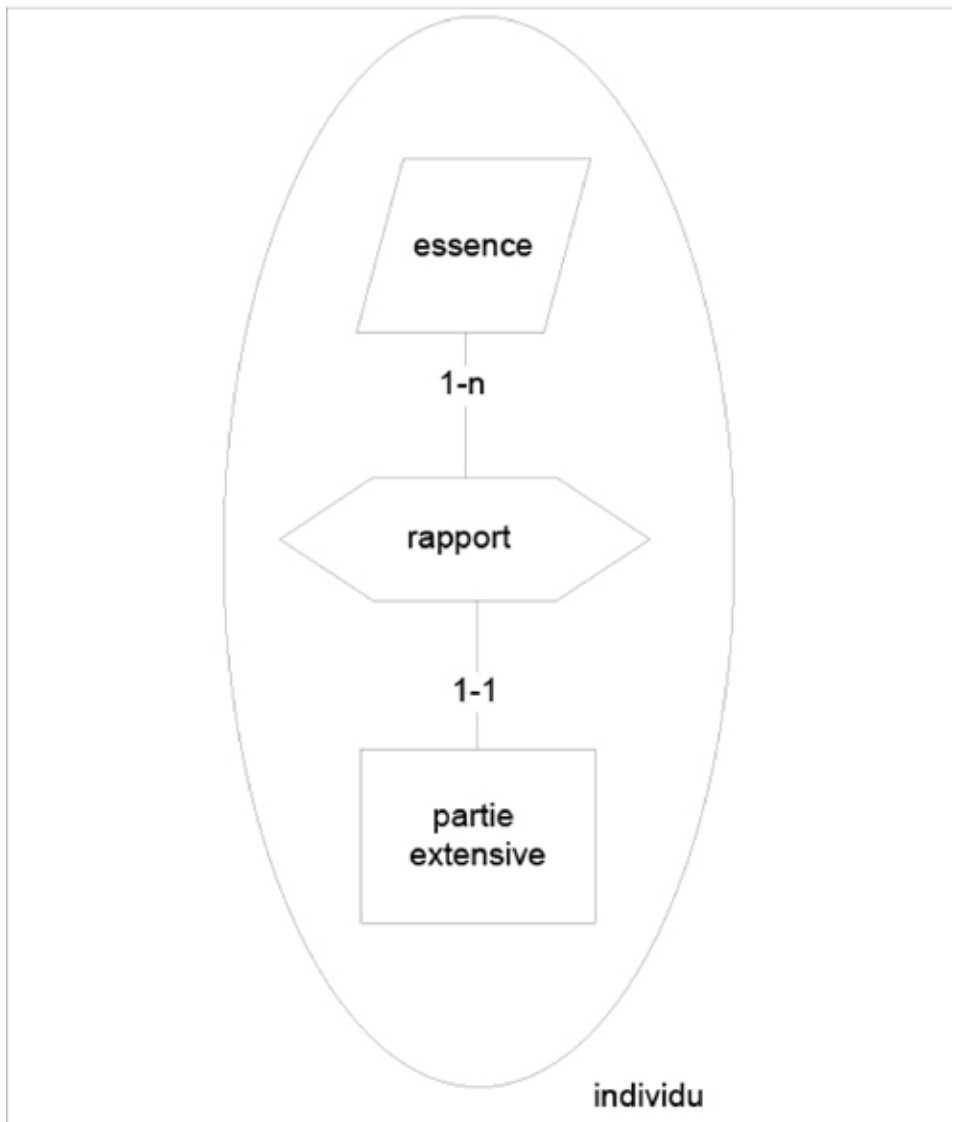


Figure 1: Représentation de l'individu

## 2.2 Schématisation des relations entre entités

Spinoza relève trois genres de connaissances qu'un individu peut avoir sur un autre (étant entendu que l'individu peut tout aussi bien être un phénomène naturel). Le premier type de con-

naissance est constitué des affects, ce que Spinoza nomme les «connaissances inadéquates». Cette connaissance résulte pour lui de la simple rencontre (choc) entre les parties extensives des deux individus (Deleuze prend l'exemple du barbotage pour illustrer cette connaissance : la relation à l'eau y est purement sensuelle et affective). Il s'agit d'une connaissance des effets du contact entre entités. Le second type de connaissance (ex : la science) porte sur les rapports. Cette connaissance aboutit à un savoir-faire qui permet de composer les rapports caractéristiques qu'un individu entretient avec ses parties extensives avec les rapports que l'autre individu entretient avec ses propres parties extensives, de sorte que le contact ne se réduit plus à un simple choc subit entre parties extensives, mais consiste en une composition de rapports (Deleuze prend l'exemple de la nage, activité dans laquelle les rapports que j'entretiens avec mes membres se composent avec les rapports que l'essence de la vague entretient avec ses molécules). Le troisième genre de connaissance (connaissance intuitive) porte sur les essences. Les 2° et 3° genres forment ensemble les «connaissances adéquates».

On peut schématiser de deux façons les relations entre individus qui aboutissent à ces connaissances. La première (Figure 2) est une schématisation purement ensembliste qui ne détaille pas les relations entre individus. La seconde (Figure 3) est proche d'un modèle entité/association en ce qu'elle relie les composantes d'un individu. Elle s'en éloigne toutefois en ce que la relation de composition relie des associations (les rapports) et non des entités et en ce que les entités «essence» sont toutes complémentaires les unes des autres et constituent ensemble le grand tout divin.



Figure 2: schématisation ensembliste globale des relations entre individus

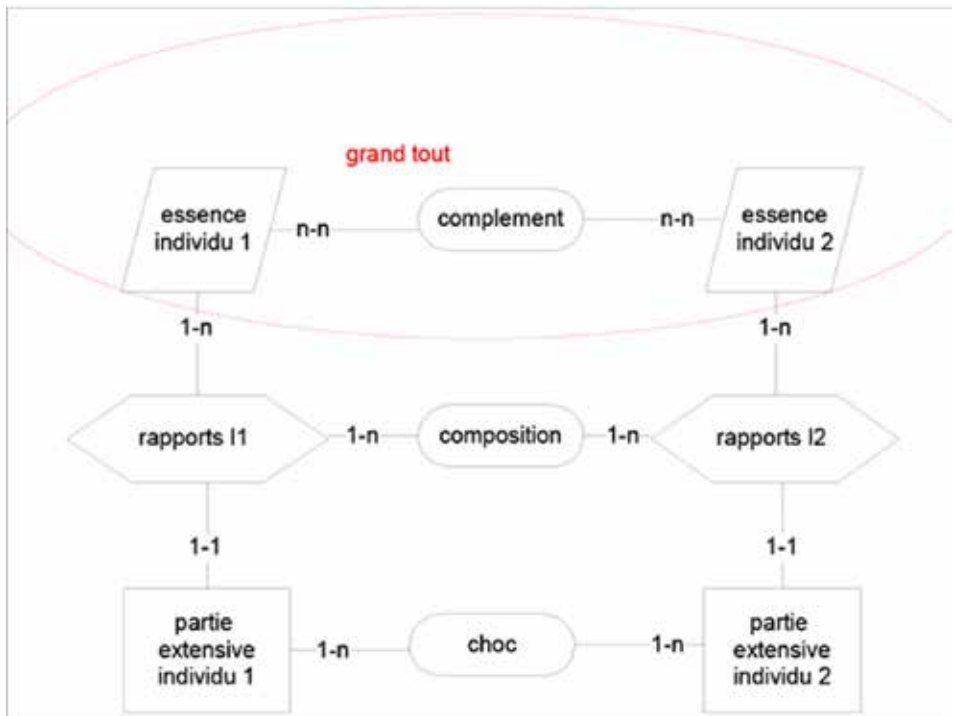


Figure 3: schématisation E/A des relations entre individus

## 2.3 schématisation de l'évolution diachronique d'un individu

### 2.3.1 extension d'un individu

L'ajout d'un axe des temps permet d'établir une représentation diachronique des deux phénomènes majeurs de l'ontologie spinoziste : l'extension d'un individu et son cycle de vie.

Suite au contact avec un objet B, un individu peut étendre ses parties extensives en créant un rapport caractéristique entre cet objet et son essence (Figure 4).

