



HAL
open science

Fabula.org : usages d'un site scientifique littéraire (1999-2001)

Marin Dacos

► **To cite this version:**

Marin Dacos. Fabula.org : usages d'un site scientifique littéraire (1999-2001). <http://www.msh-paris.fr>, 2001. edutice-00000025

HAL Id: edutice-00000025

<https://edutice.hal.science/edutice-00000025>

Submitted on 19 Jun 2003

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Fabula.org : usages d'un site scientifique littéraire (1999-2001)

Rapport rédigé par Marin Dacos

avec la collaboration d'Alexandre Gefen et François Lermigeaux

« *Fabula, ae, f. : 1. propos de table, conversation. 2. Récit sans garantie historique, légende. Fabulae : contes, chansons, sonnettes.* ». Le site Fabula.org s'ouvre sur cette citation qui définit le champ d'étude cultivé par ses fondateurs : la théorie littéraire. Né au début de l'année 1999, ce site de ressources et de recherche en ligne destiné aux chercheurs et visant à la diffusion de l'information scientifique et de la connaissance en matière de théorie et de poétique littéraire ¹ cherche à utiliser le support électronique comme nouvel espace de recherche, de réflexion et de travail collectif. L'observation des pratiques d'écriture et de lecture scientifique issues de ce projet montrent la naissance d'une nouvelle relation des chercheurs avec leur matière. Sans constituer de rupture avec les règles scientifiques élaborées par des siècles d'histoire intellectuelle, elle opère une mutation réelle des pratiques de l'écrit. Cette mutation est caractérisée par une atténuation des distances entre chercheurs et présente la particularité de n'avoir pas encore achevé sa genèse : l'ensemble est toujours en cours de formation et la communauté des chercheurs participe à son invention au quotidien.

GENESE D'UN PROJET ELECTRONIQUE ET SCIENTIFIQUE

Comme bien des innovations, Fabula provient d'une population qui n'occupe pas le centre de la société dans laquelle elle gravite ², ici la communauté scientifique littéraire.

¹ Littérarité, fictionalité, histoire des poétiques, problématiques de la représentation et des genres littéraires, théories de l'auteur, de la réception et de la lecture.

² Everett M. ROGERS, *Diffusion of innovations. Fourth edition*, Paris, The Free Press, 1995, 518 p.

Un groupe informel

Pour la plupart ATER, allocataires de recherche, PRAG ou simples doctorants, les premiers membres de Fabula.org n'avaient pas encore de position définitive dans la communauté scientifique ni obtenu de poste à durée indéterminée dans une Université ou au CNRS. La précarité de leur statut leur permet alors certaines libertés qu'ils mettent en œuvre pour inventer *Fabula*. Ils se sont pour l'essentiel rencontrés par voie électronique et ne sont donc pas classiquement issus d'une camaraderie universitaire. Ce fait est lourd de conséquence. Tout d'abord, ce groupe est géographiquement éclaté (plusieurs villes de France, Canada). Pour continuer à exister, il doit transcender les distances et s'appuyer sur les outils les moins coûteux et les plus efficaces de communication. Or, la plupart des moyens de débat et de coopération disponibles ne permettent pas de dépasser l'échange entre deux personnes (téléphone, courrier), sont extrêmement coûteux (train et avion) ou généralement inaccessibles aux membres du groupe en raison de leur faible implantation institutionnelle (organisation de colloques, organisation de séminaires, participation active à une revue scientifique). Deuxième conséquence : informel, le groupe n'a pas à rendre de comptes aux organismes de tutelle qui habilite et financent habituellement les laboratoires de recherche. Il peut donc prendre le risque de se tromper et prendre le temps dont il dispose pour multiplier les essais, les égarements, les fausses routes. L'organisation du groupe, largement informelle, repose dès lors essentiellement sur l'investissement et les compétences spécifiques de chacun, même si dès l'origine le fondateur initial assure la cohérence de l'ensemble et assume une forme de responsabilité éditoriale, afin d'éviter que le projet ne prenne des directions contradictoires. Enfin, dernière conséquence, le groupe est né à un moment où le nombre d'initiatives visant à favoriser le développement de l'édition électronique scientifique était proche de zéro. A cette époque en effet, l'Internet francophone littéraire comprenait des sites personnels de professeurs ou d'érudits passionnés, des sites portails orientés vers le grand public ou encore des sites institutionnels. Les pages personnelles réactives et parfois riches étaient le plus souvent très spécialisées et ne pouvaient pas avoir une audience large. Les sites portails ne pouvaient prétendre à un statut scientifique et les sites institutionnels n'avaient qu'une très faible réactivité et des cycles de mise à jour et d'évolution très lents. A l'inverse, Fabula se présentait dès lors comme un site techniquement novateur (pour la gestion du travail collaboratif et sa très rapide mise jour) et fondé par un noyau dur suffisamment large pour le faire évoluer rapidement et le diversifier.

Un centre de recherche en ligne

Cela signifie que les outils informatiques n'existaient pas encore pour de tels projets ou relevaient d'une logique commerciale impliquant des coûts totalement hors de proportions avec les moyens

minuscules d'un projet financé par les seuls dons de ses membres... L'ensemble des logiciels utilisés aujourd'hui par Fabula aurait coûté des centaines de milliers de Francs si ceux-ci n'avaient pas été élaborés *ad hoc*. Par ailleurs l'absence quasiment totale de demande de la part de la communauté scientifique a limité dans les faits la taille de l'équipe initiale. Cette faiblesse originelle a en réalité favorisé le développement du site car sa petite taille et les compétences techniques élevées des membres permirent une très grande souplesse, favorable à l'invention et à l'adaptation permanente. Ce qui explique sans doute le succès de Fabula et la croissance forte qui lui a permis de prendre une place importante dans le champ disciplinaire qui est le sien n'est donc paradoxalement non pas seulement la valeur scientifique de l'équipe éditoriale, mais surtout sa complémentarité. Erudits, mais seuls ou démunis techniquement, les membres auraient été à l'origine d'un site spécialisé connu de leurs seuls pairs. En revanche, la capacité technique de l'équipe fondatrice lui a permis de compenser son déficit institutionnel en proposant des outils et des fonctionnalités qui ont fédérés à leur tour de nouvelles compétences. En d'autres termes, c'est la nature évolutive et collaborative du site qui lui a assuré une croissance naturelle : les visiteurs n'étant plus de simples usagers ont pu alors à leur tour mettre au service du groupe leurs compétences.

L'objectif consistait à créer un centre de recherche en ligne, en misant sur une exploitation de toutes les technologies généralement réservées aux sites à but lucratif, permettant d'imaginer un site favorisant l'échange scientifique, l'actualisation fréquente de l'information et surtout la complète maîtrise des contenus par des rédacteurs non informaticiens. Afin d'organiser le travail collectif en l'absence de toute rencontre physique et dans le but de pouvoir coopérer avec des chercheurs quel que soit leur niveau en informatique, l'équipe a élaboré ses propres outils, profitant des compétences informatiques poussées de certains de ses membres (en développement informatique et en administration de systèmes), mais également en puisant dans la vaste librairie des logiciels libres et en unissant ses forces à des projets issus d'autres disciplines mais ayant des besoins très similaires (Revue.org en particulier ³). Le résultat est un site en perpétuelle évolution, tant du point de vue technologique (les versions de logiciels se succédant à grande vitesse) que du point de vue éditorial (les divers collaborateurs n'ayant pas à suivre de formation aux langages informatiques utilisés sur le Web, même pas le plus simple d'entre eux : le HTML).

³ <http://www.revues.org>

La multiplication des services

Emancipé des lourdeurs du papier (poids, coût), le groupe se dote très vite d'une lettre d'information à périodicité régulière, qui informe gratuitement ses abonnés de toutes les nouveautés publiées sur le site depuis la lettre précédente. Cette activité initiale permet de diffuser des annonces de colloques ou des appels à contribution, offrant à Fabula une position fédératrice dans son domaine de spécialité tout en procédant à son élargissement, puisqu'il ne s'agissait plus du seul domaine de la théorie de la fiction. L'arrivée de nouveaux collaborateurs et le succès de la lettre ont amené la création d'une série d'outils de travail en ligne : annuaires, forums, bibliographies, etc. L'intérêt de l'ensemble de ces outils est leur fonctionnement contributif et leur constante évolutivité, à la différence de services équivalents publiés sur papier, qui sont coûteux et dont la mise à jour est un perpétuel casse-tête. La curiosité à l'égard de l'instrument aidant, Fabula en est venu à héberger plusieurs revues et projets littéraires indépendants, à diffuser chaque semaine plusieurs dizaines d'annonces de parution, d'offres de postes ou d'appels à contribution. Il a également organisé et publié des colloques (dont certains ont été publiés en version papier) : un colloque ayant pour thème les « frontières de la fiction », un autre sur « L'effet de fiction » et un débat consacré à l'actualité de Roland Barthes.

