



Bibliothèque
nationale de France



Usages des bibliothèques électroniques en ligne

Projet BibUsages - Rapport final

Houssem Assadi (responsable du projet), Thomas Beauvisage
- France Télécom R&D -
France de Charentenay, Thierry Cloarec, Catherine Lupovici, Hélène Toutut,
- Bibliothèque Nationale de France -

Avec la collaboration de : J.P. Savary, C. Ziemlicki, (FTR&D), L. Micault et D. Maillet (BnF)

Etude réalisée avec le soutien du RNRT

25 juillet 2003 – version 1.1

Contact : houssem.assadi@francetelecom.com

Introduction

Le présent rapport est le livrable final du projet BibUsages, mené en partenariat entre France Télécom R&D (Direction des Interactions Humaines / laboratoire Usages, Créativité, Ergonomie) et la Bibliothèque Nationale de France (Département de la Bibliothèque Numérique). Ce projet a bénéficié d'un soutien financier¹ du Réseau National de Recherches en Télécommunications (RNRT), il s'est déroulé pendant l'année 2002.

Le projet BibUsages a pour objectif l'étude des usages des bibliothèques électroniques en ligne. De tels usages sont innovants mais ils s'insèrent dans des pratiques stabilisées, en particulier au sein de la population des enseignants et des chercheurs, mais également auprès du grand public. L'accès immédiat à un corpus volumineux d'œuvres permet à des chercheurs d'envisager des études inédites, car techniquement impossibles auparavant. Par ailleurs, les enseignants, du collège à l'université, trouvent dans les bibliothèques électroniques une ressource pédagogique inestimable.

L'objectif principal du projet est de décrire les usages des bibliothèques en ligne et en particulier ceux de Gallica, la bibliothèque électronique en ligne de la BnF, en les croisant avec les caractéristiques de la population des utilisateurs. Cette étude permet également de mettre en évidence la manière dont des usages émergents infléchissent et modifient des pratiques bien établies ; dans le cas présent, la recherche académique et l'enseignement, ainsi que les pratiques de lecture et de manipulation de textes en ligne en général (lecture à l'écran, téléchargement de textes et d'ouvrages, etc.).

D'une part, il s'agit d'expliquer, par des méthodes issues des sciences sociales et cognitives, des usages déjà largement diffusés (il existe déjà plusieurs bibliothèques électroniques librement consultables sur le Web) mais dont une compréhension plus rigoureuse permettrait d'élargir la palette des fonctionnalités proposées tout en s'adaptant aux besoins et aux caractéristiques des utilisateurs.

D'autre part, cette étude met en œuvre des technologies innovantes de capture et d'analyse de trafic associées à une approche qualitative à base d'entretiens. Il s'agit de développer une approche centrée utilisateur complète, qui reste rarement mise en œuvre dans les études d'usages d'envergure sur le Web.

Le présent rapport est organisé en quatre sections : la première décrit globalement le déroulement du projet et la méthodologie adoptée, la seconde donne une description générale des trois groupes d'utilisateurs observés selon les méthodologies retenues. La troisième section présente les résultats de l'analyse du trafic et des entretiens relatifs aux usages du Web et des bibliothèques électroniques. Dans la quatrième section, nous nous penchons, à travers les éléments issus des entretiens, sur le statut des documents numériques et leur insertion dans les pratiques établies, et identifions des profils types d'usages et d'utilisateurs sur la base des différentes analyses pratiquées.

¹ Subventions n° 01 S 0165 (BnF) et 01 S 0166 (FTR&D). Nous tenons ici à remercier M. Ilarion Pavel (Ministère de la Recherche) pour son soutien et l'intérêt qu'il a montré pour le projet.

1. Déroutement du projet et méthodologie

1.1 Présentation du projet et objectifs

Le projet BibUsages vise à étudier et analyser les usages des bibliothèques électroniques. Il répond à trois objectifs principaux :

- Décrire les usages des bibliothèques électroniques en ligne en les croisant avec les caractéristiques de la population d'utilisateurs.
- Mettre en évidence la manière dont des usages émergents infléchissent et modifient des pratiques bien établies.
- Avoir une vue prospective sur les services liés aux contenus numériques sur des réseaux haut-débit.

Par bibliothèques électroniques, nous désignons en premier lieu les fonds numérisés mis en ligne par les institutions culturelles, au premier rang desquelles la Bibliothèque Nationale de France. Son site Gallica, point d'entrée vers les fonds électroniques, propose près de 70 000 textes en consultation et téléchargement (romans, essais, traités, périodiques, etc.), dont la majorité est numérisée en mode image, ainsi que des documents sonores et des fonds d'images (photographies, cartes, etc.). L'accès à ces collections est possible soit en suivant des parcours thématiques guidés, soit en interrogeant le moteur de recherche interne du site.

Si l'étude accorde une place importante à l'accès à Gallica, elle ne s'y limite pas : nous avons souhaité voir comment l'accès à Gallica s'articulait à l'usage d'autres fonds numérisés, comme les fonds privés ou associatifs (ex. : le site de l'ABU, Association des Bibliophiles Universels) et les sites de revues en ligne (ex : revues.org). Plus encore, nous avons étendu cette recherche aux sites proposant de manière générale des « contenus à lire » : journaux en ligne, sites de vente de livres récents et anciens, édition électronique.

Pour cela, nous avons mis en place un dispositif d'analyse des usages permettant à la fois d'obtenir des informations quantitatives et qualitatives : d'une part, questionnaire et entretiens permettent d'explicitier les intentions, les attentes et les pratiques hors Web des utilisateurs ; d'autre part, un dispositif de recueil de trafic permet de connaître pour chaque utilisateur l'ensemble des pages qu'il visite et offre ainsi une vue objective et quantifiée de ses usages du Web. Ces deux dispositifs complémentaires fonctionnent de manière convergente, et permettent de dresser un portrait à grain fin des utilisateurs et des usages des bibliothèques électroniques, ainsi que des pratiques qui y sont liées.

1.2 Phases du projet et méthodologie

Le projet s'est déroulé en trois phases :

1. Enquête de cadrage *via* un questionnaire en ligne (mars 2002).

2. Constitution d'un panel et recueil de trafic Web pour ce panel (avril-décembre 2002).
3. Enquête qualitative par entretiens (octobre 2002).

1.2.1 Phase 1 - Enquête de cadrage (mars 2002)¹

Un questionnaire a été soumis aux visiteurs du site Gallica en mars 2002 durant trois semaines. Il a permis à la fois d'avoir une connaissance plus précise du public de Gallica, et de recruter les volontaires pour faire partie du panel d'utilisateurs dont le trafic Web a été enregistré. Le questionnaire utilisé est fourni en annexe (§6.1).

Outre les caractéristiques socio-démographiques des répondants, le questionnaire s'articule autour de deux thématiques principales : d'une part, l'usage de Gallica (fréquence des visites, rubriques consultées, etc.), et d'autre part les usages d'Internet en général (intensité d'usage, services utilisés, types de sites visités, etc.). A la fin du questionnaire, les répondants se sont vus proposer de participer au panel d'utilisateurs mis en place.

Au terme de cette première étape, 2340 personnes ont répondu au questionnaire, et 589 ont donné leur accord de principe pour faire partie du panel d'utilisateurs, soit près d'un quart.

1.2.2 Phase 2 - Constitution d'un panel, recueil de trafic (avril-décembre 2002)

Dans un deuxième temps, nous avons constitué un panel représentatif de la population totale des répondants au questionnaire. 97 volontaires ont été retenus, au sein desquels 72 ont effectivement téléchargé et installé le dispositif de recueil de trafic. Le recrutement des volontaires a suivi plusieurs étapes, chacune engendrant une déperdition du nombre de participants :

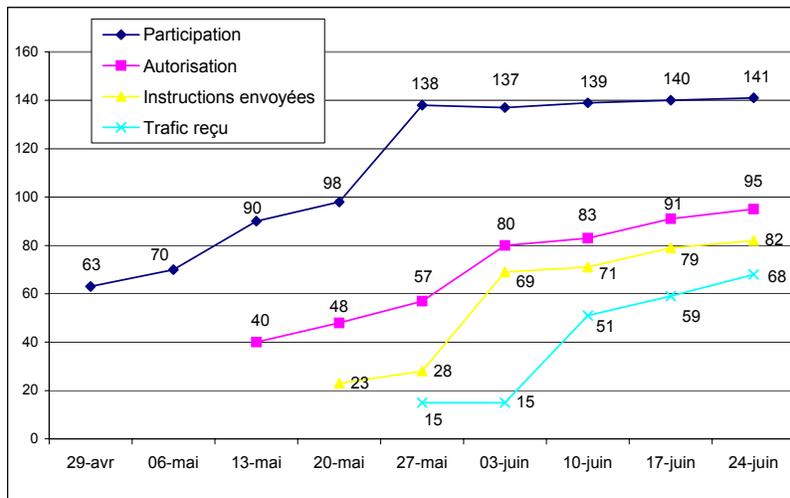
- Envoi d'un e-mail de demande de confirmation de participation.
- Envoi par courrier d'une autorisation de recueil de trafic, à retourner signée par les participants.
- Envoi par e-mail des instructions de téléchargement et d'installation du dispositif de recueil de trafic.

Les premières données de trafic nous parviennent fin mai 2002, mais l'installation du dispositif de recueil n'est achevée sur la quasi-totalité des postes que début juillet (voir Figure 1). A partir de juillet, l'ensemble des volontaires avaient installé le dispositif de recueil de trafic sur leur poste ; leur activité Web a été enregistrée pendant six mois, jusqu'en décembre 2002².

¹ Cette phase du projet a bénéficié de la collaboration de la société Ipsos. Nous tenons à remercier Laurence Lelièvre pour l'efficacité de son action.

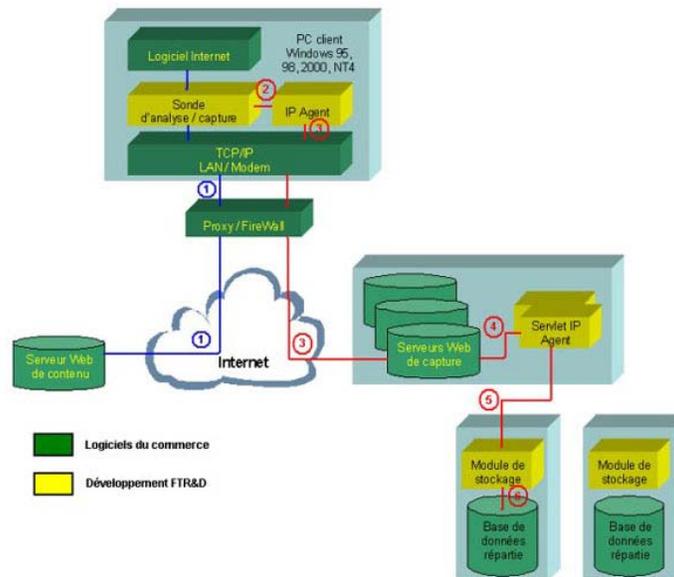
² Nous tenons particulièrement à remercier les personnes qui ont accepté de participer à cette étude, en acceptant d'installer le dispositif de recueil de trafic et en participant aux entretiens que nous avons menés. C'est grâce à leur collaboration et au temps qu'ils ont consacré à ce projet que celui-ci a pu être mené à bien.

Figure 1 - Récapitulatif des inscriptions



Le dispositif de recueil de trafic « Audinet »¹ fournit la liste horodatée de l'ensemble des URL visitées par chaque utilisateur, soit l'ensemble du trafic Web. La Figure 2 résume l'architecture suivie par Audinet.

Figure 2 - Architecture du dispositif « Audinet »



Le logiciel client est installé sur la machine des internautes ; durant l'installation, la sonde d'Audinet est insérée dans le cœur du système d'exploitation Windows (Windows 98, 2000 , NT 4 et XP sont supportés). Elle devient automatiquement active lorsqu'un logiciel client (*Internet Explorer* ou *Netscape Navigator* par exemple) accède à Internet. L'ensemble des échanges réseau est alors analysé, puis le résultat de l'analyse transmis

¹ Logiciel développé par Laurent Rabret (FTR&D/DAC) que nous tenons à remercier ici pour sa disponibilité pour le support technique du logiciel « Audinet ».

vers un serveur de collecte via le protocole HTTP ou HTTPS du Web. La sonde qui capture les flux Internet a été optimisée pour avoir un impact minimal sur les logiciels utilisés par les clients. Les informations recueillies sont ensuite envoyées vers une application locale qui, à son tour, transmet les données vers le serveur de collecte. Notons que ces données ne sont transmises que lorsque le trafic réseau est faible. Ainsi, le client ne constate aucune dégradation de performance du réseau après avoir installé « Audinet ».

In fine, les données recueillies, stockées sous formes de tables dans une base de données, contiennent pour chaque protocole la trace de chaque requête, c'est-à-dire l'heure exacte de l'action, l'URL demandée, le *Referer*, le code HTTP retourné par le serveur et les volumes de données transmis. Une plateforme d'outils de Web Mining développée à France Télécom R&D permet de reformater ces données et de les préparer pour l'analyse : identification de sessions, reconnaissance de portails, typage des sites et des parcours en sont les éléments principaux.

1.2.3 Phase 3 - Enquête qualitative sur un sous-panel (octobre 2002)

En complément des données de trafic, nous avons mené des entretiens semi-directifs auprès de 16 des 72 participants à l'expérimentation. Les entretiens ont été organisés en juillet et réalisés en octobre 2002. Sur les 72 membres actifs du panel, 50 ont été contactés afin d'obtenir leur accord de participation aux entretiens, et ce en s'intéressant aux membres les plus actifs en fonction de leur trafic Internet.

Le questionnaire de Gallica de mars 2002 et le trafic enregistré par « Audinet » permettent de dégager des profils d'usages pour chaque utilisateur interviewé, en particulier autour des types de contenus visités et de l'intensité des pratiques. Néanmoins les modalités et les contextes d'utilisation restent à approfondir. Les entretiens comblent ce vide en permettant d'une part de confirmer les pratiques telles qu'elles émergent de l'analyse du trafic et d'autre part de les inscrire dans leur contexte (usages d'Internet en général – et pas seulement du Web – et les pratiques hors-ligne). Il s'agit ainsi de voir comment la consultation des bibliothèques numériques s'inscrit dans une pratique générale d'Internet dans un contexte donné, et de mieux connaître les différents types d'usages, les motivations et les modalités de la pratique avec comme appui les données de trafic.

Pour mener ces entretiens, une grille d'entretien reprenant en filigrane ces objectifs, fournie en annexe (§6.2), a été élaborée autour de trois axes :

1. Une première partie centrée sur l'utilisation générale d'Internet permet de mieux cibler le profil général de l'internaute dans ses pratiques : durée, motivations, contexte de l'utilisation, modalités des recherches et traitement de l'information. En outre, cela permet d'avoir des renseignements sur les usages hors Web (chat, mail, forums, *peer-to-peer*...) que la sonde Audinet, dans la version utilisée pour cette étude, n'enregistre pas.
2. La deuxième partie de la grille se concentre sur l'usage des bibliothèques électroniques et de Gallica en particulier. Il s'agit de connaître les contextes d'utilisation des fonds numériques, les méthodes de recherche et les modalités

de traitement de l'information. Dans la discussion, on cherche également à pressentir des difficultés et à obtenir des propositions d'amélioration dans la conception du site Gallica.

3. La troisième partie de l'entretien se concentre sur les pratiques « off-line » : il s'agit ici de relier l'utilisation des bibliothèques électroniques à celle des bibliothèques classiques et, plus largement, aux pratiques de lecture et aux pratiques culturelles des interviewés.

Pour chaque entretien, une fiche descriptive du panéliste a été élaborée à partir de ses réponses au questionnaire de Gallica (mars 2003) et des statistiques de son trafic Internet déjà recueilli *via* Audinet. En s'appuyant sur ces informations, l'entretien va permettre alors de confirmer et d'explicitier les pratiques observées.

La richesse de la méthodologie suivie réside dans l'exploitation conjointe et croisée des données issues des trois phases du projet. Le questionnaire donne une description précise de la population d'ensemble et nous permet ainsi de bien situer l'analyse plus fine des phases 2 et 3 dans un cadre général. L'approche qualitative permet de valider des hypothèses issues de la phase d'analyse de trafic, de même que l'analyse de trafic permet de consolider les conclusions issues de l'analyse des entretiens, en les appuyant sur des mesures objectives des usages.

1.3 Etat de l'art

1.3.1 Etudes d'usage des bibliothèques en ligne

Le public des bibliothèques traditionnelles est relativement bien connu, grâce aux études et enquêtes menées par les principales bibliothèques institutionnelles auprès de leurs lecteurs. En revanche, il n'existe pas à notre connaissance d'étude générale des utilisateurs distants des bibliothèques numériques, population constituée d'universitaires, chercheurs et étudiants, mais également des professeurs et élèves du secondaire, et de particuliers menant des recherches à titre personnel.