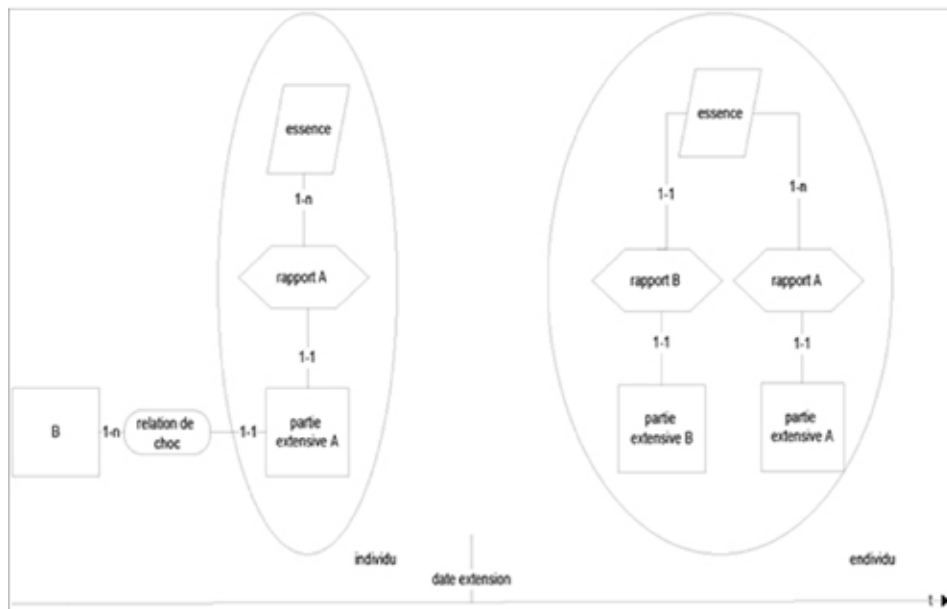


Figure 4: schématisation diachronique de l'extension d'un individu

Après extension, la partie extensive B et la partie extensive A sont alors en relation interne indirecte par l'intermédiaire de l'essence de l'individu et des rapports qu'elles entretiennent avec cette essence commune. Cette relation se traduit par un schéma simple E/A sous la forme

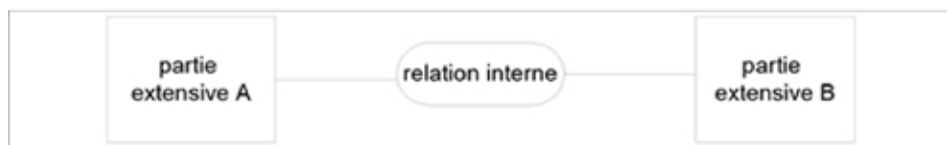


Figure 5: relation implicite entre parties internes à une entité

Ainsi donc, il n'est pas nécessaire de schématiser le fonctionnement diachronique en introduisant un axe des temps, il suffit de considérer qu'en cas d'extension toute relation entre les parties extensives d'un individu et l'objet incorporé par extension représente également, par convention, la relation interne constituée par cette extension. Cela ne signifie pas que les deux relations (le choc avant extension et la relation interne après extension) soient de même nature, cela signifie simplement que la description schématique demeure la même (Figure 6) avant et après l'extension, donc à n'importe quel moment. Afin de ne pas perdre les informations sur l'événement «extension», nous attribuons des attributs à la relation : la date, le lieu et l'id de cette extension. Cette convention de sché-

matisation qui transforme un axe des temps en une relation avec attributs sera nommée «équichronisme schématique».

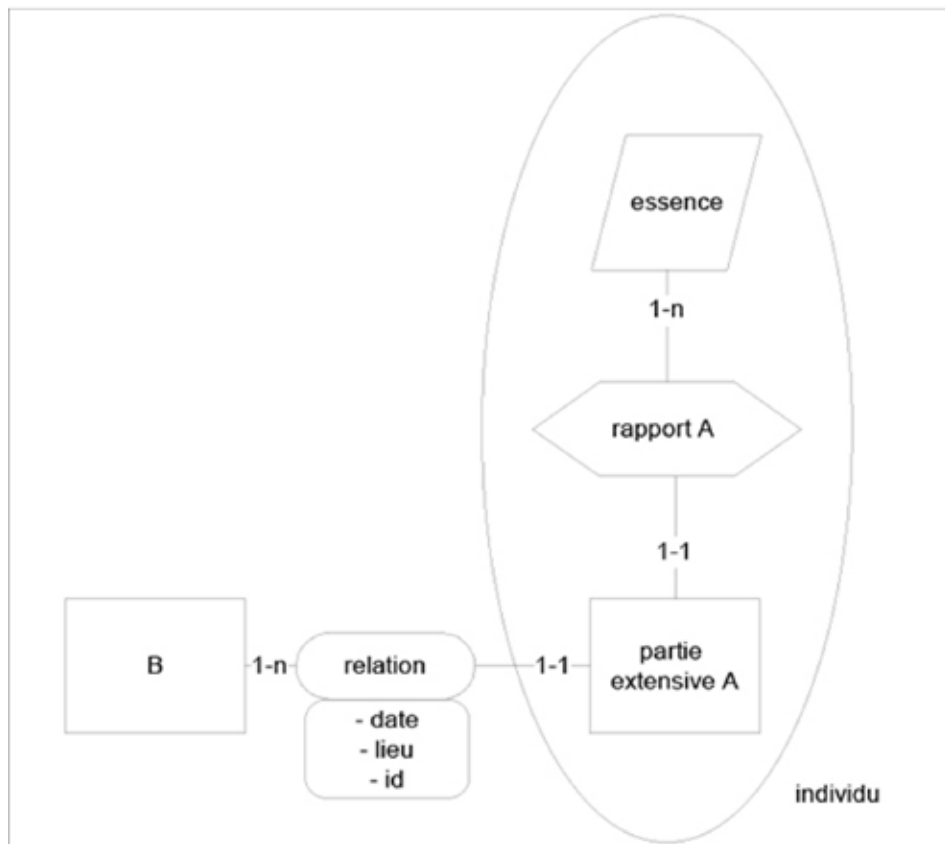


Figure 6: schéma équichronique complet de la relation d'un individu à un objet

Dans la pratique, nous nous intéresserons au contact entre un acteur et l'œuvre. L'objet B sera alors une partie extensive de l'œuvre et l'individu A, l'acteur. Nous n'allons pas nous intéresser à tout l'individu mais uniquement aux documents qu'il produit en résultat de ce contact. Il n'est donc pas utile de faire apparaître sur le schéma le détail de l'individu A et nous utiliserons la schématisation simplifiée de la figure suivante :



Figure 7: schéma équichronique simplifié du contact entre une partie extensive B de l'œuvre et un acteur

### 2.3.2 cycle de vie d'un individu

Spinoza indique que la naissance d'un individu consiste en la création de rapports caractéristiques entre objets qui le constituent, ce qui les transforme en parties extensives, à une certaine date et en un certain lieu, alors que la mort est la destruction de ces rapports, les objets constituant les anciennes parties extensives devenant parties extensives d'autres individus par création d'autres rapports.

Un tel cycle peut également se représenter dans notre schématisation en introduisant l'axe des temps. En nommant A l'objet physique qui subit les transformations, on obtient :



Figure 8: vie et mort de l'individu 1

On peut également s'affranchir de la schématisation de l'axe des temps et adopter une schématisation équichronique en ajoutant des attributs de date (début et fin) et de lieu (de naissance) au rapport qui relie l'essence aux parties extensives durant la vie de l'individu. Le rapport est donc repéré par 4 attributs : ces trois là et, bien sûr, son nom qui constitue un indice de son essence.

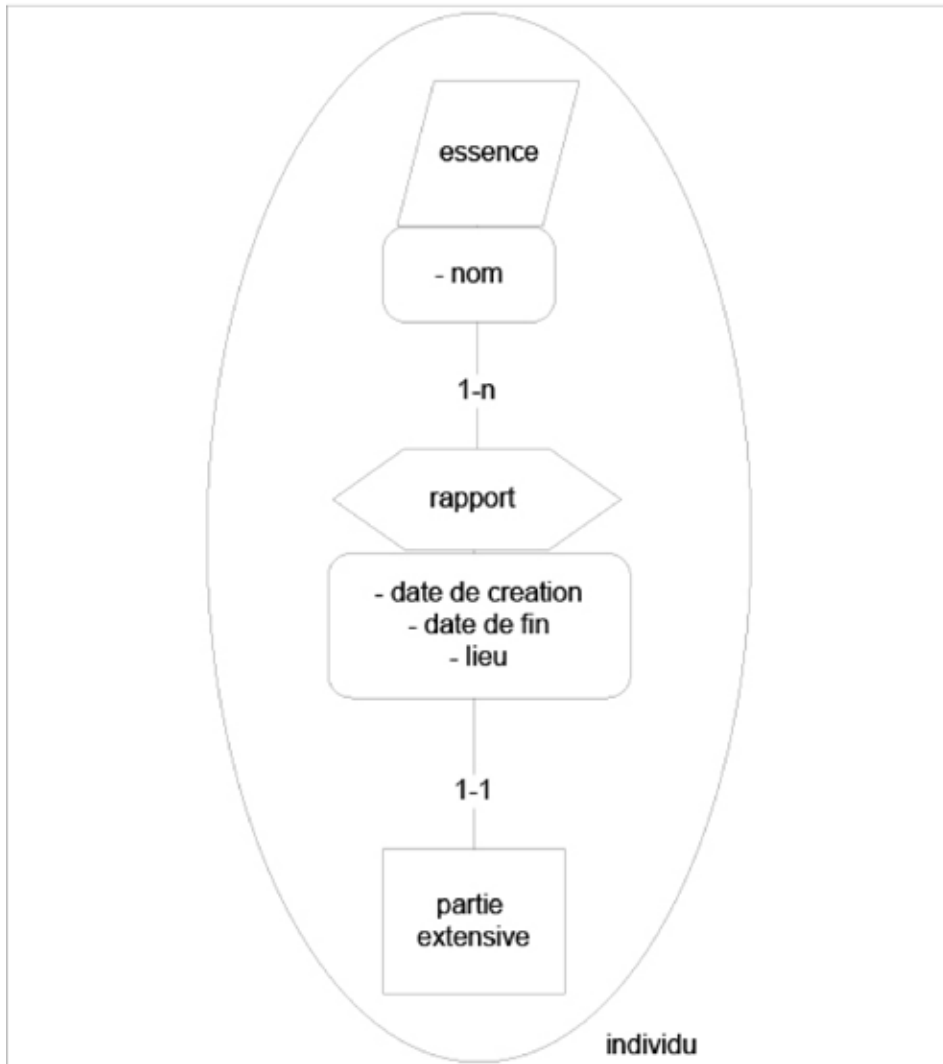


Figure 9: représentation équilibrée d'un individu

Notons que les parties extensives ne sont parties, justement, que parce qu'elles entretiennent un certain rapport avec l'essence. Ce rapport, en quelque sorte, les définit selon un certain point de vue qui est celui de l'essence et il n'est pas possible de les définir autrement que par ce point de vue.