La mise en ligne des documents a deux conséquences fondamentales pour le chercheur en sciences humaines. D'une part les circuits de publications sont renouvelés et multipliés de sorte que virtuellement tous les acteurs de la recherche peuvent publier l'intégralité de leurs travaux, d'autre part, la nature numérique des textes produits entraîne leur dispersion dans une topographie que personne ne pourrait tracer. On se trouve donc là devant un cas étrange où l'information et sa circulation sont facilitées à un degré tel que ce mouvement entraîne aussi la disparition de toute hiérarchie et l'affaïssement de la densité de l'information. Plus encore, cette liberté entraîne aussi un élargissement des lieux de publication et l'affaiblissement du rôle des éditeurs classiques. Cela entraîne une relative confusion, car certains travaux deviennent difficiles à trouver tandis que se côtoient de plus en plus de travaux très inégaux. Il faut donc offrir des lieux où les informations et les documents peuvent être non seulement catalogués, mais aussi clairement identifiés et éventuellement scientifiquement évalués. Sans cela, l'internaute en est réduit à interroger des sites de recherche généraux qui lui donneront des résultats peu probants, lui indiquant des textes peu fiables scientifiquement ou encore des sites aux homonymes inattendus, une entreprise « Balzac ⁴ » ou un village « Stendhal ». Le réseau est généraliste et

⁴ Voir par exemple le site internet [<http://www.balzac.com/>] : « BALZAC is a marketing communications agency with a no-nonsense approach to combining the strategies and tactics of marketing, public relations, advertising and graphic design to achieve the marketing goals of our client ».

sans frontière, il faut donc pouvoir le circonscrire. Cette tâche de recension et de validation des sites a été très tôt un des objectifs majeurs du groupe Fabula, mais puisqu'il existait déjà de nombreux autres sites qui proposaient des catalogues de liens ayant trait à la littérature, son but a été de ne prendre en compte que des sites d'intérêt majeurs, universitaires ou non, proposant de véritables contenus ⁵. Pour compléter cet effort de recension de l'Internet littéraire, Fabula a aussi développé un moteur de recherche comparable aux grands outils disponibles sur le web (*Google*, *Altavista*, *Voila*, etc.) mais spécialisé dans l'indexation de sites littéraires : Aleph ⁶. La recherche sur des moteurs généralistes n'est pas toujours efficace et le bruit généré par une requête est parfois considérable. Comme Aleph n'indexe que des sites spécialisés, il permet d'obtenir des résultats beaucoup plus ciblés et utilisables.

Le panorama des ressources proposées au menu du site au moment de sa croissance est donc finalement assez diversifié. D'un côté des ressources scientifiquement très solides : bibliographies, textes, articles et colloques. Mais aussi des volets plus larges qui identifient le site à un portail littéraire et universitaire : moteur de recherche, répertoire de liens, nouvelles publiées quotidiennement. Se pose alors la question du rapport entre ces publications hétérogènes. Le groupe de recherche a semble-t-il été amené naturellement à mettre à la disposition des visiteurs un éventail de ressources qui couvre des domaines généraux et des domaines spécialisés. Ce qui fonde alors sa valeur est une fertilisation croisée de ces domaines que l'on ne peut percevoir qu'à travers l'étude des usages de ces ressources par la communauté scientifique.

ELEMENTS DE METHODE

Dans le monde du livre, la possession est totalement déconnectée de l'utilisation : la détention d'un ouvrage ou d'une revue scientifique ne signalent rien sur l'usage qui en est fait, surtout lorsqu'il s'agit de documents comportant plusieurs articles dont nul ne sait lequel intéresse l'acheteur ; même les bibliothèques sont relativement démunies pour connaître les usages de leur fond, surtout quand elles le proposent en accès libre. En revanche, les publications diffusées sur Internet sont toujours demandées au point initial de diffusion, et cela avec beaucoup de précision : l'internaute ne demande pas un ouvrage ou une revue, il consulte un article ou une information. De plus, nous disposons d'informations précieuses sur ce lecteur, ce qui permet d'ébaucher quelques profils de lecture. La matière première d'une telle information est constituée par un fichier

⁵ <http://www.fabula.org/carnet>

⁶ <http://www.aleph.ens.fr>

très lourd et complexe, le fichier Log ⁷ (ou fichier de Log). Exactement comme un bibliothécaire note qui emprunte quoi au moment d'un prêt, les ordinateurs qui « servent » des documents sur Internet collationnent scrupuleusement de précieuses informations lorsqu'ils envoient un document à un autre ordinateur. Cela donne le fichier de Log, qui contient des centaines de milliers de lignes, chacune correspondant à une demande formulée par un internaute.

Les fichiers de Log et leur pertinence

Que se passe-t-il lorsqu'un internaute consulte un site web ? Il tape en général une adresse ou clique sur un lien hypertexte. L'une ou l'autre de ces opérations revient à donner au navigateur l'ordre d'afficher le contenu d'une URL ⁸ de ce type :

http://www.fabula.org/cgaccueil.php

Le navigateur va alors décoder cette URL pour être en mesure d'afficher le document demandé. Décomposons son analyse :

http://

Avant la suite de signes «:// », figure le protocole à utiliser pour réaliser l'opération ⁹. Dans le cas présent, il s'agit d'utiliser le protocole *Hyper Text Transfert Protocol*, donc à dialoguer avec une machine distante (un serveur) pour afficher des pages web. Après avoir identifié le protocole, le navigateur poursuit sa lecture de l'URL de gauche à droite :

//www.fabula.org/

Entre « // » et « / » figure le nom d'un serveur ¹⁰ : un ordinateur constamment allumé et connecté à Internet dont la fonction est d'envoyer des pages web lorsqu'on le lui demande. Ce nom est unique de façon à éviter toute ambiguïté sur le réseau mondial. Grâce à cette information, le navigateur va pouvoir appeler cet ordinateur, exactement comme le fait un individu qui décrocherait son téléphone et composerait un numéro précis. D'ailleurs, toute machine connectée à Internet doit disposer, au moins pendant le temps de sa connexion, d'un nom unique pour que les échanges

⁷ Le terme est issu de l'expression anglaise *to log in* qui désigne l'action de s'identifier sur un ordinateur en donnant son nom et son mot de passe. L'habitude avait été prise de consigner les informations relatives à ces *login* dans des fichiers indiquant qui s'était connecté et quand. Depuis, le terme de Log est appliqué à tous les fichiers d'informations générés par les logiciels avec une visée de surveillance.

⁸ URL (Anglais : Uniform Resource Locator). Adresse Internet exploitée par les navigateurs. Structure de base d'une URL : *protocol://server/directory/document*.

⁹ Sans entrer dans les détails, *ftp://* correspond à une opération de téléchargement de fichier et *http://* correspond à la consultation d'un site web. Bien d'autres protocoles sont possibles.

¹⁰ En réalité, ce nom se décompose lui-même en *.org*, puis *fabula*, puis *www*, à l'image d'un numéro de téléphone : 01 (Paris), puis 44 (tel quartier), puis le reste du numéro.

soient possibles. Une fois l'ordinateur-serveur localisé – et il peut être localisé à n'importe quel endroit de la planète – il ne reste plus au navigateur qu'à formuler une requête précise. Celle-ci est en réalité contenue dans le reste de l'URL :

/cgaccueil.php

En langage courant cette requête est la suivante : « *Bonjour, www.fabula.org, je souhaiterais obtenir par le biais du protocole http, le document qui correspond à l'adresse : /cgaccueil.php* »¹¹. La chaîne de caractère « */cgaccueil.php* » correspond donc au chemin du fichier sur l'ordinateur serveur. L'URL <http://www.fabula.org/actualites/article1231.php> demanderait quant à elle un fichier qui est dans le répertoire « *actualites* » et dont le nom est *article1231.php*. Cette demande et la réponse du serveur seront scrupuleusement archivées dans le fichier *Log*, qui sert à consigner l'ensemble des opérations effectuées par le serveur. Voici ce qui s'y inscrit lors de la demande : <http://www.fabula.org/cgaccueil.php> :

D91.noos.fr - - [13/Jun/2001:17:19:52 -0400] "GET /cgaccueil.php HTTP/1.1" 200 12565 "-" "Mozilla/4.0 (compatible; MSIE 4.01; Windows 98)"

Ces lignes sont exprimées selon un code très simple que le tableau ci-dessous synthétise.

