



L'île des savoirs choisis

D'HyperNietzsche à ScholarSource : pour une infrastructure de recherche sur le Web

Paolo D'Iorio



Édition électronique

URL : <https://journals.openedition.org/rechtrav/121>

DOI : [10.4000/recherchestravaux.121](https://doi.org/10.4000/recherchestravaux.121)

ISSN : 1969-6434

Éditeur

UGA Éditions/Université Grenoble Alpes

Édition imprimée

Date de publication : 15 juin 2008

Pagination : 279-301

ISBN : 978-2-84310-125-0

ISSN : 0151-1874

Référence électronique

Paolo D'Iorio, « L'île des savoirs choisis », *Recherches & Travaux* [En ligne], 72 | 2008, mis en ligne le 15 décembre 2009, consulté le 11 octobre 2023. URL : <http://journals.openedition.org/rechtrav/121> ; DOI : <https://doi.org/10.4000/recherchestravaux.121>

Le texte et les autres éléments (illustrations, fichiers annexes importés), sont « Tous droits réservés », sauf mention contraire.

**L'île des savoirs choisis
D'HyperNietzsche à ScholarSource :
pour une infrastructure de recherche sur le Web**

Pour étudier les sources primaires nécessaires à ses travaux, un chercheur se rend habituellement dans des lieux d'archives ou dans des bibliothèques ; en bibliothèque il lit non seulement les sources primaires mais aussi les sources secondaires, revues, monographies, éditions critiques, qu'il peut dans un certain nombre de cas se procurer également chez les libraires ; les cours à l'université, les séminaires et les colloques lui donnent l'occasion d'échanger ses connaissances avec ses collègues et de les transmettre à ses élèves ; enfin notre chercheur publie ses propres études critiques chez des éditeurs et les diffuse sur le marché du livre selon les règles établies par le droit d'auteur. Tous ces éléments matériels et organisationnels – archives, bibliothèques, librairies, universités, cours, colloques, maisons d'édition, droit d'auteur et autres encore – constituent l'infrastructure traditionnelle de la recherche en sciences humaines telle qu'elle s'est formée lentement au cours des siècles passés et au sein de laquelle évoluent encore aujourd'hui nos modes de recherche.

Une infrastructure de recherche sur le Web devrait permettre aux chercheurs d'effectuer la totalité ou du moins la plupart des activités scientifiques qu'ils accomplissent dans le cadre de l'infrastructure traditionnelle, de travailler de manière plus souple et efficace et d'avoir accès à de nouveaux services et à des outils caractéristiques d'un environnement électronique.

Malgré l'existence de nombreux projets d'édition sur le Web, d'archives institutionnelles, de journaux en ligne, il n'existe pas encore une infrastructure électronique qui réunisse de manière cohérente la version électronique des

éléments matériels et organisationnels que nous venons d'évoquer et qui permette aux chercheurs de consulter sur le Web à la fois les sources primaires et les sources secondaires; de publier dans l'infrastructure électronique leurs propres études critiques et d'avoir à leur disposition un système performant d'enseignement et de partage des connaissances sur le Web.

À l'instar de beaucoup d'autres projets, le projet HyperNietzsche a essayé de répondre à la question de savoir s'il était possible d'utiliser Internet pour la recherche en sciences humaines, en particulier pour établir, commenter, analyser les grands textes de la tradition littéraire et philosophique. Puisque j'ai déjà présenté ailleurs les objectifs et les acquis de ce projet¹, je vais présenter ici ce qu'HyperNietzsche n'est pas et ce qu'il va être à l'avenir.

Ce qu'HyperNietzsche n'est pas

1. HyperNietzsche n'est pas un portail

Il fut une époque vers la fin des années 1990 où tout le monde pensait que l'avenir d'Internet allait être déterminé par le développement des portails. Un portail est un site qui sert de point d'accès à l'ensemble du Web ou à toutes les ressources Web concernant un sujet donné. Dans les portails littéraires, cela se traduit normalement en une série d'informations concernant un auteur ou un ensemble d'auteurs : bibliographie, portraits, photos d'époque, quelques extraits des œuvres, des nouvelles concernant des publications ou des colloques récents, et des liens vers d'autres ressources similaires. Cela est fort utile mais ne correspond pas à ce qu'HyperNietzsche entend être.

Toutefois pour rendre plus patent qu'HyperNietzsche n'est pas un portail, l'équipe qui en est responsable a créé un portail de *News* qui concernent l'univers nietzschéen : le *Nietzsche News Center* (NNC). Vous pouvez y lire et annoncer des nouvelles concernant les conférences, colloques, publications, événements, articles dans la presse, et suivre des liens vers d'autres sites nietzschéens, de même que vous abonner à la *newsletter* et recevoir tous les quinze jours les nouvelles dans votre boîte aux lettres électronique. Le *Nietzsche News Center* est basé sur un système de gestion de contenu nommé Matrioska, logiciel libre et gratuit

1. Voir *HyperNietzsche. Modèle d'un hypertexte savant sur Internet pour la recherche en sciences humaines. Questions philosophiques, problèmes juridiques, outils informatiques*, P. D'Iorio éd., PUF, 2000 (version électronique librement disponible à l'adresse <http://www.hypernietzsche.org/doc/puf/>); P. D'Iorio, «Principes de l'HyperNietzsche», *Diogène*, 196, oct.-déc. 2001, PUF, p. 77-94; «HyperNietzsche», *Expériences d'éditions savantes hypertextuelles*, A. Crasson éd., Louvain-la-Neuve, Academia Bruylant, 2007, p. 265-288; «Nietzsche on New Paths : The Hyper Nietzsche Project and Open Scholarship on the Web», *Friedrich Nietzsche. Edizioni e interpretazioni*, M. C. Fornari, Sergio Franzese éd., Pise, ETS, 2008 (version électronique disponible à l'adresse : <http://www.hypernietzsche.org/doc/files/new-paths.pdf>).

qui peut donc être librement utilisé pour construire d'autres portails de nouvelles². Tout en étant un service utile pour ceux qui s'intéressent à Nietzsche, le *Nietzsche News Center* n'est pas HyperNietzsche et HyperNietzsche n'est pas un portail.

2. HyperNietzsche n'est pas un objet fermé

HyperNietzsche n'est pas non plus un objet fermé, ce que sont en revanche un livre, un CD-ROM ou leur reproduction sous forme de site Web. HyperNietzsche ressemble davantage à une bibliothèque qui contient beaucoup de documents et de livres de différents genres et qui est toujours prête à en accueillir de nouveaux. Si on le compare aux systèmes traditionnels de conservation et de transmission des connaissances, on peut définir HyperNietzsche comme la réunion d'archives publiques permettant l'accès libre aux sources primaires (œuvres, manuscrits, lettres), d'une bibliothèque publique donnant libre accès aux contributions scientifiques (éditions critiques, articles, monographies) et d'une maison d'édition universitaire sans but lucratif, toujours prête à publier de nouveaux textes selon ses procédures rigoureuses d'évaluation par les pairs et par son comité scientifique. Hyper Nietzsche voudrait donc être un premier exemple d'infrastructure de recherche pour les sciences humaines.

