

AKpédia - Annexe

Anna K. mythologies personnelles

Anna K, mythologies personnelles, est un projet de création littéraire qui s'écrit en deux endroits : une fiction (publication) et une création en ligne (diffusion web). Ces deux espaces travaillent une même matière - une collection de data autour de la vie et l'œuvre de la romancière britannique Anna Kavan - pour une mise en récit singulière entre langage naturel et langage de programmation.

Le projet a reçu le soutien du Centre national du livre. Ce crédit de préparation a permis à l'auteure d'effectuer en 2013 une série de rencontres auprès de personnes-clés (éditeur, agent littéraire, journaliste, archiviste, blogger, écrivain, artiste, chercheur) ainsi que des visites de terrain dans des institutions en possession d'archives sur Anna Kavan : Anna Kavan Society (Londres), Bethlem Royal Archives & Museum (Londres), National Library of Wales (Aberystwyth), McFarlin Library (Tulsa), Harry Ransom Research Center (Austin).

Anna K, mythologies personnelles, a reçu une aide au développement du Centre national du cinéma (DICRÉAM - Dispositif pour la Création Artistique Multimédia). En 2013, le projet a été accueilli en résidence à Nantes (PiNG), Bruxelles (Constant/Variable) et Villeneuve-lès-Avignon (La Chartreuse).

En 2014, le projet entre dans une nouvelle phase de production et diffusion, porté par les éditions HYX et livrera trois propositions intimement liées :

- une œuvre en ligne : création d'une base de données narrative entre web sémantique, web-folklore, data-visualisation et archivage actif (automne 2014)

- un atelier d'édition & écriture collaborative : wiki-sprint autour des pages multilingues d'Anna Kavan sur Wikipedia - recherche Semanticpedia (automne 2014)

- une fiction : « Anna K, mythologies personnelles » (titre provisoire) à paraître aux éditions HYX (printemps 2015)

Intentions

Écriture fictionnelle hybride : texte, code, data

L'œuvre de Kavan reste encore méconnue, pour le moins peu diffusée et critiquée en France. Pourtant, sa trajectoire et ses écrits, son style autobiographique et surréaliste, représentent moins un héritage littéraire à démontrer qu'une passionnante œuvre-puzzle à recomposer.

L'énigme Anna Kavan et son style singulier, fêlé, entre folie, fragilité, obsession et réparation sont au cœur de l'écriture fictionnelle du projet *Anna K. mythologies personnelles*.

Et parce que la vie littéraire bouge à l'ère de l'Internet : auteurs, journalistes, bloggers, chercheurs, critiques, tous pratiquent de nouveaux espaces de publication qui bouleversent les conditions de production et de réception de la création littéraire, l'intention de cette écriture fictionnelle se cristallise finalement autour d'une question : comment vit l'œuvre de Kavan aujourd'hui ?

Kavan, inconnue de tous, et pourtant étonnamment présente sur la toile... De la découverte de nombreux billets de blogs, des couvertures de livres de Kavan traduits dans différentes langues et épinglés de manière disparate, des citations sur Tumblr qui raflent plus de 400 likes, des thèses publiées récemment, des cut-up sensibles entre extraits de journal intime et littérature kavanienne, des tableaux de Kavan exposés au Freud Museum de Londres... Qui lit Kavan ? Quelle est sa présence dans le flux de data ? Qui poste quoi ? Qu'est-ce que cela raconte ? Est-ce un nouveau portrait, une nouvelle mythologie qui émergent ?

Si la bourse du Centre National du Livre a permis à Catherine Lenoble d'étudier les archives personnelles de Kavan au Royaume-Uni et aux États-Unis, la fiction s'élabore en confrontant ces archives matérielles avec un ensemble de data collectées en ligne. À la manière des folkloristes au début du 20e siècle qui collectent, étudient, préservent des productions collectives qui se transmettent de génération en génération par l'oralité (contes, récits, chants, croyances) – le web-folklore traduit cette production collective immatérielle qui habite la toile et se diffuse sur le réseau.

Cette quête et cette collecte sont à la fois la matière qui infuse la fiction et la matière dépouillée par l'auteure avec la complicité du duo <stdin> pour accompagner et enrichir le texte dans sa version papier d'une création en ligne. Pour faire parler les *data*.

Le récit mis en place se compose en grande partie d'éléments textuels en langage naturel, cependant d'autres éléments viennent enrichir la narration, en dialogue avec la création en ligne. La fiction intègre des images et des éléments textuels en langage de programmation : des portions de code (snippets), de courts programmes (scripts Python, NLTK), de la syntaxe du web sémantique (vocabulaire RDF), des définitions contextuelles (requêtes Sparkl) ou encore des graphes (GraphViz) et des visualisations (NLTK).