Code	Signification
<i>D91.noos.fr</i>	Correspond au nom de la machine qui a demandé les pages, le plus souvent ce nom dépend du fournisseur d'accès internet (ici Noos, par le câble à Paris).
<i>[13/Jun/2001:17:19:52 -0400]</i>	La date et l'heure de la demande, ici le 13 juin 2001, à 17h19, avec un décalage horaire de moins 4 heures, soit 23h19 pour la France.
<i>"GET /cgaccueil.php HTTP/1.1"</i>	La demande précise du navigateur, qui demandait le fichier <i>/cgaccueil.php</i> selon le protocole HTTP.
<i>200</i>	Code ¹² qui indique que l'opération s'est terminée avec succès.
<i>12565</i>	Taille du transfert de fichier exprimée en octets.
<i>"Mozilla/4.0 (compatible; MSIE 4.01; Windows 98)"</i>	Chaîne de caractère indiquant au serveur quel logiciel de navigation est utilisé par l'internaute, dans le cas présent, Internet Explorer, version 4.01, utilisé sous Windows 98.

Dans le cas de l'analyse qui suit nous sommes appuyés sur un fichier rassemblant les données échangées entre le site *fabula.org* et les différents utilisateurs pendant un mois (mai 2001). L'exemple précédent nous donne une idée satisfaisante des informations offertes par ce fichier : nous pouvons facilement savoir ce qui a été lu, combien de fois chaque page a été demandée, quand les échanges ont eu lieu et qui les demandait. Il est alors très simple de procéder à des analyses statistiques¹³ qui auront pour charge d'établir par exemple : combien de fois chaque fi-

¹¹ Soit une requête aisément comparable à « Bonjour, librairie de Paris, je souhaiterais recevoir par la poste le livre dont l'ISSN est xxx ».

¹² On peut aussi citer les codes les plus connus : 404 (page non trouvée) ou encore 304 (page déjà contenu dans le cache du navigateur).

¹³ De très nombreux logiciels fournissent des rapports donnant une vision globale de la fréquentation d'un site web.

chier du site a été demandé, combien de machines différentes ont demandé des fichiers, combien de temps reste un internaute lors d'une visite, quel est le nombre moyen de page vue en une visite, quelle est la provenance géographique des visiteurs, etc. Cette analyse est néanmoins difficile pour deux raisons. La première tient au fait que le nom d'une machine ne livre que bien peu d'informations sur l'utilisateur : on a des difficultés à en déduire où il est géographiquement situé. A cela s'ajoute le fait que souvent les utilisateurs demandent leur page à un « proxy », ordinateur-relais destiné à accélérer l'affichage des pages web : dans ce cas, c'est le nom de ce dernier qui apparaît et non celui de la machine de l'utilisateur. Enfin, et cela n'est pas secondaire, bien des machines sont utilisées par plusieurs personnes, tout particulièrement dans le monde universitaire (bibliothèques, bureaux partagés, etc). Ici encore, aucun moyen ne nous permet de reconstituer le nombre exact des différents visiteurs.

Bien des perspectives sont offertes par l'exploitation de ces données, mais le fonctionnement d'Internet, profondément automatisé, génère de très nombreux Logs qui ne sont pas tous pertinents pour l'analyse. Avant toute interprétation, un long travail de tri est nécessaire. Prenons le cas de la page d'accueil du site fabula : elle comporte, outre le texte, des graphiques (logos, icônes utilisées pour la mise en page, image à l'arrière-plan). Le dialogue instauré entre le navigateur et l'ordinateur-serveur va donc contenir un nombre très important de lignes, alors qu'une seule page a été vue. Le mécanisme qui donne lieu à cette sur-représentation n'est pas foncièrement différent de celui que nous venons de décrire. Dans un premier temps, lorsque l'utilisateur va saisir l'adresse de départ : <http://www.fabula.org/>, le serveur va lui envoyer le fichier correspondant, soit, par convention, `index.php`. Une fois qu'il a reçu le fichier, `index.php`, le navigateur va le parcourir pour l'analyser. C'est à ce stade qu'il relève les fichiers complémentaires qu'il doit aussi télécharger pour pouvoir afficher convenablement la page. Il demande ensuite l'ensemble de ces données. Le fichier de Log va donc contenir plusieurs entrées, alors qu'une seule page a été effectivement lue.

Le problème des images et des moteurs : double correction

Le premier travail de l'analyse revient donc à faire le tri entre ces différentes entrées pour ne retenir que celles qui correspondent à des textes réellement lus. Pour cela les statistiques à établir doivent isoler dans un premier temps tous les fichiers de type « image » puis tous les fichiers dont l'utilité n'est que technique et qui ne se traduisent pas par l'affichage d'un texte dans un navigateur (feuille de style, scripts divers). Il est d'usage de distinguer alors le nombre de « hits » (soit le nombre de demandes enregistrées, quelle que soit leur nature) et le nombre de « pages vues », qui lui correspond logiquement à de vraies pages. La proportion des entrées non pertinentes dans un fichier de Log est très importante : sur 600 000 hits enregistrés sur fabula en mai 2001, près de

300 000 ne sont que des images et près de 90 000 sont des demandes de fichiers techniques. Il ne reste donc plus que 210 000 pages réellement lues, soit un tiers seulement des *hits* enregistrés au fil du mois.

Ces pages ont-elles été lues par de vrais internautes ? En observant avec attention l'ensemble des statistiques, on peut s'apercevoir que très peu de pages ont été demandées moins de 40 fois dans le mois. Cette limite inférieure en deçà de laquelle très peu de pages descendent révèle une réalité que connaissent tous les sites Internet : la visite régulière des moteurs de recherche. En effet, les moteurs que nous pouvons interroger par l'intermédiaire de mots-clés sont obligés de mimer le comportement de réels internautes pour enregistrer puis indexer les pages de tous les sites qu'ils visitent. Leur fonctionnement est relativement simple : à partir d'une URL de départ, ils suivent mécaniquement tous les liens qu'ils trouvent sur les pages, à l'image d'un internaute qui déciderait de cliquer systématiquement pour lire l'intégralité d'un site. Cette méthode génère donc à son tour une batterie de *hits* qui ne sont liés à aucune visite réelle, même s'ils concernent des pages et non des images. Il est assez difficile de filtrer simplement les fausses demandes des moteurs de recherche. Lorsqu'ils indexent un site, ils ne le signalent pas autrement qu'en donnant leur nom à la place du nom du navigateur. Les lignes qu'il faut éliminer ressemblent donc à cela (exemple de Google) :

```
216.239.46.75 - - [05/May/2001:01:08:14 -0400] "GET /actualites/article1859.php HTTP/1.0" 200 13363 "-" "Googlebot/2.1 (+http://www.googlebot.com/bot.html)"
```

Pour éliminer ces fausses entrées, nous pouvons procéder à l'analyse exhaustive d'une page qui a été très peu lue et estimer alors le nombre de passages de robots d'indexation pendant le mois ¹⁴. Par sondage successifs sur différentes pages, on retrouve régulièrement le chiffre de 38 passages de robots ¹⁵. Lorsque l'on sait que certaines pages sont demandées plusieurs centaines voire plusieurs milliers de fois cette correction de 38 peut sembler négligeable, elle l'est moins lorsqu'il s'agit de mesurer le poids relatif des sections les unes par rapport aux autres. Pour faire ce type de comparaisons, on ajoute le plus souvent le nombre de pages vues en les regroupant par type. Pour connaître le nombre de pages lues concernant la revue des parutions littéraires, *acta fabula*, on ajoute les pages vues des différentes URLs que comprend cette section. Sur le serveur du site, tous les fichiers concernant la revue sont classés dans un dossier appelé « re-

¹⁴ Il est impossible de faire un filtrage général du fichier, car les moteurs ont des noms qui ne sont pas toujours repérables.

¹⁵ Ce chiffre assez élevé est la conséquence d'une indexation régulière de *fabula* par le moteur de recherche Aleph (<http://aleph.ens.fr/>). Ce moteur, dirigé par l'équipe de *fabula*, met à jour presque quotidiennement ses données et parcourt ainsi le site systématiquement.

vue » de sorte que l'on peut facilement regrouper et ajouter les nombres de pages vues pour chaque url en utilisant ce type de classement :

1020 pages vues pour /revue/
660 pages vues pour /revue/cr/91.php
556 pages vues pour /revue/cr/85.php
547 pages vues pour /revue/cr/94.php
538 pages vues pour /revue/cr/88.php
524 pages vues pour /revue/cr/73.php
etc.