La conférence "*The economics and usage of digital library collections*", hébergée en mars 2001 par la Librairie Universitaire de l'Université du Michigan, tente d'apporter quelques éléments de réponse, principalement sur la base de questionnaires adressés aux publics des fonds numériques. Les conclusions de cette conférence résument assez bien la situation présente : l'accès aux fonds numérisés accroît indiscutablement l'usage des documents, mais la variété des usages qui y sont liés demeure insaisissable. Qui utilise ces documents ? Quels sont les objectifs des utilisateurs ? Que représente cet usage pour eux ? Ce sont autant de questions qui restent largement sans réponse.

L'ARL (*Association of Research Libraries*) a initié en octobre 1999 un programme intitulé *New Measures Initiative* afin d'encourager les projets d'études statistiques sur les usages des bibliothèques électroniques. La communauté des bibliothécaires travaille ainsi à l'élaboration d'outils de mesure standardisés pour l'analyse des usages des services disponibles en ligne. Les approches retenues proposent de combiner l'analyse

des journaux (*logs*) de transactions enregistrés sur les serveurs (compte de sessions, de téléchargements, *clickstream*, etc.) et les mesures côté utilisateur (interaction avec le système, satisfaction de l'utilisateur, usage local/distant, etc.).

Des études existantes ont déjà montré les informations qui peuvent être recherchées par la méthodologie se fondant sur l'analyse des fichiers *logs* sur les serveurs. Dans [7] et [8], Jones *et al* proposent une analyse de *logs* de transactions de la bibliothèque numérique de Nouvelle-Zélande. Ils ont ainsi observé que les sessions d'utilisateurs sont très courtes, et comportent très peu de requêtes, celles-ci étant souvent simples ; ces observations amènent les auteurs à proposer des améliorations de conception pour l'interface de recherche. Des études similaires ont été menées concernant d'autres bibliothèques numériques, voir en particulier l'étude récente de Sfakakis et Kapidakis [9].

Bien qu'elles apportent des informations utiles et détaillées sur les usages des bibliothèques numériques, les approches « centrées serveur » laissent de côté deux questions : elles n'ont aucune information sur les caractéristiques des visiteurs (âge, sexe et caractéristiques socio-démographiques), et elles ne savent rien des contextes d'usages (ceux du Web et plus généralement ceux des pratiques culturelles) dans lesquels l'accès aux fonds numérisés se situe.

La revue de travaux effectuée par Brian-Kinns and Blandford donne un point de vue plus large des études d'usages des bibliothèques électroniques [2] ; elle montre la grande diversité des méthodes d'investigation mises en œuvre, des approches purement qualitatives jusqu'aux analyses de volumineux fichiers de transactions (*logs*). Mais on peut constater à la lecture de cette revue que la plupart des études citées traitent de la conception des bibliothèques numériques (interfaces homme-machine, outils de recherche) plutôt que des usages et des objectifs des utilisateurs.

En combinant dans notre étude questionnaire en ligne, données de trafic centrées-utilisateur et entretiens, nous souhaitons par cette méthodologie articulée combler ce manque et donner une description élargie et contextualisée des usages et des utilisateurs des bibliothèques électroniques.

1.3.2 Etudes d'usage d'Internet

Très peu d'études ont comme matériau des données de navigation enregistrées du côté de l'utilisateur d'Internet, que celui-ci soit individuellement identifié ou non ; nous en comptons au total quatre.

La première étude travaillant sur l'analyse des parcours et des usages du Web sur la base de données côté utilisateur est celle de L. Catledge et J. Pitkow en 1995, *Characterizing browsing strategies in the World-Wide-Web* ([3]). Travail fondateur, cet article présente l'analyse de données recueillies au niveau du navigateur XMosaic pendant trois semaines pour 107 utilisateurs. Ces données contiennent non seulement les URL visitées par l'utilisateur et la date, mais aussi l'ensemble des actions sur le navigateur (*Selection of hyperlink in document, Go back one document, Go to the Home document*, etc.). L'étude montre un certain nombre de résultats intéressants :

1. découpage en sessions : le temps moyen entre deux actions sur le navigateur est de 9,3 minutes. Partant, le temps d'inactivité retenu pour définir la fin d'une session est de 25,5 minutes.

2. séparation par protocole : 80 % de la navigation se fait par HTTP, dont 4 % générés par CGI¹ (contenu dynamique).
3. méthode d'interaction : 52 % des actions de navigation sont faites par le suivi de liens dans les pages ; 41 % par des "Back". Les raccourcis clavier ne sont presque jamais employés.
4. séquences répétées : une corrélation linéaire est observée entre le nombre moyen de pages vues sur un site dans une visite et le nombre de visites du site : beaucoup de sites sont vus sur une faible longueur, peu sur un longueur importante.
5. à l'intérieur d'un site, les auteurs constatent une stratégie « spoke and hub », c'est-à-dire une très forte utilisation du ``Back" pour des séries d'avant-arrière autour d'une page-pivot. Ceci suggère que cette forme de navigation est indépendante du nombre de liens proposés dans une page.
6. autre méthode de navigation souvent observée : les pages personnelles sont utilisées comme une sorte d'index vers des pages intéressantes.

Ces résultats commencent à être un peu anciens en regard de l'évolution du Web et de la diffusion de son accès ; cela étant, les résultats et la méthode n'en sont pas moins intéressants, en particulier dans la capacité à lier les pages vues à des modes d'interaction sur le navigateur, ainsi qu'à tenir compte de la temporalité de la navigation.

A la même époque, C. Cuhna et A. Bestavros et M. Crovella présentent dans [5] les résultats de l'analyse de traces de navigation côté-client sur près d'un million de requêtes ; comme dans l'étude de Catledge et Pitkow, le navigateur Mosaic sert de support au recueil de données : 37 ordinateurs partagés (stations de travail Sun) sont équipés du dispositif de recueil dans des salles d'une université d'informatique. Entre autres résultats intéressants, nous notons particulièrement l'observation que le trafic répond à une loi de Pareto (« power-law distribution ») : une faible part des documents concentre la majorité des requêtes, tandis qu'un grand nombre de pages ne sont vues que très rarement. Cette loi est également observée en ce qui concerne la taille des fichiers, ce qui intéresse particulièrement les deux auteurs dont la perspective est d'améliorer les techniques de *cache* afin d'augmenter la rapidité d'accès aux pages Web.

La troisième étude, présentée en 1997 par L. Tauscher et S. Greenberg ([10]) se penche sur la « revisite » de pages Web : l'objectif est d'élaborer des modèles de revisite, et d'en tirer des conclusions pour la conception des systèmes d'historique des navigateurs. Les données analysées sont constituées par les traces de navigation de 23 utilisateurs pendant six semaines. L'analyse montre qu'en moyenne 58 % des requêtes pour un utilisateur donné pointent vers une page qu'il a déjà visitée, en même temps que le « vocabulaire » des URL ne cesse de croître avec le temps. Des entretiens ont été menés par la suite avec les utilisateurs, qui montrent que les causes de la visite de nouvelles pages sont essentiellement : 1) le besoin de nouvelles informations ; 2) le désir d'explorer un site en particulier ; 3) la page est recommandée par un collègue, et 4) la page a été trouvée en cherchant autre chose. L'étude croise également ces données avec un enregistrement des actions sur le navigateur similaire à celui pratiqué dans le travail de Catledge et Pitkow.

¹ *Common Gateway Interface*

Enfin, en 2001, A. Cockburn et B. McKenzie présentent dans *What do Web users do ? An empirical analysis of Web use* ([4]) une analyse des usages du Web sur la base de traces d'usage. La méthode de recueil de données consiste à utiliser les fichiers d'historique et de favoris de Netscape pour 70 utilisateurs durant 4 mois ; pour chaque page, sont notés, outre l'URL, les dates de premier et de dernier accès et le nombre de visites, avec une collecte chaque jour. Dans les données récupérées, n'apparaissent que les URL explicitement demandées par l'utilisateur, les paramètres des CGI se trouvent tronqués, et les *frameset* sont regroupés sous une seule page. Cockburn et McKenzie font plusieurs constats :

- Croissance assez régulière du vocabulaire (les pages) dans le temps, de manière globale ainsi que pour chaque utilisateur.
- Forte corrélation entre le nombre de visites et le vocabulaire. En moyenne, quatre pages sont revisitées pour une page nouvelle.
- Taux de revisite : sur l'ensemble du trafic, ce taux est de 81 %.
- Pour chaque utilisateur, peu de pages sont visitées très régulièrement (en moyenne, 24 % du vocabulaire de chaque individu) : on retrouve un comportement de type « loi de Zipf ». Souvent, les utilisateurs ont des raccourcis vers les deux pages les plus vues (page de démarrage ou favori).
- La majorité des pages sont vues 1 seconde ou moins : « browsing is a rapidly interactive activity ».
- Utilisation des favoris : ils sont nombreux chez tous les sujets. Dans le temps, le nombre d'ajouts est supérieur au nombre de suppressions, de sorte que la taille des favoris ne cesse d'augmenter.
- Les utilisateurs, bien qu'appartenant au même département de l'université, voient des espaces très différents sur le web : 91 % des pages visitées hors du site de l'université n'ont été vues que par un seul utilisateur.

Les auteurs concluent sur les implications de ces résultats pour la conception de navigateurs.

Ces études sont très précieuses ; le faible nombre de travaux « centrés utilisateurs » tient à la rareté des données de ce type, et ces quatre travaux fournissent d'importants résultats en termes statistiques et méthodologiques. Deux reproches peuvent toutefois leur être faits : d'une part, les données concernent toujours des catégories particulières d'utilisateurs (étudiants en informatique le plus souvent), et pour une durée relativement limitée (quatre mois au maximum).

Le laboratoire « Usages, Créativité, Ergonomie » de France Télécom R&D mène, depuis 1999, des actions de recherche dans le domaine des usages d'Internet dans une approche centrée utilisateur [1]. Ces recherches s'appuient sur des données issues du panel résidentiel de la société de mesure d'audience sur Internet Netvalue¹. Notre étude BibUsages se situe dans le cadre méthodologique mis en place par ces recherches.

¹ Actuellement, ces recherches se poursuivent dans le cadre du projet RNRT SensNet, partenariat entre FTR&D, NetValue, CNRS/LIMSI et université Paris 3 : http://www.telecom.gouv.fr/rnrt/projets/res_01_39.htm

2. Données de cadrage

Nous présentons dans cette partie les caractéristiques de nos trois groupes d'utilisateurs : les 2340 répondants au questionnaire, les 72 participants au panel, et les 16 personnes interviewées. Il ne s'agit pas ici d'une analyse exhaustive, mais d'un résumé des caractéristiques les plus marquantes de la population étudiée, en la situant en particulier dans le cadre général de la population d'internautes français.

2.1 Profil général à l'issue du questionnaire

En premier lieu, on note une part importante de visiteurs étrangers : parmi les répondants à l'enquête, 31,4% déclarent résider à l'étranger ; en outre, pour 60% d'entre eux, le français n'est pas leur langue maternelle (mais ils la pratiquent assez pour répondre au questionnaire). Gallica s'impose ainsi comme un point d'accès à des fonds francophones depuis l'étranger.

2.1.1 Caractéristiques socio-démographiques

Dans le "lectorat" de Gallica, tel qu'il ressort de l'enquête, on constate les tendances fortes suivantes :

- Une représentation majoritaire des 50-54 ans (13,2%), mais les jeunes, scindés en deux groupes (moins de 25 ans et 25-29 ans) sont également très présents (11,8% dans les deux cas).

Tableau 1 – Répartition des répondants au questionnaire par classe d'âge

Age	% Population
Moins de 25	11,8 %
25 à 29 ans	11,8 %
30 à 34 ans	11,3 %
35 à 39 ans	10 %
40 à 44 ans	11 %
45 à 49 ans	11,4 %
50 à 54 ans	13,2 %
55 à 59 ans	9,5 %
60 ans et plus	10 %

- Une représentation majoritaire des cadres de la fonction publique, très loin devant les autres catégories (l'enseignement supérieur est le secteur d'activité le plus représenté).
- Un niveau d'études Bac+3, Bac+4 très homogène (voir Tableau 2)

Tableau 2 - Répartition des répondants par niveau d'études

Niveau d'études	% Population
ST Niveau primaire ou secondaire	8,8 %
Aucun diplôme / certificat d'études primaires	2,7 %
BEPC, Brevet des collèges	2,8 %
CAP, BEP	3,3 %
Baccalauréat, Capacité en droit	9,1 %
Diplôme universitaire de 1er cycle (Bac+2) : Deug, BTS, DUT ...	11,2 %
Diplôme universitaire de 2ème cycle (Bac+3 ou Bac+4) : Licence, Maîtrise	22,6 %
Diplôme universitaire de 3ème cycle (Bac+5) : DESS, DEA	17,5 %
Diplôme d'Ingénieur	6,2 %
Diplôme d'une Grande Ecole autre qu'une Ecole d'Ingénieur	3,8 %
Doctorat, Thèse	20,8 %

2.1.2 Les « gallicanautes » par rapport aux internautes français

En ce qui concerne l'usage de l'Internet, les personnes interrogées ont un haut degré de pratique, utilisant les services Web de façon très régulière (recherche d'informations ou opérations bancaires ou boursières, achats en ligne), tout cela majoritairement à partir d'un accès à domicile. En même temps, il faut remarquer que ces mêmes utilisateurs sont d'anciens internautes (37% d'entre eux sont connectés à Internet depuis 1997 ou antérieurement).

On voit donc que le profil général des utilisateurs rencontrés dans cette enquête se situe dans un contexte de régularité et de fidélité, de connaissance de l'outil, avec un usage personnel à partir du domicile et dans des sessions plutôt longues.

Pour analyser la spécificité du public de Gallica par rapport aux internautes français, nous recourons aux chiffres donnés par NetValue en décembre 2001.

Tableau 3 – « Gallicanautes » résidant en France et internautes français

		France - données NetValue décembre 2001	Questionnaire BibUsages en ligne (mars 2002)
Homme/femme		58% / 42%	69,3% / 30,7%
Urbain		81,0%	86,4%
Age	Moins de 25 ans	29,3%	9,8%
	25-34 ans	23,1%	19,2%
	35-44 ans	31,8%	21,1%
	45-54 ans	12,5%	25,6%
	Plus de 55 ans	3,3%	24,3%
Bas débit / haut débit		91,1% / 8,9%	67,2% / 32,8%

Les caractéristiques générales (voir Tableau 3) montrent que les gallicanautes sont, par rapport aux internautes français, plutôt des hommes, globalement plus âgés (sur-représentation des tranches 45-54 ans et plus de 55 ans). L'écart le plus important concerne le type d'équipement Internet : le haut débit est fortement sur-représenté, avec un taux d'équipement de près de 33% chez les utilisateurs de Gallica, contre 9% en général à la même époque. Les gallicanautes sont également des internautes plus « anciens », avec 67,2% des répondants résidant en France disposant d'une connexion depuis 1999 au moins.

Tableau 4 - Spécificités des gallicanautes résidant en France : PCS

France - données NetValue (décembre 2001)		Questionnaire BibUsages en ligne (mars 2002)	
Retraité(e)s	6,2%	Retraité	10,1%
Sans professions, au foyer	9,4%	Au foyer	1,3%
Temporairement sans emploi	1,6%	Temporairement sans emploi	2,5%
Étudiant(e)s	23,3%	Etudiant	13,1%
Ouvriers, ouvrières	4,4%	Etudiant ayant une activité rémunérée	5,3%
Agriculteurs, pêcheurs	0,3%	Ouvrier	0,3%
Employé(e)s	18,8%	Agriculteur, exploitant	0%
Professions intermédiaires	24,2%	Commerçant, artisan	0,5%
Professions libérales	1,5%	Employé, personnel de service	6,5%
Dirigeants	8,4%	Technicien, agent de maîtrise, cat. B	10,0%
Propriétaires d'une entreprise	1,5%	Cadre du secteur privé	13,8%
		Cadre de la fonction publique (cat. A)	26,5%
		Profession libérale	6,0%
		Chef d'entreprise, cadre dirigeant	4,0%

L'examen des professions et catégories socio-professionnelles des répondants au questionnaire (voir Tableau 4) complète l'analyse en montrant une sur-représentation des CSP supérieures (cadres, enseignants) et des étudiants, en particulier en 2^e et 3^e cycle, parmi les visiteurs de Gallica.

2.2 Composition du panel

Le panel a été composé de telle manière qu'il reflète le plus fidèlement possible les caractéristiques des répondants au questionnaire présenté sur Gallica en mars 2002. Nous présentons ici les caractéristiques de ce panel, qui reprend de manière générale celles des « gallicanautes » présentées précédemment.

Profil socio-démographique

Le panel essentiellement masculin (3/4 d'hommes). La moyenne d'âge est de 46 ans ; toutefois, cette moyenne ne met pas en évidence la sur-représentation des panélistes de plus de 55 ans, avec plus de 31 % du panel se situant dans les deux dernières tranches d'âge contre 19% pour l'enquête de Gallica.

Les panélistes exercent majoritairement une activité professionnelle (82%) et pour 45% d'entre eux, occupent un poste cadre de la fonction publique, dans le secteur de l'enseignement ou de la recherche.

Dans le cadre des panélistes n'ayant pas d'activité rémunérée (18 %), plus de 60 % sont des retraités, alors qu'ils n'étaient que 23,4% dans l'enquête générale. De plus il faut noter que 20% d'entre eux sont des étudiants de deuxième cycle.