### 2.3.3 Théorème de l'équichronisme schématique

On peut récapituler les transformations décrites ci-dessus dans le théorème de l'équichronisme qui permet de schématiser une évolution diachronique des relations entre individus et objets extérieurs (ou autres individus) sans avoir besoin de décomposer le schéma en diverses étapes grâce aux conventions de schématisation suivantes:

convention de constitution de l'individu : l'ajout au rapport d'attributs de dates (création et suppression) ainsi que de lieux de vie indique que le diagramme E/A qui relie essence et parties extensives par l'association du rapport en question n'est valide que durant l'intervalle de dates mentionné et aux lieux considérés. En dehors de ces conditions, les rapports n'existent plus et les parties extensives reprennent leur identité d'objets isolés (éventuellement liés à d'autres essences par d'autres rapports)

convention d'extension : les associations entre les parties extensives d'un individu et des objets externes ou des parties extensives d'autres individus sont représentés sur le schéma avant l'extension mais en ajoutant les attributs permettant de caractériser l'extension. Elles traduisent également les relations internes entre ces parties qui apparaissent après extension par le fait qu'elles sont alors liées à la même essence selon des rapports différents. Il n'est donc pas nécessaire d'établir un nouveau schéma pour décrire graphiquement la situation après extension dès lors qu'on ne s'intéresse qu'aux parties extensives et non aux rapports. Dans le formalisme équichronique, toutes les relations sont schématisées avant extension, parce que, dans un contact, chaque individu crée des rapports d'extension avec l'autre, ce qui compliquerait singulièrement la représentation.

## 3. Le modèle procédural

Le modèle procédural<sup>4</sup> fournit un cadre permettant d'appliquer le modèle spinoziste à la poésie numérique. Il stipule que les parties extensives de l'oeuvre sont constituées d'un "source" situé dans le domaine de l'auteur et d'un "transitoire observable" situé dans celui du lecteur. Les objets matériels de l'oeuvre sont alors intégrés au domaine de l'acteur selon les rapports interprétatifs que peut faire l'acteur. Ces rapports correspondent à une extension de l'acteur.

<sup>4</sup> On trouvera une description plus précise du modèle dans Philippe Bootz, «The Problem of Form. Transitoire Observable, a Laboratory for Emergent Programmed Art», Peter Gendolla & Jörgen Schäfer (eds.) *The Aesthetics of Net Literature. Writing, Reading and Playing in Programmable Media*, Transcript-verlag, Bielefeld, 2007 : 89- 103.

Les rapports interprétatifs d'un auteur en contact avec le source transforment ce source en un texte-auteur défini comme l'ensemble des éléments qu'il manipule et qui sont signifiants pour lui. Le texte-auteur est ainsi la forme que prend l'œuvre pour l'auteur dans le processus d'extension de l'auteur. Il ne correspond pas aux fichiers binaires manipulés par la machine mais à leur représentation manipulable par l'auteur, à savoir au code source tel qu'il apparaît dans l'outil auteur et aux données images et son avant compression éventuelle. Par exemple, une vidéo de qualité très dégradée conçue pour une faible bande passante ne correspond pas au texte-auteur de cette vidéo.

Le transitoire observable est le résultat (souvent multimédia) perceptible produit par la machine lors de l'exécution du programme. Ce n'est pas un signe mais un ensemble média susceptible d'être capturé par un capteur ad hoc (micro, caméra...). C'est, dans le langage de Spinoza, une partie extensive de l'œuvre. La lecture va avoir pour conséquence de poser des filtres et déformations liées au mécanisme percepto-cognitif de l'interprétation sur ce transitoire observable, qui va, dès lors, être intégré au domaine du lecteur en tant que «texte-à-voir». Par exemple, souvent, les éléments de l'interface ne seront pas considérés comme pertinents par le lecteur, encore moins l'interface du bureau. Ils ne font alors pas partie du texte-à-voir alors qu'ils font partie du transitoire observable. Le texte-à-voir est donc l'expression du rapport qui s'établit entre le transitoire observable et le lecteur dans le processus d'extension du lecteur lors de la lecture.

Ainsi donc programme, données et transitoire observable constituent les parties extensives de l'individu «œuvre» alors que le texte-auteur et le texte-à-voir constituent des parties extensives respectivement de l'auteur et du lecteur. Le modèle fonctionnel détaille la communication qui s'établit entre les acteurs selon deux modes descriptifs : une description structurelle (Figure 10) qui met en évidence tous les éléments et relations et un modèle fonctionnel (Figure 11) qui définit l'œuvre comme un domaine interface entre celui de l'auteur et celui du lecteur.

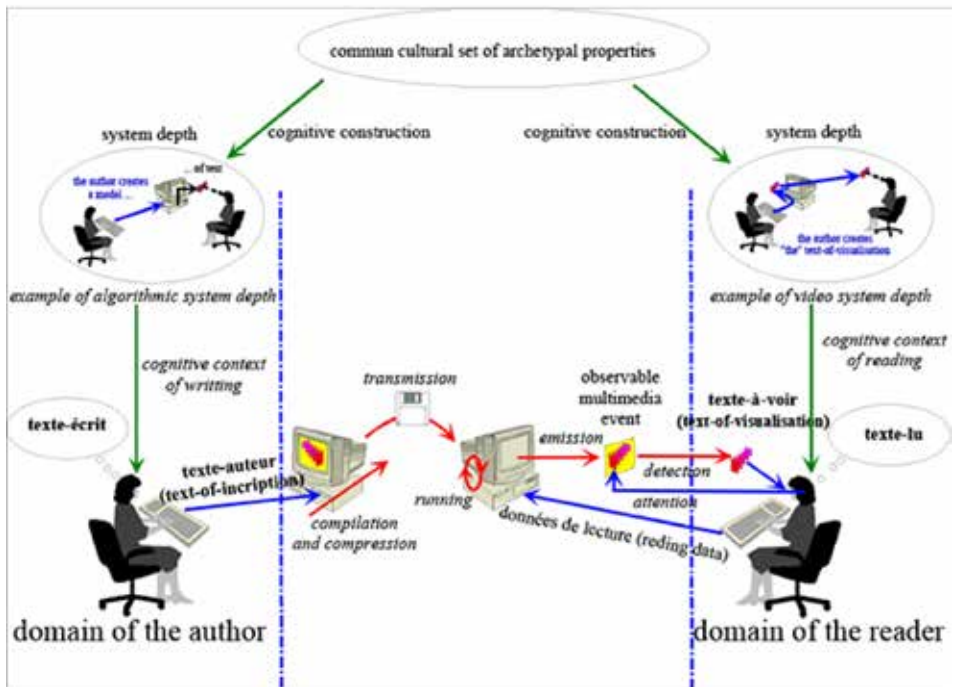


Figure 10: schéma structurel du modèle procédural

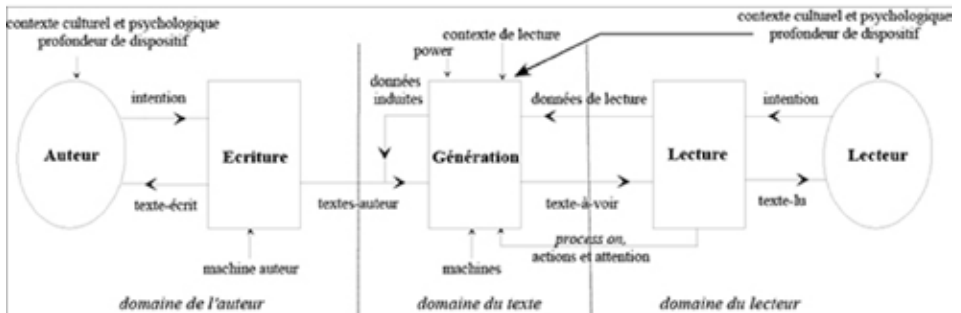


Figure 11: schéma fonctionnel du modèle procédural

Dans ce modèle, le lecteur et l'auteur ne sont pas des personnes mais des rôles, des points de vue portés sur l'œuvre qui reposent sur un accès possible ou interdit aux diverses parties extensives de celle-ci. Notamment, le lecteur n'accède pas à la source, l'auteur n'accède pas au transitoire observable et aucun n'accède au programme en cours d'exécution. Dernièrement le modèle a permis d'ajouter d'autres rôles, notamment des rôles de méta-lecteurs qui sont à cheval sur les trois domaines cités. Un méta-lecteur accède aux diverses parties extensives de l'œuvre ainsi qu'aux rapports qui s'établissent entre le lecteur et l'œuvre.