Une édition critique représente un exemple typique d'objet fermé. Or HyperNietzsche n'est pas une édition critique pour différentes raisons : tout d'abord parce qu'il ne contient pas que des éditions, mais aussi d'autres types de contributions scientifiques (par exemple des essais), ensuite et surtout parce qu'il peut contenir plus d'une édition critique à la fois et peut permettre de les confronter. Toutefois, là aussi, pour montrer comment HyperNietzsche pourrait être utilisé pour publier des éditions électroniques, mon équipe a établi, dans l'HyperNietzsche, sa propre édition génétique de deux œuvres de Nietzsche : *Le Voyageur et son Ombre* et *Aurore*. Dans le cadre de cette édition nous avons utilisé nos propres critères philologiques, notre propre langage d'encodage, nos propres interfaces de visualisation. D'autres spécialistes de Nietzsche seront libres d'utiliser *leurs* propres critères philologiques, leurs propres langages d'encodages, leurs propres modalités de visualisation et de publier leurs éditions à côté de la nôtre, *dans* l'HyperNietzsche. Et d'autres systèmes Hyper pourront *a fortiori* utiliser des critères philologiques, des langages d'encodage, des modalités de visualisation différentes et néanmoins être capables de partager leurs connaissances à l'intérieur du réseau des Hyper. Un principe fondamental d'HyperNietzsche est que la multiplicité des approches

2. Voir <http://www.matrioska.org>.

doit être encouragée. La cohérence est exigée au niveau de chaque contribution, mais n'a pas de sens pour l'ensemble des contenus publiés dans un Hyper, pour ne pas parler du réseau des Hyper. Tout comme on ne demande pas à une bibliothèque que tous ses livres utilisent la même mise en page, le même système de citations ou, encore moins, les mêmes critères scientifiques. HyperNietzsche a la cohérence d'une bibliothèque et non celle d'une édition. HyperNietzsche n'est pas une édition, il n'est pas un objet fermé.

3. HyperNietzsche n'est pas un système d'encodage et de fouille textuelle

Le fonctionnement d'HyperNietzsche et sa manière de structurer l'information ne sont pas basés sur un système d'encodage des textes et encore moins sur un langage d'encodage en particulier. Au contraire, il permet de publier des contributions scientifiques utilisant différents formats : HTML, XML et ses dérivés comme TEI pour les textes ; PDF, JPEG, PNG pour les images ; MP3 et MPEG4 pour les contributions audiovisuelles, etc.

Nous pouvons même affirmer qu'HyperNietzsche représente une étape dans la transition vers un nouveau paradigme. En tout cas, le travail sur les manuscrits nous pousse vers l'élaboration d'un nouveau paradigme qui dépasse l'approche traditionnelle de l'*humanities computing* basée sur la linguistique computationnelle, et vise à créer des environnements électroniques de recherche et de publication en ligne qui ne sont pas centrés sur la recherche textuelle à l'intérieur de banques de données textuelles, mais sur la publication de documents de différents types et formats (textes, images, audio, vidéo) qui sont générés et gérés directement par les chercheurs. Cela ne veut pas dire ajouter une nouvelle discipline à la structure traditionnelle des sciences humaines, comme l'a fait la linguistique computationnelle, mais essayer de transposer *l'ensemble* de la structure et des pratiques scientifiques en sciences humaines dans un environnement électronique. Dans le paradigme de l'*humanities computing*, le texte était au centre et tous les autres éléments (fac-similés des éditions originales ou des manuscrits, commentaires critiques) tournaient autour du texte. Le nouveau paradigme que nous proposons met en revanche en son centre les sources primaires, les documents dans toute leur matérialité, et toutes les autres contributions critiques, y compris les textes des différentes éditions qui ont été établies, tournent autour des sources primaires. Dans le vieux paradigme, la fonction principale était la recherche textuelle ; le nouveau paradigme ne renonce certes pas à la recherche textuelle, mais il privilégie la mise en contexte hypertextuelle où chaque document est accompagné par les contributions scientifiques corrélées. De la sorte, d'une consultation selon des arbres hiérarchiques on passe à la navigation dans une structure à graphe, au sens mathématique du terme, où chaque élément est lié aux autres

selon différents critères de sens. Conceptuellement, l'ancien paradigme est formé d'une masse indifférenciée de textes plus ou moins balisés et d'un moteur de recherche qui sert d'instrument d'interrogation textuelle. Dans le nouveau paradigme, en revanche, on balise les références contenues dans les articles des chercheurs, qui sont utilisées comme système de navigation dans la connaissance : ainsi on navigue suivant les relations de sens que les chercheurs ont établies entre différents passages textuels au lieu de feuilleter des listes d'occurrences de mots³. Finalement, quand l'ancien paradigme basé sur la linguistique computationnelle utilise le Web, c'est du Web 1.0 qu'il s'agit, alors que le nouveau paradigme s'inscrit complètement dans le Web 2.0 avec tout ce que cela implique en termes de technologie, mais aussi de logique des *social networks*.

Et dans ce cas aussi, tout en réaffirmant que l'HyperNietzsche ne repose pas sur un système d'encodage de texte, mon équipe ne s'est pas privée de concevoir et d'écrire un nouveau langage d'encodage spécifiquement conçu pour représenter les processus génétiques à l'œuvre dans les manuscrits de notre philosophe : le *HyperNietzsche Markup Language* (HNML). Ce langage d'encodage, basé sur l'XML et inspiré de la TEI, nous a entre autre permis de baliser et de représenter à l'écran les différentes strates présentes sur chaque page des manuscrits de Nietzsche, correspondant à diverses campagnes d'écriture⁴.

En tout cas, même si notre nouvelle édition électronique de Nietzsche est encodée en HNML, l'HyperNietzsche n'est basé ni sur l'HNML ni sur un autre langage d'encodage en particulier. Les contributeurs peuvent utiliser le langage de leur choix. Un certain nombre de formats et de langages d'encodage sont déjà gérés par HyperNietzsche, mais il est toujours possible d'ajouter le support pour un nouveau format ou pour un nouveau langage d'encodage. En effet, je le répète, HyperNietzsche veut être comme une bibliothèque avec beaucoup de rayons pour des livres de tous formats, écrits dans différents langages, et donc il doit être en mesure de gérer un grand

3. Voir *infra*, l'explication du fonctionnement de la contextualisation dynamique.

4. Voir M. D'Alfonso et H. Saller, «Kodierung und Darstellung von Schreibschriften in der elektronischen Edition des Druckmanuskripts zu *Der Wanderer und sein Schatten*», *Literatur und Literaturwissenschaften auf dem Weg zu den digitalen Medien – Eine Standortbestimmung*, M. Stolz, L. Marco G. et J. Loop éd., Zürich, Germanistik.ch, 2007, p. 117-126 (version électronique à l'adresse suivante : http://www.germanistik.ch/buch.php?id=Literatur_Neue_Medien). Nous avons dû créer ce nouveau langage, parce qu'il y a cinq ans, quand nous avons commencé à travailler à notre nouvelle édition de Nietzsche, il n'existait aucun langage d'encodage standard pour les manuscrits. Maintenant il existe une section de la TEI consacrée aux manuscrits, même s'il n'y a pas encore de nombreux exemples d'utilisation, ni de feuilles de style XSL et CSS disponibles ; voir <http://www.tei-c.org/release/doc/tei-p5-doc/en/html/MS.html>.

nombre de formats et de langages d'encodage. Il serait ridicule que ceux qui veulent construire une bibliothèque prévoient des rayons uniquement pour les livres *in octavo* écrits en anglais parce que cela serait censé être un standard...

L'île des savoirs choisis

HyperNietzsche n'est pas un portail, il n'est pas non plus un objet fermé comme une édition, il n'est pas non plus un langage d'encodage : qu'est-il alors ? HyperNietzsche, c'est une île, une île bienheureuse dans la vaste mer du Web. Et pourquoi devrions-nous visiter cette île ou même contribuer à la consolider ? Parce que dans cette île, on peut accéder à des savoirs choisis, des savoirs auxquels se fier car ils ont été validés par la communauté des spécialistes et représentent l'état de l'art dans ce domaine. Il semblerait qu'Umberto Eco ait dit un jour : « Je trouve un tas de choses très intéressantes sur Internet et après je vais voir dans les livres pour savoir si elles sont vraies ou non. » Hé bien, notre objectif est de simplifier la vie au grand savant italien en mettant à sa disposition, sur notre île, « un tas de choses très intéressantes » qui soient au moins aussi vraies que celles qu'il peut trouver dans les livres.

Mais comment pouvons-nous construire cette île des savoirs choisis ? Par des valeurs et des procédures ; et aussi à l'aide d'une technologie qui nous permette de leur donner corps, dans un environnement électronique vivant, sur le Web. Voici une petite liste de valeurs et de procédures qui forment notre île : 1) publication ouverte ; 2) évaluation par les pairs ; 3) stabilité et durabilité du texte et de l'adresse Web ; 4) un certain nombre de fonctions techniques parmi lesquelles ce que nous appelons « contextualisation dynamique ». Analysons-les plus en détail.