Total : 23 904 pages vues pour la revue *acta fabula* au mois de mai 2001.¹⁶

La même opération pour la section *Actualité* de site ¹⁷, nous donne un total de pages vues de 101 905, soit près de 5 fois plus que la revue et près d'un tiers du total des pages vues. En réalité, pour comparer ces deux chiffres de manière fiable, nous devons alors introduire la correction induite par les 38 visites de moteur recensées. Car, comme chaque URL a été indexée 38 fois de manière mécanique, plus le nombre de pages d'une section est élevé et plus la correction sera importante. Dans le cas de la section « *Actualité* » qui comporte 1958 pages, il faut retrancher à ce total 74 404 (1958x38) : seuls 27 501 demandes réelles ont donc été faites pour la section actualité. De même, la correction à faire pour la section de la revue *Acta fabula*, contenant 59 URL, fera passer le chiffre de 23 904 à 21 662. Dès lors, la fréquentation réelle de ces deux section est beaucoup plus proche que l'on ne pouvait l'imaginer au premier regard. En répétant cette opération, on peut dresser le tableau suivant qui nous donne le poids respectif de chaque section.

¹⁶ On trouvera en annexe le récapitulatif des données en fonction des différentes section du site, mais notre but n'est pas ici de le commenter, nous cherchons simplement à en expliquer la constitution statistique.

¹⁷ Soit les URL du type <http://www.fabula.org/actualités/articlexxx.php>, où xxx est un chiffre quelconque.

Les corrections successives aboutissant au nombre de pages vues

Sections du site	Pages vues par section (sans les images)	Nbre d'URL dans la section	Nbre de page corrigé	Pourcentage
Nouvelles publiées sur fabula	101905	1958	27501	24,11%
Acta	23904	59	21662	18,99%
Mini-colloque barthes	6600	91	3142	2,75%
Colloque99	6594	116	2186	1,92%
Annuaire	5608	6	5380	4,72%
Colloque 2001	4444	66	1936	1,70%
VLIB (carnet de liens)	4220	47	2434	2,13%
Ressources (sections spécialisées)	3700	14	3168	2,78%
Mobiles	3349	10	2969	2,60%
Forums	1412	27	386	0,34%
Conférences Roland Barthes	1094	4	942	0,83%

Le problème des *proxys* et de l'évaluation du nombre de visiteurs

Grâce aux deux corrections précédentes on peut évaluer convenablement le nombre de demandes relatif à un fichier particulier ou à un groupe de fichiers. Cependant, il est aussi intéressant de pouvoir connaître le nombre de visiteurs qui se cache derrière le nombre de demandes. L'analyse se heurte alors à un nouveau problème technique, celui des serveurs cache (ou *proxy*). Ces serveurs jouent le rôle d'intermédiaire entre un internaute et les sites qu'il veut visiter. Leur rôle est d'accélérer la visualisation des pages. Lorsqu'un navigateur désire afficher une URL, il peut faire appel à un *proxy* plutôt que d'interroger directement le serveur. A l'insu de l'internaute, ce serveur joue alors le rôle d'intermédiaire : il vérifie d'abord s'il a stocké la page demandée dans son disque dur ; si c'est le cas, alors il évite de la demander au serveur d'origine et transmet au navigateur la copie locale du document dont il dispose ; si ce n'est pas le cas, il la demandera au serveur d'origine avant de la transmettre au navigateur. Les *proxy* permettent ainsi d'accélérer grandement l'affichage de pages très demandées par les internautes et limitent les transferts sur le réseau mondial, puisque les échanges se font alors entre des machines qui sont géographiquement proches. Ce procédé est mis en œuvre dans les options du navigateur. Or, le plus souvent les fournisseurs d'accès Internet donnent aux internautes des navigateurs qui sont paramétrés ainsi par défaut, par souci d'économie ¹⁸. De ce fait, on peut considérer qu'une majorité des internautes navigue sur internet par le biais de serveur de cache. Cette réalité est lourde de conséquences pour évaluer le nombre de visites. En effet, dans le fichier log, c'est le nom du serveur proxy qui va apparaître et non le nom de la machine de l'internaute. Imaginons que trois clients d'AOL se connec-

tent dans la même journée sur le site Fabula et qu'ils utilisent le même serveur cache : le fichier Log ne comptera pour ces trois visiteurs qu'un seul nom de machine et les logiciels d'analyse des fichiers Log en déduiront qu'il n'y a eu qu'une seule visite. Les universités utilisent elles aussi très souvent un serveur de cache de sorte que tous les utilisateurs peuvent n'apparaître dans le fichier Log que sous un seul nom. Face à cette difficulté, nous en sommes donc réduits à des interprétations modélisées. Les solutions alternatives existantes, qui passent notamment par le biais de *cookies* permettant le traçage des visites, ont toujours semblé trop intrusives aux membres de Fabula.

Malgré les limites que nous venons d'énoncer, les informations révélées par les fichiers de Log sont extrêmement précises. Pour s'en convaincre, il suffit de s'interroger sur ce que nous connaissons des pratiques de lecture du livre (en dehors des sondages dont le coût de réalisation est élevé et la fréquence structurellement faible). Qui imaginerait pouvoir connaître, le nombre d'exemplaires diffusés, mais encore le nombre de pages lues, le nombre de pages lues plusieurs fois, le nombre de personnes qui ont eu l'occasion de lire un exemplaire ? Potentiellement, toutes ces données nous sont accessibles au sujet des sites Web et offrent aux chercheurs une occasion nouvelle d'avoir une information sur leurs pratiques de lecture et de publication.

LES ROLES DE L'INTERNAUTE

Le modèle centralisé du livre, duquel s'inspire fortement le vertical Minitel, est bousculé par l'introduction d'Internet, dont le fonctionnement est potentiellement beaucoup plus circulaire que centralisé, horizontal que vertical, participatif que passif. Comme nous allons le voir, l'internaute endosse plusieurs costumes au fil de sa navigation : tour à tour lecteur, prescripteur, critique, abonné, source d'information, rédacteur, il participe à la vie d'une communauté scientifique.

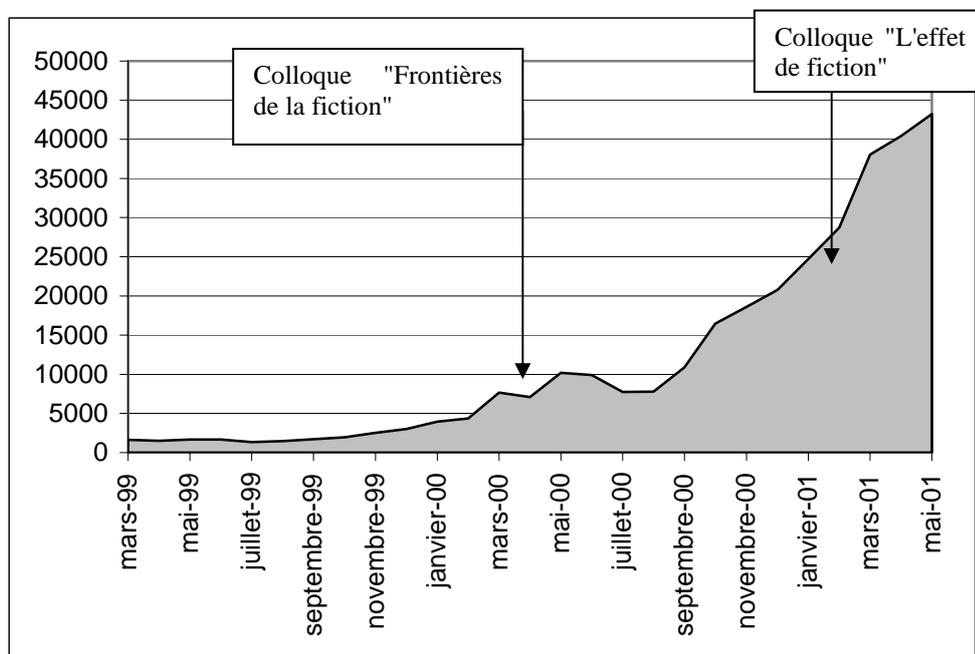
Mesure des usages et de la fréquentation

L'analyse de l'évolution chronologique montre le succès croissant du site. En mars 1999, le nombre de visites s'élevait à 1611. Il dépasse les 2 500 visites en novembre, les 5 000 à l'occasion de son premier anniversaire, les 10 000 en septembre 2000, les 25 000 en février 2001. Il faut donc au site environ six mois pour chaque doublement du nombre de visites. Cela est lié à l'augmentation de l'offre de contenu (publication de colloques, amélioration du carnet de lien,

¹⁸ Les fournisseurs d'accès payent eux-mêmes leur accès à d'autres fournisseurs. La facture est fonction du volume des transferts en octets. Si leurs clients utilisent des *proxy* pour naviguer sur internet, leur trafic sera moindre.

augmentation rapide du nombre et de la diversité des nouvelles) autant qu'à un phénomène de résonance sur la Toile et ailleurs. Il est difficile de distinguer les deux phénomènes, mais l'ensemble participe bien à une dynamique générale réelle.