En ce qui concerne la provenance géographique du panel, la diversité des régions françaises a été respectée avec une sur-représentation de Paris et la région parisienne. Cependant, en raison de contraintes d'organisation, le panel est exclusivement composé de personnes résidant en France : les répondants étrangers à l'enquête n'y sont donc pas du tout représentés.

Profil Internet et Gallica

Plus de 47% des panélistes utilisent Internet depuis 1997, de manière quotidienne et dans la majorité des cas de leur domicile (plus de 62%). Ils sont majoritairement équipés d'accès haut-débit, puisque 29 des 72 participants ont une connexion par ADSL ou Câble.

Le questionnaire en ligne nous apprend que les panélistes utilisent principalement Internet pour rechercher des informations (30 %) et échanger à partir des différents modes de communication (mail, chat, forum). Leurs principaux centres d'intérêts sur le Web sont assez hétérogènes, avec une prédominance de l'art et de la littérature, des sciences sociales et de la recherche documentaire ou bibliographique.

Le panel est une population d'internautes fidèles à Gallica : plus de 47% déclarent en être à plus de dix visites. Ils fréquentent le site de manière régulière et pour la plupart de chez eux. Leurs visites sont d'une durée supérieure à 10 minutes : 38 % restent entre 10 à 30 minutes et 32% plus d'une demi-heure. Les panélistes ont connu Gallica par différents moyens : brochure BnF, lien d'un autre site et par un moteur de recherche ou un annuaire Web.

2.3 Profil général des personnes interrogées en entretien

Le profil socio-démographique des personnes interrogées reste dans la lignée du panel et de la population générale de l'enquête de Gallica (voir Tableau 5). La plupart des personnes interrogées occupent des postes de cadre de la fonction public ou du secteur privé ; la moyenne d'âge se situe autour de 48 ans avec une fourchette allant de 33 à 76 ans, et la majorité habitent en des milieux urbains.

Les participants ont en général installé « Audinet » chez eux. Pour certains, cet ordinateur est un ordinateur familial dont ils sont l'utilisateur majoritaire ; des entretiens, il ressort que l'utilisation par les autres membres du foyer reste occasionnelle. Les choix individuels se sont portés sur l'équipement le plus utilisé pour l'installation de la sonde, c'est pourquoi seules deux personnes ont installé ce logiciel sur leur lieu professionnel.

Dans le cadre de l'utilisation d'Internet, on remarque qu'il s'agit d'anciens internautes, avec en moyenne déjà plus de 3 ans de pratiques, et ayant un usage fréquent et une consommation importante (déclarée dans le questionnaire en ligne, et confirmée par les données de trafic).

Utilisateur	Sexe	Département	Tranche d'âge	Profession	Type de connexion à domicile	Ordinateur partagé avec d'autres membres de la famille	Nbre total de sessions réalisées	Lieu d'installation d'Audinet Domicile/ Travail	Ancienneté Internet
Utilisateur A	F	16	55/59	Enseignante du supérieur	Haut débit	Oui	905	D	1998
Utilisateur B	F	75	30/34	Indépendante Correctrice	Haut débit	Oui	110	D/T	1997
Utilisateur C	F	35	55/59	Cadre de la fonction publique	Haut débit	Non	pb. install.	D	1997
Utilisateur D	H	94	35/39	En Formation Sciences Sociales	Haut débit	Oui	642	D	1997
Utilisateur E	H	83	70/74	Retraité (cadre du privé)	RTC	Oui	61	D	2000
Utilisateur F	H	77	50/54	Enseignant du primaire	RTC	Oui	150	D	1998
Utilisateur G	H	75	40/44	Enseignant (Histoire)	Haut débit	Oui	626	D	1999
Utilisateur H	H	75	45/49	Cadre du privé (informatique)	Haut débit	Oui	28	D	1999
Utilisateur I	H	83	45/49	Employé, personnel de service/Commerce et distribution	RTC	Oui	667	T	1998
Utilisateur J	H	33	50/54	Cadre de la fonction publique	Numéris	Non	278	D	2000
Utilisateur K	H	27	30/34	Cadre du privé	Haut débit	Non		T	1997
Utilisateur L	H	69	30/34	Cadre de la fonction publique (CNRS)	RTC	Non	218	T	1997
Utilisateur M	H	59	50/54	Enseignant du supérieur	Haut débit	Oui	381	D	1997
Utilisateur N	F	92	55/59	cadre du privé (assurance)	RTC	Non	262	D	1997
Utilisateur O	H	91	75/79	Retraité (comptable)	RTC	Non	658	D	2000
Utilisateur P	H	15	65/69	Retraité (enseignant du secondaire, Mathématiques)	Haut débit	Non	101	D	1997

Tableau 5 - Récapitulatif des profils des panélistes interrogés en entretien

3. Analyse des usages

Nous présentons dans cette partie les principaux résultats issus de l'analyse de trafic Web recueilli auprès de notre panel de 72 utilisateurs de juillet à décembre 2002 (soit 6 mois d'observation), en même temps que les données qualitatives issues des entretiens menées avec seize d'entre eux.

3.1 Utilisation globale d'Internet

3.1.1 La messagerie

L'usage de la messagerie n'arrive pas dans les premières considérations. Elle est certes importante mais ne rentre pas dans les premières motivations personnelles d'accès à Internet. Elle permet d'échanger essentiellement avec la famille géographiquement éloignée et les spécialistes rencontrés dans le cadre des recherches. Dans cet usage, on peut noter, l'abonnement pour certains à des *newsletters* et à des listes de diffusion¹.

*"Je vais sur des sites, je correspond avec des supposés, des présumés cousins".
(Utilisateur I)*

"J'ai pu établir des contacts avec quelques conservateurs ou anciens conservateurs de musées..." (Utilisateur N)

Seuls deux panélistes utilisent les *chats* ou les messageries instantanées ; ces outils sont utilisés majoritairement comme moyen de communication synchrone et interpersonnelle avec des correspondants éloignés, et s'apparentent à des outils de substitution du téléphone.

"Si de temps en temps, je fais du chat avec la famille... On se sert également du chat pour téléphoner au Canada" (Utilisateur K)

"Soit par webcam soit par messagerie instantanée, vous voyez ? Surtout maintenir des contacts commodes à distance. Oui des choses comme ça (...)" (Utilisateur P)

3.1.2 Usages du Web : données de cadrage

Nous présentons ici les résultats relatifs au trafic global et à l'activité des panélistes ; nous les mettons en perspective avec des données issues d'un panel général

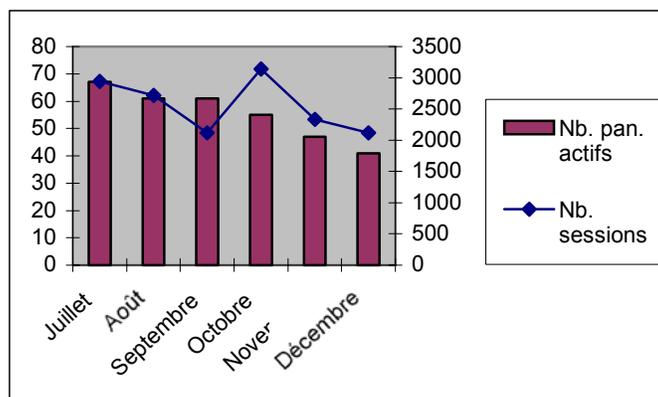
¹ Nous n'avons pas particulièrement mis l'accent sur les usages des outils de communication dans cette étude : le dispositif de recueil de trafic utilisé n'était pas paramétré pour enregistrer cet usage ; par ailleurs, nous n'avons pas cherché à développer cet aspect lors des entretiens.

d'utilisateurs représentatif de la population des internautes français fourni par la société NetValue, panel que nous nommons par la suite « panel NetValue »¹. Nous pouvons ainsi montrer les spécificités du panel BibUsages par rapport à l'ensemble des internautes français.

Activité du panel

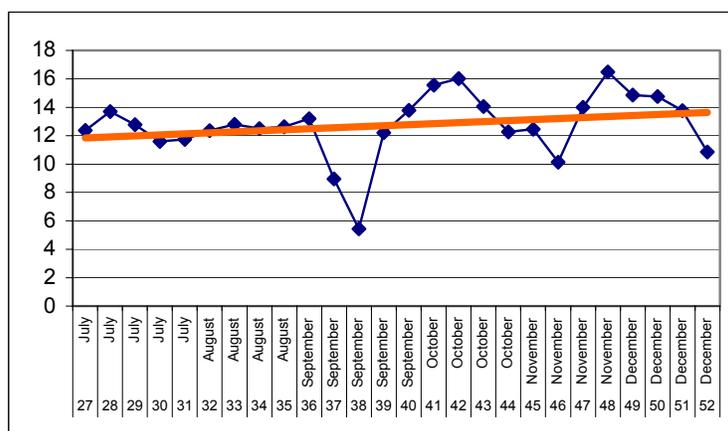
Les données d'usage du Web recueillies concernent 72 personnes observées pendant 6 mois, de juillet à décembre 2002. A l'examen du trafic généré durant cette période (Figure 3 et Figure 4), on observe une lente érosion du panel de juillet à décembre : en juillet, 68 panélistes actifs sur 72 (90%), tandis qu'en décembre, seuls 41 ont généré du trafic. Toutefois, si le nombre de sessions diminue globalement, le nombre de sessions par utilisateur a lui tendance à croître : les utilisateurs présents durant toute la période sont les plus actifs du panel.

Figure 3 - Activité du panel par mois



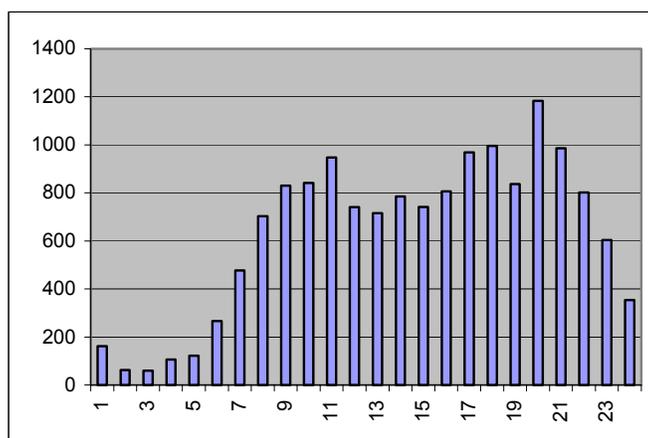
¹ Ces données concernent le trafic d'un panel de 3400 internautes résidentiels en 2001 et 2002 ; elles ont été mises à la disposition de France Télécom dans le cadre du projet RNRT SensNet (2002-2004), partenariat impliquant NetValue, France Télécom R&D, Paris III et le LIMSI.

Figure 4 - Nombre moyen de sessions par actifs et par semaine (courbe et tendance)



En ce qui concerne la répartition horaire de l'activité (Figure 5), les sessions¹ sont beaucoup plus réparties dans la journée (pas de creux en milieu de journée) par rapport au panel résidentiel NetValue. Ceci est essentiellement dû au fait que le panel BibUsages est mixte résidentiel / professionnel, à quoi il faut ajouter l'activité professionnelle à domicile.

Figure 5 - Nombre de sessions par heure de début de session



Intensité d'usage du Web

La répartition des panélistes selon l'intensité d'usage du Web en nombre de sessions révèle des écarts importants au sein du panel. Le Tableau 6, qui présente la répartition en trois groupes en fonction du nombre de sessions total effectué durant la période (juillet – décembre 2002), montre que les deux tiers des sessions sont réalisées par le tiers le plus actif du panel, tandis que 9% des sessions sont réalisées par les 13 utilisateurs les moins actifs.

¹ Une session est définie comme une période d'activité sur le Web sans période d'inactivité de plus de 30 minutes.

Tableau 6 - Répartition des panélistes par intensité d'usage (nombre de sessions)

Tiers-tile	Nombre de panélistes	Nombre de sessions	Part du total de sessions	Nb sessions min / max	Nb moyen de sessions par semaine et par personne
1	24	1186	9,0%	29 / 138	3,1
2	24	3144	23,8%	141 / 363	8,3
3	24	8866	67,2%	407 / 1041	23,5

On retrouve ici une répartition classique de l'activité sur le Web, où une minorité des internautes fait une majorité du trafic. Toutefois, nous avons globalement affaire à de gros utilisateurs du Web : avec 3,1 sessions par semaine en moyenne, les moins actifs du panel sont, par rapport à la population globale des internautes français, des utilisateurs intensifs du Web¹.

La durée moyenne des sessions est en revanche similaire à celui de l'ensemble des internautes : 30 minutes pour le panel BibUsages pour 35 minutes chez les internautes français du panel NetValue en 2002. La différence en intensité d'usage du Web se fait donc pour les gallicanautes en termes de nombre de sessions, et non en termes de durée de session.

3.2 Contenus visités

3.2.1 Audience des sites

Les portails généralistes arrivent en tête des sites visités par notre échantillon, aux côtés des outils de recherche d'information, et ce en nombre de sessions total comme en nombre d'utilisateurs (Tableau 7). Ceci va dans le sens de ce que nous apprenons au cours des entretiens, où beaucoup d'interviewés évoquent des pratiques de recherche d'information.

¹ En 2002, le tiers le plus intensif des utilisateurs du panel NetValue 2002 réalisait 3,9 sessions Web par semaine en moyenne, et les moins intensifs 1,4 sessions.

Tableau 7 - Top 20 des sites en nombre de sessions, et en nombre de panélistes

Site	Nombre de sessions	Nombre de pan.	Site	Nombre de sessions	Nombre de pan.
Wanadoo (fr)	5175	58	Google	3953	61
Google	3953	61	Wanadoo (fr)	5175	58
Voila (fr)	1289	53	www.microsoft.com	611	57
Yahoo (fr)	1248	53	membres.lycos.fr	570	56
MSN	980	48	BNF-Gallica	822	54
Yahoo (com)	882	50	Voila (fr)	1289	53
www.palmgear.com	837	2	Yahoo (fr)	1248	53
BNF-Gallica	822	54	Amazon	345	51
Le Monde	665	35	BNF-Autres	577	51
Free	618	21	www.ifrance.com	250	51
www.microsoft.com	611	57	Yahoo (com)	882	50
BNF-Autres	577	51	www.multimania.lycos.fr	330	49
membres.lycos.fr	570	56	MSN	980	48
fr.groups.yahoo.com	510	19	download.macromedia.com	130	45
GeneaNet	457	15	www.hebdotop.com	228	43
fr.rd.yahoo.com	440	28	www.pagesjaunes.fr	395	43
Caramail	438	13	Fnac	274	41
www.pdacool.com	397	2	service.bfast.com	183	38
www.pagesjaunes.fr	395	43	fr.kelkoo.com	117	37
Tiscali	376	33	SNCF	109	37

On remarque également que Gallica arrive parmi les sites les plus visités, ce qui atteste la spécificité du panel. Par ailleurs, le site du journal *Le Monde* arrive en 9^e position en nombre de sessions, et ceux d'Amazon et de la Fnac sont bien placés en nombre de panélistes. Cette première vue semble montrer un lien global entre la fréquentation des bibliothèques électroniques et les outils de recherche d'information d'une part, et les sites de « contenus à lire » et de commerce de biens culturels d'autre part.

3.2.2 Audience par types de portails et de sites

L'examen des *types* de sites et de portails visités par le panel permet d'affiner cette hypothèse. Le logiciel *&catService*, développé à FTR&D, permet d'identifier, de décrire et de regrouper, au sein de la base de trafic des utilisateurs, les adresses de pages correspondant à des portails sélectionnés manuellement. Nous pouvons ainsi analyser finement les usages des services sur tel ou tel type de site et de service sur le Web.

Tableau 8 - Audience par types de sites et de portails (juillet – décembre 2002)

Type de site	Nombre de sessions	Présents dans n% des sessions	Nombre de panélistes
Portail généraliste	8005	60,3%	71
Moteur	4183	30,1%	67
Site « perso »	3142	21,1%	69
Bibliothèque électronique	1195	8,0%	56
Media / Presse	1144	7,6%	49
e-commerce / biens culturels	833	5,5%	63
Généalogie	684	4,6%	32
WebMail	443	3,1%	15
e-commerce / Banque Bourse	439	2,9%	30
Media / Radio	305	2,0%	33
Media / TV	219	1,5%	38
e-commerce / tourisme	141	0,9%	46
e-commerce / courses	9	0,1%	4

Ici encore, nous observons un double effet « outils de recherche » et « contenus culturels » (voir Tableau 8). En premier lieu, portails généralistes et moteurs de recherche occupent une place prépondérante dans l'activité Web, et touchent la quasi-totalité du panel. D'autre part, le panel se caractérise par une très forte fréquentation des portails « culturels » : les bibliothèques électroniques, les sites de commerce de biens culturels (Alapage, Fnac, Amazon...) et les sites de médias (presse, radio, télévision) occupent une place privilégiée. En outre, les sites consacrés à la généalogie ont une place importante dans les usages, et leur présence témoigne de centres d'intérêt personnels très forts ; à cela faut-il sans doute ajouter l'audience particulière des sites personnels, présents dans 21% des sessions, qui sont d'importantes sources d'informations pour les recherches généalogiques sur le Web.

