

Une ontologie pour archiver et donner accès aux œuvres numériques littéraires

Yan Rucar, Jean-Gabriel Ganascia

Résumé : nous avons conçu une ontologie destinée à indexer le corpus des œuvres numériques littéraires. Cette ontologie prend en compte les dimensions esthétiques et matérielles des œuvres de sorte que l'analyse critique et la production d'archives se croisent dans la tâche de conservation.

Depuis le premier générateur de textes conçu par Theo Lutz en 1959 [1], la littérature créée sur ordinateur n'a cessé de produire des œuvres marquées par le caractère éphémère des logiciels et des ordinateurs. Sitôt le support obsolète, l'objet se désintègre. Les traces laissées par l'œuvre sont le plus souvent frustrantes, à la façon d'une fenêtre à peine ouverte sur un paysage riche en textures, et bientôt close. Les textes critiques comprennent un bref descriptif de l'œuvre, augmenté parfois de quelques captures d'écran. Le lien entre les mots et les images, le dynamisme de dispositifs en constant mouvement, le fonctionnement du programme sont perdus. Jamais les exégèses ne font figurer le code-source, la vie de l'organisme fébrile de l'œuvre numérique est dès sa désagrégation insituable. Les traces laissées par les œuvres numériques sont aussi frustrantes que ces chefs d'œuvres du cinéma muet aujourd'hui seulement disponibles sous forme d'une poignée de photographies, d'un scénario et d'une composition musicale. Ce qui animait et faisait le film manque, et c'est du coup le cœur du dispositif qui n'est plus qu'une vaste lacune, un centre absent. Le parallèle avec le cinéma est pertinent dans la mesure où l'œuvre numérique est mobile, caractérisée par la transformation. La littérature électronique est déjà historiquement riche, tout en étant trouée d'oubli. La bibliothèque est constituée pour une bonne partie d'ouvrages fantômes.

L'obsolescence des supports n'est pas le seul problème d'accès, l'incompatibilité des équipements est également un obstacle important à la lecture des œuvres. Ainsi les chefs d'œuvres de la fiction hypertextuelle *Afternoon, a story* [2] de Michael Joyce et *Patchwork girl* [3] de Shelley Jackson ne sont lisibles que sur des ordinateurs Mac. D'autre part, les différences entre ordinateurs peuvent complètement changer la perspective sur l'œuvre, en l'éloignant de l'intentionnalité de l'auteur.

Dès lors se fait jour la nécessité de concevoir un point de référence à ces œuvres changeantes et éphémères, un site électronique qui par-delà l'obsolescence prévisible des logiciels et l'absence d'interopérabilité des systèmes, puisse proposer des traces pertinentes de ces objets fuyants. En somme, la tâche à accomplir revient à archiver une œuvre existante, à la concevoir aujourd'hui même comme un présent révolu, à la considérer dans sa mortalité potentielle. Quand l'archivage est accompli après la désagrégation des œuvres, leurs vestiges consistent en des descriptions lacunaires dans des articles critiques, des images qui sont des fragments des objets numériques, parfois quelques mots sur les programmes. Ces traces manquent de pertinence car elles sont le fruit d'actes isolés, dispersés, et tout à fait fortuits. Des aspects essentiels des œuvres ne sont pas couverts par ces archives parcellaires. En créant l'archivage a priori paradoxal d'une œuvre encore accessible, on se rend maître des traces en cessant d'être soumis aux effets du hasard, on trie les faits saillants d'une esthétique au détriment de propriétés secondaires : « Plutôt que l'idée de conserver certains aspects, il faut se demander quels aspects abandonner » [4]. Quand l'œuvre disparaîtra, ses traits définitoires seront encore consultables dans le site d'archives.

Un tel site aura plusieurs finalités :

1. constituer une mémoire de l'objet numérique
2. disposer des substrats pouvant servir de supports pédagogiques
3. proposer une pluralité de points de vue sur une œuvre protéiforme
4. contribuer à un discours critique visant à la définition d'un genre littéraire

En effet, les archives pourront être utilisées dans un cursus universitaire afin de pallier aux difficultés d'accès aux œuvres, et leur diversité corrigera les distorsions d'ordinateur à ordinateur, qui tendraient à une perspective faussée sur l'objet d'étude.

