

Hypertexte : de l'accès aux documents vers leur structuration

Eric Bruillard

IUFM de Basse Normandie et GREYC

eric.bruillard@caen.iufm.fr

Introduction

Le programme Colisciences conduit à mettre à disposition un corpus numérisé et indexé de textes de grands auteurs du XIX^{ème} siècle ayant contribué à la naissance de la science moderne du vivant. L'indexation réalisée permet la mise au point de différents parcours sémantiques de navigation pour faciliter la recherche d'informations et la « compréhension » des idées qui y sont contenues. La question se pose, à la suite de premières expérimentations, d'imaginer et de tester des contextes et des situations d'apprentissage intégrant ce dispositif.

N'ayant pas travaillé de manière approfondie sur ce corpus, mon propos se limitera à donner un point de vue sur les questions d'apprentissage et d'hypertexte puis à apporter quelques éléments de réflexion et à livrer des pistes de travail.

En première approche, intégrer un tel dispositif dans un contexte éducatif « classique » m'apparaît comme une tâche quasiment insurmontable. En effet l'offre Colisciences est complexe tant du point de vue des contenus (des textes du siècle dernier difficiles à lire) que des activités susceptibles d'y être associées et à leur caractère interdisciplinaire presque obligatoire. Mais, l'intérêt de la notion même d'hypertexte, maintenant beaucoup décrite mais à mon sens encore peu comprise, justifie amplement de s'y plonger.

Dans une première partie, je vais m'attacher à identifier certains obstacles. D'abord la conception actuelle dominante de l'hypertexte, plus axée sur l'accès à des informations et à des documents qu'à leur organisation et leur structuration. Ensuite, sur le manque d'instrumentation réellement disponible pour aider à cette structuration. Le risque est de se restreindre à de simples tâches de recherche d'informations dans un cadre éducatif, s'éloignant de l'intérêt fort de l'hypertexte en éducation. Cela me permettra alors de souligner les enjeux d'un travail éducatif sur l'hypertexte et de suggérer des pistes d'expérimentation.

L'hypertexte : une notion à réinvestir

L'hypertexte est une notion relativement ancienne qui s'est imposée dans les champs d'usage des technologies depuis une quinzaine d'années. De nombreuses activités ont été proposées et testées, des produits se sont diffusés, notamment Hypercard en 1987, qui a fortement contribué à populariser la notion d'hypertexte, en l'associant à un environnement de programmation.

Mais le formidable succès du web a changé les choses. Cela a consacré la notion d'hypertexte, mais en la trahissant d'une certaine manière, en la réduisant, dans l'esprit de bon nombre d'utilisateurs, à une simple technologie d'abord de mise à disposition puis d'accès à des informations ou des documents. En effet, si l'hypertexte peut se définir comme un dispositif informatisé permettant l'interconnexion de documents de divers types, non sur la base d'un modèle hiérarchique ou relationnel, mais par des mécanismes associatifs sous contrôle de l'utilisateur (la navigation), ce n'est pas ainsi qu'il est en général véritablement compris.

Ainsi, quand on demande à des enseignants, pourtant utilisateurs au moins occasionnels si ce n'est réguliers d'internet, d'expliquer cette notion, ils en donnent rarement une définition cohérente. Le plus souvent, ils font comme les enfants, ils se placent mentalement dans une situation habituelle mais fictive et disent ce qu'ils font. En gros, « je clique et j'accède à un nouveau document ». Cela se résume à une action dans une perspective locale : des pages sont reliées entre elles et on se « déplace » de l'une à l'autre en cliquant sur des mots ou expressions mises en exergue ou des zones de l'écran « réactives ». En fait, la vision dominante n'est pas tant celle d'un dispositif dans sa globalité que d'une action ponctuelle. Cela n'existe que dans l'*ici et maintenant* de la consultation.

Ce phénomène a certainement de multiples causes, mais il est renforcé par la manière dont se présentent les systèmes d'exploitation des ordinateurs actuels et la généralisation de l'accès à des ressources et des programmes par simple clic sur des icônes (interfaces graphiques avec manipulation directe, du moins en partie). Si cela facilite indubitablement la mise en oeuvre par des usagers, surtout s'ils sont novices, cela peut conduire, sans formation complémentaire, à des blocages difficilement franchissables. On en reste au niveau sensori-moteur (de la perception immédiate) sans pouvoir développer le moindre langage, on a besoin de voir (s'aider de ce que montre l'écran) pour refaire les actions faites, sans pouvoir se construire de représentation globale cohérente (voir Normand et Bruillard, 2001).