Ainsi, au sein d'un panel fréquentant régulièrement les bibliothèques électroniques, l'accès aux fonds numérisés est à mettre en parallèle avec la fréquentation de sites proposant « des contenus à lire ». L'usage des portails de type « Média / Presse » est révélatrice à cet égard (Tableau 9), avec un accès particulièrement intensif au site du journal *Le Monde*.

Tableau 9 - Fréquentation des sites de type "média-presse"

Portail	Nb. panélistes	Nb. sessions	Nb sessions par panéliste	Temps moyen par session (min.)
Le Monde	37	719	19,4	8,1
Libération	24	86	3,6	3,9
Les Echos	17	118	6,9	5,3
Nouvel Obs	14	131	9,4	4,9
Telerama	14	107	7,6	6,3
Le Figaro	13	148	11,4	14,2
New York Times	6	81	13,5	7,7
Le Point	4	12	3,0	3,4

La fréquentation des bibliothèques électroniques est également à mettre en parallèle avec la fréquentation de sites de e-commerce de biens culturels, dont le Tableau 10 montre le détail : l'usage des portails de biens culturels montre l'importance de la bibliophilie, avec une fréquentation importante des sites Chapitre.com et Livre-rare-book. Nous sommes ici probablement en présence d'une manifestation des liens vertueux pouvant exister entre les Web marchand et non-marchand.

Tableau 10 - Fréquentation des sites de vente de biens culturels

Portail	Nombre de panélistes	Nombre de sessions	Nombre de sessions par panéliste	Temps moyen dans la session (min.)
Amazon	55	386	7,0	4,1
Fnac	42	299	7,1	2,4
Alapage	29	107	3,7	2,2
Chapitre.com	26	160	6,2	9,3
Livre-rare-book	17	92	5,4	5,0
Galaxidion	6	73	12,2	6,4
Librissimo	6	7	1,2	4,5
Numilog	5	14	2,8	4,7
Eyrolles	3	4	1,3	1,0
CNRS Editions	2	8	4,0	8,6
Barnes & Noble	1	1	1,0	0,4
Cylibris	1	1	1,0	0,2
Edibook	1	1	1,0	0,1
Les Introuvables	1	1	1,0	0,3
Litraweb	1	1	1,0	1,3

Les entretiens corroborent ce constat : la plupart des personnes interrogées ont déclaré avoir déjà expérimenté l'achat en ligne, mais jugent cette pratique occasionnelle. Les thèmes récurrents sont les voyages (snf.com), les biens culturels (chapitre.com, fnac.com, alapage.com...), la VPC (laredoute.fr, camif.fr...) et l'informatique. L'importance du choix de produits proposés – y compris la disponibilité de produits rares, tels que les objets de collection, en particulier les livres anciens – et la possibilité de comparer les prix sont les éléments les plus mis en avant pour motiver le recours aux sites de e-commerce :

"Alors ça, c'est essentiellement des produits qu'on trouve difficilement en pharmacie pour dire (...) C'est un vieux produit. Bon ben ça, on le trouve pas (...) difficilement, donc on le commande." (Utilisateur I)

"Il y a la facilité en ce qui concerne certains achats, comme je disais pour les livres ou pour les voyages ou pour autre chose." (Utilisateur G)

Ces éléments généraux issus principalement de l'analyse de trafic confirment la spécificité du panel constitué pour cette étude : anciens internautes, consommateurs de contenus culturels, acheteurs en ligne, tout cela au sein de pratiques stabilisées.

3.2.3 Recherche d'information

Utilisation des Favoris

Certains interviewés utilisent les Favoris (ou Signets) comme lieu de stockage de l'information. Les rigoureux classent et hiérarchisent l'information, et se constituent ainsi une sorte de centre de documentation personnel. Suite aux entretiens, nous avons proposé à certains interviewés de nous envoyer leurs Favoris. A l'examen de ces données, trois éléments à noter :

- Un nombre élevé, allant de 50 à 1000 adresses répertoriées.

- Une diversité des sites enregistrés : sites institutionnels, sites à fort potentiel de contenu, informations pratiques.
- Un classement structuré et hiérarchisé de l'information : dossiers, sous-dossiers, sites en attente de classement avant confirmation de la qualité du contenu.

Les sites enregistrés ne sont pas simplement des sites du Web français mais sont aussi en provenance de pays étrangers.

"Les favoris, ben disons que c'est vrai que pour moi le réseau, enfin le réseau sur Internet, c'est quasiment vivant. Si vous me demandez, je vais pas vous faire de la pub, mais c'est vrai qu'avec Gallica, euh, ou que ce soit l'ANPE ou l'APEC, enfin tous les sites les plus usuels, c'est vrai que je vais regarder mes favoris et c'est vrai que je vais en avoir 400 ou 500." (Utilisateur K)

"Je les mets dans des signets, puis après, je me fais des dossiers... des classements dans mes signets." (Utilisateur N)

"Je regarde si ça m'intéresse ou pas, éventuellement, si l'information m'intéresse, je suis amené à revenir dessus, je la mets dans mes favoris." (Utilisateur D)

Ces observations vont dans le sens de la singularité du panel, composé d'anciens internautes aux pratiques avancées.

Dans l'utilisation des favoris, d'autres cas apparaissent : d'une part ceux qui ne les utilisent pas du tout, et d'autre part ceux qui les utilisent de manière plus sélective, en enregistrant simplement les pages particulières, les pages génériques étant facilement retrouvées via les moteurs.

Choix du moteur

Les outils de recherche occupent, nous l'avons vu, une place très importante dans les usages du panel en terme d'audience. Au cours des six mois d'observation, 16 moteurs de recherche différents ont été utilisés ; Google arrive nettement en tête, présent dans plus de 75% des sessions comportant l'utilisation d'un moteur de recherche, tandis que Yahoo et Voila arrivent en deuxième et troisième position (présents dans 13,7% et 11,3% des sessions avec moteur). Les méta-moteurs sont assez fréquemment utilisés : vingt-quatre panélistes y ont eu recours au moins une fois au cours de la période, dix les utilisent de manière régulière.

Cette utilisation massive de Google atteste une fois de plus l'ancienneté des pratiques des internautes qui composent le panel : chez les internautes français en général en 2002, Google n'est présent que dans 40% des sessions avec moteur, et son utilisation massive est plutôt l'apanage des internautes expérimentés.

Les entretiens permettent de retracer les trajectoires d'apprentissage et d'appropriation des outils de recherche. Globalement, les interviewés ont essayé plusieurs moteurs et éprouvés les qualités de chacun :

"Alors il m'arrive de trouver par hasard aussi des choses. Non, c'est-à-dire que j'ai eu des... c'est-à-dire que quand on a un peu, un peu d'expérience, on se souvient pas des difficultés qu'on a eues au début, quoi ; au début on patouille puis après c'est comme sur Google, on se retrouve au début avec un million de réponses, puis au bout d'un certain temps ça arrive plus ça. On sait ce qui faut pas faire, on sait pas toujours ce qui faut faire, mais on sait ce qui faut pas faire." (Utilisateur O)

"Alors un que je n'utilise jamais, c'est Yahoo. Je le trouve trop commercial. En revanche

j'utilise Google, qui a un esprit synthétique et en plus je dirais un peu plus « configurable » sinon j'utilisais Altavista. Au début j'utilisais Lycos parce qu'il était très agréable, maintenant je le trouve un peu « glan-glan », excusez-moi, un peu trop pour les personnes qui ne savent pas trop où chercher. Avant il était très synthétique, et puis il est devenu trop commercial. Maintenant j'utilise Google. Exclusivement Google." (Utilisateur M)

Pour certains, différents moteurs sont mobilisés en fonction des contextes et des contenus de recherche ; ainsi Voila est-il le plus souvent utilisé dans le cadre de recherche de type « vie pratique » tournées vers la vie hors du Web (site de l'ANPE, de la SNCF, etc.) et Google pour des recherches plus pointues.

"Peut-être que je comprends mieux comment ça fonctionne, mais c'est vraiment avec Google que je trouve le plus de choses." (Utilisateur N)

"Sinon ben tout ce qui est un peu plus spécifique je recherche par Google." (Utilisateur L)

"Pour mes recherches personnelles, ça va être Kartoo par exemple parce que je ne sais pas exactement comment formuler ma recherche." (Utilisateur H)

Contenu des requêtes

L'examen des requêtes adressées aux moteurs de recherche révèle des thèmes de recherche très spécifiques selon les utilisateurs (ex : « saint-dupont ») si l'on examine le classement en nombre de sessions (Tableau 11). Dans le classement en nombre d'utilisateurs (Tableau 12), on retrouve certes les thèmes courants chez les internautes en général et relevant du Web comme usuel et comme fournisseur de services (« sncf », « anpe », « fnac », « Michelin », « mappy »), mais également des éléments qui attestent la spécificité du panel : « bnf » arrive ainsi en troisième position.

Tableau 11 - Top 20 des requêtes en nombre de sessions

Requête	Nombre de sessions	Nombre de panélistes
villemoisson sécheltes	41	1
casanova à l'opéra	38	1
casanova saint-dupont	29	1
villemoisson-sur-orge	25	1
NOGAREDE	23	1
sncf	22	14
saint-dupont	22	1
fnac	19	8
caramail	18	2
TF1	17	3
bnf	16	6
internet explorer	14	5
villemoisson sur orge	14	1
casanova opéra	13	1
montpellier	13	1
genealogie	12	5
sète	12	1

On peut supposer que, pour des utilisateurs connaissant déjà bien les sites de la BnF et de Gallica, une telle requête reflète un moyen rapide d'arriver au site lorsque celui-ci n'est pas dans leurs Favoris. La requête « généalogie » semble en revanche attester un

véritable intérêt pour ce domaine ; on retrouve cet engouement à travers des requêtes très particulières et précises souvent adressées par un seul utilisateur, et qui consistent en des noms propres (patronymes, noms de lieu). La généalogie apparaît ainsi comme un centre d'intérêt prépondérant pour une part importante du panel ; dans ce cadre, Gallica apparaît comme une source d'information dans le cadre d'une recherche de ce type.

Tableau 12 - Top 20 des requêtes en nombre de panélistes

Requête	Nombre de sessions	Nombre de panélistes
sncf	22	14
fnac	19	8
bnf	16	6
google	10	6
internet explorer	14	5
genealogie	12	5
Michelin	7	5
gallica	6	5
darty	8	4
hp	8	4
anpe	6	4
camif	6	4
pages jaunes	6	4
ratp	6	4
mappy	5	4
paris	5	4
sexe	5	4

Dans le cadre des entretiens, certains interviewés disent se servir des opérateurs (par exemple, les guillemets pour recherche des expressions exactes ou les opérateurs booléens : "ET", "OU", "-"...), d'autres mettent en avant l'importance de la connaissance du contenu de la recherche.

"Ah ben oui, le problème c'est de trouver, c'est de trouver une astuce pour pas avoir dix mille réponses, hein. Bien souvent il faut utiliser un autre mot. Je commence à avoir un entraînement, parce que quand on est sur un sujet limité comme ça, on sait plus ou moins limiter, pas toujours." (Utilisateur O)

"Comme on disait, comment je cherche, au début de la recherche, c'est quelques mots, et plus on va dans le temps et plus on augmente son savoir, et puis on élargit le spectre de connaissances, mieux la question est posée." (Utilisateur H)

"C'est pas tellement une question de technique, c'est une question de connaissance connexe à l'outil. Parce que par exemple quelqu'un qui voudrait faire sur l'histoire de l'Algérie, il faut qu'il ait une connaissance dans sa tête de l'histoire de l'Algérie, sinon, il ne trouvera rien." (Utilisateur P)

Les panélistes mobilisent ainsi une double compétence : celle issue du maniement des outils de recherche, qui amène à utiliser au mieux chaque moteur en fonction de ses spécificités, et celle relative à la connaissance du thème de recherche, souvent précis et bien maîtrisé.

3.3 Usages des bibliothèques électroniques

3.3.1 Audience des bibliothèques électroniques

Au sein du logiciel *&catService*, 23 sites ont été identifiés comme relevant de la catégorie « bibliothèques électroniques », dont la liste est donnée dans le Tableau 13.

Tableau 13 - Liste des sites identifiés comme « Bibliothèques électroniques »

BNF-Gallica
ABU
Alex Catalogue
American Memory
Arob@ase
Athena
Berkeley DL
Bibelec
Bibliopolis
Bibliothèque de Lisieux
BN Canada - Numérique
ClicNet
CNUM
Electronic Text Center
eLibrary
Gutenberg project
Inalf
LiNuM
Mozambook
NZ Digital Library
Online Books Page
Revue.org
UMDL

Notons ici que nous avons retenu une acception large de cette catégorie en y incluant, outre les versions numériques des bibliothèques patrimoniales, des sites d'éditeurs, des bibliothèques numériques régionales ou municipales, ainsi que des collections numériques issues d'initiatives associatives. Nous constituons ainsi une catégorie assez vaste de « collections de textes d'ouvrages accessibles en ligne ».

L'audience détaillée de ces différents sites (voir Tableau 14) montre que Gallica arrive en tête, en nombre de sessions comme de panélistes, mais que l'audience des autres sites est loin d'être anecdotique. Si la richesse du fond explique la fréquentation plus forte que sur les autres sites, on peut penser, à la lumière des entretiens, que Gallica figure, pour les internautes, comme une source de textes parmi d'autres.

Tableau 14 - Audience des bibliothèques électroniques

Portail	Nombre de sessions	Nombre de panélistes	Temps moyen dans une session
BNF-Gallica	822	54	21,4
BNF-Autres	577	51	6,5
ABU	31	13	3,0
Revue.org	22	12	1,8
Bibliothèque de Lisieux	20	11	3,9
Bibliopolis	15	6	8,0
Athena	14	10	1,1
ClicNet	14	8	0,6
Online Books Page	10	4	1,7
Electronic Text Center	8	5	0,7
BN Canada - Numérique	5	3	1,0
American Memory	4	2	6,3
Arob@ase	4	3	0,4
Berkeley DL	3	1	7,3
eLibrary	3	3	0,5
Gutenberg project	2	2	0,8
Alex Catalogue	1	1	0,8
Bibelec	1	1	0,0

3.3.2 Les bibliothèques électroniques en contexte

L'examen des types de sites et de services visités au sein des sessions comportant un accès à une bibliothèque électronique permet de voir comment l'usage des fonds numériques s'articule avec les autres contenus et services disponibles sur le Web, et permet de cerner plus finement les contextes d'usages. Pour cela, nous comparons la présence des différents types de sites identifiés dans *&catService* au sein des sessions avec bibliothèque électronique et dans l'ensemble des sessions (voir Tableau 15).

Tableau 15 - Présence des différents types de sites et services dans les sessions avec accès à une bibliothèque électronique et dans l'ensemble des sessions

	Présence dans les sessions bib. élec.	Présence dans l'ensemble des sessions	Variation de la présence dans les sessions bib. élec.
Bibliothèque électronique	100,0%	8,0%	+ 88,2%
e-commerce / biens culturels	10,4%	5,5%	+ 40,8%
Sites personnels	29,6%	21,1%	+ 40,8%
Moteur	40,8%	30,1%	+ 35,7%
Généalogie	5,2%	4,6%	+ 13,4%
e-commerce / tourisme	1,0%	0,9%	+ 7,7%
Media / Radio	2,0%	2,0%	- 0,8%
Portail généraliste	55,0%	60,3%	- 8,9%
Media / Presse	5,5%	7,6%	- 27,6%
e-commerce / Banque Bourse	1,8%	2,9%	- 37,2%
Media / TV	0,8%	1,5%	- 43,0%

Nous constatons ainsi que les moteurs de recherche sont sur-représentés dans les sessions Bibliothèques électroniques, où ils sont 1,8 fois plus présents que dans l'ensemble des sessions. Cet usage fort des moteurs semble montrer l'importance de l'utilisation des bibliothèques électroniques dans un contexte de recherche d'information, où le fond numérisé est une source parmi d'autres pour trouver de

l'information. Il faut sans doute ajouter à cela les sites personnels, sur-représentés dans les sessions avec Gallica, et qui s'imposent comme sources de données sur des sujets pointus, comme la généalogie.

L'usage des sites de e-commerce de biens culturels, également sur-représentés, semble correspondre à un effet de catalogue ou de « test avant achat », ce que confirment les entretiens :

"Je vais sur le site de la BNF parce que j'y trouve quelque chose, j'ai le souvenir que je pourrai trouver quelque chose, et que je vais me consacrer un peu de temps pour voir quelques images, ou quelques documents que je ne pourrai pas consulter, ou être sûr que je vais acheter un ouvrage ; je parlais de [nom d'auteur], je me suis acheté en consultant la BNF, ça m'a beaucoup intéressé, car j'ai acheté un ouvrage qui coûtait quand même 1 200 F." (Utilisateur H)

Il s'agit ici de feuilleter un livre ancien avant de l'acquérir sur un site spécialisé de bibliophilie.