Dans cette perspective nous avons conçu une ontologie pour indexer les œuvres à partir de leurs différentes caractéristiques, qui constituent autant de points d'entrées aux œuvres littéraires numériques. L'ontologie a été construite à l'aide de l'éditeur Protégé [5] et elle existe dans le format OWL. La photographie d'écran qui suit en donne une image partielle.

The screenshot shows a web browser window with the URL `http://www.semanticweb.org/jean-gabriel/ontologies/2016/9/untitled-ontology-18`. The browser tabs include 'Active Ontology', 'Entities', 'Individuals by class', 'OWLviz', and 'DL Query'. The main content area is divided into several panels:

- Class hierarchy:** Shows a tree structure starting from `owl:Thing`. It includes categories like `Auteurs`, `Concepts` (with sub-classes like `Critique_politique_de_discours`, `Immersion_dans_le_langage`, etc.), `Fonctionnement`, `Genre_littéraire` (with sub-classes like `Animation_textuelle`, `Générateur_de_textes`, etc.), `Mode_de_conservation`, `Oeuvres`, and `Propriétés_des_oeuvres`. The class `Poésie_visuelle` is highlighted.
- Annotations: Poésie_visuelle:** A panel for viewing annotations on the selected class.
- Description: Poésie_visuelle:** Shows relationships for the selected class:
 - Equivalent To:** A plus sign (+) to expand.
 - SubClass Of:** Lists `Genre_littéraire` and `oeuvre some denepas` (repeated twice). Each entry has control icons: a question mark (?), an at-sign (@), an 'x' in a circle, and a circle with a dot.
 - General class axioms:** A plus sign (+) to expand.
 - SubClass Of (Anonymous Ancestor):** A plus sign (+) to expand.
 - Instances:** A plus sign (+) to expand.
 - Target for Key:** A plus sign (+) to expand.

At the bottom of the interface, a status bar reads: "No Reasoner set. Select a reasoner from the Reasoner menu" and has a checked checkbox for "Show Inferences".

Cette ontologie comporte des descripteurs précisant les concepts de la littérature numérique, les fonctionnements des programmes informatiques associés, les genres littéraires, les modes de conservation possibles, et les propriétés multi-médiatiques des œuvres. Les descripteurs font à la fois état de l'esthétique et de la matérialité des œuvres. Le versant esthétique est représenté par les catégories des concepts et des genres littéraires, qui tentent de circonscrire l'intentionnalité des dispositifs. Le concept est le moteur, la justification de l'entité formelle. Bien que les descripteurs soient nombreux (voir tableau 1), on peut dégager trois familles conceptuelles, allant du complet abandon

aux tourbillons de l'écran à la recherche du contrôle sur le texte numérique. La contemplation de l'action de la machine, la fascination technologique, représente le premier ensemble conceptuel, au sein duquel le spectateur suit du regard les découpages textuels opérés par le générateur. Le descripteur « immersion dans le langage » décrit bien une seconde tendance visant à entourer le lecteur/spectateur/internaute de mots à peine visibles, qui vont par exemple se télescoper, un mouvement dont l'effet est de faire perdre à l'utilisateur une vision surplombante par rapport au langage pour faire de cet instrument de sens un facteur de désorientation. La tendance inverse, celle de la recherche d'un contrôle, est bien résumée par le concept de littérature ergodique forgé par Espen Aarseth [6]. La compréhension d'un mécanisme complexe permet l'accès au texte. Cette dernière modalité correspond dans le tableau des genres littéraires au jeu vidéo (voir tableau 2). Les genres littéraires peuvent être classés en quatre regroupements : la poésie visuelle, l'animation, le générateur de textes et le net-art. La poésie visuelle et l'animation se fondent sur des effets immersifs dans le langage, l'interactivité accroît cette dimension. Le générateur de textes délègue à la machine la production littéraire, et le spectateur du net-art s'abandonne à un tissage des flux mis en scène par le programme.