1. Publication ouverte

La publication ouverte est une *valeur* qui se décline au moins sous trois formes : 1) libre accès aux sources primaires ; 2) libre accès aux contributions scientifiques ; 3) libre accès au code source des logiciels utilisés.

1) Par *libre accès numérique aux sources primaires*, nous entendons la possibilité de donner un accès universel, à travers Internet, aux fonds documentaires patrimoine de l'humanité (textes, images, sons, objets artistiques ou archéologiques), conservés par des bibliothèques, des archives ou des musées publics. Source d'intérêt pour toute personne de culture, ces matériaux représentent les sources primaires pour le travail des chercheurs ; la diffusion sur Internet de leur reproduction numérique est une forme extraordinaire de valorisation qui de surcroît respecte les exigences de conservation des originaux.

Trop souvent pourtant, même dans le cas de fonds anciens tombés depuis longtemps dans le domaine public et conservés par des institutions publiques, les projets de publication sur le Web sont bloqués par la politique restrictive de certaines d'entre elles qui ne permettent pas aux chercheurs de publier une copie numérique de leur fonds sur Internet ou qui font payer pour cela des «droits de reproduction» exorbitants. Publier sur Internet une page venant des collections de la Bibliothèque nationale de France coûte 124 euros ; le tarif est de 107 livres pour 6 mois et 324 livres pour plus de dix ans à la British Library, 150 euros à la Bayerische Staatsbibliothek de Munich, etc.⁵ Sans compter que les «droits de reproduction» se cumulent avec les frais de numérisation dont le prix pratiqué par les bibliothèques est normalement beaucoup plus élevé que le prix du marché. La justification juridique des «droits de reproduction» est bien expliquée par la British Library :

Pourquoi dois-je obtenir une autorisation pour utiliser des documents de la Bibliothèque qui sont dans le domaine public et libres de droits ? Les ouvrages originaux sont certes dans le domaine public, mais les copies que fournit la Bibliothèque sont protégées par le droit d'auteur (*copyright*) parce que ce sont de nouvelles copies des documents originaux, et ce sont les copies qui sont protégées par le droit d'auteur (*copyright*). C'est pourquoi vous devez obtenir une autorisation⁶.

C'est exact d'un point de vue juridique : celui qui numérise un texte ou une image détient un droit d'auteur sur sa propre reproduction et même si le texte original était dans le domaine public, la reproduction est protégée. Mais rien n'oblige les bibliothèques à monnayer ces droits : plusieurs bibliothèques permettent d'ailleurs à leurs lecteurs de numériser directement leur fonds ou cèdent gratuitement les droits de reproduction pour des usages non commerciaux, par exemple toutes les bibliothèques italiennes, mais aussi la *Staatsbibliothek* de Berlin et un grand nombre de bibliothèques françaises (d'après leur site Web). D'autre part, réclamer des «droits de reproduction» est difficilement justifiable aussi bien d'un point de vue social que culturel. La conservation et la diffusion maximales du patrimoine sont en effet la mission principale des archives et des bibliothèques publiques et c'est la raison pour laquelle l'État, avec l'argent du contribuable, finance ces institutions et paie les salaires des conservateurs et des bibliothécaires. Si des chercheurs, avec leurs propres financements, s'engagent à numériser et à diffuser gratuitement des fonds patrimoniaux sur le Web à des fins d'enseignement et de recherche, ils ne font qu'aider les bibliothèques à accomplir leur mission statutaire et elles devraient les remercier au lieu de les

5. Voir <http://www.bnf.fr/pages/accedocu/redev.htm> ; <http://www.bl.uk/imaging/permission.html> ; <http://www.bsb-muenchen.de/Reproduktionsgebuehren.793.o.html>.

6. Voir <http://www.bl.uk/imaging/permisfaq.html>.

taxer. Et d'un point de vue culturel, il est bien curieux de s'insurger contre l'influence de la culture américaine sur le Web et de plaider pour une diffusion des contenus français comme l'a fait l'ancien président de la Bibliothèque nationale de France, et en même temps d'avoir recours à un «bureau de la redevance» qui freine toute entreprise d'édition numérique du patrimoine français. La politique des bibliothèques américaines, grâce au *fair use* qui permet d'utiliser librement les contenus numériques à des fins d'enseignement et de recherche, est beaucoup plus efficace. Et les résultats sont tangibles : il suffit de citer l'exemple du site *American Memory* de la Bibliothèque du Congrès⁷.

Les projets de numérisation nationaux et le projet de la bibliothèque numérique européenne, inspiré entre autres par le plaidoyer de Jean-Noël Jeanneney, sont bien évidemment un important moyen pour augmenter la quantité de contenus français et européens sur le Web ; mais un simple dispositif législatif national ou européen pourrait produire un effet aussi important, en permettant la floraison de toute une série d'initiatives émanant de la communauté des chercheurs qui actuellement sont bloquées. Il suffirait de statuer que tout projet qui vise à publier sur le Web en accès libre et gratuit des fonds patrimoniaux conservés dans des institutions publiques soit exempté du paiement de tout «droit de reproduction» et puisse effectuer la numérisation des documents concernés rapidement et au prix du marché⁸.

En attendant, l'accès numérique au patrimoine se négocie au cas par cas entre les chercheurs et les directeurs des bibliothèques. Dans le cas de Nietzsche, après une série d'études sur les droits de son œuvre⁹, nous avons eu la chance d'être écoutés par le président de la *Klassik Stiftung Weimar* qui a signé en 2002 une convention autorisant l'association HyperNietzsche à numériser et publier sur le Web en libre accès à des fins non commerciales tous les documents concernant Nietzsche conservés aux archives et à la Bibliothèque de Weimar : manuscrits, œuvres, correspondance, iconographie, livres de sa bibliothèque personnelle. Cette convention était une première dans le panorama allemand et peut d'ailleurs servir de modèle d'accord entre les chercheurs et les archives qui désirent mettre à disposition sur Internet des sources primaires pour le travail scientifique.

7. Voir J.-N. Jeanneney, *Quand Google défie l'Europe : plaidoyer pour un sursaut*, Mille et une nuits, 2005 ; bureau de la redevance : <http://www.bnf.fr/pages/accedocu/redev.htm> ; American Memory : <http://memory.loc.gov/ammem/>.

8. Une première discussion de ce problème a été proposée au cours du colloque *L'Open Source dans les sciences humaines. Modèles ouverts de recherche et de publication sur Internet* organisé à l'École normale supérieure de Paris par la Maison des sciences de l'Homme (programme «Numérisation pour l'enseignement et la recherche»), l'association Fabula et l'association HyperNietzsche en janvier 2002 (voir <http://www.hypernietzsche.org/events/os/>).

9. Voir par exemple l'article de S. Clausing : <http://www.hypernietzsche.org/events/lmu/clausing.html>.

2) Le libre accès aux contributions scientifiques produites par les chercheurs (articles, monographies, commentaires, éditions) est désormais un thème de grande actualité. Il suffit de rappeler le texte de la *Déclaration de Berlin sur le libre accès à la connaissance en sciences exactes, sciences de la vie, sciences humaines et sociales* qui a été signée par un grand nombre de chercheurs et d'institutions dans le monde et qui commence à produire ses effets¹⁰. Ce mouvement est mené surtout par des scientifiques et représente une réponse au marché de monopole des périodiques scientifiques créé par certaines maisons d'édition. Un article éclairant de Jean-Claude Guédon¹¹ retrace l'histoire de ce marché de monopole comme effet pervers du mécanisme d'évaluation quantitative du travail des chercheurs inventé dans les années 1950 par Eugene Garfield : le *Science Citation Index* (SCI). Le SCI comportait la définition d'un ensemble de «*core journals*», vers lesquels s'orientaient ensuite les pratiques d'acquisition des bibliothèques. Au début des années 1970, certains éditeurs se sont rendu compte que ce processus allait déboucher sur la création d'un marché de monopole et ont cherché à l'accélérer. À la fin des années 1980, le nouveau marché du périodique scientifique était en place : les prix avaient commencé à monter, complètement déconnectés des coûts de production et de distribution, tandis que la demande restait à peu près stable. Pour les bibliothèques et les centres de recherche, cela prit la forme du cauchemar nommé «*the serial pricing crisis*» : la crise du prix du périodique. À la fin du XX^e siècle, les scientifiques ont essayé d'y réagir en lançant différentes initiatives¹² et en utilisant systématiquement Internet pour la libre diffusion de leurs travaux scientifiques. Malgré les succès de ces initiatives, pour l'instant le groupe d'intérêt formé par les grandes maisons d'édition et les chercheurs qui siègent dans les comités d'évaluation des revues et qui veulent garder leur rôle de «*gatekeepers*» (gardiens d'accès), résiste.