Dans ce contexte, on peut rechercher des informations, on est loin d'avoir l'idée que l'hypertexte fournisse différentes structurations et puisse aider à en développer de nouvelles. C'est une technologie d'accès, pas de structuration.

A ce propos, on peut réfléchir à un glissement de sens sans doute récent autour de la notion de transparence. Auparavant, rendre transparent consistait à permettre de regarder à l'intérieur pour voir un processus afin de le comprendre. C'était la métaphore de la boîte de verre, opposée à la boîte noire, dont on ignorait le fonctionnement. Ce sens est conservé avec l'idée de traçabilité, par exemple pour des produits alimentaires, afin de rassurer le consommateur sur la provenance de produits qu'il pourrait juger suspects. L'informatique actuelle consacre une idée diamétralement opposée de la transparence : on rend transparent, c'est-à-dire on cache, ce qui n'est pas jugé utile pour que l'utilisateur n'ait pas à s'en préoccuper et qu'il puisse se concentrer sur ce qu'il désire faire. On développe notamment la notion de transparence « opérative » : on ne donne à voir que ce qui est jugé nécessaire à l'accomplissement d'une tâche. Dans un but d'apprentissage, décider ce qui est utile ou non n'est jamais aussi simple mais surtout, lorsque cela ne marche pas comme prévu, l'utilisateur est démuni, il ne dispose d'aucun moyen pour comprendre ce qui a pu se passer et la transparence apparaît comme un obstacle majeur à la compréhension. La transparence, dans ce dernier sens, me paraît être une illusion dangereuse.

Pour en revenir à l'hypertexte, les différentes manières d'accéder à des fragments de texte, à reconstituer des développements, en suivant et/ou en créant des parcours multiples doivent être prises en compte et non cachées, dans une transparence favorisant la traçabilité et l'intelligibilité.

Notons que cette idée correspond bien à l'hypertexte tel qu'il se développe depuis son origine. Pas tant un instrument de recherche, un outil d'accès à des informations qu'une manière de les lier entre elles et de favoriser la construction de structures de différentes natures.

Si l'histoire de l'hypertexte est bien documentée (voir par exemple Bruillard, 1997, chapitre 6 et les quelques références en anglais en fin d'article), c'est bien celle d'ingénieurs et de chercheurs voyant dans la technologie des moyens d'augmenter les capacités humaines. Du *Memex* (*Memory Extender* ou extenseur de mémoire) de Vannevar Bush au grand projet Xanadu de Ted Nelson, en passant *NLS / Augment* (système en ligne d'*augmentation*) de

Douglas Engelbart, il s'agit de faciliter l'accès, l'organisation et la structuration de documents liés sémantiquement entre eux, pour ceux qui les lisent et les interprètent. Que ce soit le chercheur (Bush), les « travailleurs de la connaissance » (Englebart après Licklider) ou tout le monde pour Nelson.

L'association d'idées (Bush), le travail coopératif (Engelbart), l'interconnexion des notes personnelles (Nelson) font partie intégrante des réflexions autour de l'hypertexte et se concrétisent dans des associations sémantiques et des chemins, souvent tracés ou balisés par des experts.

Dans un texte un peu polémique, Nelson (1967) reproche aux experts et praticiens de la recherche documentaire d'avoir déformé les idées initiales de Bush en ne retenant que les questions d'indexation. Cette tension est de nouveau d'actualité avec la pression très forte autour de questions de normalisation. Se mettre d'accord au niveau mondial sur des formes de description de ressources (cf par exemple Grandbastien, 2002, pour ce qui concerne les ressources éducatives) est un enjeu important, mais cela tend à laisser de côté la question plus complexe des formes de définition de parcours d'interprétation humains, l'un des aspects régulièrement oublié de l'hypertexte. Ce n'est pas le moindre mérite du projet Colisciences de nous le rappeler.

Un indexation, aussi fine et sophistiquée soit-elle, est-elle suffisante ? Rechercher des paragraphes particuliers est certes important, mais c'est loin de suffire pour reconstituer la « pensée » d'un auteur. C'est dans cette direction qu'il faut sans doute réfléchir aux apports principaux de la notion d'hypertexte.