Les descripteurs des fonctionnements (tableau 3), des modes de conservation (tableau 4) et des propriétés multi-médiatiques des œuvres (tableau 5) quittent le champ conceptuel pour s'intéresser à la matérialité des écrans. Les fonctionnements (tableau 3) sont génératifs (le programme produit du texte), interactifs, ou bien une séquence est déroulée (film vidéo, logiciel cartographique), enfin l'ordinateur s'ouvre aux réseaux (liens avec des journaux en ligne). Cinq principaux types de documents sont conservés (tableau 4) : le texte, l'image, le son, le parcours sous forme d'enregistrement, le programme (code-source). La dimension textuelle des œuvres (tableau 5) s'articule autour des notions d'appropriation et de transformation, de fragmentation, de couplage entre le signifié et le signifiant graphique. La dimension sonore peut être centrale ou bien un arrière-plan, les descripteurs « musique liée au texte » et « reconnaissance vocale du texte » désignent les œuvres électroniques régulées par ou régulant des fichiers sonores. La dimension visuelle peut être inscrite dans l'apparence du signifiant graphique du texte, peut être autonome mais liée sous forme d'images évoluant au rythme des mots, ou bien l'iconicité est indépendante du langage.

Tableau 1. Descripteurs de concepts

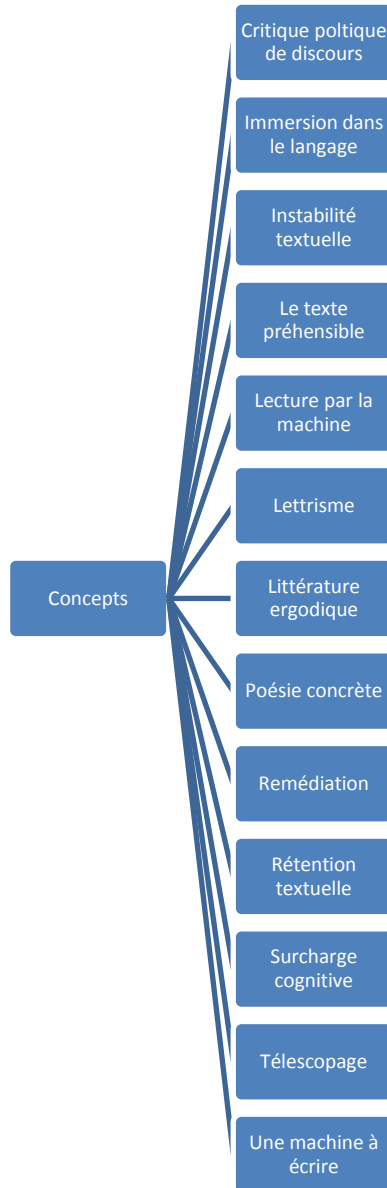


Tableau 2. Descripteurs des genres littéraires