La fiction proposée par Catherine Lenoble dans le fond et dans la forme fait ainsi converger le roman et l'édition d'artiste, la littérature et le design graphique, l'écriture numérique et l'esthétique du code. La rencontre avec les éditions HYX qui explore, dans le domaine de la littérature, les nouvelles formes de lecture et d'écriture avec la collection « Graphes » prend ici tout son sens.

Création numérique

Esthétique du code, design fiction, archivage actif

Lors des deux résidences rendues possibles en 2013 avec l'aide au développement du CNC/Dicream, Catherine Lenoble et <stdin> ont effectué une série d'expérimentations. Ci-dessous sont listées les quatre « briques » de recherches qui seront poursuivies et finalisées dans la phase de production, en vue d'intégrer une oeuvre en ligne pour une diffusion à l'automne 2014.

Indexation

La manière dont nous souhaitons organiser et articuler les différents contenus qui alimentent la fiction print et web (collection de data Zotero, graphes et visualisations, archéologie des médias, matériauèque cf. bilan, etc) repose sur une plate-forme de type wiki inspirée du wiki Active Archives permettant d'accéder à des ressources qui restent actives en ligne, modifiant le paradigme du téléchargement de ressources depuis un serveur centralisé.

Cette plate-forme est l'extension de la fiction print, un territoire fictionnel qui agrège, structure, facilite la navigation dans les différentes briques d'expérimentations et data présentées dans ce dossier. Le framework que nous envisageons est un système hétéroclite combinant des pages HTML avec des données sémantiques. L'interface d'édition de type wiki étant malléable, cela nous permet de produire aussi bien du HTML que du RDF ou des métadonnées. Techniquement, nous ne pensons pas en terme de base de données SQL centralisée mais plutôt à une collection de pages avec au-dessus un système d'indexation venant « grapiller » les données présentes sur les différentes pages. Le système est donc décentralisé, pas nécessairement unifié, si ce n'est par un design graphique et web commun reliant l'ensemble des pages.

Un des challenges est de pouvoir créer des interactions et passages entre les deux supports, l'édition-livre et l'oeuvre en ligne, notamment en ayant recours à des processus d'indexation, à mi-chemin entre l'index que l'on retrouve traditionnellement en fin de livre et les systèmes de regroupements de mots-clés facilitant la navigation sur un site, façon nuage de tags. Les nouveaux usages du Web invitent en effet à repenser la manière classique d'indexer les contenus du livre imprimé, pour envisager de mettre en exergue les sources mineures habituellement non-visibles.

Nous envisageons ainsi de mettre en place un index, dans l'édition imprimée, situé en fin d'ouvrage et élaboré sur la base des mots, ou couples de mots insolites présents dans le roman. Cette liste sera générée à l'aide de Natural Language Toolkit (NLTK est un ensemble d'outils logiciels et de corpus de textes pour l'analyse lexicale) permettant de révéler certains motifs difficiles à relever lors d'une lecture sans médiation et de mettre en avant la singularité de la langue du roman.

Sur ce modèle, d'autres listes de mots seront extraites afin de proposer en ligne une multiplicité d'entrées ainsi qu'une organisation sémantique inspirée des wikis.

À titre d'exemple, nous souhaitons exposer une « bibliographie visuelle » en s'appuyant sur une collection de couvertures d'ouvrages d'Anna Kavan et l'idée d'une archéologie des livres en superposant les couvertures lorsqu'il existe plusieurs versions d'une même oeuvre. Cette collection permettrait de visualiser différentes couches bibliographiques en y présentant les livres d'Anna Kavan par ordre de publication, enrichis de leur parcours éditorial multiple (éditeurs, ré-éditions, traductions...). Chaque livre de la collection, ayant fait l'objet d'une

expérimentation sur le site, pourra être tagué et pointer vers les résultats de cette expérimentation, créant ainsi un nouveau mode d'indexation.

Base de données narrative

Depuis deux ans, Catherine Lenoble fouille le web et collectionne les data (URLs, images, PDFs, articles de blogs, etc.) autour de la vie et l'œuvre d'Anna Kavan. L'outil en ligne Zotero a été choisi pour collecter et organiser ces références dans un seul et même endroit en ligne et construire une base de travail partagée en ayant recours à un système de marquage commun (tags). C'est un greffon logiciel pour navigateur web qui permet à ses utilisateurs de récolter et d'annoter des références bibliographiques au fur et à mesure de leurs lectures sur le web. Zotero permet d'exporter les bibliographies dans un certain nombre de formats pour le web sémantique.