Le mouvement pour le libre accès aux publications scientifiques est presque inexistant dans les sciences humaines pour les raisons décrites par Peter Suber¹³ et pour d'autres encore. Cela n'empêche que les sciences

10. Voir <http://oa.mpg.de/openaccess-berlin/berlindeclaration.html>.

11. J.-Cl. Guédon, «In Oldenburg's Long Shadow : Librarians, Research Scientists, Publishers, and the Control of Scientific Publishing», *Creating the Digital Future*, Proceedings of the 138th Annual Meeting of the Association of Research Libraries, 2001, disponible sur Internet à l'adresse : <http://www.arl.org/resources/pubs/mmproceedings/138guedon.shtml>; imprimé également sous forme de livre imprimé par l'Association of Research Libraries en 2005.

12. Voir par exemple la Public Library of Science : <http://www.publiclibraryofscience.org>; la Budapest Open Access Initiative : <http://www.soros.org/openaccess>; la Scholarly Publishing and Academic Resources Coalition : <http://www.arl.org/sparc>, etc.

13. P. Suber, «Promoting Open Access in the Humanities», <http://www.earlham.edu/~peters/writing/apa.htm> (publié également dans la revue de l'American Philological Association, *Syllecta Classica*, University of Iowa, n° 16, 2005, p. 231-246).

humaines versent de plus en plus dans une situation de crise concernant la publication des résultats de leurs recherches. Si dans les sciences de la nature il existe un marché de monopole, dans les sciences humaines bientôt il n'y aura plus de marché du tout. Premièrement parce que les bibliothèques publiques, en raison de la hausse du prix des périodiques scientifiques, ont de moins en moins d'argent pour acheter des monographies et des revues en sciences humaines. D'autre part les sciences humaines sont depuis toujours dans une situation de crise pour ce qui concerne la publication des recherches érudites ou des éditions volumineuses, notamment les éditions génétiques, les éditions fac-similés, et en général dans tous les cas où les exigences de la science et celles du marché du livre entrent en conflit. Le résultat est que les chercheurs sont de plus en plus obligés de verser une bonne partie de leur maigre budget pour financer la publication de leurs livres.

Plaidant pour la publication sur Internet non seulement des sources primaires, mais également des articles, livres, éditions produits par les chercheurs, le projet HyperNietzsche s'est préoccupé dès le début d'assurer le contrôle de qualité des textes publiés et la protection juridique contre le plagiat. Le contrôle de qualité est traité au point suivant. Du point de vue de la protection juridique, le projet HyperNietzsche, après Richard Stallman mais avant *Creative Commons*¹⁴, a souligné que pour ce qui concerne les publications savantes il fallait séparer la partie du droit d'auteur qui intéresse l'auteur (respect de la paternité intellectuelle, protection du plagiat) de celle qui intéresse les maisons d'édition (rémunération, interdiction de la copie). À cette fin, les juristes de l'HyperNietzsche ont rédigé en l'an 2000 trois licences appelées *OpenKnowledge* qui donnent aux chercheurs la possibilité juridique de publier les résultats de leurs recherches sur Internet en autorisant la lecture et la copie, mais en interdisant le plagiat¹⁵.

Pourquoi les chercheurs devraient-ils publier leurs articles sur le Web ? Tout d'abord pour une raison éthique : étant payés avec de l'argent public, ils devraient avoir l'obligation, au moins morale, de mettre les résultats de leurs travaux gratuitement à disposition du public. Ensuite pour une raison pratique, car les articles publiés sur Internet sont généralement plus lus et plus cités que les articles publiés en revue et assurent donc une grande visibilité à leur auteur et un facteur d'impact élevé à l'intérieur de leur discipline. Les chercheurs qui veulent communiquer leurs idées et leurs découvertes au monde entier savent que le Web est l'instrument le meilleur pour le faire. Mais, certes, la publication en revue ou en livre est encore indispensable pour

14. Voir <http://www.stallman.org> et <http://creativecommons.org>.

15. Voir P. Chevet, «L'application du droit d'auteur au monde contemporain de la recherche», *HyperNietzsche, op. cit.*, p. 37-64.

la carrière d'un jeune chercheur. Cela produit un paradoxe, car les jeunes chercheurs qui sont plus proches d'Internet et voudraient l'utiliser pour la diffusion de leurs travaux sont en même temps ceux qui ont le plus besoin de publier en livre pour avancer dans leur carrière, alors que les chercheurs confirmés qui pourraient se permettre le luxe de publier sur Internet sont généralement assez éloignés de la culture du Web.

Je ne veux pas entrer ici dans la vieille question de la mort du livre, mais simplement faire observer qu'il ne s'agit pas du choix entre deux supports : ce qui est différent est le cadre social, académique et économique qui les entoure actuellement. Du point de vue du support, il est évident que l'infrastructure numérique ne va pas complètement remplacer l'infrastructure traditionnelle : nous n'allons pas jeter les livres et les manuscrits une fois qu'ils sont numérisés et nous ne cesserons pas non plus de lire sur papier pour tout lire sur nos écrans. Entre le papier et le numérique, entre les divers systèmes de diffusion de la connaissance, se passera ce qui s'est passé entre les moyens de locomotion : nous n'avons pas cessé de prendre le train parce que nous disposons de l'avion, ni d'utiliser nos vélos ou nos voitures. Je ne peux pas me rendre de Grenoble à Lyon en avion, mais je peux le faire en train ; le train toutefois ne serait pas l'idéal pour une promenade dans le parc Paul Mistral, il vaudrait mieux un vélo ou de bonnes chaussures ; mon vélo pourrait aussi m'emmener de Grenoble à Paris, mais avec le train cela va plus vite... De la même manière, c'est bien évidemment inconfortable de lire 400 pages de philosophie à l'écran, et il serait utopique de prétendre qu'un de mes livres pourra être distribué dans tous les pays du monde : ce serait comme de vouloir se rendre en Amérique en carrosse. Il faudrait pouvoir utiliser chaque support selon ses qualités : l'ergonomie de lecture des livres et la rapidité de diffusion du Web, et lever tous les obstacles de type technique, mais surtout social, académique et juridique qui nous empêchent de le faire et qui obligent certains chercheurs à financer eux-mêmes la publication d'un livre qui ne sera tiré qu'à 200 exemplaires mal ou pas distribués et dont l'éditeur aura le copyright jusqu'à 70 ans après leur mort. Au contraire il faudrait faire en sorte que les Umberto Eco de l'avenir puissent dire : « je trouve un tas de choses vraies sur Internet et après je vais les imprimer et les relier sous la forme d'un livre pour pouvoir mieux les lire ».

3) Donner *libre accès au code source des logiciels* que nous avons créés pour bâtir notre infrastructure de recherche est une manière de la rendre utilisable et modifiable par d'autres projets et d'en optimiser le coût de développement. Le modèle de développement du *software en open source* et l'effet «chaudron magique» qui en découle n'ont pas besoin d'être expliqués : il suffit de renvoyer

aux écrits classiques d'Eric Raymond¹⁶. Les avantages principaux sont que : 1) les projets qui n'ont pas de budget informatique peuvent télécharger le logiciel et l'utiliser gratuitement; 2) les projets qui souhaitent effectuer des modifications et développer de nouvelles fonctions peuvent le faire librement et les résultats de leur travail seront versés au pot commun; 3) la nécessité de standardiser et de généraliser certaines fonctions impose une certaine discipline et évite au projet de se renfermer sur lui-même comme cela arrive à beaucoup de projets dans nos disciplines. Nous ne pouvons pas encore affirmer que le chaudron magique fonctionne déjà à plein régime dans le cas d'HyperNietzsche, mais nous avons posé les jalons de ce processus dans le cadre du deuxième groupe de travail de l'action COST «Communautés savantes sur le Web¹⁷».