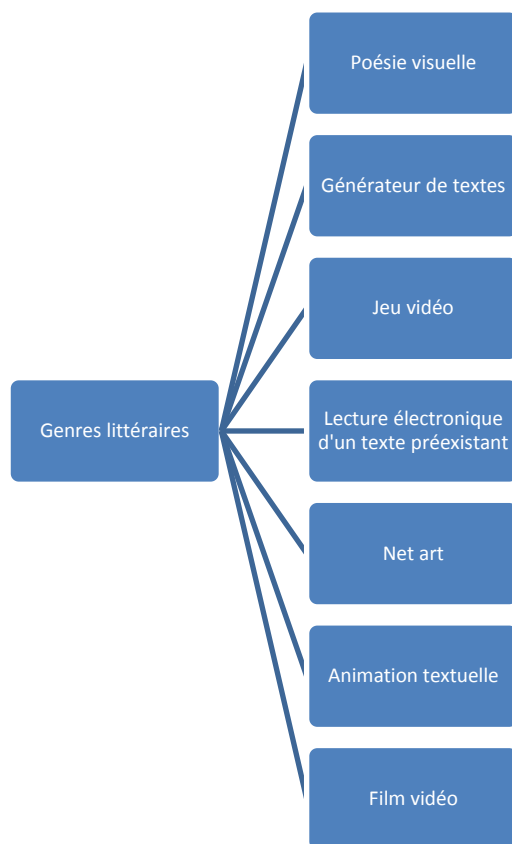


Tableau 3. Descripteurs des fonctionnements

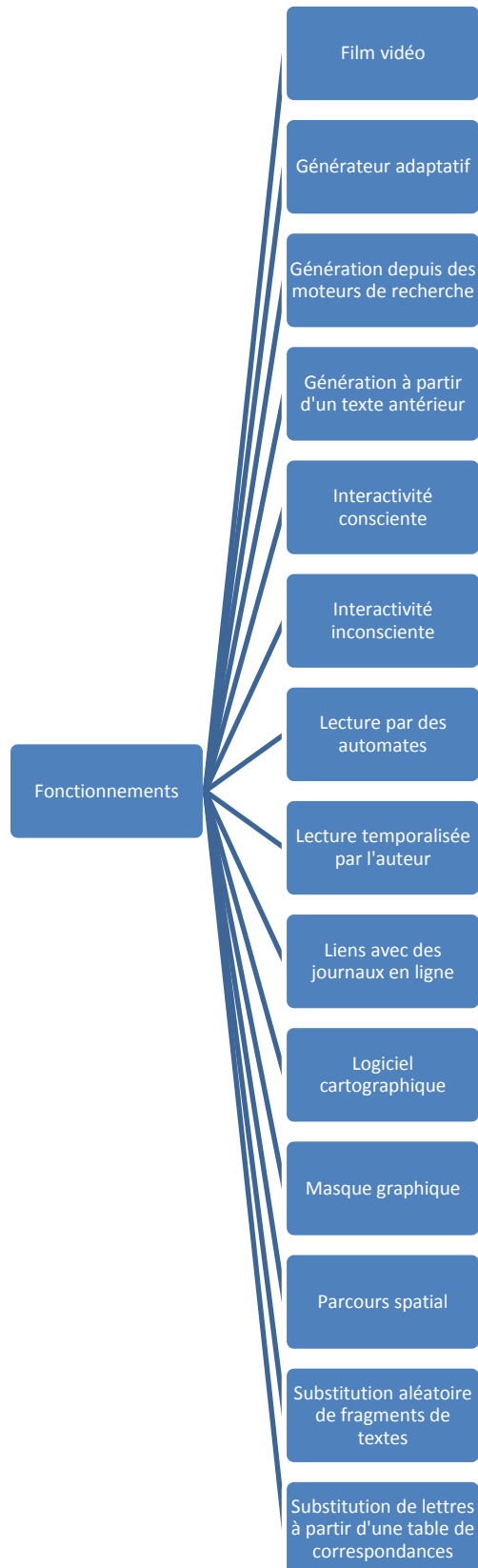


Tableau 4. Descripteurs des modes de conservation

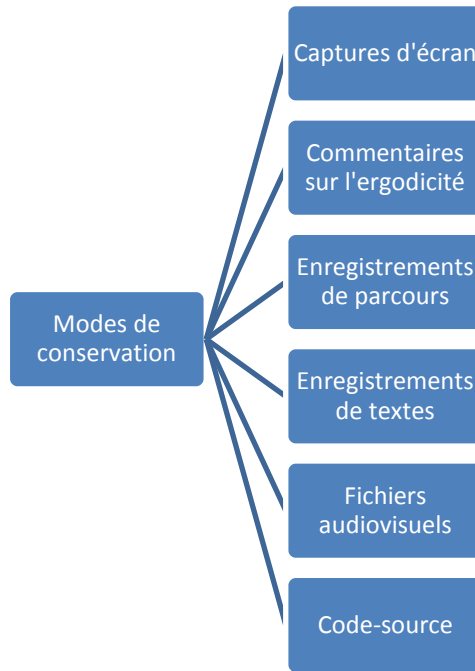
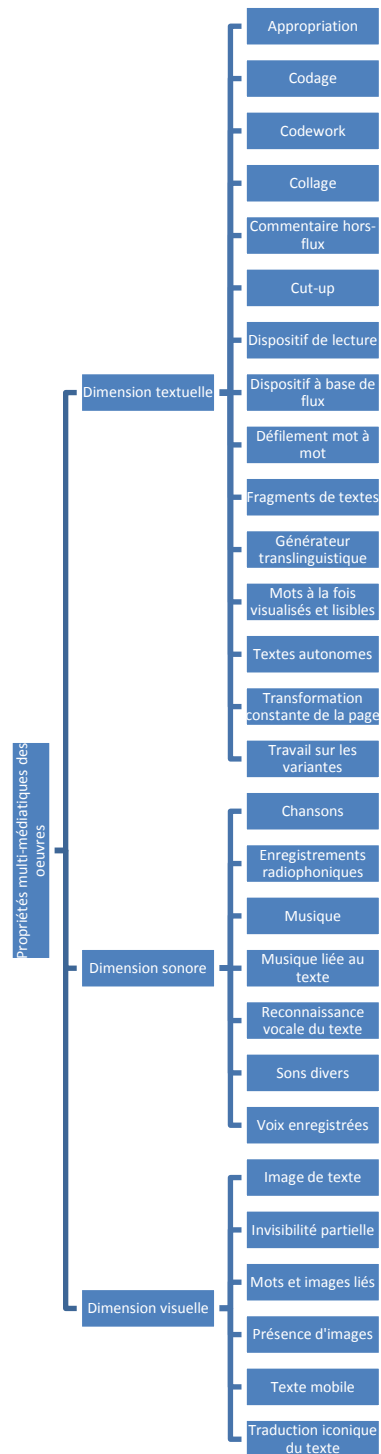


Tableau 5. Descripteurs des propriétés multi-médiatiques des œuvres



Cette ontologie est le premier pas vers la création de l'interface d'un site d'archives, dans lequel coexisteront des substrats, tels que décrits dans le tableau 4, et des points de vue critiques. Alexandra Saemmer souligne cette nécessaire mise en perspective : « La capture n'est pas suffisante, mais c'est une trace, un document, valorisable par une contextualisation » [7]. Parallèlement à l'étude critique des œuvres, qui situera la pertinence des documents, une recherche de toutes les expériences lectorales, spectatorales, perceptuelles, interactives, sera menée à travers le web. En raison des différences d'ordinateurs, il est nécessaire de puiser dans les nombreux blogs et forums de discussion toutes les descriptions de rencontres et d'interactions avec les œuvres numériques étudiées. En glanant et en organisant ces expériences distinctes voire divergentes, on rend compte de l'œuvre fuyante et protéiforme qu'est l'œuvre numérique, et on sauve de l'oubli des témoignages reposant sur des supports éminemment friables. Les postulats de ces œuvres éphémères dépendant de programmes, seront rendu pérennes par des documents figés qui seront autant d'artefacts de type muséal.