Les sites de type « Média – presse » sont, au contraire, moins présents dans les sessions bibliothèques électroniques que dans l'ensemble : si les utilisateurs des bibliothèques électroniques sont, comme nous l'avons déjà vu, globalement de gros consommateurs de journaux en ligne, les accès à ces deux types de sites ne correspondent pas aux mêmes pratiques, et se font dans des sessions et des contextes différenciés. Ce résultat illustre bien l'apport d'une analyse fine en termes de sessions, rendue possible grâce à la technologie de recueil de trafic sur le poste de l'utilisateur mise en œuvre dans ce projet ; en effet, l'analyse de l'audience (cf. Tableau 8) a montré que les usages des bibliothèques électroniques et des sites de type « média / presse » étaient globalement liés, en ce qu'ils sont particulièrement importants pour la population étudiée. L'analyse fine en termes de sessions, faite dans la présente section, apporte une information complémentaire importante : même si, globalement, la population étudiée a un intérêt et un usage fort pour ces deux types de sites, les contextes d'usage au sein de sessions de navigation sont en revanche bien distincts.

3.4 Usages de Gallica et du site de la BnF

3.4.1 Conditions de la fréquentation de Gallica

Gallica et le site général de la BnF

L'usage du site général de la BnF (adresse : <http://www.bnf.fr>) reste modéré et lors des entretiens, une confusion entre ce site et Gallica (adresse : <http://gallica.bnf.fr>) a été parfois soulignée. Lors des entretiens, certains panélistes sont allés vérifier quel site était enregistré dans leurs favoris :

"D'ailleurs je ne vais pas sur le site de la BNF, je ne vais que sur Gallica." (Utilisateur K)

"Euh j'ai mon favori, alors je ne sais plus, je vais vous dire lequel il est : « BN Opale

plus » et « Gallica », j'ai les deux." (Utilisateur B)

Dans certains cas, le site général de la BnF est utilisé comme voie d'accès à Gallica. Chez d'autres, son utilisation ne vient qu'en complément de l'utilisation de Gallica ; dans ce cas, seul le catalogue est utilisé comme base et vérification bibliographique. Enfin, les panélistes lecteurs préparent leur visite via BN Opale.

"Et sinon, je l'utilise essentiellement pour faire des recherches bibliographiques. Par le catalogue et par Opale. Catalogue Opale et voir s'il y a des livres en ligne, pour avoir des références et éventuellement par la suite, quand je vais à la BNF, localiser un livre, on sait exactement dans quel endroit il se trouve, dans quelle salle." (Utilisateur D)

"Sur le site général, alors quand je fréquentais la Bibliothèque Nationale, j'allais, je préparais mes visites. [...] Alors je m'en sers toujours pour regarder s'il existe quelque chose sur un sujet." (Utilisateur O)

Fréquence des visites

Une partie du questionnaire en ligne comportait des questions précises relatives aux usages de Gallica, les réponses à ces questions nous permettent de comprendre dans une première approche la manière dont les internautes utilisent les bibliothèques numériques.

L'usage de Gallica se caractérise d'abord par la *fidélité* : plus de 50% des répondants au questionnaire déclarent avoir l'intention de « visiter régulièrement le site Gallica dans les 6 mois prochains mois ». Ensuite, nous notons une *régularité* dans les visites : plus de 46% des utilisateurs déclarent « avoir visité le site Gallica plus de 10 fois au cours des 6 derniers mois ».

<i>Combien de fois avez-vous déjà visité le site Gallica au cours des 6 mois ?</i>	
C'est ma première visite	28,5%
moins de 5 fois	13,5%
5 à 10 fois environ	15,6%
10 à 20 fois environ	10,3%
Plus de 20 fois	30,9%
(vous ne savez pas)	1,1%

<i>Diriez-vous qu' au cours des 6 prochains mois vous consulterez le site Gallica</i>	
Régulièrement	52,1%
De temps en temps	41,0%
Rarement	3,2%
(vous n'aurez plus l'occasion de le consulter)	0,3%
(vous ne savez pas)	3,3%

Les utilisateurs de Gallica s'y connectent majoritairement depuis leur domicile : à la question « D'où vous êtes-vous connecté aujourd'hui pour consulter le site Gallica ? », ils sont 75% à avoir déclaré se connecter à partir de leur domicile.

<i>D'où vous êtes-vous connecté aujourd'hui pour consulter le site Gallica ?</i>	
De chez vous	75,6%
De votre lieu de travail	34,0%
De l'université / d'une école d'ingénieurs, de commerce...	9,1%
D'un lycée ou collège	3,5%
D'un lieu public tel qu'une bibliothèque, un cybercafé	3,2%
D'un autre lieu	1,3%

Enfin, l'usage de Gallica se caractérise par des temps de consultation exceptionnellement longs : 32% des utilisateurs déclarent passer « habituellement plus de 30mn » sur le site Gallica.

Des profils atypiques par rapport au lectorat classique de la BnF

Le lectorat de Gallica apparaît comme singulièrement différent des visiteurs traditionnels de la BnF. Ainsi, 44% des répondants au questionnaire déclarent visiter Gallica pour un usage exclusivement personnel, et s'inscrivent dans un profil de chercheurs « amateurs ».

<i>Consultez-vous le site Gallica pour un usage principalement... ?</i>	
Personnel exclusivement	43,9%
Professionnel exclusivement	12,5%
Etudes exclusivement	9,6%
Personnel et professionnel	16,1%
Personnel et études	8,5%
Professionnel et études	2,8%
Les trois	6,6%

Nous retrouvons ce constat dans les entretiens : seules trois personnes sur les seize interrogés fréquentent de manière occasionnelle les salles de lecture. Leurs usages s'inscrivent dans une recherche personnelle dans un domaine particulier comme l'histoire, la généalogie, les sciences, l'histoire de l'art... Il s'agit plus précisément d'amateurs avertis qui ont développé au fil des années des activités de recherche. Cette activité va pour certains, bien au-delà d'un simple loisir car les résultats ont donné lieu parfois à des publications déjà éditées ou en cours d'édition, ou à la constitution d'archives personnelles, ou bien encore à la mise à disposition des résultats de ces recherches par la constitution d'un site perso.

Peu utilisent Gallica à des fins professionnelles mais cet usage n'est pas un usage restrictif. Gallica permet ainsi à un panéliste correcteur de profession de vérifier noms propres, édition et orthographe ; pour un professeur d'enseignement supérieur, d'étayer son cours. Quelques exemples dans ce sens font apparaître une frontière floue entre les activités personnelles et professionnelles. Lors des entretiens, nous avons pu mieux nous rendre compte du temps consacré aux recherches. Il ne s'agit pas de trouver de la documentation complémentaire pour étayer son cours mais de la constitution de véritables archives sur son domaine d'enseignement (utilisateur M).

3.4.2 Navigation sur Gallica

Sessions Gallica

Dans les données recueillies *via* « Audinet », l'accès à Gallica représente 5,3% des 15 500 sessions enregistrées entre juillet et décembre, et concerne un nombre important d'utilisateurs : 53 panélistes sur le panel général de 72 personnes (soit les $\frac{3}{4}$) sont allés au moins une fois sur Gallica

L'intensité d'usage de Gallica répond à une répartition similaire à celle du Web : un tiers des visiteurs de Gallica fait plus de 77% des sessions comportant un accès au site, tandis qu'un autre tiers fait seulement 4% de ces sessions.

Plus encore, on observe que les utilisateurs les plus intensifs de Gallica sont aussi des utilisateurs intensifs du Web : sur la période, ils comptent 400 sessions en moyenne, pour une moyenne de 270 sessions au total pour les autres visiteurs de Gallica, et 150 sessions pour ceux qui n'y sont jamais allés.

Ainsi, par son mode de recrutement à partir d'un questionnaire affiché sur le site Gallica, le panel est constitué d'utilisateurs intensifs de ce site, et ce sont eux qui tirent le trafic Web général du panel vers le haut sur l'ensemble de la période, sur Gallica comme sur les autres sites.

Navigation sur Gallica

Les 819 sessions comportant un accès Gallica sont globalement plus longues que les autres : 1 h 01 min. en moyenne, contre 28 minutes pour les autres sessions. Par ailleurs, dans une session comportant un accès à Gallica, le temps total passé sur ce site proprement dit est en moyenne de 24 minutes, soit presque la durée moyenne d'une session sans Gallica. La visite de la bibliothèque numérique de la BnF s'apparente ainsi à une activité de longue durée.

Plus encore, la consultation de Gallica s'impose comme une activité excluant la visite alternée d'autres sites. Nous avons en effet analysé, dans les sessions, l'alternance de la visite des différents sites. Nous observons que dans plus de la moitié des cas, la navigation sur Gallica occupe une seule séquence, elle n'est pas alternée avec la visite d'un autre site ; et pour 22 % des sessions seulement, on compte deux séquences distinctes sur Gallica. Le « multi-tâches » est donc rarement pratiqué, et Gallica induit une activité longue et monolithique.

Rubriques visitées

Au sein de ces longues séquences sur le site de Gallica, la consultation de documents (texte, image, audio) est majoritaire, et s'appuie sur le duo Recherche – Feuilletage (voir Tableau 16). Le téléchargement d'ouvrages, au format TIFF ou PDF, en retrait par rapport au feuilletage, est très important : dans 38% des sessions Gallica, l'utilisateur est amené à sauvegarder une copie locale d'un document sur son poste, pour un total de plus de 2000 ouvrages téléchargés.

Tableau 16 - Services visités sur Gallica

Service	Nombre de sessions	Part des sessions Gallica	Nombre de panélistes
Consultation	628	76,7 %	48
Page Accueil	617	75,3 %	53
Recherche	596	72,8 %	50
Feuilletage d'un ouvrage	548	66,9 %	47
Téléchargement d'un ouvrage (tout ou partie)	314	38,3 %	34
Découverte	92	11,2 %	30
Dossiers	89	10,9 %	31
Aide	20	2,4 %	13

La consultation des dossiers reste marginale en nombre de sessions, mais concerne la moitié du panel, ce qui dénote un effet de « visite par curiosité » pour cette rubrique, élément confirmé dans les entretiens. Les dossiers thématiques de Gallica apparaissent ainsi comme des éléments d'information à part qui nécessitent une autre démarche.

"Je trouve que c'est un peu « gadgetique » et je trouve ça un peu limité. Puis bon... Comme mon domaine est ciblé, si vous voulez... Ça porte souvent sur des petits aspects qui sont pas dans mes préoccupations générales." (Utilisateur H)

"Je suis allé voir ce qu'il y avait dans les dossiers pour savoir si ça m'intéressait ; il y en a je sais plus combien ; pour l'instant il y en a pas qui sont dans mes centres d'intérêt pour le moment, mais je me suis promis un jour où j'aurais du temps d'approfondir un petit peu ça." (Utilisateur F)

"J'y vais systématiquement rien que pour avoir le plaisir de voir quelques belles enluminures comme ça. Si vous voulez, c'est plus de la recherche, c'est le plaisir de voir de belles choses." (Utilisateur P)

Pour la plupart des interviewés, la visite des Dossiers dans Gallica s'inscrit explicitement dans une autre logique de parcours : la flânerie des expositions numériques est opposée à l'efficacité de la recherche ciblée *via* le moteur de recherche interne au site. Les deux contextes d'usage sont ainsi clairement différenciés, le mode « prédateur » étant le plus fréquent.

3.4.3 Ergonomie du site

Module de recherche

Figure 6 - Interface du module de recherche de Gallica

L'utilisation de Gallica se fait essentiellement par le module recherche (Figure 6). Ce mode recherche demande une certaine expérience : beaucoup ont mis en avant les difficultés de départ à trouver les modalités pertinentes pour accéder rapidement l'information recherchée. Malgré tout, une fois l'information trouvée, ils en sont satisfaits.

"D'ailleurs, mes premières recherches sur « Gallica », ça a été un véritable calvaire." (Utilisateur N)

Bien souvent, les panélistes connaissent déjà précisément le document recherché. Ils interrogent alors le module de recherche de manière classique par indice bibliographique, à partir de l'auteur ou du titre du document.

"J'ai soit un titre, soit un auteur que je recherche, et puis je vais dans la rubrique recherche et je mets les mots clés pour trouver comme ça." (Utilisateur G)

On observe néanmoins parfois des stratégies de recherche combinatoire : à partir du module de recherche, les personnes interrogées développent une stratégie de recherche pour ratisser au plus large l'information. Sujet, auteur, mot du titre, variation orthographique, toutes les tactiques sont mobilisées pour aller retrouver un document.

"Là, je connais le nom de l'auteur alors je mets le nom de l'auteur. Bien ! Sinon, j'explore, je sais pas... Je m'intéresse à l'Algérie, mais il faut penser à chercher le mot Algérien ou Algérienne ou barbarie ou berbérie ou Maghreb ou... Et quelques fois, ça peut être dans Afrique etc. Donc il faut ne pas hésiter à mon avis, à balayer au maximum, utiliser le moteur de recherche tous azimuts pour essayer de trouver. Voilà comment je procède." (Utilisateur P)

"Par thème puis par auteur. Pour affiner c'est par auteur. Sinon je trouve ça très intéressant parce qu'on fait des recoupements. Quand vous commencez par les mots du

titre... Donc il ne faut pas oublier les guillemets ; parce que dans l'exemple qui est donné il n'y a pas de guillemets. Quand on envoie ça textuellement, on vous dit : zéro requête. Donc il faudrait peut-être leur dire de ne pas oublier de mettre les guillemets. Cela vous donne un certain nombre de choses, et ensuite lorsque vous reprenez ces idées-là et que vous les mettez dans les sujets, vous en avez d'autres ; donc ensuite on fait des recoupements, et on se dit : tiens celui-là est plus intéressant finalement j'aurais peut-être dû le mettre là-haut. Bon." (Utilisateur M)

"On ne peut pas dire que [le module de recherche de Gallica] me comble d'aise, par exemple, pour les mots clés, mais ça je n'ai pas la formation voulue pour pouvoir l'utiliser ; donc je fais de la recherche libre, j'y passe beaucoup de temps, mais le champ est libre puisque j'ai pas la télé. Et que j'ai l'ADSL. Et là je pioche, et puis j'ai un peu l'esprit d'escalier. Donc on fait des combinatoires, et puis l'ordinateur suffit pas, il faut un papier et un crayon...Et puis prendre à la volée, des éléments qui semblent pertinents et puis être un peu audacieux en disant on va jouer sur des critères multiples pour arriver à trouver." (Utilisateur H, à propos du module recherche)

Améliorations possibles suggérées par les utilisateurs

Parmi les propositions d'améliorations du site Gallica énoncées par les interviewés, on retiendra en particulier :

- Le désir d'un mode d'emploi et d'une aide à l'utilisation plus développée : trucs et astuces, impression, format des documents.
- La possibilité d'être informé des campagnes de numérisation.
- Une idée originale de legs en ligne des textes re-transformés en mode texte pour les autres usagers.
- La possibilité de convertir les documents pour les rendre lisibles sur PDA (Palm, Pocket PC).
- Rendre plus accessible les tables de matière et les index.
- Favoriser la navigation par des tables de matière systématiques.
- Des messages restés sans réponse.
- Une amélioration de l'ergonomie :

"Pour ma part, je trouve ça très bien fait qu'on puisse mettre en PDF les pages que l'on veut, euh très bien. Le seul problème que je dirais, mais bon, c'est pas évident, c'est quand on met un mot-clé on a une liste de livres. A chaque fois, faut cliquer pour avoir le résumé. Et après revenir sur le document pour avoir les autres. A la limite, faudrait peut-être tous les résumés d'un coup ça serait peut-être plus simple, je sais pas Ça me semble pas évident à ce niveau-là." (Utilisateur I)

Le reproche le plus souvent adressé est relatif au format PDF des documents, la longueur de téléchargement. Ces points négatifs sont souvent relativisés par les panélistes eux-mêmes qui ont conscience des difficultés techniques des opérations.

Nous avons mené dans le cadre du projet une expertise ergonomique en utilisant une méthode conçue par Jean-Pierre Savary¹, destinée à améliorer l'accessibilité des sites

¹ Ergonome à FTR&D – Direction des Interactions Humaines – Laboratoire Usages, Créativité, Ergonomie, il a mis au point une méthode permettant d'améliorer la conception des sites dans le but d'en faciliter l'accès aux personnes non-voyantes et mal-voyantes. En général, l'application de ces recommandations a pour effet d'améliorer l'ergonomie et l'accessibilité des sites à l'ensemble des utilisateurs et pas seulement aux personnes ayant des déficiences visuelles.

Web. Cette intervention a permis d'améliorer certains points, en particulier la page de recherche de Gallica.

3.5 Conclusion

La population étudiée est composée d'internautes expérimentés, fortement équipés en haut-débit. En termes socio-démographiques, nous retrouvons la population « attendue » pour ce type d'usage (CSP+, niveau de diplôme élevé...), mais avec une sur-représentation des seniors par rapport à la population générale des internautes en France.

Les utilisateurs mettent en œuvre des stratégies avancées de recherche d'information sur le Web, ils savent surmonter les difficultés liées à cette activité, manient avec dextérité plusieurs moteurs de recherche et arrivent à repérer des sites de référence dans leur domaine d'intérêt.

Nous notons également ici un usage du haut-débit un peu particulier, centré sur la consommation de « contenus numériques à lire », nécessitant de lourds téléchargements (documents souvent en mode image).