2. Évaluation par les pairs

L'évaluation par les pairs est la deuxième valeur sur laquelle nous développons notre île. Dans la science comme dans beaucoup d'autres activités humaines, un chercheur qui aspire à faire partie de la communauté des savants et à participer aux recherches collectives doit d'abord produire des travaux jugés intéressants par ses collègues. Mais plus que dans d'autres activités, l'évaluation scientifique est censée être la plus impartiale et transparente possible et basée sur des évidences incontestables. Naturellement, cela est très difficile à réaliser et les procédures d'évaluation ne sont jamais exemptes de critiques.

Le Web pourrait représenter l'occasion d'organiser l'évaluation scientifique de manière plus simple, plus transparente et plus efficace. Sur le Web, plusieurs voies sont envisageables : créer un environnement numérique qui reproduise les procédures traditionnelles d'évaluation par les pairs; utiliser la logique du réseau et déterminer le facteur d'impact à partir du balisage sémantique effectué par les usagers de sites comme *Citeulike* ou *Connotea*¹⁸; ou encore faire recours à des indicateurs de popularité comme la quantité de citations ou de téléchargements d'un article, etc. La revue *Nature*, à travers son blog *Peer-to-Peer*, a commencé une réflexion sur différents modèles d'évaluation par les pairs sur le Web accompagné de l'expérimentation concrète d'un nouveau modèle de revue en ligne : *NaturePrecedings*¹⁹.

16. E. S. Raymond, *The Cathedral and the Bazaar, Musings on Linux and Open Source by an Accidental Revolutionary*, O'Reilly, 1999 (qui contient également l'essai intitulé «Le chaudron magique»), disponible en ligne en plusieurs langues à l'adresse : www.catb.org/~esr/writings/cathedral-bazaar.

17. Voir <http://www.cost-a32.eu/wg-2.html>.

18. Voir <http://www.citeulike.org>; <http://www.connotea.org>.

19. Voir <http://blogs.nature.com/peer-to-peer/>; <http://precedings.nature.com>. Pour une discussion critique de ces différentes expériences, voir l'intéressant article de F. di Donato,

Dans le cas de l'HyperNietzsche, nous avons choisi de reproduire les procédures d'évaluation traditionnelles avec quelques modifications significatives. L'évaluation a été placée non pas dans les mains d'une maison d'édition, mais au sein de la communauté des spécialistes de Nietzsche réunis dans le cadre juridique de l'*association HyperNietzsche* (constituée en France selon la loi 1901). Concernant le type de représentation trois modèles étaient possibles : 1) évaluation directe : tous les membres de la communauté valident tout ce qui est proposé pour publication ; 2) évaluation représentative : tous les membres de la communauté délèguent l'évaluation à un nombre restreint de collègues qui forment le comité scientifique ; 3) évaluation absolutiste, éventuellement, *hopefully*, éclairée : une personne décide de manière autonome de la publication (cela arrive même dans les meilleures maisons d'édition). Nous avons choisi la deuxième solution, qui a été inscrite dans les statuts de l'association, avec élection directe des membres du comité scientifique tous les deux ans par les membres de l'association²⁰. Quant aux procédures d'évaluation, l'HyperNietzsche gère les procédures de vote, d'écriture et de lecture des rapports et de publication ou de rejet des contributions qui sont évaluées en double anonymat : les membres du comité scientifique ne connaissent pas l'identité des auteurs, les auteurs ne savent pas qui parmi les membres du comité scientifique a évalué leurs contributions. Le passage d'HyperNietzsche à Nietzsche Source (voir *infra*) nous donnera probablement l'occasion de repenser le modèle d'évaluation en introduisant une plus grande participation des membres de l'association HyperNietzsche et des lecteurs.

3. Stabilité et durabilité du texte et de l'adresse Web

La possibilité de citer des documents de manière stable et durable est une des conditions essentielles de la recherche scientifique. En effet, la science est une conversation basée sur des arguments, des démonstrations, des preuves et dans les sciences humaines ces preuves sont souvent contenues dans des documents. Le fait qu'Emma Bovary ingère de l'arsenic est au moins aussi sûr et incontestable que l'existence des électrons, mais pour vérifier ce fait il faut pouvoir citer l'édition originale de *Madame Bovary*. Sans la possibilité de citer, il n'est pas possible de vérifier les faits, ni de produire des interprétations ou d'en discuter publiquement.

Le système éditorial traditionnel est normalement en mesure d'assurer la stabilité du texte et des données bibliographiques, sauf dans les cas particuliers

«Come si valuta la qualità nella Repubblica della Scienza? Una riflessione sul concetto di peer review», *Bollettino telematico di filosofia politica*, juillet 2007, <http://purl.org/hj/bfp/164>.

20. Les statuts de l'association HyperNietzsche sont publiés dans Ph. Chevet, «Le statut juridique de l'HyperNietzsche», *HyperNietzsche, op. cit.*, p. 75-104.

où deux livres avec les mêmes références bibliographiques contiennent des textes qui sont partiellement différents et, au contraire, deux textes identiques apparaissent sous des données bibliographiques différentes. Sur le Web, en revanche, la citation est une entreprise difficile voire désespérée parce que les textes aussi bien que leurs adresses sont habituellement très instables. Avec le Web, pour la première fois, tout le monde a eu la possibilité de diffuser des contenus de manière globale et, en même temps, de les modifier constamment. Le résultat est un formidable instrument de communication, vivant et décentralisé, où les pages changent tous les jours, disparaissent, apparaissent sous d'autres noms et d'autres adresses : phénomènes troublants dans l'univers de la transmission du savoir scientifique. Mais comme qui peut le plus peut le moins, il n'est pas impossible de concevoir des applications Web qui forment un espace dans le Web où les adresses et les textes restent immuables. Pour ce qui concerne la stabilité des adresses, nous pouvons citer des initiatives comme DOI (*digital identifier for any object of intellectual property*) ou PURL (*Persistent Uniform Resource Locator*²¹), sans oublier que déjà les URL (*Uniform Resource Locator*) ou, pour être plus précis, les URI (*Uniform Resource Identifier*) sont capables d'identifier des documents de manière stable et univoque²². La vérification de l'intégrité et de la stabilité du texte, à son tour, peut être effectuée par des algorithmes de somme de contrôle (*checksum*). Dans les deux cas toutefois, il faut être conscient que les solutions techniques ne suffisent pas à assurer la stabilité du texte et de l'adresse Web si elles ne sont pas accompagnées par la volonté explicite des responsables des sites sur lesquels sont enregistrées les données : il ne s'agit pas d'une question technique, mais d'une décision de politique scientifique.

Pour cette raison, le projet HyperNietzsche ne se limite pas à proposer des solutions techniques, mais aussi des structures d'organisation et d'autogouvernement des chercheurs à travers la création de communautés savantes sur le Web selon le modèle ancien des Académies des sciences qui ont accompagné la naissance de la science moderne. Il s'agit de libres associations de savants qui travaillent sur un auteur ou sur un domaine particulier et qui garantissent à la fois la qualité des publications, la paternité intellectuelle et la stabilité des textes publiés dans les plateformes Hyper qu'elles gèrent directement. Pour faciliter ce processus, l'HyperNietzsche s'est fait promoteur d'une action au niveau européen qui vise à soutenir d'un point de vue scientifique, technique, et juridique, la formation de communautés savantes sur le Web²³.

21. <http://www.doi.org>; <http://www.purl.org>.

22. Voir *Architecture of the World Wide Web*, W3C Recommendation 15 December 2004, <http://www.w3.org/TR/webarch/#identification>.