Bien qu'il n'ait pas été introduit dans cette optique, le rôle de méta-lecteur permet tout naturellement de traiter des rôles d'analyste : un analyste (critique ou chercheur) est un méta-lecteur doté d'un objectif particulier distinct de celui des autres acteurs. Le modèle procédural définit les rôles auteur et lecteur par des procédures de construction de représentations mentales. Est dans un rôle auteur celui que se construit une représentation mentale de l'œuvre dénommée texte-écrit. Est lecteur celui qui construit une représentation de l'œuvre, dénommée texte-lu, par un processus d'interprétation percepto-cognitif du transitoire observable. Quelque soit la situation (œuvre collective, participative, individuelle, performance...) il n'y a donc jamais d'ambiguïté sur les fonctions ni de mélange des genres : un lecteur n'est jamais auteur et un auteur n'est jamais lecteur. Le modèle ne manipule donc pas de concept d'écrit-lecteur. On pourrait toutefois le reconstruire en constatant qu'un individu peut passer dans la pratique très rapidement d'un rôle à un autre. Par extension, on peut définir un rôle d'analyste comme un méta-lecteur qui poursuit l'objectif d'acquérir une connaissance scientifique précise sur un élément du domaine général couvert par le modèle de communication. Cet objectif est réalisé grâce à une observation des rapports entre les autres acteurs et l'œuvre. Dans le processus d'extension d'acteur que constitue son activité, l'analyste crée ainsi avec l'œuvre un rapport qui la met en relation interne (Figure 6) avec tout un ensemble d'autres productions ou événements qui dépendent de son objectif (œuvres d'avant-gardes pour une analyse historique, œuvres littéraires numériques ou non pour une analyse littéraire...) et qui constituent le corpus implicite ou explicite de son domaine d'expertise, alors que le lecteur la met en relation avec l'ensemble de sa culture personnelle et à des fins de construction qui lui sont propres et ne visent pas un savoir (ce qui n'exclut pas la possibilité d'une lecture savante) alors que l'auteur met l'œuvre en rapport avec d'autres productions au sein d'une démarche créative.

Ce modèle fournit un cadre permettant de construire une ontologie du domaine qui repose sur le contact entre individus, les individus étant ici l'œuvre et chacun des rôles repérés. On peut donc utiliser le schéma ensembliste du contact (Figure 2) pour schématiser très simplement les positions respectives des divers acteurs dans le processus de communication par l'œuvre. Le modèle permet également d'introduire d'autres rôles en définissant les objectifs qui identifient les points de vue et les relations de contact qui s'établissent. La Figure 12 offre l'exemple d'une représentation ensembliste d'une communication réalisée par analyse d'une performance. Le performeur est celui qui met l'œuvre dans un rapport de spectacle à destination de lecteurs. Il crée ainsi un rapport qui passe, à travers l'œuvre, dans un rapport à autrui : l'œuvre est placée en situation d'association et non plus d'objet.

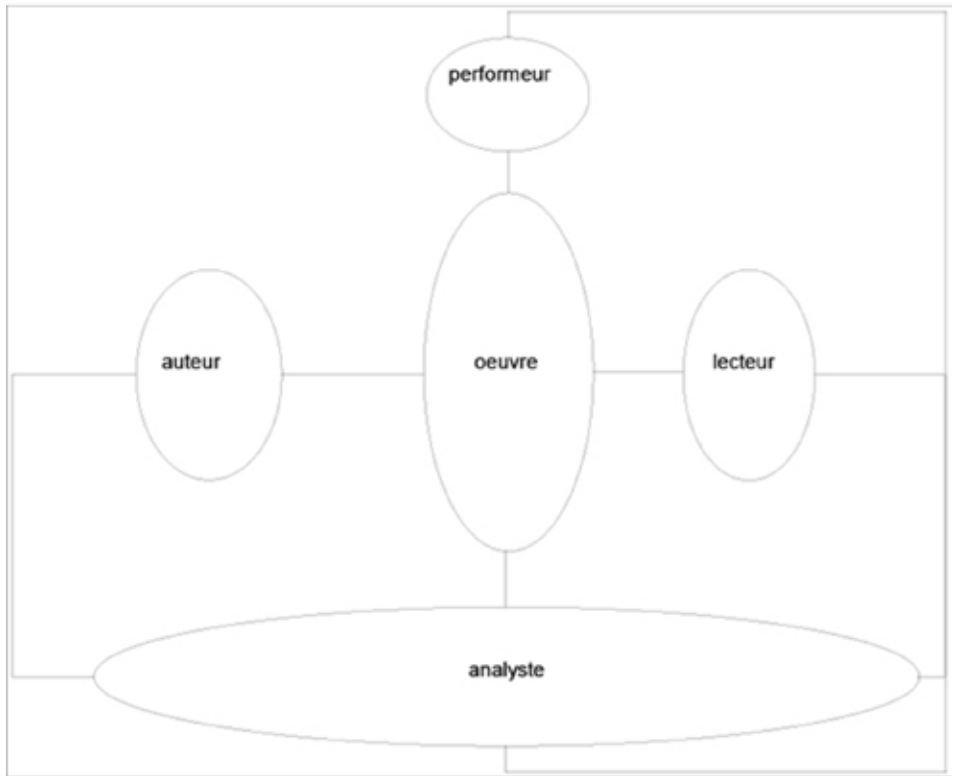


Figure 12: relations de contact entre individus du domaine

Chaque acteur fournit un «point de vue» sur l'œuvre. On peut imaginer autant d'acteurs que nécessaire, notamment autant de types d'analystes (sémioticien, historien, sociologue, technicien, critique, analyste littéraire...).

L'œuvre elle-même connaît un phénomène d'extension lors de la lecture. Elle peut s'approprier les parties extensives du lecteur et transformer la lecture en une activité signifiante interne à l'œuvre. Lors de ce contact entre le transitoire observable et le lecteur, le transitoire observable et la partie extensive du lecteur qui sont en contact perdent leur identité pour constituer une nouvelle partie extensive de l'œuvre à laquelle le modèle n'a d'ailleurs pas donné de nom. Cette partie peut être observée par un méta-lecteur qui lui donne sens par la méta-lecture<sup>5</sup> ou par

<sup>5</sup> La métalecture consiste à observer la lecture d'autrui tout en ayant une connaissance sur le texte-auteur qui provient d'un autre canal que l'œuvre.

le lecteur lui-même qui lui donne sens par la double lecture<sup>6</sup> et peut être pensée par l'auteur qui lui donne également sens dans le cadre de l'esthétique de la frustration, le plus souvent le sens d'un signe iconique ou d'une activité métaphorique.

#### 4. Ontologie de l'œuvre

L'essence de l'œuvre est l'ensemble des concepts intrinsèques à cette œuvre mais elle ne nous est pas directement atteignable. On ne peut que la pointer à l'aide d'indices qui sont d'une part le titre de l'œuvre et d'autre part la connaissance qu'on en a et qui constitue l'objet de l'indexation. L'essence de l'œuvre s'incarne dans les objets de l'œuvre qui en sont les parties extensives.

L'œuvre peut connaître des processus d'extension et se manifester par des versions différentes qui peuvent être de natures différentes : étude, version initiale, mises à jour, changement de contexte... Chaque version implémente des propriétés spécifiques de l'œuvre, pas forcément toutes à la fois, dans des objets (fichiers, transitoire observable, programme source, situation de lecture et périphérique dans le cas de performances...). Ces objets sont associés ou non à l'œuvre en des dates et des lieux spécifiques des versions. Chaque version sera donc repérée par un numéro de version (si elle est officialisée) ou un identifiant, une date de création et, s'il est besoin, une durée et éventuellement un ou des lieux de vie (notamment dans le cas de performances).

Les objets qui constituent les versions entretiennent entre eux des relations de dépendance qui définissent une structure ordonnée descriptible selon un modèle E/A. Ils sont chacun dotés d'attributs spécifiques qu'il faudra repérer. Chaque objet, indépendamment de cette hiérarchie, est dans une relation particulière à l'essence de l'œuvre à travers la version (Figure 14). Ces relations sont donc des rapports spécifiques inclus dans la version et peuvent subsister d'une version à une autre.

Dans la représentation choisie (Figure 13), les rapports entre essence et partie extensive sont synthétisés dans la seule définition de la version.

<sup>6</sup> La double lecture consiste pour le lecteur à donner un sens dans l'œuvre à son activité de lecture, à «lire sa lecture». Notons que par la double lecture le lecteur accède à son propre rapport à l'œuvre, c'est-à-dire qu'il peut parler de la construction du texte-à-voir à partir du transitoire observable.