4. Le document numérique dans les pratiques

Dans cette section, nous étudions plus avant la place du document numérique dans les pratiques : comment ce document est utilisé compte tenu de son format, de sa disponibilité, de ses modes d'accès ; comment il est appréhendé par rapport au support papier traditionnel. Nous examinons ensuite comment, étant donné ces modes d'appréhension des documents, de nouveaux profils de lecteurs et de chercheurs émergent au sein du public de Gallica, qui diffèrent du lectorat traditionnel de la BnF. Nous faisons ici majoritairement appel aux entretiens pour répondre aux interrogations que nous formulons ici.

4.1 Traitement des documents numériques

Téléchargement

Nous l'avons déjà dit, 38% des sessions sur Gallica comportent une action de téléchargement d'un ou plusieurs ouvrages, ce qui représente un total de plus de 2000 documents téléchargés par 34 des 53 personnes du panel ayant fréquenté Gallica au cours des six mois d'observation.

Les entretiens ont permis de mettre en évidence l'usage régulier voir systématique du téléchargement. Quand le contenu correspond aux besoins de la recherche, que ce soient des documents issus de bibliothèques électroniques ou plus largement de contenus accessibles sur le Web ; les documents sont enregistrés sur le disque dur puis sur des CD pour certains. L'ensemble de cette documentation est alors stocké et classé pour constituer un fond personnel.

Les motivations du téléchargement correspondent à :

- un besoin de conservation du document :

"Certains sites sont éphémères et je m'en rappelle une fois un site que j'ai téléchargé pour mémoire" (Utilisateur K)

- un désir de pouvoir le consulter ultérieurement :

"J'ai vraiment l'habitude d'enregistrer sur le disque dur, de les garder en mémoire" (Utilisateur L)

- un désir de se constituer des archives :

"Le plus difficile c'est ensuite d'archiver et d'avoir le temps de regarder les archives." (Utilisateur M)

Au sein de recherches souvent très ciblées, l'information précise contenue dans un document n'est pas seule sauvegardée, et l'entour de ces informations est également

objet de conservation. Le téléchargement va de l'enregistrement de documents avec conservation des indications d'auteur et de provenance, à l'aspiration de sites entiers.

"La même chose, par contre je dirai, j'enregistrerai peut-être pas forcément uniquement l'info qui me semble intéressante mais j'enregistrerai la page entière ou si je trouve que le site est intéressant ben j'avoue que je copie le site sur le disque dur, ou une partie du site." (Utilisateur L)

"Si c'est quelque chose que je vais consulter par la suite, donc là, je l'enregistre et je le garde pour après." (Utilisateur G)

Un nombre important d'interviewés déclare constituer ainsi des sortes de fonds numériques personnels centrés sur leurs centres d'intérêt propres, dont certains les documentent et en assurent la pérennité en les gravant sur CD-Rom, ce qui atteste une véritable politique de conservation à long terme :

"Donc ensuite sous [...] Access, je remets : nom d'auteur, titre, et le numéro du Cdrom sur lequel, parce que c'est encore archivé avec des numéros, et quand par exemple je travaille actuellement sur l'évolution de la mécanique, je prends par exemple les trucs de Pierre Duhem qui sont numérisés sauf le volume titre, c'est pas grave, et après en fonction de ça je compulse, alors je compulse directement sur l'ordinateur. Je n'imprime pas." (Utilisateur M)

"Je manque que de place sur son ordinateur. Je télécharge et puis je m'en... Je le stocke. Même, parfois, je ne l'édite pas. Mais je sais que ça existe et puis bon, le jour où je vais... rédiger une partie de ce que je fais, bien à ce moment-là, je l'imprimerai ou je... je le regarderai". (Utilisateur N)

Question de format

Ce politique de conservation se situe dans le cadre d'un travail plus général d'accumulation systématique de documents sur un sujet donné. Dans cette perspective, le mode image prépondérant dans le fonds proposé par Gallica est pour certains un frein à la manipulation des documents, et certains procèdent à leur retranscription, manuelle ou par reconnaissance optique de caractères.

"Alors quand c'est des textes, c'est facile, avec la recherche on peut voir, on va chercher un mot, hein [...] Si vous voulez parler des médecins, vous cherchez médecins, vous allez trouver quelque chose hein. Comme il apparaît un peu de tout, mais seulement on repère tout de suite, hein. En images, allez voir !" (Utilisateur O)

"Alors, je le télécharge, puis je fais une reconnaissance de caractères, puis je lis tout le texte et je corrige ma reconnaissance de caractères qui va comporter énormément de fautes, bien sûr, puisque les ouvrages donnés en ligne par la Bibliothèque Nationale, ils ont raison de les faire suffisamment légers, il y a un compromis entre la qualité du texte et puis le poids qu'ils vont avoir. Donc ce poids ne permet pas d'avoir quelque chose de parfait, mais ça ne fait rien ! A ce moment là, cet ouvrage là, en même temps que je le lis, dans mon texte, je corrige. Et après, l'ouvrage, ça y est, j'en fais un autre document PDF mais ça n'a rien à voir avec celui de la Bibliothèque Nationale, il doit être parfait. Je respecte la mise en page du bouquin bien sûr." (Utilisateur P)

Ces éléments nous amènent à penser que le statut des documents électroniques est plus proche de celui de l'usuel que de l'œuvre, leur utilité étant définie dans le cadre d'une pratique ciblée et d'interrogations précises, ce que confirment les entretiens :

"Euh, au début j'imprimais, j'imprime moins, hein. Au début j'imprimais beaucoup, euh... j'imprime ce qui est utile, hein, par exemple je viens de vous donner un exemple pour Casanova ; quand j'ai repéré, je lis à l'écran, c'est fatigant mais je lis à l'écran, et quand je repère les pages qui vont me servir, alors là je les imprime." (Utilisateur O)

La conservation des documents numérisés et la constitution de fonds électroniques personnels peuvent être ainsi vues comme l'accumulation de documents de référence pour un usage ultérieur, même s'ils ne sont pas effectivement consultés par la suite. L'attitude est bien ici celle de chercheurs balayant leur champ de recherche, et il est intéressant de constater qu'elle concerne non seulement des enseignants et des chercheurs universitaires, mais également une population de « chercheurs amateurs », qui trouvent avec Gallica et les bibliothèques électroniques en général le moyen d'accéder aux fonds habituellement réservés aux universitaires et aux chercheurs professionnels.

Lecture sur écran vs. impression

Lors des entretiens, la lecture en ligne a été le plus souvent évoquée dans le cas des recherches ponctuelles et précises. Elle intervient généralement pour la lecture de petits documents ou pour localiser l'information avant enregistrement. Un cas particulier est à mettre en avant, l'utilisation de e-books et de lecture sur PDA¹ :

"Lire sur l'écran, d'abord on s'habitue, et puis il y a un mode de saisie que je retrouve un peu sur le Palm en petit format. On a une lecture qui est d'un certain point de vue accélérée parce que le coup d'œil est synthétique. (...) On a une saisie immédiate, évidemment un petit nombre de lignes, il faut que le reader soit bon, mais s'il est bon, c'est au contraire très confortable, on a lu la page d'un coup. Je me souviens d'avoir relu la Chartreuse de Parme de cette façon, avec un plaisir considérable, parce que c'était assez voluptueux comme ça, petites pages par petites pages de saisir le texte de manière très visuelle, presque comme au cinéma. Et ça donne une sensation particulière." (Utilisateur A)

Hormis ce cas exceptionnel, la lecture sur écran est peu pratiquée : le mot "fatigant" est le plus souvent employé. Quand elle est utilisée, elle intervient dans le cadre d'une recherche très précise d'informations, de citations ou de termes définis pour des corrections orthographiques, et concerne des portions très ciblées de documents.

Malgré cela, l'impression reste minoritaire, elle n'intervient que pour de petits documents ou des extraits, le coût important étant la cause première invoquée. L'impression permet entre autres d'extraire des parcelles de document et de les traiter de manière plus aisée.

"Sinon j'imprime la page qui m'intéresse avec les références." (Utilisateur I)

"Mais je vais imprimer le chapitre qui m'intéresse, c'est plus facile à visionner sur un papier que sur écran." (Utilisateur L)

"Euh, au début j'imprimais, j'imprime moins, hein. Au début j'imprimais beaucoup, euh... j'imprime ce qui est utile, hein, par exemple je viens de vous donner un exemple pour Casanova ; quand j'ai repéré, je lis à l'écran, c'est fatigant mais je lis à l'écran, et quand je repère les pages qui vont me servir, alors là je les imprime." (Utilisateur O)

¹ Personal Digital Assistant

D'autres imprimeront systématiquement, mais cette pratique reste anecdotique sur l'ensemble des panélistes interrogés.

4.2 Bibliothèques classiques et bibliothèques électroniques

4.2.1 Fréquentation des bibliothèques classiques

Sur l'ensemble des personnes interrogées, il est difficile d'identifier un comportement général en matière de fréquentation des bibliothèques classiques. Peu ont des cartes de lecteur de la BnF et rares sont ceux qui fréquentent leur bibliothèque municipale. En revanche, les utilisateurs des bibliothèques classiques fréquentent les bibliothèques patrimoniales pour accéder aux fonds anciens. Ceux-là en ont une pratique développée et utilisent fréquemment le prêt inter-bibliothèque.

À l'exception d'un interviewé, ce sont de vrais lecteurs assidus achetant régulièrement des livres, pour certains *via* le Web, pour les autres en librairie. Leur lecture est généralement centrée sur leur domaine de recherche. Elle est un complément d'information.

Les livres achetés sont récents ou anciens. Pour les livres anciens, ils sont commandés *via* des sites de commerce en ligne spécialisés tels chapitre.com, abebook... Certaines personnes interrogées utilisent ses sites en complément du site de la BnF. Les informations collectées par les deux biais sont alors complémentaires.

D'après les entretiens, deux types d'utilisateurs des documents semblent ressortir :

- Ceux qui se satisfont de la dimension numérique du document, en intégrant le plaisir de pouvoir accéder au document en tant que contenu. Cette forme leur suffit en tant que telle.
- Les "tout-papier" qui gardent ce rapport affectif à l'objet. Pour eux l'utilisation des documents de Gallica, leur convient en tant qu'outil de travail, mais n'est pas en soi une source de plaisir puisqu'il n'y a pas de contact à l'objet.

"Je préfère, le bouquin, c'est quand même un objet, c'est fabriqué, normalement ça doit être solide, je dis bien normalement. [...] Le bouquin, c'est quand même mieux ; il y a une pérennité aussi, ça se conserve, ça c'est le côté objet que l'on possède."
(Utilisateur F)

4.2.2 Apports des bibliothèques numériques

Facilité d'accès aux documents

Dans le cadre des recherches menées par les utilisateurs, l'apport des bibliothèques numériques et en particulier de Gallica est inestimable. Elle représente un gain de temps considérable et permet de collecter des documents originaux. Le retour à la source de l'information et de la connaissance sont des notions souvent évoquées.

Un apport important est à souligner : la possibilité de rebondir dans ses recherches via les recherches par mot clé. Cette possibilité permet souvent de découvrir des documents insoupçonnés et cet apport est très intéressant pour les utilisateurs interrogés.

"Q : L'apport des bibliothèques numériques et d'Internet, dans vos recherches, ça a été... plutôt un atout très positif... ?

R : Ah bien, fondamental, parce que je pense que si je n'avais pas trouvé certaines choses sur « Gallica », euh... quand je me suis dit « Je vais faire un livre », euh... ce que j'avais, je trouvais que c'était trop... simple." (Utilisateur N)

"J'étais à la recherche d'un site comme la BNF. Pourquoi, parce que comme je travaille à domicile, je peux, autrefois j'allais beaucoup dans les bibliothèques, je. Ça m'est moins facile. Et pour moi, c'était vraiment, y a une libération." (Utilisateur J)

Dans ce cadre, les bibliothèques électroniques sont mises au même rang que l'ensemble des autres sources d'information sur le Web en termes d'accessibilité. Dès lors, le Web apparaît comme un vaste corpus de sources diverses, au sein duquel les fonds numériques se distinguent par le choix qu'elles proposent et la qualité et l'originalité des documents :

"Je suis un picoreur mais un picoreur qui sait ce qu'il veut. C'est-à-dire que j'essaie de ne pas me, comment dire, de ne pas trop me disperser. Parce que moi, je suis pris par le côté utilitaire ; je suis pas un vagabond, j'aimerais bien mais je n'ai pas le temps. Je me sers d'Internet, en fait comme d'un outil. C'est un outil, c'est un outil, bon ça peut être des fois un outil culturel, donc c'est pour s'amuser, et c'est surtout un outil pour faire des choses, pour trouver de la documentation." (Utilisateur F)

L'information « à la source »

Les personnes interrogées ne sont pas des lecteurs de la BnF mais des internautes ayant des pratiques Web bien identifiées. Menant chacun à sa manière des recherches à titre le plus souvent personnel, tous les interviewés utilisent Gallica comme un site de référence, un usuel dans lequel ils retrouvent « l'information à sa source ».

La bibliothèque numérique de la BnF retrouve ici sa vocation patrimoniale, bien que cette dimension ne soit pas explicitement évoquée par les interviewés.

Cette notion de document à la source a été développée par plusieurs panélistes qui ne se satisfont pas d'une information relatée, copiée et déformée par d'autres auteurs.

"Ça a toujours été l'objectif fondamental : tout ce qui est dit, on doit donner son auteur, et on doit pouvoir être en mesure de prouver ce que l'on dit en montrant le manuscrit [...]. Ça ne veut pas dire nécessairement que un autre n'a pas écrit ça et que l'autre a plagié, mais au moins on a une preuve comme quoi ça a été écrit et dit ; donc c'est toujours le souci du scientifique derrière moi." (Utilisateur M)

"Moi ce qui m'intéresse c'est réellement de choper l'information à la source." (Utilisateur K)

"J'ai consulté beaucoup Gallica, il y a un certain temps, où je cherchais des sources primaires pour ma recherche." (Utilisateur G)

La fréquentation de Gallica est évolutive, elle se fait au rythme des nécessités de la recherche.

4.3 Usages-types

Sur la base des traces d'usages issues du trafic des utilisateurs ainsi que des entretiens menés avec seize d'entre eux, trois "portraits-types" d'utilisateurs ont été identifiés :

- Le chercheur d'information.
- Le bibliophile.
- Le « lecteur à l'écran ».

Il ne s'agit nullement ici de donner une segmentation des utilisateurs, mais plutôt de dresser des portraits-types correspondant à des usages récurrents observés dans notre échantillon.

4.3.1 Le chercheur d'information

Ce qui est mis en avant dans ce type d'usage, c'est bien le rôle de médiation de Gallica à l'intérieur d'Internet. Gallica est une source de documents (de référence) sur le Web, qui permet souvent d'accéder à des informations attestées et de poursuivre sa recherche. On utilise le moteur de Gallica (presque) comme tout autre moteur de recherche sur Internet et les objets visés – les contenus recherchés – sont de nature essentiellement électronique.

Dans notre panel, nous avons rencontré des utilisateurs menant des recherches dans un cadre professionnel, mais surtout, une majorité de personnes menant des recherches à titre personnel autour d'un hobby, d'une passion, un sujet de recherche très présent étant la mémoire familiale et régionale.

"J'étais financier, en fait, euh... dans des industries. [...] Je suis passé sur les recherches... sur la famille de Lorraine et la Lorraine... euh... Uniquement... D'abord, par le biais de ma famille. Il se trouve que ma famille a été liée à la Maison de Lorraine depuis... le XVIIème siècle jusqu'au XIXème, alors c'est comme ça que je suis passé... Parce que j'avais des papiers de famille là-dessus, et... Je suis passé sur les histoires de la famille de Lorraine, euh... par ce biais-là, si vous voulez." (Utilisateur E)

Ces recherches personnelles occupent une partie importante de leur temps, et donnent parfois lieu soit à des publications, à la constitution d'archives personnelles (CD-ROM) ou à la création de sites personnels.

"Mon intention, c'est de rédiger quelque chose, alors je ne sais pas si ça intéressera mon éditeur, parce que ce sera beaucoup moins illustré que mon livre précédent." (Utilisateur N)

"J'ai publié un petit bouquin il y a une dizaine d'années ; je vais en republier un avant la fin de l'année, un livre pour enfants, il y a le polar qui arrive et qui, je pense, sera pour le mois de mai." (Utilisateur F)

Dans le cadre de recherches à titre personnel, l'élaboration d'un site est perçue par les personnes interrogées comme un moyen de mettre en valeur les résultats de ses recherches.

"Bon alors le principe c'est que je raconte la vie des trois [nom de famille], un peu des cousins...Oui oui, non mais tout un tas de choses ; alors et à partir de cela pour m'amuser, j'ai fait des, ce que j'ai appelé des à-côtés, c'est ce que j'ai découvert à l'occasion de mes recherches, j'en ai fait des à-côtés. " (Utilisateur O)

4.3.2 Le bibliophile

Contrairement au premier cas, Gallica joue ici un rôle de médiation vers le monde "réel" (y compris la sphère marchande). Pour le bibliophile, Gallica est utilisé en "pré-achat", le but ultime reste l'objet-livre.