Créer et structurer des chemins

Les environnements actuels de traitement de l'information permettent de conduire des activités nouvelles de type empirique voire quasi expérimental, de confronter des théories et des données même dans des disciplines sans rapport historique avec l'expérimentation (comme la littérature ou les langues étrangères). Parce que le système permet de mémoriser de grands ensembles de données et des suites d'actions, il devient possible de faire des manipulations complexes sur de grands corpus, de faire varier de manière volontaire et réversible certains paramètres et d'observer les résultats de différents points de vue. Un des risques de ce type d'approche est évidemment celui du "zapping", de la randonnée aléatoire dans une masse confuse de données dont le sens ne pourrait être construit faute d'objectifs réalistes ou de représentations adéquates (Baron et Bruillard, 2001). C'est dans ce domaine que la maîtrise des concepts liés à l'hypertexte s'avère certainement essentielle.

Si on s'accorde sur l'idée que l'hypertexte ne peut se limiter à faciliter l'accès à des ressources, il faut essayer d'identifier ce qui permet d'aller plus loin. Pour cela, il s'agit sans doute de favoriser des processus de compréhension et d'analyse en multipliant les objectivations, les représentations, notamment sous forme graphique. Pour cela, deux éléments sont à prendre en compte :

- Faciliter toutes les possibilités de visualisation globale du corpus
- Aider à garder, représenter et organiser les parcours en essayant de traduire le dynamique (temporel) par du spatial

Sur le premier point, donner des formes d'appréhension globale d'un corpus apparaissent importantes en permettant d'y projeter des éclairages, des points de vue. Visualiser tous les paragraphes ayant une caractéristique donnée (traitant d'une notion, explicitant une relation, etc.), le spatial donnant à la fois le quantitatif et la localisation, avec une facilité d'accès au contenu. Donner des modes d'interrogation sophistiqués : quelles sont les relations les plus

nombreuses, les plus rares, etc. Penser aussi à la visualisation des relations en jouant sur les couleurs, etc. De nombreuses possibilités sont encore à explorer.

Sur le second point, représenter un chemin sous la forme d'un objet manipulable est une piste essentielle à creuser. Conserver l'historique d'une consultation est une première étape, il faut également pouvoir le rejouer, l'organiser, insérer des notes, etc. Le logiciel ARI (assistant de recherche internet), développé à l'INRP, fournit un exemple d'outil à développer.

Il faut aussi pouvoir comparer aisément deux parcours, les projeter sur la visualisation du corpus, les qualifier vis-à-vis des notions présentes et des relations retenues, etc.

Pour résumer, l'idée est de développer une forme d'exercisation (en pensant à des activités se déployant dans un cadre éducatif) proprement hypertextuelle. Alors que dans le cursus scolaire, on passe beaucoup de temps à acquérir des formes de présentation écrite très spécifiques comme la dissertation ou le commentaire composé, il y a certainement d'autres formes nouvelles à développer. En particulier, les formes de catégorisation et de structuration rigides, depuis les tableaux à double entrée, occupent une place importante dans la scolarité. L'informatisation introduit la possibilité de réfléchir et de mettre en place d'autres moyens, plus flous et plus graphiques¹, dans des types d'hypertextualisation à scolariser. Il ne s'agit donc pas de se limiter à inventer quelques activités particulières, mais, de façon plus globale, à réfléchir sur les activités préparatoires possibles dans une perspective génétique tout au long de la scolarité.

Quelques pistes pour une utilisation éducative

Outre ce qui concerne directement les fonctionnalités hypertextuelles, vues de manière un peu indépendante du contenu, des activités centrées sur ce dernier sont également à développer. Considérant la complexité intrinsèque de ce qui est proposé, à la fois en termes de contenu et en termes d'exploitation de celui-ci, ainsi que l'absence d'une tradition d'exercisation dans ce domaine, il paraît impossible de trouver des formes d'utilisation éducative suffisamment proches de ce qui peut être fait dans les formations institutionnelles. Cela conduit soit à imaginer des activités plus simples centrées sur des points précis (par exemple autour de la recherche d'informations et de restitutions demandant peu de retraitement), soit des scénarios plus sophistiqués, sortant des cadres classiques.