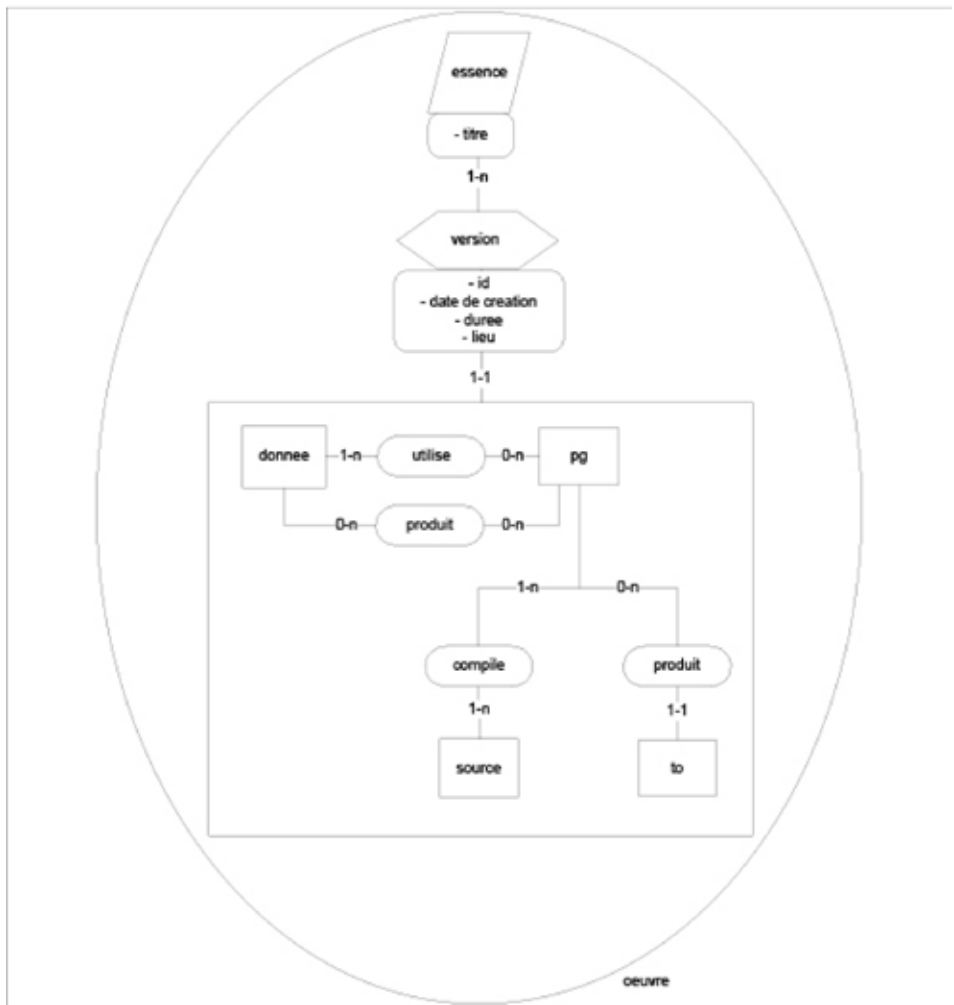


Figure 13: représentation équilibrée d'une œuvre: cas d'une œuvre purement numérique

Dans cette figure, le terme *to* désigne le transitoire observable (phénomène multimédia mesurable produit à l'exécution), le terme "données" désigne n'importe quel fichier de donnée ou champ de Base de donnée, le terme «source» regroupe les fichiers source du programme en langage de programmation auteur (fichiers *swa*, fichier *.dir*, fichier *.html*...) et le terme *pg* regroupe les fichiers exécutables (*.exe*, *.dll*, *.xtra*, *.com*...).

Remarquons que l'exécution du programme crée des relations d'actualisation des données dans le *to* et du *source* dans le *to*. Elle intervient dans le rapport entre ces parties et l'essence de l'œuvre.

Une même version peut se décliner selon plusieurs instances (une instance on-line, une instance performative, une instance sur cédérom, une édition papier...) qui sont autant de variantes de la version. Les instances constituent des mises en contexte de l'œuvre. Elles peuvent différer par leur modalité de réception, leur modalité de diffusion ou par les technologies et programmes mis en œuvre (reprogrammation, simulation) mais réfèrent toutes aux mêmes concepts de l'œuvre, c'est-à-dire qu'elles sont dans le même rapport (la version) à l'essence de l'œuvre. Alors que les versions peuvent développer des concepts différents, même si elles possèdent un jeu conceptuel commun (sinon elles constitueraient d'autres œuvres). Certaines instances peuvent ne laisser percevoir que des concepts spécifiques de l'œuvre. Les diverses instances sont donc complémentaires dans leurs rapports à l'œuvre. Les instances sont donc des sous-rapports à l'œuvre, la version est le rapport construit par l'ensemble de ces sous-rapports.

En toute généralité, l'œuvre peut ainsi se décliner en versions et instances numériques ou non numériques (vidéos, photos exposées en galeries, livres). Les parties extensives numériques peuvent comporter des objets spécifiques non numériques en relation avec le programme, par exemple pour une performance.

Compte tenu de ce que l'essence d'une œuvre n'est pas descriptible autrement que par les connaissances qu'on en a, il est inutile de la schématiser. Elle n'est maintenue sur la Figure 14 que pour repérer les deux instances. Pour gagner en lisibilité, nous n'avons pas reporté sur la Figure 14 les informations de la figure Figure 13.

Les deux figures suivantes schématisent les rapports (Figure 14) et les parties extensives (Figure 15) d'une œuvre constituée d'un programme publié sur le Web (instance A) et présentée en performance de lecture lors d'une soirée (instance B).



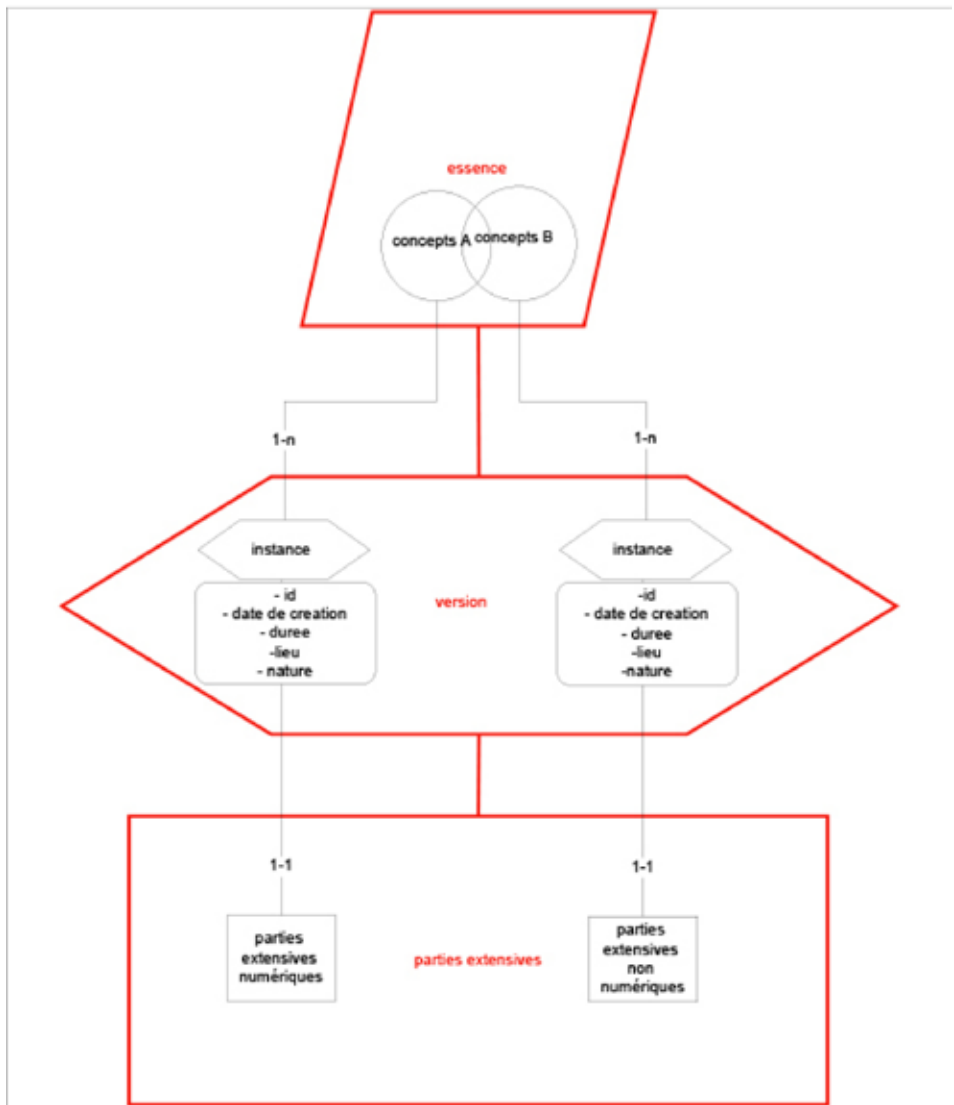


Figure 14: ontologie générale de l'œuvre constituée par une instance numérique (A) et une autre imprimée (B) d'une même version

On peut détailler les parties extensives de chaque instance à l'aide de relations internes entre parties extensives, dans un modèle E/A comme nous l'avons fait sur l'exemple précédent. On obtient alors le diagramme suivant des parties extensives. Notons que chaque partie extensive repérée est un document qu'on peut lui-même subdiviser et décrire comme tous les autres documents (cf partie suivante)

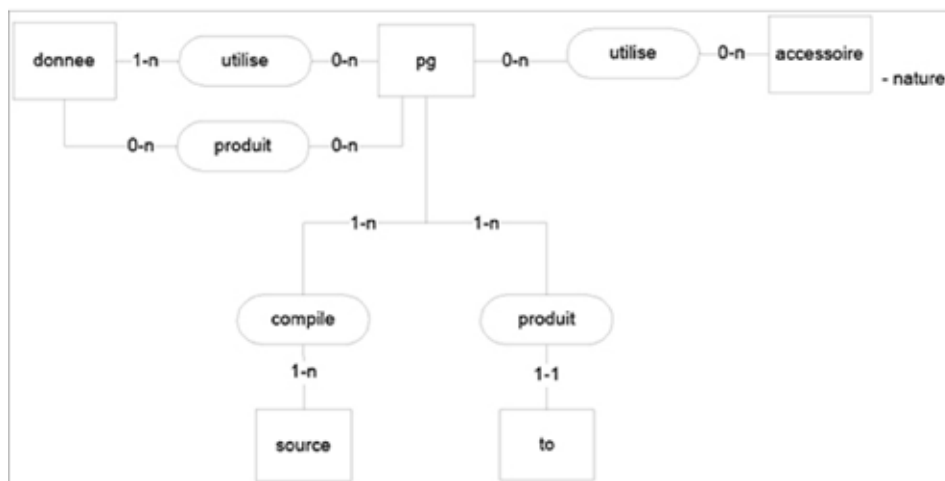


Figure 15: détail des parties extensives de l'œuvre, exemple d'une œuvre purement numérique manipulée en performance

## 5. Ontologie des documents sur l'oeuvre.

Les documents produits par les personnes peuvent être décrits par le même modèle dont l'éléments principal est la section. Une section est une partie extensive d'un acteur issue du contact que cet acteur a eu avec l'œuvre. Un document physique devra donc être décomposé en sections, chacune d'entre elles n'étant attribuable qu'à un seul rôle, c'est-à-dire ne développant qu'un point de vue. La granularité de la décomposition dépend donc de la définition de l'acteur donnée par l'indexeur. On peut par exemple simplement distinguer dans un document académique les sections qui traduisent une activité de lecture de celles qui traduisent une activité d'analyse de quelque nature que ce soit. La situation de communication correspondante à cette indexation est alors celle de la Figure 19. On peut également affiner la décomposition en distinguant les divers types d'analystes.