Chronologie du nombre de visites (sessions)



Une comparaison avec les rares statistiques précises publiées au sujet de sites d'information plus généralistes donne quelques ordres de grandeur. *Le Monde*, quotidien de référence français, recevait en juillet 2000 1,8 millions de visites tandis que Fabula en recevait 7 768 ; il en recevait 3,2 millions en juillet 2001 tandis que Fabula en recevait près de 50 000. La disproportion évidente entre les résultats s'explique par le public et le rythme de publication : quotidien vénérable vendu à plusieurs centaines de milliers d'exemplaires dans son format papier, *Le Monde* a un public beaucoup plus large que Fabula, qui touche une communauté scientifique très spécialisée, dont les publications sont moins fréquentes et dont les ventes ne dépassent que très rarement le demi-millier d'exemplaires. Ce qui est en revanche significatif, c'est que la croissance du *Monde* en douze mois correspond à une multiplication par 1,79 alors que la croissance de Fabula correspond à une multiplication par 6,5. Il ne faudrait cependant tirer aucune conclusion mécanique de ce constat car les deux publications ne sont pas au même moment de leur histoire électronique. Notons cependant, en comparant avec le jeune hebdomadaire *Marianne*, que la croissance d'un site –même issu d'une publication papier– n'est pas une constante, puisque *Marianne en ligne* stagne à un niveau inférieur à celui de Fabula. La qualité médiocre du site concerné explique ce résultat.

Nombre de visites mensuelles de quelques grands journaux nationaux

Source : Mediametrie ¹⁹	Juillet 2000	Juillet 2001
<i>Le Monde</i>	1 826 770 visites	3 279 890 visites
<i>Libération</i>	1 446 220 visites	1 920 070 visites
<i>Marianne en ligne</i>	25 440 visites	24 590 visites
<i>L'Express</i>	---	295 860 visites

Ce dynamisme est dû à l'attrait de diverses sections (voir tableau ci-dessous), au premier rang desquelles on compte *Acta fabula*. Le taux de consultation par page de cette jeune publication est en effet largement supérieur à celui de toutes les autres section (367 consultations mensuelles par page). Le système de nouvelles affiche un score très honorable, avec 137 consultations ²⁰. Le score d'*Acta fabula* démontre que les textes inédits spécifiquement universitaires, relatifs à des publications ayant pour support principal le papier, suscitent l'intérêt. On constate ainsi que c'est la tradition universitaire la plus ancrée (le support papier mais également l'exercice du compte-rendu de lecture) qui fonctionne le mieux. La nouveauté du support n'exclut pas le classicisme de l'exercice, bien au contraire, puisque le premier conquiert rapidement la légitimité de l'autre, et le second subit un vigoureux rajeunissement grâce à la souplesse du précédent. Il faut noter que les succès d'*Acta* et du système de nouvelles s'expliquent en grande partie par une adéquation parfaite au calibre d'Internet : par leur petite taille, les articles se prêtent aisément à une lecture à l'écran ou à une impression rapide. L'édition électronique est aussi une affaire de rythme, de vitesse de modem, de domestication de la lecture à l'écran. La force de ces deux publications s'explique enfin par les choix de mise en page, qui les valorisent en les plaçant au sommet des deux colonnes principales. De ce point de vue, *Acta* jouit d'un avantage supplémentaire : le rythme de publication de nouveaux comptes-rendus étant relativement lent, chaque article reste présent sur la page d'accueil pendant plusieurs semaines, à la différence de l'actualité scientifique, dont le rythme est quotidien et où une nouvelle chasse l'autre. Enfin, le public auquel *Acta* et les nouvelles se destinent, est par nature plus large. Leur visée est certes, d'alimenter le débat interne à la recherche sur la fiction, mais des non spécialistes peuvent facilement utiliser ces outils pour avoir un panorama rapide des travaux en cours dans un domaine qui est proche du leur.

Émerge un deuxième groupe, qui correspond à des productions intellectuelles de l'équipe de *Fabula* : *Virtual Library* et les colloques. Si parmi eux l'annuaire de liens a plus de succès, c'est logiquement en raison de son degré de spécialisation moindre : son public est structurellement plus

¹⁹ <http://www.mediametrie.fr/web/resultats/cybermetrie/print.php?id=348>

²⁰ Le taux de consultation obtenu par les nouvelles n'est pas calculé sur l'ensemble des 2000 nouvelles publiées. Un grand nombre d'entre elles sont sorties de l'actualité et sont devenues caduques, aussi faut-il ne pas en tenir compte pour le calcul du taux. On admet que seules les 200 dernières nouvelles (ce nombre est arbitraire) sont valables. Il s'agit de l'équivalent de trois mois de publications de nouvelles.

vaste. On ne peut pas s'étonner qu'un colloque portant sur la théorie littéraire draine moins de public qu'un compte-rendu de lecture ; ce phénomène n'a rien de spécifiquement électronique.

Enfin, un troisième groupe émerge avec des taux relativement faibles : l'annuaire de chercheurs et les *forum*. Ces résultats sont le fruit d'une pratique particulière de l'annuaire, vers lequel on ne se dirige qu'en cherchant une information très précise et qu'on ne feuillette pas à la différence d'un colloque. Le cas des *forum* est plus délicat : il s'agit d'une pratique nouvelle, qui n'est pas encore totalement entrée dans les mœurs universitaires. Même si elle a donné lieu à une participation honorable ainsi qu'à une fréquentation honnête, il semble que le statut de cette formule (pourtant très prisée dans d'autres milieux) ne soit pas clair pour une grande partie des utilisateurs. Sorte de publication écrite de ce qui relevait auparavant des débats oraux lors des colloques, la formule du *forum* démarre et ne semble pas avoir encore trouvé sa forme et son public définitifs. Il en est de même pour les listes de discussions. Là aussi, le statut du débat écrit pose question, empêchant le développement généralisé de cet outil. Il est notable qu'un site qui génère une fréquentation très importante suscite proportionnellement peu de réactions en ligne, malgré des efforts réitérés pour les encourager. Tout se passe comme si le nombre important de pages et la qualité des publications en ligne assurées par le site induisaient un comportement de l'internaute similaire à celui qu'il a devant des publications papier.

Taux de consultation par page (mai 2001)

Section		Nombre de pages vues par page disponible
<i>Acta</i>	367	
Nouvelles	137	
Vlib	52	
Colloque Barthes	35	
Colloque 2001	29	
Annuaire	16	
<i>Forum</i>	14	

Cet intérêt quantitatif pour le lectorat de Fabula ne doit pas devenir un intérêt pour son audience. Le nombre importe en effet moins en ce domaine que la qualité de la lecture, difficilement mesurable. On peut cependant l'approcher par quelques indices, glanés dans les fichiers de Log. Premier enseignement : il s'agit d'un public qui souhaite revenir. Au cours du mois de mai, 614 internautes ont ajouté le site dans leurs favoris ²¹, ce qui donne un potentiel de fidélisation de 7000 par an. Comparé à la taille de la communauté de recherche littérature, ce chiffre est élevé. De plus,