Références bibliographiques

[1] Theo Lutz, 1959. Cette œuvre fondatrice a été reprogrammée et est disponible en ligne : http://auer.netzliteratur.net/0_lutz/lutz_original.html La conservation du programme du générateur originel a permis sa remise en fonction.

[2] Joyce, Michael. *Afternoon, a story*, Watertown, US: Eastgate systems, 1987

[3] Jackson, Shelley. *Patchwork girl*, Watertown, US: Eastgate systems, 1995

[4] Rouffineau, Gilles, « Le plaisir du paratexte », in Pamal, preservation & art – media archeology lab et Labex Arts-H2H. *Cette pièce est en cours de maintenance. Merci de votre compréhension – Préservation des écritures numériques*, Journée d'étude, Paris, 27 octobre 2016

[5] Kapoor, B., Sharma, S., A comparative study ontology building tools for semantic web applications. *International journal of Web & Semantic Technology (IJWesT)*, 1(3), 2010

[6] Espen J. Aarseth, *Cybertext : perspectives on ergodic literature*, Baltimore : Johns Hopkins University Press, 1997

[7] Saemmer, Alexandra, « Conclusion », in Pamal, preservation & art – media archeology lab et Labex Arts-H2H. *Cette pièce est en cours de maintenance. Merci de votre compréhension – Préservation des écritures numériques*, Journée d'étude, Paris, 27 octobre 2016