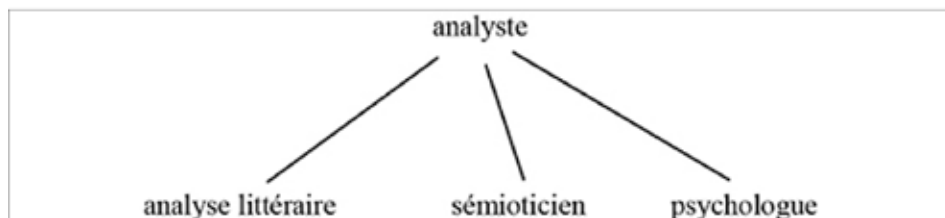


Figure 16: exemple de décomposition hiérarchique de l'acteur analyste

Les concepts développés dans une section appartiennent à l'ontologie de l'acteur. On en déduit que le degré le plus fin de granularité dans l'indexation sera obtenu en définissant des sections qui ne concernent qu'un concept développé par un seul acteur dans la hiérarchie des acteurs.

La structuration des relations entre documents physiques et sections est la même pour toutes les sections de tous les documents de toutes les ontologies d'acteurs. C'est là un point essentiel de l'extensibilité du modèle : il suffit de savoir en créer une pour que le modèle soit extensible, puisque l'extension consiste en la duplication de la structure sur toutes les ontologies d'acteurs. Je propose la structuration suivante qui s'appuie uniquement sur une structure physique du document. Les informations de version de l'ontologie d'acteur sont, comme pour l'œuvre, données sous forme d'attributs du rapport «développe». Elles permettent de déterminer l'ontologie utilisée pour l'indexation par un identifiant (par exemple son nom), une information de date et de lieu de constitution de cette ontologie.

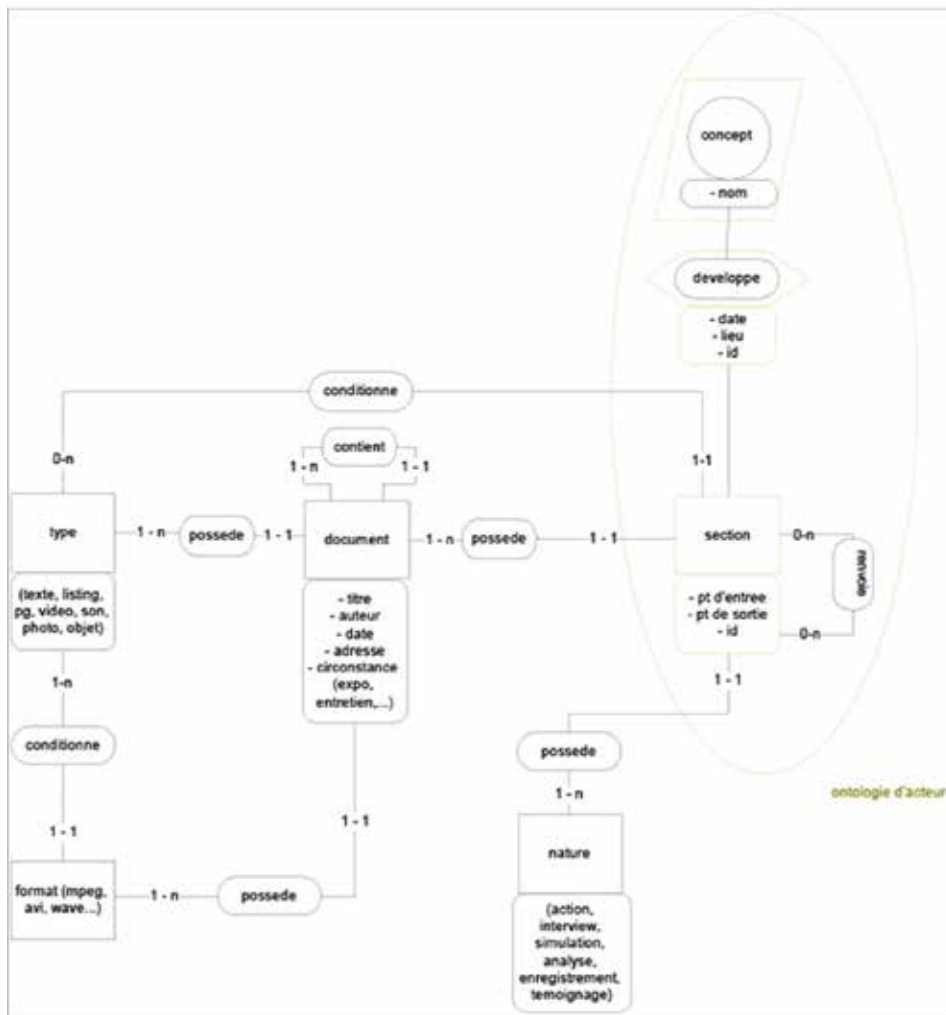


Figure 17: ontologie du document

L'ontologie d'acteur (ou rôle) est construite par la donnée de l'ensemble des concepts et des relations qu'ils entretiennent entre eux. On peut alors transformer cette description en un modèle E/A classique en introduisant explicitement l'entité rôle dans le modèle. Le type d'acteur est : lecteur, auteur..., et son id peut être le nom de l'ontologie d'acteur, c'est pourquoi le rapport «développe» suffit à caractériser l'ontologie d'acteur. On peut alors supprimer la schématisation ensembliste et récupérer les attributs de l'ontologie d'acteur sur une fonction «définit» qui décrit cette ontologie sous forme d'une association. On obtient alors le modèle E/A suivant qui servira de base à la construction de l'ontologie du domaine.

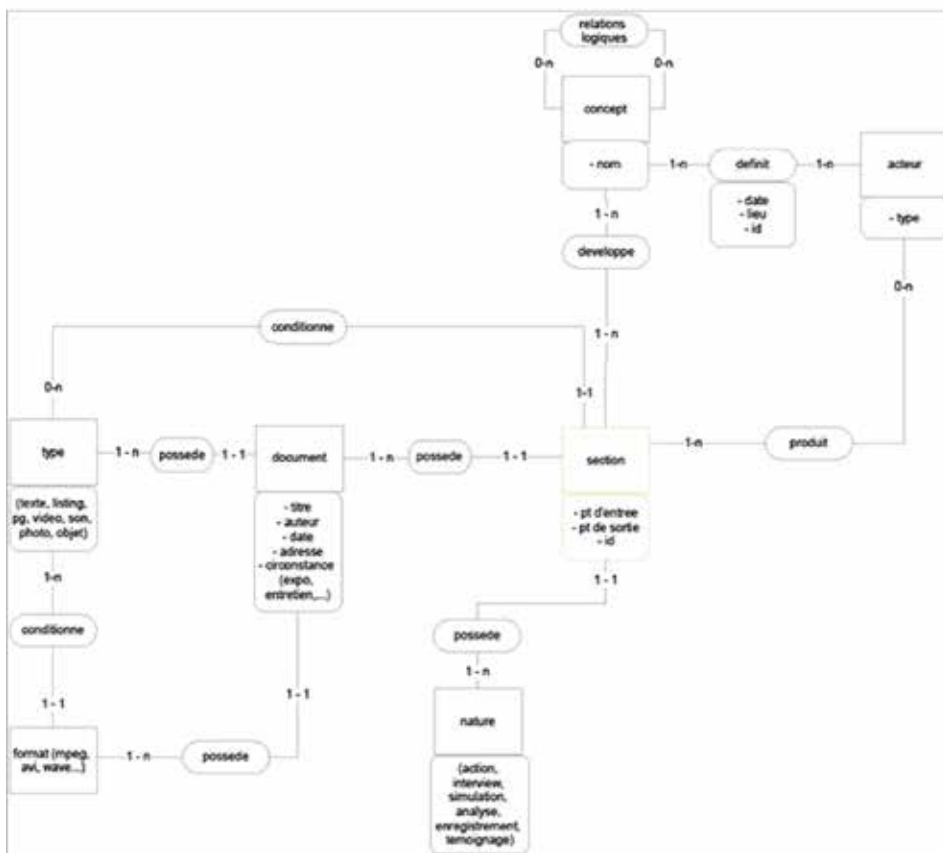


Figure 18: modèle E/A du domaine

Ce modèle doit être complété par la donnée des rapports éventuels entre concepts développés par des acteurs différents, ce qui correspond à l'ontologie de l'acteur de rang supérieur dans la décomposition hiérarchique de la Figure 16.