plus de 800 personnes consultent une page d'information sur le site Fabula (601 accès à la page « à propos », 229 accès à la page équivalente en anglais, 423 accès à la FAQ). Il ne s'agit donc pas d'un lectorat qui consomme Internet, mais qui s'intéresse à la source de l'information qu'il collecte, afin de pouvoir la mettre en contexte et de savoir quel est l'esprit qui anime ce projet. Deuxième enseignement : la durée moyenne des visites est de 14 minutes, ce qui est relativement élevé puisqu'il ne s'agit que d'une moyenne et que les habitués effectuent des micro-visites de moins d'une minute, afin de vérifier d'un coup d'œil les nouveautés du site. Il faut également noter qu'au moins 36 688 nouvelles scientifiques et 2 892 articles de colloques ont été imprimés en mai 2001 ²², ce qui marque le retour au papier des productions électroniques et signifie la volonté de l'internaute de conserver une trace de sa visite. Pour aller plus loin dans l'analyse du comportement du visiteur, il faut extraire des fichiers Log un troisième enseignement, contenu dans les expressions recherchées dans les moteurs de recherche et débouchant sur Fabula.org (dans l'ordre d'importance : Google.com, Yahoo.com, Altavista.fr, Lycos.fr, Voilà.fr, MSN.fr, Google.de). Le premier constat (voir tableau ci-dessous) qui s'impose est la très grande variété des mots-clefs : près de 18.000 expressions différentes. Parmi celles-ci, les plus importantes peuvent se regrouper en deux catégories. Tout d'abord des recherches très généralistes avec des expressions génériques comme « *littérature* », « *littérature française* », « *french literature* », « *literatura francesca* » ou bien des concepts critiques comme « *autofiction* », « *génétique* », « *naturalisme* », « *auteur* ». Ces internautes recherchent des ressources littéraires sur Internet de façon exploratoire. On trouve également des demandes d'information beaucoup plus précises qui sont la trace d'une recherche active et motivée par un besoin très précis : un auteur (« *Agrippa d'Aubigné* »), un titre d'œuvre (« *Chartreuse de Parme* »), voire un extrait sont aussi l'occasion de découvrir le site Fabula qui pourtant n'est pas un site spécialisé. Si d'aussi nombreuses requêtes aboutissent ainsi à Fabula, c'est que la variété et le nombre des nouvelles lui donnent une forte visibilité dans les différents moteurs de recherche. Dans ce deuxième type de visite, l'internaute qui cherchait des renseignements sur « *Marguerite Yourcenar* » aboutira probablement au texte d'une nouvelle. Cette lecture sera alors pour lui l'occasion de découvrir Fabula en s'écartant de sa recherche initiale, ou bien tout simplement de repartir en arrière pour trouver un site réellement spécialisé. Ainsi, un internaute arrive sur Fabula en provenance de Yahoo, où il a utilisé la fonction de moteur de recherche (fournie par Google) et tapé « *Colloque de littérature française* ». Son arrivée sur Fabula étant motivée par des raisons très proches de l'objectif que s'est donné le site, il lui trouve manifestement un réel intérêt, puisqu'il consulte 67 pages différentes en 1h10 (rappelons que 67 pages sur Internet peuvent correspondre à environ 30 à 300 pages A4). Une telle navigation est exception-

²¹ Cette estimation représente un minimum, car les utilisateurs de Netscape, Opera, Mozilla et Lynx ne sont pas pris en compte. Seuls les utilisateurs d'Internet explorer sont repérés.

²² Les impressions n'utilisant pas la fonction "Imprimer" offerte par Fabula ne sont pas comptabilisées.

nelle, car elle est longue, fouillée, maîtrisée. L'étude ponctuelle de ce type d'information sur les visites permet de dépasser l'information massive qui a été exploitée jusqu'à présent. En effet, les comptages donnent un effet de nombre mais n'approchent pas la démarche de l'internaute de façon suffisamment proche pour qu'on imagine son itinéraire, ses attentes et ses compétences. C'est la raison pour laquelle nous avons cherché à traquer dans l'énorme fichier de Log des navigations imputables à un individu, excluant les moteurs de recherche ainsi que les *proxy*.

Principales expressions cherchées dans un moteur de recherche généraliste débouchant sur une viste de Fabula.org (mai 2001)

	Expressions	Nombre	% du total
1	fabula	229	1.29%
2	romantism	122	0.68%
3	french literature	71	0.4%
4	roland barthes	70	0.39%
5	georges perec	67	0.37%
6	dictionnaire	60	0.33%
7	bonheur d'occasion	53	0.29%
8	perec	48	0.27%
9	don quichotte	47	0.26%
10	la chartreuse de parme	42	0.23%
11	cache:oavrvvf6g5c:www.fabula.org/ fabula.org	41	0.23%
12	génétique	40	0.22%
13	jardin des plantes	39	0.22%
14	littérature	38	0.21%
15	literature	37	0.2%
16	marguerite yourcenar	37	0.2%
17	walter benjamin	36	0.2%
18	moschos m.	34	0.19%
19	sodome	33	0.18%
20	cent mille milliards de poemes	33	0.18%
21	tahar ben jelloun	31	0.17%
22	literature francaise	31	0.17%
23	academie de musique	31	0.17%
24	genetique	27	0.15%
25	auteur	27	0.15%
26	philippe soupault	27	0.15%
27	anamorphoses	26	0.14%
28	jean tardieu	26	0.14%
29	naturalisme	26	0.14%
30	la princesse de clèves	26	0.14%
31	michel foucault	25	0.14%
32	bartolomeo delbene	25	0.14%
33	adramandoni	24	0.13%
34	cache:tpfiydi6r3i:www.aief.eu.org/congres/liio_congres/liio_congres. jean marc loyay	24	0.13%
35	romain gary	24	0.13%
36	lettred' amour	24	0.13%
37	les liens pipe coquette	24	0.13%
38	cache:091qgmxlzqk:www.aief.eu.org/cahiers/tables/b/b. abraham bosse saisons	24	0.13%
39	cache:539abba430ab964d:www.aief.eu.org/cahiers/tables/l/l. charondas le caron	24	0.13%
40	cache:www.aief.eu.org/cahiers/tables/j/j. jeannot et colin	24	0.13%
41	galerie du genie paris	24	0.13%
42	antoine de la roche chandieu	24	0.13%
43	cache:http://www.aief.eu.org/cahiers/tables/f/f. fables phèdre	24	0.13%
44	cache:www.aief.eu.org/cahiers/tables/r/r. soudier	24	0.13%
45	cache:www.aief.eu.org/congres/liio_congres/liio_congres. jean marc loyay	24	0.13%
46	the vagabond, by maupassant	24	0.13%
47	cache:u5q6pdcrlk0:www.aief.eu.org/cahiers/tables/l/l. charondas	24	0.13%
48	cache:http://www.fabula.org1054.	23	0.12%

	Expressions	Nombre	% du total
49	chartreuse de parme	23	0.12%
50	barthes	23	0.12%
Sous-total d'expressions trouvées pour les lignes ci-dessus		2,816	15.88%
Total d'expressions trouvées dans le fichier journal		17,727	100 %

La démarche du visiteur

Nous avons isolé plusieurs visites différentes, dont nous décrivons ici trois exemples ²³. Le 13 mai 2001 à 10h07, un individu se connecte à Fabula à partir de l'Algérie. Il semble avoir tapé l'adresse lui-même sur son navigateur, sans avoir utilisé de moteur de recherche ni activé un *bookmark*. Il consulte le site pendant une minute et dix secondes, le temps de se rendre compte qu'il correspond bien à son attente : il décide alors d'utiliser la fonction « Signaler à un ami », imaginant sans aucun doute que cet ami sera heureux de découvrir un tel site. Pour des raisons déontologiques évidentes, seul l'appel à la fonction est conservé et le serveur ne conserve pas de trace du message ni de l'adresse du destinataire. Nous ne pouvons donc pas aller plus loin dans l'observation de cette attitude, qui montre bien, cependant, la force de ce que l'on appelle communément le « bouche à oreille » : l'effet de réseau. Après avoir à nouveau consulté pendant trois minutes la page d'accueil, il se met à consulter des pages du système de nouvelles : l'article 2080 (un événement scientifique : « *L'espace en toutes lettres* »), sur lequel il ne s'attarde pas (à peine 4 secondes), puis l'article 1424 (un autre événement scientifique : « *Ecriture de soi et lecture de l'autre dans la littérature du XXe siècle* »). Il ne s'attarde que 25 secondes, mais il peut avoir enregistré ou imprimé la page pendant ce temps. Il décide alors de procéder de façon plus méthodique, en consultant la page de calendrier. Là, la richesse de l'information le retient pendant près de cinq minutes. Ayant appréhendé l'offre avec plus de précision, il se dirige sur des appels à contribution, pensant sans doute pouvoir répondre à l'un d'entre eux. Il consulte ainsi l'article 2110 (« *Fictions of dialogue* »), l'article 2170 (« *Le genre des territoires* ») et l'article 1944 (« *Patrimoine culturel et langues au Maghreb* »). Il reste sur ces pages environ 30 secondes, puis 30 secondes à nouveau, enfin plus de quatre minutes. Manifestement, l'appel sur le patrimoine culturel et les langues au Maghreb l'intéresse. Sa visite est de moins en moins précipitée, il semble avoir atteint ce qu'il cherche, prenant désormais le temps de lire en ligne. C'est sans doute la raison pour laquelle il se propose de participer au projet, décidant de consulter pendant plus de sept minutes la page « *Discuter, proposer, contribuer* » de la revue des parutions en théorie littéraire ²⁴. C'est alors que nous

²³ Le tirage au sort des exemples choisis a été effectué dans un échantillon représentant trois journées : le jeudi 3 mai, le dimanche 13 mai et le mercredi 23 mai 2001

²⁴ <http://www.fabula.org/revue/contribuer.php>

perdons sa trace. Il reviendra peut-être plus tard. Sa visite a duré environ 25 minutes, a donné lieu à un signalement à un ami ainsi qu'à la lecture de six articles scientifiques et une notice technique. Les choix techniques de Fabula.org (refus des *cookies*) empêchent de savoir s'il est revenu.