23. Il s'agit de l'action COST A32, «Open Scholarly Communities on the Web. From HyperNietzsche to Hyper-Learning», voir <http://www.cost.esf.org> et <http://www.cost-a32.eu> en particulier les groupes de travail un et trois.

Les réponses que nous avons données demandent toutefois d'être complétées si l'on se place dans une perspective à long terme. Nous travaillons encore aujourd'hui avec des documents originaux qui parfois remontent au Moyen Âge et nous pouvons encore lire les textes écrits par les Grecs et les Romains. Mais pouvons-nous être sûrs que nos documents électroniques nous survivront ? Qu'en sera-t-il de notre serveur, de nos éditions et de nos articles électroniques dans vingt ans, dans cent ans, dans quelques siècles ? Cette question est bien présente dans le débat autour des nouvelles technologies de l'information²⁴ ; pour ma part je voudrais me limiter ici à une simple réflexion de principe. Il est évident que la manière la meilleure de conserver un texte n'est pas d'en déposer l'unique copie dans un endroit « sûr », mais plutôt d'en faire des milliers de copies et de les éparpiller un peu partout dans le monde. *Lots of Copies Keep Stuff Safe* (LOCKSS) est le nom d'un important projet international pour la sauvegarde du patrimoine numérique²⁵. Le principe de LOCKSS est confirmé par trente siècles d'histoire culturelle : tout ce que nous n'avons pas copié, nous l'avons perdu. Pourquoi avons-nous perdu presque toutes les œuvres d'Héraclite et conservé presque toutes celles d'Aristote ? Parce qu'Héraclite avait une politique de DRM (*Digital Rights Management*, gestion numérique des droits) : il confia une copie de son œuvre au temple d'Artémis, le lieu le plus sûr de son époque, afin que ses pensées soient conservées, mais pas copiées et diffusées. Aristote, au contraire, était un tenant du *copyleft* et avait livré tout son savoir à ses élèves qui en firent des copies ; et copie après copie, l'œuvre d'Aristote est arrivée jusqu'à nous. Cela n'est qu'une boutade, mais il est vrai que les pires ennemis de la conservation des textes sont le *copyright*, les DRM, les formats propriétaires et tous les moyens juridiques ou techniques d'interdiction de la copie ; tandis que ses meilleurs amis sont le libre partage, le *copyleft*, les formats ouverts.

Pour résumer : la durabilité des contenus électroniques est une question technologique qui ne concerne pas tellement le type de support, mais plutôt le format des données, sachant que les formats ouverts auront plus chances d'être lisibles à l'avenir ; mais surtout elle est une question juridique, car les contenus soumis à des licences qui permettent la copie auront davantage de chances de se multiplier, de se propager et de survivre. Pour cette raison, HyperNietzsche utilise uniquement des formats ouverts et des licences ouvertes, aussi bien pour les logiciels que pour les documents. Le fait que tout le monde puisse dupliquer librement le contenu du serveur augmente la

24. Pour un panorama de la question, on peut consulter : U. M. Borghoff, P. Rödiger, J. Scheffczyk, L. Schmitz, *Long-Term Preservation of Digital Documents : Principles and Practices*, Berlin-Heidelberg, Springer, 2006.

25. Voir <http://www.lockss.org>.

probabilité que dans mille ans, de copie en copie, quelque part dans le monde on aura gardé au moins une partie de l'HyperNietzsche ; et les formats ouverts sont la meilleure garantie que les données seront encore lisibles.

4. Contextualisation dynamique

À côté des valeurs et des procédures dont nous venons de parler, nous trouvons également dans notre petite île HyperNietzsche quelques innovations techniques. Je ne vais pas en donner les détails ici, mais je mentionnerai seulement une modalité d'organisation de l'information et de navigation dans le site qui est à la base du projet : la « contextualisation dynamique²⁶ » (appelée également mise en contexte hypertextuelle). Tandis que l'utilisateur navigue dans le site, passant d'une page à l'autre, cette forme de navigation savante rend disponibles les contributions scientifiques qui concernent la page sélectionnée. Par exemple, si l'utilisateur sélectionne une page manuscrite, le système rend immédiatement accessible, sans nécessiter des recherches supplémentaires compliquées, la liste de tous les fac-similés, transcriptions et traductions disponibles concernant cette page, de même que tous les chemins génétiques et les essais critiques qui s'y réfèrent. De la même façon, si l'utilisateur sélectionne un essai critique, le masque de contextualisation présentera une liste d'hyperliens vers tous les textes et les manuscrits de Nietzsche cités par l'essai ainsi que la liste de tous les articles d'autres auteurs cités, ou qui citent l'essai consulté. En outre, les systèmes Hyper sont conçus de manière à pouvoir communiquer entre eux à l'intérieur de la Fédération des Hyper. Imaginons un chercheur qui, dans un essai sur Nietzsche contenu dans l'HyperNietzsche, cite un passage de Wittgenstein. Le lecteur doit être en mesure de passer, par un simple clic, de l'HyperNietzsche à l'HyperWittgenstein, en suivant un lien croisé qui lui permette de retrouver le contexte originel de la citation, de la voir commentée par des spécialistes de Wittgenstein ou de remonter jusqu'aux brouillons manuscrits.

Ce principe n'est pas encore complètement implémenté dans la version actuelle d'HyperNietzsche et il a été d'ailleurs à l'origine de nombreux retards technologiques dans le développement du site²⁷. Pourtant, nous considérons qu'il est nécessaire à la construction de notre infrastructure, si elle veut être en mesure de répondre aux exigences de la recherche. Pourquoi le système traditionnel était-il si performant ? Grâce à deux instruments : le classement et la

26. P. D'Iorio, « Nietzsche Open Source », § 6, *HyperNietzsche, op. cit.*, p. 21-36.

27. Pour une description du modèle informatique, voir M. Barbera et R. Giomi, « The Pearl-Diver Model. The HyperNietzsche Data Model and its Caching System », *Jahrbuch für Computerphilologie*, 5, Paderborn, Mentis Verlag, 2004, p. 191-200, <http://computerphilologie.tu-darmstadt.de/jgo3/barbera-giomi.html>.

référence. Tandis que le classement forme un système d'identifiants simples et stables, la référence est le système de navigation d'un chercheur dans cette ontologie. Songez un moment à un chercheur dans une bibliothèque de l'ère pré-électronique. Que fait-il ? Il cherche un livre ou une revue dans le catalogue ou dans une bibliographie, il prend un livre sur l'étagère, consulte un dictionnaire ou un thésaurus. Mais quel est son système de navigation ? Quand on a transposé la bibliothèque, son catalogue, le thésaurus dans un environnement électronique, on a pensé à utiliser des systèmes d'indexation et de recherche documentaire pour relier tous les documents entre eux. Mais cela n'est pas suffisant. Normalement, un universitaire pratique 20 % de recherche documentaire et 80 % de navigation contextuelle, c'est-à-dire de lecture. On a songé à encoder le document avec une grille conceptuelle de métadonnées. Mais une seule grille est insuffisante pour les nécessités d'un spécialiste qui, au cours de sa lecture, suit une vingtaine de grilles conceptuelles en même temps et essaie de développer sa propre grille d'analyse. Un chercheur, en réalité, navigue dans un ensemble de documents grâce à deux critères : la structure des documents (par exemple l'ordre des pages dans un livre ou dans un manuscrit) et les relations entre documents indiquées par d'autres spécialistes à travers les références. La contextualisation dynamique pourrait être la forme que prendra le référencement savant à l'époque numérique, plus puissante que l'ancienne car bidirectionnelle (le système est en mesure non seulement de pointer vers un lieu textuel mais aussi de remonter dans le sens contraire, à l'origine de la référence), et dynamique (la liste des référencements reçus se met automatiquement à jour sans besoin de dépouiller manuellement les revues, comme c'était le cas dans le *Science Citation Index*). Cela permet, entre autres, d'établir automatiquement des *relevés bibliométriques* sans avoir recours à un ensemble de *core journals* définis arbitrairement et à dépouiller manuellement. Si ce système se généralisait, ce serait le libre jeu des citations des chercheurs, recensées automatiquement à l'intérieur du réseau, qui déciderait de la réputation des chercheurs et non pas un petit nombre de *core journals* forcément généralistes et souvent en langue anglaise comme cela arrive, par exemple, dans le cas du *European Reference Index for the Humanities* promu par l'*European Science Foundation*. L'utilisation des descripteurs RDF (Web sémantique) pour la construction du réseau donne en outre la possibilité de peaufiner la détermination du facteur d'impact d'un article en distinguant par exemple les mentions favorables des mentions négatives.