## 6. Contacts entre l'œuvre et les acteurs

Les acteurs ne vont nous intéresser qu'à travers les sections de documents qu'ils produisent. Ces sections constituent des parties extensives de ces acteurs réalisées à la suite de leur contact avec l'œuvre. Le contact entre l'acteur et l'œuvre est schématisé dans le modèle ontologique par un réseau E/A et le modèle procédural décrit ce réseau en imposant des restrictions. Ainsi, le lecteur ne peut pas accéder au source... Comme nous l'avons vu, le texte-à-voir et le texte-auteur nous constituent des parties extensives de l'acteur lors de son extension par assimila-

tion des parties extensives de l'œuvre qui lui sont accessibles. Ces éléments sont accessibles à un analyste qui observe l'acte créatif ou l'activité de lecture. Ces extensions, dans le formalisme équichronique, n'étant pas représentées par des entités mais par des associations (Figure 7), la représentation équichronique de la situation de communication est un modèle E/A qui présente des associations entre associations. Ainsi, les relations « crée » qui relie l'auteur à l'œuvre en Figure 19 concernent le texte-auteur et la relation « observe » qui s'établit entre le lecteur et le transitoire observable concerne le texte-à-voir. La représentation finale ainsi obtenue (Figure 19) constitue le cadre du modèle ontologique du domaine. Pour des raisons de clarté, j'ai porté en marron clair les relations « analyse » qui lient l'analyste aux autres objets du modèle. Elles sont toutes de cardinalité 0-n / 0-n non portée sur le schéma et, bien sûr, c'est l'analyste qui analyse. Le schéma est construit dans le point de vue de l'œuvre, c'est-à-dire que les parties extensives des acteurs concernées par les contacts ne sont pas représentées. Les sections des documents d'acteurs sont des parties extensives de ces acteurs qui résultent de ces contacts.

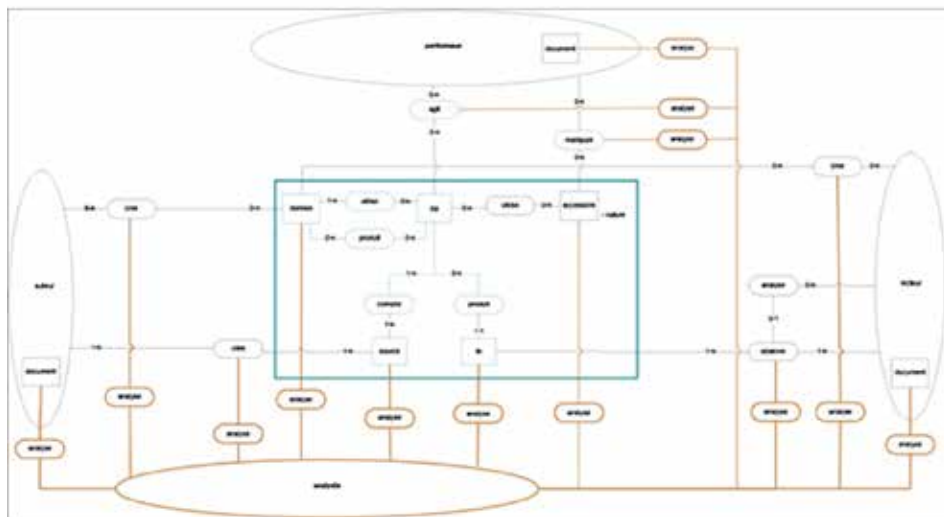


Figure 19: représentation équichronique du contact des acteurs aux objets du domaine

Les sections de documents produites par un acteur sont le résultat du contact entre l'acteur et des parties extensives de l'œuvre, celles dont elles traitent. L'analyse des concepts présents dans les sections des documents fournit une connaissance de l'œuvre, Celle-ci peut être une connaissance inadéquate, un affect, ou une connaissance adéquate décrite par des concepts (A, B, C dans la Figure 20) décrivant l'ensemble (rapports+essence) de l'œuvre tel qu'il est perçu dans le point de vue de l'acteur. Rappelons que cet ensemble demeure étranger aux documents. Le contact entre un acteur et l'œuvre est alors décrit par le schéma suivant qui va nous permettre de documenter dans l'indexation la relation entre la section et l'œuvre. Notons

qu'une section de document peut traiter d'une partie particulière d'une œuvre comme de l'ensemble des parties de toutes les œuvres si la section du document embrasse tout le champ de la poésie numérique. C'est ce que décrit la cardinalité de l'association «traite» (Figure 20).

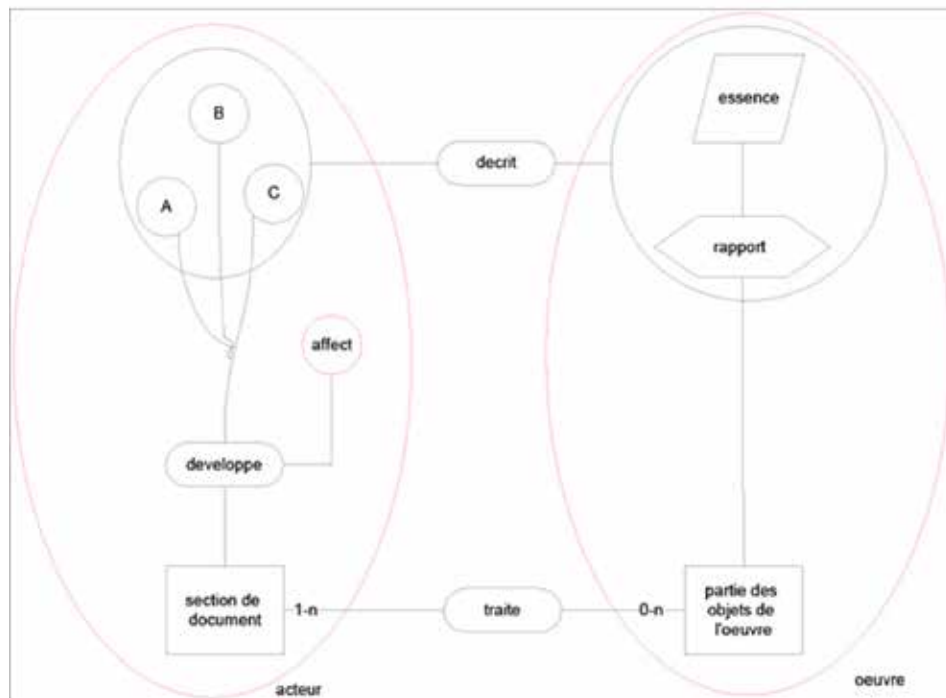


Figure 20: relation entre l'ontologie des documents et l'ontologie de l'œuvre

## 7. Processus d'indexation et visualisation des connaissances.

L'indexation d'un document revient à définir les différentes sections, à les associer à un acteur, à repérer les parties de l'œuvre dont elles traitent et les concepts et affects de l'acteur qu'elle énonce. Le schéma de la Figure 19 ou une partie de ce schéma peut alors servir d'interface graphique pour déterminer les acteurs et relations à l'œuvre concernées. L'indexation n'est pas une opération neutre qui serait exempte de tout point de vue (principe du tiers inclus oblige). Elle doit donc également être prise en compte dans le modèle et l'outil. L'indexation résulte d'un contact entre un document et un acteur indexeur. Son schéma est donc similaire à celui de la Figure 7. L'indexeur est un rôle et il convient d'indiquer le nom de la personne physique qui le tient. Le schéma de l'indexation est alors le suivant:

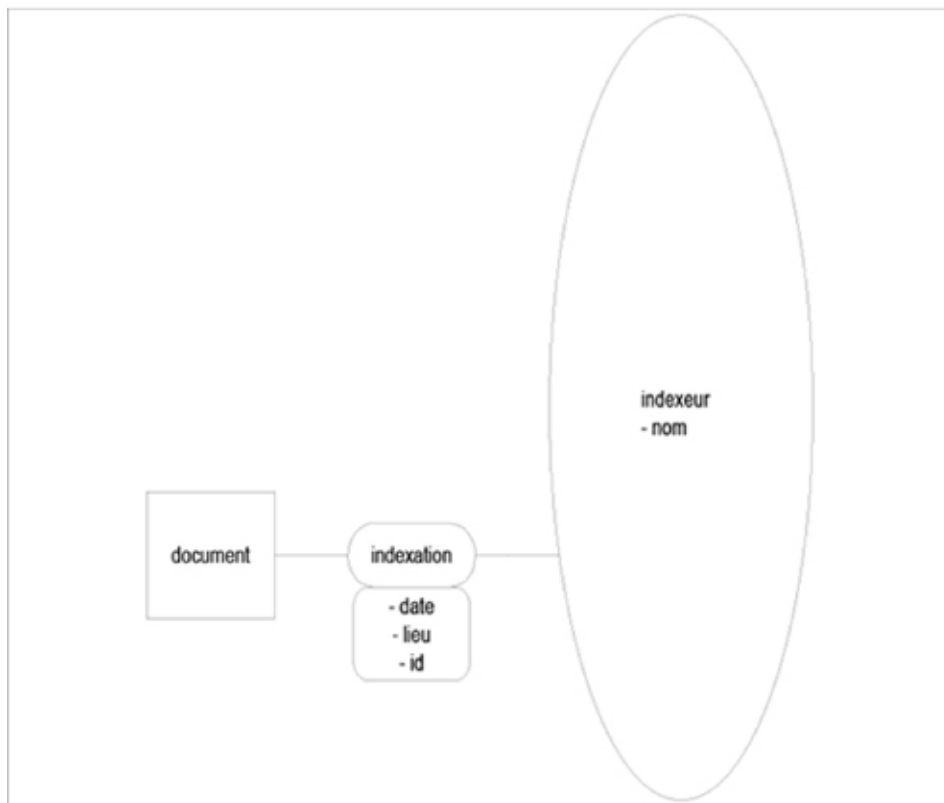


Figure 21: schéma de l'indexation

Lorsque l'indexation est achevée, on peut visualiser les connaissances portées par les documents sous forme de diagrammes de Vienn (Figure 22 et Figure 23). Dans ces schémas, les ontologies d'acteurs sont repérées par des ovales noirs, les concepts développés dans les sections documents par des ronds ou ovales verts et les sections des documents par des points bleus. Deux visualisations sont possibles : une visualisation surplombante qui permet de mieux visualiser les proximités sémantiques entre points de vue et une visualisation latérale qui permet de mieux visualiser les étendues conceptuelles de chaque section. Supposons donné un ensemble de sections indexées comme suit :