Le 3 mai 2001 à 14h42, un internaute se connectant depuis la Roumanie a trouvé sur le site Auteurs.net un lien vers Fabula sur lequel il clique. Il prend cinq minutes pour consulter la page d'accueil, puis il décide de lancer une requête « *Paul Claudel* » qui le mène simplement sur la nouvelle 420 (« *L'inspiration scripturaire dans le théâtre et la poésie de Paul Claudel* »). Le moteur de recherche interne a été efficace puisqu'il lui permet d'accéder tout de suite à une page correspondant précisément au sujet auquel il s'intéresse. Après une minute de consultation de la page, il décide de l'imprimer en appelant la fonction créée à cet effet dans Fabula. Il revient ensuite à la nouvelle 420 pour la signaler à un ami en utilisant la fonction dédiée à cet effet. Il revient alors à nouveau à cet article pour taper la même requête sur le moteur de recherche : « *Paul Claudel* ». Il accède alors à la nouvelle 627 (« *L'écriture de l'exégèse de Paul Claudel* »). Là encore, satisfait, il signale cette page à un ami. Il lance alors à nouveau la même requête, cette fois pour atteindre l'article 1474 (« *Histoires littéraires n°4* », qui publie notamment des lettres de Paul Claudel). Il consulte la page pendant 13 minutes. Il est possible qu'une telle durée soit due à une prise de notes ou à une autre occupation, l'ordinateur attendant que son propriétaire revienne. Celui-ci décide alors de consulter l'annuaire de chercheurs, consultant d'abord la liste complète de tous les chercheurs enregistrés, puis préférant chercher parmi les chercheurs ceux dont la fiche comporte le terme « *Paul Claudel* ». Six minutes de recherche n'ayant rien donné, il opte le carnet de liens, dans lequel il consulte les catégories 1003 (« *Ecrivains* » : on imagine qu'il cherche des sites au sujet de Paul Claudel, 1001 (« *Centres de recherche* ») et 2002 (« *Index et annuaires de sites* »). Il passe en tout douze minutes sur ces quatre pages. Il y trouve sans doute des sites intéressants, puisqu'il interrompt là sa visite : profitant d'un lien hypertexte fertile, il file sans doute vers d'autres terres claudelliennes. Il est d'ailleurs possible qu'au cours des douze minutes, il ait visité rapidement certains sites signalés par Vlib. Le 23 mai, une nouvelle visite provient du même fournisseur d'accès roumain... qui consulte l'article 2245 (« *La passion de Claudel* ») puis 1436 (« *Catalogue des thèses françaises en ligne* »). Sans avoir les moyens de le démontrer, il est probable que ce soit le même internaute spécialiste de Paul Claudel qui soit revenu. Cette double visite roumaine est plus longue (44 minutes) et surtout, elle manifeste un plus grande maîtrise de la part de l'internaute. Celui-ci utilise en effet les outils dédiés à une recherche précise (moteur de recherche, annuaire de chercheurs, carnet de lien) et exploite les fonctionnalités du site (fonction d'impression, signalement d'une page à un ami à plusieurs reprises). Enfin, sa visite est très précise parce qu'il sait exactement ce qu'il cherche. Son retour probable le 23 mai montre la satisfaction de la visite du 3 et confirme l'efficacité des outils de recherche qui lui permettent d'éviter d'avoir à feuilleter des centaines de pages avant de trouver celles qui l'intéressent. Ainsi, sa visite

du 23 est brève (2 minutes) mais efficace, puisqu'il a trouvé une nouvelle page concernant Claudel (publiée le 21 mai, elle n'était pas présente le 3).

Le 13 mai 2001, un individu se connecte à 10h du matin sur le site de Fabula. Une lecture de deux minutes lui suffit pour penser que ce site l'intéresse et il s'abonne immédiatement à la lettre. Il se rend ensuite sur la page « calendrier » des nouvelles, consultant cinq nouvelles différentes pendant une dizaine de minutes (« *L'espace en toutes lettres* », « *Écriture de soi et lecture de l'autre dans la littérature du XXe siècle* », « *Fictions of dialogue* », « *Le genre des territoires* », « *Patrimoine culturel et langues au Maghreb* », « *Le temps des oeuvres. Mémoire et préfiguration* »), puis quitte le site pendant quelques minutes, pour en revenir en provenance du serveur gérant les emails chez Yahoo.com : de toute évidence, il a suivi le conseil donné par le gestionnaire d'abonnement, qui consiste à aller vérifier qu'il a bien reçu confirmation de son abonnement. Il revient ensuite pour consulter une dernière page : « Contribuer ». Cet itinéraire s'apparente à une approche générale du site, effectuée par quelqu'un qui ne cherche sur le moment rien de très précis, mais qui souhaite mettre en place un système de veille documentaire (en s'abonnant à la lettre de Fabula), qui s'assure que sa démarche est couronnée de succès et enfin qui se demande s'il va contribuer à la revue *Acta fabula*. Ici encore, et malgré les consultations enthousiastes des pages « collaborer », le silence des internautes est éloquent. S'ils souhaitent réagir, c'est en s'intégrant dans l'équipe et non pas en publiant directement une réaction par le biais d'un formulaire. La crédibilité scientifique du site fonctionne comme un frein à une expression spontanée. Exactement comme dans un cadre universitaire, l'internaute semble chercher à conquérir un statut avant d'apporter sa pierre à l'édifice.

La participation des internautes

Choisis au hasard, ces trois itinéraires montrent le caractère international du lectorat de Fabula.org, confirmé par une mesure plus générale qui attribue grossièrement ²⁵ 21% des visites à la France, 16% à une origine inconnue et 63% à une origine internationale. Ils éclairent plus encore l'extrême diversité des comportements d'internautes, qui empruntent toujours de itinéraires différents, utilisent seulement les outils dont ils comprennent l'intérêt et dont ils ont besoin sur le moment, effectuent un feuilletage curieux ou une recherche pointue pour un cours ou une recherche urgente... La conception du site se doit de répondre à ces différentes démarches, qui ont toutes leur légitimité et leur sens. Enfin, ils montrent également le rôle actif que joue l'internaute, décidant de s'abonner (ou de se désabonner), choisissant de conserver l'adresse du

²⁵ On ne peut que souligner la fragilité de telles mesures, qui surestiment le rôle des Etats-Unis, en raison notamment de l'ambiguïté qui règne depuis les origines sur le « .com ».

site dans ses signets, signalant à ses amis l'existence de Fabula (450 pages ont ainsi été signalées en mai 2001), bref, participant à son tour –de près ou de loin– à la vie du site. Fondée à l'origine sur la logique du partage et de la netiquette * (il y a quelques années circulaient encore des recommandations sur la taille des signatures ajoutées en bas des courriers électroniques afin d'économiser de la bande passante), la communauté des internautes a-t-elle conservé l'impulsion initiale ou a-t-elle transformé l'édition électronique en une édition classique, reprenant la logique pyramidale et centralisée de l'économie du livre ? L'étude des pratiques de partage et d'échange gravitant autour de Fabula est d'autant plus utile qu'elle constitue le premier pas vers l'études pratiques de production scientifique proprement dites, que nous ne pourrions pas aborder ici. Cela implique une étude à part entière, s'appuyant sur des sources très différents : entretiens oraux et questionnaires écrits pour l'ensemble des collaborateurs, associées à des statistiques sur les propositions, candidatures, participations aux forums, etc. Dans le cadre de cet article, nous pouvons cependant donner quelques éclairages factuels sur l'interactivité du support électronique, si légendaire qu'on finit par ne plus savoir si elle existe et quelle forme concrète elle prend.