HyperNietzsche, somme toute, est un projet conservateur. Avant d'introduire de nouvelles fonctions, je pense que les infrastructures de recherche en réseau devraient être capables de reproduire dans ce nouveau *medium* la structure de la recherche telle qu'elle s'est construite dans les deux derniers millénaires.

Mais je dois dire qu'une des difficultés principales pour la réalisation du projet a été de comprendre avec clarté et cohérence – pour nous l'expliquer d'abord à nous-mêmes et ensuite à nos collègues informaticiens – ce que nous, les humanistes, nous faisons tous les jours.

D'HyperNietzsche à NietzscheSource

La mise en contexte hypertextuelle est un concept cohérent et rigoureux, mais difficile à représenter dans une interface intuitive et facilement navigable. Le problème principal de toutes les versions d'HyperNietzsche était donc la difficulté de navigation : les lecteurs n'arrivaient pas à comprendre comment avoir accès à un fac-similé, à la transcription d'un manuscrit de Nietzsche ou comment visualiser un rhizome de la genèse, et parfois avaient même des difficultés à lire un essai. Pour cette raison, à partir de la version 0.4, nous avons introduit toute une série de «vues» qui permettaient de suspendre de manière temporaire la contextualisation afin de naviguer avec plus d'agilité. Fin 2007 nous avons finalement décidé de changer radicalement l'interface du site et de marquer ce tournant avec un changement de nom : de HyperNietzsche à NietzscheSource.

D'abord le changement de nom : «hypertexte» a toujours été un concept vague et nébuleux et les tentatives de le rendre plus rigoureux n'ont pas eu un grand succès. Dans les années 1990, il était nouveau et séduisant ; désormais il n'est plus que fumeux et vieilli. «Source», au contraire, est un concept clair, antique et toujours actuel, qui accompagne toute activité de recherche. Il a également un sens technologique (*code source*) et politique (*Open Source*), mais son sens principal concerne les sources de la connaissance en général et les sources philologiques en particulier. Il renvoie au caractère concret et documenté d'une recherche et il indique ainsi que dans les sites Web qui portent ce nom se trouvent les sources primaires et secondaires utiles voire indispensables à tous ceux qui veulent approfondir l'étude de la vie et de la philosophie d'un auteur ou d'un autre sujet.

Nous n'avons cependant pas affaire à un simple changement de nom, l'objet lui-même a été repensé : l'interface a été complètement redessinée en séparant la navigation de la contextualisation et en divisant le site en deux parties, appelées «mode simple» et «mode savant». Le premier s'adresse à des consultations rapides d'un texte, d'un fac-similé, d'un essai, et on y évolue dans des interfaces simples où tous ceux qui ont une pratique normale du Web devraient se sentir à l'aise ; la navigation y est intuitive car elle utilise des modèles standard, réduit les fonctions au minimum et ne prévoit pas de contextualisation. Le mode savant, en revanche, permet une comparaison

savante des différents documents et commentaires concernant un certain lieu textuel grâce à un masque synoptique divisible en plusieurs colonnes. Il dispose également d'un système de navigation dans le catalogue des sources primaires basé exclusivement sur les sigles standard utilisés par les spécialistes et d'un masque de recherche avancée qui permet d'extraire des informations en fouillant dans tous les types de documents publiés sur le site. De chaque page en mode simple un lien permet de passer à la page correspondante en mode savant, et vice-versa.

NietzscheSource introduit en outre le concept de macrocontributions. Pour expliquer ce qu'est une macrocontribution, rappelons brièvement le concept de contribution scientifique tel qu'il avait été pensé par le projet HyperNietzsche. Prenons le cas d'une édition critique imprimée : elle contient normalement des contributions de différentes sortes – éditions de textes, commentaires philologiques et critiques, références croisées, références bibliographiques, essais introductifs, etc. – et ces contributions coexistent de manière indifférenciée sur le support papier. Le support électronique, en revanche, permet de différencier les divers types de contributions scientifiques pour mieux les extraire et les assembler selon des configurations différentes, par exemple pour chercher les éditions d'un même texte contenues dans différentes éditions critiques, pour disposer sur un diagramme toutes les références croisées qui concernent une certaine œuvre de Nietzsche, ou pour séparer les commentaires philologiques des commentaires philosophiques. HyperNietzsche avait donc entrepris un recensement des types de sources primaires et secondaires utilisées pour la recherche sur Nietzsche, et avait établi une ontologie bibliographique correspondante. Ce que nous n'avions pas compris est que cette manière d'organiser la connaissance, qui est excellente pour ce qui concerne la structure de la base de données sous-jacente au site Web, ne peut pas servir pour construire une interface de navigation acceptable. Bref, après avoir «démonté» la machine de la science dans ses parties constitutives – qui restaient cachées dans la forme du livre –, il fallait maintenant la réassembler de manière à ce que les lecteurs ne soient pas dépaysés, retrouvent les formes familières de transmission du savoir, et puissent naviguer dans une interface simple et intuitive. L'instrument que nous utiliserons dans NietzscheSource pour cette tâche est le concept de macrocontribution.

Une macrocontribution est formée par un ensemble de contributions homogènes scientifiquement corrélées. Les macrocontributions ont une double fonction : d'une part elles permettent une navigation plus intuitive au sein d'un ensemble cohérent de documents. D'autre part elles permettent d'éviter l'effet «en taches de léopard» de l'HyperNietzsche, où le lecteur ne savait jamais quelles étaient les parties du site déjà remplies par les contributions des

chercheurs et lesquelles restaient «à remplir», et passait son temps à naviguer dans des listes de sigles accompagnées de quelques fac-similés. NietzscheSource gardera intacte l'ontologie bibliographique de l'HyperNietzsche au niveau de la base de données, qui sera visible et également exploitable pour des comparaisons et des recherches sophistiquées dans le mode savant. Mais, dans le mode simple, NietzscheSource utilisera le concept de macrocontribution pour recréer, en les améliorant si possible, les formats traditionnels de la communication scientifique. Nous prévoyons d'utiliser les six macrocontributions suivantes, qui correspondront à autant de modules de l'interface :

1) *L'édition critique Colli/Montinari* est l'édition de référence pour les œuvres, les fragments posthumes et les lettres de Nietzsche. Puisque le droit d'auteur allemand prévoit que les éditions critiques tombent dans le domaine public vingt-cinq ans après leur publication, nous avons maintenant la possibilité de la mettre à disposition de tous les chercheurs avec une interface qui permette de lire le texte, de le citer, de chercher des mots ou des phrases dans toute l'édition ou dans certaines parties, de visualiser et d'imprimer facilement des listes d'occurrences. Le texte est identique à l'édition imprimée et corrige de nombreuses coquilles présentes dans l'édition des œuvres et des fragments posthumes disponible en CD-ROM.

2) *L'édition fac-similé D'Iorio* permet d'accéder pour la première fois à l'intégralité du legs de Nietzsche conservé à Weimar : cahiers manuscrits, premières éditions des œuvres, lettres, documents biographiques, iconographie. Le lecteur pourra feuilleter les fac-similés, les agrandir, les faire tourner et bien sûr les imprimer ou les télécharger.

3) *L'édition génétique D'Iorio* permet d'explorer deux œuvres de Nietzsche, le *Voyageur et son Ombre* et *Aurore*, en prenant en compte deux types de processus génétiques : la genèse synchronique et la genèse diachronique. La genèse synchronique se réfère aux strates d'écriture présentes sur la même page, que cette édition permet de cerner et de représenter à l'écran. La genèse diachronique suit le développement d'un texte de la première ébauche dans un carnet, à travers les cahiers, le manuscrit pour l'imprimeur, et les épreuves jusqu'à la version publiée.