**Indexation :**

- Nom indexeur : Bootz
- Id : test
- Date : 2008
- Lieu : at home



**Document :**

- Type : livre
- Format : imprimé
- Titre : Matières textuelles sur support numérique
- Auteur : Alexandra Saemmer
- Date : 2007
- Lieu : Publications de l'Université de Saint-Etienne
- Circonstance : VOID

**Sections :**

- Id : section 1:
- entrée : «Pierre Lévy» (p. 125)
- sortie : «préceptes de l'ouliipo»
- nature : analyse
- Document de l'oeuvre traité : to
- Ontologie d'acteur : analyse littéraire
- Date : 2008
- Lieu : at home
- Id : test
- concepts traités : potentiel-virtuel , ouliipo
- Id : section 2:
- document : ouvrage matières textuelles sur support numérique
- Entrée : «ce poème est transitoire observable»
- sortie : «parties intégrantes»
- nature : analyse
- Document de l'oeuvre traité : to
- ontologie : analyse sémiotique du numérique
- Date : 2008
- Lieu : at home
- Id : test
- concepts traités : transitoire observable – programme
- Id : section 3:
- document : ouvrage matières textuelles sur

- support numérique
- Entrée : «voilà ce qui a été perçu» (p. 126)
- sortie : «qu'une chimère»
- nature : analyse
- Document de l'oeuvre traité : to
- ontologie : lecteur
- Date : 2008
- Lieu : at home
- Id : test
- concepts traités : transitoire observable – programme, animation
- Id : section 4:
- document : ouvrage matières textuelles sur support numérique
- Entrée : «comme un palindrome»
- sortie : «s'efface par animation»
- nature : analyse
- Document de l'oeuvre traité : to
- ontologie : analyse littéraire
- Date : 2008
- Lieu : at home
- Id : test
- concept traité : figure rhétorique du palindrome
- Id : section 7
- document : ouvrage matières textuelles sur support numérique
- Entrée : «la série des U estbasée»(p. 127)
- Sortie : «déduites de l'observation qui assurait cette cohérence» (p. 128)
- nature : analyse
- Document de l'oeuvre traité : source
- ontologie : auteur
- Date : 2008
- Lieu : at home
- Id : test
- concept traité : génération adaptative

La visualisation des connaissances présentées dans cet ensemble pourrait prendre les formes suivantes qui s’obtiennent par rotation dans l’espace 3D et dans lesquelles les identités de concepts entre rôles d’acteurs ont été fusionnées par un diagramme de Venn de type «ou». Les ontologies d’acteurs sont repérées par des ovales noirs, les concepts développés dans les documents par des ovales verts et les sections par des points bleus.

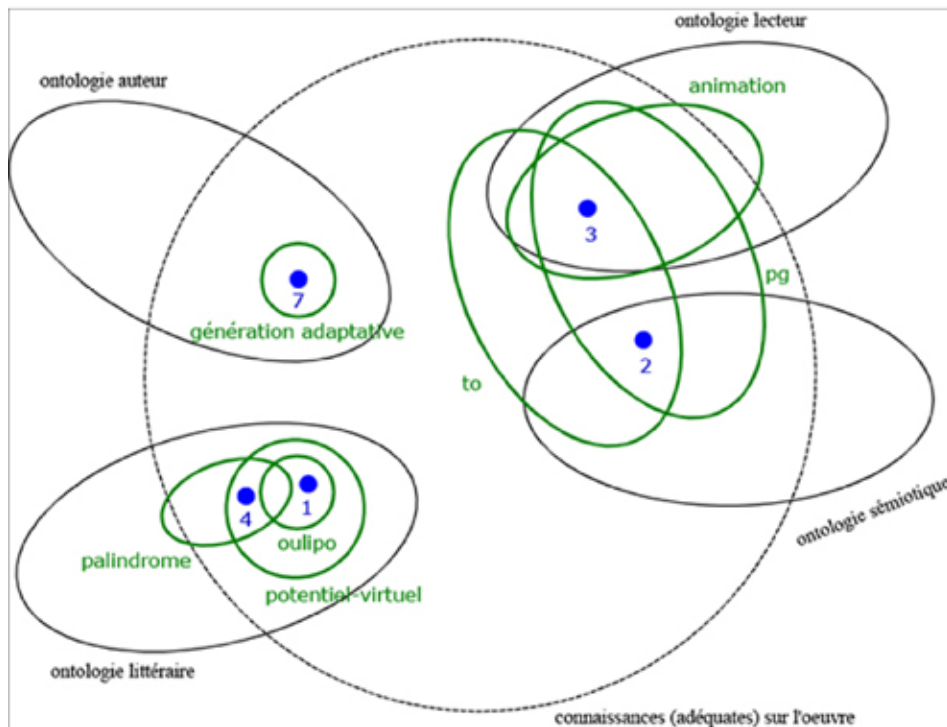


Figure 22: visualisation surplombante des connaissances

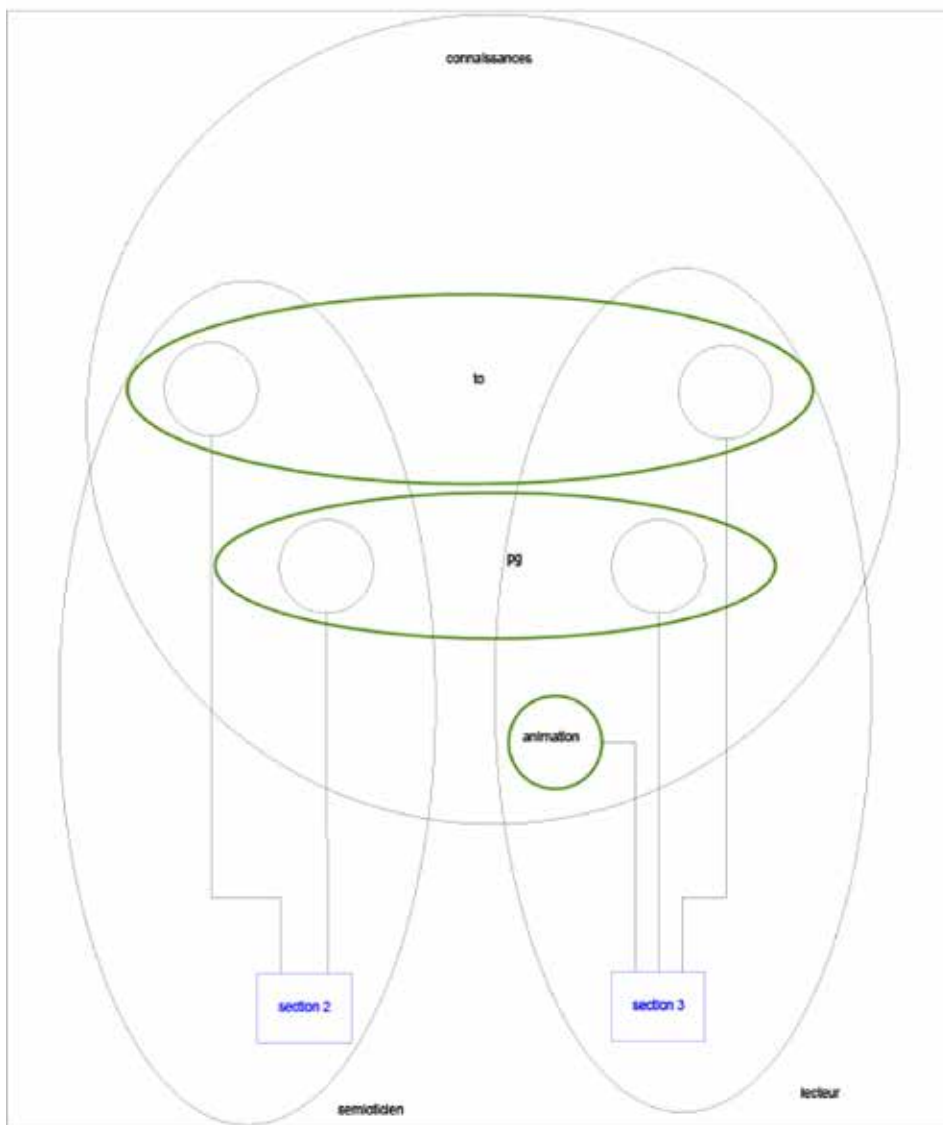


Figure 23: vue «latérale» des connaissances

## 8. conclusion

L'outil d'indexation sera réalisé en technologies Web opensource alliant XUL et SVG. Ce projet d'indexation largement ouvert réclame pour être mené à bien une large coopération multidisciplinaire qui pourrait être menée dans le cadre d'un projet européen.