On rêve d'évaluer clairement le rôle exact joué par la communauté dans l'alimentation du site en nouvelles et suggestions de sites pour le carnet de liens, de suivre à la trace la transformation d'un internaute curieux en internaute assidu, et d'un internaute assidu en internaute passionné... et d'un internaute passionné en collaborateur plus ou moins régulier du site. Une telle étude demande le déploiement d'une méthodologie complètement différente. On peut cependant évoquer rapidement les réactions des visiteurs transmises par courrier électronique, qui sont l'expression d'une communauté de lecteurs. Eclairante par sa diversité, le *corpus* des centaines de lettres reçues est surtout édifiant par sa masse : les publications scientifiques classiques ne suscitent pas une telle variété de réaction et un courrier aussi important.

On constate d'abord que les internautes préfèrent l'usage du courrier électronique aux formulaires qui lui sont proposés. Même pour soumettre une nouvelle, la plupart utilisent la formule non formée qu'est le courrier électronique. C'est sans doute le signe d'une incompréhension de la puissance de certains formulaires, qui font gagner du temps à l'équipe. C'est en particulier le cas pour le système d'actualité, qui récupère les données du formulaire directement dans une base de données. Cette relative rigidité de l'internaute se retrouve dans ses formulations, où l'on constate qu'il a parfois du mal à cerner ce nouveau support scientifique. L'hésitation est grande, entre un service public d'informations (alors que Fabula est une association relevant de la loi de 1901), un site se résumant à un personne individuelle, une association caritative destinée à rendre service aux chercheurs éloignés, un service de soutien scolaire, une revue et parfois un groupe de recherche. Cette difficulté n'est pas réellement surprenante car on n'invente pas souvent de nou-

veaux supports scientifiques. Le flou est d'ailleurs accentué par l'évolution constante du projet, qui se modèle peu à peu en fonction des évolutions techniques et surtout humaines.

L'inévitable demande d'aide est bien entendu fréquente, qu'elle soit proche de l'univers scientifique de Fabula (« *J'attends votre aide pour faire connaître mes travaux sur X²⁶ et les présenter aux abonnés de votre liste en document attaché pour envoyer mon texte aux intéressés* ») ou que le lien soit ténu (« *J'ai écrit plus de 400 pages lors d'un voyage d'un an sur les routes de France. Seule avec un chien, j'ai vécu de façon très insolite. Si ce style de roman vous intéresse dites-le moi et je vous le faire parvenir* »). Porte d'entrée réelle à une sphère de l'Internet littéraire, Fabula est alors faussement perçu comme un lieu propre à donner des conseils. D'autres messages n'ont d'autre objectif que de donner un avis sur un sujet qui les intéresse (« *J'avais demandé à mon directeur de thèse la présence [de monsieur X] dans le jury. Celui-ci m'a répondu que monsieur X ne pouvait être membre d'un jury de thèse, ce qui me paraît regrettable pour la recherche* ») : quelle réponse apporter à ce type de message ? S'y ajoutent des sollicitations commerciales diverses, plus ou moins reliées au projet littéraire de Fabula.

Un autre type de message relève de demandes d'explications correspondant à une mauvaise utilisation, les renseignements demandés étant fournis clairement sur le site : « *Madame, Monsieur, Je souhaitais savoir si vous faites des comptes-rendus sur votre site internet ou si vous donner uniquement des informations sur les ouvrages qui vous sont soumis. Faut-il aussi vous fournir les informations que l'on souhaite voir diffuser par e-mail? En vous remerciant d'avance de me répondre.* ».. Cela va souvent encore plus loin, Fabula devant se transformer en service public, répondant comme un office du tourisme à tout type de demande sur son domaine : « *Voudriez-vous avoir la gentillesse de me donner des renseignements sur des sites pour recherches, projets, conférences, présentations, etc. (de préférence aux E.U. ou au Canada).* »

D'autres sollicitations sont plus constructives. De plus en plus d'organiseurs d'événements scientifiques les signalent eux-mêmes à Fabula de façon spontanée : « *Je vous envoie en fichier attaché le préprogramme de Colloque X, que j'organise avec mes collègues X et X aux universités de Leuven et d'Anvers, en X 2002. Vous serait-il possible de les intégrer dans le site fabula ?* ».. Les messages de ce type sont nombreux et participent à la constitution d'une communauté scientifique électronique : « *comme annoncé, de conserve avec X de Paris III, nous préparons l'agrégation 2002 sur le réseau JAL (Journées d'agrégation en ligne). Pouvez-vous, à l'endroit*

²⁶ Afin de respecter le secret de la correspondance, nous avons anonymisé les messages en remplaçant les informations nominatives (ou toute précision permettant de retrouver des informations nominatives) par des X.

adéquat, l'annoncer sur votre fabuleux site et faire un lien avec [notre site] ». Des correspondants étrangers, constatant l'annonce d'un de leurs événements, adressent des remerciements et en profitent pour proposer d'autres informations : « Merci beaucoup pour annoncer mon CFP X. Je n'ai pas connu votre website. C'est très intéressant pour moi et mon collègues de le Medieval Romance Society. Nous avons deux autres annonces sur le sujet de théorie et littérature pour le même congrès médiéval. Pouvez-vous annoncer ces CFP aussi a votre website, svp? Nous vous soyons très reconnaissant. Je vous les enverrai maintenant. ». D'autres participent à la qualité du site en veillant à son exactitude pour les informations qu'ils connaissent : « Ayant détecté une erreur sur la date d'échéance de l'appel à communications pour le Congrès International du X qui aura lieu à X (Espagne), je vous prie de bien vouloir corriger le 1 octobre 2001 en le remplaçant par le X. Merci. » S'y ajoutent des demandes d'autorisation pour des liens, des demandes d'utilisation du moteur Aleph sur d'autres serveurs, des demandes d'informations sur des rédacteurs de Fabula, des signalements de problèmes techniques (réels ou imaginaires). Cette attention grandissante des membres de l'Université à l'égard de Fabula indique clairement que le site est devenu un lieu du groupe et qu'il incorpore au cours de sa croissance les édifices traditionnels de la communauté scientifique. Les propositions de partenariat entre le site Fabula et des laboratoires de recherche ou des écoles doctorales, montrent bien que progressivement la marginalité des publications électroniques se réduit et qu'elles sont en passe de devenir une ressource banalisée de la recherche en sciences humaines.

L'ensemble de cette correspondance, très hétéroclite, montre que la mise en place d'un site web scientifique expose à de nombreuses sollicitations totalement hors-sujet mais que cela permet surtout de donner une caisse de résonance à une activité scientifique qui n'est pas toujours compatible avec le rythme des publications papiers.

CONCLUSION

L'usage des ressources offertes en ligne par Fabula ne peut être sérieusement évalué sans une méthode précise, explicite et simple, qui permette à l'avenir des comparaisons terme à terme, non pour mesurer des taux d'audience, mais pour évaluer et expliquer des différences de comportements sur le même site dans le temps, ou avec d'autres sites similaires. L'enseignement principal de cette étude consiste en la mise en évidence d'une grande liberté de l'internaute, à condition que le site soit construit de façon ouverte et pensé pour permettre cette liberté : entrant dans Fabula par les petites ou la grande porte, *via* des annuaires ou des moteurs, des favoris ou des courriers électroniques de signalement à un ami, l'internaute se déplace ensuite en fonction de sa sensibilité, réinventant sans doute la lecture cursive, maîtrisée et efficace du chercheur lorsqu'il se trouve confronté à la masse documentaire d'une petite bibliothèque dans laquelle il lui faut se frayer un chemin pour trouver l'information qui l'intéresse. Rien de nouveau, en somme, à un détail près : l'accès à l'information se fait à domicile ou depuis le bureau, les recherches ne portent plus seulement sur le titre d'un ouvrage mais sur le texte intégral de la publication, enfin, et surtout, il existe un véritable échange –l'air du temps parle « d'interactivité »– qui renouvelle complètement le modèle vertical de l'édition classique.

FABULA.ORG : USAGES D'UN SITE SCIENTIFIQUE LITTERAIRE (1999-2001) **1**

GENESE D'UN PROJET ELECTRONIQUE ET SCIENTIFIQUE	1
UN GROUPE INFORMEL	2
UN CENTRE DE RECHERCHE EN LIGNE	2
LA MULTIPLICATION DES SERVICES	4
ELEMENTS DE METHODE	5
LES FICHIERS DE LOG ET LEUR PERTINENCE	6
LE PROBLEME DES IMAGES ET DES MOTEURS : DOUBLE CORRECTION	8
LE PROBLEME DES <i>PROXYS</i> ET DE L'EVALUATION DU NOMBRE DE VISITEURS	11
LES ROLES DE L'INTERNAUTE	12
MESURE DES USAGES ET DE LA FREQUENTATION	12
LA DEMARCHE DU VISITEUR	19
LA PARTICIPATION DES INTERNAUTES	21
CONCLUSION	25