À partir de cette collection évolutive (à ce jour 250 items), notre travail consiste désormais à faire parler ces data. La visualisation de cette collection sous la forme d'une ligne du temps nous est apparue la solution la plus appropriée, par le biais de filtres activables afin d'identifier les pratiques de publication en *Kavanland* et d'interroger la « valeur » littéraire de l'œuvre de Kavan sur le web. Il s'agit ici de mettre en jeu la perception de l'œuvre de Kavan, l'imaginaire et le fantasme déployée autour de l'« héroïne » Anna Kavan à la fois personnage et auteure de sa propre mythologie.

Nous souhaitons associer un script pour chacune des interrogations de la présente collection. À titre d'exemple, partir de la base de données existante sur Zotero, il est possible de ressortir l'ensemble des urls qui mentionnent le livre *Ice* (qu'il s'agisse d'un article de fan ayant relayé une couverture du livre, une chronique de journaliste, un article universitaire, une citation, etc) par le biais d'un script qui vient extraire toutes les références à *Ice* et générer une page HTML désignée de façon à présenter l'ensemble des résultats. Autant de scripts que de questions seront mis en place pour couvrir l'ensemble de la bibliographie d'Anna Kavan.

Par ailleurs, l'espace d'annotation de Zotero (édition de notes pour chaque référence bibliographique stockée) jusqu'ici sous-exploité, sera le lieu de ré-injection d'une partie de la fiction (fragments de la version print) correspondant précisément à la mise en récit de ce « voyage » en *Kavanland*.

Pour supporter, contredire ou nuancer des intuitions de ce qui s'écrit et se diffuse sur Kavan en ligne par ses fans ou critiques, nous avons envie d'aller plus loin et de faire en sorte que les internautes puissent interroger la base de données tel l'Oracle : le programme fournit une réponse à la personne venue le consulter. À la manière d'Izipedia, version bêta d'un service de réponse automatique à des questions de culture générale basé sur le contenu de Wikipédia, la requête est ici proposée comme un espace conversationnel, en interaction avec les données de l'œuvre en ligne.

En recoupant des données extraites à l'aide de Zotero, il devient en effet possible de répondre à un certain nombre de ces questions, sous peine qu'elles soient reformulées dans un langage compréhensible par la machine. Nous traduisons ces interrogations et postulats en langage naturel vers le langage formel SPARQL. Nous pouvons ainsi interroger la base de données sémantique de Zotero, et l'augmenter par l'ajout manuel de références, notes et faits. En retour, nous obtenons des listes de résultats (p. ex. classements des pays/villes des bloggers) ou des réponses par l'affirmative

ou la négative. Interrogée dynamiquement à chaque lecture, la base de données s'enrichit au fur et à mesure de nouveaux postulats comme autant de prophéties ou la prédiction de mythologies.

Littérature comparée automatique

Et si l'on passait les œuvres de Kavan dans des moulinettes algorithmiques... est-ce qu'un programme pourrait détecter le changement de style littéraire d'Anna Kavan, avant/après 1940 ? Est-ce qu'un programme pourrait proposer une analyse littéraire autre que celle offerte par la lecture à l'œil nu ?

Cette expérimentation se propose d'observer l'évolution du style littéraire de la période Helen Ferguson (HF), avant 1940 avec la période Anna Kavan (AK) après 1940.

Sur la base de ces quatre œuvres :

- d'une part *Let Me Alone*, 1930 (apparition du personnage Anna Kavan) et *Stranger Still*, 1935 (suite des aventures d'Anna Kavan) pour la période HF : style conventionnel, linéarité du récit, présence de nombreux personnages nommés et toponymie ancrée dans le réel ;

- et d'autre part *Who Are You ?*, 1963 (« nouveau roman ») et *Ice*, 1967 (slipstream) pour la période AK : récit fragmenté, resserré sur trois ou quatre personnages qui ne portent pas nécessairement de nom et qui évoluent dans un décor hors-sol, hors-temps sans réelle intrigue) ;

nous cherchons à repérer les répétitions et résonances entre les œuvres (retrouver le même type de personnage, de lieu, de dramaturgie) et les contenus à caractère autobiographique (tel personnage, tel lieu, telle situation en écho à une expérience vécue par Kavan elle-même).

Nous nous sommes penchés dans un premier temps sur les romans *Laissez-moi ma solitude* et *Neige*. Les discussions avec Peter Owen Publisher, l'éditeur d'Anna Kavan, ayant abouti nous obtenons l'autorisation de travailler à partir de ces œuvres, pour certaines livrées par l'éditeur au format numérique (*Neige*, *Laissez-moi ma solitude*, *Who Are You*) pour les autres (*Stranger Still*, *Mercury*, *A Bright Green Field*) que nous scannerons nous-même par le biais d'un Book Scanner.