Une première idée est celle de la création de parcours. On peut d'abord regarder du côté des experts, essentiellement des enseignants de SVT (sciences de la vie et de la terre), des historiens des sciences et des philosophes. Ces experts sont conduits à bâtir des cours, notamment autour de notions ou de problèmes abordés ou traités dans le corpus indexé dans le cadre du projet Colis. Comparer les chemins construits, correspondant en gros à des visites guidées, et décrire comment sont utilisés les outils proposés (par exemple les relations entre les notions), est une tâche à ne pas négliger. Les utilisations des experts peuvent fournir de premières indications sur les usages possibles par des plus novices et les chemins collectés enrichissent le corpus par l'adjonction de parcours sémantiques finalisés. Dans un premier temps, ce seront peut-être avant tout ces derniers qu'il faudrait proposer aux apprenants, les activités de construction de parcours pouvant être jugés comme étant trop complexes ou trop éloignées des tâches habituelles.

Mais de tels parcours peuvent être effectués par les apprenants, élèves de lycée, étudiants ou apprentis enseignants. Il faut alors les aider à les construire et fournir une instrumentation adéquate. Le logiciel ARI déjà cité constitue un premier exemple, permettant de conserver l'historique et d'en extraire et d'en commenter un chemin spécifique. Si l'on veut aller plus

¹ Notamment les cartes de concepts et les logiciels tel Inspiration (<http://www.inspiration.com>) qui permettent de les créer et les modifier.

loin, on peut s'inspirer des travaux autour de l'argumentation (domaine qui a pris ces dernières années une place importante dans les programmes de français tout au long de la scolarité). En particulier, Toulmin et ses travaux sur la logique ont été souvent repris. Fournir des structures argumentatives vides à remplir au cours d'une exploration, puis à réorganiser est une piste à creuser. On peut également penser aux systèmes de visualisation et de structuration tels Aquanet (Marshall & Rogers, 1992 ; Marshall *et al.*, 1994). Pour des activités plus scolaires, ce qui a pu se faire en collège autour de Perseus² (voir Charpin 1994 ; 1995), constitue une source intéressante.

En fait, derrière cela, il s'agit bien de se diriger vers une autre façon de considérer les textes, de développer un autre regard, de voir comment des formes de lecture savante (ou intellectuellement guidées) pourraient être scolarisées. Comment, par exemple, donner un statut visible aux activités d'annotation (souvent considérées comme des pratiques privées et donc peu prises en considération dans le cadre scolaire).

La question posée est celle de la détermination d'activités de parcours et de construction pour faciliter la compréhension et l'analyse. L'indexation mise en place autorise de multiples navigations, de l'exploration sans grande finalité à la recherche d'une réponse à des questionnements très ciblés. Il faut certainement réfléchir dans toutes ces directions. Sans oublier les pratiques collaboratives, qui se développent dans l'éducation.

N'oublions pas que les expérimentations peuvent également avoir une autre finalité, celle d'aider à construire les instruments et à préciser les modes d'indexation intéressants. Des recherches à caractère exploratoire, par exemple sur les modes d'exploration en fonction de relations présentées graphiquement, sont à mener.

Revisiter le concept d'hypertexte apparaît donc fondamental, notamment dans le contexte scolaire actuel où la recherche d'informations, leur organisation et leur réorganisation, dans le cadre d'activités interdisciplinaires, occupe une place croissante. Cela a pour conséquence de prolonger l'instrumentation dorénavant indispensable dans le domaine de la recherche d'informations à celui de leur structuration. Si le lecteur *fait* l'hypertexte, l'architecture proposée de même que les compétences qu'il est apte à mettre en œuvre contraignent les possibles. Cela souligne l'importance d'acquérir une maîtrise suffisante d'instruments d'aide à la structuration d'ensembles d'informations, enjeu de formation au cœur même de la problématique hypertexte. Voyant ce dernier plus comme un moyen d'établir un ensemble de relations sur les textes eux-mêmes plus que d'y accéder, conduit à s'interroger sur la genèse des activités scolaires qui pourrait le mieux y préparer.

Remerciements à Georges Vignaux pour ses remarques constructives sur une première version de ce texte.