4) La partie des *essais critiques* sera retouchée, mais sans modifications profondes.

Ces quatre premiers modules seront disponibles en 2008 tandis que les deux derniers seront ouverts au public en 2009.

5) Le catalogue de la bibliothèque et des lectures reproduira la version imprimée²⁸, mais essaiera dans les années à venir de la compléter, en collaboration avec la *Herzogin Anna Amalia Bibliothek* de Weimar, avec les fac-similés des livres de la bibliothèque de Nietzsche.

6) Les modules pour les bibliographies disposeront des fonctionnalités standard d'interrogation d'une base de données bibliographiques.

Vers ScholarSource

1. Cercles

NietzscheSource sera développé à l'intérieur du projet européen *Discovery*²⁹ et représentera le noyau d'une structure composée par deux cercles concentriques.

Le *premier cercle* est formé par un ensemble de corpus philosophiques qui, en utilisant la même infrastructure que NietzscheSource, donneront naissance à un réseau philosophique appelé Philosource. Conçu pour être continuellement enrichi par les travaux des chercheurs, Philosource sera fourni dès le départ d'une masse critique de matériaux dont nous donnons ici un bref aperçu :

1) *Grèce ancienne*. Trois grands corpus seront publiés dans Philosource par l'*Istituto per il Lessico intellettuale europeo e Storia delle idee* du CNR de Rome :

a. l'édition complète des fragments des présocratiques basée sur *Die Fragmente der Vorsokratiker* de Diels et Kranz, en grec avec traductions allemande et italienne.

b. l'édition complète de tous les témoignages sur Socrate, basée sur les *Socratis et Socraticorum Reliquiae* de Gabriele Giannantoni, en grec.

c. l'édition complète des vies des philosophes de Diogène Laërce, en grec avec traduction italienne.

2) *Textes philosophiques et scientifiques de l'époque moderne*. Le CNR de Rome publiera en outre un ensemble de 52 œuvres en latin, italien, français d'auteurs modernes comme Giordano Bruno, Descartes, Spinoza, Leibniz, Vico, Baumgarten, Kant.

3) *Le dix-neuvième et le vingtième siècle* seront représentés par Nietzsche et Wittgenstein. Les archives Wittgenstein de Bergen (Norvège) publieront dans

28. G. Campioni, P. D'Iorio, M. Cr. Fornari, F. Fronterotta, Andrea Orsucci, unter Mitwirkung von R. Müller-Buck, *Nietzsches persönliche Bibliothek*, Berlin-New York, De Gruyter, 2003.

29. Lancé en octobre 2006 et prévu pour une durée de trois ans, le projet *Discovery, Digital Semantic Corpora for Virtual Research in Philosophy*, relève du programme *eContentplus* de la Commission européenne et rassemble six partenaires technologiques et philosophiques ; voir <http://www.discovery-project.eu>.

Philosource 5 000 pages de manuscrits du legs de Wittgenstein, comprenant des matériaux venant du *Big Typescript* (1929-1934), du *Brown Book* (1934-1936), de la *Lecture on Ethics* et des *Notes on Logic*.

4) *Philosophes contemporains*. La *Radiotelevisione Italiana* (RAI) publiera dans Philosource 300 interviews audio-visuelles de philosophes contemporains comme Hans Georg Gadamer, Gilles Deleuze, Gianni Vattimo, et autres.

Le *deuxième cercle* est représenté par une série d'autres projets de sites Web consacrés à des philosophes comme Schopenhauer, mais aussi à des écrivains comme Flaubert, Proust, Valéry, Virginia Woolf, Beckett, Yeats, qui sont réunis dans l'action COST «Communautés Savantes sur le Web» et qui visent à former un réseau plus vaste, appelé ScholarSource, ouvert à différentes disciplines.

2. Rêves autour de Wikipedia

Wikipedia, comme son nom l'indique, est structuré comme une encyclopédie. Les principes de son fonctionnement sont bien connus : tout le monde peut rédiger un nouvel article ou modifier les articles des autres ; pas de contrôle de qualité ; pas de stabilité du texte ; absence complète de la notion d'auteur. D'après ces caractéristiques, il est aisé de conclure que Wikipedia, tout en étant un instrument formidable pour animer une communauté, est néanmoins profondément inapproprié en tant que moyen pour accueillir le savoir produit par une communauté «savante». Ou alors il faut penser que la recherche deviendra quelque chose de très différent par rapport à ce qu'elle a été dans les siècles derniers, tant du point de vue de la validation du savoir que de celui de l'organisation, de la progression des carrières des chercheurs, etc.

Ce modèle de diffusion du savoir peut nous faire sourire, et tout bon chercheur en sciences humaines le considérera d'un regard hautain. Ce n'est pas là que nous trouverons les sources du travail scientifique et que nous publierons les résultats de nos recherches. Mais alors, demandons-nous, nous les savants, nous les dépositaires du savoir certifié, évalué, bien structuré, destiné à durer dans les siècles ; demandons-nous : quel est notre modèle de diffusion de nos savoirs ? C'est un livre publié deux ans après la conclusion de la recherche, distribué en 300 exemplaires, en payant 4 000 euros d'aide à la publication ? Le problème, bien sûr, n'est pas le livre, que nous tous aimons et qui est encore le meilleur support pour la lecture. Le problème, ce sont les deux ans, les 300 exemplaires et les 4 000 euros. Face à l'incroyable efficacité de la diffusion des savoirs démontrée par Wikipedia, par les blogs, par toutes sortes de communautés sur le Web, qu'avons-nous à proposer ? Disposons-nous d'un modèle

qui sauvegarde la complexité et la structuration nécessaires au savoir scientifique, qui tout en étant global et ouvert assure l'évaluation par les pairs, sauvegarde la paternité intellectuelle, garantit la stabilité du texte et dispose d'un système de navigation plus sophistiqué que les listes d'occurrences ou les articles d'encyclopédie? Jusqu'à maintenant l'*humanities computing*, au lieu de concevoir une nouvelle infrastructure de recherche capable d'utiliser le nouveau *medium* électronique dans tout son potentiel, n'a produit qu'une nouvelle discipline de niche et un ensemble de projets non coordonnés les uns avec les autres. Devons-nous y voir le signe d'un destin des sciences humaines qui seraient réfractaires à jamais aux grands projets de coordination et inexorablement condamnées à la création de nouvelles niches?

Wikipedia est vivante : avec ses 2 200 000 articles en anglais, 700 000 en allemand, 600 000 en français, environs 400 000 en polonais, japonais, italien, néerlandais... pour un total de 9,5 millions d'articles en 250 langues différentes; avec une croissance moyenne des contenus de 50 % par an et environ 35 000 lecteurs par jour, elle est source d'accès au savoir pour des millions de personnes. Si les chercheurs en sciences humaines voulaient encore avoir une fonction en nos temps inquiets, ce serait peut-être celle de diffuser leurs savoirs choisis et de le faire peut-être avec le *medium* le plus puissant qui existe, le Web, en s'efforçant de créer un instrument aussi vivant, utile, consulté et alimenté que l'est Wikipedia; sans renoncer aux livres, étourdis par la puissance du Web, mais sans renoncer au Web, aveuglés par l'amour des livres. Renoncer au numérique signifierait probablement condamner les sciences humaines non pas à la mort, mais à l'insignifiance, au musée, au silence. La recherche deviendrait une activité de plus en plus éloignée de la réalité de l'échange et du partage du savoir; et inversement l'information diffusée sur le Web ne bénéficierait nullement de l'apport du savoir critique.

Le Web est une grande opportunité pour nos disciplines et probablement la seule manière de préserver les valeurs et les méthodes que nous avons adoptées et perfectionnées en deux mille ans d'histoire intellectuelle, de la Bibliothèque d'Alexandrie à nos jours. Car nous devons les préserver dans le mouvement et non pas les préserver du mouvement : nous avons certes besoin de notre île bienheureuse, et même d'un archipel d'îles bienheureuses, mais immergées dans la mer vaste et mouvante du Web.