Cette expérimentation nous permet de produire des analyses des œuvres en vocabulaire RDF exportées en GraphViz. Pour rappel, RDF est le langage de base du web sémantique. Il permet de recouper des données publiées de manière décentralisées et de créer des bases de connaissance, que l'on peut ensuite interroger. Un document RDF est écrit dans une syntaxe, qui s'appuie sur différents types de vocabulaires (par exemple FOAF, Friend Of A Friend pour décrire des relations) et permet de relier et structurer des données qui sont modélisées et organisées sous forme de graphes. Nous ré-écrivons la trame narrative d'un roman de Kavan en déclarant les personnages, leurs relations, leurs situations en détournant des vocabulaires existants pour un usage littéraire, narratif, expérimental.

En parallèle, nous procédons à un travail de traitement automatique des langues, discipline à mi-chemin entre la linguistique, l'informatique et l'intelligence artificielle, en ayant recours à la bibliothèque logicielle NLTK (Natural Language ToolKit). Ceci afin d'effectuer un inventaire sur ces quatre œuvres des noms propres (sous forme de liste alphabétique les noms des villes et des personnages) et une dispersion lexicale du récit (visualisation sur la ligne de temps du récit, des présences des personnages et visualiser l'itinéraire du récit avec la position des villes par ordre d'apparition).

Avec cette prose, ces graphes, ces listes et visualisations, nous souhaitons détecter des patterns et routines d'écriture chez Kavan. Une partie de ces éléments textuels et visuels intégreront « à plat » la version print de la fiction là où la version web permettra de recueillir l'ensemble de ces éléments, de les relier, les structurer, les indexer à partir des méta-données contenues dans ces formats.

Archéologie du texte et des médias

L'archéologie des médias pose la question de la matérialité particulière des œuvres et les différentes stratégies de conservation qui s'expérimentent (documentation, émulation, migration, partition, ré-interprétation, auto-archivage, archivage automatique), d'une certaine façon, c'est aussi une invitation à de nouvelles explorations artistiques.

La création en ligne que nous proposons peut ainsi s'envisager sous l'angle du musée en tant qu'archive vivante, espace de documentation et d'exposition de la généalogie de l'œuvre de Kavan à travers des fragments d'œuvres originales et des versions ré-actualisées remixées.

Comment l'œuvre de Kavan entre-t-elle en résonance avec le patrimoine littéraire dont elle hérite et celui qui lui succède ? Sources d'inspiration, de variations, de plagiat, de remix, nous tentons de faire découvrir les œuvres de Kavan sous un œil neuf en développant des expériences re-créatrices de liens, de filiations, d'échos et de dialogue.

Avec l'accès aux fichiers sources des œuvres de Kavan (permission de l'éditeur Peter Owen et numérisation via un Book Scanner comme évoqué précédemment), nous envisageons les expérimentations suivantes :

- détection d'auto-plagiat (avec NLTK) entre les romans *Ice* (1967) et *Mercury* (publication posthume, 1994) d'Anna Kavan. On peut parler d'auto-plagiat ou self-plagiarism lorsqu'une matière (plastique, textuelle, musicale) subit de légères modifications par son auteur-e pour créer un autre matériau fort semblable. La recomposition de sa propre création est une pratique nullement illégale et souvent relevée dans le champ artistique créant volontairement une confusion dans la réception de l'œuvre ;

- détection de plagiat par anticipation (avec NLTK) entre les nouvelles *The Stoker* (*Amerika* de Kafka, 1913) et *New & Splendid* (*A Bright Green Field* de Kavan, 1958). On évoque souvent l'influence artistique dans la succession sans considérer que l'inverse soit possible. La notion de plagiat par anticipation, proposée par François Le Lionnais, membre de l'Oulipo, consiste à faire un usage créatif du plagiat en relevant des découvertes formelles illustrées dans des œuvres antérieures. Et si Kafka avait plagié sa nouvelle *The Stoker* publiée en 1913 sur celle de Kavan *New and Splendid* éditée dans le recueil *A Bright Green Field* en 1958 ?

Et si la lecture d'une œuvre nous engage à la replacer dans un réseau de temporalités (passé, présent, futur), l'écriture aussi se relit en dialogue avec des revenants et des survenants (ceux qui viendront après). Nous ajoutons par ailleurs à ce travail, une galerie qui recueillera un ensemble de créations et de médias déposés en ligne par des artistes, chercheurs, lecteurs que Kavan inspire. Ils et elles recomposent une nouvelle œuvre en s'inspirant de son travail, lui rendant hommage ou tout simplement en se faisant passeurs de l'héritage de la romancière.