Références sur les fondateurs de l'hypertexte

Bush et le Memex

Brève présentation de Vannevar Bush : <http://www.iath.virginia.edu/elab/hfl0034.html>

As we may think : <http://www.theatlantic.com/unbound/flashbks/computer/bushf.htm>

ou <http://www.isg.sfu.ca/~duchier/misc/vbush/>

² <http://www.perseus.tufts.edu/>

Une image du Memex : http://www.kerryr.net/pioneers/memex_pic.htm

Douglas Engelbart

Brève présentation : <http://www.iath.virginia.edu/elab/hfl0035.html>

Démonstration en 1968 : <http://sloan.stanford.edu/MouseSite/1968Demo.html>

Interview pour les enfants :

<http://www.superkids.com/aweb/pages/features/mouse/mouse.html>

Le travail coopératif : <http://www.bootstrap.org/>

Augmenting Human Intellect: A Conceptual Framework (<http://www.histech.rwth-aachen.de/www/quellen/engelbart/ahi62index.html>)

Bush et Engelbart (pour l'histoire des interfaces graphiques).

http://www2.kenyon.edu/people/adamsal/gui/bush_engelbart.htm

Ted Nelson et l'hypertexte

Brève présentation : <http://www.iath.virginia.edu/elab/hfl0155.html>

Projet Xanadu : <http://xanadu.com/>

Sa home page : <http://www.sfc.keio.ac.jp/~ted/>

Ted Nelson et le mot hypertexte en 1965 : http://iberia.vassar.edu/~mijoyce/Ted_sed.html

Memex et au-delà : <http://www.cs.brown.edu/memex/>

Autres références

Baldner Jean-Marie, Bruillard Eric (2000). *L'usage des manuels scolaires et des ressources technologiques dans la classe*. Rapport de la première année de la recherche INRP 40124, 121 p. Première partie reprise en Point de Vue dans *Sciences et Techniques Educatives*, vol. 7, n° 2, Hermès, p. 443-480.

Baron Georges-Louis, Bruillard Eric (2001). Une didactique de l'informatique ? *Revue Française de Pédagogie*, n° 135, p. 163-172.

Bruillard Eric (1997). *Les machines à enseigner*. Editions Hermès, Paris, 320 p.

Bruillard Eric, Grandbastien Monique (eds.) (2001). *Education et informatique. Hommage à Martial Vivet*. Sciences et Techniques éducatives, vol. 7, n° 1, Hermès Science, 300 p.

Bruillard Eric, de La Passardière Brigitte et Baron Georges-Louis (eds.) (1998). *Le livre électronique*. Sciences et Techniques Educatives, vol. 5, n° 4, Hermès Science, Paris.

Bruillard Eric, de La Passardière Brigitte. (1998). Fonctionnalités hypertextuelles dans les environnements d'apprentissage, in Tricot A. et Rouet J.-F. (dir.) *Les hypermédias, approches cognitives et ergonomiques*, Hermès, Paris, p. 95-122.

Charpin François (1994). Analyse littéraire et civilisation avec l'ordinateur, *La pratique de l'informatique dans l'enseignement des langues anciennes*, LITALA, Paris, p. 83-89.

Charpin François (ed.) (1995). *La pratique de l'informatique dans l'enseignement des langues anciennes*, LITALA, Paris.

Grandbastien Monique (2002). Quelques questions à propos de l'indexation et de la recherche de ressources pédagogiques sur le WEB in *Les technologies en éducation : perspectives de recherche et questions vives*, Baron Georges-Louis, Bruillard Eric (eds.), INRP, Paris.

Marshall C., Shipman F., Coombs J. (1994). VIKI: spatial hypertext supporting emergent structure in *Proceedings ECHT'94*, p. 13-23.

- Marshall C.C., Rogers R.A. (1992). Two Years before the Mist: Experiences with Aquanet, *Proc. 4th ACM Conference on Hypertext, ECHT'92*, Milan, p.53-62.
- Nelson Theodor H. (1967). Getting it out of our system in Schecter George (ed.), *Information Retrieval: A Critical View*. Thompson Book Co.
- Normand Sylvie, Bruillard Eric (2001). Que révèlent les discours de futurs enseignants sur leur compréhension du fonctionnement des applications informatiques. Point de vue, *Sciences et Techniques éducatives*, vol. 8, n° 3-4, Hermès Science, p. 435-445.