"A la première manière, elle est volontaire, je vais sur le site de la BNF parce que j'y trouve quelque chose, j'ai le souvenir que je pourrai trouver quelque chose, et que je vais me consacrer un peu de temps pour voir quelques images, ou quelques documents que je ne pourrai pas consulter, ou être sûr que je vais acheter un ouvrage ; je parlais de [nom d'auteur], je me suis acheté en consultant la BNF, ça m'a beaucoup intéressé, car j'ai acheté un ouvrage qui coûtait quand même 1 200 F." (UtilisateurH)

4.3.3 Le lecteur à l'écran

Enfin, nous avons choisi d'évoquer ici un portrait-type plutôt atypique, correspondant à un usage rare d'après nos observations. Il s'agit d'utilisateurs ayant une pratique intensive de lecture à l'écran. Dans ce cas, Gallica – ou tout autre site offrant des collections de textes et d'ouvrages – est considéré comme un lieu de "consommation" et non pas de médiation. Nous avons également noté, chez un utilisateur que nous avons interviewé, un déplacement des pratiques de lecture intensive du papier vers l'écran, accompagné d'une redécouverte du plaisir lié à la lecture et à l'appréhension de l'objet-livre.

"J'écume les sites où on peut télécharger des textes, c'est ma passion." (Utilisateur A)

"La motivation de la fréquentation est toujours le téléchargement de textes ; je recherche bien sûr des textes classiques ; je suis agrégée de lettres au départ. Donc c'est un peu normal. Donc c'est par rafales, brusquement l'idée me vient, je suis assez fantasque, l'idée me vient tout d'un coup de relire Scaron ; je l'ai dans ma bibliothèque, mais maintenant je lis de la sorte sur écran, et j'ai besoin du texte tout de suite, et donc je vais voir s'il est disponible à la BNF, et si l'ouvrage que je recherche n'est pas là, je vais chercher ailleurs." (Utilisateur A)

4.4 Conclusion

Dans le cadre des pratiques observées, le statut du document numérique semble plus proche de l'usuel que de l'œuvre. Sa consultation et son usage s'inscrivent le plus souvent dans le cadre de recherches personnelles sur des sujets précis : les textes

consultés sont majoritairement envisagés comme support de recherche bien plus que comme objet de lecture. Il est frappant de constater combien ces recherches personnelles sont menées avec rigueur et de méthode : démarche de recherche systématique (avec en particulier un usage avancé des moteurs de recherche sur le Web), classement rigoureux des documents allant jusqu'à la constitution de fonds documentaires personnels, tendance à privilégier les éditions de référence (ce qui représente un des motifs du recours à Gallica). En ce sens, les fonds numériques tendent à rendre disponibles, au sein des ressources accessibles sur Internet, des documents et des possibilités de recherche nouveaux, et élargissent ainsi le public des chercheurs à l'amateur et à l'érudit.

Conclusion

En premier lieu, le projet BibUsages a permis de mieux appréhender les utilisateurs des bibliothèques électroniques. Celles-ci attirent un public qui n'est pas nécessairement habitué aux bibliothèques, mais qui y vient par le biais de recherches spécifiques : dans les entretiens autant que dans le trafic observé chez les participants de l'étude, les fonds numérisés apportent la possibilité de disposer de manière simple et rapide de documents de référence – difficilement trouvables – et qui s'inscrivent dans le cadre de contextes de recherche précis. Ce public semble assez différent de celui des bibliothèques classiques, et les chercheurs « professionnels » y sont comparativement peu représentés. Les plus de quarante ans, actifs ou retraités, sont majoritaires dans la population observée et les bibliothèques électroniques sont avant tout pour eux une source d'informations dans le cadre de recherches personnelles. L'intensité d'usage est ici bien supérieure à celle de la population générale des internautes français et va de pair avec un très fort taux d'équipement en haut débit (câble, ADSL). Nous découvrons ici une population d'internautes seniors fortement équipés en haut débit, et dont les centres d'intérêt, outre l'offre de services et de communication classique, gravitent autour des contenus « culturels ». Cette population atypique dans le paysage des internautes français constitue en elle-même un groupe d'autant plus intéressant qu'il est mal identifié dans les études d'usages à large panorama car minoritaire.

Le projet BibUsages a également permis d'appréhender les contextes d'usages des fonds numérisés. Il apparaît que si d'une manière générale les utilisateurs des bibliothèques électroniques sont également de forts consommateurs de « contenus à lire » (journaux en ligne en particulier), au sein des sessions, l'usage des bibliothèques numériques est fortement corrélé à celui des moteurs de recherche d'une part, et des sites de vente de biens culturels d'autre part. Deux profils se dégagent : celui du « chercheur amateur », dont les centres d'intérêt sont pointus et déjà bien connus de l'utilisateur, et celui du bibliophile pour qui Gallica fait office de catalogue avant achat. Dans les deux cas, la lecture en ligne est rare, tout autant que l'impression des documents téléchargés, et la lecture s'apparente à la recherche de fragments ciblés au sein de vastes collections laissant de côté la totalité des œuvres. Dans ce cadre, le statut des documents en ligne semble remis en cause : tandis que l'édition papier reste du côté de l'œuvre, l'édition électronique s'apparente à l'usuel.

Ensuite, l'étude montre les passerelles entre Web marchand et non marchand pour les utilisateurs. Alors que les acteurs d'Internet (fournisseurs de contenus et d'accès) perçoivent une dichotomie forte entre sites marchands et non marchands, les internautes passent indifféremment d'un type de site à un autre, et l'on doit plutôt parler d'enrichissement mutuel entre sites marchands et non marchands dès lors qu'on les envisage sous l'angle des pratiques. M. Gensollen avait déjà mis en évidence les liens existants entre ces deux parties d'Internet dans un article de 1999 [6].

Dès lors, attirant de nouveaux publics, induisant de nouveaux modes d'appréhension des textes, s'inscrivant dans des parcours de lecture inédits, les bibliothèques

électroniques, loin d'être une simple version numérisée des fonds traditionnels, s'apparentent à un nouvel espace de lecture et de consultation aux côtés des bibliothèques classiques.

5. Bibliographie

1. Beaudouin V. et Licoppe C. (Eds), « Parcours sur Internet », *Réseaux* Vol. 20, n° 116, 2002.
2. Bryan-Kinns N. et Blandford A., « A survey of user studies for digital libraries », Working paper, 2000.
3. Catledge L. et Pitkow J., « Characterizing browsing strategies in the World-Wide Web », *Computer Networks and ISDN Systems*, 27-6, 1995, pp. 1065-1073.
4. Cockburn A. et McKenzie B., « What do Web users do ? An empirical analysis of Web use », *International Journal of Human-Computer Studies*, 54-6, 2001, pp. 903-922.
5. Cunha C., Bestavros A. et Crovella M., « Characteristics of WWW Client-based Traces », rapport n°1995-010, Computer Science Department, Boston University, 1995.
6. Gensollen M., « La création de valeur sur Internet », *Réseaux*, 17-97, 1999, pp. 15-76.
7. Jones S. *et al.* « A Transaction Log Analysis of a Digital Library », *International Journal on Digital Libraries*, 3-2, 2000, pp. 152-169.
8. Jones S. *et al.*, « An Analysis of Usage of a Digital Library », *European Conference on Digital Libraries*, 1998, pp. 261-277.
9. Sfakakis M. et Kapidakis S., « User Behavior Tendencies on Data Collections in a Digital Library », *ECDL 2002*, LNCS 2458, pp. 550-559.
10. Tauscher L. et Greenberg S., « How people revisit Web pages : empirical findings and implications for the design of history systems », *International Journal of Human Computer Studies*, 47-1, 1997, pp. 97-138.

6. Annexes

6.1 Questionnaire en ligne

Présentation de l'étude : La Bibliothèque Nationale de France réalise une étude afin de mieux connaître les visiteurs de son site *gallica.bnf.fr* et mieux cerner la façon dont ils utilisent ce site Gallica. Soyez assuré(e) que toutes les réponses resteront strictement confidentielles. Ce questionnaire durera environ 10 minutes.

Q1. Approximativement, combien de fois avez-vous déjà visité le site « gallica.bnf.fr » au cours des 6 derniers mois ?

1. C'est ma première visite
2. moins de 5 fois
3. 5 à 10 fois environ
4. 10 à 20 fois environ
5. Plus de 20 fois
6. (vous ne savez pas)

Q2. Diriez-vous qu' au cours des 6 prochains mois vous consulterez le site « gallica.bnf.fr »

1. Régulièrement
2. De temps en temps
3. Rarement
4. (vous n'aurez plus l'occasion de le consulter)
5. (vous ne savez pas)

Q3. D'où vous connectez-vous pour consulter Le site Gallica ?

(Plusieurs réponses possibles)

1. De chez vous
2. De votre lieu de travail
3. De l'Université (pour les étudiants) ou d'une école supérieure d'ingénieurs, de commerce, ...
4. D'un lycée ou collège
5. D'un lieu public tel qu'une bibliothèque, un cybercafé ...
6. D'un autre lieu

Q4. (si Q3=6) De quel autre lieu ?

(question ouverte)

Q5. Comment avez-vous découvert notre site ?

1. Vous avez vu l'adresse du site Gallica dans une brochure ou documentation de la Bibliothèque Nationale de France
2. Vous l'avez trouvé par un lien à partir d'un autre site ou dans un message (e-mail, forum, ...)

3. Vous l'avez trouvé par un moteur de recherche comme Yahoo!, Voilà, Google
4. Par un ami, une relation
5. Autre mode

Q6. (si Q5=5) Lequel ?
(question ouverte)

Q7. Approximativement, combien de temps avez-vous l'habitude de passer sur Gallica ?

1. Moins de 5 minutes
2. 5 à 10 minutes environ
3. 10 à 30 minutes environ
4. Plus de 30 minutes
5. (vous ne savez pas)

Q8. Habituellement, pourquoi venez-vous sur Gallica ?
(question ouverte)

Q9. Parmi les rubriques suivantes, quelles sont celles que vous avez consultées au cours de vos dernières visites ?

(plusieurs réponses possibles)

1. Gallica Découverte : Thèmes
2. Gallica Découverte : Chronologies
3. Gallica Découverte : Iconographies, monnaies
4. Gallica Découverte : Dictionnaires
5. Gallica Découverte : Mode texte
6. Gallica Recherche
7. Les dossiers de Gallica : Classique
8. Les dossiers de Gallica : Utopie
9. Les dossiers de Gallica : Proust
10. Les dossiers de Gallica : La voix
11. Sociétés Savantes
12. Voyages en France
13. Aide : Les questions/réponses
14. Aide : Assistance
15. (Je suis resté(e) uniquement sur la page d'accueil)

Q10. (Si Q9 pas 15) Etes-vous d'accord avec les phrases suivantes à propos du site Gallica ?

(une seule réponse par ligne)

	Tout à fait d'accord	Plutôt d'accord	Plutôt pas d'accord	Pas du tout d'accord	Ne sait pas
Ce site a un contenu de qualité	1	2	3	4	5
En général, on trouve l'information que l'on cherche	1	2	3	4	5
Le « look » de ce site est réussi	1	2	3	4	5
La présentation des pages est claire	1	2	3	4	5
L'information est présentée de façon attractive	1	2	3	4	5
On trouve FACILEMENT l'information que l'on cherche	1	2	3	4	5
On sait à tout moment où l'on se trouve dans le site	1	2	3	4	5
Le temps de chargement des pages est acceptable	1	2	3	4	5
Le site donne envie de revenir	1	2	3	4	5

Q11. Que souhaiteriez-vous trouver sur le site Gallica ?
(question ouverte)

Q12. Avez-vous enregistré l'adresse du site parmi vos favoris ou bookmarks ?

1. Oui
2. Non

PROFIL INTERNET DES VISITEURS

Q13. Depuis quand utilisez-vous, vous-même, Internet?

1. Depuis 2002
2. Depuis 2001
3. Depuis 2000
4. Depuis 1999
5. Depuis 1998
6. Depuis 1997 et avant

Q14. A quelle fréquence utilisez-vous, vous-même, Internet ?

1. Tous les jours
2. 2 à 5 fois par semaine
3. Environ une fois par semaine
4. 1 à 3 fois par mois
5. Moins souvent

Q15. D'où avez-vous l'habitude de vous connecter à Internet ?

(plusieurs réponses possibles)

1. De chez vous
2. De votre lieu de travail
3. De l'Université (pour les étudiants) ou d'une école supérieure d'ingénieurs, de commerce, ...
4. D'un lycée ou collège
5. D'un lieu public tel qu'une bibliothèque, un cybercafé ...
6. D'un autre lieu

Q16. (Si Q15=6) De quel autre lieu ?
(question ouverte)

Q17. (Si Q15=1) Quel type de connexion Internet avez-vous à domicile ?
(Une seule réponse possible)

1. Une connexion par modem
2. Une connexion Numéris
3. Une connexion Haut-débit ADSL
4. Une connexion Haut-débit Câble
5. Autres connexions
6. (vous ne savez pas)

Q18. (si Q17=5) Quel autre type de connexion Internet avez-vous ?
(question ouverte)

Q19. Quel usage avez-vous d'Internet ?
(plusieurs réponses possibles)

1. Recherche d'information
2. Communication : chat, groupes de discussion, messagerie, emails.....
3. Achat en ligne
4. Opérations et consultations bancaires ou boursières
5. Téléchargement de musique et/ou de vidéo
6. Téléchargement de logiciels
7. Jeux en ligne
8. Autre usage

Q20. (si Q19=8) Que faites-vous d'autre(s) sur Internet ?
(question ouverte)

Q21. Quel(s) sont vos PRINCIPAUX centres d'intérêt sur le Web ?
(plusieurs réponses possibles)

1. Actualités
2. Banque et finances
3. Economie et entreprise
4. Institutions et service public
5. Emploi, stage (recherche ou offre)
6. Autres informations économiques ou institutionnelles
7. Sciences Humaines et sociales
8. Art et Littérature
9. Recherche documentaire ou bibliographique
10. Sciences et technologies
11. Informatiques et multimédia
12. Autres informations culturelles
13. Sorties, divertissements
14. Voyages, tourisme
15. Sports
16. Jeux
17. Autres Loisirs
18. Communication
19. Autres centres d'intérêt

Q22. Consultez-vous des sites Web appartenant aux catégories suivantes ?

(plusieurs réponses possibles)

1. Sites d'Université et/ou centres de Recherche
2. Sites de bibliothèques
3. Sites d'établissements culturels (musée, galerie)
4. Sites de journaux, de magazines en ligne
5. Sites de e-commerce de biens culturels (tels que la fnac, amazon, alapage)
6. Aucun site appartenant à ces catégories

Q23. Consultez-vous le site Gallica pour un usage principalement ... ?

(plusieurs réponses possibles)

1. Personnel
2. Professionnel
3. Dans le cadre de vos études

Q24. Quel(s) type(s) d'ordinateur utilisez-vous habituellement pour vous connecter à Internet ? S'agit-il...

(Plusieurs réponses possibles)

1. D'un PC
2. D'un Mac
3. D'une station de travail

Q25. Quel est le système d'exploitation de votre/vos ordinateurs ?

(plusieurs réponses possibles)

1. Windows
2. Mac OS X version 10 ou supérieure
3. Mac OS Système 9 ou antérieur
4. Linux
5. Un autre système Unix
6. (vous ne savez pas)

Q26.A (Si Q15=1) Partagez-vous votre ordinateur à domicile avec d'autres personnes ?

1. oui
2. non

Q26B. (si Q15=2) Partagez-vous votre ordinateur avec d'autres personnes sur votre lieu de travail ?

1. oui
2. non

Q26C. (Si Q15=3) Partagez-vous votre ordinateur avec d'autres personnes à l'Université ou dans votre école d'ingénieurs, de commerce ... ? oui/Non

1. oui
2. non

Q26D. (Si Q15=4) Partagez-vous votre ordinateur avec d'autres personnes au lycée ou au collège ?

1. oui
2. non

PROFIL SOCIO-DEMOGRAPHIQUE DES VISITEURS

Q27. Etes-vous :

1. Un homme
2. Une femme

Q28. Quel est votre âge ?

(question quantité)

Q29. Quelle est votre situation familiale ?

1. Célibataire
2. Vit maritalement
3. Marié(e) ou remarié(e)
4. Divorcé(e) ou séparé(e)
5. Veuf(ve)
6. (Vous ne souhaitez pas répondre)

Q30. Avez-vous une activité professionnelle rémunérée ?

1. Oui
2. non

Q31. (Si Q30=1) Quelle est votre activité professionnelle?

1. Agriculteur exploitant
2. Commerçant, artisan,
3. Chef d'entreprise, cadre dirigeant
4. Profession libérale
5. Cadre du secteur privé
6. Cadre de la fonction publique (catégorie A)
7. Technicien, agent de maîtrise, contremaître, catégorie B de la fonction publique
8. Employé, personnel de service
9. Ouvrier
10. Etudiant ayant une activité rémunérée

Q32. (Si Q30=1) Quel est votre secteur d'activité?

(une seule réponse possible)

1. Agriculture, chasse, exploitation forestière, pêche
2. Industries mécaniques, électroniques, chimiques, agro-alimentaires, production et distribution d'énergie et d'électricité, Imprimerie et autres industries
3. Batiments, Travaux publics
4. Commerce et distribution
5. Transport (terrestres, eau, aériens) et Télécommunications
6. Hotellerie, restauration
7. Etude, conseil, services aux entreprises
8. Informatique
9. Banque, assurance, immobilier
10. Santé et action sociale
11. Arts, spectacles
12. Professions de l'information
13. Professions des Bibliothèques, musées et archives
14. Ecrivain
15. Métiers du livre
16. Enseignement du Primaire
17. Enseignement du Secondaire

18. Enseignement du Supérieur
19. Recherche
20. Administration publique
21. Services aux personnes (blanchisserie, coiffure, soins de beauté, pompes funèbres, activités thermales et de thalassothérapie...)
22. Autres services (Assainissement, voirie et gestion des déchets, services domestiques)

Q33. (Si Q30=2 ou Q31=10) **Quelle est votre situation ?**

1. A la recherche d'un emploi
2. Femme/homme au foyer
3. Elève ,lycéen ou étudiant de 1^{er} cycle
4. Etudiant de 2^{ème} cycle
5. Etudiant de 3^{ème} cycle en DEA/DESS
6. Etudiant de 3^{ème} cycle en Doctorat/Thèse
7. Service militaire
8. Clergé, religieux
9. Membre d'une association
10. Retraité
11. Autre situation

Q34. (Si Q30=2 ou Q33=3 ou 4 ou 5 ou 7) **Quelle est l'activité professionnelle du chef de famille ?**

1. Agriculteur exploitant
2. Commerçant, artisan,
3. Chef d'entreprise, cadre dirigeant
4. Profession libérale
5. Cadre du secteur privé
6. Cadre de la fonction publique (catégorie A)
7. Technicien, agent de maîtrise, contremaître, catégorie B de la fonction publique
8. Employé, personnel de service
9. Ouvrier

Q35. (Si Q30=2 ou Q33=3 ou 4 ou 5 ou 7) **Quel est le secteur d'activité du chef de famille ?**
(une seule réponse possible)

1. Agriculture, chasse, exploitation forestière, pêche
2. Industries mécaniques, électroniques, chimiques, agro-alimentaires, production et distribution d'énergie et d'électricité, Imprimerie et autres industries
3. Bâtiments, Travaux publics
4. Commerce et distribution
5. Transport (terrestres, eau, aériens) et Télécommunications
6. Hôtellerie, restauration
7. Etude, conseil, services aux entreprises
8. Informatique
9. Banque, assurance, immobilier
10. Santé et action sociale
11. Arts, spectacles
12. Professions de l'information
13. Professions des Bibliothèques, musées et archives
14. Ecrivain
15. Métiers du livre
16. Enseignement du Primaire
17. Enseignement du Secondaire
18. Enseignement du Supérieur

19. Recherche
20. Administration publique
21. Services aux personnes (blanchisserie, coiffure, soins de beauté, pompes funèbres, activités thermales et de thalassothérapie...)
22. Autres services (Assainissement, voirie et gestion des déchets, services domestiques)

Q36. Quel est votre niveau d'études ?

1. Aucun diplôme / certificat d'études primaires
2. BEPC, Brevet des collèges
3. CAP, BEP
4. Baccalauréat, Capacité en droit
5. Diplôme universitaire de 1^{er} cycle (Bac+2) : Deug, BTS, DUT
6. Diplôme universitaire de 2^{ème} cycle (Bac+3 ou Bac+4) : Licence, Maîtrise
7. Diplôme universitaire de 3^{ème} cycle (Bac+5) : DESS, DEA
8. Diplôme d'Ingénieur
9. Diplôme d'une Grande Ecole autre qu'une Ecole d'Ingénieur
10. Doctorat, Thèse

Q37. (Si Q33=3) Quels sont les revenus annuels de votre foyer ?

1. Moins de 15 245 euros (moins de 100 000 FF)
2. De 15 245 à moins de 30 490 euros (de 100 000 à moins de 200 000 FF)
3. De 30 490 à moins de 45 735 euros (de 200 000 à moins de 300 000 FF)
4. 45 735 euros ou plus (300 000 FF ou plus)
5. (vous ne savez pas ou ne souhaitez pas répondre)

Q38. Résidez-vous

(une seule réponse possible)

1. En France métropolitaine
2. Dans les Dom-Tom
3. Dans un pays francophone
4. Dans un autre pays

Q39. (Si Q38=3 ou 4) Le français est-il votre langue maternelle ?

1. Oui
2. Non

**Q40. (Si Q38=1) Dans quel département vivez-vous ?
(question quantité)**

Q41. (Si Q38=1) Habitez-vous.....

1. Paris ou agglomération parisienne
2. Dans une agglomération de **plus de 200 000** habitants : *Angers, Avignon, Béthune, Brest, Bordeaux, Clermont-Ferrand, Dijon, Douai-Lens, Grenoble, Le Havre, Lille, Lyon, Marseille/Aix-en-Provence, Metz, Montpellier, Mulhouse, Nancy, Nantes, Nice, Orléans, Rennes, Reims, Rouen, St-Etienne, Strasbourg, Toulon, Toulouse, Tours, Valenciennes*
3. Dans une agglomération de **100 000 à 200 000** habitants : *Annecy, Amiens, Angoulême, Genève/Annemasse, Bayonne/Biarritz, Besançon, Caen, Calais, Chambéry, Dunkerque, Limoges, Le Mans Lorient, Montbéliard, Nîmes, Pau, Perpignan, Poitiers, La Rochelle, St-Nazaire, Thionville, Troyes, Valence*
4. Dans une agglomération de **moins de 100 000 habitants**

5. Dans une **commune rurale** (moins de 2000 habitants)

Q42. (Si Q38=2) **Habitez-vous.....**

Items 1 ET 2 : A Conserver mais à Filtrer systématiquement

1. Paris ou agglomération parisienne
2. Dans une agglomération de plus de 200 000 habitants
3. Dans une agglomération de plus de **100 000** habitants : Pointe-à-Pitre/Les Abymes (Guadeloupe), Saint-Denis ou Saint-Pierre (la Réunion), Fort de France (Martinique)
4. Dans une autre agglomération
5. Dans une **commune rurale** (moins de 2000 habitants)

PARTICIPATION AU PANEL D'UTILISATEURS GALLICA

Conditions pour poser la Question Q43 sinon fin du questionnaire

-> **CONNEXION INTERNET A DOMICILE/SUR LIEU DE TRAVAIL (Q15= 1, 2, 3 ou 4)**

ET

-> **TRAVAILLER SUR UN SYSTEME D'EXPLOITATION WINDOWS (Q25= 1)**

ET

-> **DOMICILIES EN France métropolitaine et DOM-TOM (Q38= 1 ou 2)**

Q43. La Bibliothèque nationale de France souhaite constituer, en partenariat avec France Télécom, un panel d'utilisateurs du site « gallica.bnf.fr ». France Télécom fournira un logiciel SECURISE à installer sur votre ordinateur, permettant à la BnF de suivre la manière dont vous utilisez le site Gallica et d'autres sites comparables : les types de recherche que vous effectuez, vos parcours de consultation sur le site. Cette étude est destinée à améliorer le contenu et les performances de Gallica en fonction de vos usages et de vos attentes. Les informations recueillies resteront strictement confidentielles.

Si vous souhaitez des informations complémentaires concernant cette étude, n'hésitez pas à nous contacter à l'adresse suivante: bibusages@voila.fr .

1. Je souhaite faire partie de ce panel

2. Je ne souhaite pas faire partie de ce panel

Q44. (Si Q43=1 ; sinon, fin du questionnaire) Nous vous remercions de votre participation à ce panel d'utilisateurs Gallica. Merci de nous laisser vos coordonnées afin de vous contacter d'ici quelques semaines pour la mise en place du logiciel. Elles resteront confidentielles, conformément à la loi informatique et libertés.

Votre Nom :

Prénom :

Au moins l'un des 2 numéros suivants pour vous contacter :

Téléphone personnel :

Téléphone professionnel :

Adresse e-mail (information obligatoire) :

Merci d'avoir accepté de participer à cette étude.

6.2 Grille d'entretiens

Utilisation générale d'Internet

Usages et Fréquence

Sur votre utilisation générale d'Internet...

Bref rappel sur l'équipement

Comment êtes-vous arrivé à Internet ?

- Date de première utilisation
- Dans quel cadre ?
- Dans quel objectif ?
- Modalités d'apprentissage

Vous servez-vous beaucoup d'Internet ? Pour y faire quoi ?

- Fréquence d'utilisation – régularité d'utilisation – usages courants, besoins ponctuels – partage mail/Web/chat/jeu/autres
- Distinguer les différents type d'usages personnels et professionnels ?
Quelle fréquence ?
Interactivité entre les deux usages (pro et personnel) ?

Comment utilisez-vous le mail et les autres moyens de communication ?

- Nb adresses – mail classique, webmail
- listes de discussion –
- fréquence de consultation
- nombre de correspondants
- chat, forums

Avez-vous un site Web / participez-vous à la conception d'un site ?

- Date de création du site – Contenu et but du site – Fréquence des modifications
- Evolution du site (contenu, présentation, public visé) –
Connaissance de la fréquentation du site – Motivation et objectifs pour la création du site

Modalités des usages

Qu'allez vous voir sur le Web ?

Informations – Utilisations particulières (achat en ligne, réservation...)
téléchargement de programmes

Avez-vous des sites privilégiés que vous visitez régulièrement ?

Fréquence et mode d'utilisation
Favoris (nombre, organisation thématiques...)

Y a-t-il des sites que vous avez fréquentés intensément dans des contextes précis ?

Evénements particuliers (préparation de vacances, recherche d'appart, programme de cinéma, événement particulier, actualités ... etc)

Comment trouvez-vous de nouveaux sites ?

Utilisation des moteurs, des annuaires
Utilisation de liens de sites vers d'autres sites
Adresses recommandées par d'autres gens (ami, mailing list, ...)

Quels moteurs de recherche utilisez-vous ?

Si vous en avez un moteur privilégié, pourquoi l'avez vous choisi ?
Si plusieurs, comment les utilisez-vous ?

Quelle méthode employez-vous pour effectuer votre recherche à partir d'un moteur ?

Réflexion préalable sur un mot clé, combinaison, affinage, requêtes successives...

Que faites-vous quand vous trouvez une page intéressante ?

Sauvegarde locale – bookmark – comment la retrouver – conseil du lien à d'autres gens
Modalités du traitement de l'information. (Téléchargement, impression...)

Site de la BnF

Nous allons maintenant en venir aux bibliothèques électroniques...

Vous connaissez sans doute le site de la BnF
Comment l'avez-vous connu ?

Quel est la motivation première de votre visite ?

- La curiosité : intérêt culturel, suivre l'actualité de la BnF, recherche au gré du surf sur les différents dossiers
- Intérêt professionnel : recherche déterminée (utilisation du catalogue, Gallica, dossier pédagogique)
- Intérêt personnel : recherche personnelle sur un sujet particulier/ préparation et réservation des documents pour votre prochaine visite à la BnF.

Quelle est votre fréquence d'utilisation du site de la BnF ?

- Fréquence en fonction des différents types de navigation.
- Quotidien, hebdomadaire, plusieurs fois par mois.

Quelle rubrique visitez-vous le plus ?

Catalogue en ligne, Collection, Informations pratiques, service aux lecteurs, Gallica, programme culturel, dossier pédagogique ...

Plus généralement, vous direz que vous êtes satisfait, plutôt satisfait ou peu satisfait du site ?

Navigation, lisibilité des pages, se repérer dans l'espace, code couleur.
Points positifs et points négatifs

Gallica

Comment avez-vous découvert Gallica ?

Date de découverte –
Mode : par un moteur, un annuaire, par quelqu'un, par une publicité, par hasard - Via le site de la BnF

De manière générale comment vous connectez-vous à Gallica ?

Il est dans vos bookmarks.
Vous passez par le site de la BnF.
Vous tapez directement l'adresse dans votre navigateur.
Variation de fréquences en fonction des différents besoins ponctuels (recherche particulière sur tel thème, tel auteur...)

Pourquoi passez-vous par le site de la BnF?

Parce qu'il est enregistré dans vos bookmarks
Vous ne connaissez pas l'adresse directe de Gallica
Parce qu'il correspond à une stratégie de recherche, du catalogue à Gallica ou vice et versa.

Quelle est la motivation première de votre visite ?

Vous avez une recherche précise à effectuer
Vous cherchez une documentation sur un domaine
La curiosité, vous vous laissez guider au fil des pages

Quel procédé de recherche utilisez-vous ?

Catalogue, dossiers thématiques

Modalités d'utilisation des documents de Gallica

Comment utilisez-vous les documents trouvés ?

Lecture en ligne, Téléchargement, Impression.

Pourquoi ?

Vous les stockez pour les lire ultérieurement
Vous les imprimez pour un meilleur confort de lecture
Vous recherchez une information précise que vous détectez à l'écran

Une fois téléchargés, comment utilisez-vous les documents ?

Impression, lecture, réutilisation d'extraits

Combien de temps les gardez-vous en mémoire ?

Feuilletez-vous les documents en lignes ?

A quoi correspond ce feuilletage ?

Vous ne trouvez pas l'information recherchée, vous recherchez une information précise, ce mode-là vous convient, le feuilletage permet d'évaluer l'importance du document pour votre recherche avant téléchargement ou impression.

Si vous les lisez en ligne, cette lecture vous paraît-elle satisfaisante ?

D'un point de vue ergonomique, la navigation sur Gallica vous paraît-elle aisée ?
Repérage dans l'espace, code couleur, facilité d'accès au document...

Bibliothèques numériques, bibliothèques classiques

Connaissez-vous d'autres bibliothèques numériques ?

Les utilisez-vous ?

Quelles en sont vos utilisations ?

Que vous apportent-elles ?

Quels autres sites avez-vous l'habitude de fréquenter pour vos recherches ? Pourquoi ?

Achetez-vous (quand c'est possible) les ouvrages que vous consultez sur bibliothèque numérique ?

Lisez-vous sur écran ?

Gardez-vous une préférence pour le support papier (dans quels cas) ? Quelles différences percevez-vous entre les deux supports ?

Quelles sont vos habitudes en matière de fréquentation des bibliothèques classiques ?

Fréquence de visite

Consultation vs. emprunt

Types d'ouvrages consultés/empruntés

Nombre et type de bibliothèques fréquentées (universitaires, municipales, BPI, BNF...)

Achetez-vous beaucoup de livres ? Quel type ?

Posséder vs. Consulter – Collectionneurs

Que vous apporte l'usage des bibliothèques numériques par rapport aux bibliothèques classiques ?

Le développement des bibliothèques numériques a-t-il changé votre pratique des bibliothèques classiques et vos pratiques professionnelles

Table des matières

Introduction	2
1. Déroulement du projet et méthodologie	3
1.1 Présentation du projet et objectifs.....	3
1.2 Phases du projet et méthodologie.....	3
1.2.1 Phase 1 - Enquête de cadrage (mars 2002).....	4
1.2.2 Phase 2 - Constitution d'un panel, recueil de trafic (avril-décembre 2002).....	4
1.2.3 Phase 3 - Enquête qualitative sur un sous-panel (octobre 2002).....	6
1.3 Etat de l'art.....	7
1.3.1 Etudes d'usage des bibliothèques en ligne.....	7
1.3.2 Etudes d'usage d'Internet	8
2. Données de cadrage	11
2.1 Profil général à l'issue du questionnaire	11
2.1.1 Caractéristiques socio-démographiques.....	11
2.1.2 Les « gallicanautes » par rapport aux internautes français.....	12
2.2 Composition du panel	13
2.3 Profil général des personnes interrogées en entretien	14
3. Analyse des usages	16
3.1 Utilisation globale d'Internet	16
3.1.1 La messagerie	16
3.1.2 Usages du Web : données de cadrage	16
3.2 Contenus visités	19
3.2.1 Audience des sites	19
3.2.2 Audience par types de portails et de sites	20
3.2.3 Recherche d'information	22
3.3 Usages des bibliothèques électroniques	26
3.3.1 Audience des bibliothèques électroniques	26
3.3.2 Les bibliothèques électroniques en contexte.....	27
3.4 Usages de Gallica et du site de la BnF	28
3.4.1 Conditions de la fréquentation de Gallica.....	28
3.4.2 Navigation sur Gallica	31
3.4.3 Ergonomie du site.....	32
3.5 Conclusion	35
4. Le document numérique dans les pratiques	36
4.1 Traitement des documents numériques	36
4.2 Bibliothèques classiques et bibliothèques électroniques.....	39
4.2.1 Fréquentation des bibliothèques classiques	39
4.2.2 Apports des bibliothèques numériques	39
4.3 Usages-types	41
4.3.1 Le chercheur d'information	41
4.3.2 Le bibliophile.....	42
4.3.3 Le lecteur à l'écran	42
4.4 Conclusion	42
Conclusion	44
5. Bibliographie	46

6. Annexes.....	47
6.1 Questionnaire en ligne	47
6.2 Grille d'entretiens	56
Table des matières	59